The 6th International Conference on Biochemistry and Molecular Biology

Phylogenetic and nucleotide sequence analysis of *yueB* gene obtained from *Bacillus subtilis* isolated from Thai natto

Pornthip Chawapun[1], Padchanee Sangthong[2], Pairote Wongputtisin[3], Keitarou Kimura[4], Supawadee Sriyam[5,*]

[1]Master's Degree Program in Biotechnology, Graduate School, Chiang Mai University, Chiang Mai, Thailand
[2]Division of Biochemistry and Biochemical Technology, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand
[3]Program in Biotechnology, Faculty of Science, Maejo University, Thailand
[4]Applied Microbiology Unit, Food Research Institute, National Agriculture and Food Research Organization, Japan
[5]Department of Agro-Industry, Faculty of Science and Agricultural Technology, Rajamangala University of Technology Lanna, Nan, Thailand
*E-mail: padchanee.sangthong@cmu.ac.th

Abstract

Thai natto is a conventional household spontaneously fermented soybean that is commonly found in northern Thailand. The fermentation process relies on certain microbial activities, especially those of *Bacillus* species. However, contamination of Thai natto with bacteriophages can rapidly decrease *Bacillus* species growth and the quality of any related products. The bacteriophage infection occurred through the YueB membrane receptor protein on the bacterial host cell surface that was coded by the *yueB* gene. Twenty strains including *Bacillus* spp. were isolated from Thai natto. Classification of *Bacillus* was done and the specimens were then grouped using 16S-23S rRNA gene nucleotide sequencing for the purposes of phylogenetic analysis. The phylogenetic relations of the twenty Bacillus strains were generated in comparison with those obtained from the GenBank database. *Bacillus amyloliquefaciens*, *B. aryabhattai*, *B. cereus*, *B. licheniformis*, *B. sonorensis*, *B. subtilis* and *B. vallismortis* were identified. A nucleotide sequence similarity of 99% in 16S rRNA was reported. It was found that a result of 97 to 99% was the highest percent identity when compared to the relevant detail in 23S rRNA gene sequences. Then, *B. subtilis* isolated TM4, as well as TM5, TK7 and MR4 obtained from *B. subtilis* were selected to determine the types of *yueB* gene that were present. Approximately 3.2 kb of the *yueB* gene were amplified by PCR. After which, nucleotide sequencing was performed. The translation of proteins retrieved 1,059 amino acid residues in all the isolated strains. The similarity alignment protein sequence of *B. subtilis* 168 as a reference strain revealed maximum identity scores of 81% and 1% of the gaps on the protein BLAST alignment tools. Interestingly, the *yueB* gene is essential for bacteriophage SPP1 infection. Only the isolated TM4 and TM5 genes obtained from B. Subtilis that displayed specific infection by bacteriophage SPP1 were observed. In analysis, the *yueB* gene present in *B. subtilis* will serve as the constructed knockout strain to investigate the function of yueB-

bacteriophage infection using phage sensitivity assay and complementation of *yueB* gene disruption.

Introduction

This fermented soybean product is traditionally produced and widely consumed in the northern provinces of Thailand (4, 5). Thai natto is produced from soybeans and is often used as a major ingredient in condiments that are served with several traditional Thai dishes. Thai natto is not only an important food and condiment, but is also considered a low cost protein dietary supplement (6). Thai natto is brownish in color, has a unique flavor, is layered as a slightly sticky mass, and gives off a strong ammonia-like odour. The *Bacillus* species is usually used as a microorganism in the fermentation process of various soybean products. *B. subtilis* is the predominant bacteria that is used from the beginning to the end of the soybean fermentation process. However, fermented soybean products are generally prepared on a household scale and depend on natural microflora, which are then mixed with the culture in the environment and at the level of localized production. Thai natto contains a mix of cultures of *Bacillus* strains derived from the environment including *B. subtilis*, *B. cereus*, *B. megaterium* and *B. pumilus*. However, Japanese fermented soybeans (or Japanese natto) use a pure culture strain of *B. subtilis* to produce fermented soybean products yielding unique aromas and flavors (8, 15). Other related soybean fermentation products are found in Asia, such as *natto* from Japan, *kinema* from India, *schuidouchi* from China, *dawadawa* from Nigeria, and *chungkukjang* from Korea (3, 7). The quality of the Thai natto fermentation process is critical to the end product. However, the process is at risk from both spoilage and contamination by bacteriophage microorganisms during the course of the fermentation process.

Bacteriophages are prokaryote viruses that infect the host bacteria by recognizing a membrane specific receptor of the host cell for viral transfection and multiplication (2). Bacteriophage infection is initiated by a specific interaction of the virion with receptors at the host cell surface. Bacteriophage SPP1 is a siphovirus that specifically infects soil bacterium such as *B. subtilis* (9). With regard to membrane receptor proteins, YueB protein families are the main proteins that are involved in the phage binding process. *YueB* gene is coded for the YueB protein, which is found in Gram-positive bacteria. It contains specific binding sites for bacteriophage infection that occurs only for Gram-positive bacteria (14). During Thai natto fermentation, contamination by other microbes and bacteriophages from the air and environment can be observed. These can directly affect the growth of *Bacillus* strains and the quality of Thai natto products. Based on the mechanism of bacteriophage infection *via* the YueB membrane receptor protein, disruption of the *yueB* gene in *B. subtilis* can prevent the bacteriophage transmission into the bacterial membrane. This also can be used as a model study for the purposes of developing a bacteriophage biocontrol model and can be applied with other bacteria in further research.

In this research study, the screening and identification of twenty *Bacillus* species in the production of Thai natto was done using full-length 16S rRNA and partial 23S rRNA genes. The *yueB* gene derived from *B. subtilis* isolated strains was investigated in terms of the nucleotide sequences and amino acid sequences. Bacteriophage sensitivity assay was used to analyze the *yueB* gene of *B. subtilis* in Thai natto. The disruption *yueB* gene strain can be used in Thai natto fermentation to improve the quality of Thai natto products, such as in terms of colour, odor and taste. These three factors of quality are the unique features of Thai natto that can be affected during the production process. Moreover, in this study, Thai natto was

fermented without the risk of spoilage due to bacteriophage contamination in the fermentation process. The necessary *Bacillus* type strains are only grown in Thailand and are used in Thai natto production. Consequently, the risk of spoilage was decreased resulting in natto products that can be stored for longer periods of time and better preserved than in the household spontaneous fermentation process. The findings of this research study could be applied in other fermentation processes and could possibly be used as a bacteriophage biocontrol for other bacteria in further analysis.

Methodology

Bacterial strains and growth conditions

Bacillus spp. strains were isolated from Thai natto that was kindly provided by Dr. Pairote Wongputtisin (Maejo University, Chiang Mai, Thailand). Twenty *Bacillus* spp. strains are listed in Table 1. All strains were cultivated in Luria-Bertani (LB) agar plates (2.0% [w/v] LB agar). A single colony was inoculated in LB medium (11) and incubated at 37°C with incubator shaking (160 rpm) for 24 h. The cell pellets were then harvested by centrifugation for 5 min at 8,000 ×g and the supernatant was removed. The cell pellets were stored at -80°C until analysis. *Escherichia coli* DH5α cells that were transformed with pTA2 cloning vector (Toyobo, Japan) were selected in medium supplemented with ampicillin (at a final concentration of 100 µg/mL).

Genomic DNA was extracted using NucleoSpin® Tissue kits (Macherey-Nagel, Germany) according to the manufacturer's instructions. The quality of the genomic DNA was determined by 0.8% (w/v) agarose gel electrophoresis and the genomic DNA was stored at -20°C until further analysis.

Table 1. Bacterial strains, bacteriophage and plasmid vector that were used in this study.

Bacterial strain	Isolate/Strain	Source/Reference
Bacillus spp.	BP1-10, BP5-1, BP6-10, BP7-3, BP8-23, BP9-12, BP10-5, MR4, MR6, MR9, MT3, PW10, TK2, TK5, TK6, TK7, TM2, TM3, TM4, TM5	Program in Biotechnology, Faculty of Science, Maejo University, Chiang Mai, Thailand
B. subtilis	168	Department of Agro-Industry, Faculty of Science and Agricultural Technology, Rajamangala University of Technology Lanna, Nan, Thailand
Escherichia coli	DH5α	Toyobo, Japan
Bacteriophage	SPP1 (ATCC® 27689-B1™)	American Type Culture Collection (ATCC®), USA
Vector	pTA2	Toyobo, Japan

PCR amplification of the 16S rRNA and 23S rRNA genes

The DNA of the bacterial isolates and control strains were used as DNA templates to amplify 16S rRNA and 23S rRNA genes using universal primers. The 16S rRNA gene derived from *Bacillus* spp. was amplified using 20F-1500R primers set and tested for PCR reaction at gradient annealing temperatures (from 48°C to 59°C) in order to optimize the annealing temperature. For the 23S rRNA gene, the 118V-1037R universal primer that was complementary to the conserved regions was used, and the annealing temperature was optimized between 50°C and 61°C. The primers used for 16S rRNA and 23S rRNA

amplification and sequencing are listed in Table 2. PCR amplification was conducted in a total reaction volume of 50 µL containing 2.5 µL of genomic DNA, 1.25 µL of 10 µM of each primer, 20 µL of sterile deionized water, and 25 µL of 2X PCR master mix solution (i-*Taq*™) (iNtRON Biotechnology, Korea). The PCR conditions were applied in automate thermal cycling (Biometra, Germany): initial denaturation at 94°C for 2 min, followed by 30 cycles at 94°C denaturation for 30 sec, and annealing temperature for 30 sec, and extension at 72°C for 2 min, with a final extension step at 72°C for 7 min and then cooled and held at 10°C. The PCR products were determined by 1.2% (w/v) agarose gel (Bio Basic Inc., Canada) stained with RedSafe™ Nucleic Acid Straining Solution (iNtRON Biotechnology, Korea) in 1X TAE electrophoresis buffer and visualized under UV transilluminator. The 16S and 23S rRNA gene PCR products of *Bacillus* spp. isolated from Thai natto were then purified using NucleoSpin® Gel and PCR clean-up kit (Macherey-Nagel, Germany) according to the manufacturer's instructions.

Table 2. Nucleotide sequences of oligonucleotide primers were used for amplification and sequencing of 16S rRNA, 23S rRNA and *yueB* gene

Primer name	Sequence (5' to 3')	Position	Target	Sources/references
20F	GAGTTTGATCCTGGCTCAG	9-27	16S rRNA	[19]
1500R	GTTACCTTGTTACGACTT	1509-1492	16S rRNA	[19]
1500R/520F	CAGCAGCCGCGGTAATAC	519-536	16S rRNA	[19]
20F/520R	GTATTACCGCGGCTGCTG	536-519	16S rRNA	[19]
920F	AAACTCAAATGAATTGACGG	907-926	16S rRNA	[19]
920R	CCGTCAATTCATTTGAGTTT	926-907	16S rRNA	[19]
118V	CCCAATGGGGAAACCCA	114-130	23S rRNA	[20]
1037R	CGACAAGGAATTTCGCTAC	1953-1971	23S rRNA	[20]
11a	GGAACTGAAACATCTAAGTA	190-209	23S rRNA	[17]
62ar	GGGGCCATTTTGCCGAGTTC	1736-1717	23S rRNA	[17]
M13F(-20)	GTAAAACGACGGCCAGT	600-616	*yueB* gene	This study
M13R(-24)	GGAAACAGCTATGACCATG	828-846	*yueB* gene	This study
yueB765F	CCATCCAAGCACAGCAAAAGGCG	764-786	*yueB* gene	This study
yueB2570R	GCACGAGTTTCGTGGAGGCG	2592-2573	*yueB* gene	This study
yueBFBamHI	AAGATCGATgaatccATGACAGAACAACGAAAAAGCTTG	1-24	*yueB* gene	This study
yueBRSmaI	AAGATCGATcccgggTCACGATTCATACGTTTCATCGC	3209-3231	*yueB* gene	This study

Nucleotide sequencing of 16S rRNA and 23S rRNA genes

Full-length 16S rRNA and partial 23S rRNA genes were amplified using genomic DNA obtained from twenty strains of *Bacillus* spp. that were used as templates, and then all PCR products were purified by Nucleospin® Gel and PCR clean-up kit. The purified 16S rRNA and 23S rRNA gene PCR products underwent nucleotide sequencing using 1500R/520F, 20F/520R, 920F and 920R primers for full length sequencing of the 16S rRNA genes and the 11a and 62ar primers were used for 23S rRNA gene partial sequences. DNA sequencing was performed by Big Dye terminator technique (First BASE Laboratories Sdn Bhd, Malaysia).

Nucleotide sequences and phylogenetic analysis

The sequencing data of twenty complete full-length nucleotide sequences of 16S rRNA genes and partial sequences of 23S rRNA genes derived from *Bacillus* spp. were retrieved and aligned by ClustalW tools (16) using Biological sequence alignment editor (BioEdit) program (version 7.2.5). The 16S rRNA and 23S rRNA gene nucleotide sequences were then carried out by nucleotide BLAST search program analysis of the NCBI database. The 16S rRNA gene and 23S rRNA gene sequences data were submitted to the GenBank public database and published with accession numbers at the closest sequence. Aligned sequences were analyzed using the

Molecular Evolutionary Genetics Analysis (MEGA) program (version 6.06). A phylogenetic tree was generated using the neighbor-joining method to classify the 16S rRNA and 23S rRNA genes of *Bacillus* spp. isolated from Thai natto. Bootstrap confidence values were generated using 1,000 permutations of the data set for the 16S rRNA and 23S rRNA genes to derive the nucleotide sequence similarities.

Molecular cloning of yueB gene

The *Bacillus* spp. were classified using phylogenetic analysis from 16S rRNA and 23S rRNA genes. The *B. Subtilis*, which were obtained from phylogenetic tree identification including *B. subtilis* isolates MR4, TK7, TM4 and TM5, were selected to be investigated and were then used to identify the *yueB* gene. The conserved regions present in the *yueB* gene of *B. subtilis* 168 (Genbank® accession no. NR102783.1) that were obtained from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/BLAST/) database were then used to design a set of primers to amplify the *yueB* gene. Two sets of primers for PCR amplification, which included yueBFBamHI and yueBRSmaI primers, contained a unique restriction site for *Bam*HI and *Sma*I, respectively. The PCR conditions were applied as follows: suspension in a final volume of 50 µL that consisted of 2.5 µL of genomic DNA isolated from *B. subtilis*, 1.5 µL, 1.25 µL of 10 µM of each primer, 20 µL of sterile deionized water, and 25 µL of 2X PCR master mix solution (i-*Taq*™). The PCR conditions included an initial denaturation step at 95°C for 2 min, followed by 30 cycles of denaturation at 95°C for 30 sec, annealing for 30 sec, and extension at 72°C for 1 min. A final step at 72°C for 7 min was performed to end DNA synthesis and the holding temperature was 10°C. The efficiency of *yueB* gene amplification was determined by 1.0% (w/v) agarose gel electrophoresis.

The *yueB* gene was amplified by PCR and directly purified from agarose gel using NucleoSpin® Gel and PCR clean-up kit. DNA fragments of *yueB* genes from *B. subtilis* isolated MR4, TK7, TM4 and TM5, which contained 3'A overhang were ligated to pTA2 cloning vector. The transformation of pTA2-yueB recombinant DNA into *E. coli* DH5α as the competent cells was performed. *E. coli* carrying recombinant plasmid was grown in the LB agar plates that were supplemented with ampicillin. The transformants expressing β-galactosidase were selected in X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) and induced by IPTG (isopropyl-beta-D-thiogalactopyranoside) and supplemented on the LB agar plates. Recombinant DNA purification was carried out using TIANprep Rapid Mini Plasmid Kit (Tiangen Biotech, Beijing) and the orientation of *yueB* gene was confirmed by endonuclease restriction enzyme analysis.

A total of four *B. subtilis* isolates of the *yueB* gene were amplified by PCR as has been described previously. Purified plasmid of pTA2-yueB recombinant DNA was used for nucleotide sequencing. The recombinant plasmids were then sequenced using the sequencing primers yueBFBamHI, yueBRSmaI, yueB765F, yueB2570R, M13F(-20) and M13R(-24). The pTA2-yueB recombinant DNA was submitted to First BASE Laboratories Sdn Bhd, Malaysia for nucleotide sequencing. The nucleotide sequences were aligned and translated to amino acid residues using BLASTX program analysis on the NCBI database.

Bacteriophage SPP1

The wild type bacteriophage SPP1 were propagated in *B. subtilis* isolated MR4, TK7, TM4, and TM5. *B. subtilis* 168 was used as a reference strain. Bacteriophage was processed by

the soft-agar overlay and plaque-counting method. *B. subtilis* isolates were inoculated into LB medium with incubation at 37°C and shaking at 160 rpm for 18 h. Preparation of 2% (w/v) LB agar plate containing sterilized 10 mM $CaCl_2$ solution was done for underlay. The stock of the bacteriophage was suitably used to produce 10-fold several dilutions (10^{-1} to 10^{-10}) using SM (Saline-magnesium) buffer plus gelatin (5.8 g NaCl, 2.0 g $MgSO_4 \cdot 7H_2O$, 50 mL of 50 mM Tris-base pH 7.5, 5 mL of 2.0% (w/v) gelatin and the volume was adjusted to 1-liter with deionized water) (13). One hundred microliters of each bacteriophage dilution was transferred to each microcentrifuge tube containing 100 µL cell suspension of *B. subtilis* isolates and mixed well. Subsequently, 200 µL of bacteriophage-*B. subtilis* suspension was incubated for 30 min at 37°C. After the incubation phase, 200 µL of bacteriophage-*B. subtilis* suspension was added to test tubes containing 3 µL of 0.5% (w/v) LB semi-solid medium which contained sterile 10 mM $CaCl_2$ solution for overlay and the contents were gently mixed. The overlay medium was immediately poured onto the underlay surface and the overlays were allowed to sit for 10 min at room temperature for solidification before incubation. The culture plates were incubated at 37°C for 24 h. The plaques were then visualized as zones of clearing in the bacteria lawn.

Results and Discussion

16S rRNA and 23S rRNA gene sequence analysis

The 16S rRNA and 23S rRNA genes have been widely used for identification and classification of microorganism nucleotide sequences analysis, which is a simple and commonly used method. In this study, 16S rRNA and 23S rRNA genes obtained from twenty *Bacillus* spp. isolates were sequenced. The nucleotide sequences were published and are available on the NCBI database. Complete 16S rRNA and partial 23S rRNA gene sequences were determined using universal primers. It was found that 55.5°C was the optimal annealing temperature to amplify 1.5 kb of the 16S rRNA gene without non-specific bands and high PCR product yield. The amplification of the 23S rRNA gene was done at 50.3°C to obtain annealing temperature and the PCR product was found to be approximately 900 bp in size. The complete full-length nucleotide sequences of 16S rRNA genes and the partial nucleotide sequences of 23S rRNA genes were aligned and used to construct a phylogenetic tree. The position differences of the nucleotide sequences of 16S rRNA and 23S rRNA genes are shown in Figures 1 and 2, respectively.

The phylogenetic tree of 16S rRNA genes obtained from twenty *Bacillus* spp. isolates using bootstrap values based on 1000 replications is shown in Figure 3A. The 16S rRNA gene sequences of *Bacillus* spp. strains revealed more than 97 to 98% similarity when compared with the *B. subtilis* 168 reference strain. The twenty *Bacillus* spp. isolated from Thai natto strains were separated in to six clusters, which included *B. amyloliquefaciens*, *B. aryabhattai*, *B. cereus*, *B. licheniformis*/*B. sonorensis*, *B. subtilis* and *B. vallismortis*. Moreover, *Bacillus* spp. isolates TK6, TK2, TM2 and MT3 were found to be closely related to *B. licheniformis* and *B. sonorensis*. These strains could not be identified in the related *Bacillus* spp. group. The partial 23S rRNA gene was then used to confirm the identification of twenty *Bacillus* spp. isolated from Thai natto. Approximately 900 bp of the 23S rRNA gene sequences were amplified using universal primers. Sequencing of the twenty *Bacillus* spp. strains was obtained from partial nucleotide sequences. Genetic relationship analysis of partial 23S rRNA genes was used to verify that *B. amyloliquefaciens*, *B. anthracis*, *B. aryabhattai*, *B. cereus*, *B. licheniformis*, *B. megaterium*, *B.*

sonorensis, *B. subtilis* and *B. vallismortis* were separated (Figure 3B). From the results, *Bacillus* spp. isolated TK6 and MT3 were confirmed as being similar to *B. sonorensis* and *Bacillus* spp.; isolated TK2 and TM2 were included in the *B. licheniformis* group with 99% similarity. Analysis of all *Bacillus* spp. strains revealed more than 97-99% similarity to the 23S rRNA gene sequences as has been referenced with the Genbank® database.

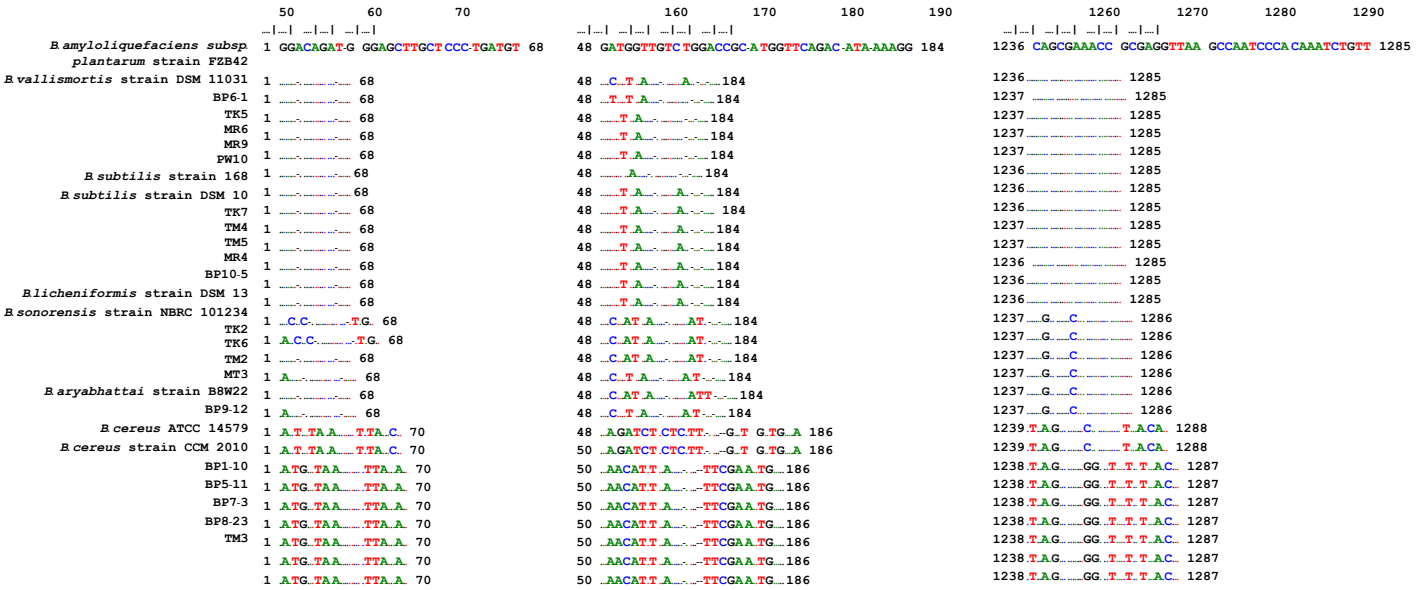

Figure 1. Position differences in full-length 16S rRNA gene nucleotide sequences using ClustalW alignment analysis of the *Bacillus* species and twenty isolates from Thai natto

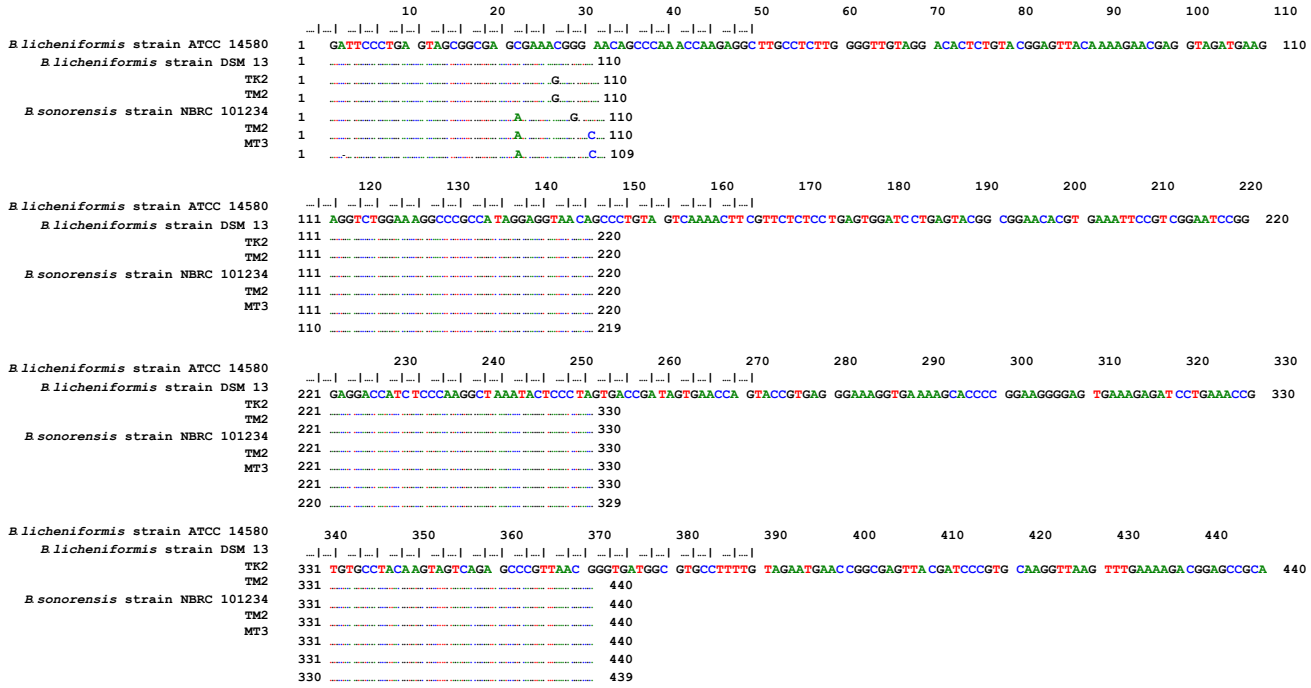

Figure 2. Position differences in partial 23S rRNA gene nucleotide sequences using ClustalW alignment analysis of the *B. licheniformis*, *B. sonorensis* and *Bacillus* species isolated from Thai natto

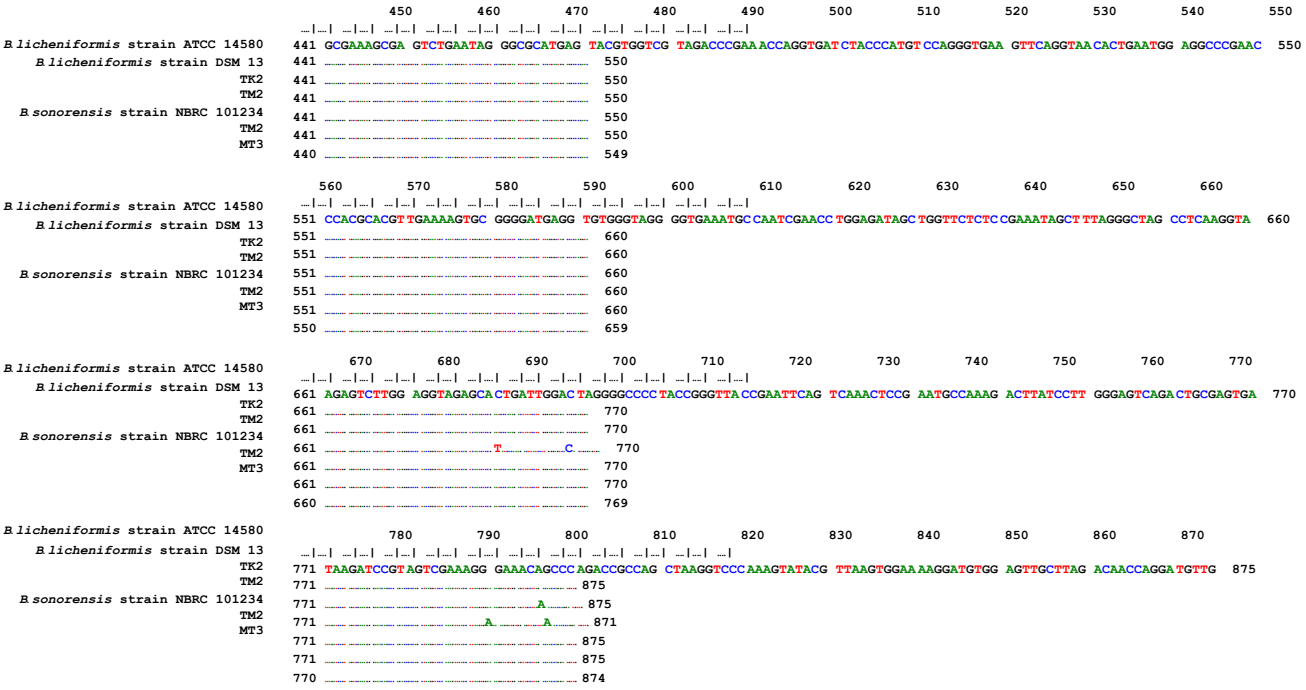

Figure 2. *(continued)* Position differences in partial 23S rRNA gene nucleotide sequences using ClustalW alignment analysis of the *B. licheniformis*, *B. sonorensis* and *Bacillus* species isolated from Thai natto

In the previous study, the taxonomy of the members of the *Bacillus* species group was determined using 16S rRNA and 23S rRNA gene sequence analysis. Analysis of the *Bacillus* group strains indicated that the three groups consisted of *B. cereus*, *B. thuringiensis*, and *B. mycoides*. Based on 16S rRNA and 23S rRNA gene nucleotide sequences, the *B. cereus* group was divided into seven sub-groups including *B. anthracis*, *B. cereus* A and B, *B. mycoides* A and B and *B. thuringiensis* A and B (1). In addition, the *B. subtilis* was used to generate a phylogenetic tree from the 16S rRNA gene. Phylogenetic analysis revealed the division of the *Bacillus* group into four clusters: cluster I contained *B. subtilis*, *B. vallismortis* and *B. mojavensis* strains; clusters II and III contained strains of *B. atrophaeus* and *B. amyloliquefaciens*, respectively; and cluster IV contained *B. sonorensis* and *B. licheniformis*. All *Bacillus* strains were recorded at more than 98 % similarity (18).

The Thai natto fermentation process is similar to the process used to produce Japanese natto, which employed the *Bacillus* species microorganism in the fermentation process. Thai natto was produced using a traditional fermentation method that used natural microflora, whereas Japanese natto was fermented using a pure culture of *B. subtilis*. Leejeerajumnean, 2003 (10) reported that Thai natto fermentation produced at 0 to 36 h involved natural microflora of *B. subtilis*, *B. megaterium* and *B. Cereus* that were produced over a period of 72 h. The *B. subtilis* was the predominant microorganism from the beginning to the end of the fermentation process employing the traditional method. In the previous study, Thai natto was collected in the North of Thailand and LPO-2 isolated strains were identified as *Bacillus* species and were closely related to the *B. subtilis* group. The phylogenetic tree using 16S rRNA gene sequencing indicated that strain LPO-2 belonged to the genus *Bacillus* and was closely related to *B. tequilensis* in the cluster of *B. subtilis* with 99.72% similarity (12). In this research study, we found *Bacillus* isolates MR4, TK7, TM4 and TM5 were closely related to those of *B. subtilis* recorded in the Genbank® database.

A)

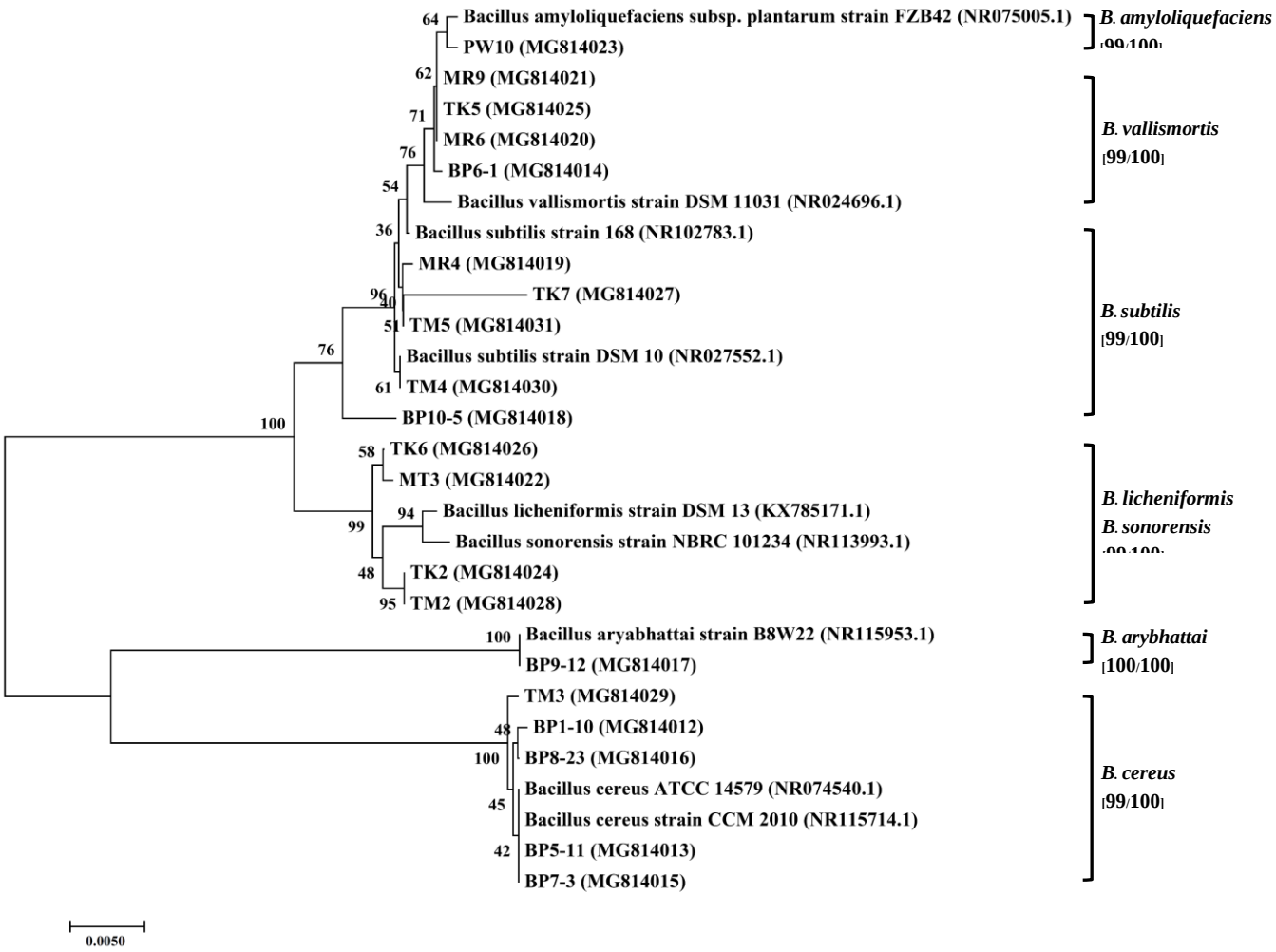

Figure 3. Genetic relationships among the twenty *Bacillus* spp. strains are presented based on A (16S rRNA gene and B (23S rRNA gene. The phylogenetic tree was constructed using the Neighbor joining method to demonstrate the classification of *Bacillus* spp). based on 1000 bootstrap replications).

B)

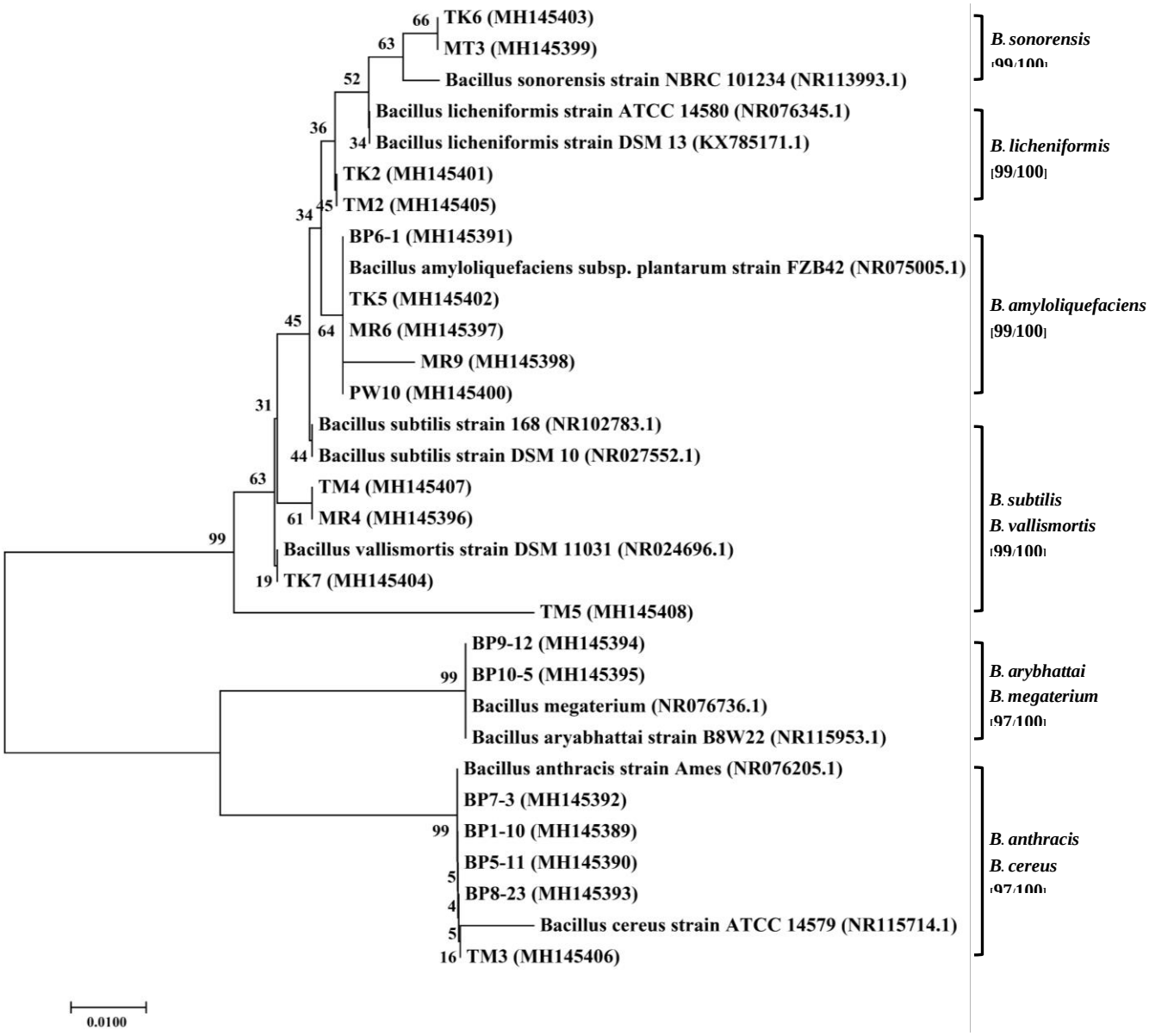

Figure 3) *.continued(* Genetic relationship among the twenty *Bacillus* spp .strains based on A (16S rRNA gene and B (23S rRNA gene .The phylogenetic tree was constructed using the Neighbor joining method to demonstrate the classification of *Bacillus* spp) .based on 1000 bootstrap replications).

Nucleotide sequences and amino acid sequences of yueB gene

The YueB membrane receptor protein was coded by the *yueB* gene, which was the essential gene involved in the specific irreversible binding of bacteriophage SPP1 to *B. subtilis*. *B. subtilis* isolated MR4, TK7, TM4 and TM5 were selected to investigate the *yueB* gene. As a result, *yueB*-analysis full-length nucleotide sequences obtained from *B. subtilis* isolated MR4, TK7, TM4 and TM5 were related to *B. subtilis* strain 168 with a maximum identity score of 86% according to nucleotide BLASTN search program analysis. The nucleotide deletion of *B. subtilis* isolated MR4, TK7, TM4 and TM5 were observed at the position 1,160-1,165)ΔGAGAAGA(, 1,351-1,356) ΔTCTCCTAATA(, 1,369-1,376)ΔCCGTCGAA(, 1,444-1,447)ΔGATA(, 1,481-1,489)ΔGAAACGGGAC(, 1,567-1,624)ΔAATCAAAGTGATGGGCTA .(

Moreover, only three strains containing *B. subtilis* isolated MR4, TM4 and TM5 were observed to display the nucleotide deletion at position 1,363 (ΔG) and position 176 (ΔA) of *B. subtilis* isolated MR4. The nucleotide insertions were found in *B. subtilis* isolated TM5 at positions 2,232 (C), 2,264 (C) and 2,273 (T), while *B. subtilis* isolate TK7 revealed the insertion at position 1,363 (C) (Supplemental data1). Nucleotide sequences were then translated to amino acid sequences using BLASTX program on NCBI GenBank® database. The complete protein translation retrieved 1,059 amino acid residues in all isolated strains, while the nucleotide sequence obtained from the *B. subtilis* strain 168 encoded 1,076 amino acid residues (Figure 4). Similarly, the alignment protein sequence of *B. subtilis* 168 as the reference strain revealed a maximum identity score of 81% and only 1% of the gaps on the protein alignment tools. The protein alignment containing four regions of 387-388, 458-462, 488-491 and 546-541 gap position were observed.

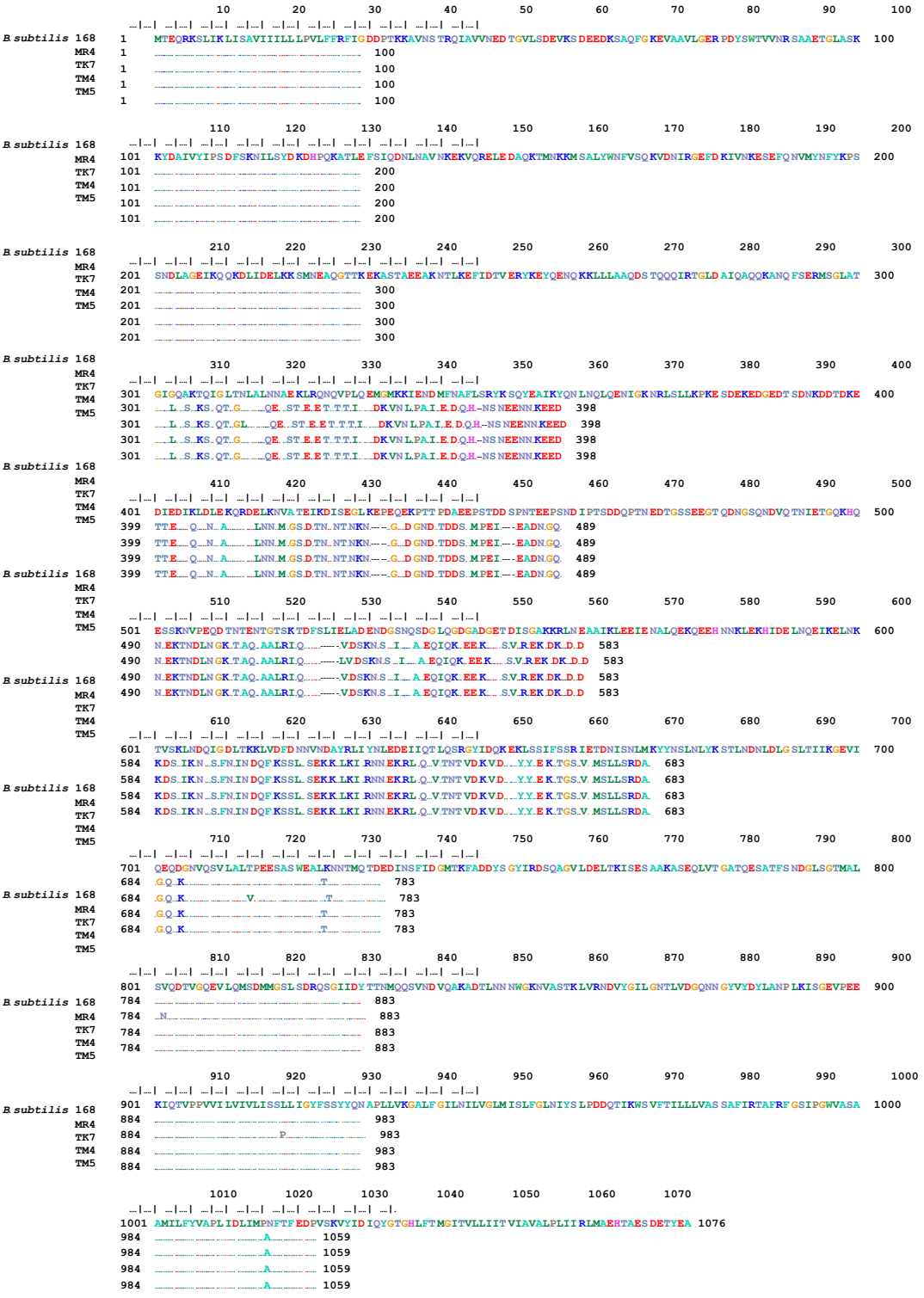

Figure 4. *Yue*B gene amino acid sequence alignment of *B. subtilis* isolates MR4, TK7, TM5 and TM5. *B. subtilis* 168 was used as the reference strain.

YueB-bacteriophage SPP1 infection

Bacteriophage SPP1 sensitivity of the *yueB* gene obtained from *B. subtilis* isolated MR4, TK7, TM4, TM5 and *B. subtilis* 168 was used as a positive control to employ the bacteriophage infection assay. The bacteriophage sensitivity was determined by counting the plaques (zones of clearing). Subsequently, the plaque forming units were evaluated. The results showed that only *B. subtilis* isolate MR4, TM4 and TM5 strains were infected by bacteriophage SPP1 when

compared with *B. subtilis* 168. In addition, only *B. subtilis* isolate TK7 was not infected by bacteriophage SPP1. This indicated that the differences in the nucleotide sequences and amino acid sequences of the *yueB* genes in all four strains might affect the bacteriophage SPP1 infection via different pathways.

In the previous study, the *yueB* genes were identified as the essential genes that are involved in the irreversible binding of bacteriophage SPP1 to *B. Subtilis*, while *YueB gene*-bacteriophage SPP1 adsorption of the wild-type and mutation *B. subtilis* cells were observed. Gene organization of the *yueB* and surrounding genes (*yueBA*, *yueC*, *yueD*, and *yueK*) were cloned in pMUTIN4, an integration vector specifically for *B. subtilis*. Bacteriophage SPP1 was used to efficiently plate the presence of IPTG. As a result, bacteriophage SPP1 plaque efficiency was found to be close to the wild-type strain in all the integrant strains, except for the *yueB* disrupted strain. Differences in plaque size were observed and smaller plaques were observed than in the wild-type strain of the integrants. These results revealed that only the *yueB* gene was necessary for irreversible binding bacteriophage SPP1 (14).

YueB membrane receptor protein in the *B. subtilis* strain was an essential target for bacteriophages SPP1 infection. The bacteriophages adsorption to the microorganism was involved in the specific interaction and the binding in a reversible or irreversible manner. Thai natto is a fermented soybean product that utilizes the *Bacillus* species as the predominant microorganism in the fermentation process. With regard to the bacteriophage infection of the YueB membrane receptor protein, information of the *Bacillus* species in Thai natto fermentation has not been fully reported. In this research study, bacteriophage SPP1 biocontrol was used in Thai natto fermentation using *yueB* gene disruption from *B. subtilis*. The integrant *yueB* gene can be used to improve the quality of Thai natto and may be applied in other fermentation processes.

Conclusion

In summary, twenty strains of *Bacillus* species isolated from Thai natto were successfully investigated, classified and presented in a phylogenetic tree. Nucleotide sequences of full-length 16S rRNA and partial 23S rRNA genes were published in terms of the sequences available *via* the Genbank® database. The results of phylogenetic analysis indicated that twenty *Bacillus* species isolated from Thai natto were closely related to the six clusters of *B. amyloliquefaciens*, *B. aryabhattai*, *B. cereus*, *B. licheniformis*, *B. sonorensis*, *B. subtilis* and *B. vallismortis*. *Bacillus* strain isolated MR4, TK7, TM4 and TM5 were similar to *B. subtilis*. The nucleotide sequences of the *yueB* gene obtained from *B. subtilis* isolated MR4, TK7, TM4 and TM5 were found of approximately 3.2 kb in size. The complete protein translation was performed and 1,059 amino acid residues were retrieved in all the isolate strains, whereas *B. subtilis* 168 encoded 1,076 amino acid residues. Similarity alignment protein sequence with a maximum identity score of 81 %was observed and 1 %of gaps on protein alignment tools were found.

References

1. Bavykin SG, Lysov YP, Zakhariev V, Kelly JJ, Jackman J, Stahl DA, Cherni A. J Clin Microbiol. 2004; 42: 3711-3730.
2. Bertin A, De Frutos M, Letellier L. J Microbiol. 2011; 14: 492-496.
3. Chantawannakul P, Oncharoen A, Klanbut K, Chukeatirote E, Lumyoung S. Sci Asia. 2002; 28: 241-245.
4. Chukeatirote E, Chainun C, Siengsubchart A, Moukamnerd C, Chantawannakul P, Lumyoung S, Boontim N, Thakang P. J Microbiol. 2006; 1: 38-44.
5. Chukeatirote E, Thakang P. J Sci. 2006; 33: 243-245.
6. Chukeatirote E, Dajanta K, Apichartsrangkoon A. J Biol Sci. 2010; 10: 581-583.
7. Dajanta K, Chukeatirote E, Apichartsrangkoon A. Chiang Mai J Sci. 2012; 39: 562-574.
8. Inatsu Y, Kimura K, Itoh Y. Jpn Agric Res Q. 2002; 36: 169-175.
9. Jakutytė L, Lurz. R, Baptista C, Carballido- Lopez R, São- José C, Tavares P, Daugelavičius R. Arch Virol. 2012; 422: 425-434.
10. Leejeerajumnean A. Silpakorn Univ Int J. 2003; 3: 277-292.
11. Mamiatis T, Fritsch ET, Sambrook J. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory. New York, NY; 1982: pp 68-73.
12. Meannui N, Riebroy S, Tangwatcharin P, Hong JH, Sumpavapol P, Chaijan M. Chiang Mai J Sci. 2017; 44: 1257-1269.
13. Sambrook J, Russell DW. The condensed protocols from molecular cloning: A Laboratory Manual. Cold Spring Harbor Laboratory. New York, NY; 2006: pp 691-716.
14. São-José C, Baptista C, Santos MA. J Bacteriol. 2004; 186: 8337-8346.
15. Suppadit T, Sangla L, Pintasean S, J Environ Manage. 2008; 4: 29-37.
16. Thompson JD, Higgins DG, Gibson TJ. Nucleic Acids Res. 1994; 22: 4673-4680.
17. Van camp G, Chappelle S, Wachter RD. Curr Microbiol. 1993; 27: 147-151.
18. Wang LT, Lee FL, Tai CJ, Kasai H. Int. J Syst Evol. Microbiol. 2007; 57: 1846-1850.
19. Yamada Y, Katsura K, Kawasaki H, *et al*. Int. J Syst Evol. Microbiol. 2000; 50: 823-829.
20. Zhang Y, Arias CR. Syst Appl Microbiol. 2007; 30: 93-101.

Acknowledgement

The authors are grateful for the financial support received from the Faculty of Science and Agricultural Technology, Rajamangala University of Technology Lanna, Nan, Thailand, and the Division of Applied Microbiology, National Food Research Institute, Japan. We also express our sincere appreciation to Dr. Pairote Wongputtisin, Program in Biotechnology, Faculty of Science, Maejo University for kindly providing the culture collection of the *Bacillus* type strains. We would also like to extend our gratitude to the Graduate School, and the Department of Chemistry, Faculty of Science, Chiang Mai University for providing the facilities and the equipment used in this research study.