CHARACTERIZATION OF BACILLUS ISOLATES USING WHOLE GENOME SEQUENCING ANALYSIS AND APPLICATION AS A POTENTIAL FOOD PROBIOTIC

Miss Gauri Khullar

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Food Science and Technology
Department of Food Technology
FACULTY OF SCIENCE
Chulalongkorn University
Academic Year 2019

การศึกษาลักษณะของบาซิลลัสไอโซเลทโดยการวิเคราะห์จีโนมแบบสมบูรณ์และความเป็นไปได้ในการประยุกต์ใช้เป็นโพรไบโอติกในอาหาร

น.ส.เการี คูลลาร์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์และเทคโนโลยีทางอาหาร ภาควิชาเทคโนโลยีทางอาหาร
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2562

Thesis Title	CHARACTERIZATION OF BACILLUS ISOLATES USING WHOLE GENOME SEQUENCING ANALYSIS AND APPLICATION AS A POTENTIAL FOOD PROBIOTIC
By	Miss Gauri Khullar
Field of Study	Food Science and Technology
Thesis Advisor	CHEUNJIT PRAKITCHIWATTANA
Thesis Co Advisor	PINIDPHON PROMBUTARA

Accepted by the FACULTY OF SCIENCE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

..... Dean of the FACULTY OF SCIENCE
(POLKIT SANGVANICH)

THESIS COMMITTEE

..... Chairman
(SUMATE TANTRATIAN)

..... Thesis Advisor
(CHEUNJIT PRAKITCHIWATTANA)

..... Thesis Co-Advisor
(PINIDPHON PROMBUTARA)

..... Examiner
(SARN SETTACHAIMONGKON)

..... External Examiner
(Pawinee Deetae)

เการี คูลลาร์ : การศึกษาลักษณะของบาซิลลัสไอโซเลทโดยการวิเคราะห์จีโนมแบบสมบูรณ์และความเป็นไปได้ในการประยุกต์ใช้เป็นโพรไบโอติกในอาหาร. (CHARACTERIZATION OF BACILLUS ISOLATES USING WHOLE GENOME SEQUENCING ANALYSIS AND APPLICATION AS A POTENTIAL FOOD PROBIOTIC) อ.ที่ปรึกษาหลัก : ชื่นจิต ประกิตชัยวัฒนา, อ.ที่ปรึกษาร่วม : พินิตพล พรหมบุตร

จากการประเมินสมบัติการเป็นโพรไบโอกติกของ Bacillus ไอโซเลทที่มีสมบัติเบื้องต้นต้านจุลินทรีย์ โดยการวิเคราะห์จีโนมแบบสมบูรณ์ (WSG) ระบุสายพันธ์ของ 6-2, 63-11 และ 78-1 ได้เป็น *Bacillus velezensis* (98.16%), *Bacillus infantis* (91.21%) and *Bacillus amyloliquefaciens* (99.06%) ตามลำดับ เมื่อทำนาย K-mer resistance พบโปรตีน cfr(B) และ tet(L) resistance (6-2, 78-1) พบยีนสร้างแบคเทอริโอซิน และสารเมตาบอไลท์ (6-2, 78-1) terpenoid (63-11) hemolysin III (6-2, 78-1) และ hlyIII homolog (63-11) extracellular protease (6-2, 63-11) และ cell-bound protease (78-1) เมื่อวิเคาะห์จีโนมพบโปรตีน F1F0 ATP Acid tolerance, chaperonin (groEL, groES), และ general stress response (DnaK) ที่เป็น EAL domain (สร้างไบโอฟิล์ม) สังเคราะห์แฟลเจลลิน และเมมเบรนโปรตีน (63-11, 78-1) ผลจากหลอดทดลองพบว่า 78-1 ทนกรดน้ำย่อย 63-11 ทนกรดน้ำดี อย่างมีนัยสำคัญ 6-2 จับกับคลอโรฟอร์มได้ดี 63-11 และ 78-1 จับกับไซลีนได้ดี 6-2 และ 78-1 ไวปานกลาง และ 63-1 ไวสูง ต่อแอมปริซิลิน คลอแรมฟินิคอบ และเตตราไซคลีน 63-11 ต้านต้านจุลินทรีย์ก่อโรคปานกลาง ส่วน 6-2 และ 78-1 ต้านสูง ทุกไอโซเลทแสดงผลย่อยเม็ดเลือดแดงเป็นลบ ยกเว้น 78-1 ผลการวิเคราะห์ของทั้งสองวิธีในการศึกษานี้ทำให้ได้ข้อมูลเพียงพอที่จะใช้ในการบ่งชี้ศักยภาพ และข้อดีในการประยุกต์ใช้ *Bacillus* ในทางการค้า โดยเฉพาะ *B. infantis* (63-11) ที่แสดงสมบัติของการเป็นโพรไบติกทีมีศักยภาพมากที่สุด

สาขาวิชา	วิทยาศาสตร์และเทคโนโลยีทางอาหาร	ลายมือชื่อนิสิต .....
ปีการศึกษา	2562	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

6172601823 : MAJOR FOOD SCIENCE AND TECHNOLOGY

KEYWORD: Bacillus, Whole-genome, Probiotics, Food, Antimicrobial activity

Gauri Khullar : CHARACTERIZATION OF BACILLUS ISOLATES USING WHOLE GENOME SEQUENCING ANALYSIS AND APPLICATION AS A POTENTIAL FOOD PROBIOTIC. Advisor: CHEUNJIT PRAKITCHIWATTANA Co-advisor: PINIDPHON PROMBUTARA

Whole-genome sequencing (WGS) performed to evaluate potential probiotic properties of Bacillus species (6-2, 63-11 & 78-1) pre-reflecting antimicrobial properties, identified them as *Bacillus velezensis* (98.16%), *Bacillus infantis* (91.21%) and *Bacillus amyloliquefaciens* (99.06%) respectively. It also predicted K-mer resistance to cfr(B) and tet(L) proteins (6-2 & 78-1); bacteriocin and metabolite synthesis (6-2 & 78-1), terpenoid gene (63-11); hemolysin III (6-2 & 78-1) and hlyIII homolog (63-11); extracellular protease (6-2 & 63-11) and cell-bound protease (78-1) genes. WGS gut adaption F1F0 ATP, chaperonin (groEL, groES) and general stress response proteins (DnaK); EAL domain protein (biofilm), flagellin synthesis, and putative integral membrane proteins (63-11 & 78-1) were also annotated. The in-vitro assessment demonstrated a significant effect on growth and performance under gastric (78-1) and bile acid conditions (63-11) along with high hydrophobicity to chloroform (6-2) and xylene (63-11 & 78-1). It also reflected moderate (6-2 & 78-1) to high (63-11) susceptibility towards Ampicillin, Chloramphenicol and Tetracycline, and moderate (63-11) to high (6-2 & 78-1) antagonistic effects towards pathogens, with no hemolytic activity (except 78-1). Based on the findings of both the tests, this study provides enough evidence to support the potential benefits and commercial applications of *Bacillus* species, with *B. infantis* (63-11) being the most potential probiotic candidate.

Field of Study: Food Science and Technology

Academic Year: 2019

Student's Signature .....

Advisor's Signature .....

Co-advisor's Signature .....

ACKNOWLEDGEMENTS

I would like to express my immense gratitude to my advisor Dr. Cheunjit Prakitchaiwattana Associate Professor at Chulalongkorn University for her immense knowledge, support, encouragement, timely guidance, and patience during my Master research and thesis writing and my co-advisor Dr. Pinidphon Prombutara, for his support and valuable counseling throughout my WGS training and thesis writing.

I would like to express my appreciation to all the Master thesis committee members: Assoc. Dr. Sumate Tantratian, Dr. Sarn Settachaimongkon, and Asst. Prof. Dr. Pawinee Deetae for their educative and constructive suggestions for improving my Master thesis.

My research and Master thesis would also not have been possible without the Research Program for Development of Small and Medium Enterprise Researchers (2019) from the National Research Council of Thailand, the Thailand Research Fund, and the Thailand Research Organizations Network (TRON).

I am extremely grateful to the Faculty of Science, Chulalongkorn University for accepting my application for the master's degree program in Food Science and Technology. Many thanks to Dr. Rachatida Det-udom, PostDoc at Faculty of Science, Chulalongkorn University, and other lab mates who gave me memorable experiences and unlimited help during my studies.

Lastly, I would like to express my respect and regards to my parents for always encouraging me to pursue my dream of building a career in R&D and strive for excellence in life.

Gauri Khullar

TABLE OF CONTENTS

Page

EXTENDED ABSTRACT (ENGLISH)

The objective of this study was to characterize and evaluate potential probiotic properties of three bacterial species, pre-reflecting biocontrol agent properties, using whole genome sequencing (WGS) and in-vitro analysis. The three strains (6-2, 63-11 and 78-1) were identified as *Bacillus velezensis* (98.16%), *Bacillus infantis* (91.21%) and *Bacillus amyloliquefaciens* (99.06%) respectively, based on ANI & DDH values and phylogeny. WGS based K-mer resistance analysis predicted cfr(B) and tet(L) resistance proteins in 6-2 and 78-1, but no AMR genes in 63-11 and secondary metabolites analysis demonstrated Polyketide, NRPS and dipeptides genes (bacteriocin and antifungal agents) in 6-2 and 78-1 and non-mevalonate terpenoid mva (pigment) gene in 63-11. In-vitro assessment demonstrated moderate (6-2 & 78-1) to high (63-11) susceptibility towards Ampicillin, Chloramphenicol and Tetracycline and moderate (63-11) to high (6-2 & 78-1) antagonistic effects towards both Gram positive and Gram-negative pathogens. Toxicity analysis detected hemolysin III in (6-2 & 78-1) and hlyIII homolog in (63-11) and genomic analysis revealed extracellular protease gene (6-2 & 63-11) and cell bound protease gene (78-1). In-vitro screening established no hemolytic activity (except 78-1), with lowest protease activity in 78-1 and highest in 63-11. In-vitro survival rate of selected species in GIT analyzed by gastric and bile acid resistance showed moderate (6-2 & 63-11) to high (78-1) tolerance in acidic pH and moderate (6-2 & 78-1) to high (63-11) tolerance to bile salt. Genomic analysis also identified F1F0 ATP Acid tolerance proteins and chaperonin (groEL, groES), Cold shock (CspB) and general stress response protein (DnaK) that provide protection from cellular degradation. Adherence to epithelial cells analyzed by in-vitro hydrophobicity analysis reflected relatively high affinity of 6-2 to chloroform and high affinity of 63-11 and 78-1 strains to xylene and ethyl acetate. WGS annotated results also revealed EAL domain protein synthesis (key components of biofilm formation processes), flagellin synthesis and putative integral membrane proteins (Lipid transport) synthesis in 63-11 and 78-1 but none in strain 6-2. Based on the comparative findings of both the tests, this study provides enough evidence to support the potential benefits and commercial applications of *Bacillus* species, mainly *Bacillus infantis* (63-11) as most potential probiotic candidate.

TABLE OF CONTENT TABLES & FIGURES

CHAPTER 1 - INTRODUCTION

1.1 PROBIOTICS AND HEALTHY LIFESTYLE

Used for the primary purpose of ‘‘keeping a healthy gut’’ [2], probiotic microorganisms are known to confer clinically proven health benefit to its host when temporarily grow in the gastrointestinal tract under a specific environment; Often by inhibiting the growth of pathogenic organisms while maintain a symbiotic relationship with the natural gut microbiota habituating in the same environment [3] and improving the intestinal microbial balance [4]. FAO/WHO defines them as “live microorganisms which, when administered in adequate amounts confer a health benefit on the host” [5, 6].

The most important and documented beneficial effects include regulation of lactose digestion, lipid and oxalate metabolism; control of chronic intestinal inflammatory, protection against allergic diseases, and reduction of risk factors of respiratory tract infections, *Helicobacter pylori*, UTI, osteoporosis and many others [7]. Human intervention studies have also shown that probiotics enhance innate immunity including natural killer cell activity, phagocytic activity, and respiratory burst [8]. However, there is insufficient data to fully elucidate the mechanism of action, interaction and adaptation of these bacteria with the human gut and its native microflora [9-11].

A total transit time of approximately 3–8 h is required for the probiotic product to overcome the many barriers [12]. During this transit, the probiotic microorganism should be able to resist gastric acid and bile salts present in the upper gastrointestinal track and be able to adhere and colonize onto the lower gastrointestinal mucosal walls. Strains that do not pass through and do not survive different physicochemical, enzymatic, and microbial stresses of the upper sections of the GIT are not able to settle in the lower sections and perform their health functions.[13]. Hence, FAO guideline clearly state that probiotics should be able to survive the passage through the upper intestinal tract and arrive at their site of action, irrespective of how they are delivered [5]. Thus, it is essential to evaluate all probiotic products for the following properties - general aspects (origin, identity, resistance to mutation), technical aspects

(growth and survival under in vitro and during processing, and viability during transport and storage), physiological traits (performance under low pH (2.5), gastric juice, bile acid, pancreatic juice; adhesion potential to intestinal epithelium; resistance against environmental stress and antimicrobial factors prevailing in the upper GIT), functional properties (adhesion, colonization and proliferation potential on the mucosa epithelial cells; competitiveness towards native microflora; antimicrobial antagonism against gut pathogens while maintaining growth of native microflora; stimulation of immune response; selective stimulation of beneficial autochthonous bacteria) and safety evaluation (lack of potential invasive, transferable and virulence genes; no resistance against therapeutic antibiotics) [5].

Commercial probiotic preparations are generally comprised of mixtures of *Lactobacillus* and *Bifidobacterium* species, although Gram positive *Bacillus* spp. (Bacteria) have also been utilized [14] as probiotics, some for over 50 years [2]. Ubiquitous in nature but found in higher concentrations in soil, water and some traditional food products; *Bacillus* species are known to have a plant origin [15] and are commonly found in some fermented foods derived from animal [16]. They have a long history of extensive use in fermented foods largely in the African and Asian regions, but have recently become more prominent in global probiotic products. Their ability to produce a wide range of active substances derived from secondary metabolism that provides protective action against toxigenic microorganisms [17], and their ability to form endospores [18] has allowed them to have several functional benefits over the conventional probiotic products. Most have shown to survive under stress condition of food processing, and better suited to survival in a variety of food products compared to the more typical probiotic species. In addition to the spore-coat that provides protection from UV radiation, heat, solvents, hydrogen peroxide and enzymes, and allows them in their spore form to be stored indefinitely on the shelf, [19] *Bacillus* species have hydrophobic cell surface properties [20] and can survive low pH of the gastric barrier [21], which gives them the ability to adhere to and resist the

flux of the intestinal content; and compete with gut pathogen to colonize in the GI tract [22]. Having the additional advantage of a longer and stable shelf-life, bacterial spores have shown more success in colonizing, with a larger number of viable cells in the gastrointestinal tract [2, 23] as compared to other genus.

Used mainly as animal feed supplementation to improve the animal welfare and their product quality [24], *Bacillus* genus has proven to be an excellent source of biocontrol agents. However, the lack of extensively study in their use in the Food Industry has created a gap and inspires a novel approach to their potential applications. According to properties described above, *Bacillus* species have shown to be a potential candidate for their use as probiotics.

Multiple studies on *Bacillus* species and their probiotic potential in aquaculture and medicine have been conducted using both in-vitro and in-vivo models. These studies have given an in-depth knowledge on the functional properties and the species interaction with the animal and human models however they have been inconclusive when determining their possible mechanism of action with various pathways in the human body and their reliability to perform under stress conditions. Molecular analysis using Whole genome sequencing has brough a novel approach to understanding the possible phenomenon. This approach has been increasingly gaining appreciation and the combines results from both genomic and in-vitro & in-vivo studies has given a more concrete evaluation of *Bacillus* species and their probiotic properties.

1.2 OBJECTIVE

This study is focused on characterizing the three novel strains of *Bacillus* species isolated from traditional Thai salted fermented fish, by conducting *in-vitro* analysis and partial safety assessment tests as defined by the FAO guidelines, in addition to whole-genome sequencing. The aim of this study is to identify the species, demonstrate the species efficacy under stressful conditions, evaluate the level of pathogenicity and its potential applications as probiotic food supplement. The purpose of *in-vitro* analysis

is to both quantify and qualify necessary functional properties the bacterial isolates possess. Genomic evaluation will be conducted to reconfirm the data collected from *in-vitro* analysis and to identify specific genes responsible for horizontal/vertical gene transfer, acquired antimicrobial resistance and pathogenicity and virulence factors.

CHAPTER 2 - LITERATURE REVIEW

2.1 BACTERIAL PROBIOTICS

Some probiotic bacteria such as *Lactobacillus* and *Bifidobacterium* are considered members of the normal colonic microflora and hence, are not viewed as being overly pathogenic; Widely utilized in yogurts and other dairy products, they can retain viability during storage and possess the ability to survive passage through the gastrointestinal track [25]. However, since these probiotics do not permanently colonize the host, they need to be ingested or applied regularly for any health promoting properties to persist. Spore-forming bacteria of the *Bacillus* genus on the other hand, are not considered resident members of the gastrointestinal microflora, however, exhibit numerous benefits over the traditionally used *Lactobacillus* and *Bifidobacterium* bacteria [26-30].

The spore-coat compromising of bacterial endospores containing condensed and inactive chromosome at its core and peptidoglycan-rich cortex and proteinaceous material on the outside [31] protects the spores of the *Bacillus* spp. from UV radiation, heat, solvents, hydrogen peroxide and enzymes such as lysozyme [32], enabling them to be stored indefinitely on the shelf in a desiccated form without any deleterious effect on their viability. They also assist with surviving the low pH of the gastric barrier [26, 30] and if administrated in a specified amount, the entire dose of ingested bacteria has the potential to reach the small intestine intact. Furthermore, they can not only adhere but also colonize in the human gut eliminating the need for constant consumption [33], while simultaneously creating a microbial-based barrier and resisting the colonization of pathogens in the large intestines. In addition, *Bacillus* species produce a large number of antimicrobials/secondary metabolites that include bacteriocins and bacteriocin-like inhibitory substances (BLIS) as well as antibiotics and responsible for the bacteria's antimicrobial effects; inhibit the growth of pathogenic organisms [34, 35].

Bacterial spores produced in nature can survive under extreme environmental and can revert from germination stage to vegetative stage, if exposed to appropriate nutrients and growth conditions. Theoretically the upper intestinal region is considered rich in nutrients that can induce germination by allowing the water to enter the spore, breaking and removing the spore-coats, and resuming growth as vegetative cells; a reversible process that does not require de novo protein synthesis [36]. This is demonstrated in the schematic of the spore cycle of *Bacillus* species (Fig. 2.1). *Bacillus subtilis*, a facultative aerobe, for example, according to recent studies can survive under appropriate conditions and grow anaerobically if able to utilize nitrate or nitrite as an electron acceptor or by fermentation in the absence of electron acceptors [37]. The subspecies *Bacillus subtilis* var. natto has also been shown to germinate in the GIT of mice [38].

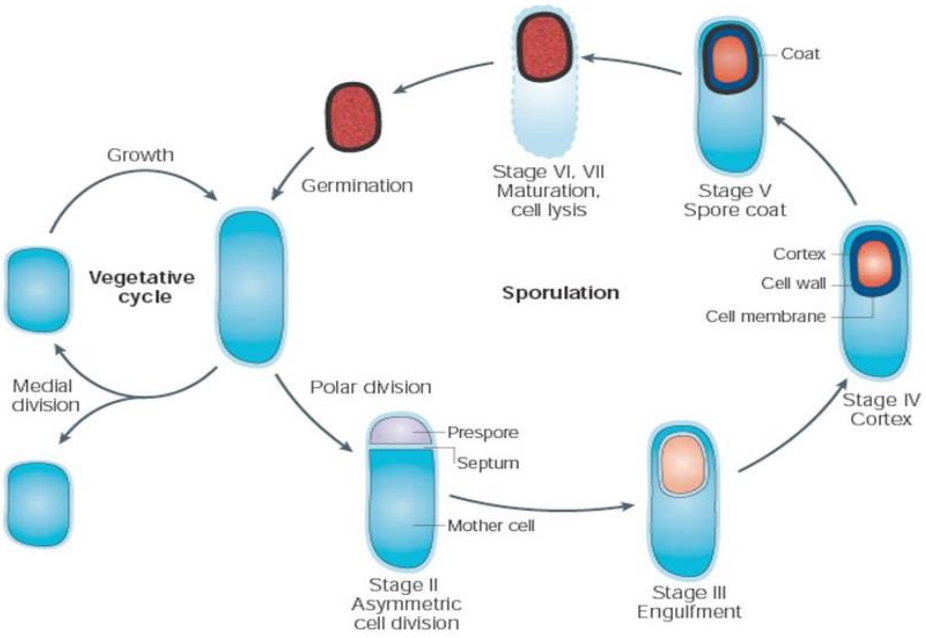

Figure 2.1 A schematic representing the sporulation life cycle of bacterial spore formers [1].

2.1.1 *Bacillus* probiotic for Human use

Used primarily in their spore form, *Bacillus* probiotic products for human use fall under two categories - those for prophylactic use and those sold as health food supplements

or novel foods. Although, many carry poorly define or invalid species [39] and due to the regulations most of them are categorized and commercially developed as food or dietary supplements (United States, Europe), natural health products (Canada), or as food for specific health uses (Japan) [40].

Prophylactic products are marketed for gastrointestinal disorders particularly childhood diarrhea or as an adjunct to antibiotic use. They are mostly available over the counter (OTC), and often recommended by a physician. Although, some countries such as the US do not permit their use, in Europe they are quite common with Italy being a major user since the 1950s. In SE Asia there is a history of extensive antibiotic usage and as a result it is common practice to use bacterial probiotics as an adjunct. One example of such product is Enterogermina®, which carries a mixture of four strains of antibiotic-resistant *Bacillus clausii,* an alkaliphilic species able to tolerate high pH 7–14 [2, 41]. The product claims to enhance the body's immune system by inducing IFN-c synthesis in murine spleen cells [42] following germination of the spores in the small intestine. Studies show all *Bacillus* clausii strains induce proliferation of CD4+ T cells in the presence of irradiated APC spleen cells [43]. Several studies in both humans and animal models have also provided strong evidence that oral administration of spores stimulates the immune system. Some studies have shown that orally administered *Bacillus subtilis* leads to a rapid induction of interferon production by mononuclear cells in the peripheral blood, which stimulated the activity of both macrophages and NK cells [44]. Additionally, reports supporting the use of *Bacillus* probiotics for therapeutic purpose have shown beneficial effects of the bacteria on urinary tract infections as well [45].

Health food supplements and novel foods claiming of enhancing the well-being of consumers by restoring the natural microflora to the gut and reducing risk of various diseases are openly sold over the internet. One example of such food is the Japanese product Natto: fermented soybeans with *Bacillus subtilis* (natto) or *B. subtilis* var. natto. Known for decades for its health benefits, the strain is thought to stimulate the

immune system [46], producing vitamin K2 and having anti-cancer properties.[47]. Furthermore, Nattokinase, a serine protease secreted from vegetative cells of all strains of *Bacillus subtilis,* with highest production in Natto has also shown to reduce blood clotting by fibrinolysis [48, 49]. The enzyme has GRAS status and is purified and sold as a health supplement worldwide [18].

2.1.2 *Bacillus* probiotic for Animal use

Bacillus probiotic products for animal use fall under the Scientific Committee on Animal Nutrition [50], according to which a complete ban on antibiotics for the use of animal husbandry has been issued by 2006 [51]. Implemented due to the concern over the spread of antibiotic resistance genes, failure to identify new antibiotics and inherent problems with developing new vaccines; The absence of antibiotic usage in animal feed good husbandry has led to the implementation of prebiotics, probiotics and synbiotics in animal feed for improved digestibility and immune health [2].

The use of *Bacillus* species in aquaculture [52] has expanded rapidly and become quite familiar to most researchers in the last two decades. Larval forms of most fish and shellfish are sensitive to gastrointestinal disorders because they are released into the environment at an early stage before their digestive tract and immune system has fully developed. Shrimps in particular have a non-specific immune response and vaccination that can only provide short-term protection against pathogens. Probiotic treatments on the other hand provide long-term protection and displayed antagonistic effect on pathogens. Bacterial supplements mainly referring to the *Bacillus* spp. have been used as probiotics and biocontrol agents in products such as Biostart® and Liqualife® [53, 54].

2.2 *BACILLUS* SPECIES IN FOOD PRODUCTS

The results from global research studies indicate a trend of moving the use of probiotic strains away from the pharmaceutical sectors and towards the functional health food sectors, as the probiotic bacteria supplied with food show more easy adaptation to the conditions in the GIT and delivery of enhanced beneficial health effects as

compared to pharmaceutical products [55]. Some probiotics have also been found to improve feed digestibility and reduce metabolic disorders [56]. This has stimulated the incorporation of probiotics into matrices based on milk, fruits, vegetables, cheeses, and meat products; Probiotic strains that can successfully be manufactured and incorporated into food products where they can retain their viability and function to create pleasant flavors, extended shelf life, as well as have a positive impact on human health.

The use of food matrices as carriers of probiotic bacteria enables regular ingestion of probiotics and assures that their beneficial effects are maintained. Probiotic strains selected for the production of food must meet the necessary safety, functionality, and technological criteria [57]. Since different probiotic species show different sensitivities towards the acidity, temperature and moisture conditions of the GI track, it is important to determine the metabolic activity and viability of bacteria to ensure whether a probiotic should be added to the food material.

Creating probiotic products require precise methodology to be followed, from characterizing the species at the strain level as a great majority of recognized effects are strain-dependent and many functional properties being restricted to the subspecies level; over to testing of functional capacities such as resistance to acid and bile, mucosal adherence, and adhesion stability, and finally to documenting viability throughout a set storage period monitored under strict conditions. Maintaining a standard minimum level of 10^6–10^7 CFU/mL or CFU/g viable probiotic cells at the expiry date is strongly recommended [58, 59].

Traditionally probiotics have been delivered through fermented dairy product such as a yogurt, kefir, cheese, buttermilk ice-cream, baby food, whey-based beverages, sour milk, [25, 60-62] by directly adding probiotic strains to intensively heat-treated milk. Although it is an uncontrolled fermentation process having its disadvantages including specific refrigeration requirements and short shelf life [63]. A secondary more controlled method for production of fermented foods is based on the use of starter

culture. This method has contributed to the microbial safety and offered technological, nutritional, sensorial, and health benefits [64].

There has been an extensive history of isolating *Bacillus* spp. naturally present on the surface of plant leaves and/or source raw material from Bacillus fermented foods (BFFs) found in the Asian and African region. These including natto (Japan), chungkookjang (Korea), kinema (Nepal) etc., geographically distributed across Asia and West Africa contain *Bacillus subtilis, Bacillus amyloliquefaciens, Bacillus licheniformis*, *Bacillus circulans, Bacillus pumilus*, and *Bacillus brevis* species. They have different names in these countries despite the many similarities in their manufacturing process and appearances [65-67].

Although many of the BFFs found in the Southeast Asian and Africa still need to be characterized, metabolomics analyses and experiments performed with animal models for some Asian fermented soybean foods such as natto and chungkookjang support their health claims [68].

2.2.1 Manufacturing of Probiotic food products

Fermentation is globally recognized as one of the most popular and oldest preservation methods in food technology and has played a wide role in the development of novel functional foods with a wide diversity of discriminants of sensory quality and enriched nutritional qualities, the production of fermented foods due to their strict refrigeration requirements and short shelf life have taken a back seat. Even though both dairy and nondairy substrates are considered as good vehicles for delivering probiotic bacteria to the human GIT and are often used as the base for the development of probiotic foods [69], the development of nondairy non-fermented products such as chocolate, cereal-based products, oat-based products, fruit juices, using probiotic microorganisms has been steadily gaining popularity [25, 60-62, 70] over the traditional dairy products.

Along with the manufacturing procedures such as fermentation, encapsulation, and storage conditions like temperature, humidity, and pH etc., the composition and the

nature of a food matrix of a food substrate plays an equally important role in the growth and survival, as well as the sensorially acceptance of the probiotic product [71] during the GIT transit [59, 72]. For instance, cheese which has relatively high pH and fat content, a solid consistency, and a higher buffering capacity may provide a protective barrier to the probiotic bacteria to safely pass through the GIT [73].

2.3 GUIDELINE FOR SAFETY EVALUATION OF PROBIOTIC (FDA & EFSA)

To identify and characterize a microorganism down to the strain level and investigate its functional properties - resistance to acid and bile, mucosal adherence, antibiotic susceptibility; safety, efficacy and viability under a given storage period, and functional health claims, strict guidelines need to be followed. Each country has their own set of guidelines that branch out from a standardized guideline established by a globally recognized regulatory body. FDA and EFSA are examples of such.

In the United States, probiotic products mostly fall under the Food and Drug Administration (FDA) with the granted GRAS status (generally recognized as safe), that specify guidelines for the evaluation of probiotics in food to confirm their health claims and benefits. The "Joint FAO/WHO Expert Consultation on Evaluation of Health and Nutritional Properties of Probiotics in Food", guideline is a Global Standard for evaluation of Probiotics. It describes the procedure in four main steps: (1) Strain identification (by phenotypic and genotypic methods), (2) Functional characterization and safety assessment, (3) Health claims validated through human studies (DBPS) for control of chronic disorders, protection against allergic diseases and reducing risk of respiratory tract infections, cancer etc., and (4) Proper labeling of genus, species, strain designation, minimum viable numbers of bacteria at end of shelf-life, storage conditions as well as corporate contact details for consumer information [74]. Since nonspecific structure/functioning claims (nonspecific claims) do not require governmental approval in the United States, they are frequently used with probiotic products (2019) and even though Good Manufacturing Practice guidelines issued by

the FDA are required to be followed for manufacturing dietary supplements, the end product's quality and efficacy are not a concern [75].

In Europe, probiotic-containing foods and food supplements fall under the European Union (EU) regulation covered by the Food Products Directive and Regulation [76]. The European Food Safety Authority (EFSA) is the responsible agency for food supplements, and therefore for majority of probiotic products evaluation. It evaluates all the safe microbial cultures according to a QPS-list (Qualified Presumption of Safety), designed solely for the safety assessment of biological agents. To satisfy the QPS status, any new culture must: i) be identified at the strain and species level; ii) be absent of transferable antimicrobial resistances and iii) lack toxigenic activity. Furthermore, EFSA is responsible for the assessment of health claims made on foods (including food supplements and probiotics) [77].

Both these guidelines outline a number of parallel recommended in-vitro safety tests for probiotic assessment for gaining knowledge on strain's mechanism of action, but also highlight the importance of conducting In-vivo studies to determine whether the probiotic induces any undesirable physiological effects, as these tests are not sufficient for describing the strain as a probiotic, or fully adequate to predict the functionality of the probiotic microorganisms in human body. Both the guidelines also recommend that probiotic strains, even among a group of bacteria that is Generally Recognized as Safe (GRAS), be characterized and tested to assure safety.

Taking these regulatory guidelines into consideration, [78] published a decision tree that summarizes in a systematic stepwise order on conducting a thorough safety assessments of microbial cultures intended for human and animal consumption.

2.4 PROBIOTIC CANDIDATE STRAIN CHARATERISTICS

2.4.1 In vitro screening and safety evaluation

Irrespective of the probiotic product being a food, food supplement, prophylactic or drug, the probiotic microorganism must be present in enough predefined quantity up

till the end of the shelf-life, to effectively provide the host with suggestive health benefits. This quantity can reduce when the probiotic microorganism passes through the gastrointestinal tract and is unable to resist acid and bile salts or colonize in the gut. According to the FAO/WHO Guidelines for the Evaluation of Probiotics in Food [74], the currently used *in-vitro* tests are - Resistance to gastric acidity; Bile acid resistance; Adherence to mucus and/or human epithelial cells and cell lines; Antimicrobial activity against potentially pathogenic bacteria; Ability to reduce pathogen adhesion to surfaces and Bile salt hydrolase activity. These tests are not adequate to fully comprehend the functionality of the probiotic microorganisms in human body. They can however, give useful insight to the possible mechanism of the microorganism's probiotic effect.

To ensure the safety and efficacy of the probiotic food product, it is also essential that the bacterial probiotic not carry any transmissible antibiotic resistance genes [79] or pathogenic or toxin genes. Antimicrobial activity is also an important criterion for the selection of probiotic microorganism. Antimicrobial activity targets the enteric undesirables and pathogen microbes [80] that may release bacteriocins and/or bacteriocin-like substances that have deleterious effects on the host immune system as well as stability of food supplement. The probiotic product should be safe for consumption and most importantly be contamination-free. Lastly, the viability of the probiotic, another important criterion as per the FDA regulations can also be checked through in-vitro screening by simply growing the viable microbes on various nutrient agar mediums [81]. While in-vitro tests have their benefits when it comes to preliminary screening, they also have some flaws. It is possible to get false negative results or inconclusive results. The artificial gastric juice and bile salts need to be freshly prepared before the experiment needs to be conducted. Each stage must be re-checked, and the bacterial plates should not be older than one month to ensure viable cell growth. Antibiotic tests of the bacterial isolates should be screened for a range of standard antibiotics to ensure there is no horizontal/vertical gene transfer or acquired antibiotic resistance. It is not possible to completely rely on the in-vitro tests; therefore, it is

important to run similar tests under in-vivo model to determine the reliability and accuracy of the experiments. Moreover, extensive genomic study of all isolates should be conducted to understand and quantify the expression of specific genes that may contribute to the microbe's probiotic properties [74].

2.4.2 Whole Genome Sequencing and Bioinformatic Tools

Although there is enough evidence to claim the functional benefits of the bacterial isolate, it is equally important to comprehend its mechanism of action, molecular evolution and the basis for its potential probiotic and health-promoting activities. Capillary sequencing technique was the first approach to successfully sequence a nearly full human genome; however, it is too expensive and time consuming for commercial purposes. It has been progressively displaced by next-generation sequencing (NGS), also known as Whole-genome sequencing [82]; A comprehensive method for rapid analysis of entire genomes of any organism at a single time, having the ability to generate accurate reference genomes, microbial identification, and other comparative bioinformatic studies. This novel approach has proven to be a more accurate genotypic method for bacterial identification and is adequate for effectively discriminating between the many *Bacillus* species, as compared to Capillary sequencing or 16S ribosomal RNA sequencing. It determines the order of all the nucleotides in an individual's DNA and can determine variations in any part of the genome [83]. It also enables fast and accurate *De novo* sequencing, useful for characterization of functional properties for any novel species. This is done by comparing the annotated sequence against the NCBI Database and by calculating the ANI value (Average nucleotide identity) and/or the DNA-DNA hybridization. Other programs such as resistance gene finder, Antimicrobial resistance finder, Pathogen finder and Virulence finder can be run for further analysis such as hemolytic activity and pathogenicity of the bacterial isolate. Any species can also be screened for genes responsible for epithelial adhesion, gastric and bile salt tolerance, bacteriocin production and general cell membrane composition.

However, WGS has its flaws too. The biggest problem comes down to the reference genome. NCBI has a huge database and WGS gives a plethora of information. This information is screened against the reference genome of choice and can vary from one genome to another. While de novo sequencing is possible for when a reference is unavailable, it can lead to more errors since we have nothing concrete to compare to. Moreover, the annotation results of any strain only identify the presence or absence of specific genes. For example, we can determine if BSH gene required for bile salt tolerance is present in our species or not. However, to quantify the level of expression, each gene needs to be specifically screened against a known reference. There is no doubt that WGS method achieves high resolution data and can identify small discrepancies for the sequence of interest, however, it is extremely time consuming and an expensive method.

Over the last decade, numerous studies have been conducted to investigate and characterize probiotic properties of *Bacillus* candidate strains, as summarized in Table 2.4. In these researches, in vitro analysis has been reported as the main methods used along with PCR-based 16S sequencing for genotypic based characterization of specific gene for evaluating probiotic properties and safety inspection of *Bacillus* species. WSG technology, even though has the added advantage of identifying functional genes involved in key/general probiotic properties of candidate strains and for in-depth safety evaluation of candidate strains, has not been wildly used in probiotic characterization and is still a novel approach. This technique, however, has greatly accelerated the possible understanding of the diversity of the potential probiotic bacteria, but provided insights into the interaction and adaptation into the human gut and its native microflora.

Table 2.4 Some of the recent reports on various probiotic characterization techniques for Bacillus isolates

Origin	Strain	Protocol	Significant point of works	Reference
India	*Bacillus velezensis* strain DU14	16S rDNA gene sequencing; Std. In-vitro analysis; FTIR analysis	Significant tolerance 2–9 pH range and 1% (w/v) of bile salt; sensitivity against both broad and narrow spectrum antibiotics; antagonistic against both Gram + and – (*Bacillus cereus* ATCC-11778) & (*Escherichia coli* ATCC-25922) pathogens. Non hemolytic; CFSC showed no cytotoxicity against mouse liver cells. FTIR analysis proved exopolysaccharide synthesis.	[82]
USA	*Bacillus coagulans* GanedenBC 30TM	In vitro bacterial reverse mutation assay; in vitro chromosomal aberration assay; micronucleus assay in mice; acute and 90-day sub chronic repeated oral toxicity studies in rats, acute eye and skin irritation studies in rabbits.	Strain does not demonstrate mutagenic, clastogenic, or genotoxic effects. Results of the acute and 90-day sub chronic oral toxicity studies in rats resulted NOAEL greater than 1000 mg/kg per day, giving a safety factor ranging from 3173 to 95,200 times. Hence GanedenBC30TM considered safe for chronic human consumption.	[83]
China	*Bacillus velezensis* K2	RNA isolation and real-time quantitative PCR; cluster analysis gyrB sequence. In-vitro assay. ACP, AKP and C3 activity analysis, Challenge test	Antimicrobial spectrum against fish pathogens - *Vibrio harveyi, Vibrio alginolyticus, Aeromonas hydrophila, Aeromonas veronii, Aeromonas caviae, Enterococcus casseliflavus* and *Lactococcus garvieae*; intraperitoneal injection of K2 in healthy grouper cause no pathological abnormality or death; increased serum acid phosphatase (ACP) activity ($P < 0.05$). Up-regulation of innate cellular and humoral immune responses (lysozyme gene, piscidin, IgM and MyD88), enhancing the resistance to V. harveyi.	[84]
Japan	*Bacillus subtilis* strains (BFFs)	Genomic sequencing & Metabolomic profile analysis, Experiments with animal models	Genetic variations commonly found are required for soybean fermentation. Metabolomics analyses and experiments with animal models support health claims of BFFs.	[68]

China	*Bacillus velezensis* JW	RNA isolation and real-time quantitative PCR. Antimicrobial assay - ACP, AKP and GSH-PX activity analysis, Challenge test	Showed antimicrobial activity against a broad range of fish pathogenic bacteria -*Aeromonas hydrophila, Aeromonas salmonicida, Lactococcus garvieae, Streptococcus agalactiae*, and *Vibrio Parahemolyticus*; Increased acid phosphatase (ACP), alkaline phosphatase (AKP), and glutathione peroxidase (GSH-PX) activity; Four bacteriocins, three Polyketide Synthetase (PKS), and five Nonribosomal Peptide-Synthetase (NRPS) gene clusters identified;	[85]
Korea	*Bacillus subtilis* P229	Detection of enterotoxin genes by PCR and electrophoresis and std. in-vitro analysis	Vegetative cells did not produce beta glucuronidase, were sufficiently susceptible to antibiotics, and adhered strongly to human intestinal epithelial cells. The strain did not induce hemolysis or carry enterotoxin genes and showed high levels of autoaggregation, and coaggregation with pathogens depended on the species involved. It did not produce hazardous biogenic amines from histidine and ornithine.	[86]
India	*Bacillus subtilis* (KX756706), *Bacillus cereus* (KX756707), *Bacillus amyloliquefaciens* (KX775224)	Bio-safety assay in in-vivo model & std. In-vitro tests. DNA extraction, PCR amplification using universal primer. colony morphological, biochemical characterization and 16S rRNA sequencing	Not pathogenic to the host fish. able to survive in acidic and alkaline conditions, higher tolerance to bile salt, high surface hydrophobicity to solvents, and were found to tolerate in gastric juice. All three isolates exhibited notable amylase, proteolytic, lipase activity and susceptibility to various antibiotics. Only *Bacillus amyloliquefaciens* (KX775224) exhibited a good antagonistic activity against three fish pathogens Viz: *Aeromonas hydrophila, Acinetobacter* sp. and *Acinetobacter tandoii* & positive results for biofilm formation assay. None exhibit any hemolytic activity	[87]
India	*Bacillus clausii* UBBC07	Acute and subacute studies in rats, whole genome sequencing, Antibiotic resistance analysis	No Observed Adverse Effect Level (NOAEL) found to be 1000 (126 billion cfu) mg/kg body weight/day by oral route. no lethality or toxic clinical symptoms in the experimental rats. Does not produce lecithinase and it is non-hemolytic.	[88]

			Absence of toxin genes and transferable antibiotic resistance genes	
France	*Bacillus subtilis* CU1	16S rDNA and gyrB nucleotide analyses. RAPD PCR and PFGE analyses 16-week randomized, double-blind, placebo-controlled, parallel-arm study	No antibiotic resistance greater than existing regulatory cutoffs against clinically important antibiotics, no induce hemolysis or production of surfactant factors, absence of toxigenic activity in vitro. Safe and well-tolerated in the clinical subjects without undesirable physiological effects on markers of liver and kidney function, complete blood counts, hemodynamic parameters, and vital signs	[89]
India	*Bacillus amyloliquefaciens* AMS1	16S rRNA gene sequence and in-vitro analysis. Scanning electron microscopy studies.	Showed potential probiotic characteristics as well as a significant cellulolytic activity in vitro. Survived harsh physio-chemical conditions present in gastrointestinal tract. Degraded CMC, maize straw and filter paper within 96 hours of incubation, conferring cellulolytic potential	[90]
India	*Bacillus infantis* KADR2	In vitro intestinal condition based on resistance to bile tolerance, low pH, hydrophobicity, catalase activity and antibiotics susceptibilities. Partial 16S rRNA gene sequencing	Showed resistance to acid, bile salt, gastric juice condition, as well as a good capacity for adherence to hydrocarbon, to pathogens, and higher antagonistic effect against fish pathogens. Partial 16S rRNA gene sequencing showed 99% homology with *Bacillus infantis* supported by morphological and physiological characterization.	[91]
Turkey	*Bacillus indicus* HU36	Viable bacteria count Color analysis Sensory profiling	Survival rate between 88 and 91% in maltodextrin and lemon fiber. Bacteria and dietary fiber addition did not show any negative effects on product sensory and color properties;	[92]

CHAPTER 3 - MATERIALS AND METHODS

3.1 MATERIAL

Table 3.1a Chemicals and reagents used in this study

No.	Chemicals and reagents	Sources
1.	Nutrient Broth (Powder)	Himedia, India
2.	Molecular NaCl	Univar, United State
3.	Glycerol/Glycine	Kemaus, Australia
4.	Nutrient Agar Powder	Himedia, India
5.	Nucleic Acid Extraction Kits (Bacteria)	Vivantis, Malaysia
6.	Pepsin	Sigma-Aldrich, United State
7.	0.1 M HCl	Fisher Chemical, UK
8.	Tryptic Soy Broth (TSB)	Himedia, India
9.	Bile salt powder	Sigma-Aldrich, United State
10.	Ampicillin powder	Vivantis, Malaysia
11.	Tetracycline disc	BD, United State
12.	Chloramphenicol disc	BD, United State
13.	Molecular water	BD, United State
14.	Xylene	Fisher Chemical, UK
15.	Chloroform	QRec, New Zealand
16.	Ethyl Acetate	QRec, New Zealand
17.	Potassium Chloride (KCl)	Univar, New Zealand, Australia
18.	Disodium phosphate (Na2HPO4)	Carlo Erba, Italy
19.	Potassium dihydrogenphosphate (KH_2PO_4)	Univar, United State
20.	DNA Tag polymerase	Vivantis, Malaysia
21.	ethidium bromide	Applichem, Spain
22.	azocasein	Sigma-Aldrich, United State
23.	Tris-HCl	Sigma-Aldrich, United State
24.	Trichloroacetic acid (TCA)	Sigma-Aldrich, United State

Table 3.1b Instruments used in this study

No.	Instruments	Sources
1.	Incubator	Blinder FED 400, Germany
2.	Biosafety Cabinet (BSL-2)	Telstar BioUltra 4, Japan
3.	pH Indicator	Mettler Toledo, United State

4.	Optical Microscope	Olympus BX51, Japan
5.	Digital Camera	Canon, Japan
6.	Centrifuge	Hettich MIKRO 22R, Germany
7.	ELISA plate reader/ Spectrophotometer	Biochrom UVM 340, United Kingdom
8.	Fume Hood	Extractor, Fume Hood, Thailand
9.	Microwave	LG Electrnics, Thailand
10.	Autoclave (High Pressure Steam Sterlizer	Meditop Tomy SX-700, Japan
11.	Hot Air Oven	Heraeus, Germany
12.	Weighing Balance	Mettler Toledo, United State
13.	Hot water Bath	Scientific Promotion Co. Ltd., Thailand
14.	Shaker	Wisd cube, Germany
15.	Refrigerator	Freezer SM-H112, Thailand
16.	DNA thermal cycler	BioRad T100TM, Singapore
17.	Electrophoresis gel chamber	Electrophoresis gel chamber; HU413L, United Kingdom
18.	Electrophoresis power supply	Amersham pharmacia, Bitech, Sweden
19.	UV transluminatior	Vilber Lourmat, France

3.2 METHODOLOGY

3.2.1 Bacterial strain source, isolation, and growth conditions

A total of 124 halophile bacterial species were isolated from a traditional Thai salted fermented fishes (Pla-ra) through different processes and fermentation periods by Prakitchaiwattana and research group (2017) in several parts of Thailand. The isolates were stored at -80°C in nutrient broth supplemented with 50% glycerol. Preliminary identification of isolates was done by DNA sequencing. DNA of isolates was extracted following the method described by [95]. Conserved regions of the 16S rRNA region of isolates was chosen and PCR amplification performed using primer set 338F/519R [96], in 50 µL reaction mixtures compromising of 2 µL DNA (10-50 ng/µL), 0.1 mM of each primer, 1.5mM MgCl2, 0.1mM dNTPs mix and 2 µL of DNA Tag polymerase. Samples were subjected to an initial cycle of denaturation (94°C for 2 min), followed by 35 cycles of denaturation (94 °C for 30 s), annealing (55°C for 30 s) and elongation (72°C for 30 s), ending with extension at 72°C for 7 min, in the DNA thermal cycler. Ten µL

of PCR products were directly applied onto 1.5% (w/v) agarose gel in 1% TAE buffer containing 2M Tris base, 1M glacial acetic acid, 0.5 M EDTA, pH 8.0 to 1000mL distilled water and electrophoresis was performed at a constant voltage of 200V for 30 minutes. On completion the gel was the stained with 1% ethidium bromide and observed under the UV transilluminator. The amplified data was then sent to commercial sequencing facility (Macrogen, Korea) after cleaning and analyzed using nucleotide BLAST program of NCBI Database. Out of the 124 strains, three novel strains - *Bacillus velezensis* (6-2), *Bacillus infantis* (63-11), *Bacillus amyloliquefaciens* (78-1) were selectively used in this study, by cultivating them on nutrient agar plates at 37°C for 24 h before use.

3.2.2 Whole Genome Sequencing and Annotation of Bacterial isolate

Bacterial DNA of strain 63-11 was extracted using the Bacterial DNA extraction kit (Vivantis). Library preparation and sequencing of the isolates was done at Omics Science and Bioinformatics Center, Chulalongkorn University, using Nextera XT DNA prep kit and Illumina Miseq sequencer. Raw reads quality was checked using FASTQC software. Adaptors and poor-quality reads were removed using Trim Galore, and the filtered reads were used as an input for Unicycler, genome assembly program. Annotation of assembled genome was done using Prokka Version 1.13 [96, 97].

3.2.2.1 Species identification and genome comparison

ANI (Average nucleotide identity) value detects the level of similarity of the total genomic sequences between two or more strains based on the identification of homologous fragments of fixed length using the BLAST algorithm. Similarly, DDH values determine the relatedness between strains and is considered an important criterion in the delineation of bacterial species. The quantitative relationship between DDH value and ANI value can give precise information of the identification of a novel genomic sequence. Strains with ANI > 95% and DDH > 70% value are considered as belonging to the same species [98].

The ANI and DDH value between the three *Bacillus* isolated strains (6-2, 63-11, 78-1) and reference genomes - *Bacillus amyloliquefaciens* DSM 7 = ATCC 23350 (Accession NC_014551), *Bacillus amyloliquefaciens* EGD-AQ14 (Accession NZ_AVQH01000059.1), *Bacillus amyloliquefaciens* strain Y2 (Accession CP003332), Bacillus anthracis CZC5 (Accession AP018443), *Bacillus atrophaeus* strain BA59 (Accession CP024051.1), Bacillus cereus ATCC 14579 (Accession NZ_CP034551), *Bacillus firmus* DS1 (Accession APVL01000002.1), *Bacillus infantis* NRRL B-14911 (Accession NC_022524.1), *Bacillus licheniformis* DSM 13 = ATCC 14580 (Accession NC_006270), *Bacillus pumilus* strain 145 (Accession CP027116.1), *Bacillus subtilis* subsp. subtilis (Accession NC_000964), *Bacillus thuringiensis* strain C15 (Accession CP021436.1), *Bacillus velezensis* strain NRRL B-4257 (Accession NZ_LLZB01000000), *Bacillus velezensis* YAU B9601-Y2 (Accession NC_017061.1), as taken from the NCBI database were calculated and compared using JspeciesWS web server tool [99] and Genome-Genome Distance Calculator (GGDC) 2.1 BLAST [100]. Additionally, circular map was constructed using CGView Server to characterize specific genomic regions or genes between the bacterial isolates and the closest related species [101].

The phylogenetic tree of the three isolates and reference genomes was constructed using Type (Strain) Genome Server (TYGS), to analyze both Whole Genome sequencing and 16s RNA sequencing based results, and understand the molecular evolution and diversity between the *Bacillus* species [102].

3.2.2.2 Screening for antibiotic and antibiotic resistance genes of *Bacillus* isolates

Resistance gene identifier (RGI) tool of Comprehensive Antimicrobial Resistance Database (CARD) [103], ResFinder tool of Center for Genomic Epidemiology [104] and, BLAST analysis [105, 106] of Erythromycin ribosomal methylase (erm), Aminoglycoside O-nucleotidyltransferase (aadD2) and Chloramphenicol acetyltransferase (cat) gene of *Bacillus clausii* against *Bacillus* isolates was conducted to screen for DNA-based intrinsic and acquired antibiotic/antimicrobial resistance genes that may have been

present in the three isolates. The threshold for %ID of Resfinder was set to 90% and Minimum length to 60%. The refence genomes for BLAST analysis - *Bacillus clausii* KSM-K16 DNA (GenBank: AP006627.1), *Bacillus clausii* NR aadD2 gene for aminoglycoside O-nucleotidyltransferase ANT(4')-Ib (NCBI Reference Sequence: NG_047392.1) and *Bacillus clausii* DSM8716 putative leader peptide and ribosomal methylase Erm34 genes (GenBank: AY234334.1) were taken from the NCBI Database. Lastly KmerResistance tool of CGE [104] was used to analyze AMR genes based on k-mer sequences of the three *Bacillus* isolates [107]. The identity threshold was set to 70% and dept correlation threshold to 10%.

3.2.2.3 Screening for toxic and virulent genes of *Bacillus* isolates

PathogenFinder tools of Center for Genomic Epidemiology [108] was used for surface screening of pathogen genes found in all bacteria and VFanalyzer (Virulence Factors analyzer) tool was used for in-depth screening of *Bacillus* isolates against known pathogenic *Bacillus* species (*B. anthracis* str. Ames(pXO1- pXO2-), *B. anthracis* str. Sterne(pXO1! pXO2-), *B. cereus* ATCC 10987, *B. cereus* ATCC 14579, *B. licheniformis* DSM 13 (ATCC 14580), *B. subtilis* subsp. subtilis str. 168) for presence of any toxin genes. The reference genomes for VFanalyser were preinstalled on the VFBD database [109].

Bacillus isolates were also compared against hemolysin (hly), cytotoxin K (cyt), diarrheal toxin (bce), hemolytic enterotoxin (hbl) and nonhemolytic enterotoxin (nhe) genes of *Bacillus cereus* species using MegaBLAST tool of NCBI database [110]. The reference virulent genes were taken from the NCBI database with the following accession numbers - *Bacillus cereus* strain 5.39 hemolysin BL (hblA) gene, partial cds (GenBank: KF681259.1), *Bacillus cereus* strain R1 HblB toxin gene, complete cds (GenBank: MK268740.1), *Bacillus cereus* strain EC303 hemolysin BL component L2 (hblC) gene, partial cds (GenBank: JQ039144.1), *Bacillus cereus* partial hblD gene for hemolysin BL binding component L1, strain BC (GenBank: AJ937194.1), *Bacillus cereus* nheA gene, strain MHI 1761 (GenBank: FN825684.1), *Bacillus cereus* partial nheB gene for Enterotoxin B, strain BK (GenBank: AJ937178.1), *Bacillus*

cereus nheC gene, strain MHI 1672 (GenBank: FN825685.1), *Bacillus cereus* bceT gene for diarrheal enterotoxin, complete cds (GenBank: D17312.1), *Bacillus cereus* strain TIAC959 CytK (cytK) gene, partial cds (GenBank: KP409163.1), *Bacillus cereus* E33L, complete genome (GenBank: CP000001.1)

In addition to the above tools, a fourth computation tool - BlastKOALA (KEGG Orthology And Links Annotation) was used to further analyze the virulence genes of the three isolates against both completely sequenced and some partial sequenced genomes present in the KEGG database [111].

3.2.2.4 Screening for secondary metabolite gene clusters

Secondary metabolite gene clusters commonly found in bacterial species are considered an important source of antagonistic compounds such as - antibiotics, anti-cancer agents, immunosuppressants etc. that help comprehend how microorganisms adapt to various ecological niches. The SeMe gene clusters of the three bacterial strains were identified using an online tools, NP.searcher [112] and antiSMASH (version 5.1.2) [113]; Set to briefly quantify the SeMe gene as well as detect well-defined clusters containing all required parts and partial clusters missing one or more functional parts, from the draft genomes of the three *Bacillus* isolates.

3.2.2.5 Biochemical reactions analysis

Organism-specific pathways occurring in the cells of *Bacillus* species were analyzed using KEGG (Kyoto Encyclopedia of Genes and Genomes) databases along with the enzymes secreted by the three *Bacillus* isolates using the KEGG ENZYME, an extension of KEGG LIGAND database.

Secondly, protein in *Bacillus* isolates were analyzed by Prokka program of Galaxy Version 1.14.5 and Interproscan functional predictions of program of ORFs of Galaxy Version 5.0.0. using Pfam database to identify genes responsible for survival and proliferation in human GIT.

3.2.3 In-vitro Screening of probiotics properties

3.2.3.1 Gastric juice tolerance

To assess gastric acid tolerance, protocol from [114] was followed by inoculationg single colony of each strain in 5 mL of NB broth at 37°C for 24 h. 1 mL of the overnight suspension of all three strains (6-2, 63-11 and 78-1) was inoculated into 9 mL of artificial gastric juice comprising of 0.3% w/v pepsin (Sigma-Aldrich), adjusted with 0.1 M HCl; pH 2.5, and incubated for 3 h at 150 rpm. Viable cells were counted by pouring on NA plates and incubating for 24 h. Tolerance rates of the isolate in the presence of gastric juice was represented in CFU ml $^{-1}$ and non- inoculated NB Broth was used as negative control.

3.2.3.2 Bile tolerance

Bile acid resistance of the isolated strains 6-2, 63-11 and 78-1 was measured by inoculating a single colony in 5 mL of NB broth at 37°C for 24 h, following the protocol from [114] with a slight modification. 1 mL of incubated suspension of each strain was inoculated with 9 mL of TSB containing 0.3% w/v bile salt (Sigma-Aldrich) for 3 h, 6 h and 24 h with shaking at 150 rpm. Viable cells were counted by pouring on NA plates and by measuring OD at 600nm. Non-inoculated NB Broth was used as negative control.

3.2.3.3 Hydrophobicity assay

Hydrophobicity of the three isolates was measured following the protocol from [92]as decribed by [115] with slight modifications. Three organic solvents - xylene (apolar), chloroform (polar acidic solvent) and ethyl acetate (polar basic solvent) were used. Overnight grown culture in NB was pelleted (6000 gX, 5 min) and the cell pellet was washed twice with phosphate buffer and re-suspended in phosphate buffer. The absorbance of suspended pellet of the isolate was measured at 600 nm. The cell suspension was then mixed with equal volume of organic solvents and vortexed for 2 min. The two phases were allowed to separate for 1 h and the absorbance of the aqueous phase was measured at 600 nm. The hydrophobicity of bacterial adhesion to the solvent was calculated using the formula:

$$\%\ Hydrophobicity = \frac{1-A1}{A0} * 100 \qquad (1)$$

Where A0 represents initial absorption before mixing with hydrocarbon sources and A1 represents final absorption after mixing with hydrocarbon sources.

3.2.3.4 Antibiotic susceptibility assay

Antibiotic susceptibility of the three isolates strains were identified by disc diffusion method from [88], as described by [116] with a few modifications. Three antibiotics were used: ampicillin (10 µg), tetracycline (30 µg) and chloramphenicol (30 µg). Overnight inoculated bacterial cultures were spread onto NA plates and antibiotic-impregnated paper dicks were placed on the plate and incubation at 37°C for 24 h. The standard antibiotic discs for 30 µg Chloramphenicol and 30 µg tetracycline were supplied by (Becton, Dickinson and Company, USA). The 10 µg Ampicillin discs were prepared from stock solution compromising of 10 mg of Ampicillin powder (Vivantis) in 1 ml Molecular water. Discs impregnated with molecular water were used as negative control. Zones of inhibition were measured after 24 h of incubation.

3.2.3.5 Antimicrobial activity

Antimicrobial activity against potentially pathogens was conducted using spot on lawn assay, by inoculating the NA plates with *Staphylococcus aureus* ATCC 25923, *Bacillus cereus* ATCC 6633 and *Escherichia coli* ATCC 25922, following protocol from [116] with slight modification. The three bacterial isolates were potted into the pre-inoculated plates and the zone of inhibition was measures. A minimum of 2mm zone of inhibition against potential pathogens is essential to justify for its antimicrobial activity.

3.2.3.6 Hemolytic activity

Hemolysis was tested by the protocol of described by [88] with some modifications. The three isolates were streaked on pre-prepared blood-agar plates (supplied by Department of medical Science, Ministry of Public Health) supplemented with 5% human blood and incubated for 24–48 h at 37°C to detect patterns of hemolysis.

3.2.3.7 Protease activity

Qualitative preliminary screening of proteolytic activity of cultures protocol was modified from [117] and observed in NA medium containing 1% casein with 5% (w/v) total salts. Clear zones around the colonies appearing over the next 48 h were taken as evidence of proteolytic activity. Quantitative screening of proteinase activity was determined using azocasein (Sigma-Aldrich) as substrate with modifications following the protocol modified from [118]. OD600 of overnight-NB isolate suspension was adjust to 0.5 and incubated at 37°C for 18 h. before cell debris removal was done using centrifugation at 12000 gX, 5 min at 4°C. Aliquots of 100 µl of cell-free supernatant were added to a mixture of 50 ul of Tris-HCl, pH 8.0 and 50 ul of 1% azocasein solution (w/v). 500 µL of 5% saline solution was used as negative control. After 15 min of incubation at 37°C, the reaction was stopped by adding 100 µL of 10% trichloroacetic acid and the samples were centrifuged at 10,000 gX for 10 min. Determination of enzyme activity was done by spectrophotometry at 405 nm.

3.2.4 Halophilic property assessment

Halophilic (Salt tolerance) properties of bacterial isolates was primarily screened in Nutrient Broth (NB) supplemented with 1,5,10,15 and 20% NaCl [119] using 96 well micro titre plate. The culture turbidity was then measured at 600nm, after incubation at 37°C for 24 hours.

3.2.5 Statistical analysis

All the experiments were performed in triplicates, and the results were subjected to one-way analysis of variance (ANOVA). The significance of the differences between treatments was compared by Duncan tests ($P<0.05$). Data were analyzed by SPSS for Windows version 22.0 (MD5: 490f47f1e1a20469e499c7a75aac4385). The P values of less than 0.05 were considered statistically significant.

CHAPTER 4 - RESULTS AND DISCUSSION

4.1 16S RNA AND WHOLE GENOME SEQUENCING

A total of 124 halophile isolates identified from the Plara project conducted by Cheunjit Prakitchaiwattana and research group, 2017, were preliminarily screened for their inhibitory activity against some pathogens. The strains that tested positive were then subjected to 16s RNA sequencing analysis and compared against known genomes accessible through the NCBI Database. A large percentage of these isolates were found to belong to the *Bacillus* genus, mainly - *Bacillus subtilis, Bacillus pumilus, Bacillus amyloliquefaciens, Bacillus licheniformis, Bacillus atrophaeus* and *Bacillus altitudies*. Preliminary screening as shown in Table 4.1a investigated by Chhetri, Prakitchaiwattana, & Settachaimongkon [120] established that Plara samples containing *Bacillus* spp. had low count or absence of *Staphylococcus* spp., indicating the antagonistic relationship between this two groups and potential bio-control agent properties of *Bacillus* isolates under halophilic environment. Based on these biosafety results, three strains (from 16s RNA analysis) including two strains of *B. subtilis* and one of *B. infantis*, were selected for evaluation of their potential probiotic properties in this study. These strains were subjected to Whole genome sequencing and In-vitro assessment for potential probiotic characterization.

The 16s RNA sequencing technique requires specific primers to get a vague idea of the taxonomy of a microbiome. WGS on the other hand does not require primers and can analyze the complete genome instead of just a small conserved region, giving accurately taxonomic identification of the microorganism down to the strain level, as well as evaluating its possible functional traits. Whole-genome sequencing technique has also proven to give a more accurate phylogenic identification of all samples belonging to a metagenome. This novel approach has proven to be adequate for effective discriminating between the many *Bacillus* species, AMR gene and virulence gene identification, and metabolic capacity analysis, as compared to 16S ribosomal RNA sequencing [122]. Results mentioned in Table 4.1a show species identity analyzed by 16s RNA and WGS were not similar. Two *Bacillus* strains identified with over 99%

homology to *B. subtilis* by 16S rRNA sequencing, were identified as *Bacillus velezensis* (6-2) and *Bacillus amyloliquefaciens* (78-1) by WGS. This proves WGS is a more accurate and comprehensive method for genotypic strain identification.

Table 4.1a Comparison of 16s RNA sequencing with Whole Genome Sequencing

Isolates	16S rRNA sequencing			Whole genome sequencing		
	Closest relatives (GenBank)	Identity	Accession number*	Closest relatives (GenBank)	Identity	Accession number*
6-2	*Bacillus subtilis*	99.37%	MH010140.1	*Bacillus velezensis*	97.59%	NZ_LLZB01000000
		98.16%	GU434362.1		98.16%	NC_017061.1
63-11	*Bacillus infantis*	98.14%	MN243631.1	*Bacillus infantis*	91.21%	NC_022524.1
		98.14%	MK281522.1			
78-1	*Bacillus subtilis*	100.00%	MH010139.1	*Bacillus amyloliquefaciens*	99.06%	NZ_AVQH01000059.1
		100.00%	GU434362.1		97.99%	CP003332

Contigs are a set of overlapping DNA segments that are important for assembling and mapping of a complete genome [123], and for taxonomical analysis along with individual gene identification. A set of contigs are known as scaffolds. The large number of contigs after assembly of short-read sequences cause major problems as genes are lost between the fragmented sequences on the contig boundaries. Hence, smaller number of contigs defines well assembles genome [124]. By extension longer contig lengths suggest less fragmentation and reduces the chance of gene loss.

N50 contig value is determined by sorting all contigs in a descending order of size, and then adding the contigs until the total added size equals at least half of the total size of all assembled contigs. The smallest contig size used in this addition process represents the N50 value. The larger the N50 value, the better is the assembly [125].

Coverage or the depth of a DNA sequencing refers to the number of unique reads (a specific nucleotide) read in a reconstructed sequence. The higher the coverage, the better as if confirms the inserted nucleotide and reduced the possibility of mismatched sequence.

Based on these definitions, sequence reads and genome annotation of WGS assay of the three bacterial isolates, as shown in Table 4.1b were analyzed. The results indicate that *Bacillus velezensis* (6-2) was the best assembled genome, closely followed by *Bacillus infantis* (63-11) and *Bacillus amyloliquefaciens* (78-1) species. *Bacillus velezensis* (6-2) isolate had the lowest number of contigs and longest contig length, hence least fragmentation; highest N50 value; highest coverage, hence more accurate reading.

Table 4.1b Summary of sequence reads and genome annotation

	6-2	63-11	78-1
GC (%)	46.21	45.92	45.72
Number of contigs	17	25	74
Longest contig length	1,613,000 bp	865270 bp	445548 bp
Total bases (genome size)	3,951,373	4803916	4085300
N50	1,016,459	287413	198661
Average coverage	112X	**150X**	109X
Predicted coding sequences	3884	4844	4081

4.1.1 Species identification and genome comparison

Numerous bioinformatic tools and methods can be utilized to indicate strain identity, characteristics, closest evolutionary relation etc. based on WGS analysis. Average nucleotide identity (ANI), DNA-DNA hybridization (DDH) and circular mapping are examples of such. The ANI value gives a measure of nucleotide-level genomic relatedness among prokaryotic strains based on pairwise comparison (similarity) between coding regions of two genomes. A cutoff score of >95% indicates that the respective isolates belong to the same species. The DDH method is another reliable in silico method that gives a universal and accurate delineation of prokaryotic species and sub-species. The Genome-Genome Distance Calculator (GGDC) 2.1 BLAST gives both DDH value and G+C content difference, making the comparisons much easier. Percentage similarity based on DDH value is represented with 70% or above species boundary, and above 79% sub-species boundary. The percent difference in the genomic G+C content between distinct species have a value close to 0 and for same

species not more than 1 [126]. The ANI and DDH values (Table 4.1.1a, 4.1.1b) of draft genome sequences of all three strains was quantified using JspeciesWS web server tool and Genome-Genome Distance Calculator (GGDC) 2.1 BLAST, respectively. The reference genomes taken from the NCBI database were - *Bacillus amyloliquefaciens* DSM 7 = ATCC 23350 (Accession NC_014551), *Bacillus amyloliquefaciens* EGD-AQ14 (Accession NZ_AVQH01000059.1), *Bacillus amyloliquefaciens* strain Y2 (Accession CP003332), *Bacillus anthracis* CZC5 (Accession AP018443), *Bacillus atrophaeus* strain BA59 (Accession CP024051.1), *Bacillus cereus* ATCC 14579 (Accession NZ_CP034551), *Bacillus firmus* DS1 (Accession APVL01000002.1), *Bacillus infantis* NRRL B-14911 (Accession NC_022524.1), *Bacillus licheniformis* DSM 13 = ATCC 14580 (Accession NC_006270), *Bacillus pumilus* strain 145 (Accession CP027116.1), *Bacillus subtilis subsp. subtilis* (Accession NC_000964), *Bacillus thuringiensis* strain C15 (Accession CP021436.1), *Bacillus velezensis* strain NRRL B-4257 (Accession NZ_LLZB01000000), *Bacillus velezensis* YAU B9601-Y2 (Accession NC_017061.1).

The ANI value with 95% species boundary (Table 4.1.1a) and DDH value with 70% species boundary (Table 4.1.1b) show high degree of similarity of 6-2 strain with *Bacillus velezensis* strain NRRL B-4257 and *Bacillus velezensis* YAU B9601-Y2; 63-11 strain with *Bacillus infantis* NRRL B-14911 and 78-1 strain with *Bacillus amyloliquefaciens* EGD-AQ14 and *Bacillus amyloliquefaciens* strain Y2; The results indicating that these strains most likely belonging to the mentioned species. Also, none of the three isolates show any similarity with *Bacillus anthracis* CZC5 or *Bacillus cereus* ATCC 14579, known pathogenic strains: supporting evidence of their non-pathogenicity. Additionally, a high degree of homology was observed between the genome of 6-2 isolate and 78-1 isolate, indicating overlapping or extremely closely related. Since both these isolates belong to *Bacillus velezensis* and *Bacillus amyloliquefaciens* genus and studies have shown that *Bacillus subtilis, Bacillus amyloliquefaciens, Bacillus licheniformis,* are phylogenetically and phenotypically close species, commonly referred to as *Bacillus subtilis* species complex and *Bacillus amyloliquefaciens* subsp.

plantarum is a later hetero-typic synonym of *Bacillus velezensis* [127, 128]; This could explain the high similarity between the two isolates.

The circular map constructed using CGView Server; A rapid Bacterial genome visualization and browsing tool for circular genome mapping that uses BLAST to compare primary sequence to up to three reference genomes for identification of conserved regions, evaluates horizontal gene transfer and differences in gene copy number, and visualizes regions of a known genome covered by novel sequences [102] showed close similarity between genome of the strain 78-1 (*Bacillus amyloliquefaciens*) and BLAST 1,2,3 of *Bacillus velezensis* NRRL B-4257, 6-2 (*Bacillus velezensis*) and *Bacillus amyloliquefaciens* DSM 7=ATCC 23350 respectively (Fig. 4.1.1a). The innermost black circle and GC content represents 78-1 (*Bacillus amyloliquefaciens*). This further confirms the results from ANI and DDH table and supports the theory that *Bacillus velezensis* and *Bacillus amyloliquefaciens* are not as evolutionary distinct species.

A second circular map constructed for analysis of comparison between 63-11 (*Bacillus infantis*) isolate and *Bacillus infantis* NRRL B-14911 genome (Fig 4.1.1b), and validate the results of ANI & DDH values showed that isolate 63-11 had high similarity with *Bacillus infantis* NRRL B-14911 strain, indicating same species and possible different subspecies, as the genome of 63-11 isolate was partially incomplete as can be seen from the gaps in the map. Furthermore, it is a likely possibility that some fundamental genes might have been lost during DNA extraction or WGS analysis and further analysis is required.

Table 4.1.1a Average Nucleotide Identity - ANI (%) calculated for Bacillus velezensis (6-2), Bacillus infantis (63-11) and Bacillus amyloliquefaciens (78-1), against reference genomes taken from NCBI Database

	6-2	63-11	78-1	Seq 1	Seq 2	Seq 3	Seq 4	Seq 5	Seq 6	Seq 7	Seq 8	Seq 9	Seq 10	Seq 11	Seq 12	Seq 13	Seq 14
6-2	*	67.47	98.97	93.49	95.62	98.47	66.09	76.65	66.11	66.00	67.52	71.62	69.42	76.09	66.09	95.04	98.47
63-11	67.2	*	67.22	67.17	65.20	67.24	65.90	67.08	65.86	69.81	91.13	67.53	66.59	67.15	65.92	67.84	67.24
78-1	99.03	67.42	*	93.66	95.20	98.36	66.01	76.74	66.03	65.93	67.36	71.54	69.47	76.14	65.97	94.90	98.36
Seq 1	93.9	67.99	93.74	*	90.23	93.77	66.92	77.05	66.84	65.76	67.99	72.10	69.93	76.43	66.80	91.21	93.77
Seq 2	98.58	65.66	99.06	93.65	*	99.02	63.19	75.10	63.19	66.74	65.92	69.61	67.32	74.04	62.83	*	99.02
Seq 3	98.18	68.15	97.99	93.25	96.43	*	66.80	76.76	66.80	66.20	68.10	72.24	69.85	76.37	66.77	97.21	100.00
Seq 4	66.59	66.79	66.53	66.78	63.99	66.67	*	67.26	91.28	65.55	66.59	66.52	67.42	67.26	91.28	65.52	66.66
Seq 5	76.94	67.79	76.93	76.91	74.24	76.94	67.18	*	67.02	65.53	67.80	72.23	70.23	79.20	67.16	75.90	76.94
Seq 6	66.66	66.72	66.63	66.81	63.59	66.82	91.20	67.41	*	65.79	66.66	66.68	67.54	67.23	98.71	65.33	66.8
Seq 7	65.68	70.33	65.61	66.06	66.97	65.53	65.54	65.79	65.39	*	70.39	65.66	65.66	65.70	65.24	66.10	65.53
Seq 8	67.75	91.21	67.75	67.83	65.43	67.75	66.48	67.74	66.46	69.88	*	67.90	67.21	67.72	66.51	67.57	67.75
Seq 9	71.89	67.90	71.87	71.92	69.26	71.99	66.38	72.41	66.26	65.39	67.75	*	69.62	72.17	66.22	70.97	71.99
Seq 10	69.93	67.39	69.96	69.94	66.72	69.91	67.60	70.35	67.50	66.03	67.49	69.71	*	70.41	67.45	68.76	69.91
Seq 11	76.34	67.83	76.38	76.28	72.67	76.46	67.25	79.35	67.29	66.25	67.98	72.18	70.27	*	67.20	75.58	76.46
Seq 12	66.84	66.62	66.86	67.23	63.88	66.96	91.26	67.56	98.71	66.44	66.80	66.92	67.45	67.29	*	65.39	66.94
Seq 13	97.59	68.67	97.56	93.27	*	99.62	65.79	76.73	65.71	64.95	68.55	71.63	69.09	76.54	65.63	*	99.62
Seq 14	98.16	67.87	97.95	93.16	95.56	99.99	66.72	76.67	66.73	65.40	67.89	72.00	69.74	76.29	66.72	96.75	*

Seq 1 Bacillus amyloliquefaciens DSM 7 = ATCC 23350; Seq 2 Bacillus amyloliquefaciens EGD-AQ14;Seq 3 Bacillus amyloliquefaciens strain Y2; Seq 4 Bacillus anthracis CZC5; Seq 5 Bacillus atrophaeus strain BA59; Seq 6 Bacillus cereus ATCC 14579; Seq 7 Bacillus firmus DS1; Seq 8 Bacillus infantis NRRL B-14911; Seq 9 Bacillus licheniformis DSM 13 = ATCC 14580; Seq 10 Bacillus pumilus strain 145; Seq 11 Bacillus subtilis subsp. subtilis; Seq 12 Bacillus thuringiensis strain C15; Seq 13 Bacillus velezensis strain NRRL B-4257; Seq 14 Bacillus velezensis YAU B9601-Y2

**Default program settings applied

Table 4.1.1b DNA-DNA hybridization (DDH) values calculated between Bacillus velezensis (6-2), Bacillus infantis (63-11) and Bacillus amyloliquefaciens (78-1), against reference genomes taken from NCBI Database

	6-2					63-11					78-1				
Reference genome	DDH	Distance	>= 70%	> 79%	G+C diff	DDH	Distance	>= 70%	> 79%	G+C diff	DDH	Distance	>= 70%	> 79%	G+C diff
Seq 1	56.1	0.0587	38.84	8.64	0.12	12.8	0.9825	0.00	0.00	0.16	55.9	0.0592	37.96	8.39	0.37
Seq 2	**90.5**	0.0116	**95.96**	66.03	0.78	12.5	1.00	0.00	0.00	1.05	92.6	0.0093	**96.58**	69.21	1.26
Seq 3	**89.2**	0.013	**95.52**	63.98	0.35	12.8	0.9829	0.00	0.00	0.07	88.2	0.0141	**95.15**	62.38	0.14
Seq 4	29.6	0.1441	0.09	0.04	10.83	12.7	0.9876	0.00	0.00	10.56	29.4	0.1453	0.08	0.08	10.35
Seq 5	20.9	0.2098	0.00	0.00	3.10	12.7	0.9857	0.00	0.00	2.82	21	0.2096	0.00	0.00	2.62
Seq 6	30.3	0.1407	0.12	0.05	10.92	12.7	0.9864	0.00	0.00	10.65	30.2	0.141	0.11	0.05	10.44
Seq 7	24.9	0.1753	0.01	0.01	5.31	12.6	0.9924	0.00	0.00	5.03	24.8	0.1757	0.01	0.01	4.82
Seq 8	26.1	0.1664	0.02	0.01	0.17	77.1	0.1468	**87.69**	52.65	0.11	25.7	0.1693	0.01	0.01	0.32
Seq 9	18.8	0.2334	0.00	0.00	0.01	12.8	0.9845	0.00	0.00	0.27	18.9	0.2327	0.00	0.00	0.48
Seq 10	18.3	0.2397	0.00	0.00	5.04	12.7	0.9876	0.00	0.00	4.76	18.4	0.2385	0.00	0.00	4.55
Seq 11	20.5	0.2142	0.00	0.00	2.69	12.8	0.9825	0.00	0.00	2.41	20.6	0.2135	0.00	0.00	2.20
Seq 12	30.2	0.1413	0.11	0.05	10.78	12.7	0.9876	0.00	0.00	10.51	30.1	0.1414	0.11	0.05	10.3
Seq 13	**87.8**	0.0145	**94.99**	61.68	0.50	12.6	0.9961	0.00	0.00	0.23	87.5	0.0149	**94.85**	61.14	0.02
Seq 14	**89.2**	0.013	**95.51**	63.93	0.34	12.8	0.9829	0.00	0.00	0.07	88.2	0.0141	**95.14**	62.35	0.14

Seq 1 Bacillus amyloliquefaciens DSM 7 = ATCC 23350; Seq 2 Bacillus amyloliquefaciens EGD-AQ14;Seq 3 Bacillus amyloliquefaciens strain Y2; Seq 4 Bacillus anthracis CZC5; Seq 5 Bacillus atrophaeus strain BA59; Seq 6 Bacillus cereus ATCC 14579; Seq 7 Bacillus firmus DS1; Seq 8 Bacillus infantis NRRL B-14911; Seq 9 Bacillus licheniformis DSM 13 = ATCC 14580; Seq 10 Bacillus pumilus strain 145; Seq 11 Bacillus subtilis subsp. subtilis; Seq 12 Bacillus thuringiensis strain C15; Seq 13 Bacillus velezensis strain NRRL B-4257; Seq 14 Bacillus velezensis YAU B9601-Y2

**Default program settings applied

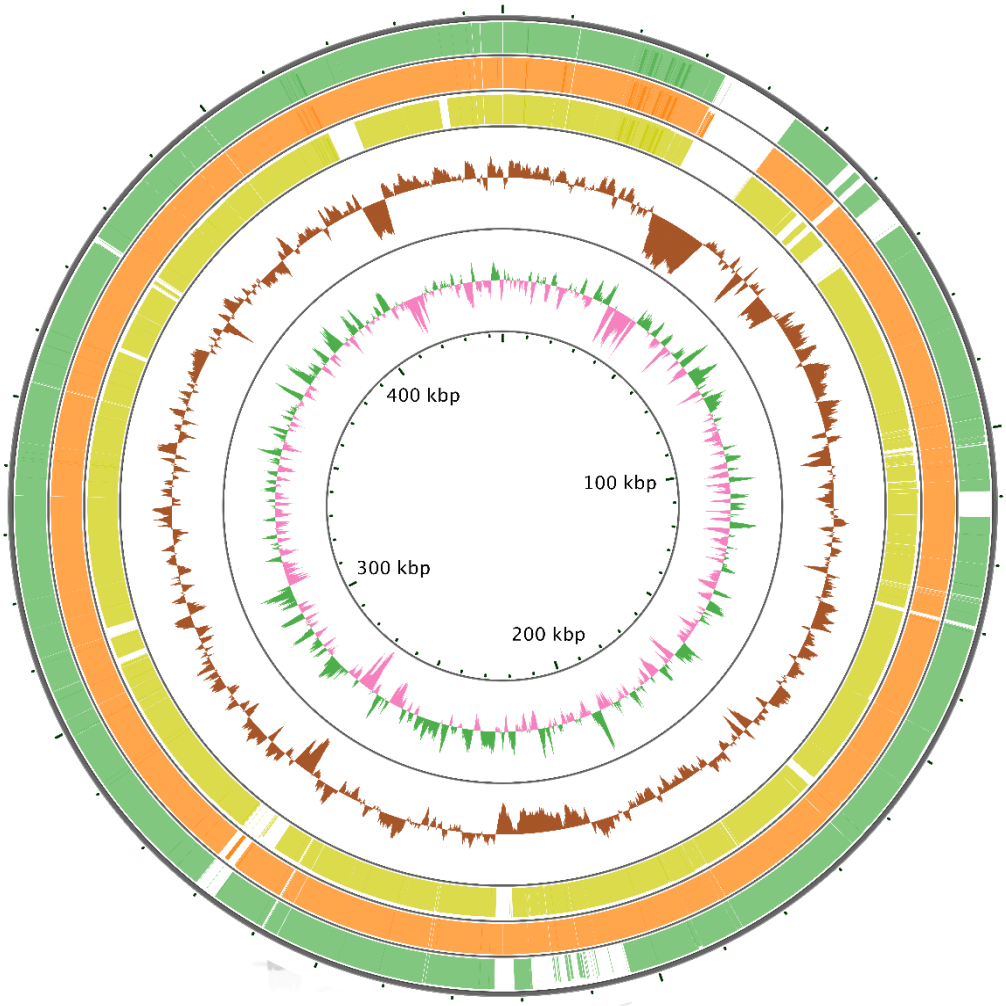

Figure 4.1.1a Comparisons between the genome of 78-1 (Bacillus amyloliquefaciens) and its closest Bacillus species - Bacillus velezensis NRRL B-4257, 6-2 (Bacillus velezensis) and Bacillus amyloliquefaciens DSM 7=ATCC 23350. Isolate 78-1 is represented by the GC content and the innermost circle. Constructed using CGView Server.

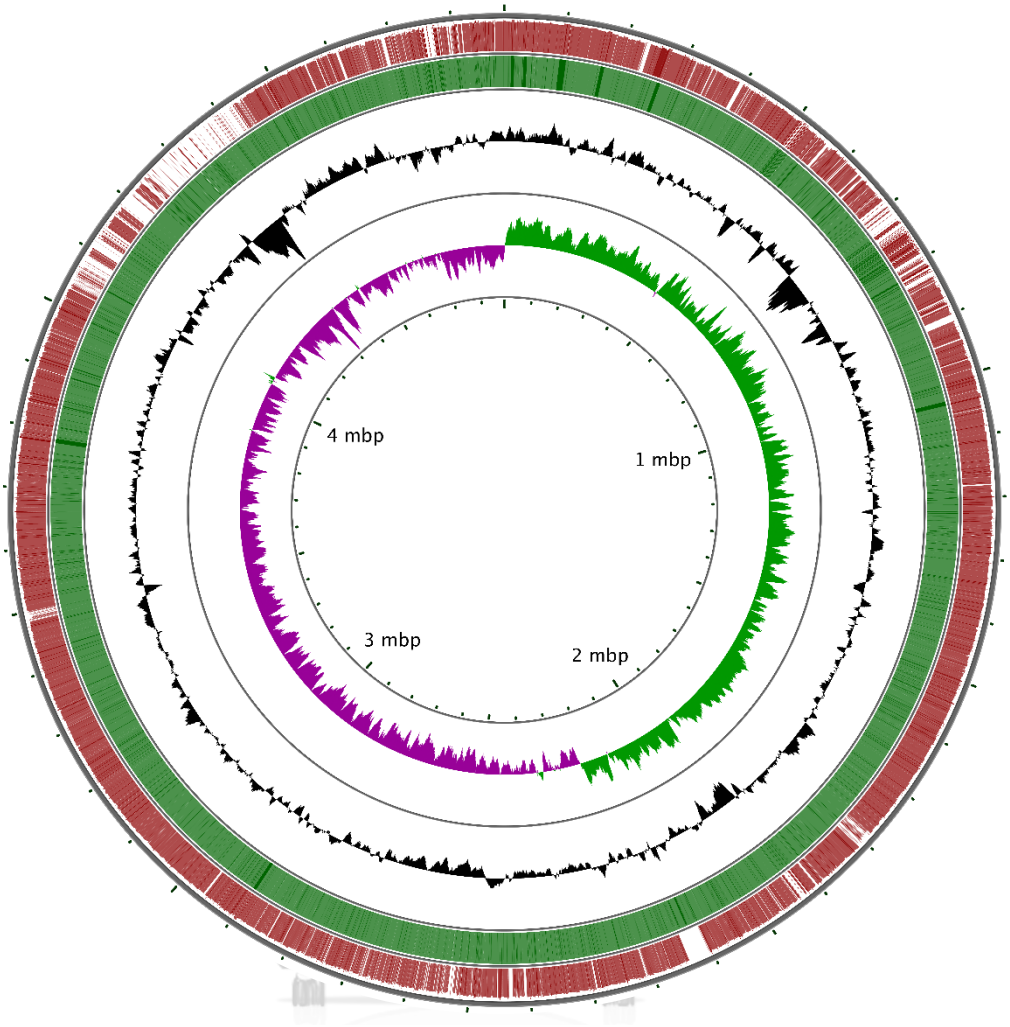

Figure 4.1.1b The comparison between 63-11 - Bacillus infantis (BLAST 1) isolate and Bacillus infantis NRRL B-14911 (BLAST 2) genome constructed using CGView Server tool.

To further verify the results of ANI and DDH methodology and circular map, evolutionary relationship of all the three isolates was determined by constructing a phylogenetic tree. The phylogenetic analysis of the three isolates and reference genomes was done using Type (Strain) Genome Server (TYGS), that analyzed the results based on both Whole Genome sequencing and 16s RNA sequencing (Fig 4.1.1c). Determination of closest type strain genomes was done in two complementary ways. First, the isolates and reference genomes were compared against all type strain genomes available in the TYGS database via the MASH algorithm [129]. Second, an additional set of ten closely related type strains extracted from the isolates and

reference genomes were determined via the 16S rDNA gene sequences, using RNAmmer [130] and each sequence was subsequently BLASTed [131] against the 16S rDNA gene sequence of each of the currently 11767 type strains available in the TYGS database. This was used as a proxy to find the best 50 matching type strains for each genome and to subsequently calculate precise distances using the Genome BLAST Distance Phylogeny approach (GBDP) [132].

The GBDP approach indicated Whole Genome sequencing based-results (Fig. 4.1.1c-1) to be more accurate and in parallel with the ANI and DDH values, as compared to 16s RNA sequencing-based results (Fig. 4.1.1c-2). Isolates 6-2 and 78-1 located in the same cluster, showed proximity to each other, and closely related to *Bacillus velezensis* and *Bacillus amyloliquefaciens* species, supporting the results of circular map (Fig. 4.1.1a). Isolate 63-11 showed highest affinity to *Bacillus infantis* NRRL B-14911 species through both Whole Genome and 16s RNA based sequencing, also supporting circular map results in (Fig 4.1.1b).

A zoomed in view of the GBDP tree of *Bacillus* isolates of both the Whole-genome sequence-based and 16S rDNA gene sequence-based as constructed by TYGS is shown in (Fig. 4.1.1d) to get a better understanding of the phylogenetic evolution of the three strains against the reference genomes.

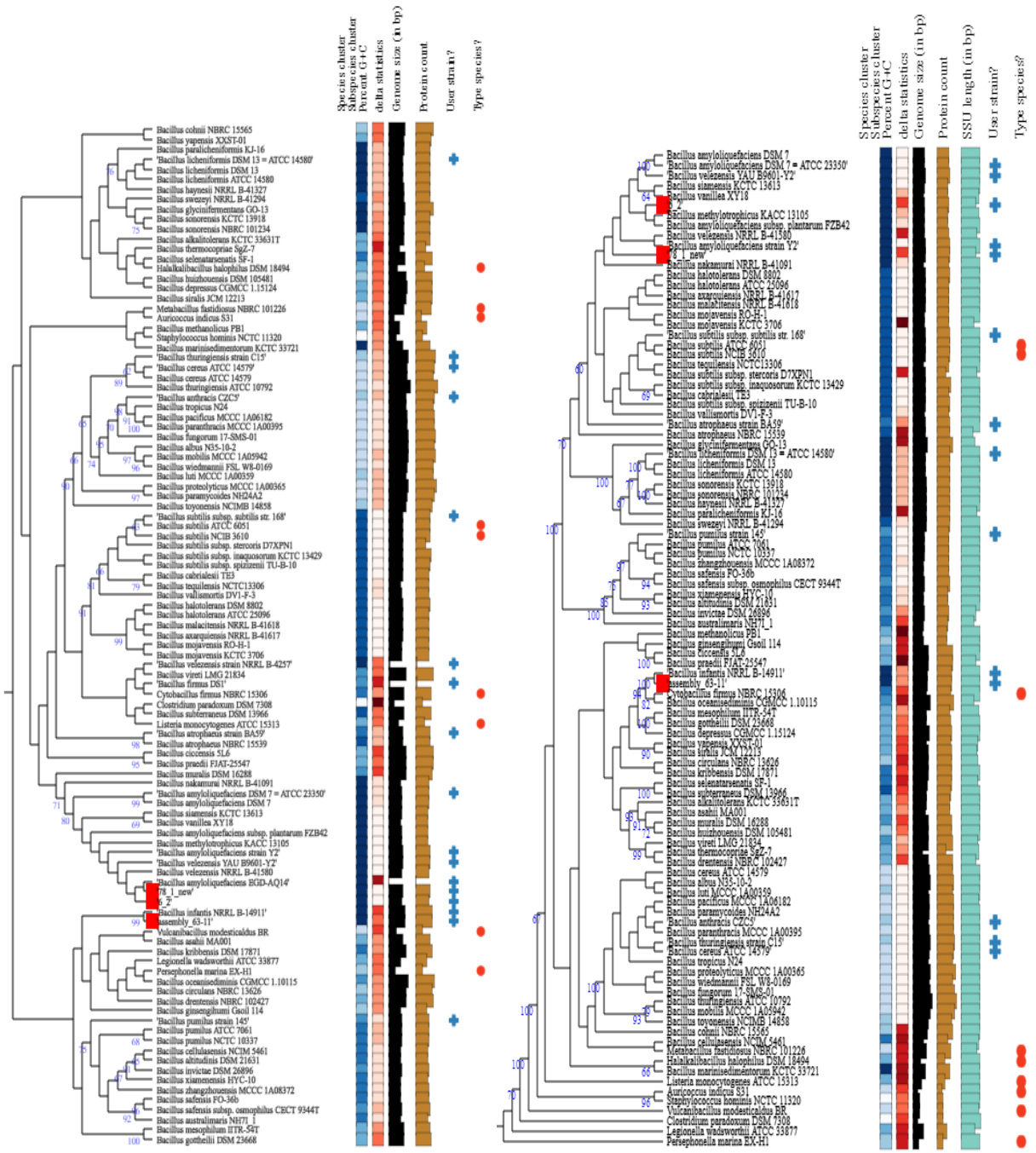

Figure 4.1.1c Genome BLAST Distance Phylogeny tree of Bacillus isolates – (1) Whole-genome sequence-based and (2) 16S rDNA gene sequence-based; Constructed using Type (Strain) Genome Server (TYGS).

**Additional references of TYGS server used to widen the search

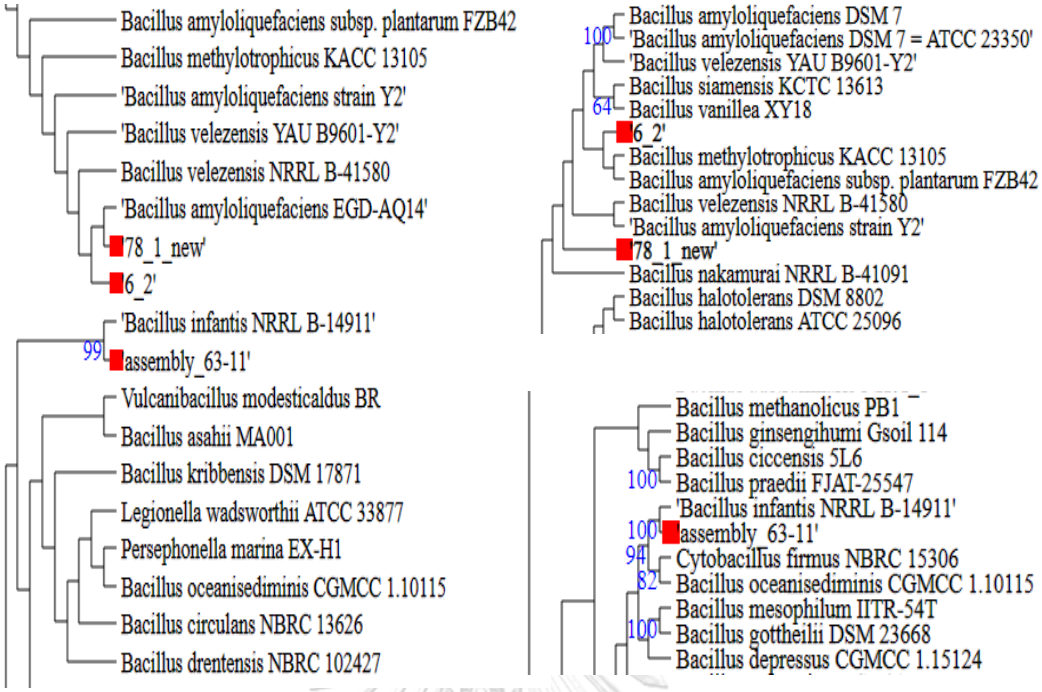

Figure 4.1.1d Close-up of Genome BLAST Distance Phylogeny tree of Bacillus isolates – (1) Whole-genome sequence-based and (2) 16S rDNA gene sequence-based; Constructed using Type (Strain) Genome Server (TYGS).

4.1.2 Screening for antibiotic and antimicrobial resistance genes

Antimicrobial resistance encompasses of resistance to drugs that fails to treat microbial infections caused by bacteria, parasites, viruses, and fungi [133]. It has increasingly become a serious threat to global public health. As a results microorganism have developed resistance to drugs and medicines have become ineffective, increasing the risk of spread to others. Some bacterial species have intrinsic antimicrobial resistance towards specific drugs that is non-transferable, and some have extrinsic resistance caused by mutation or external factors that are indeed transferable. Also known as acquired resistance, these microorganisms show resistance to specific antibiotics. Certain organisms commonly referred to as “superbugs” show resistance towards all drugs and are extremely virulent [134].

Computational screening for intrinsic and acquired antimicrobial resistant (AMR) was conducted by protein homolog analysis using Comprehensive Antimicrobial Resistance Database (CARD), DNA-based analysis using ResFinder tool of Center for Genomic

Epidemiology and BLAST analysis of *Bacillus* isolates against Erythromycin ribosomal methylase (*erm*), Aminoglycoside O-nucleotidyltransferase (*aadD2*) and Chloramphenicol acetyltransferase (*cat*) genes of *Bacillus clausii*. Lastly, KmerResistance tool of CGE was used to analyze AMR genes based on k-mer sequences to overcome the uncertain results that might have been caused by missing genes or draft genomes of the three *Bacillus isolates* [108].

Comprehensive Antimicrobial Resistance Database (CARD), is an excellent source of curated DNA and protein reference sequences that confers or contributes to resistance to various antibiotics. The tool identifies the antimicrobial resistance gene based on two phenomenon's - BLAST and RGI analysis. The RGI (Resistance Gene Identifier) predicts resistome(s) from protein, genome, or metagenomics data based on homology and SNP models. The RGI tool can sequence both high quality assemblies (includes contigs > 20,000 bp), as well as low quality/coverage assemblies (<20,000 bp), with the option of excluding/including prediction of partial genes respectively. CARD analyses AMR genes based on protein homolog model detects a protein sequence based on its similarity to a curated reference sequence and assigns it a BLASTP bitscore cutoff score to match its strength. If the sequence is 100% identical to the reference sequence along its entire length, it is deemed a "perfect match". If the match is not identical but the bitscore of the matched sequence is greater than the curated BLASTP bitscore cutoff, it is deemed "strict match" [104]. The results analyzed through CARD as shown in Table 4.1.2a predicted a strict match with clbA (Macrolide, Lincosamide, Streptogramin, Oxazolidinone, Phenicol, Pleuromutilin) and tet (Tetracycline) antibiotic resistance genes in isolate 6-2 (*Bacillus velezensis*) and 78-1 (*Bacillus amyloliquefaciens*). No antibiotic resistance genes were found in 63-11 (*Bacillus infantis*) isolate.

Table 4.1.2a Analysis of antibiotic resistance genes using Comprehensive Antimicrobial Resistance Database (CARD)

RGI Criteria	AMR gene	Detection Criteria	AMR gene family	Drug Class	Resistance Mechanism	% Identity	Bitscore	Cut-off
6-2								
Strict	clbA	protein homolog model	Cfr 23S ribosomal RNA methyltransferase	macrolide antibiotic, lincosamide antibiotic, streptogramin antibiotic, oxazolidinone antibiotic, phenicol antibiotic, pleuromutilin antibiotic	antibiotic target alteration	98.85	695.3	600
Strict	tet (45)	protein homolog model	major facilitator superfamily (MFS) antibiotic efflux pump	tetracycline antibiotic	antibiotic efflux	75.05	696	450
63-11								
No data available								
78-1								
Strict	clbA	protein homolog model	Cfr 23S ribosomal RNA methyltransferase	macrolide antibiotic, lincosamide antibiotic, streptogramin antibiotic, oxazolidinone antibiotic, phenicol antibiotic, pleuromutilin antibiotic	antibiotic target alteration	98.57	693.7	600
Strict	tet (45)	protein homolog model	major facilitator superfamily (MFS) antibiotic efflux pump	tetracycline antibiotic	antibiotic efflux	75.05	696	450

**Strict criteria settings applied to identify non-exact matches as well.

Since the results from protein homolog model of CARD identified AMR genes in two of the isolates, the isolates were further examined using another tool (Resfinder) to confirm presence of any acquired AMR genes in total or partial DNA-based sequence of bacteria isolates. The acquired antibiotic resistance genes database of ResFinder is compiled from existing databases, e.g., the ARDB, and is reasonably complete as new genes are continuously being added. The tool is able to identity phages that are vehicles responsible for horizontal gene transfer (HGT) between bacteria within the same [135, 136] or different species [137], as well as transfer of antibiotic resistance genes [138, 139] usually by the process of generalized transduction. The ResFinder tool of Center for Genomic Epidemiology was set at 90 % ID threshold and 60 % minimum length and the results in Table 4.1.2b showed no AMR genes present in any of the three *Bacillus* isolates.

Table 4.1.2b Acquired antimicrobial resistance gene results using Resfinder tool of CGE (**Default program settings applied)

Resistance gene	Gene code	Hits		
		6-2	63-11	78-1
Rifampicin	Rifr	-	-	-
Oxazolidinone	Optr	-	-	-
Nitroimidazole	nim	-	-	-
Fosfomycin	fos	-	-	-
Macrolide	emr / mef / mrea	-	-	-
Tetracycline	tet	-	-	-
Glycopeptide	van	-	-	-
Phenicol		-	-	-
Trimethoprim	dfr	-	-	-
Quinolone		-	-	-
Beta-lactam	AmpC / ESBLs	-	-	-
Colistin	mcr	-	-	-
Fusidicacid	fusA	-	-	-
Sulphonamide	sul	-	-	-
Aminoglycoside		-	-	-

Since antibiotic resistance can also sometimes refer to as extrinsic antimicrobial resistance that occurs due to external change in the genome of a bacteria (mutation), additional genome analysis for establishing antibiotic resistance of whole genome sequence of all three isolates (6-2, 63-11, 78-1) was carried out by BLAST analysis of the three isolates against Erythromycin ribosomal methylase (*erm*), Aminoglycoside O-nucleotidyltransferase (*aadD2*) and Chloramphenicol acetyltransferase (*cat*) gene of *Bacillus clausii*.

Erm proteins are part of the RNA methyltransferase family and methylate A2058 (*E. coli* nomenclature) of the 23S ribosomal RNA conferring degrees of resistance to Macrolides, Lincosamides and Streptogramin b [104]. aadD2 is an aminoglycoside enzyme specific for streptomycin/streptidine synthesis [140]. Studies have shown aadD2 gene shares 47% identity with ant (4′)-Ia from Staphylococcus aureus, which encodes an aminoglycoside 4′-O-nucleotidyltransferase, that conferred resistance to kanamycin, tobramycin, and amikacin [141].Chloramphenicol acetyltransferase (CAT) gene encodes for Chloramphenicol resistance caused due to enzymatic inactivation by acetylation and is found to be chromosomally located in all four resistant *Bacillus clausii* strains [142].

According to BLAST results showed in Table 4.1.2c, erm, *aadD2* and *cat* gene of *Bacillus clausii* were identified in 6-2 (*Bacillus velezensis*) and 78-1 (*Bacillus amyloliquefaciens*), however since the Evalue (number of expected hits of similar quality) was not less than 0.1 the results were considered not significant. BLAST and MegaBLAST analysis of isolate 63-11 (*Bacillus infantis*) showed no significant hits against all three antibiotic resistance genes.

Table 4.1.2c BLAST and MegaBLAST analysis of Bacillus isolates against erm gene, aadD2 and CAT gene of Bacillus clausii.

	BLAST			MegaBLAST
	Description	Evalue	% Identity	
6-2				
erm	*Bacillus clausii* DSM8716	0.32	84.85%	No significant similarity found
aadD2	*Bacillus clausii* NR	0.26	91.30%	No significant similarity found
cat	*Bacillus clausii* KSM-K16	0.17	88.46%	No significant similarity found
63-11				
erm	No significant similarity found			No significant similarity found
aadD2	No significant similarity found			No significant similarity found
cat	No significant similarity found			No significant similarity found
78-1				
erm	*Bacillus clausii* DSM8716	1.5	100.00%	No significant similarity found
aadD2	*Bacillus clausii* NR	0.10	100.00%	No significant similarity found
cat	*Bacillus clausii* KSM-K16	0.82	94.44%	No significant similarity found

**Default program settings applied

All three tools – RGI tool of CARD, Resfiner and MegaBLAST, identified AMR genes either by assembling raw reads and comparing against reference database or by mapping the reads directly to reference sequences. Possibility of false positives results increases as some genes can be missed if split over two or more contigs. This reduces the overall quality and sensitivity of the results. To maintain the sensitivity and keep low false positive results, *k*-mers approach can be used to map the raw WGS data against reference databases and identify resistance genes as well as determine the species. Mapping against the species reference can then be used to normalize the antimicrobial resistance prediction. *k-mers* are DNA fragments of length "*k"*, occuring in a genome, that measure the coverage and depth of predicted species and ensure "no match found" in case of any mismatch or indel in the query genome [108].

The results from KmerResistance tool of CGE in Table 4.1.2d identified isolate 6-2 (*Bacillus velezensis*) and 78-1 (*Bacillus amyloliquefaciens*) contained cfr(B) and tet(L) gene with a very high query coverage (percentage that overlaps the reference/template sequence). The cfr(B) gene encodes for Ribosomal RNA large subunit methyltransferase Cfr of *Staphylococcus aureus*; Methylates position 8 of adenine 2503 in 23S rRNA that confers resistance to some classes of antibiotics - chloramphenicol, florfenicol, clindamycin and linezolid [143]. tet(L) encodes for Tetracycline resistance protein Tet(L) of *Streptococcus thermophilus* and is able to transport across a lipid bilayer but not the nuclear membrane [144].

The results in Table 4.1.2d, are highly significant as they show a coverage (number of times a unique nucleotide is read) of nearly 100% and confirm that these two isolates indeed have AMR genes in their genome. None of the techniques used in the study predicted the presence of any AMR genes in isolate 63-11 (*Bacillus infantis*). Hence, we can assume that this isolate is relatively safe.

Table 4.1.2d k-mer resistance results for Bacillus velezensis, Bacillus amyloliquefaciens and Bacillus infantis isolate

Strain	Template	Score	Template length	q value	p value	Template coverage	Query coverage	Depth
6-2	NZ_CP023075.1 *Bacillus velezensis* strain K26 chromosome, complete genome	132336	150593	132316.22	1.0e-26	88.55	89.66	0.88
	cfr(B)_3_KR610408	681	1050	679.87	1.0e-26	100.00	100.00	1.00
	tet(L)_5_X08034	1020	1377	1019.01	1.0e-26	94.34	106.00	1.23
63-11	NC_022524.1 *Bacillus infantis* NRRL B-14911, complete genome	52367	191254	52318.07	1.0e-26	27.34	27.79	0.27
78-1	NZ_CP023075.1 *Bacillus velezensis* strain K26 chromosome, complete genome	129269	150593	129238.73	1.0e-26	86.39	84.42	0.86
	cfr(B)_3_KR610408	687	1050	685.87	1.0e-26	100.00	100.00	1.00
	tet(L)_5_X08034	1020	1377	1019.00	1.0e-26	95.64	104.56	1.24

**Default program settings applied

4.1.3 Screening for toxin and virulent genes

Computational screening for pathogenic genes was conducted using PathogenFinder tool of Center for Genomic Epidemiology and toxin genes using VFanalyzer (Virulence Factors analyzer) tool of VFDB. Additionally, the isolates were MegaBLAST against hly, cyt, bce, hbl and nhe genes of *Bacillus cereus* species. Following reference genomes - *Bacillus cereus* E33L, hly gene (GenBank: CP000001.1), *Bacillus cereus* strain TIAC959 CytK (cytK) gene (GenBank: KP409163.1), *Bacillus cereus* bceT gene for diarrheal enterotoxin (GenBank: D17312.1), *Bacillus cereus* strain 5.39 hemolysin BL (hblA) gene (GenBank: KF681259.1), *Bacillus cereus* strain R1 HblB toxin gene (GenBank: MK268740.1), *Bacillus cereus* strain EC303 hemolysin BL component L2 (hblC) gene (GenBank: JQ039144.1), *Bacillus cereus* partial hblD gene for hemolysin BL binding component L1, strain BC (GenBank: AJ937194.1), *Bacillus cereus* nheA gene, strain MHI 1761 (GenBank: FN825684.1), *Bacillus cereus* partial nheB gene for Enterotoxin B, strain BK (GenBank: AJ937178.1), *Bacillus cereus* nheC gene, strain MHI 1672 (GenBank: FN825685.1), were taken from the NCBI Database for comparative search. Gene expression profile was also analyzed by using BlastKOALA tool of KEGG (Kyoto Encyclopedia of Genes and Genomes) Database.

PathogenFinder tool predicts bacteria's pathogenicity towards human hosts by sequencing assembled genome/ contigs against all bacterial genomes - both pathogenic and non-pathogenic families. The results obtained from PathogenFinder tool shown in Table 4.1.3a, matched 6-2 and 78-1 strain with no pathogenic families and 63-11 strain to one pathogenic family, *Streptococcus suis* 05ZYH33 (Accession CP000407) with a percentage identity of 84.21%; Present in 30S Ribosomal protein S21, a small subunit “split protein” that are selectively removed from 30S subunits under low salt conditions [145] Since all three isolates are halophilic, presence of this protein is not a concern. Besides, evidence suggests that in prokaryotes, the peptidyl transferase reaction is performed by the large 23S rRNA subunit [146].

Furthermore, the probability of a genome/contig being a human pathogen falls between a range of 0 to 1 and since the probability of all the strains as predicted by Pathogenfinder was very close to zero, all three strains were predicted as be non-human pathogens.

Table 4.1.3a Pathogenic gene findings (**Default program settings applied)

	6-2	63-11	78-1
Probability of being a human pathogen	0.227	0.286	0.226
Matched Pathogenic Families	0	1	0
Matched Not Pathogenic Families	69	7	68

Virulence factor database (VFDB) is a comprehensive database of curating information about virulence factors of all bacterial pathogens. It identifies virulence factors (VFs) that enable a microorganism to establish itself on or within a host of a species and enhance its potential to cause disease. These virulence factors include bacterial toxins, cell surface proteins that mediate bacterial attachment, cell surface carbohydrates and proteins that protect a bacterium, and hydrolytic enzymes that may contribute to the pathogenicity of the bacterium. VFanalyzer systematically classifies known/potential VFs in given complete/draft bacterial genomes by comparing them against genomes preexisting in the VFDB. The virulence results shown in Table 4.1.3b predicted no significant hits for hemolytic enterotoxin and/or non-hemolytic enterotoxin genes for any of the three isolates, with the exception of hemolysin III gene in isolate 6-2 and 78-1 and hemolysin III homolog in isolate 63-11 (*Bacillus infantis*)

Hemolysins are extracellular toxic proteins which are produced by many gram-positive and gram-negative, which possess a certain pathogenic potential [147]. Hemolysin III gene is a pore-forming hemolysin determinant of *Escherichia coli* that causes temperature-independent erythrocyte lysis [148]. However, other bacterial components such as adhesins, serum resistance, capsules, iron transport systems and other virulence factors are also connected with the pathogenicity of these bacteria.

Hemolysin proteins do not solely define a bacterium pathogenic and are not specifically located on transmissible plasmids [149].

Although, VFanalyzer only screens those genomes that belong to the genera of bacterial pathogens with full information in VFDB. This means if the isolate belongs to a different bacterium or is only partially sequenced, the predicted results might not be as accurate [110]. Hence, *Bacillus* isolates were screened against hemolysin (*hly*), cytotoxin K (*cyt*), diarrheal toxin (*bce*), hemolytic enterotoxin (*hbl*) and nonhemolytic enterotoxin (*nhe*) genes of *Bacillus cereus* species using MegaBLAST tool of NCBI database.

MegaBLAST, unlike BLAST screens for "highly similar" sequences. The results from MegaBLAST analysis in Table 4.1.3b showed absence of all standard toxin genes in all three isolates. The tool also has its disadvantages as it is designed to align sequences that are nearly identical, differing by only a few percent from one another [150]. Since the reference genomes taken from NCBI belonged to *Bacillus cereus*; a known virulent *Bacillus* species and not *Bacillus velezensis, Bacillus infantis* or *Bacillus amyloliquefaciens,* it is possible the sequences might have been misaligned and some protein functions lost during the translation.

As a result, a third computation tool (BlastKOALA) was used to analyze the virulence genes of the three isolates. BlastKOALA (KEGG Orthology And Links Annotation) tool of KEGG Database, is an annotation tool that assigns K number to query genes by BLAST search against nonredundant set of KEGG GENES [112]. The KEGG GENES database is a collection of all the completely sequenced and some partial sequenced genomes with up-to-date annotation of gene functions. It not only evaluates sequence similarity of individual genes but also examines whether an organism contains a complete set of genes representing a higher order biological function and if those genes are physically coupled on the chromosome or not [151, 152]. The results of BlastKOALA shown in Table 4.1.3b, indicate presence of hemolysin III gene in all the bacterial

isolates. Given that this result coincides with the result of VFanalyser, it is essential that animal and human-model bases studies be conducted along with WGS to evaluate the degree of pathogenicity of these three bacterial isolates.

*Table 4.1.3b Predicted Toxin genes of Bacillus isolates (**Default program settings applied)*

	VFanalyser			MegaBLAST			BlastKOALA		
	6-2	63-11	78-1	6-2	63-11	78-1	6-2	63-11	78-1
Hemolysin II (hlyII)	-	-	-	-	-	-	-	-	-
Hemolysin III (hlyIII)	+	-	+	-	-	-	+	+	+
Cytotoxin K (cytK)	-	-	-	-	-	-	-	-	-
Diarrheal toxin (bceT)	N/A	N/A	N/A	-	-	-	-	-	-
Hemolytic enterotoxin (hbl genes)									
hblA	-	-	-	-	-	-	-	-	-
hblB	-	-	-	-	-	-	-	-	-
hblC	-	-	-	-	-	-	-	-	-
hblD	-	-	-	-	-	-	-	-	-
Nonhemolytic enterotoxin (nhe genes)									
nheA	-	-	-	-	-	-	-	-	-
nheB	-	-	-	-	-	-	-	-	-
nheC	-	-	-	-	-	-	-	-	-

4.1.4 Screening for secondary metabolite gene clusters

Bacillus species, mainly *Bacillus velezensis* species are widely known to synthesize various bacteriocin or bacteriocin like inhibitory substances (Smitha & Bhat, 2013) that encode for polyketides, lipopeptides and dipeptides with high similarity to the closest related *Bacillus* species [153, 154]. These substances are in general known as secondary metabolites that demonstrate broad range of antagonistic (antibiotic, antimicrobial, antiviral) properties against pathogens [155]. These genes can also sometimes give false pathogenic/virulent results due to their inhibitory properties.

Secondary metabolite biosynthesis for the three isolates used in this study was searched through NP.searcher tool, which gave a brief prediction of the possible gene clusters as shown in Table 4.1.4. Secondary metabolites synthesis searched through antiSMASH (bacterial version) tool was able to detects well-defined clusters containing all required parts as well as partial clusters missing one or more functional parts. The analysis was able to give a more quantified and detailed description of gene clusters present in the three isolates.

*Table 4.1.4 Secondary metabolite gene clusters searched through NP.searcher tool (**Default program settings applied)*

	NRPS	PKS	Mixed NRPS/PKS	trans AT PKSs	Mevalonate terpenoid genes	Non-mevalonate terpenoid genes
6-2	3	1	1	2	0	1
63-11	0	0	0	0	0	1
78-1	2	1	1	2	0	1

For 6-2 isolate, Polyketide gene clusters responsible for biosynthesis of difficidin, bacillaene and macrolactin H showed percentage similarity of 100% with *Bacillus velezensis* FZB42. Non-ribosomal peptide synthases (NRPS) gene clusters associated with the biosynthesis of lipopeptides like surfactin and fengycin showed 82% and 93% similarity with *Bacillus velezensis* FZB42 and commonly found dipeptides bacillibactin and bacilysin gene clusters showed 100% similarity to *Bacillus subtilis subsp. subtilis str. 168* and *Bacillus velezensis* FZB42 respectively. Since 6-2 isolate was identified as *Bacillus velezensis*, it is safe to assume that these results are accurate

For 78-1 isolate, Polyketide gene clusters for difficidin, bacillaene and macrolactin H showed percentage similarity of 86%,92%,90% respectively and NRPS gene clusters for fengycin, bacillibactin and bacilysin showed percentage similarity of 93%,100%,100% respectively to *Bacillus subtilis* subsp. subtilis str. 168 and *Bacillus velezensis* FZB42 strains. Given that 78-1 (*Bacillus amyloliquefaciens*) isolate showed proximity to isolates 6-2 and *Bacillus amyloliquefaciens* species has high evolutionary similarity to

Bacillus subtilis and *Bacillus velezensis* species, this result is most likely accurate. Furthermore, as almost identical gene clusters were identified in 6-2 and 78-2 isolates, the results allude to the close ecological relationship between the two species.

Isolate 63-11 was the only species from this study that did not show any gene clusters for PKSs or NRPS genes. It did however show presence of terpenoid genes responsible for biosynthesis of Carotenoid, with a 50% similarity to *Halobacillus halophilus* DSM 2266. Shares many functional properties with *Bacillus infantis* species - it is Gram-positive, endospore-forming bacteria that is moderately tolerant to halophile conditions and rarely pathogenic in nature [156].

4.1.5 Biochemical analysis of *Bacillus* isolates

Bacillus genus have been extensively studies for their ability to naturally synthesis antibiotics, enzymes vitamins and other metabolic components of industrial importance. Studies on alkaline protease synthesis by *Bacillus infantis* and numerous other enzymes and vitamins produced by *Bacillus subtilis* and its close related species have been conducted to better understand posttranslational mechanisms of metabolic regulation of *Bacillus* species [157]. However, this process is tedious. Genome sequencing has greatly simplified the process for elucidating complicated molecular systems in a cell and prediction of functional characterization of microorganisms [158].

KEGG (Kyoto Encyclopedia of Genes and Genomes) databases is a heavily cross-referenced, truly integrated database of all biological processes. It implements a rigorous Smith-Waterman dynamic programming algorithm that produces optimal, non-asymmetric alignment between two genes [159]. It analyzes both macro- and micro-molecule interactions that occur in the cells and reconstructs organism-specific pathways [158] for comparing and computing each gene against pathways in a human gut.

Pictographic representation of all metabolic pathways of isolates 6-2, 63-11 and 78-1 are demonstrated in (Fig 4.1.5). Both 6-2 and 78-1 isolates showed approximately similar set of 11 metabolic pathways and isolate 63-11 showed a total of 7 metabolic pathways. All three isolates showed carbohydrate, glycans, proteins and amino acids synthesis and/or metabolism, and other cellular process which implies that these isolated are able to generates ATP (energy) from major metabolic pathways including glycolysis and Krebs's (Citric acid) cycle and hence are likely to survive in the gastrointestinal track. This discovery is supported by the study conducted by [160] that demonstrated *Bacillus infantis* strain has the ability to metabolize D-xylose, galactose, glucose, fructose, mannitol, sorbitol, methyl a-D-glucoside, N-acetylglucosamine, amygdalin, arbutin, aesculin, salicin, maltose, melibiose, sucrose, trehalose, raffinose, starch, glycogen, gluconate, cellobiose, lactose and inulin. Results in Table 4.1.5 list all the enzymes secreted by the draft genome of the three *Bacillus* isolates used in this study. It is constructed using the KEGG ENZYME, an extension of KEGG LIGAND database, that assigns an EC number as per the IUBMB/IUPAC Biochemical Nomenclature Committee [161]. Based on some of the enzymes listed in the Table 4.1.5a, we can assume that these isolates may plan an important role in fermentation pathways.

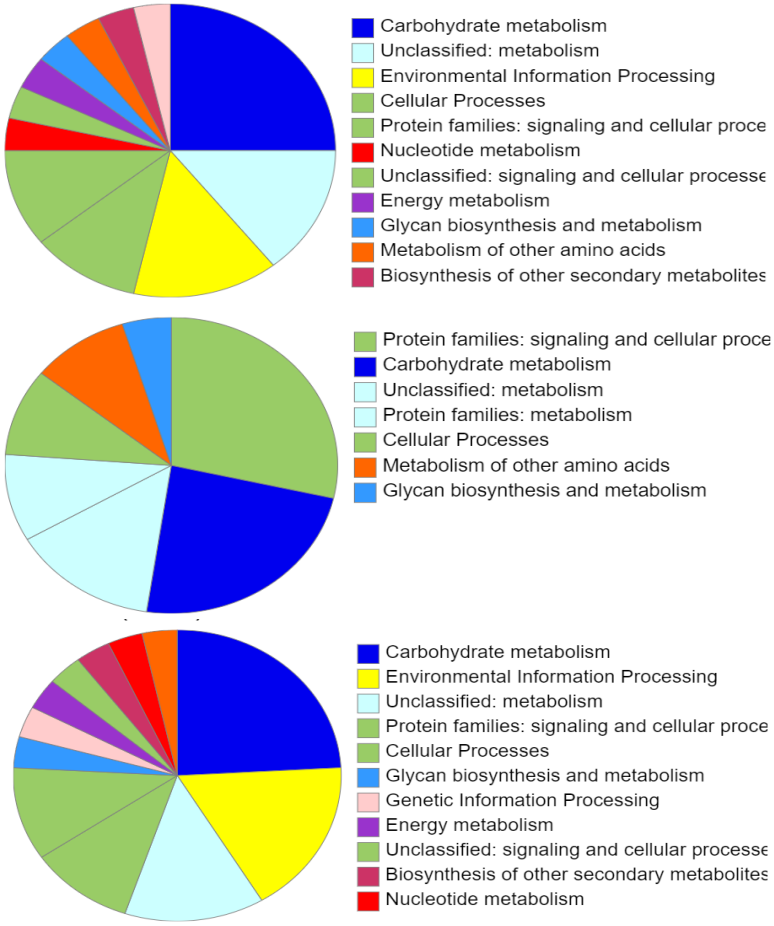

*Figure 4.1.5 Pictographic representation of all metabolic pathways of isolates 6-2 (Bacillus velezensis), 63-11 (Bacillus infantis) and 78-1 (B. amyloliquefaciens) respectively. (**Default program settings applied)*

Table 4.1.5a Predicted enzymes of *Bacillus* isolates used in this study, using KEGG ENZYME database (**Default program settings applied)

Enzyme	EC number	6-2	63-11	78-1
Oxidoreductases				
UDP-glucose 6-dehydrogenase	1.1.1.22	+	+	+
UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase	1.1.1.336	-	+	-
catalase	1.11.1.6	+	+	+
Transferases				

glutathione hydrolase	2.3.2.2.3.4.19.13	+	+	+
undecaprenyl diphosphate synthase	2.5.1.31	+	+	+
Phosphatidylglycerol prolipoprotein diacylglyceryl transferase	2.5.1.145	-	+	-
pyridoxal phosphate-dependent aminotransferase EpsN	2.6.1.-	+	+	+
adenylylsulfate kinase	2.7.1.25	+	-	+
UTP--glucose-1-phosphate uridylyltransferase	2.7.7.9	+	+	+
polyisoprenyl-teichoic acid--peptidoglycan teichoic acid transferase	2.7.8.-	+	+	+
sugar transferase EpsL	2.-.-.-	+	+	+
Hydrolases				
PPM family protein phosphatase	3.1.3.16	-	+	-
protein-tyrosine phosphatase	3.1.3.48	+	-	+
glutathione hydrolase	2.3.2.2.3.4.19.13	+	+	+
signal peptidase II	3.4.23.36	+	-	+
urease subunit alpha	3.5.1.5	+	-	+
immune inhibitor A	3.4.24.-	-	+	-
Isomerases				
UDP-glucose 4-epimerase	5.1.3.2	+	+	+
mannose-6-phosphate isomerase	5.3.1.8	+	-	+
Lignase				
gamma-polyglutamate synthase	6.3.2.-	+	-	+
Translocases				
flagellum-specific ATP synthase	7.4.2.8	+	+	+
pilus assembly protein CpaF	7.4.2.8	-	+	-

Additionally, protein in *Bacillus* isolates were analyzed by Prokka (Prokaryotic genome annotation) program of Galaxy Version 1.14.5 and Interproscan functional predictions of program of ORFs of Galaxy Version 5.0.0. Prokka annotations used COGs database for phylogenetic classification of the proteins and assigned the proteins a COG identity number; Interproscan used Pfam database (curated protein families, each represented by multiple sequence alignments and profile hidden Markov model (HMM) and assigned the proteins a PF identity tag. The distribution of protein based on these two programs, defined in Table 4.1.5b showed isolate 6-2 synthesized Environmental stress resistant and Acid tolerant proteins, isolate 78-1 synthesized Acid tolerant and Adhesive to epithelial layer proteins and isolate 63-11 synthesized all relevant and required - Environmental stress resistant, Adhesive to epithelial layer and Acid tolerant proteins (F1F0 ATP) proteins. This proves that out of the three isolates, 63-11 (*Bacillus infantis*) showed most promising results. A fourth category - Bile tolerant proteins is equally essential for the probiotics to survive the in the upper gastrointestinal track. This includes salt hydrolysate (BSH), that initiates defense mechanisms to resist the deleterious action of bile acid which can be highly toxic to non-native microorganisms.

Although intrinsic bile tolerance is likely strain-dependent, bacterial species can progressively adapt by synthesizing selective stress response proteins that may provide protection against oxidative damages and counteract some of the cellular damage induced by bile exposure [162].

Researchers have identified molecular elements that have demonstrated positive bile induced response in *Lactobacillus* and *Bifidobacterium* species [162-164]. These molecules include - Glycolytic enzymes, F0F1-ATPase, DnaK and EPS proteins. Results in (Fig 4.1.5) and Table 4.1.5b support that all three isolates synthesized enzymes responsible for glycolytic pathway as well as general stress response molecules and acid resistant proteins.

*Table 4.1.5b Genome Annotation results using Quast & Pfam analysis (**Default program settings applied)*

	Probiotic nature	Domain-	6-2	63-11		78-1	
Environmental stress resistance			Identity tag	Identity tag	E value (PF-pfam)	Identity tag	E value (PF-pfam)
	60 kDa chaperonin groEL	groL	COG0234				
	10 kDa chaperonin groES protein	groS	COG0459	COG0234			
	heat-shock	TcpE family				PF12648	4.80E-19
	C-terminal, D2-small domain, of ClpB protein	ClpB protein		PF10431	2.20E-23		
	C-terminal, D2-small domain, of ClpB protein	ClpB protein		PF10431	8.50E-26		
	AAA lid domain	ClpA/ClpB		PF17871	2.00E-31		
	AAA lid domain	ClpA/ClpB		PF17871	4.90E-37		
	ATP-dependent Clp protease ATP-binding subunit ClpE	clpE	COG0542	COG0542			
		clpP	COG0740				
		clpQ	COG5405				
		clpX	COG1219	COG1219			
		clpY	COG1220				
	Cold shock protein	cspB	COG1278				
		cspC	COG1278				
		cspD	COG1278				
	Stress response protein CsbD	csbD	COG3237			COG3237	
	Chaperone protein DnaK	dnaK	COG0443	COG0443			
	Catalase	ROS		PF00199	4.30E-171		
	Catalase	katA_1	COG0753	PF00199	1.90E-181		
	Vegetative catalase	katA_2	COG0753				
	Catalase HPII	katE	COG0753				
	Catalase-related immune-responsive	ROS		PF06628	2.80E-20		
	Catalase-related immune-responsive	ROS		PF06628	3.10E-20		
	Lysozyme-like					PF13702	4.60E-42
	Lysozyme-like					PF13702	2.50E-42

	Stress response protein NhaX	nhaX	COG0589				
	Response regulator protein VraR	vraR	COG2197				
	Stress response protein SCP2	yceC	COG2310				
Adhesion to gut	Fibronectin type III-like domain	fn3		PF14310	1.20E-25		
	Glycosyl hydrolases family 2, TIM barrel domain	TIM domain		PF02836	6.20E-128	PF02055	1.10E-32
	Flagellin	hag_1		COG1344			
	Flagellin	hag_2		COG1344			
	Flagellin	hag_3		COG1344			
	FlgD Ig-like domain	FlgD				PF13860	1.60E-07
	PEP-utilising enzyme, TIM barrel domain	TIM domain		PF02896	1.60E-127		
	Putative integral membrane proteins (Lipid transport)	MMPL family		PF03176	1.70E-59	PF03176	2.90E-74
				PF03176	3.30E-46	PF03176	1.00E-36
				PF03176	2.70E-25	PF03176	3.80E-46
				PF03176	3.10E-24	PF03176	2.10E-62
	Biofilm formation	EAL domain		PF00563	2.40E-33		
				PF00563	4.10E-77		
				PF00563	6.30E-76		
				PF00563	3.00E-71		
				PF00563	1.60E-58		
		EAL-domain associated protein		PF10388	4.30E-78		
	Putative glycosyltransferase EpsD	epsD					
	putative sugar transferase EpsL	epsL				COG2148	
	Putative acetyltransferase EpsM	epsM				COG0110	

	Putative pyridoxal phosphate-dependent aminotransferase EpsN	epsN				COG0399	
Acid tolerance	F1F0 ATP protein	atpB	COG0356	COG0356		COG0356	
		atpF	COG0711			COG0711	
		atpE	COG0636			COG0636	
		atpA	COG0056	COG0056		COG0056	
		atpD	COG0055	COG0055		COG0055	
		atpG	COG0224	COG0224		COG0224	
		atpH	COG0712			COG0712	

4.2 IN-VITRO RESULTS

4.2.1 Gastric Acid Tolerance

For the probiotics to effectively provide health benefits to the host [80] and remain viable and stable, it is essential they survive the digestion time in the body and proliferate under stressful conditions. For that they must pass through a series of barriers in the GIT. Factors such food matrix; degree of physical activity; the activity of the digestive enzymes; the pH of the digestive juices and bile salts; the composition and size of the gut microflora; and the peristaltic movements and kinetics of GIT transit affect the digestion rate and viability of the Probiotic microorganisms (ref?). The upper gastrointestinal track has pH of 2–3, and the total transit time through the GIT estimated to be 3–8 h [165], with approximately 2 h in the stomach. Gastric juices contain pepsin, rennet, and gastric lipase. Pepsin initiates protein breakdown, rennet curdles milk proteins, and gastric lipase initiates the distribution of emulsified fats [166]. Results for the gastric acid tolerance shown in Table 4.2.1 suggests high tolerance of all three isolates to gastric acid, for up to six hours of incubation, with a minimum of ¾ population surviving the low pH (pH 2.5) conditions. Surprisingly, isolate 78-1 (*Bacillus amyloliquefaciens*) demonstrated positive growth under the gastric acid conditions indicating highest tolerance as compared to the other two isolates. The results of Table 4.2.1 had a high significance of ($P < 0.05$).

Table 4.2.1 Gastric Acid Tolerance of Bacillus isolates after 6 h of incubation

Isolate Code	Isolate Name	Viable cell count (Log CFU/ml) after incubation				% survivability after 6 h
		Control	0 h	3 h	6 h	
6-2	*B. velezensis*	1.845098	1.113943	1.093422	0.835691	75.02
63-11	*B. infantis*	1.69897	0.332438	0.230449	0.278754	83.85
78-1	*B. amyloliquefaciens*	1.929419	2.464788	2.729691	2.878522	116.79

4.2.2 Bile Acid Tolerance

From the stomach, the food moves towards the lower sections of the GIT where conditions are more basic [167]. The pH of the small intestine juice is approximately 7.0 [166]. The intestine contains amylase, intestinal lactase, sucrose, aminopeptidase, and carboxypeptidase, which degrade disaccharides and polysaccharides and disrupt peptide bonds. The intestines also contain bile acid which has bactericidal effect on the survival of bacteria, as it increases the permeability of the bacterial cell membrane and causes cell lysis. The hydrolysis of bile salts by bile salt hydrolase (BSH) is a natural defense mechanism against the toxic effects of bile salts. BSH cleaves glycine or taurine moieties from conjugated bile salts [167] helping the bacteria to survive.

Previously observed studies have shown that *Bacillus* spp. are weakly tolerant to bile salt concentration [168]. Results of the bile acid tolerance in Table 4.2.2 however, show high survivability of all three isolates under bile salts even after 24 h of incubation. Isolate 63-11 (*Bacillus infantis*) shows approximately 100% survivability, closely followed by isolate 6-2 (*Bacillus velezensis*). The results had a high significance of ($P < 0.05$).

Table 4.2.2 Bile Acid Tolerance of *Bacillus* isolates after 6 h and 24 h of incubation

Incubation Period	Viable cell count (Log CFU/ml)					% survivability after 6 h	% survivability after 24 h
	Control	0 h	3 h	6 h	24 h		
6-2	5.64140	5.99430	6.63940	5.70760	5.17820	95.2	86.4
63-11	4.94944	5.12864	5.01285	5.12864	5.08453	100.0	99.1
78-1	7.08328	7.59331	7.24318	6.64769	5.40433	87.5	71.2

4.2.3 Hydrophobicity

Along with the survivability in the gastrointestinal track, it is crucial that probiotic bacterial be able to adhere to the intestinal mucosa as well as inhibit the adherence of potentially pathogenic bacteria to intestinal gut. This ability will help reduce the regular consumption of probiotics while maintaining the viable cell number for a longer duration. Hydrophobicity of the isolates of average log 4.6 CFU/ml, checked under three organic solvents – Chloroform, Ethyl acetate and Xylene (Fig 4.2.3) on average showed highest surface adhesion for xylene (apolar solvent), followed by ethyl acetate (polar aprotic solvent) and lowest for chloroform (polar aprotic solvent). Percentage hydrophobicity of B. velezensis (6-2) strain was - 45% in chloroform, 61% in ethyl acetate & 56% in xylene; Percentage hydrophobicity of B. infantis (63-11) was - 57% in chloroform, 39% in ethyl acetate & 83% in xylene and percentage hydrophobicity of B. amyloliquefaciens (78-1) isolate was - 34% in chloroform, in 70% ethyl acetate & 77% in xylene. *Bacillus infantis* (63-11) showed highest % hydrophobicity in xylene than the other two isolates, demonstrating better adherence to epithelial cell under apolar conditions. *B. amyloliquefaciens* (78-1) showed second higher % hydrophobicity in xylene and *B. velezensis* (6-2) showed highest % hydrophobicity in ethyl acetate, demonstrating better epithelial cell adherence under polar aprotic conditions. The relatively low affinities to chloroform when compared to xylene and ethyl acetate, indicate probiotic strains have nonacidic and poor electron donor property. Based on these finding, *B. infantis* (63-11) showed the most potential adherence property out of all the three isolates. The results for hydrophobicity under ethyl acetate (P=0.185) and

xylene (P=0.25) were not as significant, however the chloroform results showed high significance (P<0.05).

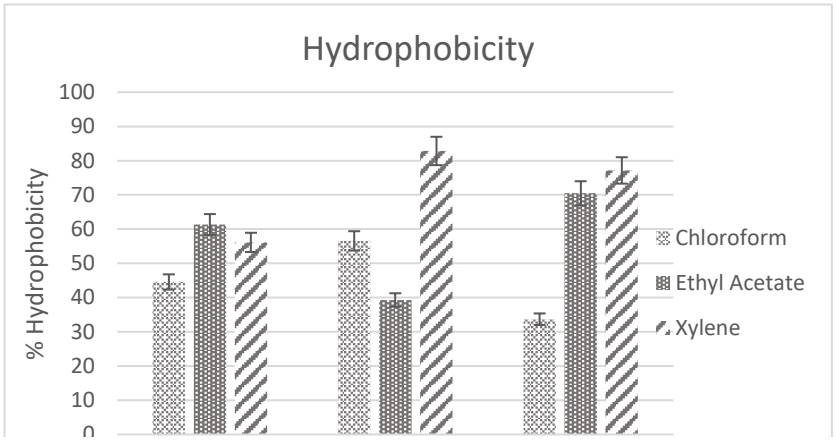

Figure 4.2.3 Cell surface hydrophobicity of bacterial isolate's against organic solvents – Chloroform (P <0.05), Ethyl Acetate (P >0.05) and Xylene (P >0.05).

4.2.4 Antibiotic Susceptibility

Along with surviving in the GIT, it is equally essential that probiotics not exhibit negative effects on native gastrointestinal microorganisms crucial for the development of the immune system, or exhibit any transmissible antibiotic resistance genes or toxin synthesis [79]. Hence, to ensure the safety and efficacy of the probiotic, screening for activity of probiotic isolates against standard antibiotics is fundamental. In the in-vitro model bacterial isolates were screened against three known antibiotics – Ampicillin (10 µg/ml), Chloramphenicol (30 µg/ml) and Tetracycline (30 µg/ml). The selection of antibiotics was done on the basis of a previous study conducted on *Bacillus infantis* strain isolated from gut of *Labeo rohita* (fish) which shown high susceptibility to (>10 mm ZOI) ampicillin and moderate susceptibility (2-4 mm ZOI) to chloramphenicol and tetracycline[93]. The ZOI of each antibiotic was measured for all the bacterial isolates of log 5 CFU/ml Table 4.2.4 shows the zone of inhibition for each antibiotics, between the range of 0-15mm. Isolate 63-11 (*Bacillus infantis*) exhibited high susceptibility (>10-15mm of ZOI) to all three antibiotics, whereas isolate 6-2 (*Bacillus velezensis*) and 78-1 (*Bacillus amyloliquefaciens*) exhibited low to moderate tolerance to the antibiotics.

European Committee on Antimicrobial Susceptibility Testing (EUCAST9), Clinical and Laboratory Standard Institute (CLSI), European Food Safety Authority (EFSA) and other similar organizations have internationally recognized methods for evaluation of antibiotic activity by calculating the MIC of each antibiotic for individual genus. MIC expressed as mg/L or µg/mL, is the measure of the lowest concentration of antimicrobial that inhibits bacterial growth. This cut-off values helps distinguish strains with acquired resistance from susceptible strains. CLSI has defined a standardized ZOI and MIC breakpoint comparison, however, unfortunately it has no comparison for *Bacillus* genus due to insufficient data. The results in Table 4.2.4 were compared against Staphylococcus spp. (Gram positive pathogen). EFSA has no conversion for ZOI into MIC value, however it does define the breakpoint for *Bacillus* genus [77]. The significance of the differences between the antibiotic and *Bacillus* isolates, compared by Duncan tests was not significant ($P > 0.05$)

Table 4.2.4 Antibiotic susceptibility of Bacillus isolates; +, ≤5mm of ZOI, ++, >5-10mm of ZOI; +++, >10-15mm of ZOI.

Antibiotics	Concentration (µg/ml)	Zone of Inhibition (mm)			ZOI and MIC breakpoint (Staphylococcus spp.)			MIC (µg/ml)
		6-2	63-11	78-1	6-2	63-11	78-1	(EFSA)
Ampicillin	10	++	+++	+	R	R	R	n.r
Chloramphenicol	30	++	+++	+++	R	I	R	8
Tetracycline	30	+	+++	++	R	I	R	8

4.2.5 Antimicrobial activity

Antimicrobial activity is another important factor in the selection criteria for probiotics. Antimicrobial activity targets the enteric undesirables and pathogen microbes [80] that may release bacteriocins and/or bacteriocin-like substances that have deleterious effects on the host immune system as well as stability of food supplement. The antagonistic effect of the isolate was tested against both Gram-positive and Gram-negative pathogens as shown in Table 4.2.5. The zone of inhibition against *Escherichia coli* ATCC 25922, *Staphylococcus aureus* ATCC 25923 and *Bacillus. cereus* ATCC 6633,

for isolate 6-2 (*Bacillus velezensis*) were 13.5, 11.0 and 16.5 mm; for isolate 63-11 (*Bacillus infantis*) were 13.8, 10.5 and 13.5 mm; for isolate 78-1 (*B. amyloliquefaciens*) were 15.5, 10.3 and 13.3 mm, respectively. On average all three isolates demonstrated moderate antagonistic effects against both Gram-positive and Gram-negative pathogens. The results have a ($P < 0.05$).

Table 4.2.5 Antibiotic susceptibility of Bacillus isolates; +, ≤10mm of ZOI, ++, >10-15mm of ZOI; +++, >15mm of ZOI

Pathogen	Zone of Inhibition (mm)		
	6-2	63-11	78-1
Escherichia coli ATCC 25922	+++	++	++
Staphylococcus aureus ATCC 25923	++	++	++
Bacillus. cereus ATCC 6633	++	++	+++

4.2.6 Hemolytic Test

Hemolysin is an important virulence factor that have the ability to lyse red blood cells (RBCs). It is classified into three types - alpha (α), beta (β) and gamma (γ). Alpha hemolysins cause a partial lysis of the RBCs, displaying a greenish zone surrounding the colonies on Blood agar plate. Beta hemolysins produce a complete lysis of the RBCs, displaying a clear transparent zone surrounding the colonies on Blood agar plate [169]. Gamma hemolysins does not induce partial or complete lysis of RBCs, and hence is known as non-hemolysin. Hemolysin molecules are extremely common in E. coli infections of the urinary tract (UTI) and other extraintestinal sites [170] and are classified into four types – hyl – I, II, III and VI (Cytotoxin K) [171]. Hemolysin III is the least characterized hemolytic toxin from the *Bacillus cereus* group. Its hly-III gene, cloned and characterized in Escherichia coli, encodes for an oligomeric pore formation toxin that acts in three steps: the protein first binds to the erythrocyte surface, monomers are then assembled to form the transmembrane pore, leading to erythrocyte lysis [172]. While the first two steps are temperature dependent, the final lysis is not [173]. The In vitro assay of hemolytic activity performed on 5% human

blood plates depicted in (Fig 4.2.6), demonstrated only strain 78-1 (B. amyloliquefaciens) had beta (β)-hemolytic activity (complete breakdown of RBSs caused by Streptolysin, an exotoxin enzyme). The other two isolates - 6-2 (*Bacillus velezensis*) and 63-11 (*Bacillus infantis*) exhibited negative signs of hemolysis.

The positive expression of 78-1 (*B. amyloliquefaciens*) isolate could be explained due the presence of cell-bound protease gene (*wpcA*) found in the isolate and/or bacteriocin synthesis in the genome, and not related to hemolysin gene function. The negative in-vitro hemolysin results of 6-2 (*Bacillus velezensis*) and 63-11 (*Bacillus infantis*) may also be due to a suppression or lack of transcription of hemolysin III and hemolysin III homolog gene found in the respective isolates. This discovery however has raised safety concerns for the use of 78-1 (*B. amyloliquefaciens*) isolate and the reliability of hemolysin test that required further intensive investigation.

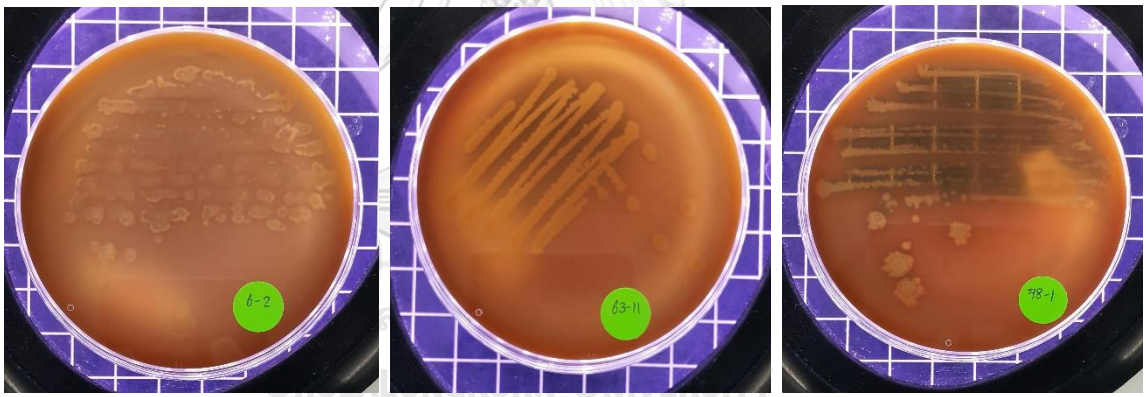

Figure 4.2.6 Left to Right: Non- hemolytic (6-2), Non- hemolytic (63-11), beta –hemolytic (78-1) due to clear transparent zone around the colonies on 5% Human blood agar plates.

4.2.7 Protease activity

Proteases is one of the largest groups of hydrolytic enzymes contributing to about 60% of total worldwide sales of industrial enzymes due to their wide applications in various industries. Microbial proteases have been preferred over the animal and plant proteases because of their fundamental features and ease in production [174]. Proteolytic bacteria produce the protease enzymes that helps break down peptide bonds in protein molecules. These enzymes can be extracellular - produced within

the cell and then released out [175] or cell bound, that cannot be released out of the cell. Although protease enzyme has its various industrial applications, it can have a co-hemolytic effect resulting in the lysis of biological membranes [176], often giving false hemolytic results.

Protease activity of the isolates was evaluated to further investigate the hemolysin activity demonstrated by the bacterial strains. For the results shown in Table 4.2.7, OD values of proteolytic activity (R1 and R2) using azocasein (nonspecific protease substrate that causes casein to hydrolyzed and azo dye released in the cell) were calculated after 15 mins of incubation at 37°C, Blank was measure using 5% saline solution as negative control. The difference was calculated and averaged to get a more precise assessment of protease activity. The results showed that 63-11 (*Bacillus infantis*) isolate had the highest protease activity and 78-1 (*Bacillus amyloliquefaciens*) isolate had the lowest. The reason for low protease activity of 78-1 (*B. amyloliquefaciens*) strain could be because of the presence of cell bound protease genes found in the bacterial genome. On the other hand, the reason for high protease activity of 63-11 (*Bacillus infantis*) strain could be because of the presence of exo-protease genes found in its bacterial genome. Another explanation for the high activity could be that *Bacillus infantis* strain in an effective protease producing bacteria, supported by the study carried out by the Department of Biotechnology, Amritsar, Punjab, India [174]. Both these results lead to two conclusions, (1) protease activity has no effect on the hemolytic activity of the isolates, (2) *Bacillus infantis* strain 63-11 has additional industrial application along with the potential probiotic activity.

Table 4.2.7 Quantitative screening of proteinase activity of Bacillus isolates using azocasein

Isolate code	Isolate Name	Protease activity				
		R1	R2	R1-blank	R2-blank	**Average**
6-2	*B. velezensis*	0.201	0.207	0.1087	0.1147	0.1117

63-11	*B. infantis*	0.304	0.268	0.234	0.198	0.216
78-1	*B. amyloliquefaciens*	0.161	0.173	0.091	0.103	0.097

4.2.8 Halophilic property assessment

According to Kushner & Kamekura [177] there are to date 18 recognized species of *Bacillus* and related genus that can be categorized as moderate halophiles or halotolerant bacteria. Moderately halophilic bacteria are those organisms that grow best in the presence of 3–15% salt, while halotolerant bacteria are non-halophilic microorganisms that can thrive in the absence of salts as well as tolerate relatively high salt concentrations. The property of *Bacillus* species has a whole range of applications in the aquaculture industry as well as salted food industry. Hence, halophilic property of all 124 isolates of Plara project was determined by Cheunjit Prakitchaiwattana and research group, 2017. Out of those 124 isolates, three were selectively chosen for this study based on their halophilic and inhibitory properties against Staphylococcus spp. Their halophilic properties are mentioned in Table 4.2.8. As per the above definition, both 6-2 (*Bacillus velezensis*) and 78-1 (*B. amyloliquefaciens*) isolates proved to be moderately halophilic in nature whereas isolate 63-11 (*Bacillus infantis*) demonstrated low halophilic property as compared to the other two strains Give that all three strains were isolated from the same source, the results allude that individual isolates have their own specific characterization and henceforth applications in both salted-fermented foods and non-salted foods.

Table 4.2.8 Assessment of halophilic property of Bacillus isolates supplemented with varying concentrations of NaCl.

Isolate code	Isolate Name	% NaCl				
		0	5	10	15	20
6-2	*B. velezensis*	G	G	G	NG	NG
63-11	*B. infantis*	G	RG	NG	NG	NG
78-1	*B. amyloliquefaciens*	G	G	G	NG	NG

4.3 Comparison of data obtained from WSG analysis and In-Vitro assay

The results from WGS analysis and In-vitro tests were compiled and analyzed simultaneously to check the reliability and repeatability of the results obtained (table 4.3). Species identification through 16sRNA sequencing identified the strains as *Bacillus subtilis* (6-2 and 78-1) and *Bacillus infantis* (63-11), which proved to be a laborious and non-specific procedure [178]. Hence, novel strain-specific approach - WGS (based on ANI & DDH values) and phylogeny evolutionary study (TYGS View) was used, that identified the strains as – *Bacillus velezensis* (6-2), *Bacillus infantis* (63-11) and *Bacillus amyloliquefaciens* (78-1). Gram staining and colony morphology also identified the strains as *Bacillus* species and revealed high similarity between isolates 6-2 and 78-1(pale yellow colonies) as compared to 63-11 (reddish-orange colonies). Following species identification, safety assessment of the isolates was conducted.

The isolates were screened for antibiotic and antimicrobial genes and WGS results predicted presence of cfr(B) gene (chloramphenicol, florfenicol, clindamycin and linezolid) and tet(L) -Tetracycline resistance protein in isolates 6-2 and 78-1. Isolate 63-11 exhibited no such genes. Although these genes had been identified, the expression of these genes had not been assayed and hence, not quantified. In-vitro antibiotic activity (against standard antibiotics) and antimicrobial activity (against potentially pathogenic bacteria) was assayed for quantification of WGS results. None of the strains reflected any resistance property to any of the standard antibiotics (ampicillin, chloramphenicol, tetracycline), however, due to insufficient data on *Bacillus* species, the values could not be compared to the MIC values established by CLSI. Antimicrobial resistance of *Bacillus* isolates revealed moderate to high resistance to both gram positive and gram-negative pathogens (*Escherichia coli* ATCC 25922, *Staphylococcus aureus* ATCC 25923 and *Bacillus. cereus* ATCC 6633). Virulence assay though WGS did not identify any cytotoxin K, hemolytic enterotoxin, and non-hemolytic enterotoxin genes in any of the isolates, hence in-vitro cytotoxicity assay on cell line systems were not conducted, but WGS did predict hemolytic III gene in isolate 6-2 and 78-1 and

hemolytic III homolog in isolate 63-11. Since many Gram-negative as well as Gram-positive bacterial species synthesize [147] this gene but may not always express it, genomic annotation results were verified through in-vitro hemolysin activity using 5% human blood agar plates. Isolate 6-2 and 63-11 displayed non-hemolytic activity but isolate 78-1 exhibited β hemolytic activity (clear transparent zone around the colonies). Hemolytic activity is not solely responsible for the overall virulence of the isolates. There are multiple features including bacteriocin and protease production that may give false hemolysis results. Since all isolates were overall predicted to be non-human pathogens, protease activity and secondary metabolites were assayed to test the hypothesis [119]. Exocellular protease genes were found in the genome of 6-2 and 63-11 strain, and cell bound protease genes were found in the genome of 78-1 strain as shown by annotation results of WGS. In-vitro testing for protease activity seconded this discovery by demonstrating highest and second highest protease activity of 63-11 and 6-2 strains, respectively. Strain 78-1 had the lowest protease activity, proving the false hemolytic activity hypothesis wrong. Therefore, secondary metabolite synthesis was checked. WGS results from antiSMASH tool and Pfam (protein) analysis identified bacteriocin producing (Polyketide, NRPS and dipeptides) genes in isolate 6-2 and 78-1. This coincided with the In-vitro antimicrobial resistance results of the two isolates. These SeMe genes could have proven the hypothesis if isolate 6-2 would have displayed hemolytic activity. The lack of hemolytic activity of 6-2 strain proved the hypothesis false, raising doubt over the virulence profile of strain 78-1 (*Bacillus amyloliquefaciens*). It is important that expression of hly gene be analyzed to quantify the hemolytic activity of 78-1 isolate. Isolate 63-11 did not demonstrated any inhibitory metabolite synthesis; It did however display synthesis of non - mevalonate terpenoid mva genes which was proved by the reddish-orange hue of 63-11 colonies.

Following the partial safety assessment evaluation, functional characteristic of the isolates as mentioned in the FDA guidelines was analyzed. In-vitro gastric acidity resistance of all isolates after 6 h of incubation was moderate – high, with isolate 78-

1 showing positive growth under acidic conditions, and bile acid resistance, after 24 h of incubation in the high range with isolate 63-11 showing approximately 100% survivability. This was supported by the annotation results from WGS data which showed F1F0 ATP Acid tolerance and universal stress response proteins in all three isolates. Studies supporting the symbiotic actions of F1F0 ATP genes and DnaK proteins alluded to the positive induced response of bacterial species to bile acid resistance. WGS based physiochemical analysis also identified enzymes necessary for glycolytic and citric acid pathways, further alluding to the strain's survivability in the human GIT.

Adherence to epithelial cells was assayed by analysis hydrophobic activity of the isolates under organic acid (Chloroform, Ethyl acetate, xylene) conditions. Isolates 63-11 and 78-1 showed highest % hydrophobicity (xylene), except for isolate 6-2 (chloroform). The relatively low affinities of 63-11 and 78-1 strains to chloroform when compared to xylene and ethyl acetate, addressed to their aprotic and poor electron donor property, indicating adhesive property of these isolates as compared to 6-2 strain. Genomic annotation also showed EAL domain protein synthesis (key components of biofilm formation processes), flagellin synthesis and putative integral membrane proteins (Lipid transport) synthesis in genomes of 63-11 and 78-1 strain. Genome analysis of strain 6-2 revealed no adhesion (to gut mucosa) proteins.

Overall, the cumulative findings of the tests coincided, giving an in-sight to the possible mechanism of action and the safety and efficacy of the isolated strains; Providing sufficient evidence to support the potential benefits of *Bacillus* species, mainly *Bacillus infantis* as probiotic in food. WGS assessment results on itself were adequate for identifying and characterizing genes responsible for functional properties of potential probiotic. They were however not sufficient for quantifying these potential probiotic properties. Assessment on the case-by-case expression of the predicted genes and the possible mechanism of probiotic interaction with the host under in-vivo model was able to give further insights to the safety, functional and metabolic activity assessment of the *Bacillus* isolates.

Table 4.3 Comparison of data obtained from WSG analysis and In-Vitro assay

Properties	WSG analysis			In-vitro assay		
	6-2	63-11	78-1	6-2	63-11	78-1
Species identification and Phylogeny analysis	16s RNA screening (species specific) – *Bacillus subtilis* WGS (strain specific) - *Bacillus velezensis*; Very high similarity to 78-1 strain	16s RNA and WGS - *Bacillus infantis*; Different subspecies to *Bacillus infantis* NRRL B-14911 strain	16s RNA screening (species specific) – *Bacillus subtilis* WGS (strain specific) - *Bacillus amyloliquefaciens*; Very high similarity to 6-2 strain	Gram Staining -*Bacillus* rods Similar (cell and colony) morphology to 78-1 strain. Pale yellow colonies.	Gram Staining - *Bacillus* rods. Different colony morphology - Reddish-orange colonies	Gram Staining - *Bacillus* rods Similar (cell and colony) morphology to 6-2 strain. Pale yellow colonies
Antibiotic resistance/susceptibility to standard antibiotics	KmerResistance tool - Resistant to cfr(B) gene (chloramphenicol, florfenicol, clindamycin and linezolid) and tet(L) -Tetracycline resistance protein	No antibiotic resistance genes found.	KmerResistance tool - Resistant to cfr(B) gene (chloramphenicol, florfenicol, clindamycin and linezolid) and tet(L) - Tetracycline resistance protein	Moderately susceptible (> 5-10 mm of ZOI) to ampicillin & chloramphenicol. Least susceptible (≤ 5 mm of ZOI) to tetracycline	Highly susceptible (>10-15mm of ZOI) to all three antibiotics.	Least susceptible to ampicillin; Moderately to tetracycline, and high to chloramphenicol
Antimicrobial resistance to pathogenic bacteria/AMR genes		No AMR genes found		Highly resistant (≥15 mm of ZOI) towards E. coli, and moderately resistant to B. cereus and S. aureus	Moderately resistant (10-15 mm of ZOI) to all three pathogens	Moderately resistant (10-15 mm of ZOI) towards E. coli and S. aureus. Highly resistant to B. cereus

Table 4.3 Comparison of data obtained from WSG analysis and In-Vitro assay (Continued)

Properties	WSG analysis			In-vitro assay		
	6-2	63-11	78-1	6-2	63-11	78-1
Toxicity/ Virulent factors	VFanalyser Tool - Hemolysin III gene detected.	VFanalyser Tool - Hemolysin III homolog detected	VFanalyser Tool - Hemolysin III gene detected	Non- hemolytic	Non- hemolytic	β hemolytic (clear transparent zone around the colonies)
Protease production	Detected extracellular protease gene (mpr)	Detected extracellular protease gene (mpr)	Detected cell bound protease gene (wprA)	high protease activity (OD 405nm 0.1117)	highest protease activity (OD 405nm 0.216)	Lowest protease activity (OD 405nm 0.097)
Secondary metabolite production	Presence of - ● Polyketide ● NRPS ● dipeptides gene clusters (Antimicrobial/ inhibitory activity)	Presence of Non-mevalonate terpenoid mva genes	Presence of - ● Polyketide ● NRPS ● dipeptides gene clusters (Antimicrobial/ inhibitory activity)	Coincided with the antimicrobial resistance and protease activity (possible antibiotic/bacteriocin production)	Reddish-orange colonies of endospore; Associated with lycopene production	Coincided with the antimicrobial resistance and protease activity (possible antibiotic/bacteriocin production)
Bile Acid Tolerance	Correlation to F1F0 ATP Acid tolerance and DnaK proteins	Correlation to F1F0 ATP Acid tolerance and DnaK proteins	Correlation to F1F0 ATP Acid tolerance proteins	High tolerance with 95% survivability after 6 h of incubation	High tolerance. Steady growth with 99% survivability after 24 h of incubation	Moderate tolerance with 88% survivability after 6 h of incubation

Table 4.3 Comparison of data obtained from WSG analysis and In-Vitro assay (Continued)

Properties	WSG analysis			In-vitro assay		
	6-2	63-11	78-1	**6-2**	63-11	78-1
Gastric acid Tolerance	Quast and Pfam analysis identified - chaperonin (groEL, groES), Cold shock (CspB) and general stress response protein F1F0 ATP Acid tolerance proteins present.	Quast and Pfam analysis identified - chaperonin (groEL, groES), and universal stress response protein F1F0 ATP Acid tolerance proteins present.	Quast and Pfam analysis identified – heat-shock and general stress response protein F1F0 ATP Acid tolerance proteins present.	Moderate tolerance with 75% survivability after 6 h of incubation	Moderate tolerance with 84% survivability after 6 h of incubation	High tolerance and increase in growth after 6 h of incubation
Hydrophobicity (Adhesion Property) – non-polar activity	No adhesion proteins/genes identified	Quast and Pfam analysis – • EAL domain protein (biofilm) • flagellin synthesis • putative integral membrane proteins (Lipid transport)	Quast and Pfam analysis – • EAL domain protein (biofilm) • flagellin synthesis • putative integral membrane proteins (Lipid transport)	highest % hydrophobicity in ethyl acetate, demonstrating better epithelial cell adherence under polar aprotic conditions	highest % hydrophobicity in xylene; Demonstrating epithelial cell adherence under nonacidic conditions	second higher % hydrophobicity in xylene; demonstrating epithelial cell adherence under onacidic conditions

CHAPTER 5 - CONCLUSION

A series of In-vitro and Computational tests using WGS were conducted for three *Bacillus* strains isolated from Thai fermented fish Plara. The tests were conducted as per the FDA and EFSA guidelines to establish potential probiotic properties of selected *Bacillus* isolates. The isolates were properly characterized at strain-level using WGS technique as *Bacillus velezensis* (6-2), *Bacillus infantis* (63-11) and *Bacillus amyloliquefaciens* (78-1); and the cumulative findings of these tests coincided, giving an in-sight to the possible mechanism of action and the safety and efficacy of the isolated species, as most strains showed moderate-high survival rates under duress (acidic and alkaline conditions), effective adherence to gut epithelial cells, non-virulence and non-pathogenicity along with inhibitory activities against pathogen strains. one of the strains (63-11) also demonstrated additional, multifunctional pigment producing gene(s) that may exhibit additional health benefits to the host.

Based on the results collected from both the tests this study successfully provided sufficient evidence to support the potential benefits of *Bacillus* species, mainly *Bacillus infantis* (63-11) as the most potential food probiotic candidate. Assessment on the case-by-case expression of the predicted genes and the mechanism of probiotic interaction with the host and colonizing microbes in the *in-vivo* model will give further insights to the safety and efficacy of *Bacillus* probiotic strains. Also, further investigation in the commercial applications of the pigment producing gene of *Bacillus infantis* (63-11) and its potential health benefits including potential cholesterol reducing property may further empower the *Bacillus* isolate's probiotic properties.

CHAPTER 6 - FUTURE PROSPECTIVE

Based on the results obtained from this study, *Bacillus infantis* (63-11) has been considered as the most potential probiotic candidate of choice. The applications of this isolate in a food product could essential have serious commercial value. Based on till date studies on probiotic food-based products, majority of them are incorporated in fermented food products. There are not many studies focused on probiotics being incorporated in non-fermented foods, even though studies show that are a much better carrier. Chocolate is one type of non-fermented confectionery food product that is internationally craved and highly consumed across all age groups. It is a semi-solid suspension of fat, sugar and cocoa, which on addition shows high antioxidant activities [179]. Chocolate production is a complicated process consisting of 6 main stages- mixing, refining, conching, tempering, molding and packaging [94]. Studies claim that if the probiotics is added after the conching or tempering stage, it suffers no form of mechanical or heat stress that may negative effects its viability and stability [180]. Furthermore, the cocoa butter in chocolates provides a protective effect from water and Hydrogen ions to probiotic and hence they and more likely to survive 4 times higher than in a milk-containing products [181].

Bacillus infantis (63-11) interestingly produced a reddish-orange pigment in the spore form, due to the presence of Lycopene, Lutein and β-Carotene. Since there are no studies on enrichment of white chocolate with *Bacillus species* or application of pigment producing gene as natural food coloring agent in white chocolate and nor has *Bacillus infantis* been clinically studied for human consumption, this gives the study its novel approach. Moreover, there are multiple health promoting properties associated with multifunctional pigment producing property of *Bacillus* species. These include antioxidation activity and cholesterol reducing property. *Bacillus subtilis* [182], *Bacillus coagulans* [183], *Bacillus indicus, Bacillus firmus, Bacillus altitudinis and Bacillus safensis* species have all been studies for their synthesis of carotenoids and its health benefits associated with antioxidant properties [184, 185] and prevention of degenerative diseases to enhance the immune response in both the humans and

animals. probiotics have also been used as a new alternative to reduce blood cholesterol levels and serve as a useful dietary approach. There is ample scientific evidence on positive impact of probiotic on cholesterol metabolism, for lowering blood cholesterol levels and CVD-related outcomes [56].

REFERENCES

1. Errington, J., *Regulation of endospore formation in Bacillus subtilis.* Nature reviews. Microbiology, 2003. **1**: p. 117-26.
2. Hong, H.A., L.H. Duc, and S.M. Cutting, *The use of bacterial spore formers as probiotics.* FEMS Microbiology Reviews, 2005. **29**(4): p. 813-835.
3. Giri, S.S., V. Sukumaran, and M. Oviya, *Potential probiotic Lactobacillus plantarum VSG3 improves the growth, immunity, and disease resistance of tropical freshwater fish, Labeo rohita.* Fish & Shellfish Immunology, 2013. **34**(2): p. 660-666.
4. Fuller, R., *Fuller R.. Probiotics in man and animal. J Appl Bacteriol 66: 365-378.* The Journal of applied bacteriology, 1989. **66**: p. 365-78.
5. FAO/WHO, et al., *Guidelines for the evaluation of probiotics in food.* Joint FAO/WHO Working Group Report on Drafting Guidelines for the Evaluation of Probiotics in Food, 2002: p. 1-11.
6. Group, J.F.W.W., *Drafting Guidelines for the Evaluation of Probiotics in Food*. Joint FAO/WHO Working Group: London, Ontario, Canada. p. 11.
7. Varzakas, T., et al., *6 - Innovative and fortified food: Probiotics, prebiotics, GMOs, and superfood*, in *Preparation and Processing of Religious and Cultural Foods*, M.E. Ali and N.N.A. Nizar, Editors. 2018, Woodhead Publishing. p. 67-129.
8. Klein, A., et al., *Lactobacillus acidophilus 74-2 and Bifidobacterium animalis subsp lactis DGCC 420 modulate unspecific cellular immune response in healthy adults.* European Journal of Clinical Nutrition, 2008. **62**(5): p. 584-593.
9. Ventura, M., et al., *Genome-scale analyses of health-promoting bacteria: probiogenomics.* Nature Reviews Microbiology, 2009. **7**(1): p. 61-71.
10. Turroni, F., D. van Sinderen, and M. Ventura, *Genomics and ecological overview of the genus Bifidobacterium.* International Journal of Food Microbiology, 2011. **149**(1): p. 37-44.
11. Turroni, F., et al., *Human gut microbiota and bifidobacteria: from composition to functionality.* Antonie van Leeuwenhoek, 2008. **94**(1): p. 35-50.

12. Iole Pitino, C.L.R., Giuseppina Mandalari, Alberto Lo Curto, Richard Martin Faulks, Yvan Le Marc, Carlo Bisignano, Cinzia Caggia, Martin Sean John Wickham, *Survival of Lactobacillus rhamnosus strains in the upper gastrointestinal tract.* Food Microbiology, Volume 27, Issue 8, 2010: p. 1121-1127.
13. Kiran, F., Mokrani, M., Osmanagaoglu, O.,, *Effect of encapsulation on viability of Pediococcus pentosaceus OZF during its passage through the gastrointestinal tract. .* Current Microbiology, 2015: p. 71, 95–105.
14. Sean Farmer, L.J., *Tropical compositions containing probiotic bacillus bacteria, spores, and extracellular products and uses thereof*. 2002, United States Patent CA (US)
15. Schultz, M., J. Burton, and R. Chanyi, *Use of Bacillus in Human Intestinal Probiotic Applications*. 2017. p. 119-123.
16. Floch, M., Y. Ringel, and W.A. Walker, *The Microbiota in Gastrointestinal Pathophysiology: Implications for Human Health, Prebiotics, Probiotics, and Dysbiosis*. 2016. 1-419.
17. Ripert, G., et al., *Secreted Compounds of the Probiotic Bacillus clausii Strain O/C Inhibit the Cytotoxic Effects Induced by Clostridium difficile and Bacillus cereus Toxins.* Antimicrobial Agents and Chemotherapy, 2016. **60**(6): p. 3445-3454.
18. Cutting, S.M., *Bacillus probiotics.* Food Microbiology, 2011. **28**(2): p. 214-220.
19. Nicholson, W.J., Munakata, N., Horneck, G., Melosh, H.J., Setlow, P., *Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. .* Microbiol. Mol. Biol. Rev. , 2000: p. 64, 548-572.
20. X.H. Guo., J.M.K., H.M. Nam., S.Y. Park and J.M. Kim., *Screening lactic acid bacteria from swine origins for multistrain probiotics based on in vitro functional properties. .* Anaerobe 2010: p. 27: 432-7.

21. Spinosa, M.R., Braccini, T., Ricca, E., De Felice, M., Morelli, L., Pozzi, G., Oggioni, M.R., *On the fate of ingested Bacillus spores.* Res. Microbiol., 2000: p. 151, 361-368.

22. C.C. Tsai., P.P.L., Y.M. Hsieh, *Three Lactobacillus strains from healthy infant stool inhibit enterotoxigenic Escherichia coli grown in vitro. .* Anaerobe, 2008: p. 14: 61-7.

23. Stanton, C., et al., *Fermented functional foods based on probiotics and their biogenic metabolites.* Current Opinion in Biotechnology, 2005. **16**(2): p. 198-203.

24. Elshaghabee, F.M.F., et al., *Bacillus As Potential Probiotics: Status, Concerns, and Future Perspectives.* Frontiers in Microbiology, 2017. **8**(1490).

25. Tripathi, M. and S. Giri, *Probiotic functional foods: Survival of probiotics during processing and storage.* Journal of Functional Foods, 2014. **9**: p. 225–241.

26. Barbosa, T.M., et al., *Screening for bacillus isolates in the broiler gastrointestinal tract.* Applied and environmental microbiology, 2005. **71**(2): p. 968-978.

27. Bezkorovainy, A., *Probiotics: determinants of survival and growth in the gut.* The American Journal of Clinical Nutrition, 2001. **73**(2): p. 399s-405s.

28. Graff, S., et al., *Formulations for Protecting the Probiotic Saccharomyces boulardii from Degradation in Acidic Condition.* Biological and Pharmaceutical Bulletin, 2008. **31**(2): p. 266-272.

29. Ljungh, Å., *Lactic Acid Bacteria as Probiotics.* Current issues in intestinal microbiology, 2006. **7**: p. 73-89.

30. Spinosa, M.R., et al., *On the fate of ingested Bacillus spores.* Research in Microbiology, 2000. **151**(5): p. 361-368.

31. Henriques, A.O. and J. Charles P. Moran, *Structure, Assembly, and Function of the Spore Surface Layers.* Annual Review of Microbiology, 2007. **61**(1): p. 555-588.

32. Nicholson, W.L., et al., *Resistance of Bacillus Endospores to Extreme Terrestrial and Extraterrestrial Environments.* Microbiology and Molecular Biology Reviews, 2000. **64**(3): p. 548.

33. Tuohy, K.M., et al., *Survivability of a probiotic Lactobacillus casei in the gastrointestinal tract of healthy human volunteers and its impact on the faecal microflora.* Journal of Applied Microbiology, 2007. **102**(4): p. 1026-1032.

34. Barefoot, S. and C. Nettles, *Antibiosis Revisited: Bacteriocins Produced by Dairy Starter Cultures.* Journal of dairy science, 1993. **76**: p. 2366-79.

35. Klaenhammer, T.R., *Genetics of bacteriocins produced by lactic acid bacteria*.* FEMS Microbiology Reviews, 1993. **12**(1-3): p. 39-85.

36. Moir, A., *How do spores germinate?* Journal of Applied Microbiology, 2006. **101**(3): p. 526-530.

37. Nakano, M.M. and P. Zuber, *ANAEROBIC GROWTH OF A "STRICT AEROBE" (BACILLUS SUBTILIS).* Annual Review of Microbiology, 1998. **52**(1): p. 165-190.

38. Hosoi, T., et al., *Changes in fecal microflora induced by intubation of mice with Bacillus subtilis (natto) spores are dependent upon dietary components.* Canadian journal of microbiology, 1999. **45**(1): p. 59-66.

39. Sanders, M.E., Morelli, L. and Tompkins, T.A., *Spore-formers as human probiotics: Bacillus, Sporolactobacillus, and Brevibacillus. .* Comprehensive Rev. Food Sci. Food Safety 2, 2003: p. 101–110.

40. Kolacek, S., et al., *Commercial Probiotic Products: A Call for Improved Quality Control. A Position Paper by the ESPGHAN Working Group for Probiotics and Prebiotics.* Journal of Pediatric Gastroenterology and Nutrition, 2017. **65**(1): p. 117-124.

41. Cutting, S.M., *Bacillus probiotics.* Food Microbiology, 2011: p. 214-220.

42. Urdaci, M.C., Bressollier, P. and Pinchuk, I., *Bacillus clausii probiotic strains: antimicrobial and immunomodulatory activities.* J. Clin. Gastroenterol., 2004: p. 38, S86–S90.

43. Lawrence, C.a.N., C., *Production of interleukin-12 by murine macrophages in response to bacterial peptidoglycan. .* Infect. Immunol. , 1993: p. 66, 4947–4949.

44. Kosak, T., Maeda, T., Nakada, Y., Yukawa, M. and Tanaka, S., *Effect of Bacillus subtilis spore administration on activation of macrophages and natural killer cells in mice. .* Vet. Microbiol. , 1998: p. 60, 215–225.

45. Meroni, P.L., Palmieri, R., Barcellini, W., De Bartolo, G. and Zanussi, C., *Effect of long-term treatment with B. subtilis on the frequency of urinary tract infections in older patients.* Chemioterapia 1983: p. 2, 142–144.

46. Hosoi, T. and K. Kiuchi, *Production and probiotic effects of natto.* Bacterial Spore Formers: Probiotics and Emerging Applications, 2004: p. 143-154.

47. Huynh A. Hong, L.H.D., Simon M. Cutting, *The use of bacterial spore formers as probiotics.* FEMS Microbiology Reviews 2005: p. 29,813–835.

48. Sumi, H., et al., *A novel fibrinolytic enzyme (nattokinase) in the vegetable cheese Natto: A typical and popular soybean food in the Japanese diet.* Experientia, 1987. **43**: p. 1110-1.

49. Sumi, H., et al., *Effect of Bacillus natto-fermented product (BIOZYME) on blood alcohol, aldehyde concentrations after whisky drinking in human volunteers, and acute toxicity of acetaldehyde in mice.* Aruk**ō**ru kenky**ū** to yakubutsu izon = Japanese journal of alcohol studies & drug dependence, 1995. **30**: p. 69-79.

50. Ferrinho, P. and J. Pereira-Miguel, *The Health Status of the European Union. Narrowing the Health Gap: European Commission, Health and Consumer Protection.* 2003.

51. Cartman, S.T.a.L.R., R.M., *Spore probiotics as animal feed supplements In: Bacterial Spore Formers: Probiotics and Emerging Applications.* Ricca, E., Henriques, A.O. and Cutting, S.M., Eds, 2004: p. 155–161.

52. Verschuere, L., Rombaut, G., Sorgeloos, P. and Verstraete, W., *Probiotic bacteria as biological control agents in aquaculture. .* Mic. Mol. Biol. Rev., 2000 p. 64, 655–671.

53. Gatesoupe, F.J., *The use of probiotics in aquaculture.* Aquaculture, 1999. **180**(1): p. 147-165.

54. Timmermans, L.P.M., *Early development and differentiation in fish.* Sarsia, 1987. **72**(3-4): p. 331-339.

55. Gareau, M.G., P.M. Sherman, and W.A. Walker, *Probiotics and the gut microbiota in intestinal health and disease.* Nature reviews. Gastroenterology & hepatology, 2010. **7**(9): p. 503-514.

56. Kumar, M., et al., *Cholesterol-lowering probiotics as potential biotherapeutics for metabolic diseases.* Experimental diabetes research, 2012. **2012**: p. 902917-902917.

57. Neffe-Skociń**s**ka, K., et al., *Chapter 3 - Trends and Possibilities of the Use of Probiotics in Food Production*, in *Alternative and Replacement Foods*, A.M. Holban and A.M. Grumezescu, Editors. 2018, Academic Press. p. 65-94.

58. Shori, A.B., *Influence of food matrix on the viability of probiotic bacteria: A review based on dairy and non-dairy beverages.* Food Bioscience, 2016. **13**: p. 1-8.

59. Senaka Ranadheera, C., et al., *Probiotic viability and physico-chemical and sensory properties of plain and stirred fruit yogurts made from goat's milk.* Food Chemistry, 2012. **135**(3): p. 1411-1418.

60. Shori, A.B., *The potential applications of probiotics on dairy and non-dairy foods focusing on viability during storage.* Biocatalysis and Agricultural Biotechnology, 2015. **4**(4): p. 423-431.

61. Khan, M., et al., *Meat as a functional food with special reference to probiotic sausages.* Food Research International 2011. **44**: p. 3125-3133.

62. Kandylis, P., et al., *Dairy and non-dairy probiotic beverages.* Current Opinion in Food Science, 2016. **7**: p. 58-63.

63. Lalic˘ic´-Petronijevic´, J., Popov-Raljic´, J., Obradovic´, D., Radulovic´, Z., Paunovic´, D., Petrušic´, M., Pezo, L.,, *Viability of probiotic strains Lactobacillus acidophilus NCFM® and Bifidobacterium lactis HN019 and their impact on sensory and rheological properties of milk and dark chocolates during storage for 180 days.* Journal of Functional Foods, 2015: p. 15, 541–550.

64. Katarzyna Neffe-Skociń**s**ka, A.R., Aleksandra Szydłowska, Danuta Kołoż**y**n-Krajewska, *Chapter 3, Trends and Possibilities of the Use of Probiotics in Food*

Production, in *Alternative and Replacement Foods, Handbook of Food Bioengineering*. 2018, Academic Press. p. 65-94.

65. Tamang, J.P., K. Watanabe, and W.H. Holzapfel, *Review: Diversity of Microorganisms in Global Fermented Foods and Beverages.* Frontiers in microbiology, 2016. **7**: p. 377-377.
66. Steinkraus, K.H., et al., *Handbook of Indigenous Fermented Foods.* XF2006239778, 1997. **62**.
67. Chukeatirote, E., *Thua nao: Thai fermented soybean.* Journal of Ethnic Foods, 2015. **2**(3): p. 115-118.
68. Kimura, K. and S. Yokorama, *Trends in the application of Bacillus in fermented foods.* Current opinion in biotechnology, 2018. **56**: p. 36-42.
69. Mishra, S. and H.N. Mishra, *Technological aspects of probiotic functional food development.* Nutrafoods, 2012. **11**(4): p. 117-130.
70. Laličić-Petronijević, J., *Viability of probiotic strains Lactobacillus acidophilus NCFM® and Bifidobacterium lactis HN019 and their impact on sensory and rheological properties of milk and dark chocolates during storage for 180 days.* Journal of functional foods, 2015. **v. 15**: p. pp. 541-550-2015 v.15.
71. K. Makinen, B.B., R. Bel-Rhlid, E. Ananta, *Science and technology for the mastership of probiotic applications in food products.* Journal of Biotechnology, Volume 162, Issue 4, 2012: p. 356-365.
72. Klu, Y.A. and J. Chen, *Effect of peanut butter matrices on the fate of probiotics during simulated gastrointestinal passage.* LWT - Food Science and Technology, 2015. **62**.
73. Hayes, M., et al., *Cheese as a delivery vehicle for probiotics and biogenic substances.* Australian Journal of Dairy, 2006. **61**: p. 132-141.
74. FAO/WHO, *Guidelines for the evaluation of probiotics in food.* Joint FAO/WHO Working Group Report on Drafting Guidelines for the Evaluation of Probiotics in Food, 2002: p. 1-11.
75. Kolacek, S., et al., *Commercial Probiotic Products: A Call for Improved Quality Control. A Position Paper by the ESPGHAN Working Group for Probiotics and Prebiotics.* Journal of Pediatric Gastroenterology and Nutrition, 2017. **65**(1).

76. Gragnani, M., *The EU regulation 1151/2012 on quality schemes for agricultural products and foodstuffs.* 2013. **8**: p. 376-385.
77. EFSA, *Guidance on the assessment of bacterial susceptibility to antimicrobials of human and veterinary importance.* EFSA Journal, 2012. **10**(6): p. 2740.
78. Pariza, M.W., et al., *Determining the safety of microbial cultures for consumption by humans and animals.* Regulatory Toxicology and Pharmacology, 2015. **73**(1): p. 164-171.
79. Gueimonde, M., et al., *Antibiotic resistance in probiotic bacteria.* Frontiers in Microbiology, 2013. **4**(202).
80. Conlon, M.A. and A.R. Bird, *The impact of diet and lifestyle on gut microbiota and human health.* Nutrients, 2014. **7**(1): p. 17-44.
81. Kolaček, S., et al., *Commercial probiotic products: A call for improved quality control. A Position Paper by the ESPGHAN Working Group for Probiotics and Prebiotics.* Journal of Pediatric Gastroenterology and Nutrition, 2017: p. 1.
82. Mardis, E.R., *Chapter 7 - Whole-Genome Sequencing: New Technologies, Approaches, and Applications*, in *Genomic and Personalized Medicine (Second Edition)*, G.S. Ginsburg and H.F. Willard, Editors. 2013, Academic Press. p. 87-93.
83. Directors, A.B.o., *Points to consider in the clinical application of genomic sequencing.* Genetics in Medicine, 2012. **14**(8): p. 759-761.
84. Borah, T., et al., *Probiotic characterization of indigenous Bacillus velezensis strain DU14 isolated from Apong, a traditionally fermented rice beer of Assam.* Biocatalysis and Agricultural Biotechnology, 2019. **18**: p. 101008.
85. Endres, J.R., et al., *Safety assessment of a proprietary preparation of a novel Probiotic, Bacillus coagulans, as a food ingredient.* Food Chem Toxicol, 2009. **47**(6): p. 1231-8.
86. Li, J., et al., *Effects of potential probiotic Bacillus velezensis K2 on growth, immunity and resistance to Vibrio harveyi infection of hybrid grouper*

(Epinephelus lanceolatus ♂ × E. fuscoguttatus ♀). Fish & Shellfish Immunology, 2019. **93**: p. 1047-1055.

87. Yi, Y., et al., *Probiotic potential of Bacillus velezensis JW: Antimicrobial activity against fish pathogenic bacteria and immune enhancement effects on Carassius auratus.* Fish & Shellfish Immunology, 2018. **78**: p. 322-330.
88. Jeon, H.-L., et al., *Evaluation of probiotic Bacillus subtilis P229 isolated from cheonggukjang and its application in soybean fermentation.* LWT, 2018. **97**: p. 94-99.
89. Kavitha, M., M. Raja, and P. Perumal, *Evaluation of probiotic potential of Bacillus spp. isolated from the digestive tract of freshwater fish Labeo calbasu (Hamilton, 1822).* Aquaculture Reports, 2018. **11**: p. 59-69.
90. Lakshmi, S.G., et al., *Safety assesment of Bacillus clausii UBBC07, a spore forming probiotic.* Toxicology Reports, 2017. **4**: p. 62-71.
91. Lefevre, M., et al., *Safety assessment of Bacillus subtilis CU1 for use as a probiotic in humans.* Regulatory Toxicology and Pharmacology, 2017. **83**: p. 54-65.
92. Manhar, A.K., et al., *In vitro evaluation of celluloytic Bacillus amyloliquefaciens AMS1 isolated from traditional fermented soybean (Churpi) as an animal probiotic.* Research in Veterinary Science, 2015. **99**: p. 149-156.
93. Ramesh, D. and V. Rajendran, *Probiotic assessment of Bacillus infantis isolated from gastrointestinal tract of Labeo rohita.* International Journal of Scientific and Research Publications, 2014. **4**: p. 1-6.
94. Erdem, Ö., et al., *Development of a novel synbiotic dark chocolate enriched with Bacillus indicus HU36, maltodextrin and lemon fiber: Optimization by response surface methodology.* LWT - Food Science and Technology, 2014. **56**(1): p. 187-193.
95. Dashti, A., et al., *Heat Treatment of Bacteria: A Simple Method of DNA Extraction for Molecular Techniques.* Kuwait Medical Journal, 2009. **41**.

96. Mao, D.-P., et al., *Coverage evaluation of universal bacterial primers using the metagenomic datasets.* BMC Microbiology, 2012. **12**(1): p. 66.

97. Seemann, T., *Prokka: rapid prokaryotic genome annotation.* Bioinformatics, 2014. **30**(14): p. 2068-2069.

98. Wick, R.R., et al., *Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads.* PLOS Computational Biology, 2017. **13**(6): p. e1005595.

99. Kim, M., et al., *Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes.* Int J Syst Evol Microbiol, 2014. **64**(Pt 2): p. 346-351.

100. Kaas, R.S., et al., *Solving the problem of comparing whole bacterial genomes across different sequencing platforms.* PloS one, 2014. **9**(8): p. e104984-e104984.

101. Goris, J., et al., *DNA–DNA hybridization values and their relationship to whole-genome sequence similarities.* International Journal of Systematic and Evolutionary Microbiology, 2007. **57**(1): p. 81-91.

102. Grant, J.R. and P. Stothard, *The CGView Server: a comparative genomics tool for circular genomes.* Nucleic acids research, 2008. **36**(Web Server issue): p. W181-W184.

103. Meier-Kolthoff, J.P. and M. Göker, *TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy.* Nature communications, 2019. **10**(1): p. 2182-2182.

104. Alcock, B.P., et al., *CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database.* Nucleic Acids Res, 2020. **48**(D1): p. D517-d525.

105. Cosentino, S., et al., *PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data.* PLOS ONE, 2013. **8**(10): p. e77302.

106. Madden, T., *The BLAST sequence analysis tool.* The NCBI Handbook, 2002.

107. Carver, T.J., et al., *ACT: the Artemis Comparison Tool.* Bioinformatics, 2005. **21**(16): p. 3422-3.

108. Clausen, P.T.L.C., et al., *Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data.* Journal of Antimicrobial Chemotherapy, 2016. **71**(9): p. 2484-2488.

109. Zankari, E., et al., *Identification of acquired antimicrobial resistance genes.* Journal of Antimicrobial Chemotherapy, 2012. **67**(11): p. 2640-2644.

110. Liu, B., et al., *VFDB 2019: a comparative pathogenomic platform with an interactive web interface.* Nucleic Acids Research, 2018. **47**(D1): p. D687-D692.

111. Chen, Y., et al., *High speed BLASTN: an accelerated MegaBLAST search tool.* Nucleic acids research, 2015. **43**(16): p. 7762-7768.

112. Kanehisa, M., Y. Sato, and K. Morishima, *BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences.* J Mol Biol, 2016. **428**(4): p. 726-731.

113. Li, M., et al., *Automated genome mining for natural products.* BMC bioinformatics, 2009. **10**: p. 185.

114. Weber, T., et al., *antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters.* Nucleic Acids Research, 2015. **43**(W1): p. W237-W243.

115. Jeon, E.B., et al., *Characterization of Lactobacillus plantarum Lb41, an isolate from kimchi and its application as a probiotic in cottage cheese.* Food science and biotechnology, 2016. **25**(4): p. 1129-1133.

116. Rosenberg, M., *Microbial adhesion to hydrocarbons: twenty-five years of doing MATH.* FEMS Microbiology Letters, 2006. **262**(2): p. 129-134.

117. CLSI, C.a.L.S.I., *Performance Standards for Antimicrobial Susceptibility Testing; Twenty-Fourth Informational Supplement (M100-S24).* 24th ed. ed. Vol. 34. 2014: Clinical and Laboratory Standards Institute. 230.

118. Sánchez-Porro, C., et al., *Screening and characterization of the protease CP1 produced by the moderately halophilic bacterium Pseudoalteromonas sp. strain CP76.* Extremophiles : life under extreme conditions, 2003. **7**: p. 221-8.

119. Silva, M.F., et al., *Biological and Enzymatic Characterization of Proteases from Crude Venom of the Ant Odontomachus bauri.* Toxins, 2015. **7**(12): p. 5114-5128.

120. Tanasupawat, S., et al., *Lactic acid bacteria isolated from soy sauce mash in Thailand.* J Gen Appl Microbiol, 2002. **48**(4): p. 201-9.

121. Chhetri, V., C. Prakitchaiwattana, and S. Settachaimongkon, *A potential protective culture; halophilic Bacillus isolates with bacteriocin encoding gene against Staphylococcus aureus in salt added foods.* Food Control, 2019. **104**: p. 292-299.

122. Ranjan, R., et al., *Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing.* Biochem Biophys Res Commun, 2016. **469**(4): p. 967-77.

123. Stubbs, L., *Contig*, in *Encyclopedia of Genetics*, S. Brenner and J.H. Miller, Editors. 2001, Academic Press: New York. p. 463.

124. Smits, T.H.M., *The importance of genome sequence quality to microbial comparative genomics.* BMC Genomics, 2019. **20**(1): p. 662.

125. Choudhuri, S., *Chapter 7 - Additional Bioinformatic Analyses Involving Nucleic-Acid Sequences**The opinions expressed in this chapter are the author's own and they do not necessarily reflect the opinions of the FDA, the DHHS, or the Federal Government*, in *Bioinformatics for Beginners*, S. Choudhuri, Editor. 2014, Academic Press: Oxford. p. 157-181.

126. Meier-Kolthoff, J.P., H.-P. Klenk, and M. Göker, *Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age.* International Journal of Systematic and Evolutionary Microbiology, 2014. **64**(Pt_2): p. 352-356.

127. Dunlap, C.A., et al., *Bacillus velezensis is not a later heterotypic synonym of Bacillus amyloliquefaciens; Bacillus methylotrophicus, Bacillus amyloliquefaciens subsp. plantarum and 'Bacillus oryzicola' are later heterotypic synonyms of Bacillus velezensis based on phylogenomics.* International Journal of Systematic and Evolutionary Microbiology, 2016. **66**(3): p. 1212-1217.

128. Fritze, D., *Taxonomy of the Genus Bacillus and Related Genera: The Aerobic Endospore-Forming Bacteria.* Phytopathology™, 2004. **94**(11): p. 1245-1248.

129. Ondov, B.D., et al., *Mash: fast genome and metagenome distance estimation using MinHash.* Genome Biology, 2016. **17**(1): p. 132.

130. Lagesen, K., et al., *RNAmmer: consistent and rapid annotation of ribosomal RNA genes.* Nucleic Acids Research, 2007. **35**(9): p. 3100-3108.

131. Camacho, C., et al., *BLAST+: architecture and applications.* BMC Bioinformatics, 2009. **10**(1): p. 421.

132. Meier-Kolthoff, J.P., et al., *Genome sequence-based species delimitation with confidence intervals and improved distance functions.* BMC Bioinformatics, 2013. **14**(1): p. 60.

133. WHO *Antimicrobial resistance*. 2018.

134. Shankar, P.R., *Antimicrobial Resistance: Global Report on Surveillance.* Australasian Medical Journal, 2014. **7**: p. 237.

135. Brüssow, H., C. Canchaya, and W.D. Hardt, *Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion.* Microbiol Mol Biol Rev, 2004. **68**(3): p. 560-602, table of contents.

136. Filée, J., P. Forterre, and J. Laurent, *The role played by viruses in the evolution of their hosts: a view based on informational protein phylogenies.* Res Microbiol, 2003. **154**(4): p. 237-43.

137. Chen, J. and R.P. Novick, *Phage-mediated intergeneric transfer of toxin genes.* Science, 2009. **323**(5910): p. 139-41.

138. Colomer-Lluch, M., J. Jofre, and M. Muniesa, *Antibiotic resistance genes in the bacteriophage DNA fraction of environmental samples.* PLoS One, 2011. **6**(3): p. e17549.

139. Modi, S.R., et al., *Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome.* Nature, 2013. **499**(7457): p. 219-22.

140. Doi, Y., J.-I. Wachino, and Y. Arakawa, *Aminoglycoside Resistance: The Emergence of Acquired 16S Ribosomal RNA Methyltransferases.* Infectious disease clinics of North America, 2016. **30**(2): p. 523-537.

141. Bozdogan, B., et al., *Chromosomal aadD2 encodes an aminoglycoside nucleotidyltransferase in Bacillus clausii.* Antimicrobial agents and chemotherapy, 2003. **47**(4): p. 1343-1346.

142. Galopin, S., V. Cattoir, and R. Leclercq, *A chromosomal chloramphenicol acetyltransferase determinant from a probiotic strain of Bacillus clausii.* FEMS Microbiology Letters, 2009. **296**(2): p. 185-189.

143. Kehrenberg, C., et al., *A new mechanism for chloramphenicol, florfenicol and clindamycin resistance: methylation of 23S ribosomal RNA at A2503.* Mol Microbiol, 2005. **57**(4): p. 1064-73.

144. Teo, J.W.P., T.M.C. Tan, and C.L. Poh, *Genetic determinants of tetracycline resistance in Vibrio harveyi.* Antimicrobial agents and chemotherapy, 2002. **46**(4): p. 1038-1045.

145. Agafonov, D.E., V.A. Kolb, and A.S. Spirin, *Proteins on ribosome surface: Measurements of protein exposure by hot tritium bombardment technique.* Proceedings of the National Academy of Sciences, 1997. **94**(24): p. 12892-12897.

146. Mueller, F. and R. Brimacombe, *A new model for the three-dimensional folding of Escherichia coli 16 S ribosomal RNA. II†. The RNA-protein interaction data2†Paper I in this series is an accompanying paper, Mueller & Brimacombe (1997).1Edited by D. E. Draper.* Journal of Molecular Biology, 1997. **271**(4): p. 545-565.

147. Goebel, W., T. Chakraborty, and J. Kreft, *Bacterial Hemolysins as Virulence Factors.* Antonie van Leeuwenhoek, 1988. **54**: p. 453-63.

148. Baida, G.E. and N.P. Kuzmin, *Mechanism of action of hemolysin III from Bacillus cereus.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 1996. **1284**(2): p. 122-124.

149. Hacker, J., et al., *Influence of cloned Escherichia coli hemolysin genes, S-fimbriae and serum resistance on pathogenicity in different animal models.* Microbial Pathogenesis, 1986. **1**(6): p. 533-547.

150. Wheeler, D. and M. Bhagwat, *BLAST QuickStart: example-driven web-based BLAST tutorial*. 2007.

151. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Res, 2000. **28**(1): p. 27-30.

152. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic acids research, 2000. **28**(1): p. 27-30.

153. Pereira, J.Q., et al., *Functional genome annotation depicts probiotic properties of Bacillus velezensis FTC01.* Gene, 2019. **713**: p. 143971.

154. Cai, X.-C., et al., *Genomic and metabolic traits endow Bacillus velezensis CC09 with a potential biocontrol agent in control of wheat powdery mildew disease.* Microbiological Research, 2017. **196**: p. 89-94.

155. Zhao, P., et al., *Bacillus amyloliquefaciens Q-426 as a potential biocontrol agent against Fusarium oxysporum f. sp. spinaciae.* Journal of Basic Microbiology, 2014. **54**(5): p. 448-456.

156. Seck, E.H., et al., *Halophilic & halotolerant prokaryotes in humans.* Future Microbiol, 2018. **13**: p. 799-812.

157. Schilling, O., et al., *Transcriptional and metabolic responses of Bacillus subtilis to the availability of organic acids: transcription regulation is important but not sufficient to account for metabolic adaptation.* Applied and environmental microbiology, 2007. **73**(2): p. 499-507.

158. Ogata, H., et al., *Computation with the KEGG pathway database.* Bio Systems, 1998. **47**: p. 119-28.

159. Smith, T. and M. Waterman, *Identification of Common Molecular Subsequences.* Journal of molecular biology, 1981. **147**: p. 195-7.

160. Ko, K.S., et al., *Bacillus infantis sp nov and Bacillus idriensis sp nov., isolated from a patient with neonatal sepsis.* International journal of systematic and evolutionary microbiology, 2006. **56**: p. 2541-4.

161. Kanehisa, M., *Enzyme Annotation and Metabolic Reconstruction Using KEGG.* Methods Mol Biol, 2017. **1611**: p. 135-145.

162. Ruiz, L., A. Margolles, and B. Sánchez, *Bile resistance mechanisms in Lactobacillus and Bifidobacterium.* Frontiers in microbiology, 2013. **4**: p. 396-396.

163. Alcántara, C. and M. Zúñiga, *Proteomic and transcriptomic analysis of the response to bile stress of Lactobacillus casei BL23.* Microbiology, 2012. **158**(Pt 5): p. 1206-1218.

164. Hamon, E., et al., *Comparative proteomic analysis of Lactobacillus plantarum for the identification of key proteins in bile tolerance.* BMC Microbiol, 2011. **11**: p. 63.

165. Pitino, I., et al., *Survival of Lactobacillus rhamnosus strains in the upper gastrointestinal tract.* Food Microbiol, 2010. **27**(8): p. 1121-7.

166. Minekus, M., et al., *A standardised static in vitro digestion method suitable for food - an international consensus.* Food Funct, 2014. **5**(6): p. 1113-24.

167. Vandenplas, Y., G. Huys, and G. Daube, *Probiotics: an update.* J Pediatr (Rio J), 2015. **91**(1): p. 6-21.

168. Wang, Y., et al., *Extracellular products from virulent strain of Edwardsiella tarda stimulate mouse macrophages (RAW264.7) to produce nitric oxide (NO) and tumor necrosis factor (TNF)-α.* Fish & Shellfish Immunology, 2010. **29**(5): p. 778-785.

169. Vesper, S.J. and M. Jo Vesper, *Possible Role of Fungal Hemolysins in Sick Building Syndrome*, in *Advances in Applied Microbiology*. 2004, Academic Press. p. 191-213.

170. Koronakis, V. and C. Hughes, *CHAPTER 13 - Hemolysin*, in *Escherichia Coli*, M.S. Donnenberg, Editor. 2002, Academic Press: San Diego. p. 361-378.

171. Ramarao, N., D. Lereclus, and A. Sorokin, *Chapter 59 - The Bacillus cereus Group*, in *Molecular Medical Microbiology (Second Edition)*, Y.-W. Tang, et al., Editors. 2015, Academic Press: Boston. p. 1041-1078.

172. Lindbäck, T. and P.E. Granum, *29 - Bacillus cereus phospholipases, enterotoxins, and other hemolysins*, in *The Comprehensive Sourcebook of Bacterial Protein Toxins (Fourth Edition)*, J. Alouf, D. Ladant, and M.R. Popoff, Editors. 2015, Academic Press: Boston. p. 839-857.

173. Baida, G.E. and N.P. Kuzmin, *Mechanism of action of hemolysin III from Bacillus cereus.* Biochim Biophys Acta, 1996. **1284**(2): p. 122-4.

174. Saggu, S.K. and P.C. Mishra, *Characterization of thermostable alkaline proteases from Bacillus infantis SKS1 isolated from garden soil.* PLoS One, 2017. **12**(11): p. e0188724.

175. Nascimento, W. and M. Martins, *Studies on the stability of protease from Bacillus sp. and its compatibility with commercial detergent.* Brazilian Journal of Microbiology - BRAZ J MICROBIOL, 2006. **37**.

176. Pakshir, K., et al., *Proteolytic activity and cooperative hemolytic effect of dermatophytes with different species of bacteria.* Current medical mycology, 2016. **2**(4): p. 9-14.

177. Kushner, D.J. and M. Kamekura, *Physiology of halophilic eubacteria.* Halophilic Bact, 1988. **1**: p. 109-140.

178. Chanprasartsuk, O.-o., C. Prakitchaiwattana, and R. Sanguandeekul, *Comparison of Methods for Identification of Yeasts Isolated during Spontaneous Fermentation of Freshly Crushed Pineapple Juices.* Journal of Agricultural Science and Technology, 2013. **15**: p. 1479-1490.

179. Vanzani, P., et al., *Efficiency and capacity of antioxidant rich foods in trapping peroxyl radicals: A full evaluation of radical scavenging activity.* Food Research International, 2011. **44**: p. 269-275.

180. Konar, N., et al., *Improving functionality of chocolate: A review on probiotic, prebiotic, and/or synbiotic characteristics.* Trends in Food Science & Technology, 2016. **49**: p. 35-44.

181. Possemiers, S., et al., *Bacteria and chocolate: A successful combination for probiotic delivery.* International Journal of Food Microbiology, 2010. **141**(1): p. 97-103.

182. Xue, J. and B. Ahring, *Enhancing Isoprene Production by Genetic Modification of the 1-Deoxy-D-Xylulose-5-Phosphate Pathway in Bacillus subtilis.* Applied and environmental microbiology, 2011. **77**: p. 2399-405.

183. Majeed, M., et al., *Evaluation of the in vitro cholesterol-lowering activity of the probiotic strain Bacillus coagulans MTCC 5856.* International Journal of Food Science & Technology, 2018.

184. Dawoud, T.M., et al., *Characterization and antifungal activity of the yellow pigment produced by a Bacillus sp. DBS4 isolated from the lichen Dirinaria agealita.* Saudi Journal of Biological Sciences, 2020. **27**(5): p. 1403-1411.

185. Khaneja, R., et al., *Carotenoids found in Bacillus.* J Appl Microbiol, 2010. **108**(6): p. 1889-902.

VITA

NAME	Gauri Khullar
DATE OF BIRTH	17 February 1996
PLACE OF BIRTH	Faridabad, Haryana, India
INSTITUTIONS ATTENDED	Chulalongkorn University, Bangkok, Thailand (Master of Science in Food Science and Technology) National Institute of Technology, Durgapur, India (Bachelor of Technology in Biotechnology) Garden International School, Rayong, Thailand (International Baccalaureate & IGCSE)
HOME ADDRESS	C 4-2, Eastern Star Properties, Eastern Seaboard, Ban Chang, Rayong, Thailand, 21130.
PUBLICATION	Food Innovation Asia Conference (FIAC) 2020 Proceedings - Bioinformatic Tools and in-vitro analysis for screening of potential Bacillus probiotic Publication used as a food supplement
AWARD RECEIVED	First place achieved in the Oral competition, Food Innovation Asia Conference (FIAC), 2020 Completion of Biosafety and Biosecurity Training, at Faculty of Tropical Medicine, Mahidol University (2020) Certificate of completion of Chemical Safety Training for Research Students and Researchers, at SHECU (2019) Certificate of Oral Presentation - Benefits of Bacillus probiotics and application as potential novel food probiotic, at the 6th Asia-Pacific Probiotics Workshop (APPW 2019) Education Scholarship for Master Program at Chulalonkorn University (2018)