

## Phylogenetic Tree Analysis for Bacterial Identification by BALST and MEGA 5.2 Software

Sabai Thein<sup>1</sup>, Ohn Maung<sup>2</sup>

### Abstract

The identification of bacteria is the most important technique for microbiologists. The traditional methods as morphological and biochemical ways were not enough to identification of bacteria. The advance tools were modified to identify the isolated bacteria. The principal method of bacterial identification should attempt PCR product of 16S RNA gene. The accuracy of phylogenetic analysis was improved by using the evolution of computer science. These results were examined in BLAST on webpage and the final data were check and align in MEGA software. In this research work, six bacterial strains were extracted DNA, run PCR, checked gel electrophoresis, identified by BLAST and align data in MEGA. According to phylogenetic analysis, the isolated six bacteria strains were observed as *Bacillus* sp.

**Keywords:** PCR, computer science, BLAST, MEGA, *Bacillus* sp.

### Introduction

The identification and classification of microorganisms such as bacteria, fungi are based on morphological characters such as shape, gram staining and color. Moreover, the biochemical reactions are checked as aerobic/anaerobic, enzyme production and other chemical reactions. The initial molecular techniques for bacterial identification are GC content [1], plasmid profiling, and compatibility to genetic transformation [2]. At present time, the fundamental molecular applications are greatly applied in bacterial detection and identification. These techniques involved hybridization and nucleotide sequencing. The hybridization methods are Southern [3], PCR [4], real-time PCR [5], microarray [6], universal tagging method [7]. For the microbial detection and identification, hybridization techniques are precise and specific way.

The presence and absence of target microorganisms and specific characters such as antibiotic resistance and virulence factors were detected by using specific PCR primers. The detection and identification of non-target bacteria require for another approaches. For the taxonomic analysis of different bacteria are investigated by using degenerate primers, 16SrDNA [8].

Homologs, the sequences similarities of organisms are maintained from a common ancestor. The phylogenetic analysis is available using the homolog sequences. Homologs were detected from DNA or protein database by homology search programs such as BLAST [9]. The search program, BLAST (Basic Local Alignment Search Tool) is the most frequently used for the calculating sequence similarity. This program can be found on NCBI (National Center for Biotechnology Information) webpage. BLAST, which is based on pairwise sequence alignment, uses approximate calculation for speeding up, because the query sequence is compared to huge amount of sequences in the database. First, BLAST searches for “words” shared by the query sequence and a database sequence. The default word length for nucleotide is 11, and for amino acid is 3. For proteins, BLAST searches not only identical oligopeptides but also similar oligopeptides, and adds them to the words list [10].

Three or more biological sequences including protein, DNA, or RNA are aligned by a Multiple Sequence Alignment (MSA). It calculates the best match for the selected sequences and these sequences are lined up for identities, similarities and the differences between them. The input sequences share a lineage and are descended from a common ancestor according to the evolutionary relationship. The programs as ClustalW, ClustalX, Clustal Omega, MAFFT, and MUSCLE are several popular programs for MSA [10]. MEGA 7 (Kumar S *et al.*, 2016) which includes MUSCLE is also popular software because it has a user-friendly interface and it is available to infer phylogenetic tree.

In this research, the extraction of bacterial DNA, running PCR process and checked by BLAST and align by MEGA software. The final results were expressed as four bacterial strains as *Bacillus* sp and two strains as *Lysinibacillus* sp.

---

<sup>1</sup>Assistance Lecturer, Department of Botany, Taunggoke Degree College

<sup>2</sup>Professor and Head, Department of Botany, Taunggoke Degree College

## Materials and Methods

### 2.1. Extraction of Bacterial DNA

For phylogenic analysis of bacteria, bacterial DNA extraction was prepared by using TIANamp Bacterial DNA Kit (TIANGEN BIOTECH (BEIJING) CO., LTD.)

### 2.2. PCR amplification and sequencing (16S rRNA gene)

The nucleotide sequences of selected bacterial 16S rRNA genes were performed by using universal primer set 27f (5'-GTTT GATC MTGGCTC AG-3') and 1492R (5'-TACGGYTACCTT GTT ACGACTT -3') (Lane, 1991) at 94°C – 3 min, 94°C – 30 sec, 54°C – 30 sec, 72°C – 2 min, 72°C – 7 min and 16°C – 2 min (35 cycles).

### 2.3. Nucleotide sequences analysis

In this study, the nucleotide sequences primary analysis was determined by NCBI Blast and CLUSTALW. BioEdit Sequence Alignment Editor and MEGA 5.2 software that are used to confirmed and rearrange the sequences. MEGA 5.2 software was utilized for the construction of the selected bacterial phylogenetic analysis (Lane, 1991).

## Results

### 3.1 Phylogenetic Tree Analysis of Bacteria

Firstly, the DNA results from sequencing company were prepared as text document (Figure. 1 and 2).

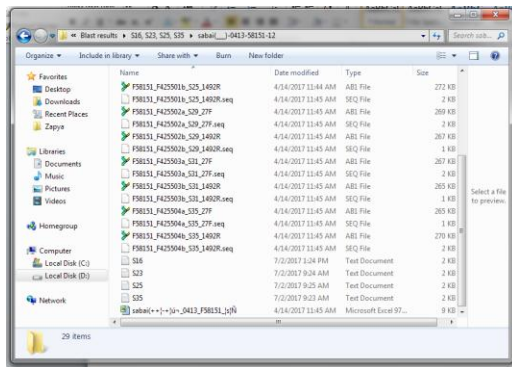


Figure 1. DNA sequence results files

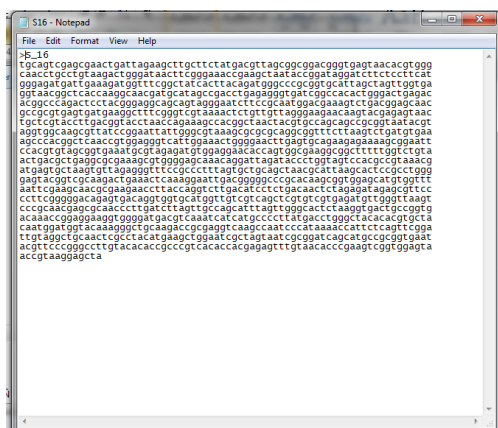


Figure 2. DNA sequence results Notepad file

### 3.2. NCBI webpage and BLAST

Secondly, open NCBI site and Click BLAST (Figure 3).

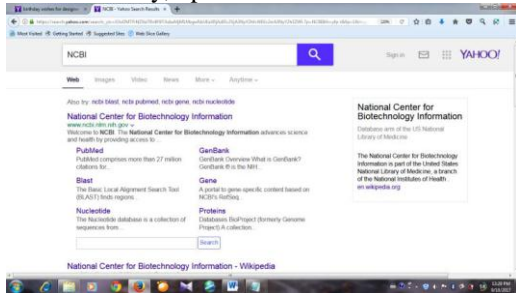


Figure 3. NCBI and BLAST site

### 3.3. Preparing text in BLAST

Copy the text document result and paste on Enter accession number (Figure 4).

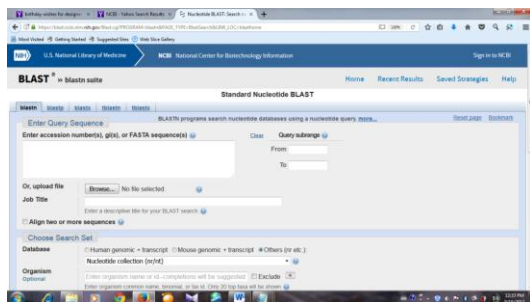


Figure 4. Preparing text in BLAST

### 3.4. BLAST results after running

The blast results can be seen after running 10 to 30 seconds (Figure 5).

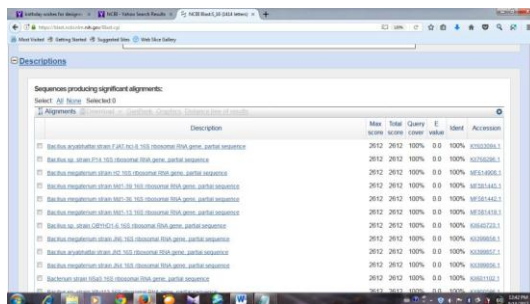


Figure 5. BLAST results from NCBI

### 3.5. Prepare the results as a text file

The BLAST results were prepared as a text document for the next steps (Figure 6).

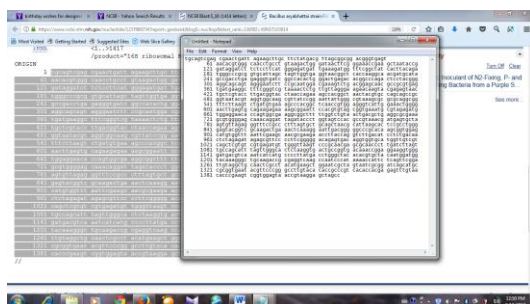


Figure 6. BLAST results prepared as text document

### 3.6. Prepare the results in one page

Both of your results and references should prepare in one page (Figure 7).

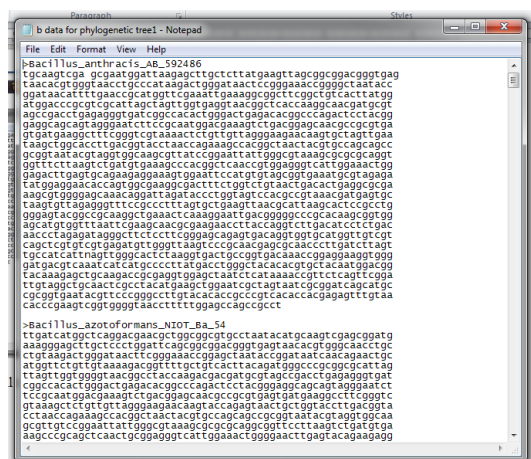


Figure 7. Preparing both results in one page

### 3.7. Open Mega 5.2 Application.

This application should open and click Align> Edit/Build Alignment (Figure 8).

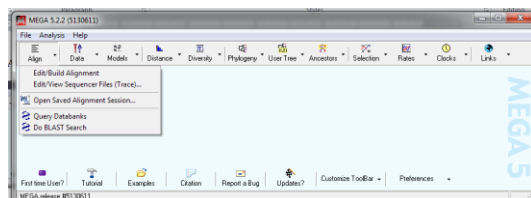


Figure 8. Using MEGA 5.2

### 3.8. Create a new Alignment

A new Alignment was create and click DNA button (Figure 9).

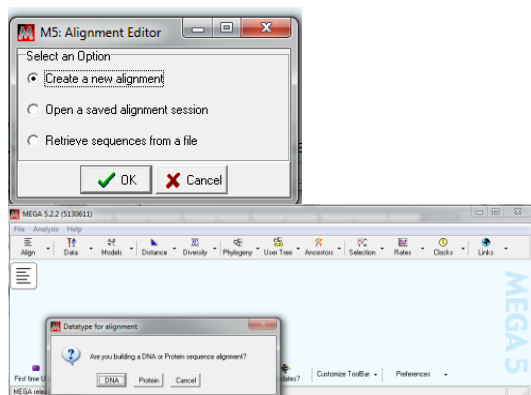


Figure 9. Preparing a new alignment

### 3.9. Using M5 Alignment Explorer

M5 Alignment Explorer was appeared and run as Edit> Insert Sequence from File and also should choose edit file (.txt) and save this file as (.mas) (Figure 10).

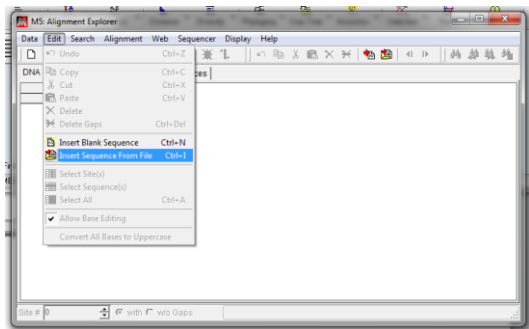


Figure 10. Preparing M5 Alignment

### 3.10. Running ClustalW Progress

When running the ClustalW progress, Alignment> Align by ClustalW and appear M5 ClustalW Parameters (Figure 11).

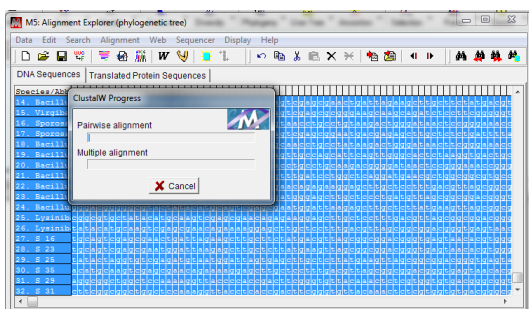


Figure 11. Running ClustalW progress

### 3.10. Constructing Neighbor-Joining Tree

The constructing the similar sequences as Mega 5.2.2 > Phylogeny> Construct/ Text Neighbor-Joining Tree and save the file (.mas) (Figure 12).

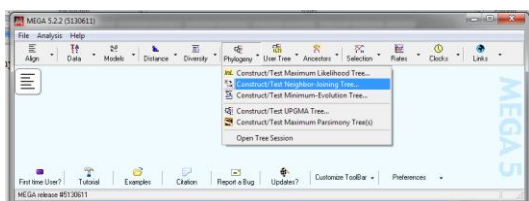


Figure 12. Constructing Neighbor-Joining Tree

### 3.11. Computing the M5 Preferences

The computing M5 Analysis Preferences was used, clicked the Compute Button and M5 Tree Explorer Result was appeared. The result was copied and paste on MS word (Figure 13, 14, 15).

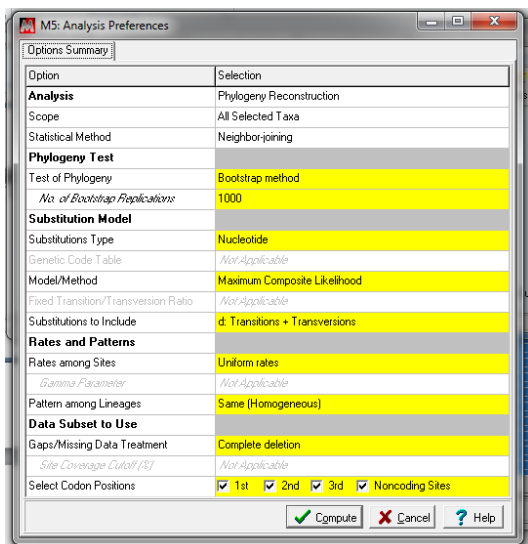


Figure 13. Computing the M5 Preferences

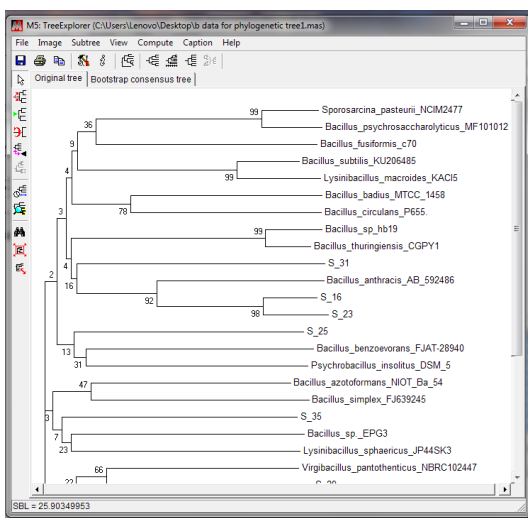


Figure 14. M5 Tree Explorer Results

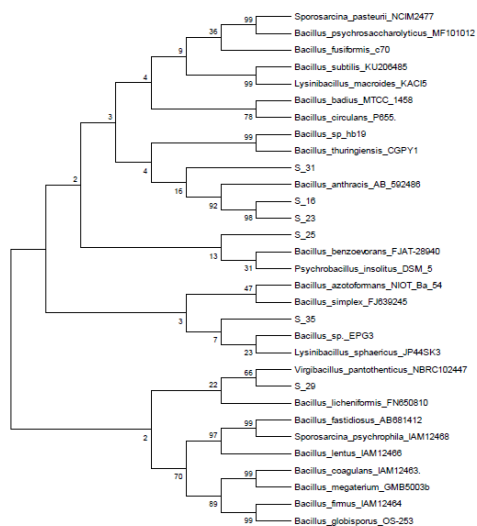


Figure 15. Final Results of M5 Tree Explorer

## Discussion and Conclusion

In this research, the extraction of bacteria DNA, running PCR and arranging computerized based phylogenetic tree were applied. The final phylogenetic tree expressed the three strains (S16, S23, S3) as *Bacillus anthracis*; one strain (S25) as *Bacillus benzoevorans*; (S29) as *Virgibacillus pantothenicus* and the rest one (S35) as *Bacillus* sp. So, all the selected bacteria were identified as *Bacillus* species.

The previous identification methods as morphological and biochemical methods were support as approximately genus level of bacteria. The advance techniques as DNA extraction, primer application and computer sciences based examinations were expressed more accurate bacterial identification. The most reliable bacterial identification is the integration of morphological methods with molecular applications. These predications were more accurate and reasonable results for the bacterial identification.

## ACKNOWLEDGMENTS

I would like to really thank Prof. Zonghua Wang and Prof. Hong Liu (Fujian Agriculture and Forestry University, China) their kind suggestion and support for this research.

## References

- [1] Sigal, N., Senez, J. C., Le Gall, J. and Sebald, M., "Base composition of the deoxyribonucleic acid of sulfate-reducing bacteria", *J Bacteriol*, 85:, 1963, pp. 1315–1318.
- [2] Perry, D. and Slade, H., "Transformation of streptococci to streptomycin resistance", *J Bacteriol* 83: 1961, pp 443–449
- [3] Nilsson, H., Taneera, J., Castedal, M., Glatz, E., Olsson, R. and Wadstro"m, T., "Identification of *Helicobacter pylori* and Other *Helicobacter* Species by PCR, hybridization and partial DNA sequencing in human liver samples from patients with primary sclerosing cholangitis or primary biliary cirrhosis", *J Clin Microbiol* 38:, 2000, pp 1072–1076.
- [4] Owen, R. H., Dymock, D., Booth, V., Weightman, A. J. and Wade, W. G., "Detection of unculturable bacteria in periodontal health and disease by PCR", *J Clin Microbiol* 37: 1999, pp 1469–1473
- [5] Richards, G. P., Watson, M. A., Fankhauser, R. L. and Monroe, S. S., "Genogroup I and II noroviruses detected in stool samples by realtime reverse transcription-PCR using highly degenerate universal primers", *Appl Envir Microbiol* 70:, 2004, pp 7179–7184
- [6] Mitterer, G., Huber, M., Leidinger, E., Kirisits, C., Lubitz, W., Mueller, M. W. and Schmidt, W. M., "Microarray-based identification of bacteria in clinical samples by solid-phase PCR amplification of 23S ribosomal DNA sequences", *J Clin Microbiol* 42:, 2004, pp 1048–1057
- [7] Yasuoka, M. O., "A multiplex polymerase chain reaction– based diagnostic method for bacterial vaginosis", *Obst Gynecol* 100:, 2004, 759–764
- [8] Verma, A., Sampla, A. K. and Tyagi, J. S., "Mycobacterium tuberculosis *rrn* promoters: differential usage and growth rate-dependent control", *J Bacteriol* 181: 1999, pp 4326–4333
- [9] Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.*, 25: 1997, 3389-3402.
- [10] Horiike, T., "An Introduction to Molecular Phylogenetic Analysis, Reviews in Agricultural Science, 4: , 2016, pp 36- 45, Doi: 10.7831/ras.4.36
- [11] Kumar, S., Stecher, G. and Tamura, K., "MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets", *Mol Biol Evol.* 33:, 2016, pp 1870-1874.
- [12] Lane, D. J., "16S/23S rRNA sequencing. Nucleic Acid Techniques in Bacterial Systematics", Eds Stackebrandt, E., Goodfellow, M. N.Y.: Wiley. 1991, pp 115–175.