A multi locus sequence analysis scheme for phylogeny of the *Bacillus subtilis* species complex and its advantages over 16S rRNA genes

Dao Nu Dieu Hong[1*], Trang Hoang Long[2], Nguyen Thi Thuy Tien[1], Dinh Anh Hoa[1], Tran Thi Phan[1], Le Thi Huynh Tram[1,2]

[1]Biotechnology Center of Ho Chi Minh City, Ho Chi Minh City, Vietnam
[2]International University, Vietnam National University-HCMC, Vietnam
[*]Corresponding author: dieuhong1791@gmail.com

ARTICLE INFO

DOI:10.46223/HCMCOUJS.tech.en.12.2.2255.2022

Received: April 25th, 2022

Revised: May 18th, 2022

Accepted: May 27th, 2022

Keywords:

a Multi Locus Sequence Analysis (MLSA approach); *Bacillus subtilis;* housekeeping genes; 16S rRNA

ABSTRACT

A multi locus sequence analysis (MLSA approach) was studied on the *Bacillus* genus, or the *Bacillus subtilis* species complex for specific, including 08 strains from four species (*B. subtilis, B. pumilus, B. licheniformis* and *B. amyloliquefaciens*) were provided by Biotechnology Center of Ho Chi Minh City. The research was based on sequences of 16S rRNA genes, the concatenation of five protein-coding housekeeping genes: *glpF, pta, purH, pycA,* and *rpoD*. After PCR amplification and sequencing the phylogenetic tree of 16S rRNA sequences, concatenate sequences (as well as the phylogenetic tree of each housekeeping gene) are constructed for comparison and discussion. The aim of this study is reach for better resolution and differentiation of strains and species within the *B. subtilis* species and to determine whether MLSA scheme show advantages in 16S rRNA gene-based studies.

1. Introduction

The genus *Bacillus* was named in 1835 by Christian Gottfried Ehrenberg, to contain rod-shaped (bacillus) bacteria (Ehrenberg, 1835). *Bacillus* was later amended by Ferdinand Cohn to further describe them as spore-forming, Gram-positive, aerobic, or facultatively anaerobic bacteria (Cohn, 1872). Since identified, *Bacillus* species have been studied worldwide for their uses in many medical, pharmaceutical, agricultural, and industrial processes that take advantage of their wide range of physiology characteristics and capacity to produce a variety of enzymes, antibiotics, and other metabolites (Turnbull & Baron, 1966). At the time of writing, there are 318 species have been named in total, with *Bacillus subtilis* as the type species (Skerman, McGowan, & Sneath, 1980). The *Bacillus subtilis* species complex is a group that presents phylogenetically and phonetically homogenous species, consisting of the important species *B. subtilis*, *B. licheniformis*, and *B. pumilus* (Fritze, 2004). Besides the three original species has been mentioned above, which were described in 1973 (Gordon, Haynes, & Pang, 1973), many novel species have been identified that belong to the *B. subtilis* species complex, which is: *B. atrophaeus* (Nakamura, 1989), *B. mojavensis* (Roberts, Nakamura, & Cohan, 1994), *B. vallismortis* (Roberts et al., 1994), *B. sonorensis* (Palmisano, Nakamura, Duncan, Istock, & Cohan, 2001), *B. velezensis* (Ruiz-García, Béjar, Martinez-Checa, Llamas, & Quesada, 2005), *B. tequilensis* (Gatson et al., 2006), *B. halotolerans* (Bacon & Hinton, 2002) and *B.*

amyloliquefacien (Priest, Goodfellow, Shute, & Berkeley, 1987). Also, in 1999 (Nakamura, Roberts, & Cohan, 1999) *B. subtilis* was subdivided into two subspecies: *B. subtilis* subsp. *subtilis*, *B. subtilis* subsp. *spizizenii* and in 2009 (Rooney, Price, Ehrhardt, Swezey, & Bannan, 2009), *B. subtilis* subsp. *inaquosorum* has been classified as the third subspecies of *B. subtilis*. For the *B. subtilis* species complex, it has been recognized that these species are hard to distinguish on the basis of traditional phenotypic methods for many years.

For over 40 years, in perspective of phylogenetic analyses of bacteria, the 16S rRNA gene is most popular genetic marker for identification of bacteria (Yarza et al., 2014). Liu's study that we use 16S rRNA gene sequence is unable to distingguish different strains (Liu et al., 2013).

The Multi Locus Sequence Analysis (MLSA) method has been increasingly used in phylogeny and taxonomy (Gevers et al., 2005; Peeters et al., 2016). In general, MLSA concatenates the sequence of gene fragments of the housekeeping genes and utilize this concatenated sequence to determine phylogenetic relationships.

Many housekeeping genes are *rpoD* (RNA polymerase sigma factor), *pycA* (pyruvate carboxylase), *pta* (phosphate acetyltransferase), *glpF* (glycerol uptake facilitator) and *purH* (phosphoribosyl aminoimidazole carboxamide formyl transferase), because (Glaeser & Kämpfer, 2015). The concatenated housekeeping genes could minimize the weight, combination and position of closely strains (Glaeser & Kämpfer, 2015; Timilsina et al., 2015). This analysis MLSA assigns a high similarity between sequences differing only a single nucleotide. As a result, this analysis is more suitable for organisms with clonal evolution and can determine phylogenic relationships between closely related species.

In this study, the new MLSA scheme was created based on five housekeeping genes with *glpF, pta, purH, pycA, rpoD* and 16S rRNA, then construct the phylogenetic tree to compare between 16S rRNA gene-based studies and MLSA scheme to show that whether MLSA approach truly better resolution than 16S rRNA sequences.

2. Materials and methods

2.1. Strains and primers

Eight strains from the *Bacillus subtilis* species complex whose names have been validly published belonging to the genus *Bacillus* were analyzed in this study (Table 1). All strains belonged to the HCMBiotech Collection of Microorganisms (HBCM) in the Biotechnology Center of Ho Chi Minh City, Vietnam.

These strains were grown aerobically on TSA agar. An addition of 35 *Bacillus* species (34 species belong to the *B. subtilis* species complex, one used as an outgroup) was used to reconstruct the phylogenetic tree. All 35 *Bacillus* species were retrieved from National Center for Biotechnology Information (NCBI) accession number can be found in Supplementary Table 1. All oligonucleotide primers used in this study are listed in Table 2. Primers's names include all information on the region amplified for each gene.

Table 1

The information of all strains used in this study include strains ID, Bacterium's name

Strains number	Strains ID	Bacterium
1	HBCM-B0029	*B. amyloliquefaciens*
2	HBCM-B0037	*B. pumilus*
3	HBCM-B0027	*B. subtilis*
4	HBCM-B0028	*B. subtilis*
5	HBCM-B0039	*B. tequilensis*
6	HBCM-B0020	*B. licheniformis*
7	HBCM-B0116	*B. subtilis*
8	HBCM-B0112	*B. amyloliquefaciens*

Source: HBCM data bacteria of collecting species

Table 2

List of primers were used in this study

Primers	Sequence (5'-3')[1]	Target genes	Source
20F	AGA GTT TGA TCM TGG CTC AG	16S rRNA	Integrated DNA Technologies
1500R	CGA TCC TAC TTG CGT AG		
glpF-F	WTG ACA GCA TTT TGG GG	glycerol uptake facilitator (*glpF*)	
glpF-R	GTA AAA TAC RCC GCC GA		
pycA-F	AAA TCA GAR GCG AAA GC	pyruvate carboxylase (*pycA*)	
pycA-R	CCT GAG CGG TAA GCC AT		
pta-F	ATA CAT AYG AAG GVA TGG AAG A	phosphotransacet ylase (*pta*)	
pta-R	TAG CCG ATR TTY CCY GCT		
purH-F	AYA TTC ACG GMG GNC TBC T	phosphoribosyl aminoimidazole carboxamide formyl transferase (*purH*)	
purH-R	TGY TCK BTC GGY TCY CTT TT		
rpoD-F	GCY GAA GAA GAA TTT GAC CTB AA	RNA polymerase sigma factor (*rpoD*)	
rpoD-R	CGT TTR CTT CTG CTH GGA TGT CT		

[1]Degenerate base codes: R = A+G; Y = C+T; M = A+C; K = G+T; W = A+T; H = A+T+C; B = G+T+C; N = A+C+G+T; V = A+C+G

Source: The researcher's data analysis

2.2. DNA extraction and annealing temperature determining

Total DNA extraction were completed by using chemical and enzymatic lysis protocol to extract DNA from bacterial biomass (Elkins, 2013). The quantity and purity of the total DNA

were measured using a Nanodrop spectrophotometer at 260/280nm. The optimum annealing temperature of each primer pair was determined by gradient PCR, using a range of eight different temperatures with Tm - 5 degrees as the middle temperature. The PCR results were analysis by 1% agarose gel electrophoresis with a DNA ladder from Thermo Fisher Scientific. The annealing temperature displaying a good amplification was selected for later PCR reactions. The base Tm of each primer was taken from the production, and was shown in Supplementary Table 2. Five different temperatures for each primer are shown in Supplementary Table 3.

2.3. Amplification and sequencing

In each strain, five gene fragments *glpF*, *pycA*, *pta*, *purH* and *rpoD* were amplified and sequenced. The 16S rRNA region was also amplified and sequenced for comparison. PCR amplification was performed on total DNA extraction. The amplification recipe followed Thermal Scientific DreamTaq DNA Polymerase according to the manufacturer's recommendation. All reactions were performed in an Eppendorf Mastercycler EP gradient thermocycler. After amplification step, PCR products were visualized by 1% electrophoresis agarose gel and stained with GelRed® Nucleic Acid Gel Stain (Biotium). Their concentration is verified by a Nanodrop spectrophotometer at 260/280nm. Then, PCR products were purified using the PCR purification kit (Axygen Scientific, Inc., USA) according to the manufacturer's instructions and collected DNA purified PCR products were. Next, purified PCR products were sequenced using the Sanger sequencing method with BigDye™ Direct Cycle Sequencing Kit (Thermo Fisher Scientific).

2.4. Individual gene analyses

Some characteristics of the individual gene was analyzed such as parsimony informative sites, GC content and sequence similarities for further discussion and comparison.

2.5. Multisequence alignment

PCR sequencing products were then imported into machine to get fully nucleotide sequences. The nucleotide sequences were edited by using ATGC software (GENETYX CORPORATION). After that, all gene sequences were analyzed using the software MEGA-X to compare. Sequences from the 16S region and five other housekeeping genes *glpF*, pycA, *pta*, *purH* and *rpoD* from different *Bacillus* isolates strains and one species from outside of the *B. subtilis* species complex are obtained from NCBI. Multisequence alignment using ClustalW with default settings to align the same region on different strains obtained. Phylogenic analysis from individual genes was conducted, then gene fragments of each species were concatenated. Phylogenic reconstructions were generated in each strain in this study using Neighbor-Joining Analysis in MEGA-X.

3. Results

3.1. Primers optimal annealing temperature

The goal of this experiment was to determine the optimal annealing temperature of five pairs of primers: *glpF, pycA, pta, purH, rpoD*. Gradient PCR was used to determine the optimal annealing temperature for PCR reactions to amplify the five genes' fragments. PCR products were visualized under agarose gel electrophoresis. The result shows that in the *glpF* fragment the optimal temperature is 48˚C. Similarly, the optimal temperature for four genes fragments *pycA*; *pta purH, rpoD* is 46˚C, 49˚C, 54˚C, and 51˚C, respectively.

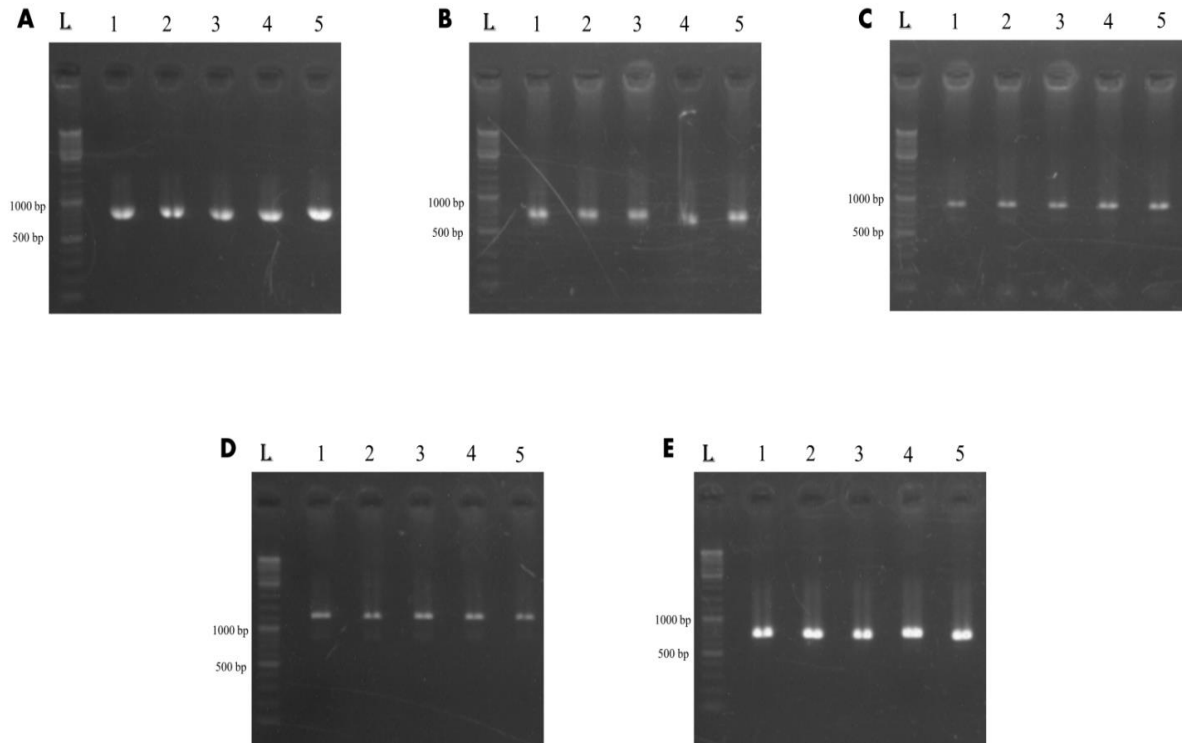

Figure 1. Gradient PCR results for determination of optimal annealing temperature (46˚C, 49˚C, 54˚C, 51˚C) for primer pairs of the five genes fragments. (A) is *glpF*, (B) is *pycA*, (C) is *pta*, (D) is *purH*, (E) is *rpoD*. The "L" is short for DNA ladder

3.2. PCR amplification and sequencing

The five gene fragments and 16S region were amplified and sequenced on eight strains.

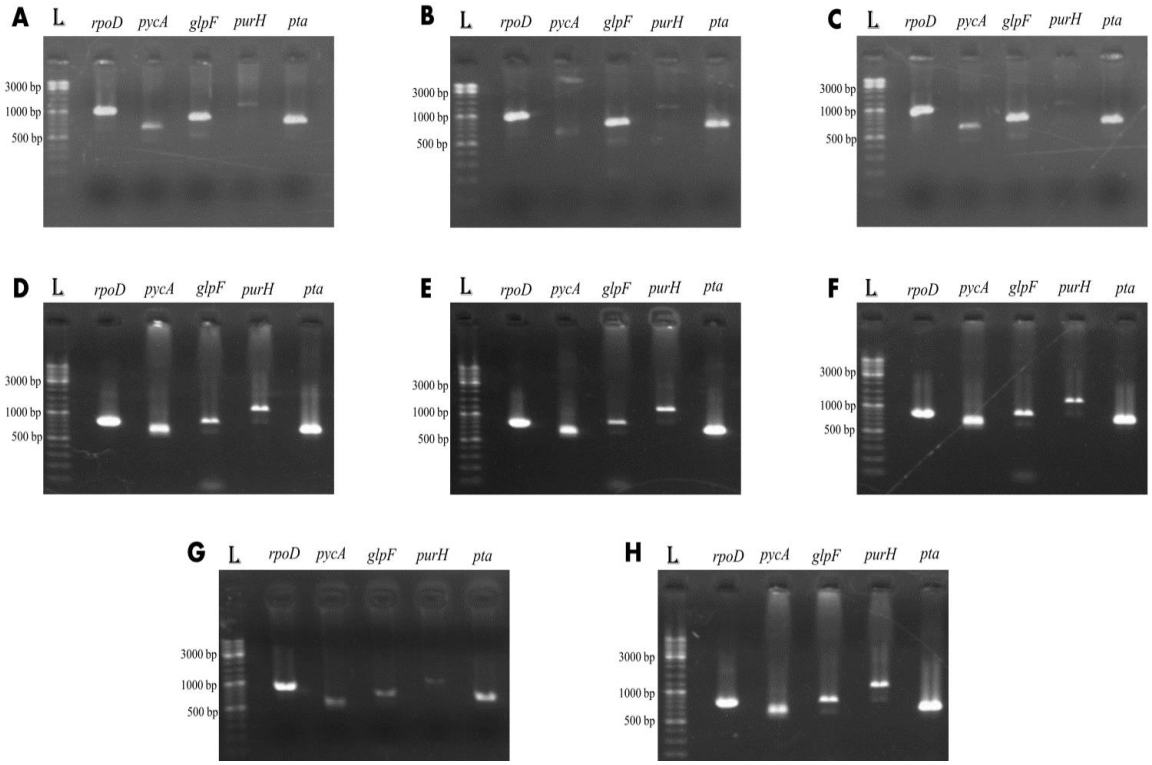

Figure 2. PCR amplification on eight strains. The "L" is short for DNA ladder. (A) to (H) are strain 1 to strain 8, respectively

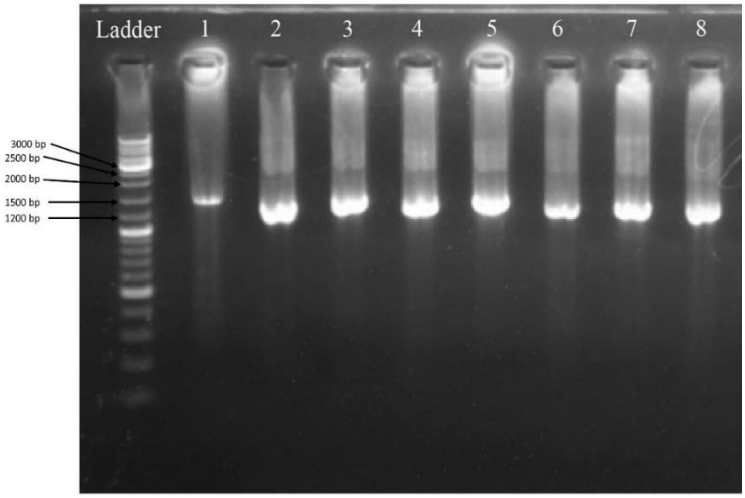

Figure 3. PCR amplification on eight strains for 16S region. Lanes 1 to 8: strain 1 to strain 8, respectively

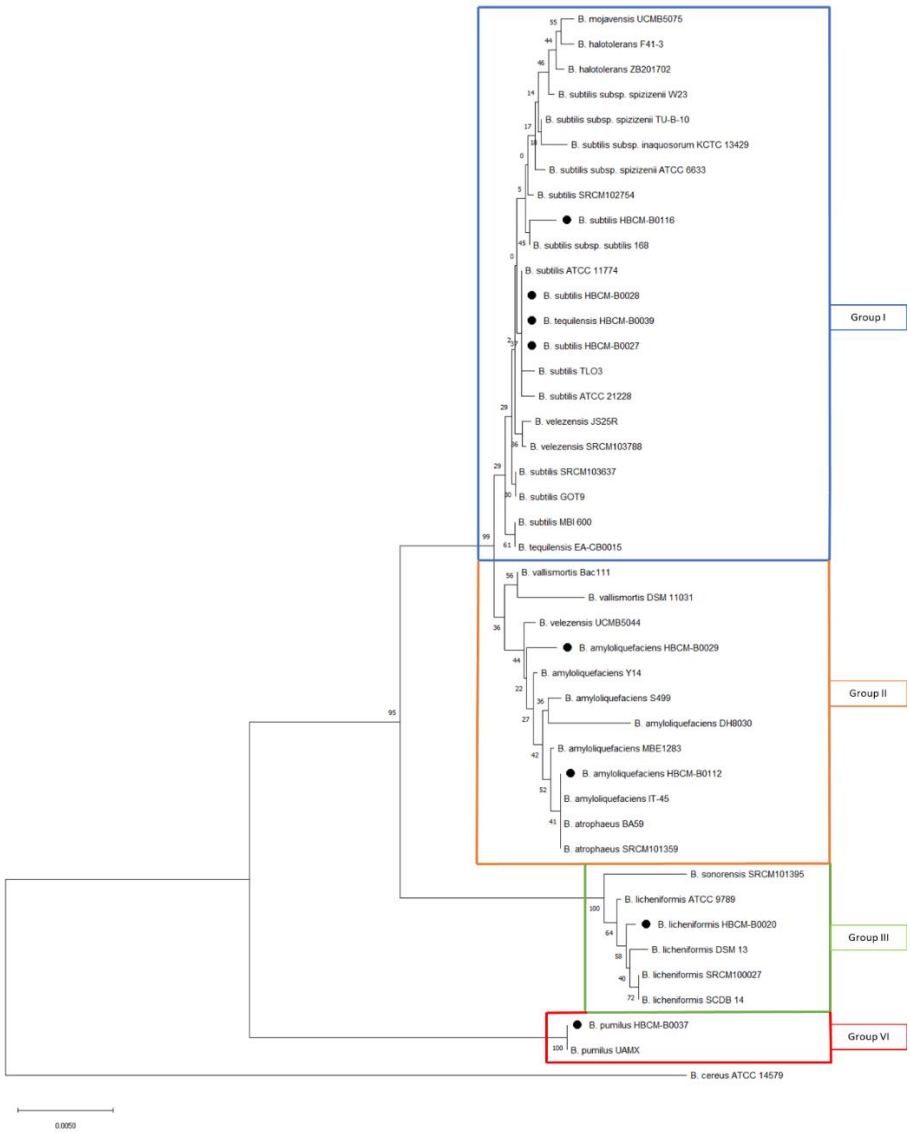

Figure 4. Phylogenetic tree of *Bacillus subtilis* species complex using 16S rRNA sequence analyzed by Neighbor-Joining method. *Bacillus cereus* ATCC 14579 was used as an outgroup. All phylogenetic trees were based on 1,000 replications (bootstraps). Scale bar = 0.0050 substitutions per nucleotide position. Eight strains which used for analysis are labeled with black dot
Source: Gordon et al. (1973)

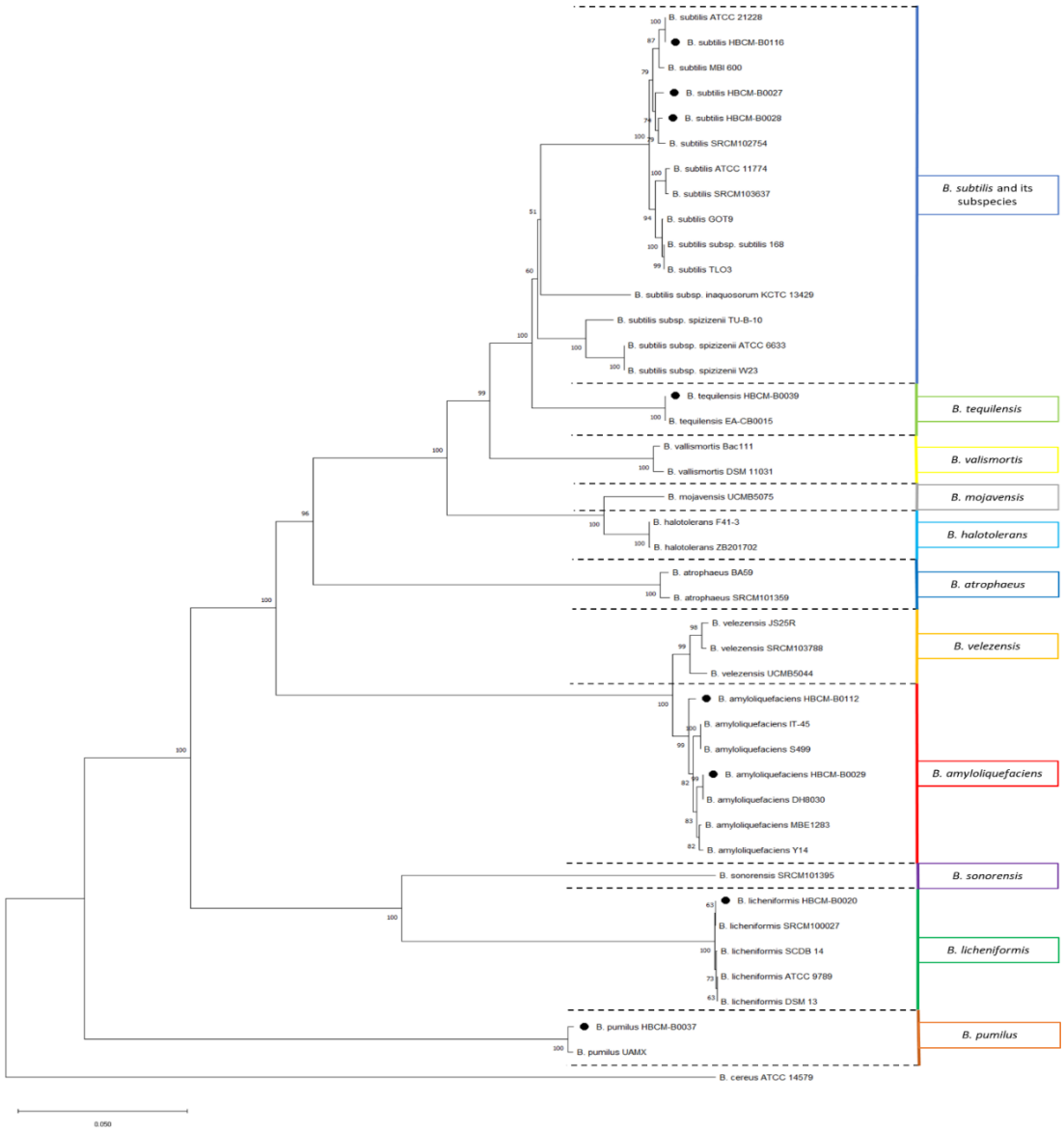

Figure 5. Phylogenetic tree of *Bacillus subtilis* species complex using MLSA scheme analyzed by Neighbor-Joining method. *Bacillus cereus* ATCC 14579 was used as an outgroup. All phylogenetic trees were based on 1,000 replications (bootstraps). Scale bar = 0.050 substitutions per nucleotide position. Eight strains dwhich used for analysis are labeled with black dot

Source: Gordon et al. (1973)

Table 3

Characteristics of the 16S rRNA gene, single housekeeping gene and the concatenated genes from all strains

Locus	Length (approximate)	Parsimony informative site		GC content (%)	Similarities (%)
		No.	%		
16S rRNA	1475	63	4.27	54.9	96.95 - 100
glpF	820	375	45.2	46.9	73.41 - 100
pta	960	329	33.8	46.4	78.42 - 100

Locus	Length (approximate)	Parsimony informative site		GC content (%)	Similarities (%)
		No.	%		
purH	1200	484	38.7	47.9	77.91 - 99.92
pycA	720	309	42.7	46	75.29 - 99.91
rpoD	1100	351	31.1	46.3	81.02 - 100
MLSA	4800	1902	38.7	47.1	78.77 - 99.66

Source: Konstantinidis and Tiedje (2005)

4. Discussion

All 16S rRNA region of eight analyzed strains is visualized under gel electrophoresis. The length was easy to determine as it goes from 1,500 base pairs to 1,600 base pairs, which match with the common length of 16S rRNA region in *Bacillus* species. All sequences are reverse complement (if needed) for the same plus strand. To maintain the right order for a further experiment, we came to the decision that the order of gene fragments concatenate would follow alphabet order. The final concatenate gene will be *glpF-pta-purH-pycA-rpoD*, and the length of the concatenate gene will be approximately 4,700 base pairs.

We suggest that within the *B. subtilis* species complex four groups can be distinguished based on the result of the 16S rRNA phylogenetic tree:

- Group (I) consisting of *B. subtilis* including its three subspecies: *B. subtilis subsp subtilis, B. subtilis subsp spizizenii*, and *B. subtilis subsp inaquosorum, B. tequilensis, B. mojavensis, B. halotolerans*
- Group (II) consisting of *B. amyloliquefaciens* and *B. vallismortis* and *B. atrophaeus*
- Group (III) consisting of *B. licheniformis* and *B. sonorensis*
- Group (IV) consisting of *B. pumilus*

B. velezensis was appeared in two group, Group (I) and Group (II).

As we had mention above, 16S rRNA gene-based studies often provides low resolution in differentiation between species that are very closely related. Many previous papers have shown that some species share over 97% or even 100% of 16S rRNA gene sequences (Das, Dash, Mangwani, Chakraborty, & Kumari, 2014). In this study, we found that there are two problems in dealing with 16S rRNA gene-based studies. The first thing is that the strains in the same species have almost 99.9% similarities between 16S rRNA gene sequences. For instance, in Figure 4, in Group (I), there are 02 strains of *B. subtilis* retrieved from NCBI (SCRM 103637 and GOT9) sharing 100% sequence similarities. This also occurred in three other groups. The second thing is 16S rRNA region between species to species also shown no significant difference. Look at Group (I) when *B. subtilis* MBI 600 shares more similarities with *B. tequilensis* EA-CB0015 than other *B. subtilis* species. Also, the sequence identity means between *B. subtilis* and *B. amyloliquefaciens* is higher than 99% (approximate 99.35%). This is a really high identity considering these two species are very hard to distinguish. Furthermore, the bootstrap value is low in many branches, the tree is not considered reliable. Combine all the observation, we can say that although 16S rRNA gene-based studies has been widely recognized for microorganisms' phylogenetic relationships identification for a very long time, it still has a downside and not good for further research to examine the accurate taxonomic position or to obtain higher resolution for phylogenetic relationships, for example, the strains of the *B. subtilis* species complex in this study.

In order to provide better resolution for relationship phylogenetic resolution, the MLSA approach has given a new path for further research and has been extensively used (Martens et al., 2008). Some species used the MLSA method based on housekeeping genes such as *Pseudomonas, Glycyrrhiza*, *Mesorhizobium* (Ampomah, Mousavi, Lindstrom, & Huss-Danell, 2017; Mulet, Lalucat, & Garcia-Valdes, 2010).

In this study, MLSA has presented some advantages over 16S rRNA gene-based studies. Although the phylogenetic tree based on MLSA shown a similar topology to the 16S rRNA genes sequences tree, the MLSA approach demonstrated more clear resolution, the branches are also longer and supported by much higher bootstrap values. Instead of the disorders species in 16S rRNA phylogenetic tree, *Bacillus* species in the MLSA approach are separated. We can easily see that 13 different species belong to the *Bacillus subtilis* species complex without any mixing. Along with that, the MLSA approach has given a new look at phylogeny relationship in the *Bacillus subtilis* species complex. Because the resolution is much clearer now, some species seem to have a different relationship, such as *B. atrophaeus* are closer to *B. subtilis* than *B. amyloliquefaciens* or *B. velezensis* now belong to other branches with *B. amyloliquefaciens.*

There are some materials that support the MLSA approach to making phylogenetic trees. The similarities between sequences in the MLSA scheme shown significantly different than sequences in 16S rRNA. In 16S rRNA, some sequences share 100% similarities, which is make researchers harder to distinguish new strains or even new species if not conduct further research. Parsimony informative sites also state that 16S rRNA gene-based studies are inadequate to discriminate closely related strains and species with only 4.27%. MLSA approach with 38.7% parsimony informative site is much more supported for better resolution in phylogeny. A phylogenetic tree of each housekeeping gene was also conducted to see if one housekeeping gene can also show better resolution than 16S rRNA. All five trees (Supplementary Figure 1 to Supplementary Figure 5) demonstrated a similar topology tree to the MLSA tree, but bootstrap values are lower, with some species still showning high similarities in sequences. The order of phylogenetic species is also not homologous with the MLSA tree, for instance, the phylogenetic tree of *pta* genes (Supplementary Figure 2) and *purH* genes (Supplementary Figure 5). Despite many reasons to support the MLSA approach, there still cannot completely reject 16S rRNA gene-base studies. GC content of 16S rRNA is higher than MLSA approach, which means 16S rRNA sequences are more likely conserved than MLSA sequences, also higher GC content contains more information about phylogenetics although in this study GC content of 16S rRNA and MLSA are considered low (< 60%). Besides that, there is something in the MLSA tree that strongly consider. B. subtilis since 2009 (Nakamura et al., 1999) as been *B. subtilis* has been subdivided into three subspecies: *B. subtilis* subsp. *subtilis, B. subtilis* subsp. *spizizenii* and *B. subtilis* subsp. *inaquosorum.* This study shown something otherwise. *B. subtilis* subsp. *subtilis* and *B. subtilis* subsp. *inaquosorum* are still in the same branch with *B. subtilis,* but MLSA approach has separated *B. subtilis* subsp. *spizizenii* into new branches, suggesting the species is not subspecies of *B. subtilis* anymore. Unfortunately, our study cannot provide enough materials to strongly prove that hypothesis, so it is still the hypothesis for considering the MLSA can give new look for phylogeny analysis.

MLSA is a practical, easy-to-do, and reliable technique for replacing 16S rRNA gene-based studies for better phylogenetic relationship resolution of bacteria species (Glaeser & Kämpfer, 2015) it still has some disadvantages. The MLSA approach does not based on one particular canonical scheme. Many housekeeping genes from different strains were used and with analytical methodologies make it difficult to unify one scheme that can access to all bacteria genera. In the formation about a public database combination are limited (Liu, Lai, & Shao, 2017), the study about genome sequencing of Dddh AND and will contribute to compensate for the shortcomings in the MLSA method (Hahnke et al., 2016).

5. Conclusion

In this study, we successfully described the better resolution and differentiation of species within the *B. subtilis* species complex, including 08 strains from four species (*B. subtilis*, *B. amyloliquefaciens*, *B. licheniformis*, and *B. pumilus*) were provided by Biotechnology Center of Ho Chi Minh City. A multi locus sequence analysis (MLSA approach with the concatenation of five protein-coding housekeeping genes: glpF, purH, pycA and rpoD) was established and performed on the *Bacillus* genus, MLSA scheme shown more advantages upon 16S rRNA gene-based studies.

Funding

This research received funding sponsored by the Biotechnology Center of Ho Chi Minh City, Vietnam.

ACKNOWLEDGMENTS

We are very grateful for the eight *Bacillus* strains used in this study from the Microbiology Department from Biotechnology Center of Ho Chi Minh City.

References

Ampomah, O. Y., Mousavi, S. A., Lindstrom, K., & Huss-Danell, K. (2017). Diverse *Mesorhizobium* bacteria nodulate native *Astragalus* and *Oxytropis* in arctic and subarctic areas in Eurasia. *Systematic and Applied Microbiology, 40*(1), 51-58. doi:10.1016/j.syapm.2016.11.004

Bacon, C. W., & Hinton, D. M. (2002). Endophytic and biological control potential of *Bacillus mojavensis* and related species. *Biological Control, 23*(3), 274-284. doi:10.1006/bcon.2001.1016

Cohn, F. (1872). Untersuchungen über Bakterien. *Beitrage zur Biologie der Pflanzen Heft 2, 1*(2), 127-224.

Das, S., Dash, H. R., Mangwani, N., Chakraborty, J., & Kumari, S. (2014). Understanding molecular identification and polyphasic taxonomic approaches for genetic relatedness and phylogenetic relationships of microorganisms. *Journal Microbiological Methods, 103,* 80-100. doi:10.1016/j.mimet.2014.05.013

Ehrenberg, C. G. (1835). Physikalische abhandlungen der koeniglichen akademie der wissenschaften zu berlin aus den jahren 1833-1835. *Königlich Bayerische Akademie,* 145-336.

Elkins, K. M. (2013). DNA extraction. *Forensic DNA Biology*, 39-52. doi:10.1016/B978-0-12-394585-3.00004-3

Fritze, D. (2004). Taxonomy of the genus Bacillus and related genera: The aerobic endospore-forming bacteria. *Phytopathology, 94*(11), 1245-1248. doi:10.1094/PHYTO.2004.94.11.1245

Gatson, J. W., Benz, B. F., Chandrasekaran, C., Satomi, M., Venkateswaran, K., & Hart, M. E. (2006). *Bacillus tequilensis* sp. nov., isolated from 2000-year-old Mexican shaft-tomb, is closely related to *Bacillus subtilis*. *International Journal of Systematic Bacteriology, 56*(7), 1475-1484. doi:10.1099/ijs.0.63946-0

Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., ... Swings, J. (2005). Opinion: Re-evaluating prokaryotic species. *Nature Review Microbiology, 3*(11), 733-739. doi:10.1038/nrmicro1236

Glaeser, S. P., & Kämpfer, P. (2015). Multilocus sequence analysis (MLSA) in prokaryotic taxonomy. *Systematic and Applied Microbiology, 38*(4), 237-245. doi:10.1016/j.syapm.2015.03.007

Gordon, R. E., Haynes, W. C., & Pang, C. H.-N. (1973). *The genus bacillus*. Washington, D.C.: United States Department of Agriculture.

Hahnke, R. L., Meier-Kolthoff, J. P., Garcia-Lopez, M., Mukherjee, S., Huntemann, M., Ivanova, N. N., … Goker, M. (2016). Genome-based taxonomic classification of *bacteroidetes*. *Frontiers in Microbiology, 7*, Article 2003. doi:10.3389/fmicb.2016.02003

Kitahara, K., & Miyazaki, K. (2013). Revisiting bacterial phylogeny: Natural and experimental evidence for horizontal gene transfer of 16S rRNA. *Mobile Genetic Element, 3*(1), Article e24210. doi:10.4161/mge.24210

Konstantinidis, K. T., & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America, 102*(7), 2567-2572. doi:10.1073/pnas.0409727102

Liu, Y., Lai, Q., & Shao, Z. (2017). A multilocus sequence analysis scheme for phylogeny of thioclava bacteria and proposal of two novel species. *Frontiers in Microbiology, 8*, Article 1321. doi:10.3389/fmicb.2017.01321

Liu, Y., Lai, Q., Dong, C., Sun, F., Wang, L., Li, G., & Shao, Z. (2013). Phylogenetic diversity of the *Bacillus pumilus* group and the marine ecotype revealed by multilocus sequence analysis. *PLoS ONE, 8*(11), Article 80097. doi:10.1371/journal.pone.0080097

Liu, Y., Lai, Q., Göker, M., Meier-Kolthoff, J. P., Wang, M., Sun, Y., … Shao, Z. (2015). Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Scientific Reports, 5*(1), Article 14082. doi:10.1038/srep14082

Macheras, E., Roux, A. L., Bastian, S., Leão, S. C., Palaci, M., Sivadon-Tardy, V., … Heym, B. (2011). Multilocus sequence analysis and rpoB sequencing of Mycobacterium abscessus (sensu lato) strains. *Journal Clinical Microbiology, 49*(2), 491-499. doi:10.1128/JCM.01274-10

Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P., & Willems, A. (2008). Advantages of multilocus sequence analysis for taxonomic studies: A case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *International Journal of Systematic Bacteriology, 58*(1), 200-214. doi:10.1099/ijs.0.65392-0

Mulet, M., Lalucat, J., & Garcia-Valdes, E. (2010). DNA sequence-based analysis of the *Pseudomonas* species. *Environmental Microbiology, 12*(6), 1513-1530. doi:10.1111/j.1462-2920.2010.02181

Nakamura, L. K. (1989). Taxonomic relationship of black-pigmented *Bacillus subtilis* strains and a proposal for *Bacillus atrophaeus* sp. nov. *International Journal of Systematic Bacteriology, 39*(3), 295-300. doi:10.1099/00207713-39-3-295

Nakamura, L. K., Roberts, M. S., & Cohan, F. M. (1999). Relationship of *Bacillus subtilis* clades associated with strains 168 and W23: A proposal for *Bacillus subtilis* subsp. *subtilis* subsp. nov. and *Bacillus subtilis* subsp. *spizizenii subsp*. nov. *International Journal of Systematic Bacteriology, 49*(3), 1211-1215. doi:10.1099/00207713-49-3-1211

Palmisano, M. M., Nakamura, L. K., Duncan, K. E., Istock, C. A., & Cohan, F. M. (2001). *Bacillus sonorensis* sp. nov., a close relative of *B. licheniformis*, isolated from soil in the Sonoran Desert, Arizona. *International Journal of Systematic Bacteriology, 51*(5), 1671-1679. doi:10.1099/00207713-51-5-1671

Peeters, C., Meier-Kolthoff, J. P., Verheyde, B., De Brandt, E., Cooper, V. S., & Vandamme, P. (2016). Phylogenomic study of *Burkholderia glathei*-like organisms, proposal of 13 novel *Burkholderia* species and emended descriptions of *Burkholderia sordidicola, Burkholderia zhejiangensis,* and *Burkholderia grimmiae*. *Frontier in Microbiology, 7*, Article 877. doi:10.3389/fmicb.2016.00877

Priest, F. G., Goodfellow, M., Shute, L. A., & Berkeley, W. (1987). *Bacillus amyloliquefaciens* sp. nov., nom. rev. *International Journal of Systematic Bacteriology, 37*(1), 69-71. doi:10.1099/00207713-37-1-69

Roberts, M. S., Nakamura, L. K., & Cohan, F. M. (1994). *Bacillus mojavensis* sp. nov., distinguishable from *Bacillus subtilis* by sexual isolation, divergence in DNA sequence, and differences in fatty acid composition. *International Journal of Systematic Bacteriology, 44*(2), 256-264. doi:10.1099/00207713-44-2-256

Roberts, M. S., Nakamura, L. K., & Cohan, F. M. (1996). *Bacillus vallismortis* sp. nov., a close relative of *Bacillus subtilis*, isolated from soil in death valley, California. *International Journal of Systematic Bacteriology, 46*(2), 470-475. doi:10.1099/00207713-46-2-470

Rooney, A. P., Price, N. P., Ehrhardt, C., Swezey, J. L., & Bannan, J. D. (2009). Phylogeny and molecular taxonomy of the *Bacillus subtilis* species complex and description of *Bacillus subtilis* subsp. *inaquosorum* subsp. nov. *International Journal of Systematic Bacteriology, 59*(10), 2420-2436. doi:10.1099/ijs.0.009126-0

Ruiz-García, C., Béjar, V., Martinez-Checa, F., Llamas, I., & Quesada, E. (2005). *Bacillus velezensis* sp. nov., a surfactant-producing bacterium isolated from the river Velez in Malaga, southern Spain. *International Journal of Systematic Bacteriology, 55*(1), 191-195. doi:10.1099/ijs.0.63310-0

Skerman, V. B. D., McGowan, V., & Sneath, P. H. A. (Eds.). (1980). Approved lists of bacterial names. *International Journal of Systematic Bacteriology, 30*(3), 225-420. doi:10.1099/00207713-30-1-225

Timilsina, S., Jibrin, M. O., Potnis, N., Minsavage, G. V., Kebede, M., Schwartz, A., … Goss, E. M. (2015). Multilocus sequence analysis of xanthomonads causing bacterial spot of tomato and pepper plants reveals strains generated by recombination among species and recent global spread of Xanthomonas gardneri. *Applied Enviroment Microbiology, 81*(4), 1520-1529. doi:10.1128/AEM.03000-14

Turnbull, P. C. B., & Baron, S. (Eds.) (1996). *Medical microbiology* (4th ed.). Galveston, TX: University of Texas Medical Branch at Galveston.

Vinuesa, P. (2010). Multilocus sequence analysis and bacterial species phylogeny. In O. Aharon & R. T. Papke (Eds.), *Molecular phylogeny of microorganisms* (pp. 45-64). Norfolk, VA: Caister Academic Press.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K. H., … Amann, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Review Microbiology, 2*(9), 635-645. doi:10.1038/nrmicro3330