JMBFS

Journal of Microbiology, Biotechnology and Food Sciences

International peer-reviewed scientific online journal

Published by Faculty of Biotechnology and Food Sciences

A COMPUTATIONAL GENOME ANALYSIS OF STRAIN *BACILLUS SUBTILIS* MIZ-8 ISOLATED FROM BEKANG REVEALS A DISTINCT CHROMOSOME AND PLASMID CONFERRING SELECTIVE ADVANTAGE

Harini Keerthana Suresh Kumar[1], Trishala Gopikrishna[2], Tapasya K[3], Kumar Perumal[4], Elavarashi Elangovan[5]*

Address(es): Dr. Elavarashi Elangovan – Corresponding author
[1] Universitätsklinikum Jena - Faculty of Medicine, Friedrich Schiller Universität Jena, Department of Molecular Medicine, Germany.
[2] Sri Ramachandra Institute of Higher Education and Research (SRIHER), Faculty of Biomedical Sciences & Technology, Department of Biotechnology, Chennai, India.
[3] Sri Ramachandra Institute of Higher Education and Research (SRIHER), Faculty of Biomedical Sciences & Technology, Department of Biotechnology, Chennai, India.
[4] Sri Ramachandra Institute of Higher Education and Research (SRIHER), Faculty of Biomedical Sciences & Technology, Department of Biotechnology, Chennai, India.
[5] Sri Ramachandra Institute of Higher Education and Research (SRIHER), Faculty of Biomedical Sciences & Technology, Department of Biotechnology, Chennai, India.

*Corresponding author: elavarashi@sriramachandra.edu.in

https://doi.org/10.55251/jmbfs.9141

ARTICLE INFO

Received 28. 6. 2022
Revised 16. 6. 2023
Accepted 26. 6. 2023
Published 1. 10. 2023

Regular article

OPEN ACCESS

ABSTRACT

Bacillus subtilis Miz-8 is isolated from Bekang, a traditional fermented soybean food of Mizoram, India. Here, we report computationally analyzed strain Miz-8 mapped to the annotated genome of the reference strain *Bacillus subtilis* subsp. natto BEST195. Miz-8 has a circular chromosome and a small plasmid with genome size of 4105264 bp and 5838 bp, respectively. Genome contains the genes responsible for synthesis of biosynthetic metabolites like surfactin, fengycin, bacillibactin, bacilysin, subtilosin A and alpha amylase production. The strain Miz-8 harbored virulence genes identical to the strain BEST195 render themselves harmless for human consumption. However, Strain Miz-8 has one intact prophage region but no integrase protein, making it incapable of lateral transfer of antimicrobials. Antibiotic resistance genes were predicted among which *tmrB* gene was on perfect hit. Plasmid of strain Miz-8 contains no prophage sequences and antibiotic resistance genes. Furthermore, there were several single nucleotide polymorphisms and 344 insertion-deletion polymorphisms. Bacillus subtilis Miz-8's genomic information unmasks its functional significance on human health.

Keywords: *Bacillus subtilis*, Bekang, fermented food, Whole genome sequencing

INTRODUCTION

"Bekang " the traditional fermented soybean food (FSF) of Mizoram, India, has been the staple diet of people in the Eastern states of India and globally consumed in East and Southeast Asian countries **(Tamang, 2020).** Numerous beneficial microorganisms present in these FS foods confer diverse properties to these foods **(Frias *et al.*, 2017)**. Soybeans are protein rich crops subjected to enzymatic conversion with predominant microorganisms like *Bacillus* species under appropriate environmental conditions like temperature, pH, and moisture content. The ethnic fermented soybean is a cost-effective source of plant protein with low fat and increased nutritive content **(Tamang, 2015)**. Several previous studies have reported that dominant microorganisms like *B. amyloliquefaciens, B. circulans, B. licheniformis, B. sphaericus, B. subtilis, and B. thuringiensis* present in FSF are involved in the process of fermentation and also produce several beneficial compounds to improve human health, such as isoflavones, lipopeptides, protein hydrolysates, and enzymes **(Beltran-Gracia *et al.*, 2017)**. Among them, *B. subtilis* is the most common microorganism present in these foods which are approved as GRAS (Generally Regarded as Safe) organisms by the Food and Drug Administration (FDA). Due to these beneficial properties, certain members of genus *Bacillus* have extensively been studied as a probiotic microorganism **(Sudan *et al.*, 2021)**

Although microbial composition in fermented food products can be thoroughly studied yet it cannot be just limited to approaches like culture-based methods, Polymerase Chain Reaction (PCR) based amplification and sequencing of 16S rRNA gene which have demonstrated to be useful for the microbiological investigation of many fermented foods. Present day advancements in Next Generation Sequencing (NGS) technologies have provided extensive knowledge about microbial diversity **(Deka *et al.*, 2021)**. Therefore, NGS technologies can be used as a platform for performing comprehensive microbial analysis at the genomic level, substantially enhancing our knowledge of the bacterial diversity present in these fermented products.

Bacillus sp., the predominant microorganisms present in FSF have shown a myriad of applications in several industries. Bioactive compounds in these FSF produced during fermentation by the *Bacillus* sp., have been reported to exhibit antioxidant, anti-tumor, anti-diabetic, antimicrobial, anti-hypertensive, anti-allergic, and immunoregulatory properties **(Sanjukta & Rai, 2016)**. To our knowledge, *Bacillus subtilis* present in fermented foods like bekang have not been much investigated using computational analysis. Therefore, the objective of the present study was to acquire genomic insight of the whole genome of *B.subtilis* strain Miz-8 isolated from bekang, based on their secondary metabolites production, occurrence of antibiotic resistant genes and virulence genes, identification of prophage clusters, SNPs and Indels, which imparts a basis for understanding the beneficial and deleterious impact on human health. A special emphasis has been made to understand the relationship between amylase production by *Bacillus subtilis* Miz-8 and digestive system.

MATERIALS AND METHODS

Sample collection and isolation

A sample of homemade fermented soybean food 'bekang' was collected from Mizoram, an Eastern state of India as depicted in figure1. The sample was wrapped in banana leaf and transported to the laboratory for processing. *Bacillus* spores are highly resistant to heat and so germination can be activated by heating to produce vegetative cells on nutrient agar. Ten grams of fresh bekang (fermented soybean food) was added to 90ml of 0.85% (w/v) sterile physiological saline (NaCl) from Jedux Parental Pvt.Ltd and homogenized in a lab-blender for 10 sec. The homogeneous mixture was kept in a water bath at boiling temperature for 2 mins and then a serial dilution of 10^{-1} to 10^{-7} was performed. The strain was isolated by plating on nutrient agar (Himedia, India) and incubated at 37°C for 24 hrs **(Chettri & Tamang, 2015)**. Based on the colony morphology, the isolate was selected from the low titre plates and stored at room temperature. Subsequently, *Bacillus subtilis* Miz-8 strain was maintained in 25% glycerol at -20^0 C.

Figure 1 "Bekang" traditional fermented soybean food of Mizoram, India

Genome sequencing, assembling and annotating

Bacillus subtilis Miz-8 was subjected to complete genome sequencing using Illumina HI-SEQ X sequencing technology at AgriGenome, Kerala. Genomic DNA was extracted using SDS from overnight stored Luria Bertani broth. DNA concentration was checked by Qubit, nanodrop and agarose gel electrophoresis. Genome library was prepared by following agilent tapestation method using NEBNext Ultra DNA Library Prep kit. Pair-end read orientation was obtained as R1 - 3,158,943 and R2- 3,158,943 with a mean read length of 150 bp. Besides, raw reads (FASTQ) of sequenced genome were subjected to pre-processing in which low quality reads were trimmed using AdapterRemoval v2. Genome assembly was performed using Burrows-Wheeler Aligner (BWA) MEM tool version 0.7.17 alignment algorithm and duplicates were removed by sambamba version 0.8.0. Genome annotation was obtained by NCBI Prokaryotic Genome Annotation Pipeline (PGAP) http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html, Pathosystems Resource Integration Center (PATRIC) https://www.patricbrc.org/. The reads were mapped to the annotated genome of the strain *Bacillus subtilis* subsp. nattoBEST195 (GenBank ID: NC_017196.2). Complete genome of *Bacillus subtilis* Miz-8 consists of a chromosome and a small plasmid which were uploaded in NCBI (GenBank ID: CP076731.1, CP076732.1), respectively. Annotated genome features were visualized by Artemis and circular graphical maps of genome and plasmid were drawn by DNAplotter **(Carver *et al*., 2009)**. For genome based prokaryote taxonomy and detailed identification of strain, Type (Strain) Genome Server (TYGS) was performed (https://tygs.dsmz.de) **(Meier-Kolthoff & Göker, 2019).**

Multilocus Sequence Typing (MLST) Analysis

Multilocus Sequence Typing of complete genome *Bacillus subtilis* Miz-8 targeting seven housekeeping genes such as *glpF, ilvD, pta, purH, pycA, rpoD, tpiA*andthe sequence type of strain *Bacillus subtilis* Miz-8 were identified by MLST database version 2.0 (https://cge.cbs.dtu.dk/services/MLST/) of *Bacillus subtilis* configuration.

Mining of secondary metabolites, antibiotic resistance genes, virulence genes, prophage clusters, Single Nucleotide Polymorphism (SNP), insertion-deletion polymorphism (Indel)

Genome mining of secondary metabolites in *Bacillus subtilis* Miz-8 was analyzed using antiSMASH version 6.0.1 http://antisMash.secondarymetabolites.org **(Blin *et al*., 2021)** and Joint Genome Institute Integrated Microbial Genomes Microbiomes (JGI IMG) https://img.jgi.doe.gov/ was utilized to analyze various other secondary metabolites using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway **(Markowitz *et al*., 2014)**. Surveillance of antibiotic resistance genes in *Bacillus subtilis* Miz-8 was determined using Comprehensive Antibiotic Resistance Database (CARD) https://card.mcmaster.ca/**(Alcock *et al*., 2020)**. Comparative pathogenomic analysis of virulence genes in *Bacillus subtilis* Miz-8 were compared with other *Bacillus* strains using Virulence Factor Database (VFDB) http://www.mgc.ac.cn/VFs/main.htm **(Liu *et al*., 2019)**. Existence of prophage regions was achieved using PHAge Search Tool Enhanced Release (PHASTER) https://phaster.ca/**(Arndt *et al*., 2016)**. BCF tools version 1.11 HTSlib were utilized to manipulate variant calls of *Bacillus subtilis* Miz-8, in comparison with the reference genome *Bacillus subtilis* subsp. natto BEST195 to identify Single Nucleotide Polymorphisms (SNP) and insertion-deletion polymorphisms (Indels). Genetic variant annotation and prediction of functional genes and proteins were analyzed by SnpEff.

Production of alpha amylase

Bacillus subtilis Miz-8 was cultured in Tryptic Soy Broth (TSB). After centrifugation, the pellet was resuspended in sterile distilled water and adjusted to 0.1 OD. The suspension was placed on spot -on - lawn assay using starch agar to determine the *α*-amylase activity. After incubation at 37ºC for 24h, for visualization of the clear zone, the plates were stained with 5 ml of Gram's iodine and the relative enzyme activity (REA) was calculated **(Latorre *et al*., 2016)**. Genes vary in sizes and function, and they are involved in many cellular pathways via their specific protein. Protein coding genes involved in alpha amylase production connected to the KEGG pathway were analyzed using the JGI IMG database **(Markowitz *et al*., 2011).**

RESULTS AND DISCUSSION

Major characteristics of complete genome *Bacillus subtilis* Miz-8

Bacillus subtilis strain Miz-8 has a circular chromosome and a plasmid with genome size 4105264 bp (GenBank ID: CP076731.1, Genome map) and 5838 bp (GenBank ID: CP076732.1, Plasmid genome map), respectively are shown in figure 2. The genome contains an average GC content of 43.50%, 4020 protein coding genes, 87 tRNA genes, 10 rRNA loci, 5 ncRNA and 207 pseudogenes. The phylogenetic taxonomy of the *Bacillus subtilis* strain Miz-8 was obtained using TYGS database. The submitted genomic sequences were clustered with the closely related species and subspecies which are available in the database. Based on species and subspecies clusters, percentage of G+C content, delta statistics, genome size, protein count and Small Subunit (SSU) length, revealed that the query sequence belongs to *Bacillus subtilis* subsp.*subtilis*, is depicted in figure 3. Multilocus Sequence Typing of *B. subtilis* Miz-8 targeting seven housekeeping genes such as *glpF, ilvD, pta, purH, pycA, rpoD*and*tpiA* uncovers that the strain belongs to Sequence Type 208 (ST-208), since the strain Miz-8 contains glpF_1, ilvD_1, pta_2, purH_6, pycA_4, rpoD_3, tpiA_1.

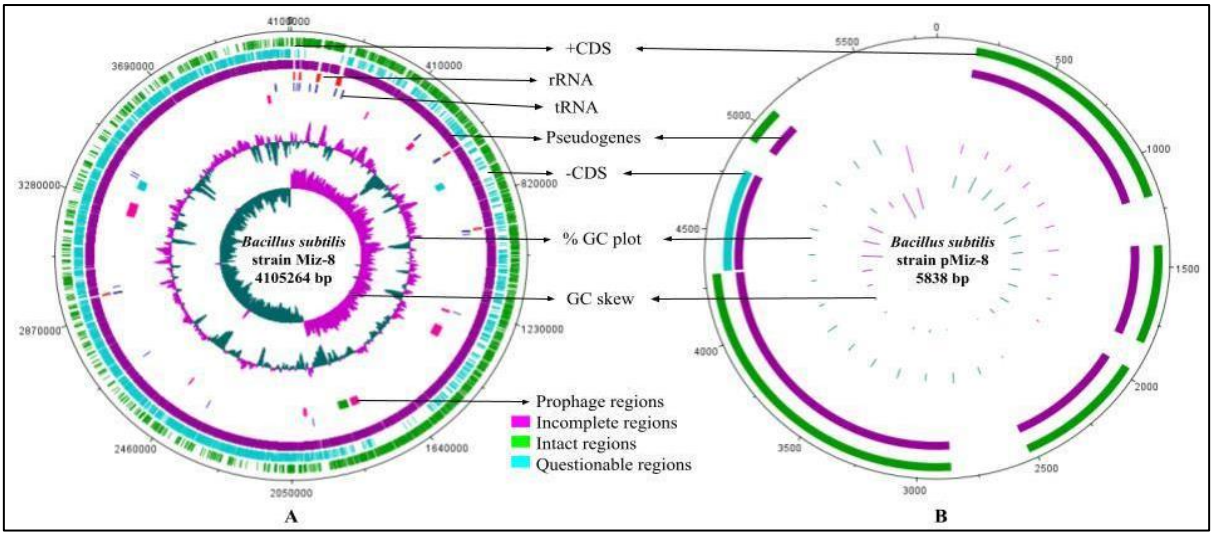


Figure 2 Circular representation of *Bacillus subtilis* Miz-8. (A) chromosome DNA labeling from outer to inner rings indicates forward CDS (green), reverse CDS (blue), pseudogenes (violet), rRNA (red in chromosome), tRNA (deep blue in chromosome), prophage regions in chromosome, GC content, GC skew (B) *Bacillus subtilis* plasmid pMiz-8

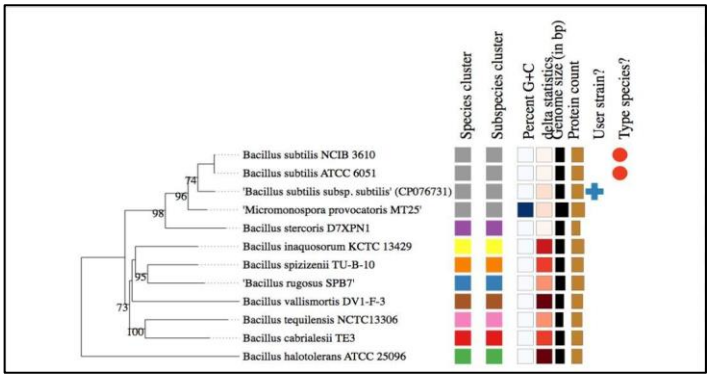


Figure 3 Phylogenetic taxonomy of *Bacillus subtilis* subsp. *subtilis* Miz-8 (CP076731)

Bioinformatic Analysis of AMR/virulence genes and toxin related genes

PHASTER, a bioinformatic investigation was used to analyze the transferability of antimicrobial resistance or virulence factors (AMR/VF) or toxin genes within the genome and plasmid of strain *Bacillus subtilis* Miz-8. One putative intact prophage (region 7), two questionable and ten incomplete phage sequences were identified in the genome of *Bacillus subtilis* Miz-8 (figure 4A), suggesting the possibility of active lateral gene transfer by phage infection. However, the prophage sequence of region 7 lacks an integrase protein, a key protein for prophage induction (figure 4B) **(Balasubramanian *et al*., 2019)**. The questionable and incomplete phage sequences were mostly codes for hypothetical proteins, phage capsid, transposase, and tail. Fortunately, the plasmid of strain Miz-8 contains no prophage sequence. Based on the outcome of bioinformatic analysis, we believe that the prophage sequences identified in the genome of *Bacillus subtilis* Miz-8 may be not inducible and hence is safe from the transferability of AMR/virulence genes or toxin genes by a prophage.

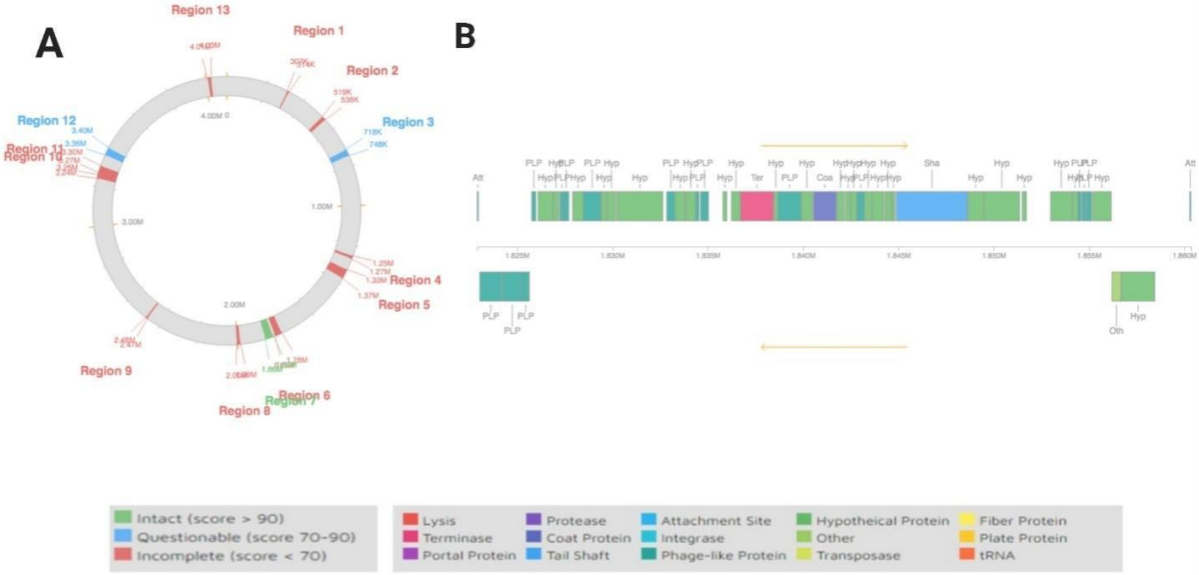


Figure 4 Bioinformatic analysis of prophage sequences in the genome *Bacillus subtilis* Miz-8. (A) the locations of putative prophage regions identified in the chromosome of *Bacillus subtilis* Miz-8 (B) the genomic map of putative prophage region 7 in panel A.

Prediction of secondary metabolites

The genes involved in the biosynthesis of antimicrobials were predicted using antiSMASH. *B. subtilis* Miz-8 genome revealed nine putative gene clusters responsible for antimicrobial metabolites biosynthesis out of which only five clusters showed similarity to some known secondary metabolites. Among them, three genes encode non-ribosomal peptides (NRPS) like surfactin, fengycin, bacillibactin and other two were subtilosin A and bacilysin. These gene clusters identified can suppress the growth of pathogens due to their biosynthesis of antimicrobial peptides. Tab 1 summarizes the antimicrobial gene clusters in the genome of *B. subtilis*Miz-8 which has similarity to certain genes producing secondary metabolites. The JGI genomic database was used to find the overall protein coding genes of *Bacillus subtilis* strain Miz-8 connected to the KEGG pathway as represented in figure 5. The KEGG pathway indicates that the strain Miz-8 is capable of interacting with various physiological and metabolic pathways. The results showed that the genome strain Miz-8 has few other genes involved in the biosynthesis of secondary metabolites, indicating that the strain may have the potential to synthesize a variety of lipopeptide substances as depicted in Tab 2.

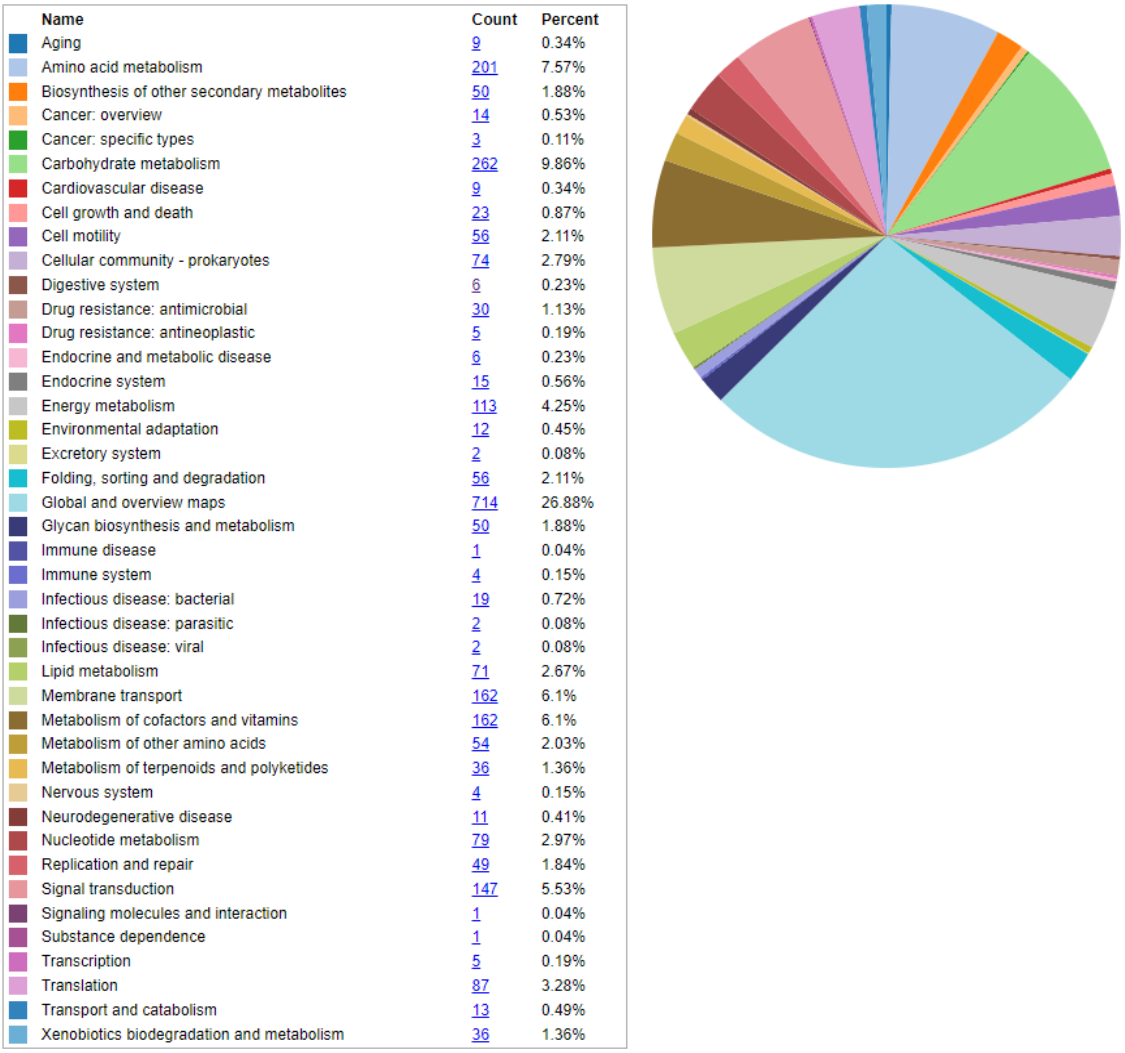
Table 1 Summary of antimicrobial gene clusters in *B. subtilis* Miz-8 using antiSMASH

Region	Predicted biosynthetic metabolite	Position within contig (nucleotide range)	% Similarity (known cluster)
Region 1	NRPS	355,574-418,358	Surfactin biosynthetic gene cluster (82% of genes show similarity)
Region 3	NRPS (Beta Lactone)	1,930,475-1,981,109	Fengycin biosynthetic gene cluster (93% of genes show similarity)
Region 6	NRPS	3,021,955-3,069,091	Bacillibactin biosynthetic gene cluster (100% of genes show similarity)
Region 7	Sactipeptide	3,697,704-3,719,315	Subtilosin A biosynthetic gene cluster (100% of genes show similarity)
Region 8	other	3,722,620-3,764,038	Bacilysin biosynthetic gene cluster (100% of genes show similarity)

**NRPS* non-ribosomal peptide synthetase

Table 2 Overview of other secondary metabolite gene clusters involved in *B. subtilis* Miz-8 using KEGG pathway analysis.

Sub-category	Subsystem	Main enzymes involved in synthesis
Beta-lactam compound	Monobactam biosynthesis	Aspartate kinase Aspartate-semialdehyde dehydrogenase sulfate adenylyltransferase 4-hydroxy-tetrahydrodipicolinate synthase MbtH protein
Pigment	Prodigiosin biosynthesis	3-oxoacyl reductase S-malonyltransferase enoyl reductase I
Aminoglycoside	Streptomycin biosynthesis	Phosphoglucomutase Myo-inositol-1 myo-inositol 2-dehydrogenase Glucokinase dTDP-glucose 4,6-dehydratase dTDP-4-dehydrorhamnose reductase dTDP-4-dehydrorhamnose 3,5-epimerase
Amino sugar compound	Kanosamine biosynthesis	kanosamine-6-phosphate phosphatase 3-dehydro-glucose-6-phosphate–glutamate transaminase glucose-6-phosphate3-dehydrogenase dihydroanticapsin dehydrogenase
Non-ribosomal peptide	Bacilysin biosynthesis	Protein BacA, BacB, Bac C, BacF, BacG.



Name	Count	Percent
Aging	9	0.34%
Amino acid metabolism	201	7.57%
Biosynthesis of other secondary metabolites	50	1.88%
Cancer: overview	14	0.53%
Cancer: specific types	3	0.11%
Carbohydrate metabolism	262	9.86%
Cardiovascular disease	9	0.34%
Cell growth and death	23	0.87%
Cell motility	56	2.11%
Cellular community - prokaryotes	74	2.79%
Digestive system	6	0.23%
Drug resistance: antimicrobial	30	1.13%
Drug resistance: antineoplastic	5	0.19%
Endocrine and metabolic disease	6	0.23%
Endocrine system	15	0.56%
Energy metabolism	113	4.25%
Environmental adaptation	12	0.45%
Excretory system	2	0.08%
Folding, sorting and degradation	56	2.11%
Global and overview maps	714	26.88%
Glycan biosynthesis and metabolism	50	1.88%
Immune disease	1	0.04%
Immune system	4	0.15%
Infectious disease: bacterial	19	0.72%
Infectious disease: parasitic	2	0.08%
Infectious disease: viral	2	0.08%
Lipid metabolism	71	2.67%
Membrane transport	162	6.1%
Metabolism of cofactors and vitamins	162	6.1%
Metabolism of other amino acids	54	2.03%
Metabolism of terpenoids and polyketides	36	1.36%
Nervous system	4	0.15%
Neurodegenerative disease	11	0.41%
Nucleotide metabolism	79	2.97%
Replication and repair	49	1.84%
Signal transduction	147	5.53%
Signaling molecules and interaction	1	0.04%
Substance dependence	1	0.04%
Transcription	5	0.19%
Translation	87	3.28%
Transport and catabolism	13	0.49%
Xenobiotics biodegradation and metabolism	36	1.36%

Figure 5 *B. subtilis* Miz-8 consists of a number of protein coding genes which are depicted in different colour codes that are connected to KEGG pathway

Antibiotic Resistance Gene Analysis of *B. subtilis* Miz-8

Antibiotic resistance genes were predicted using the Comprehensive Antibiotic Resistance Database (CARD). CARD uses the subset of the available NCBI that is relevant to antibiotic-resistant bacteria. The perfect algorithm predicts AMR proteins with exact homology (100%) with a query on the CARD reference sequences. Strict RGI genes representing previously unreported variants of antibiotics target changes were noticed through altered sequences. The loose

algorithm indicates the detection of new, emerging risks and more distant homologs of antibiotics resistance genes **(Micheal *et al.*, 2020)**. Genome of *Bacillus subtilis* strain Miz-8 holds six antibiotic resistance genes as depicted in tab 3. Among them, one was found to be a perfect hit for tunicamycin resistance protein (*tmrB*), whereas the others were strict hits. When compared with the previous studies, most of the *Bacillus subtilis* strains, commercially available probiotic *Bacillus subtilis* strains and other *Bacillus* sp. unveils strict hits of *tmrB* gene **(Brutscher *et al.*, 2022) (Olanrewaju *et al.*, 2021)**. Nevertheless, caution must be taken for *Bacillus subtilis* Miz-8 strain, as the *tmrB* gene in its genome has a perfect score. Knocking out the AMR genes can be performed to design a probiotic for commercial use **(Sleator, 2015) (Goh & Barrangou, 2019)**. Besides, plasmid of the strain Miz-8 contains no antibiotic resistance genes. Therefore, strain Miz-8 can be considered to be safe for human consumption, as it lacks the ability to transfer the resistance gene from one microbial community to others **(Abriouel *et al.*, 2019) (Aarts & Margolles, 2015)**.

Table 3 Occurrence of antibiotic resistance gene family in *B. subtilis* Miz-8

Resistance Gene Identifier (RGI) criteria	Antibiotic Resistance Ontology (ARO) term	Drug class
perfect	*tmrB*	nucleoside antibiotics
strict	*yddK*	aminoglycoside antibiotic, tetracycline antibiotic, phenicol antibiotic
strict	*ykkC*	aminoglycoside antibiotic, tetracycline antibiotic, phenicol antibiotic
strict	*vmIR*	macrolide antibiotic, lincosamide antibiotic, streptogramin antibiotic, tetracycline antibiotic, oxazolidinone antibiotic, phenicol antibiotic, pleuromutilin antibiotic
strict	*mphK*	macrolide antibiotics
strict	*aadk*	aminoglycoside antibiotic

Prediction of virulence genes in *B. subtilis* Miz-8

Genome of *Bacillus subtilis* Miz-8 was compared with other related species such as *B. anthracis, B. cereus, B. clausii, B. halodurans, B. licheniformis, B. subtilis* and *B. thuringiensis* for the prediction of virulence genes using VF analyser as depicted in Tab 4. *B. subtilis* Miz-8 harbored virulence genes which are associated with immune evasion, iron acquisition and toxin. Immune evasion genes such as *capA, capB, capC, capD* encoding polyglutamic acid capsules which are responsible for evading the host's immune response were predicted, thereby maximizing self-survival **(Sharma *et al.*, 2020)**. Genes encoding for exopolysaccharide, hyaluronic acid and polysaccharide capsule is absent in the strain Miz-8 which suggests that Miz-8 is not resistant to host defence and phagocytosis. Genes responsible for the biofilm formation were found to be absent on the strain Miz-8 as *B.anthracis* and *B.cereus* bear the gene for adherence. Studies have shown that the *bslA* gene is located in the pXO1 region which is necessary for its adherence to the host cells and causes a pathogenic effect **(Kern and Schneewind 2008)**. Enzymes produced from some pathogenic bacteria can overtake or counter attack the host immune system like immune inhibitor A metalloproteinase from *B.cereus*. This protein is encoded by the gene *inhA*, which aids the organism to escape from the immune cells **(Guillemet *et al.*, 2010)**. The absence of *inhA* gene in the strain Miz-8 makes them vulnerable to the host response. *B.anthracis* are known to be non-hemolytic but still bears the silent gene for the same, which are PC-PLC, PI-PLC and SMase **(Pomerantsev *et al.*, 2003)**. These genes are present and can be expressed in B.cereus for hemolytic activity but absent in the strain Miz-8. *B. subtilis* Miz-8 harbored *dhbA, dhbB, dhbC, dhbE, dhbF* genes encoding bacillibactin are involved in chelating ferric iron from the environment to sustain nutritional immunity **(Caza & Kronstad, 2013)**. Yet, they do not carry the *hal* gene and *IlsA* gene like the pathogenic bacterial strain. These genes are required to disturb the formation of protein by cleaving the histidine and capture iron from the environment **(Daou *et al.*, 2009)**. The strain Miz-8 also lacks petrobactin and cannot resist the oxidative stress **(Hagan *et al.*, 2018)**. There are regulator gene and secretion system to regulate the synthesis and release of the virulence factor as in many pathogenic bacteria such as pagR-XO1 **(Liang *et al.*, 2016)** and type VII secretion system **(Rivera-Calzada *et al.*, 2021)**. These genes which are responsible for the productions of toxins are absent in the strain Miz-8 and thus will not be able to produce harmful toxins. The strain Miz-8 was compared to some of the harmful toxins produced by *B.anthracis* and *B.cereus* which are responsible for disrupting the host mitochondria **(Häggblom *et al.*, 2002)** and cell membrane **(Liu *et al.*, 2017)**. Among these toxins, *hlyIII* gene encoding hemolysin III is the least characterized hemolytic toxin from the *B. cereus* group, involving pore forming toxin, leading to erythrocyte lysis **(Baida & Kuzmin, 1996)**. Intra generic comparison of *B. subtilis* Miz-8 was performed with the reference genome *B. subtilis* subsp. natto BEST195, which displayed identical virulence genes. *B. subtilis* subsp. natto BEST195 containing food products are commercially available in various retail markets in Japan and neighboring countries. Owing to the similarities between these strains, the presence of *B. subtilis* Miz-8 in bekang has no harm for human consumption.

Table 4 Summary of virulence genes predicted in *B. subtilis* strain Miz-8 and other related strains

Virulence factor	Related Genes	Strain Miz-8	*B.anthracis Str.Sterne* (pXO1! pXO2-)	*B.cereus* ATCC 14579, *B.cereus* G9241	*B.clausii*	*B.halodurans*	*B.licheniformis*	*B.subtilis subsp. subtilis*	*B.thuringiensis serovar*
BslA	bslA	-	pXO1-90	AQ16_5674	-	-	-	-	-
Immune inhibitor A metalloproteinase C	inhA	-	BAS1197, BAS0638	BC1284, BC0666, AQ16_1827, AQ16_1206	-	-	-	-	BT9727_1175, BT9727_0582
Phosphatidylcholine-preferring phospholipase C (PC-PLC)	plcA	-	BAS0643	BC0670					
Phosphatidylinositol-specific phospholipase C (PI-PLC)	Piplc	-	BAS3604	BC3761, AQ16_1823					
Sphingomyelinase (SMase)	Sph	-	BAS0644	BC0671, AQ16_1822					
B. cereus exo-polysaccharide (BPS)	bpsA, bpsB, bpsC, bpsD, bpsE,	-	--	AQ16_5620, AQ16_5619, AQ16_5617, AQ16_5616, AQ16_5615,					

Continue tab. 4	bpsF, bpsG, bpsH, bpsX			AQ16_5614, AQ16_5613, AQ16_5612, AQ16_5622					
Hyaluronic acid (HA) capsule	hasA,hasB,hasC	-	pXO1-93, pXO1-95, pXO1-94	AQ16_5677, AQ16_5679, AQ16_5678	-	-	-	-	-
Polyglutamic acid capsule	capA,capB,capC,capD,capE	orf03824, orf03826, orf03825, orf02076	-	-	-	-	BL02476, BL02478, BL02477, BL03798	BSU35880, BSU35900, BSU35890, BSU18410	-
Polysaccharide capsule	Undetermined	-	-	BC5263; BC5264; BC5265; BC5266; BC5267; BC5268; BC5269; BC5270; BC5271; BC5272; BC5273; BC5274; BC5275; BC5276; BC5277; BC5278; BC5279	-	-	-	-	BT9727_4945; BT9727_4946; BT9727_4947; BT9727_4948; BT9727_4949; BT9727_4950; BT9727_4951; BT9727_4952; BT9727_4953; BT9727_4954; BT9727_4955; BT9727_4956; BT9727_4957; BT9727_4958; BT9727_4959; BT9727_4960
Bacillibactin	dhbA, dhbB, dhbC, dhbE, dhbF	orf03256, orf03253, orf03255, orf03254, orf03252	BAS2204, BAS2207, BAS2205, BAS2206, BAS2208	BC2302, BC2305, BC2303, BC2304, BC2306, AQ16_227, AQ16_224, AQ16_226, AQ16_225, AQ16_223	-	-	BL04020, BL04023, BL04021, BL04022, BL04024	BSU32000, BSU31970, BSU31990, BSU31980, BSU31960	BT9727_2143, BT9727_2146, BT9727_2144, BT9727_2145, BT9727_2147
Hal	hal	-	BAS0520	BC0552, AQ16_1944	-	-	-	-	BT9727_0463
IlsA	ilsA	-	BAS1246	BC1331, AQ16_1159	-	-	-	-	BT9727_1220
Petrobactin	asbA, asbB, asbC, asbD, asbE, asbF	-	BAS1838, BAS1839, BAS1840, BAS1841, BAS1842, BAS1843	BC1978, BC1979, BC1980, BC1981, BC1982, BC1983, AQ16_545, AQ16_544, AQ16_543, AQ16_542, AQ16_541, AQ16_540	-	-	-	-	BT9727_1812, BT9727_1813, BT9727_1814, BT9727_1815, BT9727_1816, BT9727_1817
AcpAB	acpA,acpB,	-	-	-	-	-	-	-	-
AtxA	atxA	-	-	AQ16_5600, AQ16_5694	-	-	-	-	-
PagR-XO1	pagR-XO1	-	pXO1-109	AQ16_5594	-	-	-	-	-
PagR-XO2	pagR-XO2	-	-	AQ16_5706	-	-	-	-	-
PlcR-PapR quorum sensing	papR, plcR	-	BAS5198 BAS5199	BC5349 BC5350, AQ16_2670, AQ16_2669	-	-	-	-	BT9727_5032, BT9727_5033
Type VII secretion system	essC,esxB,esxL	-	BAS2037, BAS2042, BAS2035, BAS2036, BAS2043	-	-	-	-	-	-
Anthrax toxin	Cya,lef,pagA	-	pXO1-122, pXO1-107, pXO1-110	AQ16_5688, AQ16_5710, AQ16_5705	-	-	-	-	-

Continue tab. 4									
Anthrolysin O/Cereolysin O/Hemolysin I	alo	-	BAS3109	BC5101, AQ16_4769	-	-	-	-	BT9727_3096
Cereulide	cesA, cesB, cesC, cesD, cesH, cesP, cesT	-	-	-	-	-	-	-	-
Certhrax	cer	-	-	AQ16_5569	-	-	-	-	-
Cytotoxin K (Hemolysin IV)	cytK	-	-	BC1110, AQ16_1392	-	-	-	-	BT9727_1008
Hemolysin II	hlyII	-	BAS2353	BC3523, AQ16_4553	-	-	-	-	BT9727_2311
Hemolysin III homolog	Undetermined	-	BAS5305	BC5449, AQ16_2572	-	-	-	-	BT9727_5132
Hemolysin III	hlyIII	orf02207	BAS2087	BC2196, AQ16_339	ABC2532	BH2865	BL03308	BSU21790	BT9727_2027
Hemolytic enterotoxin HBL	hblA, hblC, hblD	-	-	BC3101, BC3104, BC3103, AQ16_4933, AQ16_4930, AQ16_4931	-	-	-	-	BT9727_2890, BT9727_2893, BT9727_2892
Insecticidal crystalline toxins	cry, cyt, vip	-	-	-	-	-	-	-	-
Non-hemolytic enterotoxin (Nhe)	nheA, nheB, nheC	-	BAS1749, BAS1750, BAS1751	BC1809, BC1810, BC1811, AQ16_660, AQ16_659, AQ16_658	-	-	-	-	BT9727_1727, BT9727_1728, BT9727_1729

SNPs and Indels corresponding to the reference genome

Emphasizing the antimicrobial activity of *B. subtilis* Miz-8 strain, three surfactin genes were studied for SNP, namely, *srfAA*, *srfAB* and *srfAC*. Majority of the remaining SNPs were associated with various metabolic processes and biosynthesis of several biomolecules such as enzymes, vitamins and proteins involved in various metabolic processes and catalytic reactions. Three genes were chosen since they are known lipopeptides with antibacterial properties. Multiple amino acid changes were observed in each of these genes. The numbers of amino acid changes for the respective genes are 17, 26 and 13. The positions of these amino acid changes in the surfactin genes ranged from 377880 to 398779. Among these positions, the amino acids with the most changes are Alanine and Leucine. Protein BLAST search using translated FASTA sequence was performed for the three surfactin genes and the percentage identity observed for all three genes were limited with higher reduction.

Genome of *B. subtilis* Miz-8 detected 344 Indels in comparison with the reference genome *B. subtilis* subsp. Natto BEST195. Eighty Indels were shortlisted based on the function of the genes that were affected. Among 80 Indels, three were focused owing to their property of antibiotic resistance, whereas the rest were related to biosynthesis and structural modifications. *B. subtilis* Miz 8 harbored a single nucleotide insertion in position 336486 causing upstream gene variant in the expression of *tmrB* gene, relating to tunicamycin resistance and a frameshift variant in position 602100 expressed *vmlR* gene, encoding dissociation of antibiotics such as virginiamycin M and lincomycin. A disruptive inframe deletion occurred in position 2497530 that causes expression of *aadK* gene encoding resistance to streptomycin. Like SNP, protein BLAST search using translated FASTA sequence was performed for three genes, *tmrB*, *vmlR* and *aadK* and the percentage identity for all three genes was not significant since the similarity was observed in limited sequences. The percentage identity for *aadK* was observed with a gradual decrement of 0.45% and 0.51% for *tmrB* and *vmlR*.

Production of alpha amylase connected to KEGG pathway

In vitro study of *B. subtilis* Miz-8 for the production of amylase was observed from the spot on lawn assay. As shown in the figure 6, the strain Miz-8 utilized starch as a carbon source in the production of amylase with the REA value of 4.7. Strain Miz-8 comprises six genes that are involved in the metabolic pathway of the digestive system (figure 5). Alpha amylase (EC 3.2.1.1) aids in biosynthesis of secondary metabolites, carbohydrate digestion and absorption, pancreatic secretion, starch and sucrose metabolism and salivary secretion. Other genes retrieved from the genome *Bacillus subtilis* Miz-8 includes tryptophan-rich sensory protein in metabolizing cholesterol, acetyl-CoA C- acetyltransferase (EC 2.3.1.9) in digestion and absorption of fat and Cu+ exporting ATPase (EC 7.2.2.8) and copper chaperone in mineral absorption. The salivary amylase known as ptyalin and pancreatic amylase aids in metabolism of starch to maltose. Breakdown of carbohydrates is an important role in the digestion process and alpha amylase are often involved in this metabolic process throughout the digestion. Therefore, based on *in vitro* and in silico analyses, we can conclude that *Bacillus subtilis* Miz-8 is a good alpha amylase producer and plays a crucial role in the digestive system.

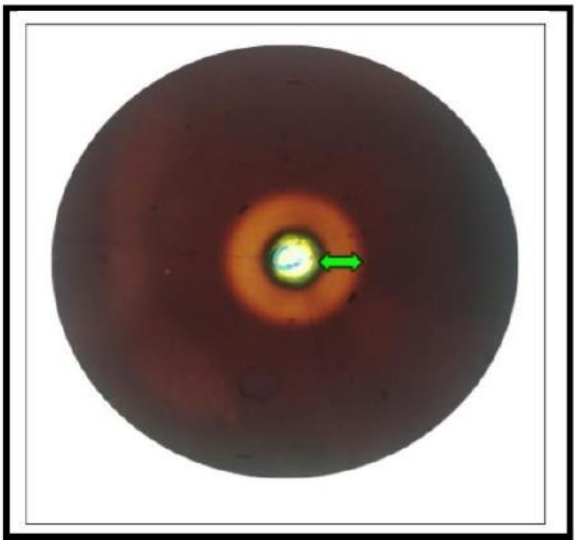


Figure 6 *B. subtilis* Miz-8 produces amylase with the zone of clearance observed on starch agar

CONCLUSION

The current study focused on bioinformatic analysis to understand the beneficial and deleterious impact of *Bacillus subtilis* Miz-8 in "bekang" fermented soybean food on human health. Next Generation sequencing has helped to gain an insightful knowledge about the genome of the strain *Bacillus subtilis* Miz-8. Further, bioinformatic study suggested that our strain has beneficial effects upon consumption and it has no harmful effects. For additional manifestation, the *in vitro* method suggested that the strain can be categorized as a good alpha amylase producer, providing in depth information about the enzyme which orchestrates certain functions in the digestion system. From the compiled results from each

category, *Bacillus subtilis* Miz 8 may have a lead role in fermenting and manufacturing essential compounds.

REFERENCES

Aarts, H., & Margolles, A. (2015). Antibiotic resistance genes in food and gut (non-pathogenic) bacteria. Bad genes in good bugs. *Frontiers in microbiology*, *5*, 754. https://doi.org/10.3389/fmicb.2014.00754

Abriouel, H., Pérez Montoro, B., de la Fuente Ordoñez, J. J., Lavilla Lerma, L., Knapp, C. W., & Benomar, N. (2019). New insights into the role of plasmids from probiotic Lactobacillus pentosus MP-10 in Aloreña table olive brine fermentation. *Scientific reports*, *9*(1), 10938. https://doi.org/10.1038/s41598-019-47384-1

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A. V., Cheng, A. A., Liu, S., Min, S. Y., Miroshnichenko, A., Tran, H. K., Werfalli, R. E., Nasir, J. A., Oloni, M., Speicher, D. J., Florescu, A., Singh, B., Faltyn, M., … McArthur, A. G. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic acids research*, *48*(D1), D517–D525. https://doi.org/10.1093/nar/gkz935

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic acids research*, *44*(W1), W16–W21. https://doi.org/10.1093/nar/gkw387

Baida, G. E., & Kuzmin, N. P. (1996). Mechanism of action of hemolysin III from Bacillus cereus. *Biochimica et biophysica acta*, *1284*(2), 122–124. https://doi.org/10.1016/s0005-2736(96)00168-x

Balasubramanian, S., Osburne, M. S., BrinJones, H., Tai, A. K., & Leong, J. M. (2019). Prophage induction, but not production of phage particles, is required for lethal disease in a microbiome-replete murine model of enterohemorrhagic E. coli infection. *PLoS pathogens*, *15*(1), e1007494. https://doi.org/10.1371/journal.ppat.1007494

Beltran-Gracia, E., Macedo-Raygoza, G., Villafaña-Rojas, J., Martinez-Rodriguez, A., Chavez-Castrillon, Y. Y., Espinosa-Escalante, F. M., … Beltran-Garcia, M. J. (2017). Production of Lipopeptides by Fermentation Processes: Endophytic Bacteria, Fermentation Strategies and Easy Methods for Bacterial Selection. InTech. doi: https://doi.org/10.5772/64236

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., van Wezel, G. P., Medema, M. H., & Weber, T. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic acids research*, *49*(W1), W29–W35. https://doi.org/10.1093/nar/gkab335

Brutscher, L. M., Borgmeier, C., Garvey, S. M., & Spears, J. L. (2022). Preclinical Safety Assessment of *Bacillus subtilis* BS50 for Probiotic and Food Applications. *Microorganisms*, *10*(5), 1038. https://doi.org/10.3390/microorganisms10051038

Carver, T., Thomson, N., Bleasby, A., Berriman, M., & Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics (Oxford, England)*, *25*(1), 119–120. https://doi.org/10.1093/bioinformatics/btn578

Caza, M., & Kronstad, J. W. (2013). Shared and distinct mechanisms of iron acquisition by bacterial and fungal pathogens of humans. *Frontiers in cellular and infection microbiology*, *3*, 80. https://doi.org/10.3389/fcimb.2013.00080

Chettri, R., & Tamang, J. P. (2015). Bacillus species isolated from tungrymbai and bekang, naturally fermented soybean foods of India. *International journal of food microbiology*, *197*, 72–76. https://doi.org/10.1016/j.ijfoodmicro.2014.12.021

Daou, N., Buisson, C., Gohar, M., Vidic, J., Bierne, H., Kallassy, M., Lereclus, D., & Nielsen-LeRoux, C. (2009). IlsA, a unique surface protein of Bacillus cereus required for iron acquisition from heme, hemoglobin and ferritin. *PLoS pathogens*, *5*(11), e1000675. https://doi.org/10.1371/journal.ppat.1000675

Deka, P., Mehetre, G. T., Lalnunmawii, E., Upadhyaya, K., Singh, G., Hashem, A., Al-Arjani, A.-B. F., Fathi Abd_Allah, E., & Singh, B. P. (2021). Metagenomic Analysis of Bacterial Diversity in Traditional Fermented Foods Reveals Food-Specific Dominance of Specific Bacterial Taxa. *Fermentation*, *7*(3), 167. https://doi.org/10.3390/fermentation7030167

Frías, J., Peñas, E., & Martínez-Villaluenga, C. (2017). Fermented Pulses in Nutrition and Health Promotion. https://doi.org/10.1016/b978-0-12-802309-9.00016-9

Goh, Y. J., & Barrangou, R. (2019). Harnessing CRISPR-Cas systems for precision engineering of designer probiotic lactobacilli. *Current opinion in biotechnology*, *56*, 163–171. https://doi.org/10.1016/j.copbio.2018.11.009

Guillemet, E., Cadot, C., Tran, S. L., Guinebretière, M. H., Lereclus, D., & Ramarao, N. (2010). The InhA metalloproteases of Bacillus cereus contribute concomitantly to virulence. *Journal of bacteriology*, *192*(1), 286–294. https://doi.org/10.1128/JB.00264-09

Hagan, A. K., Plotnick, Y. M., Dingle, R. E., Mendel, Z. I., Cendrowski, S. R., Sherman, D. H., Tripathi, A., & Hanna, P. C. (2018). Petrobactin Protects against Oxidative Stress and Enhances Sporulation Efficiency in Bacillus anthracis Sterne. *mBio*, *9*(6), e02079-18. https://doi.org/10.1128/mBio.02079-18

Kern, J. W., & Schneewind, O. (2008). BslA, a pXO1-encoded adhesin of Bacillus anthracis. *Molecular microbiology*, *68*(2), 504–515. https://doi.org/10.1111/j.1365-2958.2008.06169.x

Latorre, J. D., Hernandez-Velasco, X., Wolfenden, R. E., Vicente, J. L., Wolfenden, A. D., Menconi, A., Bielke, L. R., Hargis, B. M., & Tellez, G. (2016). Evaluation and Selection of *Bacillus* Species Based on Enzyme Production, Antimicrobial Activity, and Biofilm Synthesis as Direct-Fed Microbial Candidates for Poultry. *Frontiers in veterinary science*, *3*, 95. https://doi.org/10.3389/fvets.2016.00095

Liang, X., Zhang, E., Zhang, H., Wei, J., Li, W., Zhu, J., Wang, B., & Dong, S. (2016). Involvement of the pagR gene of pXO2 in anthrax pathogenesis. *Scientific Reports*, *6*. https://doi.org/10.1038/srep28827

Liu, B., Zheng, D., Jin, Q., Chen, L., & Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic acids research*, *47*(D1), D687–D692. https://doi.org/10.1093/nar/gky1080

Liu, X., Ding, S., Shi, P., Dietrich, R., Märtlbauer, E., & Zhu, K. (2017). Non-hemolytic enterotoxin of Bacillus cereus induces apoptosis in Vero cells. *Cellular microbiology*, *19*(4), 10.1111/cmi.12684. https://doi.org/10.1111/cmi.12684

Markowitz, V. M., Chen, I. M., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., Ratner, A., Huang, J., Woyke, T., Huntemann, M., Anderson, I., Billis, K., Varghese, N., Mavromatis, K., Pati, A., Ivanova, N. N., & Kyrpides, N. C. (2014). IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic acids research*, *42*(Database issue), D560–D567. https://doi.org/10.1093/nar/gkt963

Markowitz, V. M., Chen, I. M., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Ratner, A., Jacob, B., Huang, J., Williams, P., Huntemann, M., Anderson, I., Mavromatis, K., Ivanova, N. N., & Kyrpides, N. C. (2012). IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic acids research*, *40*(Database issue), D115–D122. https://doi.org/10.1093/nar/gkr1044

Meier-Kolthoff, J. P., & Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature communications*, *10*(1), 2182. https://doi.org/10.1038/s41467-019-10210-3

Micheal, A., Catherine, O., Adenike, A., Adeola, A., Adedamola, A., & Ademola, A. (2020). Predictive comparative antibiotic resistance (AMR) profiles of rhizobacteria genes using CARD: a bioinformatics approach. Highlights in BioScience, 3. https://doi.org/10.36462/H. BioSci.20223

Olanrewaju, O. S., Ayilara, M. S., Ayangbenro, A. S., & Babalola, O. O. (2021). Genome Mining of Three Plant Growth-Promoting Bacillus Species from Maize Rhizosphere. *Applied biochemistry and biotechnology*, *193*(12), 3949–3969. https://doi.org/10.1007/s12010-021-03660-3

Pomerantsev, A. P., Kalnin, K. V., Osorio, M., & Leppla, S. H. (2003). Phosphatidylcholine-specific phospholipase C and sphingomyelinase activities in bacteria of the Bacillus cereus group. *Infection and immunity*, *71*(11), 6591–6606. https://doi.org/10.1128/IAI.71.11.6591-6606.2003

Rivera-Calzada, A., Famelis, N., Llorca, O., & Geibel, S. (2021). Type VII secretion systems: structure, functions and transport models. *Nature reviews. Microbiology*, *19*(9), 567–584. https://doi.org/10.1038/s41579-021-00560-5

Sanjukta, S., & Rai, A. K. (2016). Production of bioactive peptides during soybean fermentation and their potential health benefits. Trends in Food Science & Technology, 50, 1–10. https://doi.org/10.1016/j.tifs.2016.01.010

Sharma, S., Bhatnagar, R., & Gaur, D. (2020). *Bacillus anthracis* Poly-γ-D-Glutamate Capsule Inhibits Opsonic Phagocytosis by Impeding Complement Activation. *Frontiers in immunology*, *11*, 462. https://doi.org/10.3389/fimmu.2020.00462

Sharma, S., Bhatnagar, R., & Gaur, D. (2020). *Bacillus anthracis* Poly-γ-D-Glutamate Capsule Inhibits Opsonic Phagocytosis by Impeding Complement Activation. *Frontiers in immunology*, *11*, 462. https://doi.org/10.3389/fimmu.2020.00462

Sleator R. D. (2015). Designer probiotics: Development and applications in gastrointestinal health. *World journal of gastrointestinal pathophysiology*, *6*(3), 73–78. https://doi.org/10.4291/wjgp.v6.i3.73

Sudan, S., Flick, R., Nong, L., & Li, J. (2021). Potential Probiotic *Bacillus subtilis* Isolated from a Novel Niche Exhibits Broad Range Antibacterial Activity and Causes Virulence and Metabolic Dysregulation in Enterotoxic *E. coli*. *Microorganisms*, *9*(7), 1483. https://doi.org/10.3390/microorganisms9071483

Tamang, J. P. (2015). Naturally fermented ethnic soybean foods of India. Journal of Ethnic Foods, 2(1), 8–17. https://doi.org/10.1016/j.jef.2015.02.003

Tamang, J. P. (2020). History and Culture of Indian Ethnic Fermented Foods and Beverages. Ethnic Fermented Foods and Beverages of India: Science History and Culture, 1–40. https://doi.org/10.1007/978-981-15-1486-9_1