CATÓLICA
ESCOLA SUPERIOR DE BIOTECNOLOGIA
PORTO

STRATEGY FOR TRACKING PROBIOTIC STRAINS IN MULTIPLE SAMPLES

by

Alexandra Lopes de Almeida Guimarães Serôdio

July 2025

CATÓLICA
ESCOLA SUPERIOR DE BIOTECNOLOGIA
PORTO

STRATEGY FOR TRACKING PROBIOTIC STRAINS IN MULTIPLE SAMPLES

Dissertation presented to Escola Superior de Biotecnologia of the Universidade Católica Portuguesa to fulfil the requirements of Master of Science degree in Applied Microbiology

---

by

Alexandra Lopes de Almeida Guimarães Serôdio

Supervisor (i3S): Doctor Ricardo Araújo

Tutor (University): Professor Doctor Ana Gomes

July 2025

Resumo

Probióticos são culturas microbianas vivas e o seu consumo, ou aplicação em quantidades adequadas em produtos alimentares, apresenta propriedades benéficas para a saúde do consumidor. Estes microrganismos são atualmente usados como estratégia para preservar o balanço das comunidades microbianas no organismo humano e, por serem semelhantes aos microrganismos que naturalmente fazem parte da nossa microflora, ou por desempenharem funções idênticas, são usados no tratamento de certas doenças. O consumo de produtos que integram estirpes probióticas tem mostrado grande potencial na prevenção de condições a nível oral, respiratório, gastrointestinal e cutâneo.

Bacillus pumilus, Bacillus subtilis, Levilactobacillus brevis (previamente designada de *Lactobacillus brevis*) e *Limosilactobacillus reuteri* (previamente designada de *Lactobacillus reuteri*) são quatro espécies de bactérias com estirpes probióticas bem conhecidas e são comummente encontradas no nosso organismo. Estas estirpes ajudam na inibição do crescimento de microrganismos patogénicos, como bactérias, fungos, protozoários e vírus, contribuem para a prevenção de doenças cardiovasculares e ósseas e para o tratamento de patologias orais, cutâneas e do trato gastrointestinal, tais como a periodontite, o eczema e inflamações crónicas do trato gastrointestinal, têm propriedades anti-inflamatórias e imunomoduladoras e são comercialmente interessantes por produzirem metabolitos como antibióticos, enzimas e antigénios.

Algumas estirpes de *Bacillus* comercialmente usadas como probióticos, apresentam vantagens tanto na sua forma vegetativa como na forma de endósporo, em comparação com estirpes do género *Levilactobacillus e Limosilactobacillus*. Estas vantagens incluem a capacidade de armazenamento por períodos extensos sob forma dissecada e à temperatura ambiente, a resistência ao calor e a estabilidade a múltiplos valores de pH, tornando-as particularmente adaptáveis às condições extremas do aparelho gástrico.

Com este projeto, pretende-se desenvolver um painel de *primers* específicos capaz de detetar estirpes concretas de *B. subtilis, B. pumilus, L. brevis* e *L. reuteri*, presentes em amostras biológicas.

Palavras-chave: Probióticos; *Bacillus*; *Levilactobacillus*; *Limosilactobacillus*; SNP.

Abstract

Probiotics are live microbial cultures, and their consumption or application in adequate amounts to food products is shown to have beneficial properties for the consumers health. These microorganisms are currently used as a strategy to preserve the balance of microbial communities in humans and, since they are similar to the indigenous microorganisms naturally found in our microbiota, or perform identical functions, they are used to treat certain diseases.

Consumption of products containing probiotic strains has shown great potential in the prevention of oral, respiratory, gastrointestinal, and skin diseases.

Bacillus pumilus, Bacillus subtilis, Levilactobacillus brevis (formally known as *Lactobacillus brevis*) and *Limosilactobacillus reuteri* (formally known as *Lactobacillus reuteri*) are four species of bacteria commonly found in our body, with well-known probiotic strains. They can help to inhibit the growth of pathogenic microorganisms, such as bacteria, fungi, protozoa, and viruses, contribute to the prevention of cardiovascular and bone diseases and to the treatment of oral, skin and gastrointestinal tract pathologies, like periodontitis, eczema and inflammatory bowel disease, have anti-inflammatory and immunomodulatory properties, and be commercially interesting since they produce metabolites such as antibiotics, enzymes and antigens.

Some strains of *Bacillus* used commercially as probiotics, offer advantages in both their vegetative and endospore forms, when compared to other *Levilactobacillus and Limosilactobacillus* strains. These advantages include the capacity to be stored for extended periods in their desiccated form and at room temperature, heat-resistance, and stability at multiple pH values, making them particularly adaptable to the extreme conditions of the gastric tract.

In this project, we aim to develop a panel of specific primers capable of detecting unique strains of *B. subtilis*, *B. pumilus*, *L. brevis* and *L. reuteri* in biological samples.

Keywords: Probiotics; *Bacillus*; *Levilactobacillus*; *Limosilactobacillus*; SNP.

Agradecimentos

Com o findar de mais um capítulo no meu percurso académico e etapa da minha vida, gostaria de agradecer a todas as pessoas que fizeram parte desta caminhada e que me ajudaram a atingir este meu objetivo de obter grau de mestre.

Antes de mais gostaria de agradecer à minha amiga e colega de curso Carolina Costa, por me ter introduzido ao meu orientador e por me ter ajudado a encontrar um projeto numa área que me suscita tanto interesse.

Ao meu orientador, Ricardo Araújo, deixo aqui um enorme agradecimento pela oportunidade criada em fazer parte deste projeto, mas também pela paciência, disponibilidade, ensinamentos que partilhou comigo ao longo deste percurso e pelo seu contributo na escrita da minha tese. Gostaria de agradecer também à *Fundação para a Ciência e a Tecnologia* que apoiou este estudo através do projeto 2023.14724.PEX.

Não podia deixar de agradecer à Escola Superior de Biotecnologia (ESB), instituição onde fiz a minha Licenciatura e agora concluo o meu Mestrado, e aos professores desta faculdade pelas bases que me transmitiram ao longo do meu percurso académico, que foram bastante importantes no desenrolar deste projeto.

O maior agradecimento dirijo-o aos meus pais, pelas oportunidades que me proporcionaram na vida, em específico a de obter um curso superior, mas também por me terem motivado a finalizar o meu Mestrado e por terem aceite e apoiado todas as minhas decisões, tanto académicas, como profissionais ao longo destes anos. Aos meus irmãos, Ricardo e Filipa, que também fizeram parte desta faculdade e que são um exemplo a nível académico e profissional para mim, agradeço todo o apoio e ajuda prestada para a conclusão desta etapa.

Às minhas amigas, Ana Catarina Sousa e Marta Silva, e às minhas afilhadas, Sofia Montalvão, Raquel Lopes e Íris Domingues, agradeço todo o apoio e motivação que me deram, mas acima de tudo a sua amizade.

Por fim, mas não menos importante, dirijo um enorme agradecimento ao meu namorado e melhor amigo, Ricardo, por toda a ajuda, apoio, companheirismo e paciência que teve nos últimos tempos. Foi um grande pilar para mim nesta jornada de altos e baixos e esteve sempre disponível para me ajudar em todos os momentos da mesma.

Acknowledgement

As another chapter comes to an end in my academic path and life, I would like to thank all of those who have been part of this journey and who have helped me achieve my goal of obtaining a Master's degree.

First of all, I would like to thank my friend and classmate Carolina Costa for introducing me to my supervisor and for helping me find a project in an area that interests me so much.

To my supervisor, Ricardo Araújo, I would like to express my gratitude for the opportunity to be part of this project, but also for his patience, availability, knowledge that he has shared with me throughout this process and for his contribution in the writing of my thesis. I would also like to thank *Fundação para a Ciência e a Tecnologia*, for supporting this study through project 2023.14724.PEX.

I could not fail to thank Escola Superior de Biotecnologia (ESB), the institution where I completed my bachelor's degree and now conclude my Master's degree, and the professors from this college for the bases they have given me throughout my academic career, which have been very important in the development of this project.

My greatest appreciation goes to my parents, for the opportunities they have given me in life, in particular the opportunity to obtain a college degree, but also for motivating me to complete my master's degree and for accepting and supporting all my decisions, both academic and professional, over the years. I would like to thank my siblings, Ricardo and Filipa, who also attended this college and set an academic and professional example for me, for all their support and help in concluding this stage.

To my friends, Ana Catarina Sousa and Marta Silva, and my goddaughters, Sofia Montalvão, Raquel Lopes and Íris Domingues, I thank you for all the support and motivation you have given me, but above all for your friendship.

Last but not least, I would like to say a huge thank you to my boyfriend and best friend, Ricardo, for all his help, support, companionship and patience in recent times. He has been a great support to me throughout this journey of ups and downs and has always been there for me every step of the way.

Contents

List of Figures

List of Tables

Abbreviations List

AT: Adenine and Thymine

bp: Base pair

CDSs: Coding DNA Sequences

DNA: Deoxyribonucleic acid

GC: Guanine and Cytosine

HGT: Horizontal Gene Transfer

kb: Kilobase pair

MLST: Multiple-locus Sequence Typing

Mt: Melting Temperature

NGS: Next-generation Sequencing

PCR: Polymerase Chain Reaction

PFGE: Pulsed-field Gel Electrophoresis

RNA: Ribonucleic acid

SNP: Single Nucleotide Polymorphism

tRNAs: Transfer RNAs

WGS: Whole Genome Sequencing

Chapter 1: Introduction

1.1 Biodiversity and Genetic Diversity

Biodiversity represents the assembly of all the different species that can be found in distinct environments. However, there is also diversity within species, resulting from genomic variation accumulated over time in each individual. Genetic diversity, therefore, refers to the set of distinct characteristics that are inherited within a species, and it is essential for individuals to adapt to environmental changes. It reflects some historical demographic processes that determine the setting, long-term persistence, and adaptive potential of populations, and for that reason there is an increasing awareness protecting hotspots for long-term persistence of diversity (1,2).

If a species exhibits high genetic diversity, it is expected to find a wide variety of traits represented in individuals from this group. This indicator, representative of a large gene pool, positively affects the resilience and function of the ecosystem, but also the ability of species to adapt, persist and survive in a given environment (3,4).

On the other hand, low genetic diversity is usually associated with reduced longevity and lack of health among individuals, as well as decreased capacity for population growth and increased susceptibility to diseases (4). The population size of a particular species can be reduced due to habitat loss, overexploitation, the impact of introduced species and pollution, reaching a point where factors such as genetics increase the risk of extinction. Endangered species usually have smaller or declining populations, making the loss of genetic diversity inevitable, and as a result the ability of populations to evolve and adapt to new selective pressures is reduced (4,5).

A study performed in genetics and in the ecology field (6) suggested that genetic diversity and species diversity are correlated in three main ways: (i) a parallel response to environmental stimulus, could be responsible for promoting a positive relationship between them; (ii) within a community or ecosystem, genetic diversity may casually control species diversity; (iii) the diversity and relative abundance of coexisting species may casually have an impact in species genetic, when the species diversity of a community affects negatively the selective system.

Quantitative indicators of changes in genetic diversity can be obtained by combining sequencing technology and genetic material isolation methodologies (1,3).

Genetic diversity underlies most of the phenotypic variability present in bacteria, as well as their geographic distribution, host specificity, pathogenicity, antibiotic resistance, and virulence (7), making it increasingly important to identify these organisms at strain level. Microorganisms are important elements in the maintenance of ecological balance in the biosphere, and extensive research on its genetic diversity have been carried out using morphological and molecular traits (8). The prokaryote genome is highly dynamic, with rates of gain and loss of genes comparable to the level of mutations in eukaryotes (9).

The acquisition of non-self-genetic material might be seen as a mechanism of diversification and adaptation to the environment, with heterogeneous environments of free-living organisms impacting the guanine and cytosine (GC) content by horizontal gene transfer (HGT) (10). A comparison of nucleotide sequences across several bacteria reveals that codon composition generally is related to genomic content and that the differences in codon composition take place to the greatest extent at the third base position (11). In order for an organism and its offspring to survive it is essential for the instability and integrity of the genome to be balanced. This dynamic leads to the evolution of the organism and can be caused by rearrangement of genetic material and point mutations, which comprise changes with multiple magnitudes and origins in the deoxyribonucleic acid (DNA) sequence (12).

Besides the endogenous source of variability, bacteria have other potential forms of diversity provided by a variety of transfer systems, such as conjugation, transformation and transduction. These allow bacteria to expand their genome through acquisition of genes from other bacteria, a mechanism that became known as HGT (10). Recent genomic comparison studies show that it occurs frequently in prokaryotes and has been playing an important role in microbial evolution, providing opportunities to improve their adaptation to the environment, as well as to the host in which they inhabit, and to enhance their competitiveness (13). Persistent gene loss, combined with periodic events of massive gene gain by HGT, dominates the genetic dynamics of prokaryotes. HGT of high mobile genetic elements, such as plasmids and genomic islands which contain multiple genes, has contributed to the evolution of interactions between prokaryotes and eukaryotes, whether by symbiosis or pathogenicity, acquiring mutualistic and infection abilities respectively (13).

Plasmids are linear or circular shaped molecules responsible for the autonomous replication of extra-chromosomal DNA inside host cells, sometimes spreading it between cells, and which have repetitive sequences and a variable number of copies (14,15). These structures are commonly found in bacteria and they do not carry any core fundamental genes for growth or multiplication of bacterial cells, although they play an important role in the virulence of these microorganisms, as they often transport and transmit genes that code for antibiotic resistance. They also present a critical importance for pathogenic bacteria, given the fact that they encode properties such as the production of toxins and colonization factors (16). Like genomic islands, plasmids are common vehicles for HGT (17) and they often have a distinct nucleotide composition from the genome they are associated with (15).

Genomic islands are clusters of known chromosomal genes or genes that are predicted to have been acquired by HGT, ranging in size from 10 to 500 kb (15). These units are able to integrate bacterial chromosomes, providing their host advantageous properties, such as new metabolic functions, antibiotic resistance, niche colonization or virulence factors (18). Genomic islands often have a different nucleotide composition compared to the rest of the genome (15).

There are two other mechanisms of HGT previously mentioned, namely transformation and transduction. The first relates to the ability of some bacteria to incorporate DNA from the environment into the cytoplasm and to integrate it into their genome, while the second relies on specific phages that sometimes encapsidate DNA fragments from the host, releasing them once they infect a new cell (19,20).

With the increased easiness in sequencing new bacterial genomes, increased efforts have been made to sample multiple genomes within a particular species in order to discover intra-specific evolutionary processes. Therefore, intra-specific diversity includes both genomic and phenotypic diversities being related to physiological versatility and ecological resilience upon climatic stress (21,22).

For example, *B. subtilis* has an unexpectedly high genomic diversity, found in the content of genes involved in antibiotic production, cell wall synthesis, sporulation and germination, more specifically the presence or absence of polymorphisms in those same genes (23,24). A unique trait of this species is its ability to trigger a production program of competent cells capable of efficiently internalizing and recombining exogenous DNA without apparent sequence specificity. This lack of specificity provides the opportunity for genomic diversification through the acquisition of new genes (9). *B. subtilis* KM and *B. subtilis* CGMCC63528 strains were

compared in a study and high-sequence variations were identified, including large insertions, deletions and single nucleotide polymorphisms (SNPs) (25).

1.1.1 Mutations

Mutations are permanent changes in genetic material and represent a major source of heritable phenotypic diversity, which is the substrate of natural selection (26). While some mutations are silent, others may lead to phenotypic variation, evolution, and specialization (27). They can be divided as beneficial, deleterious, lethal and neutral, based on their impact on fitness, and they mainly differ in two aspects, being them the frequency at which they occur and the importance they have to researchers. Beneficial mutations are rare and occur at single events (28) and are important in understanding adaptation (29), while lethal and deleterious mutations are more prevalent and are of great value in understanding gene function or molecular determinants of disease. Neutral mutations are typically used as markers for building genealogies and phylogenies, and because they lack selective effect, they have an evolutionary destiny that is completely governed by chance (26).

Except for a small number of mutations limited to the lac system and some previously described beneficial mutation involving regulatory changes in expression, beneficial mutations are uncommon, difficult to detect in the environment and to identify them it is a challenging task (30). Even in well studied bacteria such as *Escherichia coli*, there is a considerable difficulty in predicting what kind of change constitutes a beneficial mutation in a new environment (31,32).

These newly added sequences in the genetic material have the potential to alter protein function and expression, as well as impact chromosome structure, thus indirectly influencing the phenotype. When natural selection processes are not powerful enough to maintain certain genes, since they are not needed in a given environment, they are disabled and progressively disappear or get lost in large deletions (33).

1.2 Methods for evaluation of bacterial genotyping

Bacterial phenotype can be determined by the observation of colony morphology and other features on culture medium, results of biochemical tests, serology, susceptibility, and pathogenicity (34,35). But these approaches are not capable to discriminate and quantify all the diversity found within a certain species and may lead to substantial subjective judgment (32). In addition, these methods became difficult and time-consuming for the characterization of slow-growing and fastidious organisms (32).

Genotyping methods are used to identify and distinguish bacterial strains based on their genetic content, bringing significant advantages, including faster turnaround time and faster sensitivity generating genomic profiles (36). Strain refers to a genetic variant within a biological species, in which a group of organisms belonging to that same species share certain genetic traits that are not found in other members of the same species. This group of isolates, which can be distinguished from others of the same species by minimal but identifiable differences, exhibits either a distinct phenotype or genotype, or even both (37,38).

The increase of genotypic databases has been crucial to allow fast and easy access by laboratories enabling comparisons and better epidemiological surveillances to be conducted. Conventional phenotyping methods, mainly based on biochemical properties, tend to offer results similar to the ones observed by molecular biology approaches (39).

Pulsed-field gel electrophoresis (PFGE) is an electrophoresis genotyping technique based on the separation of large fragments of DNA (10 kb - 10 Mb), and their mobility through a gel matrix. By applying alternating electric fields at different angles, this technique is able to separate large molecules of DNA on agarose gel, which are formed by action of restrictive enzymes that fragment the genetic material in specific zones, with the banding patterns visualized in PFGE reflecting the DNA polymorphism in the recognition sites of selected groups of strains (40). The selection of restrictive enzymes is one of the most important factors in generating reliable PFGE types, since the site where these enzymes cleave is unique to each of them.

Bacterial isolates showing the same PFGE profiles are considered to represent the same strain, whereas a difference of one to three bands determines that the isolates are closely related (41,42). If isolates differ between four and six bands, they are considered possibly related, and if the difference in bands is greater than these values, they are considered unrelated. Although

PFGE is a widely used method, it is time and labour consuming, requires high-quality DNA, lacks resolution power to distinguish bands of nearly identical size, and is associated with higher risks of contamination due to prolonged handling of the bacterial material before being treated with the proteases and restriction enzymes (40).

Multiple-locus sequence typing (MLST) is one of the most widely used genotyping methods for characterization of bacterial strains, relying on DNA sequencing to determine allelic variations in several conserved genes (typically seven genes). This method allows the observation of multiple housekeeping genes whose sequences are constrained due to the essential function they encode, being the variation observed in these sequences neutral or nearly neutral (40). Generally, 400-500 bp fragments of the seven selected genes are sequenced and given a number, with each strain being assigned a seven-number allelic profile called sequence type. For example, *B. pumilus* MLST panel considers the housekeeping genes including 23S rRNA (RNA component found in the large 50S subunit of ribosomes), *gyrA* (DNA gyrase subunit A), *gyrB* (DNA gyrase subunit B), *pycA* (pyruvate carboxylase) and *rpoB* [ribonucleic acid (RNA) polymerase β subunit] (43).

MLST scheme does not cover intergenic regions, but it may capture small insertions and deletions in core genes that are missed by other analyses. This technique has some limitations, as the allele number system is not representative of the actual gene sequence, fails to identify variability between relatively closely strains due to the use of highly conserved housekeeping genes, and the associated costs and time consumed in sequencing seven genes can be high when large collection of isolates are used. Table 1.1 compares the advantages and disadvantages of applying these two approaches, MLST and PFGE, for genotyping.

Table 1.1 Comparison between MLST and PFGE bacterial genotyping methods (44–46).

		Advantages	Disadvantages
Bacterial Genotyping Methods	MLST	- Data electronically portable; - DNA sequences are easily stored in online databases; - It can be used as a non-culture-based typing method; - Practical for comparisons of a large number of isolates; - Suitable for long-term evolution of bacterial population structures; - Captures small insertions and deletions in core genes.	- Time and labour intensive; - High cost; - Lacking reliability (Alleles numbering system is not representative of actual gene sequence); - Use of housekeeping genes often fails to detect the variability of closely related strains; - Does not cover intergenic regions.
	PFGE	- Moderated cost; - Can separate large fragments of DNA; - Highly discriminatory and sensitive technique; - Good reproducibility.	- Restrictive enzymes with rare cutters; - Time needed to complete the analysis (may require 2 to 4 days); - Requires high quality DNA; - Lacks power of resolution to distinguish bands with nearly identical size; - Risk of laboratory-acquired contamination, due to prolonged handling of bacterial isolates.

1.3 Single nucleotide polymorphisms (SNPs)

There are millions of DNA sequence variations in the genome, as a result of different nucleotides in homologous fragments of DNA, and minor insertions and deletions in the genetic material (47). Such variations called SNPs are the most abundant source of polymorphism in the DNA. The high abundance and stability of these SNPs makes them particularly useful as DNA markers. They are the most frequent form of variability between species of microorganisms and can be used to identify and learn about the dynamics of microbiological populations. Setting up a database with a solid set of these markers will allow new approaches for investigating the causative genes of drug response or complex conditions, such has multifactorial disease, and epidemiological studies (48,49).

Continuous efforts are being made in order to create technology for large-scale analysis at low cost of SNPs and the abundant genetic variations that arise in organisms. This is particularly important for studying large microbial populations. SNP-based approaches allow maximum resolution among organisms with low levels of genetic diversity, making it the most suitable sequencing method (49), since their analysis covers intergenic regions and does not feature multiple genetic variations in single allele. Currently, there are available numerous high-throughput SNP technologies, even though there is not only one that can satisfy all users in terms of scale, accuracy, throughput and cost.

The restriction fragment length polymorphism method, capable of tracing insertions, deletions and restriction site polymorphism, was the very first technique to be used for the detection of DNA polymorphism and serves as standard for evaluating other DNA tests. It is also possible to rely on the enzyme-linked immunoassay which becomes advantageous since many clinical laboratories have the necessary instrumentation to perform it, despite its difficulties in detecting SNPs that shift a single amino acid residue.

In addition to the above-mentioned technologies, the three following methods are also used to detect SNPs. The Sequenom MassARRAY iPLEX, used to genotype species like *Mycobacterium tuberculosis, Neisseria gonorrhoeae* and *Pseudomonas aeruginosa,* emerged as an alternative platform to MLST (50), and has shown to be a method with great potential for users who seek an accurate SNP genotyping assay, with modest multiplexing and low assay setup costs as a result of unmodified oligonucleotide primers (51). Despite its many advantages,

this method has major limitations as it requires DNA sequencing of seven housekeeping genes, which makes it an expensive and time-consuming approach (50).

The other method used is based on DNA microarray technology, which is applied in the detection of nucleic acids. SNP microarrays have been used for the rapid characterization of complex clinical samples, without the use of culture, in order to determine SNPs that are associated with resistance and pathogenicity (52,53). It appears as an alternative to sequencing, since it provides a rapid method, capable of characterizing isolates from niche to strain in 24 hours, and at a fraction of the cost and work. Although sequencing provides a more detailed overview of the microbial genome, it ends up showing drawbacks on the above-mentioned aspects, and also on the time associated with library preparation, sequencing, bioinformatic analysis and data storage, which can often be obstacles to the analysis of isolates in a standard laboratory. When compared to other genotyping methods such as multiple-locus variable number tandem repeat analysis or MLST, SNP microarrays are capable of detecting SNP variations along all the existing sequences in the genome and deliver greater resolution, unlike MLST which only detects a few conserved genes, or multiple-locus variable number tandem repeat analysis which only identifies a selected number of tandem repeats (53). On a final note, about it can be mentioned the high-resolution melting technology which, of the three strategies discussed, is the one most subject to inaccuracies in genotyping, caused by background fluorescence. These anomalies can be overcome and clarified by combining this strategy with mini-sequencing assays (54,55).

SNaPshot mini-sequencing assay was introduced during the last decade as a quick, sensitive and robust technique performed using a multiplex single base extension primer strategy in the presence of fluorescent labelled dideoxynucleotides (56). These methods are capable of accurately genotyping multiple SNPs simultaneously in the same reaction and gel lane, requiring minimal optimization (56). Through the use of single base extension primers of different lengths, it allows the recognition of fragments of multiple sizes by automated capillary electrophoresis. As a result, an electropherogram is obtained with a set of peaks with different colours and lengths that represent site-specific genomic variations, by means of capillary electrophoresis (57). This approach has the advantage of targeting small polymerase chain reaction (PCR) products, facilitating the placement of markers and the absence of stutter products. However, the limited number of markers and alleles can be restrictive to some genomic analyses (58).

The majority of SNP detection methods requires a step for nucleic acid amplification as either PCR, quantitative PCR, solid phase, or isothermal amplification (59), since these provide analytical performance required for many molecular analyses. Normally, the PCR-based methods used for the detection of SNP are sorted into the following categories: (i) analysis of polymorphism and mutant alleles using primers that match the fragments containing substituted nucleotides or by use of oligonucleotides blocking undesired templates, and (ii) combined approach of melting curve analysis and real-time PCR techniques by use of hydrolysis probes, or fluorescent dyes bounded to double-stranded DNA (59). To perform SNP analysis on genomic data, raw sequence reads have to be compared with reference sequences using available tools for short-read alignment, to identify zones of variation (60). This highly informative markers are able to reveal the evolutionary history of homogeneous groups, and to detect and trace outbreaks. These are spotted by comparing the sequences of isolates of interest with reference genomes, and the nucleotides that vary between the two are recorded in the database.

1.3.1 SNP panels for microorganism identification and typing

1.3.1.1 *Lactobacillus* spp.

To date, it has proven to be quite difficult to distinguish between some closely related species within the *Lactobacillus* genus. However, there are at least three SNaPshot minisequencing tests described for the detection of *Lactobacillus* spp.

The first one was conceived to quickly identify species from the *Lactobacillus plantarum* group, with five being closely related (*Lactiplantibacillus fabifermentans, Lactiplantibacillus paraplantarum, L. plantarum* subsp. *argentoratensis*, *L. plantarum* subsp. *plantarum* and *Lactiplantibacillus pentosus*), in which *L. plantarum* subsp. plantarum and *L. pentosus* showed probiotic properties (61).

The second test was designed to identify the *Lactobacillus casei* group, which has five closely related species (*Lacticaseibacillus casei*, *Lacticaseibacillus paracasei* subsp. *paracasei*, *L. paracasei* subsp. *tolerans*, *Lactobacillus zeae* and *Lacticaseibacillus rhamnosus*). While the two subspecies of *L. paracasei* showed a complementarity of 86% in DNA

hybridization, *L. casei* and *L. zeae* presented a similarity value of 80%, proving that they belong to the same species. A few strains of the *L. casei*, *L. paracasei* and *L. rhamnosus* species have shown to be probiotics with a potential beneficial effect on our immune system and of interest to the food industry (61).

Later, a third test was developed to identify *Lactobacillus delbruekii* subsp. *bulgaricus*. This test, the mini-sequencing scheme, was suggested for targeting the *recA* gene and therefore to discriminate different species of the *Lactobacillus acidophilus* group. Only through phenotyping and genotyping techniques (e.g. 16S rDNA sequence analysis) it becomes problematic to distinguish and identify specific species and subspecies of the *L. acidophilus* group (62).

1.3.1.2 *Bacillus anthracis*

B. anthracis is a member of the *Bacillus cereus* group, which is a subgroup of the *Bacillus* species. Within it are found the following eight known species, *Bacillus cyttoxicus, Bacillus toyonensis,* the non-pathogenic bacteria, *Bacillus mycoides, Bacillus pseudomycoides* and *Bacillus weihenstephanensis*, and other opportunistic or pathogenic microorganisms to insects and mammals, *B. anthracis*, *B. cereus* and *B. thuringiensis*, which have shown to be extremely close related based on 16s rRNA gene analysis (63,64).

B. cereus is a food-borne pathogenic bacterium responsible for synthesizing toxins, that can be encoded in its chromosomes and plasmids, conferring this microorganism virulence. We can also find in some strains of *B. anthracis* and *B. thuringiensis*, several virulence factors, such as genes encoding enterotoxins, that are typically found in *B. cereus* (64). *B. anthracis*, the deadliest of these pathogenic bacteria, contains two plasmids responsible for its virulence, pXO1 and pXO2, in which pXO1 encodes three lethal toxins and pXO2 encodes a capsule composed of poly-D-glutamic acid, responsible for protecting the microorganism from being phagocytized. Only with these two plasmids present in *B. anthracis*, is the bacteria able to present virulence and infect humans (65).

The three species *B. cereus, B. anthracis* and *B. thuringiensis* display morphological and biochemical identity and share a high degree of genetic similarities. Although it has been suggested that their virulence factors and pathogenic properties play an important role in

distinguishing them down to the species level, the high degree of lateral gene transfer of these virulence factors and the fact that some of these genes are encoded on plasmids that can be easily lost by the microorganism, complicate the identification of these microorganisms as a species (66). Even the use of rapid molecular tests, like PCR, as shown not to be the right method for identifying *B. anthracis*, since this bacteria shares a high degree of genomic homology with *B. cereus* and *B. thuringiensis*. The thermal shift assay and high-resolution melting methods can be used to quickly discriminate *B. anthracis* from other identical bacteria, using two SNPs specific to this species, which are a nonsense mutation at nucleotide position 640 of the *plcR* gene and nucleotide position 1668 of the *gyrA* gene (65).

1.3.1.3 *Mycobacterium tuberculosis*

Nowadays, tuberculosis still represents a major health problem worldwide, and the microorganisms that are able to cause this infection in humans and animals include *M. tuberculosis, Mycobacterium africanum, Mycobacterium bovis, Mycobacterium canettii, Mycobacterium microti, Mycobacterium caprae, Mycobacterium pinnipedii*, *M. bovis* BCG and *Mycobacterium mungi* all forming the *M. tuberculosis* complex (67,68). Although they differ in some traits such as host tropism, phenotype and pathogenicity, all members of the *M. tuberculosis* complex are genetically close.

Distinguishing the different species of *M. tuberculosis* complex is essential in order to determine which ones are strictly zoonotic human or animal tuberculosis and consequently to initiate the appropriate therapy. In particular, it is necessary to be able to differentiate *M. tuberculosis* from *M. bovis*, as the later species is naturally resistant to treatment with the antituberculous drug pyrazinamide (67). The use of a minimum number of SNP sets has been proposed as a possible solution for distinguishing these species.

The following method has been developed to simultaneously distinguish *M. tuberculosis* complex members down to the species level and *M. tuberculosis* strains. It is a SNaPshot mini sequencing-based approach, which has a high multiplexing capacity, robustness and extreme sensitivity. In addition to these advantages, this method is simple and affordable, requiring only equipment commonly available in molecular microbiology laboratories.

In order to differentiate *M. bovis* from other *M. tuberculosis* complex species, the following SNPs serve as bases: nucleotide position 756 of *gyrB* gene and nucleotide position 1410 of *gyrB* gene. For the genotyping of *M. tuberculosis* isolates the following SNPs present in the H37Rv strain at position 1977, 3352929, 74092, 105139, 2460626, 232574, 311613, 913274, 2154724, which corresponds to *katG* at position 463 of the gene, were used as base. *M. tuberculosis* H37Rv is identified through the SNP at nucleotide position 95 of the *gyrA* gene (69).

1.3.1.4 *Brucella* spp.

Brucella is a pathogenic bacteria sorted into six species, that are distinguished according to their host specificity, being them *Brucella abortus, Brucella canis, Brucella melitensis, Brucella neotomae, Brucella ovis* and *Brucella suis*. Despite their host specificity, *Brucella* spp. proved to be difficult to differentiate using molecular techniques, since *Brucella* genomes are highly conserved, with more than 90% homology between species based on DNA-DNA hybridization, identical 16S rRNA sequence between all species and more than 98% identical genome sequence (70).

The two methods capable of distinguish species of *Brucella*, which have clonally derived populations structures, are based in SNPs that accurately describe the framework of a species. These methods are MLST of housekeeping genes and whole-genome comparisons (71).

Five housekeeping gene products (*abc*, *aroE* and *gdh*, *groEL*, *pip*, *cysW*) and the *omp25*, *rpoB* and *trpE* genes have been selected to identify species-specific SNPs required for real-time PCR testing. By using this method, 19 SNPs can be identified, including four specific to *B. abortus*, three to *B. canis* and *B. suis*, two to *B. ovis* and one to *B. melitensis* (71).

1.3.2 Whole Genome Sequencing (WGS)

Whole genome sequencing (WGS) technologies can be applied to the diagnosis of infectious diseases in the public health, animal health, and food and environmental safety sectors (60). A relevant example is *Helicobacter pylori*, a human gastric pathogen highly prevalent in the gastric microbiome of patients with chronic gastritis (72), that infects more than 50% of the world's population (73) and its colonization causes a persistent inflammatory response that results in this condition. This bacterium has a high genetic diversity, with each individual usually hosting a distinct bacterial population, and with recent advances in technology, new subpopulations have been identified showing to be composed of multiple strains of *H. pylori*. The genotypic variation of this bacteria may be related to the presence of virulence factors between strains and is associated with different outcomes of infection in distinct individuals (74). WGS technology has been mandatory in expanding our ability to detect a higher number of SNPs (75), since it covers the complete genome of the microorganism, as well as multiple insertions and deletions. There are numerous WGS approaches for quantifying the relationship between isolated bacterial strains, with the two most widely used in public health being gene-by-gene and SNP-based approaches.

SNP approaches identify unique nucleotide differences between isolates when compared to a reference genome. This reference genome must be closely related to the isolates in order to identify SNPs with true phylogenetic information for comparison (76,77). The gene-by-gene method, which can be considered an extension of the MLST scheme, identifies variations in accessory or core genes, due to mutations or recombinant events, and can be applied to various specimens without the need for high-quality reference genomes. Bacterial genomes contain both core and accessory components, and a great part of them is flexible (77). A core genome consists of all the genes shared between the strains of given species that are likely to encode functions related to the basic biology and phenotype of the species. Ideally, these genes truly reflect the genealogy within a species and do not change their presence over time (49). Some genes found in certain strains and those that are unique to each strain form the accessory genome. This group of genes are important for species' diversity and most likely provides selective advantages to bacteria, being some of this benefits antibiotic resistance, habitat adaptation and the capacity to colonize other hosts (78).

WGS has been shown to be a technique with great discriminatory power, and it is essentially used on small groups of isolates instead of large collections of them (77,79).

Although WGS is in theory the ideal way to elucidate regarding genetic variability within a bacterial species, there are however a few challenges facing the application of WGS, the major ones being data analysis due to its complexity, the need for bioinformatic support to interpret the information, and the associated costs (60).

1.4 Next Generation Sequencing (NGS)

To determine the exact order of nucleotides, the basic units of nucleic acids, in DNA and RNA, is called sequencing of genetic information. This method has considerably grown over the years, since it became more affordable and accessible for research institutes and clinical laboratories, to sequence genetic material.

The first major research around DNA sequencing was the Human Genome Project that required many years and the investment of billions of dollars (80). It involved first-generation sequencing methodologies, known as Sanger sequencing, developed in 1977 by Fred Sanger and was embraced as primary technology in laboratory and commercial sequencing applications due to its high efficiency and low radioactivity (81). Over the last decades it has been implemented in research and clinical genetics, focusing in the identification of deletions, substitutions and small insertions (82).

Next-generation sequencing (NGS) tools have been developed in the last decades showing lower associated costs and faster throughput with large volume samples (82,83). Yet, for laboratory experiments, the Sanger technique is still employed, especially for the study of regions – typically exons – where NGS lacks the capacity to provide adequate depth of coverage or produce sufficiently high-quality data (84).

The growth of NGS applications in science, medical diagnosis, agrogenomics, and forensic science, has been possible due to significant improvements in sequencing technology, such as in speed, read length, throughput, and a substantial decrease in per-base cost (85). This technology that allowed several small labs to sequence entire genomes, enables simultaneous sequencing of millions of DNA sequences, and may involve the following workflow: (i) sample preparation including extraction of genetic material by bridge amplification or emulsion PCR (76); (ii) library preparation, which normally includes mechanical or enzymatic shearing of the DNA (76,86), and clonal amplification; (iii) sequencing and alignment of short reads. Sequencing is the most direct step in the NGS process because all sequencer brands are relatively simple to operate and include comprehensive manufacturer support services. The critical part of this step is loading the proper amount of library molecules into the instrument, which can be achieved by accurate library quantification and by following appropriate quality control procedures (86). Short read technologies currently in use are collectively referred to as massively parallel sequencing and are also often mentioned as second-generation sequencing.

They generate billions of nucleotide sequences during each run, where each single genome is sequenced multiple times in small random pieces to generate very sizable data sets (76).

In order to obtain such a large number of DNA molecules, PCR amplification is required, although this process is susceptible of introducing some errors in the template sequence, as well as to the amplification bias. In addition, the amplification process increases the complexity and time associated with sample preparation, a process that may take several days (87). Furthermore, second-generation sequencing or short-read sequencing techniques cannot directly sequence DNA and RNA molecules without the need for library preparation or sequencing reagents (85). Short-read sequencing techniques tend to generate fairly fragmented genome assemblies, therefore it is needed additional work to assembly such sequences for longer reads in order to provide closed reference genomes. These longer reads are especially useful when sequencing complex genomic regions such as repeats and phages (50-75 kb in length) (76), which is why third-generation sequencing platforms have been introduced.

They directly target single DNA molecules without requiring PCR amplification, and to carry out their detection and functional analysis, instrumentation such as atomic force microscopy, single-molecule fluorescence spectroscopy/single-molecule real-time technology (Pacific Bioscience), and single-channel current measurements (MinION) has been created. These technologies have been used to screen physical properties of individual DNA strands, indicating that their use could be applied to single nucleotide resolution and rapid DNA sequence analysis. However, these new sequencers remain under development and present drawbacks regarding the reliability of the resulting sequence. They are still inaccurate in the analysis of short-sequencing read lengths which can lead to data reading errors (81), and data analysis can be very time consuming or may require extra knowledge in bioinformatics to collect the most accurate sequencing data (85). Table 1.2 compares several first-, second- and third-generation sequencing systems, listing their strengths and weaknesses, as well as their throughput range and read length potential.

Table 1.2 Characteristics, strengths and weaknesses of commonly used sequencing platforms (76,82,86).

Sequencing technology	Platform Instrument		Throughput range (Gb)	Read length (bp)	Strength	Weakness
First-generation	Sanger sequencing	ABI 3500/3730	0.0003	500 – 1,000	Produces long high-quality DNA sequences	Elevated Cost; Laborious and time consuming to sequence whole genomes (even for small size genomes) and throughput
Second-generation	Illumina	MiniSeq	1.7-7.5	75 to 150	Affordable instrument pricing; Reasonable per sample cost	Run length; Read length; Low to mid sample throughput
		MiSeq	0.3-15	50 to 300	Affordable instrument pricing; Read length; Scalability; Reasonable per sample cost; Reduced run time	Low to mid sample throughput
		NextSeq	10-120	75 to 150	High Throughput	Run length; Read length
		HiSeq (2500)	10-1,000	50 to 250	Read accuracy; High Throughput	High initial investment; Run length
		NovaSeq 5000/6000	2,000-6,000	50 to 150	Read accuracy; High Throughput	High initial investment; Run length
	Ion Torrent	PGM	0.08-2	Up to 400	Read length; Speed	Throughput; Higher per sample cost
		S5	0.6-15	Up to 400	Read length; Speed	Higher per sample cost
		Proton	10-15	Up to 200	Speed; Throughput	Higher per sample cost
Third-generation	Pacific BioSciences	PacBio RSII	0.5-1	Up to 60 kb	Read length; Speed; Longer sequence reads than second generation; Does not require amplification; Sequence extended repetitive regions	High error rate initial
		Sequel	5-10	Up to 60 kb	Read length; Speed; Longer sequence reads than second generation; Does not require amplification; Sequence extended repetitive regions	High error rate
	Oxford Nanopore	MInION	0.1-1	Up to 100 kb	Read length; Speed; Longer sequence reads than second generation; Does not require amplification; Sequence extended repetitive regions	High error rate; Run length

1.5 Bacterial Genomes

The bacterial genome is organized in two cellular units, referred as chromosomes and plasmids. It carries the information that specifies the cell genetic functions, and as a consequence determines the chemical, structural, and replicative aspects of the cell. Bacterial genomes differ in a three-fold interval in terms of adenine and thymine (AT) content versus GC content. GC content can be variable within species and strains due to factors such as genome size, mutations, and environmental conditions, including temperature, aerobiosis conditions and nitrogen abundance (88).

The high variability in GC content within the bacterial genome is of enormous intrigue, since it is assumed that changes in GC content over evolutionary time are a natural response to environmental conditions in order to confer advantages (89). There is a continuum of values that range from under 25% GC content to over 73% GC content. For instance, species of *Mycoplasma* normally carry as much as 40% GC content, while *Micrococcus* species contain GC content values as high as 75%, and the genomes of *E. coli* have nearly 50% GC content (90). Genes encoding stable RNAs (e.g., ribosomal, transfer, and small regulatory RNAs) have less variation in GC content, since the product RNA molecules rely on particular nucleotide compositions and sequences in order to maintain suitable based paired secondary structures. Protein-encoding genes range far more, however the greatest variation is found in non-coding regions and in the "wobble" position, the third nucleotide position in codons, of protein-encoding genes (91).

The average GC content of free-living bacteria is higher than the average for obligate pathogens and symbionts, representing bacteria with higher GC content the Actinomycetota (formally known as Actinobacteria), Acidobacteriota (formally known as Acidobacteria), Betaproteobacteria, and Deinococcota (formally known as *Denococcus-Thermus*). On the other hand, lower GC content is predominantly associated with Campylobacteria (formally known as Epsilonproteobacteria), Bacillota (formally known as Firmicutes), Fusobacteria, and Spirochetes.

The bacterial genome is remarkably stable from one generation to the next, and yet it is plastic on an evolutionary scale. These events can be driven by deletions, duplications, amplifications, insertions, inversions and translocations, leading to changes in the amount of information contained in the genome and the appearance of new sequences at the sites of the

events (27). Because they provide raw material for the evolution of new genes, duplications are considered very important (92). Duplications and deletions result from two main mechanisms, these being slipped-strand mispairing during DNA replication or unequal crossover between two daughter sets of duplicated sequences, resulting in one chromosome harbouring a duplication and the second a deletion. Minor deletions may lead to altered gene dosage, gene product or patterns of gene expression (27,93,94).

Bacterial mutagenesis is greatly impacted by transposable elements such as insertions sequencies and transposons as they trigger mutations at a higher rate. These mobile DNA sequences have the power to restrict or allow the mobility DNA within the cell or between genomes through reactions of excision or integration (95). The vast majority of transposable elements require transposases to identify and process the end of the elements and its short terminal inverted repeats in order to be mobile (27). Additionally, they are capable of replicating themselves in genomes without the host cell's DNA (96).

While transposons carry a variety of information, such as antibiotic resistance genes, insertion sequences contain information exclusively related to their transposition and regulation (27,97).

Genome sequencing had a great impact for the understanding of the biology, diversity and evolutionary path of bacteria, since it allows us to obtain the complete information about the genetic material of a living being, as well as to identify non-essential genes and core sets of conserved genetic functions, by means of comparative genome methods (98). The core-genome is the group of genes shared between most of the individuals within a given species, but also the essential genes unique to a certain species, including those involved in translation, ribosomal structure, transcription, and replication, whereas the accessory-genome comprises genes that are only present in a fraction of the individuals belonging to a species, or even in a single individual (99). The vast majority of genes belonging to the accessory-genome are acquired through HGT and encode for functions that are not crucial for the biology of a given species (100).

Some subsets of genes are rarely gained or lost as they encode for functions that are required for growth under nearly any condition. These are referred to as essential genes and represent a group of genes indispensable for the survival of an organism in a particular environment, meaning that they are absolutely indispensable to the viability of the microorganism and that they encode for functions considered to be the basis of life (101,102).

Typically, essential genes encode conservative functional elements that mainly contribute to DNA replication, gene translation, gene transcription and substance transport, and their inactivation or loss may lead ultimately to cell death (103). In addition, essential genes are evolutionarily more conserved than non-essential genes (104), presumably as a consequence of their higher expression levels, and they constitute a persistent set of genes, which are expected to be less subject to loss or horizontal transfer. For this very reason these genes carry fewer selective advantages, thus escaping natural selection (105).

Identifying proteins that are predicted to be encoded by an organism's genome, can reveal a lot about the behaviours and habitats in which that same organism lives. The sequencing of the *B. subtilis* 168 genome defied the previous belief that the bacterium was obligate aerobic, as genes encoding a putative respiratory nitrate reductase were found and suggested its capability to grow in anaerobic conditions using nitrate as an electron acceptor (106). From this study it was also concluded that a great part of the *B. subtilis* genome encoded genes for secondary metabolites, some of which being potent compounds that inhibit fungi and bacteria, possibly helping this bacterium competing in the human body and potentially serving as probiotics.

1.6 Probiotics

Probiotics are defined as live cultures that can be consumed or applied in the body showing beneficial properties for our health when administered in adequate quantities (107,108). Probiotics are commonly used as a strategy to preserve the balance of the microbial community in our body and to help improve the microbiota in case of disturbance or disease, as they are identical or similar to the microorganisms that naturally form part of our microflora. In some cases, these microorganisms in their inactive state and their molecular components can also provide properties similar to probiotics (109–111).

A probiotic microorganism must also fulfil the following requirements: (i) be preferably isolated from the human microbiota; (ii) be able to survive the extreme conditions of the gastrointestinal system, such as acidic pH, enzyme action and exposure to bile salts; (iii) show benefits for the body; (iv) be safe for consumption and not contain vectors capable of transferring antibiotic resistance or toxicity factors (109,112,113).

Probiotics are required to be identified by their genus, species and strain designation, i.e. to be a defined entity, since there are many different probiotics being commercialized today and with different purposes (112,114). Probiotics are commonly incorporated into a range of products, including foods and functional products such as yogurts and other fermented food, dietary supplements, pharmaceutical items, and cosmetic (108,115).

These microorganisms are most commonly found in the intestine, but there are other parts of the body with the potential to isolate probiotics, such as the skin, oral cavity, lungs, urinary tract and the vagina (116). Among the many benefits that probiotics have for our body, it is important to mention the contribution they have in the digestion of food, protection against pathogenic microbes that cause diseases, synthesis of vitamins and production of substances with desirable effects, stimulation of the immune system and control of the inflammatory response, digestion and absorption of medicines (108,117). Probiotics are used in the prevention and treatment of gastrointestinal conditions such as diarrhoea, oral infections, allergic reactions and related conditions like eczema and allergic rhinitis, and more recently the possibility of using these microorganisms as adjuncts in the treatment of metabolic disorders (e.g. obesity and type 2 diabetes) has been studied (107,108,117–119).

Probiotics main mechanism of action includes colonization and balance of the intestinal microbial community, microbial antagonism and inhibition of bacteriocin synthesis,

modulation of enzymatic activity associated with the metabolization of bile salts and inactivation of xenobiotics, production of volatile fatty acids, cell adhesion and mucin production, stimulation, modulation and regulation of the immune system and interaction with the brain-gut axis by regulation of endocrine and neurologic functions (107,109). The most common microorganisms to be used in probiotic products, in mono or mixed cultures, belong to the bacterial genus *Lactobacillus*, *Bifidobacterium*, *Enterococcus* and the yeast species *Saccharomyces boulardii*, which are isolated from natural environments (120).

Advances in human microbiome research contributed for the probiotic field growth, and also to unveil that the microbiota colonizing healthy humans usually differs from the microbiota of humans with certain diseases or health conditions (108,121). Previous research has shown that patients with atopic dermatitis, when compared to healthy patients, reveal loss of local microbiota diversity, including in species such as *Corynebacterium* spp. and *Staphylococcus epidermidis*, while showing a predominance in the amount of *Staphylococcus aureus*. As for patients with seborrheic dermatitis, they tend to suffer a reduction in the amount of *Corynebacterium* spp. and an increase in the level of *Micrococcus* spp, *Pseudomonas* spp. and *Staphylococcus* spp. (122–124).

Based on the studies conducted, it is expected that manipulating the intestinal microflora will promote an increase in the number of beneficial microorganisms, having a positive impact on the immune and digestive system, metabolism and gut-brain communication, and that potential probiotic candidate species might be isolated from other sites in the body of healthy subjects - "next generation" probiotics (108).

1.6.1 Subjects of this study

1.6.1.1 *Bacillus pumilus*

B. pumilus performs a major ecological role, since it produces compounds that are antagonistic to fungi and pathogenic bacteria, like *S. aureus* and some *Vibrio* species (125). Certain strains of *B. pumilus* synthetise highly desirable antibiotics, like amicoumacin A, that targets the ribosomes of pathogenic microorganism, like some *Lactobacillales* and *Staphylococcaceae* (126). Other strains produce extracellular compounds, such as bacteriocin and surfactin, which inhibit the bacterial activity of *Staphylococcus* through disruption of the cell membrane (127).

This species has proven to be of benefit in the intestinal flora, stimulating growth performance, innate immunity, and stress tolerance (128,129). The following attributes of certain isolates have been identified as essential to their probiotic potential: (i) colonization of gastrointestinal tract through competitive exclusion (competition for adhesion sites); (ii) immunostimulation by induction of proinflammatory responses (cytokinines that increase phagocytosis; cytotoxic cells); (iii) antimicrobial activity (secretion of antimicrobial compounds such as enterotoxins) (130).

Modulation of the microbiota can also be used in order to achieve a more mature and stable configuration of our organism at an earlier stage of life, since the host's age is a determining factor for the diversity and stability of the microbiota, being these two parameters lower in an earlier stage (129). Increasing the diversity and modifying the microbiota – by introduction of bacterial populations specific to a more mature stage of life – is important for optimal host metabolism and immune development, since they present higher resilience. *B. pumilus* vegetative and spore form has been used in probiotic formulations, both benefiting from high stability to the surrounding atmospheric conditions such as heat, gastric conditions and moisture (125).

Also, *B. pumilus* SG43 can reveal red pigmentation, presumably conferred by the production of carotenoids (C30) and riboflavin. These two compounds have antioxidant properties, and their production in both vegetative and spore forms may be of great interest in

providing benefits to human health, particularly for people suffering from inflammatory bowel disease (131).

1.6.1.2 *Bacillus subtilis*

B. subtilis is a gram-positive bacterium which is non-pathogenic for mammals, including humans, with great commercial interest as a source of several secondary metabolites such as antibiotics, fine chemicals, enzymes, as well as heterologous proteins and antigens (132). This species is able of synthesizing 66 different types of antibiotics (133). *B. subtilis* has an antimicrobial effect that is enhanced by the synthesis of antimicrobial substances, antidiarrheal effect, immunostimulatory effect, and it is responsible for the competitive exclusion of pathogens, prevention of gut inflammation, and normalization of the intestinal flora (134,135). This species contains roughly 192 essential genes, involved in metabolic processes of information processing, synthesis of cell envelope and determination of cell shape and division, and cell energetics (136).

Due to its ability to metabolize a variety of sugars, a study conducted with *B. subtilis* NCIB3610 (132) investigated the probiotic potential of this strain in the treatment of oral infections caused by *Streptococcus mutans*. In the same study, strain NCIB3610 showed a more active metabolism of sugars than *S. mutans* and its potential in the treatment of oral conditions caused by this pathogenic microorganism. *B. subtilis* has also been subject of research in reducing cholesterol levels in humans, since a correlation has been observed between the modulation of the intestinal microbiota and the prevention of risks associated with cardiovascular diseases (137).

Conditions generated by climate change, such as increased environmental temperature, with more intense, frequent, and longer lasting heat waves, have been the cause of several physiological changes resulting from heat stress, such as abnormal bone remodelling and deterioration. Probiotics such as *B. subtilis* C-3102, *B. subtilis* ATCC 6051 and *B. subtilis* PB6 have proven to prevent and reduce the negative effects caused physiologically by induced stress, thereby increasing bone growth (mineral content, weight, size and length index) (138,139).

The impact that this bacterium has on our body is immense and can be enhanced when combined with other probiotic bacteria, thus helping to increase the levels of this group of

bacteria in the microbiota and to combat pathogenic microorganisms that threaten our health (140,141).

1.6.1.3 *Levilactobacillus brevis*

L. brevis is a bacterial species known for its probiotic strains, such as *L. brevis* ATCC 8287, ATCC 14869, and FFC199, and it is capable of surviving under such conditions as pH 2, being tolerant to bile salts, revealing self-aggregating and/or hydrophobic properties, and being resistant to the antibiotic vancomycin, a trait typically shown by the *Levilactobacillus* genus (142,143).

L. brevis strains reveal antagonistic activity against some well-known pathogenic bacteria, such as *B. cereus, Listeria monocytogenes*, and *S. aureus*, resulting in the improvement of intestinal health. Some of the mechanisms involved in the antagonist activity of this strain are: (i) co-aggregation, which is the ability to aggregate with organisms of the same or different strain, in this instance a pathogen; (ii) synthesis of organic acids – lactic acid and acetic acid – leading to lower intracellular pH or intracellular accumulation of these compounds in their ionized form, resulting in the death of the pathogenic microorganism; (iii) production of antibacterial peptides – bacteriocins and antimicrobial peptides – involved in pore formation and/or inhibition of cell wall synthesis; (iv) synthesis of hydrogen peroxide and other compounds such as reuterin; and (v) Competition for available nutrients and for mucosal adhesion sites, as these probiotics share carbohydrate-binding specificities with some pathogens, thus making competition for receptor sites on host cell possible (142–144).

L. brevis has the potential to compete with these pathogenic bacteria, inhibiting and displacing the adhesion of pathogenic bacteria to the cells of our organism (145). This species is involved in the regulation and treatment of conditions such as periodontitis, through inhibition of oral pathogenic microorganisms such as the species *S. mutans*, *Porphyromonas gingivalis* and *Fusobacterium nucleatum*, production of hydrogen peroxide and alleviation of inflammation, as well as through modulatory effects on the host response - decreased expression of all proinflammatory cytokines (TNF, IL-1β, among others) implicated in inflammatory periodontal bone loss - and the periodontal microbiota – inhibition of anaerobic bacteria, predominantly associated with periodontitis, and stimulation of aerobic bacteria. The release of

enzyme arginine deiminase inhibits the synthesis of nitric oxide, which is stimulated by the production of pathogenic compounds, such as liposaccharides by *P. gingivalis*, while competing with nitric oxide synthase for the arginine substrate. Nitric oxide production induces the release of pro-inflammatory cytokines, which is reversed by the enzyme arginine deiminase, resulting in anti-inflammatory effects (146). These metabolites, such as hydrogen peroxide, arginine deiminase, and lactic acid, secreted by *L. brevis* are crucial for preventing the growth of oral pathogens and for immunomodulating our system (147).

In addition, *L. brevis* KT16-2 has been studied with the aim of preventing symptoms associated with chronic autoimmune celiac disease (148). *L. brevis* KT16-2 is able to hydrolyse gluten, therefore it is a good candidate as a probiotic supplement by individuals with this condition (148). Certain supplements (e.g. fructooligosaccharides) when combined with these probiotic bacteria can enhance their viability and antioxidant activity (149,150). These fructooligosaccharides provide a protective effect to probiotics through their interaction with the phospholipids present in the cytoplasmic membrane, therefore ensuring probiotic stability and viability (151,152).

The improvement in the properties of these lactic acid bacteria, such as adherence ability and inhibition of pathogens, shows great potential in the food industry for application in fermented dairy products in order to upgrade their quality and safety (150).

L. brevis KU200019 inhibition of foodborne pathogens occurs due to competition for nutrients and the binding sites on the epithelial cells, but also through metabolism of oligosaccharides with antagonistic action against pathogenic microorganisms, inhibiting their growth (153). In addition, these compounds – oligosaccharides – exhibit anti-adhesive activity, by mimicking the host cell receptor sites and thus preventing the binding of pathogenic microorganisms to the cells of the intestinal epithelium (150). The ability of *L. brevis* KU200019 to adhere the intestine is enabled by the fermentation of these oligosaccharides, resulting in short-chain fatty acids production, and supplying the energy needed for probiotic proliferation (150).

1.6.1.4 *Limosilactobacillus reuteri*

L. reuteri is a bacterial specie that naturally resides in the intestinal microbiota of humans and animals, and it can also be discovered in the urinary tract, skin, and even breast milk (154). In order for some *L. reuteri* stains to show their probiotic potential and provide health advantages to the consumer, these bacteria must interact with the host microbiota (155). As an example, a previous study (156) shown that *L. reuteri* I5007 influenced the gastrointestinal flora by lowering the pH level, which allowed beneficial bacteria in the body to thrive, such as *Limosilactobacillus vaginalis* and *Bifidobacterium* spp., and helped to reduce pathogenic bacteria, such as *E. coli*. This effect on the microbial community was mediated by the production of lactic acid.

L. reuteri is responsible for the synthesis of antimicrobial compounds such as organic acids and ethanol, but also converts glycerol into reuterin, a compound with a broad spectrum of antibacterial, antifungal, antiviral and antiprotozoal activity, synthesized under the pH, temperature and anaerobic conditions present in the gastrointestinal system (154,156). For probiotics it is essential for them to adhere onto the intestine epithelium cells, and this feature tends to vary between strains depending on the amount of glyceraldehyde-3-phosphate dehydrogenase that this microorganism has in its cell wall. For instance, *L. reuteri* ZJ615 shows less adhesion to intestinal epithelial cells when compared to *L. reuteri* ZJ617 (157). A higher abundance of glyceraldehyde-3-phosphate dehydrogenase in this probiotic, means a higher capacity to adhere to the mucin present in the intestine, thus preventing the development of serious conditions at the intestinal tract (158,159).

This species has some bacterial strains, that have shown to be helpful suppressing the severity of enteric infections and modulate the immune system, but also have proven to be of great interest in the treatment of conditions such as diarrhoea, infantile colic, eczema, oral conditions and *H. pylori* infections, being these strains *L. reuteri* ATCC 55730, ATCC PTA 4659, ATCC PTA 5289, ATCC PTA 6475 and DSM17938 (154). Since *L. reuteri* DSM 17938 and ATCC PTA 5289 have proven to be helpful minimizing epithelial damage, oxidative stress, and inflammation of the oral mucosa normally induced by chemotherapy, their therapeutic potential in treating conditions like oral mucositis has shown promise (160).

The following table (Table 1.3) provides information concerning the genome of our four species of study mentioned above, *B. pumilus*, *B. subtilis*, *L. brevis* and *L. reuteri*, based on reference strains found in published articles (161–165).

Table 1.3 Overview between reference genomes of four different species.

Bacterial species	***B. pumilus* SAFR-032**	***B. subtilis* 168**	***L. brevis* BSO 464**	***L. reuteri* SD2112**
Genome length (bp)	3,704,641	4,214,547	2,503,991	2,264,399
GC ratio content (%)	41.3	43.5	45.9	39.0
Coding DNA sequences (CDSs) (Total)	3,679	4,320	2,627	2,236
Coding DNA sequences (CDSs) (with protein)	3,617	4,125	2,533	2,127
Transfer RNAs (tRNAs)	72	86	49	70
Number of chromosomes	1	1	1	1
Number of Plasmids	0	0	8	4
References	(166)	(167)	(164,168)	(169)

Chapter 2: Objective

2.1 Aim of the study

The aim of this project was to develop a set of primers capable of detecting specific probiotic strains belonging to the four species mentioned above, *B. pumilus*, *B. subtilis*, *L. brevis* and *L. reuteri*, so it can be used in studies focused on the detection of such specific strains in biological samples.

Chapter 3: Materials and Methods

3.1 Genome selection for species *Bacillus pumilus*, *Bacillus subtilis*, *Levilactobacillus brevis* and *Limosilactobacillus reuteri*

Representative genomes for each target species were selected with the assistance of the software Geneious (Biomatters Ltd., Boston, United Stated of America). The selection was made using the NCBI's Genome database by collecting all complete genomes available for each bacterial species. Out of the set of genomes obtained, a single one was selected for each target species, based on the most frequently referenced species in scientific articles and was shown to be a model to be used in this study but also for presenting identical characteristics such as genome size, GC content and normal CDSs associated to each species, as referenced in Table 1.3.

3.2 Screening for genes

The selection for target genes was made using the software Geneious and the search conducted in the NCBI's Nucleotide database. Filters were applied at the level of the nucleotide sequence length, in order to obtain genomic fragments with length between 500 bp and 10,000 bp for each of the target species considered in the study, *B. pumilus, B. subtilis*, *L. brevis* and *L. reuteri*.

Following the search in Geneious software, genes were filtered for each of the four target species, based on previous research carried out to identify the most referenced genes when searching each one of our species. Housekeeping genes were also selected for each of the species studied in this project, as they are essential genes in basic bacterial functions, allowing them to survive, and also because they can be found in all species, making them good targets for our study.

3.3 Selection of polymorphic genes

From the obtained sequences of this research, their alignment ("Multiple align") was performed using the Geneious software, in which the only sequences selected to be aligned were those that belonged to our study species and which showed to have nucleotide information available when searched.

The previously obtained fragments, belonging to our genes of interest, were aligned with the reference genomes in order to analyse differences and similarities between them. For this study only fragments with polymorphisms were selected, therefore the polymorphic regions were identified in each of the previously aligned fragments, and those that showed these variations were selected as having potential. These fragments were later evaluated to determine if they contained any sequences that met the necessary conditions for the design of primers.

3.4 Primers

Following the selection of fragments with polymorphic areas, the sequences were evaluated in order to determine whether they had sequences with potential for the design of primers. This evaluation was carried out according to the following characteristics, which in case they occur in the identified sequences, turn them into potential primers. The following criteria were used: (i) primer length between 18 and 24 bp; (ii) melting temperature (Mt) ideally between 50 and 65ºC, and with 5ºC in between primers; (iii) GC content between 40 and 60%; (iv) no secondary structures formed, namely self-dimers and hetero-dimers; (v) no hairpins formation.

As a desirable but not mandatory requirement, primers should have one to three guanines or cytosines in the last five nucleotides of the fragment to promote specific binding and have a guanine or cytosine as the last nucleotide to make it more stable (GC clamp).

Chapter 4: Results

4.1 Selection of four relevant strains for the target species

Geneious software collected 214, 839, 242 and 407 genomes for the species *B. pumilus, B. subtilis*, *L. brevis* and *L. reuteri*, respectively. For each species, and based on the methodology previously described for selecting a representative genome, were obtained a set of genomes belonging to the strains *B. pumilus* SAFR-032, *B. subtilis* 168, *L. brevis* KB290 and *L. reuteri* I5007. A total of 9, 1, 10 (nine of them plasmids) and 7 (six of them plasmids) sequences were gathered for *B. pumilus* SAFR-032, *B. subtilis* 168, *L. brevis* KB290 and *L. reuteri* I5007 respectively, and for each strain was selected a genome, which represent chromosomal DNA.

General details on the genetic content of the four strains selected for our study, gathered from GenBank using Geneious software, are presented below in Table 4.1.

Table 4.1 Overview between the four strains selected for our target species.

Bacterial strains selected	***B. pumilus* SAFR-032**	***B. subtilis* 168**	***L. brevis* KB290**	***L. reuteri* I5007**
Genome length (bp)	3,704,641	4,215,606	2,395,134	1,947,706
GC ratio content (%)	41.3	43.5	46.1	39.0
Coding DNA Sequences (CDSs)	3,614	4,174	2,535	2,009
Accession number from NCBI	NC_009848	NC_000964	NC_020819	NC_021494

4.2 Screening for relevant genes within our target species

Following the analysis conducted on Geneious software and relevant articles associated with our study species, specific genes were selected for *B. pumilus*, *B. subtilis*, *L. brevis* and *L. reuteri*. These conserved genes were identified using references for each of the species, making them suitable for comparison between strains. Tables 4.2 to 4.5 show these genes for the four species, as well as information on the products they encode and their function.

Table 4.2 Selection of genes for *Bacillus pumilus*.

Gene	Encodes	Activity	References
aprE	Subsilisin E	Extracellular alkaline protease gene responsible for more than 70% of the total extracellular proteolytic activity	(170,171)
aroE	3-phosphoshikimate 1-carboxyvinyltransferase	Biosynthesis of aromatic amino acids	(172)
gyrB	DNA gyrase subunit B	Type II topoisomerase that negatively supercoils closed circular double-stranded DNA to modulate DNA topology and maintain chromosomes in a wounded state, controlling the replication initiation, transcription and chromosome segregation	(173)
mutL	DNA mismatch repair protein	Recognizes and repairs mismatched bases in new synthesized DNA daughter strands	(174)
pycA	Pyruvate carboxylase	Conversion of pyruvate into oxaloacetate	(88,175)
pyrE	Orotate phosphoribosyltransferase	Catalyses the transfer of a ribosyl phosphate group from 5-phosphoribose 1-diphosphate to orotate, leading to the formation of orotidine monophosphate	(176,177)
rpoB	RNA polymerase beta subunit	It promotes the transcription of DNA into RNA through the four ribonucleoside triphosphates used as substrates	(178)
trpB	Beta subunit of tryptophan synthase	Essential in the final stage of tryptophan synthase, where it converts idole and L-serine into tryptophan	(179)
xynB	Beta xylosidase	Xylan-degrading intracellular enzyme	(180)

Table 4.3 Selection of genes for *Bacillus subtilis*.

Gene	Encodes	Activity	References
comA	Transcriptional regulatory protein *comA*	Response regulator of the *comP*/*comA* regulatory system which is involved in a major quorum response pathway that controls the development of genetic competence	(181)
comC	Late competence gene	It is necessary for the development of genetic competence	(182)
comP	Sensor histidine kinase *comP*	It is a sensor in the *comP*/*comA* regulatory system that performs an important role in sporulation, which is partially interchangeable with the function of SpoIIJ	(183)
comQ	A protein linked to cell communication and the process of competence for genetic transformation	It is important for regulating cellular competence, thus allowing the bacteria to absorb exogenous DNA from the environment	(182)
csaA	*CsaA* chaperone	Molecular chaperon involved in protein secretion	(184)
degS	Signal transduction histidine-protein kinase/phosphatase *degS*	It is part of the *degS*/*degU* regulatory complex, crucial in the transitional growth phase. In addition, it controls the expression of various cellular activities, including the production of degradative enzymes such as neutral and alkaline proteases, as well as the formation of flagella and biofilms	(185)
degU	Transcriptional regulatory protein *degU*	It is part of the *degS*/*degU* two-component regulatory system and is involved in processes such as exoprotease production, genetic competence and motility	(186)
dnaK	Chaperon protein *dnaK*	Involved in stress response, especially heat shock, and assists in the proper folding of proteins as well as in the prevention of the aggregation of misfolded proteins under stressful conditions	(187)
groEL	*GroEL* chaperone	Together with the co-factor *groES*, it forms the chaperon *groE*	(188)
gyrA	DNA gyrase subunit A	Type II topoisomerase that negatively supercoils closed circular double-stranded DNA and also regulates the initiation of DNA replication	(189)
gyrB	DNA gyrase subunit B	Type II topoisomerase that negatively supercoils closed circular double-stranded DNA to modulate DNA topology and maintain chromosomes in a wounded state, controlling the replication initiation, transcription and chromosome segregation	(173)
pgsA	Phosphatidylglycerophosphate synthase	*PgsA* is an essential protein involved in the synthesis of phosphatidylglycerol	(190)
ptsG	Glucose permease of the phosphotransferase system	Transport of glucose as well as phosphorylation, and control of *glsT* activity	(190)
rpoB	RNA polymerase beta subunit	It promotes the transcription of DNA into RNA through the four ribonucleoside triphosphates used as substrates	(191)

Table 4.4 Selection of genes for *Levilactobacillus brevis.*

Gene	Encodes	Activity	References
dnaK	Chaperone protein *dnaK*	Involved in stress response, especially heat shock, and assists in the proper folding of proteins as well as in the prevention of the aggregation of misfolded proteins under stressful conditions	(187)
gadB	Glutamate decarboxylase	Catalyses the conversion of glutamate into gamma-aminobutyric acid	(192)
gadC	Gamma-aminobutyric acid antiporter	Enhances the availability of glutamate and, consequently, the efficiency of the glutamate decarboxylase system	(192)
gyrB	DNA gyrase subunit B	Type II topoisomerase that negatively supercoils closed circular double-stranded DNA to modulate DNA topology and maintain chromosomes in a wounded state, controlling the replication initiation, transcription and chromosome segregation	(193)
horA	Multidrug transporter *horA*	Encodes a multidrug transporter protein with a key role in hop resistance, that protects the bacteria from antimicrobial effects	(194)
recA	Recombinase *recA*	Encodes a DNA recombination protein of the repair system	(195,196)
rpoA	RNA polymerase alpha subunit	Essential in bacterial transcription	(197)

Table 4.5 Selection of genes for *Limosilactobacillus reuteri.*

Gene	Encodes	Activity	References
ddl	D-alanine-D-alanine ligase	Enzyme involved in peptidoglycan biosynthesis	(198)
dltA	D-alanyl-D-alanine carrier protein ligase	Involved in the D-alanylation of teichoic acids in the cell wall of many gram-positive bacteria and modifies lipoteichoic acids by integrating D-alanine, affecting the overall charge of the cell wall	(199)
fabZ	Beta-hydroxyacyl-ACP-dehydratase	Regulates the activity of 3-hydroxyacyl-ACP dehydratase, the step involving the dehydration of fatty acid in their elongation, thus promoting the conversion of 3-hydroxyacyl-ACP into enoyl-ACP	(200)
glnQ	Protein belonging to the glutamine transport system	It is involved in the active transport of glutamine across the cell membrane and contributes to the uptake of this amino acid by the bacteria	(200,201)
gyrB	DNA gyrase subunit B	Type II topoisomerase that negatively supercoils closed circular double-stranded DNA to modulate DNA topology and maintain chromosomes in a wounded state, controlling the replication initiation, transcription and chromosome segregation	(193)
leuS	Leucyl-tRNA synthetase	It is responsible for binding the amino acid leucine to its cognate tRNA molecules during protein synthesis	(202)
pkt	Phosphoketolase	Catalyses the conversion of xylulose 5-phosphate or fructose 6-phosphate into glyceraldehyde 3-phosphate and acetylphosphate, a crucial reaction in the phosphoketolase pathway for sugar metabolism	(202)
recA	Recombinase *recA*	Encodes a DNA recombination protein of the repair system	(195,196)

Once the genes for each species were selected, the NCBI Nucleotide database was used to search for fragments that were associated with the target species and the respective genes chosen for each one, resulting in the following number of samples displayed in Tables 4.6 to 4.9. In these same tables it is possible to see the number of fragments obtained by the program, some of which were not described as being from the desired species or gene.

In some instances, although the number of fragments obtained for a given gene was significant, they displayed many nucleotide sequences with no information available in the NCBI Nucleotide database. This limitation affected, in particular, the fragments obtained for the *comQ* gene from *B. subtilis*, as well as the *rpoB* and *gyrB* from *B. pumilus*. As a result, these genes were excluded from subsequent stages of the study.

Table 4.6 Number of sequences obtained using Geneious software, after searching the selected genes for *Bacillus pumilus.*

Research Topic	Found Sequences	Sequences with no Visible Nucleotide Information	Sequences of the Selected Species	Sequences from other Species	Sequences of the selected species and gene
aprE	3	0	1	2	1
aroE	82	0	19	63	19
gyrB	230	1	136	94	127
mutL	83	4	20	63	16
pycA	79	0	16	63	16
pyrE	81	1	18	63	17
rpoB	170	1	70	100	66
trpB	80	1	17	63	16
xynB	4	0	3	1	3

Table 4.7 Number of sequences obtained using Geneious software, after searching the selected genes for *Bacillus subtilis.*

Research Topic	Found Sequences	Sequences with no Visible Nucleotide Information	Sequences of the Selected Species	Sequences from other Species	Sequences of the selected species and gene
comA	23	6	17	6	5
comC	12	5	9	3	3
comP	69	6	66	3	46
comQ	68	5	66	2	49
csaA	1511	186	13	1498	1
degS	13	7	10	3	5
degU	15	9	9	6	5
dnaK	180	87	82	98	15
groEL	163	16	99	64	84
gyrA	859	134	613	246	558
gyrB	423	131	244	179	182
pgsA	116	95	30	86	7
ptsG	96	89	8	88	2
rpoB	521	81	345	176	329

Table 4.8 Number of sequences obtained using Geneious software, after searching the selected genes for *Levilactobacillus brevis.*

Research Topic	Found Sequences	Sequences with no Visible Nucleotide Information	Sequences of the Selected Species	Sequences from other Species	Sequences of the selected species and gene
dnaK	13	2	13	0	11
gadB	8	0	7	1	7
gadC	6	0	6	0	5
gyrB	28	0	28	0	27
horA	4	0	4	0	4
recA	45	6	21	24	19
rpoA	18	1	18	0	17

Table 4.9 Number of sequences obtained using Geneious software, after searching the selected genes for *Limosilactobacillus reuteri.*

Research Topic	Found Sequences	Sequences with no Visible Nucleotide Information	Sequences of the Selected Species	Sequences from other Species	Sequences of the selected species and gene
ddl	133	0	133	0	133
dltA	165	24	165	0	133
fabZ	68	48	68	0	0
glnQ	22	7	22	0	0
gyrB	152	17	152	0	134
leuS	144	11	144	0	133
pkt	133	0	133	0	133
recA	156	15	156	0	133

For each of the target species, were selected the genes that had a significant number of fragments, which were the *aroE*, *gyrB*, mutL, *pycA*, *pyrE*, *rpoB* and *trpB* genes from *B. pumilus*, *comP*, *comQ*, *dnaK*, *groEL*, *gyrA*, *gyrB* and *rpoB* from *B. subtilis*, *gyrB*, *recA* and *rpoA* from *L. brevis* and *ddl*, *dltA*, *gyrB*, *leuS*, *pkt* and *recA* from *L. reuteri*.

4.3 Promising fragments suitable for primers

Once selected the target genes, 19, 127, 16, 16, 17, 66, 16, 46, 49, 15, 84, 558, 182, 329, 27, 19, 17, 133, 133, 134, 133, 133 and 133 fragments were aligned according to the previous search performed for "*B. pumilus aroE*", "*B. pumilus gyrB*", "*B. pumilus mutL*", "*B. pumilus pycA*", "*B. pumilus pyrE*", "*B. pumilus rpoB*", "*B. pumilus trpB*", "*B. subtilis comP*", "*B. subtilis comQ*", "*B. subtilis dnaK*", "*B. subtilis groEL*", "*B. subtilis gyrA*", "*B. subtilis gyrB*", "*B. subtilis rpoB*", "*L. brevis gyrB*", "*L. brevis recA*", "*L. brevis rpoA*", "*L. reuteri ddl*", "*L. reuteri dltA*", "*L. reuteri gyrB*", "*L. reuteri leuS*", "*L. reuteri pkt*" and "*L. reuteri recA*", respectively. These fragments were aligned with the reference genomes of each species, *B. pumilus* SAFR-032, *B. subtilis* 168, *L. brevis* KB290 and *L. reuteri* I5007, as shown by the examples in Figures 4.1 to 4.4.

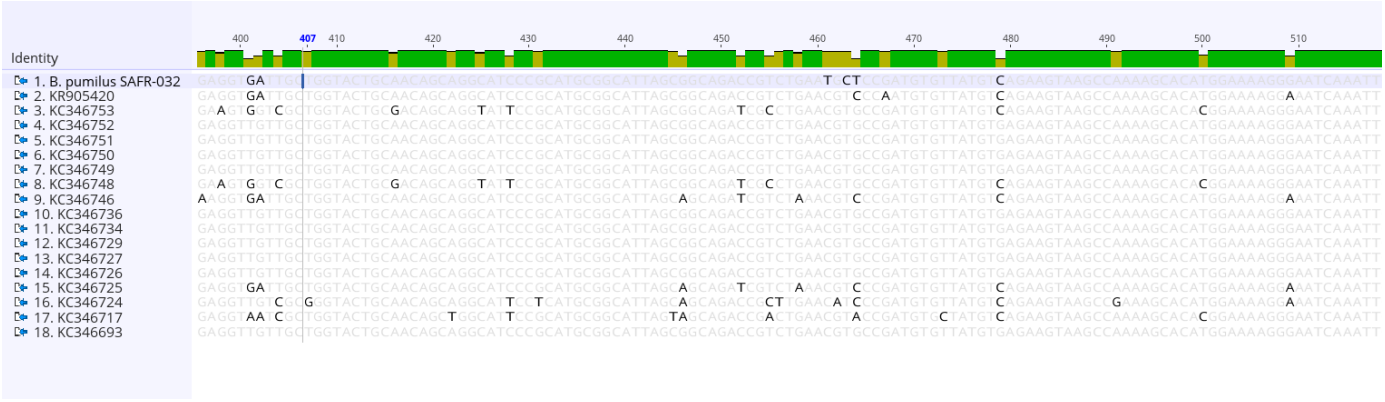

Figure 4.1 Alignment of 17 fragments of the *pyrE* target gene with the reference genome of *Bacillus pumilus* SAFR-032.

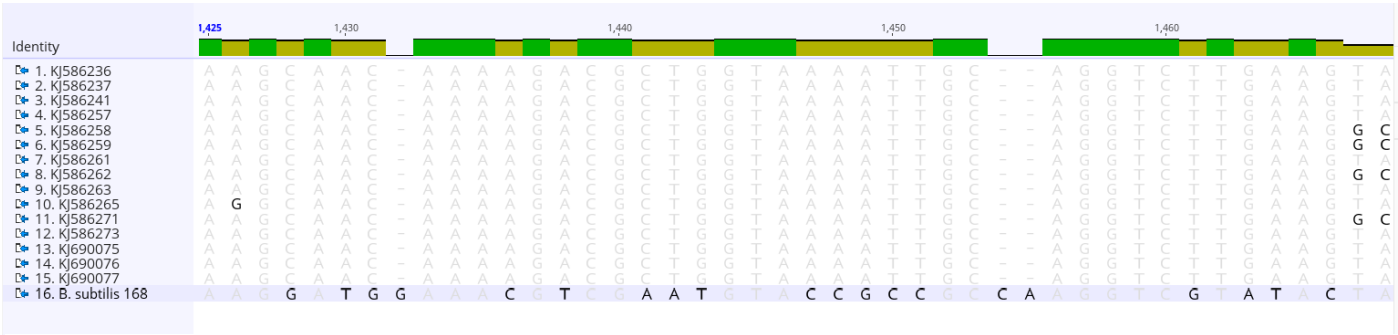

Figure 4.2 Alignment of 15 fragments of the *dnaK* target gene with the reference genome of *Bacillus subtilis* 168.

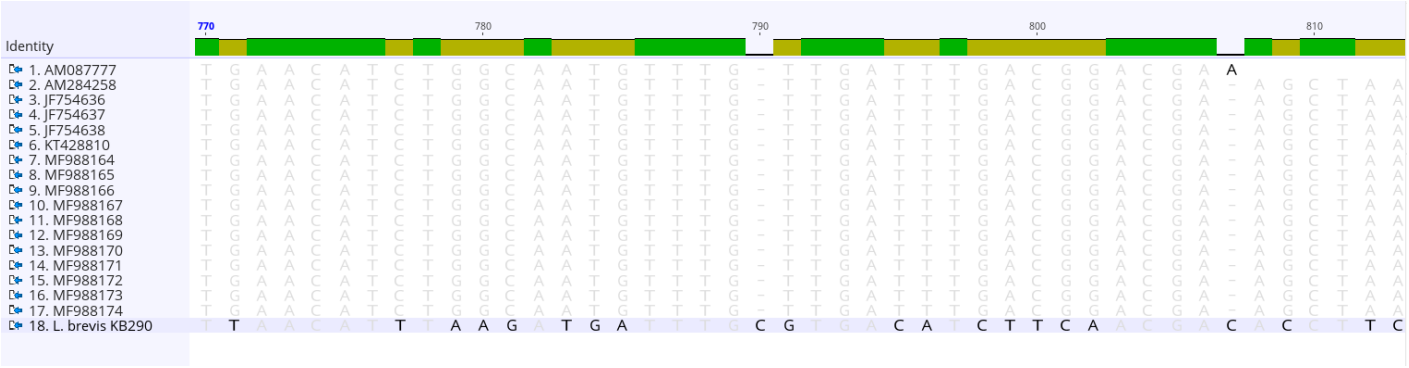

Figure 4.3 Alignment of 17 fragments of the *rpoA* target gene with the reference genome of *Levilactobacillus brevis* KB290.

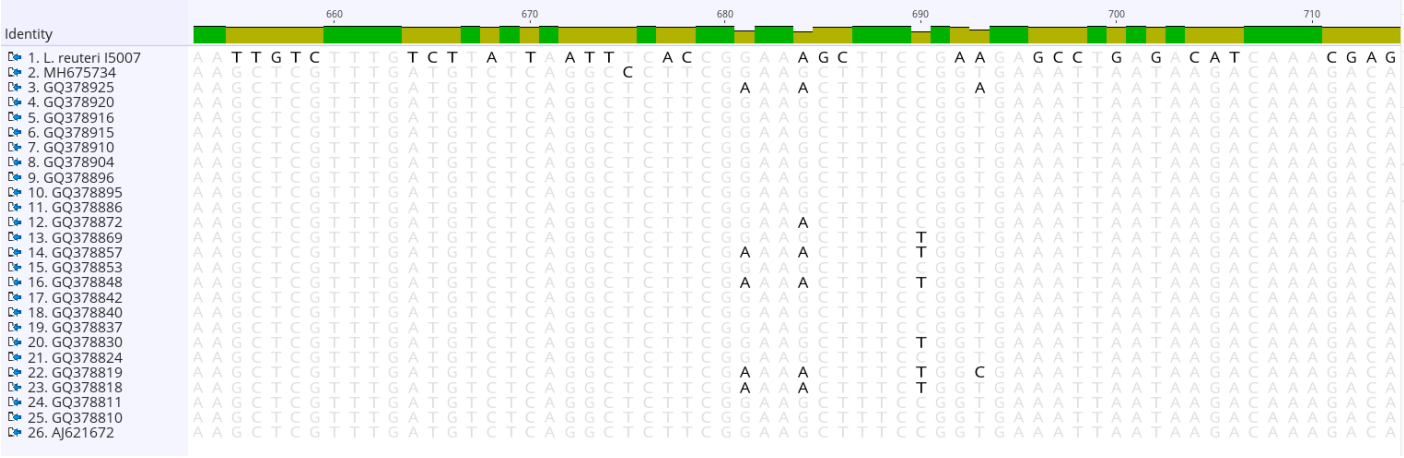

Figure 4.4 Alignment of 25 unique sequence fragments, out of a total of 133 fragments, of the *recA* target gene with the reference genome of *Limosilactobacillus reuteri* I5007.

After aligning these sequences, the identification of non-conserved regions in each reference genome, meaning regions that accumulate enough changes to specifically detect one strain and not another, was performed, since these regions could be useful and have potential for the design of our primers. In the Appendix 1 to 4, were summarized for each strain of the target species, sequences with potential for the design of our primers. Each of these fragments was selected from a panel of many others, and the ones listed in the following four tables showed that they had potential to be a primer for our study strains, according to the parameters presented in section 3.4 of this document, which summarize the criteria for a good primer selection. In this stage of the project, the parameters that were evaluated for the selection of fragments with potential for primers, and which are summarized in the tables below, were the length of the primer, value of the melting temperature, GC content, the formation of self-dimers and hairpins, with these last parameters being evaluated using the OligoAnalyzer^{TM} Tool from Integrated DNA Technologies, Inc. (IDT^{TM}).

4.4 Primer design for target strains

Among the fragments previously selected as having potential for primers, were highlighted in the tables below (Table 4.10 to 4.12) those that operate better collectively and which meet the criteria mentioned in Chapter 3.4 of this document. For each of the strains studied in this project, a set of primers specific to them was created, making it possible to uniquely identify each one of them. In this step of the project, was tested the possibility of forming hetero-dimers between the primers proposed in Appendix 1 to 4, with only those in the tables below showing a lower probability of forming these bonds or causing any problem during the amplification process. This last variable was calculated using the OligoAnalyzer^{TM} Tool from Integrated DNA Technologies, Inc. (IDT^{TM}).

Table 4.10 Selected primers for the target genes of *Bacillus pumilus* SAFR-032 and corresponding length, melting temperature and GC content value of each primer.

Gene	Primers	Primer length (bp)	Melting Temperature (ºC)	GC content (%)
gyrB	5' TATGACTGTTCTTCACGCTGGC 3'	22	57.3	50.0
	3' CCTACCTCGCGTGTAAGCTTGC 5'	22	60.9	59.1
mutL	5' TCCCGCTTCCGTTGTAAAAG 3'	20	62.6	50.0
	3' GAATGCAGTGGTTCCTTTTCAG 5'	22	61.1	45.5
	3' CGATAAATACCCTGTCGTCA 5'	20	59.8	45.0
pycA	5' GACACCTTCTTCTAAAGTAGTC 3'	22	57.9	40.9
	5' ACTCATTGTGAAACTCGTGTCA 3'	22	62.4	40.9
pyrE	5' TGCTCCGATGTGTTATGTC 3'	19	59.5	47.4
	3' CCATCAGCAGCATCTTCTAGAT 5'	22	60.0	45.5
	3' CCTACCTTCTTTGGCCAGATT 5'	20	59.3	45.0

For the *B. pumilus* SAFR-032 strain, a set of ten primers was selected, five of which being forward primers and the other five reverse primers, with a length ranging from 19 bp to 22 bp, therefore fulfilling the appropriate primer length to avoid any undesired results during product amplification. As for GC content, the primers presented values between 40.9% and 59.1%, which were within the ideal values mentioned in Chapter 3.4, and the majority of them contained two to three cytosine or guanine nucleotides at the end of the chain, giving them greater binding strength to the DNA strand to be amplified. Regarding the melting temperature of the primers, the values ranged from 57.3ºC to 62.6ºC, indicating that they had an ideal and similar Mt, with a temperature variation of approximately 5ºC between the ten primers.

Table 4.11 Selected primers for the target genes of *Bacillus subtilis* 168 and corresponding length, melting temperature and GC content value of each primer.

Gene	Primers	Primer length (bp)	Melting Temperature (ºC)	GC content (%)
dnaK	5' TTTCGCTTGCAGAAAGTCC 3'	19	53.5	47.4
	5' TCTTGTGGAGTGTAATCCTTTC 3'	22	52.3	40.9
	5' GTTCGCCGTTTTTAAATGCA 3'	20	52.4	40.0
	3' CCCAAACATTATCTAACCGCAA 5'	22	53.4	40.9
	3' CAAACAACAAGGATTCAGCTAAGG 5'	24	54.1	41.7

Regarding the *B. subtilis* 168 strain, a set of five primers was picked, three of which were forward primers and two reverse primers, presenting a length that varied from 19 bp to 24 bp, thus fulfilling the appropriate length required for a good primer. As for CG content, they showed levels ranging from 40.0% to 47.4%, which is in line with the ideal values mentioned in Chapter 3.4, and most of these primers presented two to three cytosine or guanine nucleotides at the end of the chain, conferring greater binding strength to the DNA chain intended to be amplified. When it comes to the melting temperature of the primers, the values fluctuated between 52.3ºC and 54.1ºC, showing that they had an ideal and similar Mt between them, with the temperature variation between the five primers being less than 2ºC.

Table 4.12 Selected primers for the target genes of *Limosilactobacillus reuteri* I5007 and corresponding length, melting temperature and GC content value of each primer.

Gene	Primers	Primer length (bp)	Melting Temperature (ºC)	GC content (%)
ddl	5' AACTTCCCACATTTGAGGG 3'	19	52.5	47.4
	5' GGTACACGAACAGCTCCT 3'	18	54.2	55.6
	5' TTCTTGAGGATTGGCGAT 3'	18	51.1	44.4
	3' TTACGGAACATACGGAAG 5'	18	50.2	44.4
recA	5' GGCTTCTGCATATTGAGGATCA 3'	22	54.8	45.5
	3' AAGGTTATTGTAGGCATGGT 5'	20	51.1	40.0
	3' TTCTCGGACTCTGTAGTTTG 5'	20	52.4	45.0
	3' GACGCTTTCGTTATCCAGTAGTTA 5'	24	54.1	41.7

Lastly, the *L. reuteri* I5007 was analysed by selecting a set of eight primers, four of which were forward primers and four reverse primers, presenting a length ranging from 19 bp to 24 bp and thus meeting the criteria required for a suitable primer length. In terms of GC content, the primers exhibited a percentage that ranged from 40.0% to 55.6%, thus falling within the ideal values mentioned in Chapter 3.4, and some of them contained two to three cytosine or guanine nucleotides at the end of the chain, thus making them stronger in binding to the DNA chain to be amplified. Regarding the melting temperature of the primers, the values varied between 50.2ºC and 54.8ºC.

Chapter 5: Discussion

5.1 Probiotic applications

For the present study, four strains of different bacterial species exhibiting probiotic potential were selected, being them *B. pumilus* SAFR-032, *B. subtilis* 168, *L. brevis* KB290 and *L. reuteri* I5007. These bacterial strains can contribute to human health. A study shown that strain *B. pumilus* SAFR-032 was able to produce an antimicrobial agent with effect on *E. coli* ATCC 10536, a member of one of the most common pathogenic bacterial species responsible for causing infections in the urinary tract of children and adults (203). The designation of this antimicrobial agent was not specified. The ammonium sulphate method known for extracting protein bioactive agents, was used to isolate it, and it was possible to infer that it contained a peptide chain in its structure or a particular unit with amino acids joined by peptide bonds.

Meanwhile, *B. subtilis* 168 is known for producing sublancin 168, a bacteriocin with strong antimicrobial activity against certain gram-positive bacteria, including *Priestia megaterium* (formally known as *Bacillus megaterium*) and *Streptococcus pyogenes* (204). *P. megaterium* is a bacteria that usually does not reveal significant pathogenicity, and that can be discovered in environments such as in food (e.g. honey) and soil. However, although it is uncommon, some studies report the pathogenic character of this bacteria and associate it with some cases of human infections, as mentioned in the following research. In addition to some cases of bacteraemia, meningitis, and other conditions, a case study also reported a soft tissue infection caused by this bacteria after a potential wound developed by a patient (205). In contrast, *S. pyogenes* is a pathogenic bacteria responsible for causing a wide range of infections, from mild and common, such as pharyngitis, impetigo and scarlet fever, to more serious cases of infection with high mortality rates, such as septicaemia, necrotizing fasciitis and streptococcal toxic shock syndrome (206,207). *B. subtilis* 168 has also shown its potential, in experiments carried out with *Drosophila melanogaster* and *in vitro* cell cultures, in the treatment of CaOX nephrolithiasis, commonly known as kidney stones, as it is capable of degrading oxalate, the cause of 80% of cases concerning kidney stones, and preventing the aggregation and adhesion of CaOX crystals to the cells of the kidney epithelium (208).

A third study (209) aimed to confirm that the oral administration of probiotic *Bacillus* species promoted the growth of *Lactobacillus* in the intestinal tract, by co-culturing *in vitro* strains of *L. plantarum* SDMCC050204-pL157 with *B. subtilis* 168. Strains of *L. brevis* have a high sensitivity to reactive oxygen species because they lack an efficient antioxidant system which is normally found in aerobic microorganisms. So, these bacteria carry a gene that encodes a heme-dependent catalase, which catalyses the degrading of hydrogen peroxide from reactive oxygen species, despite the fact that these microorganisms do not have the ability to synthesize heme and for this reason they are defined as catalase-negative strains. For this experiment, in which the conditions of the intestinal tract were simulated, an autolysate was added to the cultures of *L. plantarum* SDMCC050204-pL157, minimizing the hydrogen peroxide levels, protecting the DNA integrity of *L. plantarum* SDMCC050204-pL157 against oxidative stress and increasing the survival rate of this strain. The heme group, formed by the autolysis of *B. subtilis* 168, activated the heme-dependent catalase - *katA* - in *L. plantarum* SDMCC050204-pL157, improving its ability to catalyse hydrogen peroxide and thus survive oxidative stress.

This study proposed a model in which, after oral administration of the probiotic *B. subtilis* 168, this strain undergoes a process of autolysis in the intestine, releasing the heme group and possibly other factors that promote the growth and resilience of *Lactobacillus* strains, and the benefits of its administration extend beyond the direct effects, playing a supporting role by interacting with other probiotic strains.

Previous observations have indicated that *L. brevis* KB290 may be effective in the treatment of a wide range of health conditions, being some of them obesity and gastrointestinal disorders, such as irritable bowel syndrome. As for the case of obesity, supplementation with heat-killed *L. brevis* KB290, a paraprobiotic, helped to reduce indicators of obesity, such as the increased density of adipose tissue caused by a unhealthy diet and high levels of glucose in the system, and for irritable bowel syndrome, this probiotic significantly relieved abdominal pain and modified the composition of the intestinal microflora of patients after taking *L. brevis* KB290, showing a decrease in *Clostridium*, which is normally highly prevalent in irritable bowel syndrome patients, and an increase in *Bifidobacterium*, which is highly prevalent among healthy individuals (210). The difference between a paraprobiotic and a probiotic is that the first one refers to a non-viable/inactive microbial cell or a fraction of it derived from a probiotic bacterium, which confers health benefits to its consumer, when administered in the indicated doses, similar to those provided by a probiotic. Paraprobiotics have emerged as an alternative to probiotics since they bring the following advantages: they are more stable and safer to be

consumed, they have a longer shelf life, they are easier to store, the doses administered are easier to control and they are compatible with antibiotics and antifungals (211,212). *L. brevis* KB290 exhibits immunomodulatory properties that can be enhanced against influenza virus infections. Research results have shown that oral administration of the probiotic *L. brevis* KB290 plays a significant role in mitigating some of the symptoms induced by influenza virus infection, such as loss of body mass, helping increase the levels of IFV-specific immunoglobulin A and raising the synthesis of IFN-α induced by viral infection (213).

Multiple studies support the probiotic effects of *L. reuteri* I5007 on gut health, including enhancement of gut barrier integrity, modulation of immune responses, improvement of intestinal morphology, and reduction of inflammation in both animal and cell culture models.

A study has revealed that pre-treatment with *L. reuteri* I5007 helped to prevent the development of colitis in mice, and the impact was attributed to changes in the intestinal microbiota and metabolites, and their roles in improving colonic function (214). Tests performed *in vivo* with rats treated with dextran sulphate sodium, a compound commonly used to induce colitis in animal models, revealed that the *L. reuteri* I5007 pre-treatment improved the proper functioning of the colon by maintaining its consistency and minimizing the loss of length induced by treatment with dextran sulphate sodium. It also helped significantly in decelerating the weight loss of the subjects, when compared to other treated solely with dextran sulphate sodium who showed weight loss and helped recovering the number of globlet cells and reducing ulceration of the mucosa. Observation of mice colon treated with dextran sulphate sodium revealed severe inflammation of the mucosa and the near loss of this layer, when compared to the test carried out after pre-treatment with *L. reuteri* I5007 which showed an inhibition of this inflammatory response of decreased mucosal ulceration, epithelial oedema and crypt loss, and recovered globlet cells and MUC-2 expression.

In the same study, an *in vitro* test was undertaken with human colon cell line HT-29, exposed to lipopolysaccharides in order to induce an expression of inflammatory cytokines. The pre-treatment with *L. reuteri* I5007 diminished the pro-inflammatory cytokines IL-1β and TNF-α levels, which are associated with the pathogenic potential of IBD, while promoting the augmented expression of IL-10 (214). Additionally, researchers observed a modulatory effect on the gut microbiota, as pre-treatment with *L. reuteri* I5007 led to a decrease in the abundance of the *Enterobacteriaceae* family and the *Escherichia-Shigella* genus, both reported to be connected with individuals with IBD in other studies.

In a previous study (215), scientists investigated the ability of *L. reuteri* I5007 to support and protect the intestinal epithelium by modulating tight junction proteins, claudin-1, occludin and ZO-1, through an *in vitro* experiment with an IPEC-J2 cell line and an *in vivo* experiment with newborn piglets. These proteins were shown to be responsible for preserving the gut mucosal barrier structure and for regulating the transport of ions, solutes and water across it, but also for preventing and treating several illnesses regarding the intestine, through augmentation and protection of the gut barrier, when modulated by probiotics. *L. reuteri* I5007 showed its probiotic effect by upregulating tight junction protein expression (claudin-1, occluding and ZO-1), preventing epithelial barrier disruption and inflammation induced by endotoxins (e.g. LPS), resulting in the intestinal barrier protection and preserving its health.

The results of the above-mentioned experiment proved to be successful as the probiotic strain *L. reuteri* I5007 was able to modulate the expression of tight junction proteins *in vivo*, in the intestine, and *in vitro*, using IPEC-J2 cells, thus reinforcing the belief that this probiotic helps protecting and supporting the intestinal mucosal barrier.

5.2 Development of specific primers for probiotic strains

The aim of this project was to create a panel of specific primers capable of detecting the probiotic strains mentioned above. Based on the results obtained, it is possible to state that it was partially accomplished, as a set of primers was proposed for all the strains selected, except for *L. brevis* KB290. A major limitation of this study lies in the fact that the results obtained can only be considered positive on a theoretical level, as the set of primers proposed for each probiotic strain has not yet been validated. Other limitations are associated with the bioinformatic software used, Geneious, when compared with other platforms. The license cost for the use of this software is elevated, it's restricted mode blocks access to most of the program tools, it requires a higher memory and CPU power from the user's computer to work efficiently and it takes longer time to generate results, particularly when analysing large datasets.

The genomes of each of the selected probiotic strains and the gene fragments chosen for our study, obtained from the NCBI database collection, were compared with multiple fragments and allowed to analyse the sequences and determine polymorphisms between all sequences.

It is commonly found in the literature, sets of primers capable of being used to detect single microbial species and some examples are presented for *B. subtilis* and *L. brevis*. In a recent study (216), out of 17 housekeeping genes studied, the authors selected *pycA* and *aroE* as the best genes to distinguish *B. subtilis* from other *Bacillus* species, with the *pycA* gene showing the highest number of polymorphisms as it had 1075 polymorphic sites. The *pycA* gene has shown a nucleotide identity between *B. subtilis* strains of 98.9%-100%, compared to the identity shared with other *Bacillus* species of 79.7%-82.1%, and the *aroE* gene has shown a similarity between strains of 98.8%-99.9%, in contrast to other *Bacillus* species which have shown a similarity of 67.0%-74.6% in the entire sequence of this gene. After selecting the two genes, distinct nucleotide sequences were identified by comparative analysis, and then specific primers were designed for these sequences. In the end, the PCR method was able to confirm the specificity of these primers by only amplifying the DNA of *B. subtilis* and not from other *Bacillus* species. During a study performed with lactic acid bacteria, the quantitative PCR method was used for the quantitative and qualitative detection of six lactic acid bacteria isolates using the *tuf* gene, a gene whose region was more variable between species, with one of these isolates belonging to the *L. brevis* species (217).

Other studies carried out over the last couple of years have proposed different approaches to distinguishing the strains used in our study from some other strains of the same species (218,219). A study revealed that it was possible to differentiate *B. subtilis* 168 from other strains of the same species based on existing polymorphisms in the sequences of *B. subtilis* 168, *B. subtilis* NCBI 3610T, *B. subtilis* PY76 and *B. subtilis* SMY (218).

A second research conducted in the last few years was capable to distinguish *L. brevis* KB290 from other strain of the same species (219). Comparisons between *L. brevis* KB290 and another probiotic strain, *L. brevis* ATCC 367, showed that strain KB290 harbours genes encoding for 375 unique proteins when compared with strain ATCC 367. This study allowed us to observe the existing differences between these two strains of the same species, although no specific detection method has been proposed for *L. brevis* strain KB290.

Finally, for *L. reuteri* I5007, many studies have been found with methods for detecting this strain, although they have not proved to be specific. As an example, we found a study conducted with the aim of evaluating the effects of this strain on the diversity of the intestinal microbial community in pigs (156). The method used to detect and analyse the microbial community was Denaturing Gradient Gel Electrophoresis after performing PCR of products resulting from the amplification of the V6 - V8 regions of the 16S rRNA of these same microorganisms, which included, among many microorganisms, the strain *L. reuteri* I5007 tested in this thesis. This same method was used to separate the fragments produced by PCR amplification of two sets of primers.

Based on the literature, no studies or approaches previously proposed were able to uniquely distinguish specific strains from large sets of strains including closely related strains. It becomes difficult to find variations in regions of specific strains that are unique and very few comparative studies with genetically close microorganisms belonging to the same species have shown promising results in this regard (161). The challenge lies in finding genomic regions that remain stable over time, but at the same time exhibit enough variability to distinguish a given strain from others of the same species.

Given the above situations, a set of highly conserved genes, some of which housekeeping genes, were selected for comparison with other strains. Only part of these genes was used further in this project because they proved to have the unique characteristics needed to design specific primers for each of our strains. The first reason for excluding certain genes

was the low number of variability found for these genes in the NCBI database. Ideally, the availability of multiple gene sequences, variable enough to characterize the full species, would facilitate our task and get good conclusions from the comparison of genomes or sequences. The second reason for the rejection of genes in this project was that, after aligning the reference genomes with the fragments of the chosen genes, there were genes with multiple polymorphisms, such as SNPs, but they were too sparce and no enough polymorphisms were available in order to provide a differentiation of a specific strain from others based on that gene (in other words it was not possible to design specific primers).

Therefore, enough SNPs were identified in a last set of genes allowing the design of primers capable of distinguishing our targeting strains from a large number of others belonging to the same species, including closely related strains. A set of ten primers – five forward and five reverse primers, were proposed for *B. pumilus* SAFR-032; five primers – three forward and two reverse primers, were proposed for *B. subtilis* 168; and eight primers – four forward and four reverse primers, were proposed for *L. reuteri* I5007. These primers need to be tested now in multiple samples in order to verify in real samples the result observed *in silico*.

Chapter 6: Conclusion

The purpose of this study was to develop a panel of primers capable of detecting the strains *B. pumilus* SAFR-032, *B. subtilis* 168, *L. brevis* KB290 and *L. reuteri* I5007. These primers could be potentially used in the future to detect the same strains in complex biological samples with high biodiversity and genetic diversity. The goal was achieved for three strains, as a set of potential primers was proposed for the detection of the targeted strains. However, it was not possible to design a set of primers for the *L. brevis* KB290 strain as the variability of the tested genes was low. At the end of this project, a set of ten primers was selected for *B. pumilus* SAFR-032, five primers for *B. subtilis* 168 and eight primers for *L. reuteri* I5007.

Although it was not possible to deliver a set of primers for *L. brevis* KB290, the study yielded promising results. A panel of primers with potential to detect the probiotic strains *B. pumilus* SAFR-032, *B. subtilis* 168 and *L. reuteri* I5007 was successfully proposed. In contrast with previous comparative studies that failed to propose methods for the detection of the above-mentioned strains, the current work presents a methodological approach with potential applicability not only for the selected strains, but also for the detection of other strains.

Chapter 7: Future work

Considering the work carried out in this project and the fact that a set of primers was proposed for the detection of the strains *B. pumilus* SAFR-032, *B. subtilis* 168 and *L. reuteri* I5007, the next step for future research involves the experimental validation of these primers. The primers were designed *in silico* and demonstrated potential for the detection of the proposed strains among large genomic datasets, including closely related strains. Nevertheless, it will only be possible to validate the specificity of these primers through laboratorial work with the conditions proposed during this study testing clinical samples, such as intestinal mucosa, intestinal tissue, faecal or oral samples, and other samples, such as food samples, food products with probiotic effects where these strains are often included or fermented food products.

Certain improvements could also be introduced to this work in its many stages, such as a more extensive search of different genes for our study strains, which not only could provide more fragments from other strains for comparison with ours, but could also provide more unique genomic regions, which would be an important factor in the design of specific primers for our strains. In addition, the comparison of thousands of genomes for each species, including closely related strains, can be relevant insights on the best regions for polymorphic segments and therefore unique primer designs. Finally, the methodology developed in this work could be extended to other probiotic strains including the design of multiplex systems capable to detect and monitor multiple strains at once and therefore exploring the potential of multiple probiotics combined and cooperating for a better human health.

Appendixes

Appendix 1 Fragments containing polymorphic regions with potential for primers in the *Bacillus pumilus* SAFR-032 *gyrB*, *mutL*, *pycA* and *pyrE* genes.

Gene	Gene Fragment	Gene Fragment length (bp)	Melting Temperature (ºC)	GC content (%)
gyrB	5’ TATGACTGTTCTTCACGCTGGC 3’ 3’ ATACTGACAAGAAGTGCGACCG 5’	22	57.3	50.0
	3’ TCTACGGCTACACCTACCTC 5’	20	57.0	55.0
	3’ TCTACGGCTACACCTACCT 5’	19	56.2	52.6
	5’ GGATGGAGCGCACATTCGAACG 3’ 3’ CCTACCTCGCGTGTAAGCTTGC 5’	22	60.9	59.1
mutL	5’ TCCCGCTTCCGTTGTAAAAGAG 3’	22	64.5	50.0
	5’ TCCCGCTTCCGTTGTAAAAGA 3’	21	64.1	47.6
	5’ TCCCGCTTCCGTTGTAAAAG 3’	20	62.6	50.0
	5’ CGCTTCCGTTGTAAAAGA 3’	18	57.7	44.4
	5’ GCTTACGTCACCAAGGAAAAGTC 3’ 3’ CGAATGCAGTGGTTCCTTTTCAG 5’	23	63.5	47.8
	5’ CTTACGTCACCAAGGAAAAGTC 3’ 3’ GAATGCAGTGGTTCCTTTTCAG 5’	22	61.1	45.5
	5’ CTTACGTCACCAAGGAAAAG 3’ 3’ GAATGCAGTGGTTCCTTTTC 5’	20	58.5	45.0
	3’ CCGATAAATACCCTGTCGTCAAC 5’	23	63.3	47.8
	3’ CCGATAAATACCCTGTCGTCAA 5’	22	62.6	45.5
	3’ CCGATAAATACCCTGTCGTCA 5’	21	62.1	47.6
	5’ GCTATTTATGGGACAGCAGT 3’ 3’ CGATAAATACCCTGTCGTCA 5’	20	59.8	45.0
	5’ GCTATTTATGGGACAGCAG 3’ 3’ CGATAAATACCCTGTCGTC 5’	19	57.8	47.4
	5’ CTATTTATGGGACAGCAG 3’ 3’ GATAAATACCCTGTCGTC 5’	18	54.4	44.4
	3’ TACCGTCTATACACTTCGTAAAGG 5’	24	63.6	41.7
	3’ ACCGTCTATACACTTCGTAAAGG 5’	23	63.3	43.5
	5’ GTGAAGCATTTCCCGCTCGTAA 3’	22	65.0	50.0
	5’ TGAAGCATTTCCCGCTCGTAA 3’	21	64.4	47.6
	5’ TGAAGCATTTCCCGCTCGTA 3’	20	63.9	50.0
	5’ GAAGCATTTCCCGCTCGTA 3’	19	62.1	52.6
	5’ CACGCTTCTGCCGATTGGT 3' 3’ GTGCGAAGACGGCTAACCA 5’	19	65.0	57.9
pycA	5’ GACACCTTCTTCTAAAGTAGTCGG 3’	24	62.1	45.8
	5’ GACACCTTCTTCTAAAGTAGTCG 3’	23	60.2	43.5
	5’ GACACCTTCTTCTAAAGTAGTC 3’	22	57.9	40.9
	5’ GTGACACCTTCTTCTAAAGTAGTC 3’	24	60.5	41.7
	5’ CCTTCTTCTAAAGTAGTCGGAGAT 3’	24	61.3	41.7
	5’ CTTCTAAAGTAGTCGGAGAT 3’ 3’ GAAGATTTCATCAGCCTCTA 5’	20	56.0	40.0
	5’ CGGTTCTTGACACACCAACATTT 3’ 3’ GCCAAGAACTGTGTGGTTGTAAA 5’	23	64.0	43.5
	5’ GGTTCTTGACACACCAACATTT 3’ 3’ CCAAGAACTGTGTGGTTGTAAA 5’	22	61.8	40.9
	5’ ACTCATTGTGAAACTCGTGTCA 3’	22	62.4	40.9

	3’ TGAGTAACACTTTGAGCACAGT 5’			
pyrE	3’ GTAACGGCTTCCAGACTTTCTCA 5’	23	64.7	47.8
	3’ TAACGGCTTCCAGACTTTCTCA 5’	22	64.1	45.5
	3’ AACGGCTTCCAGACTTTCTCA 5’	21	63.8	47.6
	3’ ACGGCTTCCAGACTTTCTCA 5’	20	63.3	50.0
	3’ GGCTTCCAGACTTTCTCA 5’	18	58.3	50.0
	5’ CTGAATGCTCCGATGTGTTATGTC 3’ 3’ GACTTACGAGGCTACACAATACAG 5’	24	63.4	45.8
	5’ TGAATGCTCCGATGTGTTATGTC 3’ 3’ ACTTACGAGGCTACACAATACAG 5’	23	62.9	43.5
	5’ GAATGCTCCGATGTGTTATGTC 3’ 3’ CTTACGAGGCTACACAATACAG 5’	22	61.3	45.5
	5’ AATGCTCCGATGTGTTATGTC 3’ 3’ TTACGAGGCTACACAATACAG 5’	21	60.6	42.9
	5’ ATGCTCCGATGTGTTATGTC 3’ 3’ TACGAGGCTACACAATAGAC 5’	20	60.0	45.0
	5’ TGCTCCGATGTGTTATGTC 3’ 3’ ACGAGGCTACACAATACAG 5’	19	59.5	47.4
	3’ TTCCATCAGCAGCATCTTCTAGAT 5’	24	62.0	41.7
	3’ TCCATCAGCAGCATCTTCTAGAT 5’	23	61.5	43.5
	3’ CCATCAGCAGCATCTTCTAGAT 5’	22	60.0	45.5
	3’ AGAACCTACCTTCTTTGGCAGATT 5’	24	63.8	41.7
	3’ GAACCTACCTTCTTTGGCAGATT 5’	23	62.4	43.5
	3’ AACCTACCTTCTTTGGCAGATT 5’	22	61.8	40.9
	3’ ACCTACCTTCTTTGGCAGATT 5’	21	61.2	42.9
	3’ CCTACCTTCTTTGGCAGATT 5’	20	59.3	45.0
	3’ CTACCTTCTTTGGCAGATT 5’	19	56.8	42.1

Appendix 2 Fragments containing polymorphic regions with potential for primers in the *Bacillus subtilis* 168 *dnaK* and *groEL* genes.

Gene	Gene Fragment	Gene Fragment length (bp)	Melting Temperature (ºC)	GC content (%)
dnaK	5’ GTCGATTTCGCTTGCAGAAAGTCC 3’	24	58.6	50.0
	5’ GTCGATTTCGCTTGCAGAAAGTC 3’	23	56.8	47.8
	5’ GTCGATTTCGCTTGCAGAAAGT 3’	22	56.1	45.5
	5’ GTCGATTTCGCTTGCAGAAAG 3’	21	54.7	47.6
	5’ GTCGATTTCGCTTGCAGAAA 3’	20	53.9	45.0
	5’ GTCGATTTCGCTTGCAGAA 3’	19	53.5	47.4
	5’ TTTCGCTTGCAGAAAGTCC 3’ 3’ AAAGCGAACGTCTTTCAGG 5’	19	53.5	47.4
	3’ AGTTCCACTTCGCCAGGACGA 5’	21	61.2	57.1
	5’ TCAAGGTGAAGCGGTCCTGC 3’ 3’AGTTCCACTTCGCCAGGACG 5’	20	60.1	60.0
	5’ TCAAGGTGAAGCGGTCCTG 3’ 3’ AGTTCCACTTCGCCAGGAC 5’	19	57.3	57.9
	5’ TCTTGTGGAGTGTAATCCTTTC 3’ 3’ AGAACACCTCACATTAGGAAG 5’	22	52.3	40.9
	5’ TCTTGTGGAGTGTAATCCTT 3’ 3’ AGAACACCTCACATTAGGAA 5’	20	50.7	40.0
	5’ CTTGTGGAGTGTAATCCTTTC 3’ 3’ GAACACCTCACATTAGGAAAG 5’	21	50.9	42.9
	5’ GTACCCATATGACGTTTGATAG 3’ 3’ CATGGGTATACTGCAAACTATC 5’	22	50.6	40.9
	5’ GTTAGGGTTTGTAATAGATTGGCG 3’ 3’ CAATCCCAAACATTATCTAACCGC 5’	24	53.9	41.7
	5’ GGGTTTGTAATAGATTGGCGTT 3’ 3’ CCCAAACATTATCTAACCGCAA 5’	22	53.4	40.9
	5’ GGGTTTGTAATAGATTGGCGT 3’ 3’ CCCAAACATTATCTAACCGCA 5’	21	53.0	42.9
	5’ GGGTTTGTAATAGATTGGCG 3’ 3’ CCCAAACATTATCTAACCGC 5’	20	51.3	45.0
	5’ GTTCGCCGTTTTTAAATGCA 3’ 3’CAAGCGGCAAAAATTTACGT 5’	20	52.4	40.0
	5’ GTTTGTTGTTCCTAAGTCGATTCC 3’ 3’CAAACAACAAGGATTCAGCTAAGG 5’	24	54.1	41.7
	5’ GTTGTTCCTAAGTCGATTCCGAT 3’	23	54.7	43.5
	5’ GTTGTTCCTAAGTCGATTCCGA 3’	22	54.5	45.5
	5’ GTTGTTCCTAAGTCGATTCC 3’ 3’ CAACAAGGATTCAGCTAAGG 5’	20	50.6	45.0
	5’ TGTTCCTAAGTCGATTCCGAT 3’	21	53.3	42.9
	5’ TGTTCCTAAGTCGATTCCGA 3’	20	53.0	45.0
	5’ GTTCCTAAGTCGATTCCGATAACT 3’	24	54.1	41.7
	5’ GTTCCTAAGTCGATTCCGATAA 3’	22	51.8	40.9
	5’ GTTCCTAAGTCGATTCCGAT 3’	20	51.7	45.0
	5’ GTTCCTAAGTCGATTCCGA 3’	19	51.3	47.4
	5’ CCTAAGTCGATTCCGATAACT 3’	21	51.4	42.9
groEL	5’ CAAAAGTTACACGCTCAGCTCTTC 3’ 3’ GTTTTCAATGTGCGAGTCGAGAAG 5’	24	63.9	45.8
	5’ CAAAAGTTACACGCTCAGCTCTT 3’ 3’ GTTTTCAATGTGCGAGTCGAGAA 5’	23	63.4	43.5
	5’ AAAAGTTACACGCTCAGCTCTT 3’ 3’ TTTTCAATGTGCGAGTCGAGAA 5’	22	62.7	40.9

	5’ AGCTCTTCAAAACGCTGCG 3’ 3’ TCGAGAAGTTTTGCGACGC 5’	19	63.4	52.6
	5' CTGCGTCTGTAGCTGCAATGT 3’ 3’ GACGCAGACATCGACGTTACA 5’	21	64.9	52.4
	5’ TGCGTCTGTAGCTGCAATGT 3’ 3’ ACGCAGACATCGACGTTACA 5’	20	64.4	50.0
	5’ ACAAGCCAGAAGAAAACGGTG 3’ 3’ TGTTCGGTCTTCTTTTGCCAC 5’	21	63.3	47.6
	5’ CAAGCCAGAAGAAAACGGTG 3’ 3’ GTTCGGTCTTCTTTTGCCAC 5’	20	61.6	50.0
	5’ AAGCCAGAAGAAAACGGTG 3’ 3’ TTCGGTCTTCTTTTGCCAC 5’	19	60.6	47.4
	5’ AGCCAGAAGAAAACGGTG 3’ 3’ TCGGTCTTCTTTTGCCAC 5’	18	59.9	50.0

Appendix 3 Fragments containing polymorphic regions with potential for primers in the *Levilactobacillus brevis* KB290 *rpoA* gene.

Gene	**Gene Fragment**	**Gene Fragment length (bp)**	**Melting Temperature (ºC)**	**GC content (%)**
rpoA	5’ CAAACATTGCCAGATGTTCAG 3’	21	52.5	42.9
	5’ CAAACATTGCCAGATGTTCA 3’	20	51.4	40.0
	5’ AAACATTGCCAGATGTTCAG 3’	20	51.2	40.0
	5’ AACATTGCCAGATGTTCAG 3’ 3’ TTGTAACGGTCTACAAGTC 5’	19	50.7	42.1
	5’ ACATTGCCAGATGTTCAG 3’ 3’ TGTAACGGTCTACAAGTC 5’	18	50.0	44.4
	5’ CAGAACCCCGTCGATTTGAATG 3’ 3’ GTCTTGGGGCAGCTAAACTTAC 5’	22	56.3	50.0
	5’CAGAACCCCGTCGATTTGA 3’ 3’ GTCTTGGGGCAGCTAAACT 5’	19	54.9	52.6
	5’ AGAACCCCGTCGATTTGAATG 3’ 3’ TCTTGGGGCAGCTAAACTTAC 5’	21	55.3	47.6
	5’ GAACCCCGTCGATTTGAATG 3’ 3’ CTTGGGGCAGCTAAACTTAC 5’	20	54.0	50.0
	5’ AACCCCGTCGATTTGAATG 3’ 3’ TTGGGGCAGCTAAACTTAC 5’	19	53.1	47.4

Appendix 4 Fragments containing polymorphic regions with potential for primers in the *Limosilactobacillus reuteri* I5007 *ddl* and *recA* gene.

Gene	Gene Fragment	Gene Fragment length (bp)	Melting Temperature (ºC)	GC content (%)
ddl	5’ AACTTCCCACATTTGAGGG 3’	19	52.5	47.4
	5’ ACTTCCCACATTTGAGGG 3’	18	52.0	50.0
	5’ AATGCCTTGTATGCCTTC 3’ 3’ TTACGGAACATACGGAAG 5’	18	50.2	44.4
	5’ GGTACACGAACAGCTCCTAA 3’ 3’ CCATGTGCTTGTCGAGGATT 5’	20	54.1	50.0
	5’ GGTACACGAACAGCTCCTA 3’ 3’ CCATGTGCTTGTCGAGGAT 5’	19	53.7	52.6
	5’ GGTACACGAACAGCTCCT 3’ 3’ CCATGTGCTTGTCGAGGA 5’	18	54.2	55.6
	5’ TCAATTTCTTGAGGATTGGC 3’	20	50.7	40.0
	5’ TTCTTGAGGATTGGCGAT 3’ 3’ AAGAACTCCTAACCGCTA 5’	18	51.1	44.4
recA	5' TTCCAATAACATCCGTACCA 3’ 3’ AAGGTTATTGTAGGCATGGT 5’	20	51.1	40.0
	5’ AGCTTCCGAAGAGCCTGAGACAT 3’ 3’ TCGAAGGCTTCTCGGACTCTGTA 5’	23	60.1	52.2
	5’ AGCTTCCGAAGAGCCTGAGACA 3’ 3’ TCGAAGGCTTCTCGGACTCTGT 5’	22	60.2	54.5
	5’ AGCTTCCGAAGAGCCTGAGAC 3’ 3’ TCGAAGGCTTCTCGGACTCTG 5’	21	58.9	57.1
	5’ AGCTTCCGAAGAGCCTGAG 3’ 3’ TCGAAGGCTTCTCGGACTC 5’	19	56.9	57.9
	5’ AAGAGCCTGAGACATCAAACGAGC 3’ 3’ TTCTCGGACTCTGTAGTTTGCTCG 5’	24	59.1	50.0
	5’ AAGAGCCTGAGACATCAAACGAG 3’ 3’ TTCTCGGACTCTGTAGTTTGCTC 5’	23	56.8	47.8
	5’ AAGAGCCTGAGACATCAAACGA 3’ 3’ TTCTCGGACTCTGTAGTTTGCT 5’	22	56.2	45.5
	5’ AAGAGCCTGAGACATCAAACG 3’ 3’ TTCTCGGACTCTGTAGTTTGC 5’	21	55.0	47.6
	5’ AAGAGCCTGAGACATCAAAC 3’ 3’ TTCTCGGACTCTGTAGTTTG 5’	20	52.4	45.0
	3’ GACGCTTTCGTTATCCAGTAGTTA 5’	24	55.8	41.7
	3’ GACGCTTTCGTTATCCAGTAG 5’	21	53.9	47.6
	3’ GACGCTTTCGTTATCCAG 5’	18	51.0	50.0
	3’ CGCTTTCGTTATCCAGTAGTTA 5’	22	53.6	40.9
	3’ CGCTTTCGTTATCCAGTAG 5’	19	51.3	47.4
	3’ CCGAAGACGTATAACTCCTAGTTC 5’	24	55.8	45.8
	3’ CCGAAGACGTATAACTCCTAGTT 5’	23	55.1	43.5
	5’ GGCTTCTGCATATTGAGGATCA 3’ 3’ CCGAAGACGTATAACTCCTAGT 5’	22	54.8	45.5
	3’ GAAGACGTATAACTCCTAGTTC 5’	22	58.9	40.9
	3’ GACGTATAACTCCTAGTTCTCGTA 5’	24	62.9	41.7
	3’ GACGTATAACTCCTAGTTCTCGT 5’	23	62.6	43.5
	5’ CTGTCCCGCCTTGCCGTTGTAC 3’ 3’ GACAGGGCGGAACGGCAACATG 5’	22	68.9	63.6

References

1. Curry C. J., Davis B. W., Bertola L. D., White P. A., Murphy W. J., Derr J. N. (2021) Spatiotemporal genetic diversity of lions reveals the influence of habitat fragmentation across Africa. *Molecular Biology and Evolution* **38**(1):48–57. https://doi.org/10.1093/molbev/msaa174.

2. Hanson J. O., Veríssimo A., Velo-Antón G., Marques A., Camacho-Sanchez M., Martínez-Solano Í., Gonçalves H., Sequeira F., Possingham H. P., Carvalho S. B. (2020) Evaluating surrogates of genetic diversity for conservation planning. *Conservation Biology* **35**(2):642. https://doi.org/10.1111/cobi.13602.

3. Pereira L., Mutesa L., Tindana P., Ramsay M. (2021) African genetic diversity and adaptation inform a precision medicine agenda. *Nature Review Genetics* **22**(5):284–306. https://doi.org/10.1038/s41576-020-00306-8.

4. Teixeira J. C., Huber C. D. (2021) The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences of the United States of America* **118**(10). https://doi.org/10.1073/pnas.2015096118.

5. Zhang L., Zhang H. L., Chen Y., Nizamani M. M., Wu T., Liu T., Zhou Q. (2024) Assessing genetic diversity in critically endangered *Chieniodendron hainanense* populations within fragmented habitats in Hainan. *Scientific Reports* **14**(6988). https://doi.org/10.1038/s41598-024-56630-0.

6. Vellend M., Geber M. A. (2005) Connections between species diversity and genetic diversity. *Ecology Letters* **8**(7):767–781. https://doi.org/10.1111/j.1461-0248.2005.00775.x.

7. Jin Y., Li Y., Huang S., Hong C., Feng X., Cai H., Xia Y., Li S., Zhang L., Lou Y., Guan W. (2024) Whole-genome sequencing analysis of antimicrobial resistance, virulence factors, and genetic diversity of *Salmonella* from Wenzhou, China. *Microorganisms* **12**(11):2166. https://doi.org/10.3390/microorganisms12112166.

8. Chauhan P., Sharma N., Tapwal A., Kumar A., Verma G. S., Meena M., Seth C. S., Swapnil P. (2023) Soil microbiome: Diversity, benefits and interactions with plants. *Sustainability* **15**(19):14643. https://doi.org/10.3390/su151914643.

9. Brito P. H., Chevreux B., Serra C. R., Schyns G., Henriques A. O., Pereira-Leal J. B. (2018) Genetic competence drives genome diversity in *Bacillus subtilis*. *Genome Biology and Evolution* **10**(1):108–124. https://doi.org/10.1093/gbe/evx270.

10. Emamalipour M., Seidi K., Vahed S. Z., Jahanban-Esfahlan A., Jaymand M., Majdi H., Amoozgar Z., Chitkushev L. T., Javaheri T., Jahanban-Esfahlan R., Zare P. (2020) Horizontal gene transfer: From evolutionary flexibility to disease progression. *Frontiers in Cell and Developmental Biology* **8**(229). https://doi.org/10.3389/fcell.2020.00229.

11. Näsvall K., Boman J., Talla V., Backström N. (2023) Base composition, codon usage, and patterns of gene sequence evolution in butterflies. *Genome Biology and Evolution* **15**(8). https://doi.org/10.1093/gbe/evad150.

12. Comaills V., Castellano-Pozo M. (2023) Chromosomal instability in genome evolution: From cancer to macroevolution. *Biology* **12**(5):1–35. https://doi.org/10.3390/biology12050671.

13. Bamba M., Aoki S., Kajita T., Setoguchi H., Watano Y., Sato S., Tsuchimatsu T. (2019) Exploring genetic diversity and signatures of horizontal gene transfer in nodule bacteria associated with *Lotus japonicus* in natural environments. *Molecular Plant-Microbe Interactions* **32**(9):1110–1120. https://doi.org/10.1094/MPMI-02-19-0039-R.

14. Dionisio F., Zilhão R., Gama J. A. (2019) Interactions between plasmids and other mobile genetic elements affect their transmission and persistence. *Plasmid* **102**:29–36. https://doi.org/10.1016/j.plasmid.2019.01.003.

15. Maguire F., Jia B., Gray K. L., Lau W. Y. V., Beiko R. G., Brinkman F. S. L. (2020) Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic islands. *Microbial Genomics* **6**(10). https://doi.org/10.1099/mgen.0.000436.

16. Lampel K. A., Formal S. B., Maurelli A. T. (2018) A brief history of *Shigella*. *EcoSal Plus* **8**(1). https://doi.org/10.1128/ecosalplus.esp-0006-2017.

17. Jaskólska M., Adams D. W., Blokesch M. (2022) Two defence systems eliminate plasmids from seventh pandemic *Vibrio cholerae*. *Nature* **604**(7905):323-329. https://doi.org/10.1038/s41586-022-04546-y.

18. Piña-Iturbe A., Suazo I. D., Hoppe-Elsholz G., Ulloa-Allendes D., González P. A., Kalergis A. M., Bueno S. M. (2020) Horizontally acquired homologs of xenogeneic silencers: Modulators of gene expression encoded by plasmids, phages and genomic islands. *Genes* **11**(2):142. https://doi.org/10.3390/genes11020142.

19. Chiang Y. N., Penadés J. R., Chen J. (2019) Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLoS Pathogens* **15**(8):1–7. https://doi.org/10.1371/journal.ppat.1007878.

20. Borodovich T., Shkoporov A. N., Ross R. P., Hill C. (2022) Phage-mediated horizontal gene transfer and its implications for the human gut microbiome. *Gastroenterology Report* **10**:1–12. https://doi.org/10.1093/gastro/goac012.

21. Des Roches S., Pendleton L. H., Shapiro B., Palkovacs E. P. (2021) Conserving intraspecific variation for nature's contributions to people. *Nature Ecology & Evolution* **5**(5):574–582. https://doi.org/10.1038/s41559-021-01403-5.

22. Hamer J., Matthiessen B., Pulina S., Hattich G. S. I. (2022) Maintenance of intraspecific diversity in response to species competition and nutrient fluctuations. *Microorganisms* **10**(1):113. https://doi.org/10.3390/microorganisms10010113.

23. Earl A. M., Losick R., Kolter R. (2006) *Bacillus subtilis* genome diversity. *Journal of Bacteriology* **189**(3):1163-1170. https://doi.org/10.1128/jb.01343-06.

24. Wang T., Shi Y., Zheng M., Zheng J. (2024) Comparative genomics unveils functional diversity, pangenome openness, and underlying biological drivers among *Bacillus subtilis* group. *Microorganisms* **12**(5):986. https://doi.org/10.3390/microorganisms12050986.

25. Yu G., Wang X. C., Tian W. H., Shi J. C., Wang B., Ye Q., Dong S. G., Zeng M., Wang J. Z. (2015) Genomic diversity and evolution of *Bacillus subtilis*. *Biomedical and Environmental Sciences* **28**(8):620–625. https://doi.org/10.3967/bes2015.087.

26. Tenaillon O., Matic I. (2020) The impact of neutral mutations on genome evolvability. *Current Biology* **30**(10): R527–R534. https://doi.org/10.1016/j.cub.2020.03.056.

27. Darmon E., Leach D. R. F. (2014) Bacterial genome instability. *Microbiology and Molecular Biology Reviews* **78**(1):1–39. https://doi.org/10.1128/mmbr.00035-13.

28. Lavanchy E., Cumer T., Topaloudis A., Ducrest A. L., Simon C., Roulin A., Goudet J. (2024) Too big to purge: Persistence of deleterious mutations in island populations of the european Barn Owl (*Tyto alba*). *Heredity* **133**(6):437–449. https://doi.org/10.1038/s41437-024-00728-8.

29. Roberts G., Petrie M. (2022) Sexual selection for males with beneficial mutations. *Scientific Reports* **12**(12613). https://doi.org/10.1038/s41598-022-16002-y.

30. Orr H. A. (2010) The population genetics of beneficial mutations. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**(1544):1201. https://doi.org/10.1098/rstb.2009.0282.

31. Ferenci T. (2007) The spread of a beneficial mutation in experimental bacterial populations: The influence of the environment and genotype on the fixation of *rpoS* mutations. *Heredity* **100**(5):446–452. https://doi.org/10.1038/sj.hdy.6801077.

32. Tang Y. W., Ellis N. M., Hopkins M. K., Smith D. H., Dodge D. E., Persing D. H. (1998) Comparison of phenotypic and genotypic techniques for identification of unusual aerobic pathogenic gram-negative bacilli. *Journal of Clinical Microbiology* **36**(12):3674–3679. https://doi.org/10.1128/jcm.36.12.3674-3679.1998.

33. Cutter A. D., Jovelin R. (2015) When natural selection gives gene function the cold shoulder. *BioEssays* **37**(11):1169–1173. https://doi.org/10.1002/bies.201500083.

34. Ochoa-Díaz M. M., Daza-Giovannetty S., Gómez-Camargo D. (2018) Bacterial genotyping methods: From the basics to modern. *Methods in Molecular Biology* **1734**:13–20. https://doi.org/10.1007/978-1-4939-7604-1_2.

35. Nykrynova M., Barton V., Sedlar K., Bezdicek M., Lengerova M., Skutkova H. (2021) Word entropy-based approach to detect highly variable genetic markers for bacterial genotyping. *Frontiers in Microbiology* **12**(631605). https://doi.org/10.3389/fmicb.2021.631605.

36. Ahmed S. S., Alp E. (2015) Genotyping methods for monitoring the epidemic evolution of *A. baumannii* strains. *The Journal of Infection in Developing Countries* **9**(4):347–354. https://doi.org/10.3855/jidc.6201.

37. Dijkshoorn L., Ursing B. M., Ursing J. B. (2000) Strain, clone and species: Comments on three basic concepts of bacteriology. *Journal of Medical Microbiology* **49**(5):397–401. https://doi.org/10.1099/0022-1317-49-5-397.

38. Van Rossum T., Ferretti P., Maistrenko O. M., Bork P. (2020) Diversity within species: Interpreting strains in microbiomes. *Nature Reviews Microbiology* **18**(9):491-506. https://doi.org/10.1038/s41579-020-0368-1.

39. Zasada A. A., Mosiej E. (2018) Contemporary microbiology and identification of *Corynebacteria* spp. causing infections in human. *Letters in Applied Microbiology* **66**(6):472–483. https://doi.org/10.1111/lam.12883.

40. Li W., Raoult D., Fournier P. E. (2009) Bacterial strain typing in the genomic era. *FEMS Microbiology Reviews* **33**(5):892–916. https://doi.org/10.1111/j.1574-6976.2009.00182.x.

41. Tenover F. C., Arbeit R. D., Goering R. V., Mickelsen P. A., Murray B. E., Persing D. H., Swaminathan B. (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed- field gel electrophoresis: Criteria for bacterial strain typing. *Journal of Clinical Microbiology* **33**(9):2233–2239. https://doi.org/10.1128/jcm.33.9.2233-2239.1995.

42. Donelli G., Vuotto C., Mastromarino P. (2013) Phenotyping and genotyping are both essential to identify and classify a probiotic microorganism. *Microbial Ecology in Health and Disease* **24**(20105). https://doi.org/10.3402/mehd.v24i0.20105.

43. Liu Y., Lai Q., Dong C., Sun F., Wang L., Li G., Shao Z. (2013) Phylogenetic diversity of the *Bacillus pumilus* group and the marine ecotype revealed by multilocus sequence analysis. *PLoS One* **8**(11). https://doi.org/10.1371/journal.pone.0080097.

44. Ross T. L., Merz W. G., Farkosh M., Carroll K. C. (2005) Comparison of an automated repetitive sequence-based PCR microbial typing system to pulsed-field gel electrophoresis for analysis of outbreaks of methicillin-resistant *Staphylococcus aureus*. *Journal of Clinical Microbiology* **43**(11):5642-5647. https://doi.org/10.1128/jcm.43.11.5642-5647.2005.

45. Hallin M., Deplano A., Denis O., De Mendonça R., De Ryck R., Struelens M. J. (2007) Validation of pulsed-field gel electrophoresis and spa typing for long-term, nationwide

epidemiological surveillance studies of *Staphylococcus aureus* infections. *Journal of Clinical Microbiology* **45**(1):127–133. https://doi.org/10.1128/jcm.01866-06.

46. Olive D. M., Bean P. (1999) Principles and applications of methods for DNA-based typing of microbial organisms. *Journal of Clinical Microbiology* **37**(6):1661-1669. https://doi.org/10.1128/jcm.37.6.1661-1669.1999.

47. Zhu Y. L., Song Q. J., Hyten D. L., Van Tassell C. P., Matukumalli L. K., Grimm D. R., Hyatt S. M., Fickus E. W., Young N. D., Cregan P. B. (2003) Single-nucleotide polymorphisms in soybean. *Genetics* **163**(3):1123-1134. https://doi.org/10.1093/genetics/163.3.1123.

48. Moura A., Criscuolo A., Pouseele H., Maury M. M., Leclercq A., Tarr C., Björkman J. T., Dallman T., Reimer A., Enouf V., Larsonneur E., Carleton H., Bracq-Dieye H., Katz L. S., Jones L., Touchon M., Tourdjman M., Walker M., Stroika S., Cantinelli T., Chenal-Francisque V., Kucerova Z., Rocha E. P. C., Nadon C., Grant K., Nielson E. M., Pot B., Gerner-Smidt P., Lecuit M., Brisse S. (2016) Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nature Microbiology* **2**(16185). https://doi.org/10.1038/nmicrobiol.2016.185.

49. Pearce M. E., Alikhan N. F., Dallman T. J., Zhou Z., Grant K., Maiden M. C. J. (2018) Comparative analysis of core genome MLST and SNP typing within a european *Salmonella* serovar *Enteritidis* outbreak. *International Journal of Food Microbiology* **274**:1-11. https://doi.org/10.1016/j.ijfoodmicro.2018.02.023.

50. Trembizki E., Smith H., Lahra M. M., Chen M., Donovan B., Fairley C. K., Guy R., Kaldor J., Regan D., Ward J., Nissen M. D., Sloots T. P., Whiley D. M. (2014) High-throughput informative single nucleotide polymorphism-based typing of *Neisseria gonorrhoeae* using the Sequenom MassARRAY iPLEX platform. *Journal of Antimicrobial Chemotherapy* **69**(6):1526–1532. https://doi.org/10.1093/jac/dkt544.

51. Gabriel S., Ziaugra L., Tabbaa D. (2009) SNP genotyping using the sequenom MassARRAY iPLEX platform. *Current Protocols in Human Genetics* **60**(2.12):2.12.1-2.12.18. https://doi.org/10.1002/0471142905.hg0212s60.

52. Peter H., Berggrav K., Thomas P., Pfeifer Y., Witte W., Templeton K., Bachmann T. T. (2012) Direct detection and genotyping of *Klebsiella pneumoniae* carbapenemases from

urine by use of a new DNA microarray test. *Journal of Clinical Microbiology* **50**(12):3990–3998. https://doi.org/10.1128/jcm.00990-12.

53. Gardner S. N., Thissen J. B., McLoughlin K. S., Slezak T., Jaing C. J. (2013) Optimizing SNP microarray probe design for high accuracy microbial genotyping. *Journal of Microbiological Methods* **94**(3):303–310. https://doi.org/10.1016/j.mimet.2013.07.006.

54. Costa J. M., Garcia-Hermoso D., Olivi M., Cabaret O., Farrugia C., Lecellier G., Dromer F., Bretagne S. (2010) Genotyping of *Candida albicans* using length fragment and high-resolution melting analyses together with minisequencing of a polymorphic microsatellite locus. *Journal of Microbiological Methods* **80**(3):306–309. https://doi.org/10.1016/j.mimet.2010.01.002.

55. Costa J. M., Cabaret O., Moukoury S., Bretagne S. (2011) Genotyping of the protozoan pathogen *Toxoplasma gondii* using high-resolution melting analysis of the repeated B1 gene. *Journal of Microbiological Methods* **86**(3):357–363. https://doi.org/10.1016/j.mimet.2011.06.017.

56. Makridakis N. M., Reichardt J. K. V. (2001) Multiplex automated primer extension analysis: Simultaneous genotyping of several polymorphisms. *Biotechniques* **31**(6):1374–1380. https://doi.org/10.2144/01316md05.

57. Eusebio N., Pinheiro T., Amorim A. A., Gamboa F., Saraiva L., Gusmão L., Amorim A., Araujo R. (2013) SNaPaer: A practical single nucleotide polymorphism multiplex assay for genotyping of *Pseudomonas aeruginosa*. *PLoS One* **8**(6). https://doi.org/10.1371/journal.pone.0066083.

58. Caramalho R., Gusmão L., Lackner M., Amorim A., Araujo R. (2013) SNaPAfu: A novel single nucleotide polymorphism multiplex assay for *Aspergillus fumigatus* direct detection, identification and genotyping in clinical specimens. *PLoS One* **8**(10). https://doi.org/10.1371/journal.pone.0075968.

59. Matsuda K. (2017) PCR-based detection methods for single-nucleotide polymorphism or mutation: Real-time PCR and its substantial contribution toward technological refinement. *Advances in Clinical Chemistry* **80**:45–72. https://doi.org/10.1016/bs.acc.2016.11.002.

60. Labbé G., Kruczkiewicz P., Robertson J., Mabon P., Schonfeld J., Kein D., Rankin M. A., Gopez M., Hole D., Son D., Knox N., Laing C. R., Bessonov K., Taboada E. N., Yoshida C., Ziebell K., Nichani A., Johnson R. P., Van Domselaar G., Nash J. H. E. (2021) Rapid and accurate SNP genotyping of clonal bacterial pathogens with BioHansel. *Microbial Genomics* **7**(9). https://doi.org/10.1099/mgen.0.000651.

61. Huang C. H., Chang M. T., Huang M. C., Lee F. L. (2011) Rapid identification of *Lactobacillus plantarum* group using the SNaPshot minisequencing assay. *Systematic and Applied Microbiology* **34**(8):586–589. https://doi.org/10.1016/j.syapm.2011.02.006.

62. Huang C. H., Chang M. T., Huang M. C., Wang L. T., Huang L., Lee F. L. (2012) Discrimination of the *Lactobacillus acidophilus* group using sequencing, species-specific PCR and SNaPshot mini-sequencing technology based on the *recA* gene. *Journal of the Science of Food and Agriculture* **92**(13):2703–2708. https://doi.org/10.1002/jsfa.5692.

63. Patiño-Navarrete R., Sanchis V. (2017) Evolutionary processes and environmental factors underlying the genetic diversity and lifestyles of *Bacillus cereus* group bacteria. *Research in Microbiology* **168**(4):309–318. https://doi.org/10.1016/j.resmic.2016.07.002.

64. Ehling-Schulz M., Lereclus D., Koehler T. M. (2019) The *Bacillus cereus* group: *Bacillus* species with pathogenic potential. *Microbiology Spectrum* **7**(3). https://doi.org/10.1128/microbiolspec.gpp3-0032-2018.

65. Derzelle S., Mendy C., Laroche S., Madani N. (2011) Use of high-resolution melting and melting temperature-shift assays for specific detection and identification of *Bacillus anthracis* based on single nucleotide discrimination. *Journal of Microbiological Methods* **87**(2):195–201. https://doi.org/10.1016/j.mimet.2011.08.005.

66. Ahmod N. Z., Gupta R. S., Shah H. N. (2011) Identification of a *Bacillus anthracis* specific indel in the *yeaC* gene and development of a rapid pyrosequencing assay for distinguishing *B. anthracis* from the *B. cereus* group. *Journal of Microbiological Methods* **87**(3):278–285. https://doi.org/10.1016/j.mimet.2011.08.015.

67. Esteban J., Muñoz-Egea M. C. (2016) *Mycobacterium bovis* and other uncommon members of the *Mycobacterium tuberculosis* complex. *Microbiology Spectrum* **4**(6). https://doi.org/10.1128/microbiolspec.tnmi7-0021-2016.

68. Borham M., Oreiby A., El-Gedawy A., Hegazy Y., Khalifa H. O., Al-Gaabary M., Matsumoto T. (2022) Review on bovine tuberculosis: An emerging disease associated with multidrug-resistant *Mycobacterium* species. *Pathogens* **11**(7):715. https://doi.org/10.3390/pathogens11070715.

69. Bouakaze C., Keyser C., De Martino S. J., Sougakoff W., Veziris N., Dabernat H., Ludes B. (2010) Identification and genotyping of *Mycobacterium tuberculosis* complex species by use of a SNaPshot minisequencing-based assay. *Journal of Clinical Microbiology* **48**(5):1758-1766. https://doi.org/10.1128/jcm.02255-09.

70. Girault G., Perrot L., Mick V., Ponsart C. (2022) High-resolution melting PCR as rapid genotyping tool for *Brucella* species. *Microorganisms* **10**(2):336. https://doi.org/10.3390/microorganisms10020336.

71. Foster J. T., Okinaka R. T., Svensson R., Shaw K., De B. K., Robison R. A., Probert W. S., Kenefic L. J., Brown W. D., Keim P. (2007) Real-time PCR assays of single-nucleotide polymorphisms defining the major *Brucella* clades. *Journal of Clinical Microbiology* **46**(1):296-301. https://doi.org/10.1128/jcm.01496-07.

72. Kivi M., Rodin S., Kupershmidt I., Lundin A., Tindberg Y., Granström M., Engstrand L. (2007) *Helicobacter pylori* genome variability in a framework of familial transmission. *BMC Microbiology* **7**(54). https://doi.org/10.1186/1471-2180-7-54.

73. Kabamba E. T., Tuan V. P., Yamaoka Y. (2018) Genetic populations and virulence factors of *Helicobacter pylori*. *Infection, Genetics and Evolution* **60**:109-116. https://doi.org/10.1016/j.meegid.2018.02.022.

74. Hanafiah A., Lopes B. S. (2020) Genetic diversity and virulence characteristics of *Helicobacter pylori* isolates in different human ethnic groups. *Infection, Genetics and Evolution* **78**:104135. https://doi.org/10.1016/j.meegid.2019.104135.

75. Amoako K. K., Thomas M. C., Janzen T. W., Goji N. (2017) Rapid SNP detection and genotyping of bacterial pathogens by pyrosequencing. *Methods in Molecular Biology* **1492**:203–220. https://doi.org/10.1007/978-1-4939-6442-0_15.

76. Besser J., Carleton H. A., Gerner-Smidt P., Lindsey R. L., Trees E. (2018) Next-generation sequencing technologies and their application to the study and control of bacterial

infections. *Clinical Microbiology and Infection* **24**(4):335-341. https://doi.org/10.1016/j.cmi.2017.10.013.

77. Blanc D. S., Magalhães B., Koenig I., Senn L., Grandbastien B. (2020) Comparison of whole genome (wg-) and core genome (cg-) MLST (BioNumericsTM) versus SNP variant calling for epidemiological investigation of *Pseudomonas aeruginosa*. *Frontiers in Microbiology* **11**(1729). https://doi.org/10.3389/fmicb.2020.01729.

78. Tettelin H., Riley D., Cattuto C., Medini D. (2008) Comparative genomics: The bacterial pan-genome. *Current Opinion in Microbiology* **11**(5):472–477. https://doi.org/10.1016/j.mib.2008.09.006.

79. Gilchrist C. A., Turner S. D., Riley M. F., Petri Jr W. A., Hewlett E. L. (2015) Whole-genome sequencing in outbreak analysis. *Clinical Microbiology Reviews* **28**(3):541–563. https://doi.org/10.1128/cmr.00075-13.

80. Green E. D., Watson J. D., Collins F. S. (2015) Human genome project: Twenty-five years of big biology. *Nature* **526**(7571):29–31. https://doi.org/10.1038/526029a.

81. Liu L., Li Y., Li S., Hu N., He Y., Pong R., Lin D., Lu L., Law M. (2012) Comparison of next-generation sequencing systems. *Journal of Biomedicine and Biotechnology* **2012**(7):251364. https://doi.org/10.1155/2012/251364.

82. Zhong Y., Xu F., Wu J., Schubert J., Li M. M. (2021) Application of next generation sequencing in laboratory medicine. *Annals of Laboratory Medicine* **41**(1):25-43. https://doi.org/10.3343/alm.2021.41.1.25.

83. Kumar P., Choudhary M., Jat B. S., Kumar B., Singh V., Kumar V., Singla D., Rakshit S. (2021) Skim sequencing: An advanced NGS technology for crop improvement. *Journal of Genetics* **100**(38). https://doi.org/10.1007/s12041-021-01285-3.

84. Baudhuin L. M., Lagerstedt S. A., Klee E. W., Fadra N., Oglesbee D., Ferber M. J. (2015) Confirming variants in next-generation sequencing panel testing by sanger sequencing. *The Journal of Molecular Diagnostics* **17**(4):456–461. https://doi.org/10.1016/j.jmoldx.2015.03.004.

85. Van Djik E. L., Auger H., Jaszczyszyn Y., Thermes C. (2014) Ten years of next-generation sequencing technology. *Trends in Genetics* **30**(9):418–426. https://doi.org/10.1016/j.tig.2014.07.001.

86. Vincent A. T., Derome N., Boyle B., Culley A. I., Charette S. J. (2017) Next-generation sequencing (NGS) in the microbiological world: How to make the most of your money. *Journal of Microbiological Methods* **138**:60–71. https://doi.org/10.1016/j.mimet.2016.02.016.

87. Schadt E. E., Turner S., Kasarskis A. (2010) A window into third-generation sequencing. *Human Molecular Genetics* **19**(R2): R227–240. https://doi.org/10.1093/hmg/ddq416.

88. Fu X., Gong L., Liu Y., Lai Q., Li G., Shao Z. (2021) *Bacillus pumilus* group comparative genomics: Toward pangenome features, diversity, and marine environmental adaptation. *Frontiers in Microbiology* **12**(571212). https://doi.org/10.3389/fmicb.2021.571212.

89. Teng W., Liao B., Chen M., Shu W. (2023) Genomic legacies of ancient adaptation illuminate GC-content evolution in bacteria. *Microbiology Spectrum* **11**(1). https://doi.org/10.1128/spectrum.02145-22.

90. Wang Z., Feng J. X., Li X. P., Zhang J. (2021) Whole-genome sequencing of *Micrococcus luteus* MT1691313, isolated from the Mariana trench. *Microbiology Resource Announcements* **10**(23). https://doi.org/10.1128/mra.00369-21.

91. Sair Jr M. H. (2008) The bacterial chromosome. *Critical Reviews in Biochemistry and Molecular Biology* **43**(2):89–134. https://doi.org/10.1080/10409230801921262.

92. Gordo I., Perfeito L., Sousa A. (2011) Fitness effects of mutations in bacteria. *Journal of Molecular Microbiology and Biotechnology* **21**(1–2):20–35. https://doi.org/10.1159/000332747.

93. Oehler J., Morrow C. A., Whitby M. C. (2023) Gene duplication and deletion caused by over-replication at a fork barrier. *Nature Communications* **14**(7730). https://doi.org/10.1038/s41467-023-43494-7.

94. Clancy S., Shaw K. M. (2008) DNA deletion and duplication and the associated genetic disorders. *Nature Education* **1**(1):23.

95. Partridge S. R., Kwong S. M., Firth N., Jensen S. O. (2018) Mobile genetic elements associated with antimicrobial resistance. *Clinical Microbiology Reviews* **31**(4). https://doi.org/10.1128/cmr.00088-17.

96. Casacuberta E., González J. (2013) The impact of transposable elements in environmental adaptation. *Molecular Ecology* **22**(6):1503–1517. https://doi.org/10.1111/mec.12170.

97. Gordo I., Perfeito L., Sousa A. (2011) Fitness effects of mutations in bacteria. *Journal of Molecular Microbiology and Biotechnology* **21**(1–2):20–35. https://doi.org/10.1159/000332747.

98. Loman N. J., Pallen M. J. (2015) Twenty years of bacterial genome sequencing. *Nature Reviews Microbiology* **13**(12):787–794. https://doi.org/10.1038/nrmicro3565.

99. Lin Y., Zhang R. R. (2011) Putative essential and core-essential genes in *Mycoplasma* genomes. *Scientific Reports* **1**(53). https://doi.org/10.1038/srep00053.

100. Martínez-Carranza E., Barajas H., Alcaraz L. D., Servín-González L., Ponce-Soto G. Y., Soberón-Chávez G. (2018) Variability of bacterial essential genes among closely related bacteria: The case of *Escherichia coli*. *Frontiers in Microbiology* **9**(1059). https://doi.org/10.3389/fmicb.2018.01059.

101. Zhang Z., Ren Q. (2015) Why are essential genes essential? - The essentiality of *Saccharomyces* genes. *Microbial Cell* **2**(8):280-287. https://doi.org/10.15698/mic2015.08.218.

102. Dilucca M., Cimini G., Giansanti A. (2018) Essentiality, conservation, evolutionary pressure and codon bias in bacterial genomes. *Gene* **663**:178–188. https://doi.org/10.1016/j.gene.2018.04.017.

103. Liu T., Luo H., Gao F. (2021) Position preference of essential genes in prokaryotic operons. *PLoS One* **16**(4). https://doi.org/10.1371/journal.pone.0250380.

104. Kania A. (2021) Harnessing the information theory and chaos game representation for pattern searching among essential and non-essential genes in bacteria. *Journal of Theoretical Biology* **531**(110917). https://doi.org/10.1016/j.jtbi.2021.110917.

105. Chen L., Zhang Y. H., Wang S. P., Zhang Y. H., Huang T., Cai Y. D. (2017) Prediction and analysis of essential genes using the enrichments of gene ontology and KEGG pathways. *PLoS One* **12**(9):1–22. https://doi.org/10.1371/journal.pone.0184129.

106. Koo B. M., Kritikos G., Farelli J. D., Todor H., Tong K., Kimsey H., Wapinski I., Galardini M., Cabal A., Peters J. M., Hachmann A. B., Rudner D. Z., Allen K. N., Typas A., Gross C. A. (2017) Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. *Cell Systems* **4**(3):291–305. https://doi.org/10.1016/j.cels.2016.12.013.

107. Latif A., Shehzad A., Niazi S., Zahid A., Ashraf W., Iqbal M. W., Rehman A., Riaz T., Aadil R. M., Khan I. M., Özogul F., Rocha J. M., Esatbeyoglu T., Korma S. A. (2023) Probiotics: Mechanism of action, health benefits and their application in food industries. *Frontiers in Microbiology* **14**(1216674). https://doi.org/10.3389/fmicb.2023.1216674.

108. Maftei N. M., Raileanu C. R., Balta A. A., Ambrose L., Boev M., Marin D. B., Lisa E. L. (2024) The potential impact of probiotics on human health: An update on their health-promoting properties. *Microorganisms* **12**(2):234. https://doi.org/10.3390/microorganisms12020234.

109. Plaza-Diaz J., Ruiz-Ojeda F. J., Gil-Campos M., Gil A. (2019) Mechanisms of action of probiotics. *Advances in Nutrition* **10**(Suppl 1): S49-S66. https://doi.org/10.1093/advances/nmy063.

110. Bodke H., Jogdand S. (2022) Role of probiotics in human health. *Cureus* **14**(11). https://doi.org/10.7759/cureus.31313.

111. Zawistowska-Rojek A., Zaręba T., Tyski S. (2022) Microbiological testing of probiotic preparations. *International Journal of Environmental Research and Public Health* **19**(9):5701. https://doi.org/10.3390/ijerph19095701.

112. Binda S., Hill C., Johansen E., Obis D., Pot B., Sanders M. E., Tremblay A., Ouwehand A. C. (2020) Criteria to qualify microorganisms as "probiotic" in foods and dietary supplements. *Frontiers in Microbiology* **11**(1662). https://doi.org/10.3389/fmicb.2020.01662.

113. Gul S., Durante-Mangoni E. (2024) Unravelling the puzzle: Health benefits of probiotics—A comprehensive review. *Journal of Clinical Medicine* **13**(5):1436. https://doi.org/10.3390/jcm13051436.

114. Lee Y. Y., Leow A. H. R., Chai P. F., Ali R. A. R., Lee W. S., Goh K. L. (2020) Use of probiotics in clinical practice with special reference to diarrheal diseases: A position statement of the Malaysian Society of Gastroenterology and Hepatology. *Journal of Gastroenterology and Hepatology Open* **5**(1):11-19. https://doi.org/10.1002/jgh3.12469.

115. Caliendo T., Hilas O. (2021) Probiotic use in gastrointestinal disorders. *US Pharmacist* **46**(12):18–20.

116. AL-Smadi K., Leite-Silva V. R., Filho N. A., Lopes P. S., Mohammed Y. (2023) Innovative approaches for maintaining and enhancing skin health and managing skin diseases through microbiome-targeted strategies. *Antibiotics* **12**(12):1698. https://doi.org/10.3390/antibiotics12121698.

117. Rau S., Gregg A., Yaceczko S., Limketkai B. (2024) Prebiotics and probiotics for gastrointestinal disorders. *Nutrients* **16**(6):778. https://doi.org/10.3390/nu16060778.

118. Milner E., Stevens B., An M., Lam V., Ainsworth M., Dihle P., Stearns J., Dombrowski A., Rego D., Segars K. (2021) Utilizing probiotics for the prevention and treatment of gastrointestinal diseases. *Frontiers in Microbiology* **12**(689958). https://doi.org/10.3389/fmicb.2021.689958.

119. García-Santos J. A., Nieto-Ruiz A., García-Ricobaraza M., Cerdó T., Campoy C. (2023) Impact of probiotics on the prevention and treatment of gastrointestinal diseases in the pediatric population. *International Journal of Molecular Sciences* **24**(11):1–29. https://doi.org/10.3390/ijms24119427.

120. Chelliah R., Kim E. J., Daliri E. B. M., Antony U., Oh D. H. (2021) In vitro probiotic evaluation of *Saccharomyces boulardii* with antimicrobial spectrum in a *Caenorhabditis elegans* model. *Foods* **10**(6):1428. https://doi.org/10.3390/foods10061428.

121. Cullen C. M., Aneja K. K., Beyhan S., Cho C. E., Woloszynek S., Convertino M., McCoy S. J., Zhang Y., Anderson M. Z., Alvarez-Ponce D., Smirnova E., Karstens L., Dorrestein

P. C., Li H., Gupta A. S., Cheung K., Powers J. G., Zhao Z., Rosen G. L. (2020) Emerging priorities for microbiome research. *Frontiers in Microbiology* **11**(136):1–27. https://doi.org/10.3389/fmicb.2020.00136.

122. Fyhrquist N., Muirhead G., Prast-Nielsen S., Jeanmougin M., Olah P., Skoog T., Jules-Clement G., Feld M., Barrientos-Somarribas M., Sinkko H., Van Den Bogaard E. H., Zeeuwen P. L. J. M., Rikken G., Schalkwijk J., Niehues H., Däubener W., Eller S. K., Alexander H., Pennino D., Suomela S., Tessas I., Lybeck E., Baran A. M., Darban H., Gangwar R. S., Gerstel U., Jahn K., Karisola P., Yan L., Hansmann B., Katayama S., Meller S., Bylesjö M., Hupé P., Levi-Schaffer F., Greco D., Ranki A., Schröder J. M., Barker J., Kere J., Tsoka S., Lauerma A., Soumelis V., Nestle F. O., Homey B., Andersson B., Alenius H. (2019) Microbe-host interplay in atopic dermatitis and psoriasis. *Nature Communications* **10**(4703):1–15. https://doi.org/10.1038/s41467-019-12253-y.

123. Dityen K., Soonthornchai W., Kueanjinda P., Kullapanich C., Tunsakul N., Somboonna N., Wongpiyabovorn J. (2022) Analysis of cutaneous bacterial microbiota of Thai patients with seborrheic dermatitis. *Experimental Dermatology* **31**(12):1949–1955. https://doi.org/10.1111/exd.14674.

124. De Almeida C. V., Antiga E., Lulli M. (2023) Oral and topical probiotics and postbiotics in skincare and dermatological therapy: A concise review. *Microorganisms* **11**(6):1420. https://doi.org/10.3390/microorganisms11061420.

125. Yang H. L., Sun Y. Z., Hu X., Ye J. dan, Lu K. L., Hu L. H., Zhang J. J. (2019) *Bacillus pumilus* SE5 originated PG and LTA tuned the intestinal TLRs/MyD88 signalling and microbiota in grouper (*Epinephelus coioides*). *Fish & Shellfish Immunology* **88**:266–271. https://doi.org/10.1016/j.fsi.2019.03.005.

126. Terekhov S. S., Nazarov A. S., Mokrushina Y. A., Baranova M. N., Potapova N. A., Malakhova M. V., Ilina E. N., Smirnov I. V., Gabibov A. G. (2020) Deep functional profiling facilitates the evaluation of the antibacterial potential of the antibiotic amicoumacin. *Antibiotics* **9**(4):157. https://doi.org/10.3390/antibiotics9040157.

127. Zidour M., Belguesmia Y., Cudennec B., Grard T., Flahaut C., Souissi S., Drider D. (2019) Genome sequencing and analysis of *Bacillus pumilus* ICVB403 isolated from *Acartia tonsa* copepod eggs revealed surfactin and bacteriocin production: Insights on

anti-*Staphylococcus* activity. *Probiotics and Antimicrobial Proteins* **11**(3):990–998. https://doi.org/10.1007/s12602-018-9461-4.

128. Thy H. T. T., Tri N. N., Quy O. M., Fotedar R., Kannika K., Unajak S., Areechon N. (2017) Effects of the dietary supplementation of mixed probiotic spores of *Bacillus amyloliquefaciens* 54A, and *Bacillus pumilus* 47B on growth, innate immunity and stress responses of striped catfish (*Pangasianodon hypophthalmus*). *Fish & Shellfish Immunology* **60**:391–399. https://doi.org/10.1016/j.fsi.2016.11.016.

129. Bilal M., Achard C., Barbe F., Chevaux E., Ronholm J., Zhao X. (2021) *Bacillus pumilus* and *Bacillus subtilis* promote early maturation of faecal microbiota in broiler chickens. *Microorganisms* **9**(9):1899. https://doi.org/10.3390/microorganisms9091899.

130. Duc L. H., Hong H. A., Barbosa T. M., Henriques A. O., Cutting S. M. (2004) Characterization of *Bacillus* probiotics available for human use. *Applied and Environmental Microbiology* **70**(4):2161-2171. https://doi.org/10.1128/AEM.70.4.2161-2171.2004.

131. Kotowicz N., Bhardwaj R. K., Ferreira W. T., Hong H. A., Olender A., Ramirez J., Cutting S. M. (2019) Safety and probiotic evaluation of two *Bacillus* strains producing antioxidant compounds. *Beneficial Microbes* **10**(7):759-772. https://doi.org/10.3920/BM2019.0040.

132. Duanis-Assaf D., Steinberg D., Shemesh M. (2020) Efficiency of *Bacillus subtilis* metabolism of sugar alcohols governs its probiotic effect against cariogenic *Streptococcus mutans*. *Artificial Cells, Nanomedicine, and Biotechnology* **48**(1):1222–1230. https://doi.org/10.1080/21691401.2020.1822855.

133. Shleeva M. O., Kondratieva D. A., Kaprelyants A. S. (2023) *Bacillus licheniformis*: A producer of antimicrobial substances, including antimycobacterials, which are feasible for medical applications. *Pharmaceutics* **15**(7):1893. https://doi.org/10.3390/pharmaceutics15071893.

134. Suva M. A., Sureja V. P., Kheni D. B. (2016) Novel insight on probiotic *Bacillus subtilis*: Mechanism of action and clinical applications. *Journal of Current Research in Scientific Medicine* **2**(2):65–72. https://doi.org/10.4103/2455-3069.198381.

135. Lv P., Song Y., Liu C., Yu L., Shang Y., Tang H., Sun S., Wang F. (2020) Application of *Bacillus subtilis* as a live vaccine vector: A review. The *Journal of Veterinary Medical Science* **82**(11):1693-1699. https://doi.org/10.1292/jvms.20-0363.

136. Akinsemolu A. A., Onyeaka H., Odion S., Adebanjo I. (2024) Exploring *Bacillus subtilis*: Ecology, biotechnological applications, and future prospects. *Journal of Basic Microbiology* **64**(6):2300614. https://doi.org/10.1002/jobm.202300614.

137. Trotter R. E., Vazquez A. R., Grubb D. S., Freedman K. E., Grabos L. E., Jones S., Gentile C. L., Melby C. L., Johnson S. A., Weir T. L. (2020) *Bacillus subtilis* DE111 intake may improve blood lipids and endothelial function in healthy adults. *Beneficial Microbes* **11**(7):621-630. https://doi.org/10.3920/BM2020.0039.

138. Yan F. F., Wang W. C., Cheng H. W. (2020) *Bacillus subtilis*–based probiotic promotes bone growth by inhibition of inflammation in broilers subjected to cyclic heating episodes. *Poultry Science* **99**(11):5252-5260. https://doi.org/10.1016/j.psj.2020.08.051.

139. Jiang S., Yan F. F., Hu J. Y., Mohammed A., Cheng H. W. (2021) *Bacillus subtilis*-based probiotic improves skeletal health and immunity in broiler chickens exposed to heat stress. *Animals* **11**(6):1494. https://doi.org/10.3390/ani11061494.

140. Khochamit N., Siripornadulsil S., Sukon P., Siripornadulsil W. (2020) *Bacillus subtilis* and lactic acid bacteria improve the growth performance and blood parameters and reduce *Salmonella* infection in broilers. *Veterinary World* **13**(12):2663-2672. https://doi.org/10.14202/vetworld.2020.2663-2672.

141. Paytuví-Gallart A., Sanseverino W., Winger A. M. (2020) Daily intake of probiotic strain *Bacillus subtilis* DE111 supports a healthy microbiome in children attending day-care. *Beneficial Microbes* **11**(7):611–620. https://doi.org/10.3920/BM2020.0022.

142. Rönkä E., Malinen E., Saarela M., Rinta-Koski M., Aarnikunnas J., Palva A. (2003) Probiotic and milk technological properties of *Lactobacillus brevis*. *International Journal of Food Microbiology* **83**(1):63–74. https://doi.org/10.1016/S0168-1605(02)00315-X.

143. Ramos C. L., Thorsen L., Schwan R. F., Jespersen L. (2013) Strain-specific probiotics properties of *Lactobacillus fermentum*, *Lactobacillus plantarum* and *Lactobacillus brevis* isolates from Brazilian food products. *Food Microbiology* **36**(1):22–29. https://doi.org/10.1016/j.fm.2013.03.010.

144. Yakabe T., Moore E. L., Yokota S., Sui H., Nobuta Y., Fukao M., Palmer H., Yajimi N. (2009) Safety assessment of *Lactobacillus brevis* KB290 as a probiotic strain. *Food and Chemical Toxicology* **47**(10):2450–2453. https://doi.org/10.1016/j.fct.2009.07.001.

145. Hojjati M., Behabahani B. A., Falah F. (2020) Aggregation, adherence, anti-adhesion and antagonistic activity properties relating to surface charge of probiotic *Lactobacillus brevis* gp104 against *Staphylococcus aureus*. *Microbial Pathogenesis* **147**:104420. https://doi.org/10.1016/j.micpath.2020.104420.

146. Fang F., Xu J., Li Q., Xia X., Du G. (2018) Characterization of a *Lactobacillus brevis* strain with potential oral probiotic properties. *BMC Microbiology* **18**(221). https://doi.org/10.1186/s12866-018-1369-3.

147. Alfano A., Perillo F., Fusco A., Savio V., Corsaro M. M., Donnarumma G., Schiraldi C., Cimini D. (2020) *Lactobacillus brevis* CD2: Fermentation strategies and extracellular metabolites characterization. *Probiotics and Antimicrobial Proteins* **12**(4):1542–1554. https://doi.org/10.1007/s12602-020-09651-w.

148. Kunduhoglu B., Hacioglu S. (2020) Probiotic potential and gluten hydrolysis activity of *Lactobacillus brevis* KT16-2. *Probiotics and Antimicrobial Proteins* **13**(3):720–733. https://doi.org/10.1007/s12602-020-09723-x.

149. Sah B. N. P., Vasiljevic T., McKechnie S., Donkor O. N. (2015) Effect of refrigerated storage on probiotic viability and the production and stability of antimutagenic and antioxidant peptides in yogurt supplemented with pineapple peel. *Journal of Dairy Science* **98**(9):5905–5916. https://doi.org/10.3168/jds.2015-9450.

150. Kariyawasam K. M. G. M. M., Yang S. J., Lee N. K., Paik H. D. (2020) Probiotic properties of *Lactobacillus brevis* KU200019 and synergistic activity with fructooligosaccharides in antagonistic activity against foodborne pathogens. *Food Science of Animal Resources* **40**(2):297-310. https://doi.org/10.5851/kosfa.2020.e15.

151. Romano N., Marro M., Marsal M., Loza-Álvarez P., Gomez-Zavaglia A. (2021) Fructose derived oligosaccharides prevent lipid membrane destabilization and DNA conformational alterations during vacuum-drying of *Lactobacillus delbrueckii* subsp. *bulgaricus*. *Food Research International* **143**(110235). https://doi.org/10.1016/j.foodres.2021.110235.

152. Ferreira T. H., Maximiano P., Simões P. N. (2025) Effect of fructooligosaccharides in full-hydrated lactic acid bacteria membrane models during thermal stress: A molecular simulation study. *Food Research International* **200**:115475. https://doi.org/10.1016/j.foodres.2024.115475.

153. Pereira G. V. de M., Coelho B. de O., Júnior A. I. M., Thomaz-Soccol V., Soccol C. R. (2018) How to select a probiotic? A review and update of methods and criteria. *Biotechnology Advances* **36**(8):2060–2076. https://doi.org/10.1016/j.biotechadv.2018.09.003.

154. Mu Q., Tavella V. J., Luo X. M. (2018) Role of *Lactobacillus reuteri* in human health and diseases. *Frontiers in Microbiology* **9**(757). https://doi.org/10.3389/fmicb.2018.00757.

155. Frerejacques M., Rousselle C., Gauthier L., Cottet-Emard S., Derobert L., Roynette A., Lerch T. Z., Changey F. (2020) Human skin bacterial community response to probiotic (*Lactobacillus reuteri* DSM 17938) introduction. *Microorganisms* **8**(8):1–9. https://doi.org/10.3390/microorganisms8081223.

156. Hou C., Liu H., Zhang J., Zhang S., Yang F., Zeng X., Thacker P. A., Zhang G., Qiao S. (2015) Intestinal microbiota succession and immunomodulatory consequences after introduction of *Lactobacillus reuteri* I5007 in neonatal piglets. *PLoS One* **10**(3). https://doi.org/10.1371/journal.pone.0119505.

157. Deng Z., Dai T., Zhang W., Zhu J., Luo X. M., Fu D., Liu J., Wang H. (2020) Glyceraldehyde-3-phosphate dehydrogenase increases the adhesion of *Lactobacillus reuteri* to host mucin to enhance probiotic effects. *International Journal of Molecular Sciences* **21**(24):1–16. https://doi.org/10.3390/ijms21249756.

158. Zhang W. M., Wang H. F., Gao K., Wang C., Liu L., Liu J. X. (2015) *Lactobacillus reuteri* glyceraldehyde-3-phosphate dehydrogenase functions in adhesion to intestinal epithelial cells. *Canadian Journal of Microbiology* **61**(5):373–380. https://doi.org/10.1139/cjm-2014-0734.

159. Hjern A., Lindblom K., Reuter A., Silfverdal S. A. (2020) A systematic review of prevention and treatment of infantile colic. *Acta Paediatrica* **109**(9):1733–1744. https://doi.org/10.1111/apa.15247.

160. Gupta N., Ferreira J., Hong C. H. L., Tan K. S. (2020) *Lactobacillus reuteri* DSM 17938 and ATCC PTA 5289 ameliorates chemotherapy-induced oral mucositis. *Scientific Reports* **10**(16189):1–11. https://doi.org/10.1038/s41598-020-73292-w.

161. Stepanov V. G., Tirumalai M. R., Montazari S., Checinska A., Venkateswaran K., Fox G. E. (2016) *Bacillus pumilus* SAFR-032 genome revisited: Sequence update and re-annotation. *PLoS One* **11**(6). https://doi.org/10.1371/journal.pone.0157331.

162. Tirumalai M. R., Stepanov V. G., Wünsche A., Montazari S., Gonzalez R. O., Venkateswaran K., Fox G. E. (2018) *Bacillus safensis* FO-36b and *Bacillus pumilus* SAFR-032: A whole genome comparison of two spacecraft assembly facility isolates. *BMC Microbiology* **18**(57). https://doi.org/10.1186/s12866-018-1191-y.

163. Srivatsan A., Han Y., Peng J., Tehranchi A. K., Gibbs R., Wang J. D., Chen R. (2008) High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. *PLoS Genetics* **4**(8): e1000139. https://doi.org/10.1371/journal.pgen.1000139.

164. Bergsveinson J., Pittet V., Ewen E., Baecker N., Ziola B. (2015) Genome sequence of rapid beer-spoiling isolate *Lactobacillus brevis* BSO 464. *Genome Announcements* **3**(6). https://doi.org/10.1128/genomea.01411-15.

165. Bergsveinson J., Baecker N., Pittet V., Ziola B. (2015) Role of plasmids in *Lactobacillus brevis* BSO 464 hop tolerance and beer spoilage. *Applied and Environmental Microbiology* **81**(4):1234–1241. https://doi.org/10.1128/aem.02870-14.

166. *Bacillus pumilus* SAFR-032, complete sequence. 2025. Available: https://www.ncbi.nlm.nih.gov/nuccore/NC_009848 [date visited: 2/7/2025].

167. *Bacillus subtilis* subsp. *subtilis* str. 168 chromosome, whole genome shotgun sequence. 2025. Available: https://www.ncbi.nlm.nih.gov/nuccore/NZ_CM000487 [date visited: 2/7/2025].

168. *Levilactobacillus brevis* BSO 464 chromosome. 2024. Available: https://www.ncbi.nlm.nih.gov/nuccore/NZ_CP005977 [date visited: 2/7/2025].

169. *Limosilactobacillus reuteri* SD2112, complete sequence. 2024. Available: https://www.ncbi.nlm.nih.gov/nuccore/NC_015697 [date visited: 2/7/2025].

170. Toymentseva A. A., Mascher T., Sharipova M. R. (2017) Regulatory characteristics of *Bacillus pumilus* protease promoters. *Current Microbiology* **74**(5):550–559. https://doi.org/10.1007/s00284-017-1212-3.

171. Xie C. Y., Li W. J., Feng H. (2023) Tuning transcription factor *degU* for developing extracellular protease overproducer in *Bacillus pumilus*. *Microbial Cell Factories* **22**(163). https://doi.org/10.1186/s12934-023-02177-0.

172. P15770 AROE_ECOLI. 2025. Available: https://www.uniprot.org/uniprotkb/P15770/entry [date visited: 26/6/2025].

173. P0AES6 GYRB_ECOLI. 2025. Available: https://www.uniprot.org/uniprotkb/P0AES6/entry [date visited: 26/6/2025].

174. P23367 MUTL_ECOLI. 2025. Available: https://www.uniprot.org/uniprotkb/P23367/entry [date visited: 26/6/2025].

175. Jiang N., Hong B., Luo K., Li Y., Fu H., Wang J. (2023) Isolation of *Bacillus subtilis* and *Bacillus pumilus* with anti-*Vibrio parahaemolyticus* activity and identification of the anti-*Vibrio parahaemolyticus* substance. *Microorganisms* **11**(7):1667. https://doi.org/10.3390/microorganisms11071667.

176. Hollensteiner J., Poehlein A., Daniel R., Liesegang H., Vidal S., Wemheuer F. (2017) Draft genome sequence of *Bacillus pumilus* strain GM3FR, an endophyte isolated from aerial plant tissues of *Festuca rubra* L. *Genome Announcements* **5**(13). https://doi.org/10.1128/genomea.00085-17.

177. Husni A. A. A., Ismail S. I., Jaafar N. M., Zulperi D. (2021) Current classification of the *Bacillus pumilus* group species, the rubber-pathogenic bacteria causing trunk bulges disease in Malaysia as assessed by MLSA and multi rep-PCR approaches. The *Plant Pathology Journal* **37**(3):243-257. https://doi.org/10.5423/PPJ.OA.02.2021.0017.

178. P37870 RPOB_BACSU. 2025. Available: https://www.uniprot.org/uniprotkb/P37870/entry [date visited: 26/6/2025].

179. P0A879 TRPB_ECOLI. 2025. Available: https://www.uniprot.org/uniprotkb/P0A879/entry [date visited: 26/6/2025].

180. La Grange D. C., Claeyssens M., Pretorius I. S., Van Zyl W. H. (2000) Coexpression of the *Bacillus pumilus* β-xylosidase (xynB) gene with the *Trichoderma reesei* β-xylanase 2 (xyn2) gene in the yeast *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology* **54**(2):195–200. https://doi.org/10.1007/s002530000372.

181. Wolf D., Rippa V., Mobarec J. C., Sauer P., Adlung L., Kolb P., Bischofs I. B. (2016) The quorum-sensing regulator *comA* from *Bacillus subtilis* activates transcription using topologically distinct DNA motifs. *Nucleic Acids Research* **44**(5):2160–2172. https://doi.org/10.1093/nar/gkv1242.

182. Ashikaga S., Nanamiya H., Ohashi Y., Kawamura F. (2000) Natural genetic competence in *Bacillus subtilis* Natto OK2. *Journal of Bacteriology* **182**(9):2411-2415. https://doi.org/10.1128/jb.182.9.2411-2415.2000.

183. Q99027 COMP_BACSU. 2025. Available: https://www.uniprot.org/uniprotkb/Q99027/entry [date visited: 26/6/2025].

184. P37584 CSAA_BACSU. 2025. Available: https://www.uniprot.org/uniprotkb/P37584/entry [date visited: 26/6/2025].

185. Xu S., Cao Q., Liu Z., Chen J., Yan P., Li B., Yu Y. (2022) Transcriptomic analysis reveals the role of tmRNA on biofilm formation in *Bacillus subtilis*. *Microorganisms* **10**(7):1338. https://doi.org/10.3390/microorganisms10071338.

186. P13800 DEGU_BACSU. 2025. Available: https://www.uniprot.org/uniprotkb/P13800/history [date visited: 26/6/2025].

187. Matavacas J., Hallgren J., Von Wachenfeldt C. (2023) *Bacillus subtilis* forms twisted cells with cell wall integrity defects upon removal of the molecular chaperones *dnaK* and trigger factor. *Frontiers in Microbiology* **13**(988768). https://doi.org/10.3389/fmicb.2022.988768.

188. P0A6F5 CH60_ECOLI. 2025. Available: https://www.uniprot.org/uniprotkb/P0A6F5/history [date visited: 26/6/2025].

189. P9WG47 GYRA_MYCTU. 2025. Available: https://www.uniprot.org/uniprotkb/P9WG47/history [date visited: 26/6/2025].

190. Beilen J. van, Blohmke C. J., Folkerts H., de Boer R., Zakrzewska A., Kulik W., Vaz F. M., Brul S., Beek A. T. (2016) *RodZ* and *pgsA* play intertwined roles in membrane homeostasis of *Bacillus subtilis* and resistance to weak organic acid stress. *Frontiers in Microbiology* **7**(1633). https://doi.org/10.3389/fmicb.2016.01633.

191. P37870 RPOB_BACSU. 2025. Available: https://www.uniprot.org/uniprotkb/P37870/entry [date visited: 26/6/2025].

192. Gong L., Ren C., Xu Y. (2020) *GlnR* negatively regulates glutamate-dependent acid resistance in *Lactobacillus brevis*. *Applied and Environmental Microbiology* **86**(7). https://doi.org/10.1128/AEM.02615-19.

193. P0AES6 GYRB_ECOLI. 2025. Available: https://www.uniprot.org/uniprotkb/P0AES6/history [date visited: 26/6/2025].

194. Xu Z., Xu R., Soteyome T., Deng Y., Chen L., Liang Y., Bai C., Huang T., Liu J., Harro J. M., Kjellerup B. V. (2020) Genomic analysis of a hop-resistance *Lactobacillus brevis* strain responsible for food spoilage and capable of entering into the VBNC state. *Microbial Pathogenesis* **145**(104186). https://doi.org/10.1016/j.micpath.2020.104186.

195. Sharma A., Kaur J., Lee S., Park Y. S. (2019) Tracking of intentionally inoculated lactic acid bacteria strains in yogurt and probiotic powder. *Microorganisms* **8**(5). https://doi.org/10.3390/microorganisms8010005.

196. De Vuyst L., González-Alonso V., Wardhana Y. R., Pradal I. (2023) Taxonomy and species diversity of sourdough lactic acid bacteria. p. 97–160. In: Handbook on sourdough biotechnology (Eds. Gobbetti M., Gänzle M.). Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-031-23084-4_6.

197. P0A7Z4 RPOA_ECOLI. 2025. Available: https://www.uniprot.org/uniprotkb/P0A7Z4/entry [date visited: 26/6/2025].

198. 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. *Journal of Bacteriology* **200**(13). https://doi.org/10.1128/jb.00607-17.

199. Q81G39 DLTA_BACCR. 2025. Available: https://www.uniprot.org/uniprotkb/Q81G39/entry [data visited: 26/6/2025].

200. Zhu S., Zhang Y., Wang J., Zhang C., Liu X. (2022) Investigating differential expressed genes of *Limosilactobacillus reuteri* LR08 regulated by soybean protein and peptides. *Foods* **11**(9):1251. https://doi.org/10.3390/foods11091251.

201. Rajput A., Chauhan S. M., Mohite O. S., Hyun J. C., Ardalani O., Jahn L. J., Sommer M. O., Palsson B. O. (2023) Pangenome analysis reveals the genetic basis for taxonomic classification of the *Lactobacillaceae* family. *Food Microbiology* **115**(104334). https://doi.org/10.1016/j.fm.2023.104334.

202. Oh P. L., Benson A. K., Peterson D. A., Patil P. B., Moriyama E. N., Roos S., Walter J. (2010) Diversification of the gut symbiont *Lactobacillus reuteri* as a result of host-driven evolution. *The ISME Journal* **4**(3):377–387. https://doi.org/10.1038/ismej.2009.123.

203. Alanazi A. S., Qureshi K. A., Elhassan G. O., El-Agamy E. I. (2016) Isolation, purification and characterization of antimicrobial agent antagonistic to *Escherichia coli* ATCC 10536 produced by *Bacillus pumilus* SAFR-032 isolated from the soil of Unaizah, Al Qassim province of Saudi Arabia. *Pakistan Journal of Biological Sciences* **19**(5):191–201. https://doi.org/10.3923/pjbs.2016.191.201.

204. Ji S., Li W., Xin H., Wang S., Cao B. (2015) Improved production of sublancin 168 biosynthesized by *Bacillus subtilis* 168 using chemometric methodology and statistical experimental designs. *BioMed Research International* **2015**(687915). https://doi.org/10.1155/2015/687915.

205. Bocchi M. B., Perna A., Cianni L., Vitiello R., Greco T., Maccauro G., Perisano C. (2020) A rare case of *Bacillus megaterium* soft tissues infection. *Acta Biomedica* **91**(Suppl 14). https://doi.org/10.23750/abm.v91i14-s.10849.

206. Brouwer S., Rivera-Hernandez T., Curren B. F., Harbison-Price N., De Oliveira D. M. P., Jespersen M. G., Davies M. R., Walker M. J. (2023) Pathogenesis, epidemiology and control of group A *Streptococcus* infection. *Nature Reviews Microbiology* **21**(7):431-447. https://doi.org/10.1038/s41579-023-00865-7.

207. Meletis G., Ketikidis A. L. S., Floropoulou N., Tychala A., Kagkalou G., Vasilaki O., Mantzana P., Skoura L., Protonotariou E. (2023) Antimicrobial resistance rates of *Streptococcus pyogenes* in a greek tertiary care hospital: 6-year data and literature review. *New Microbiologica* **46**(1):37–42.

208. Al K. F., Daisley B. A., Chanyi R. M., Bjazevic J., Razvi H., Reid G., Burton J. P. (2020) Oxalate-degrading *Bacillus subtilis* mitigates urolithiasis in a *Drosophila melanogaster* model. *mSphere* **5**(5). https://doi.org/10.1128/msphere.00498-20.

209. Yu T., Kong J., Zhang L., Gu X., Wang M., Guo T. (2019) New crosstalk between probiotics *Lactobacillus plantarum* and *Bacillus subtilis*. *Scientific Reports* **9**(13151). https://doi.org/10.1038/s41598-019-49688-8.

210. Watanabe J., Hashimoto N., Yin T., Sandagdorj B., Arakawa C., Inoue T., Suzuki S. (2021) Heat-killed *Lactobacillus brevis* KB290 attenuates visceral fat accumulation induced by high-fat diet in mice. *Journal of Applied Microbiology* **131**(4):1998-2009. https://doi.org/10.1111/jam.15079.

211. Siciliano R. A., Reale A., Mazzeo M. F., Morandi S., Silvetti T., Brasca M. (2021) Paraprobiotics: A new perspective for functional foods and nutraceuticals. *Nutrients* **13**(4):1225. https://doi.org/10.3390/nu13041225.

212. Mehta J. P., Ayakar S., Singhal R. S. (2023) The potential of paraprobiotics and postbiotics to modulate the immune system: A review. *Microbiological Research* **275**(127449). https://doi.org/10.1016/j.micres.2023.127449.

213. Satomi S., Khanum S., Miller P., Suzuki S., Suganuma H., Heiser A., Gupta S. K. (2020) Short communication: Oral administration of heat-killed *Lactobacillus brevis* KB290 in combination with retinoic acid provides protection against Influenza virus infection in mice. *Nutrients* **12**(10):2925. https://doi.org/10.3390/nu12102925.

214. Wang G., Huang S., Cai S., Yu H., Wang Y., Zeng X., Qiao S. (2020) *Lactobacillus reuteri* ameliorates intestinal inflammation and modulates gut microbiota and metabolic disorders in dextran sulfate sodium-induced colitis in mice. *Nutrients* **12**(8):2298. https://doi.org/10.3390/nu12082298.

215. Yang F., Wang A., Zeng X., Hou C., Liu H., Qiao S. (2015) *Lactobacillus reuteri* I5007 modulates tight junction protein expression in IPEC-J2 cells with LPS stimulation and

in newborn piglets under normal conditions. *BMC Microbiology* **15**(32). https://doi.org/10.1186/s12866-015-0372-1.

216. Lee G., Heo S., Kim T., Na H. E., Park J., Lee E., Lee J. H., Jeong D. W. (2022) Discrimination of *Bacillus subtilis* from other *Bacillus* species using specific oligonucleotide primers for the pyruvate carboxylase and shikimate dehydrogenase genes. *Journal of Microbiology and Biotechnology* **32**(8):1011-1016. https://doi.org/10.4014/jmb.2205.05014.

217. Fan X., Li X., Zhang T., Xu J., Shi Z., Wu Z., Wu J., Pan D., Du L. (2021) A novel qPCR method for the detection of lactic acid bacteria in fermented milk. *Foods* **10**(12):3066. https://doi.org/10.3390/foods10123066.

218. Zeigler D. R., Prágai Z., Rodriguez S., Chevreux B., Muffler A., Albert T., Bai R., Wyss M., Perkins J. B. (2008) The origins of 168, W23, and other *Bacillus subtilis* legacy strains. *Journal of Bacteriology* **190**(21):6983–6995. https://doi.org/10.1128/jb.00722-08.

219. Salvetti E., O'Toole P. W. (2017) The genomic basis of *Lactobacilli* as health-promoting organisms. *Microbiology Spectrum* **5**(3). https://doi.org/10.1128/microbiolspec.bad-0011-2016.