Austin Journal of Proteomics, Bioinformatics & Genomics

Open Access

Austin Publishing Group

Research Article

The Amino Acids' Profiles for Searching Proteins Similarity: Bacitracin Synthase as a Model

Amara AA* and Sharaf MM

Protein Research Department, Genetic Engineering and Biotechnology Research Institute, City for Scientific Research and Technological Applications, Universities and Research Center district, Egypt

***Corresponding author:** Amro Abd Al Fattah Amara, Protein Research Department, Genetic Engineering and Biotechnology Research Institute, City for Scientific Research and Technological Applications, Universities and Research Center district, Alexandria, Egypt

Received: May 22, 2017; **Accepted:** June 20, 2017; **Published:** June 27, 2017

Abstract

The information provided by protein and nucleotide sequence database have flourished science and research. The scientific applications of biotechnology and molecular biology flourished and developed intensively in the last twenty years, due to the availability of such information. The main data input were usually DNA sequences, which were then subjected to translation using software to their related protein sequences. One missed information and its database is the amino acids % profile, which depend on the twenty amino acids sequences % rather than the amino acid names of the protein's backbone. In this study, in lab amino acid profile for Bacitracin proteins were generated to investigate an upcoming research strategy to investigate the Bacitracin synthases (or synthetases) based on their amino acids profile. 100 Bacitracin synthase sequences and other homologous proteins were collected. Fourteen highly conserved regions of Baci-tracin synthases were used to generate a hypothetical protein sequence. The non Bacitracin synthases sequences (in the 100 aa sequences) were used even –with high homology- to evaluate the efficacy of the design. The 100 aa sequences were aligned, unsuitable sequences were removed and 94 aa sequences were selected. The sequences were shortened to remove any unsuitable equality in the sequence alignment. Further clustering was generated. The best alignment was obtained, the hypothetical protein was added, the 95 aa sequences were realigned and the phylogenic tree was generated. The 95 aa sequences then are-summarized as amino acid % profiles using Mega 6 software. The amino acids % were then clustered again using Past 3 statistic software. The hypothetical protein matched with the same sequences in both trees obtained from the protein sequence alignments and the amino acid profiles. This study suggests using the amino acid % profile as a simpler candidate in searching for a particular protein. This tool could be used in lab, or could be a starting point for installing a new database.

Keywords: Amino acids % profile; Bacitracin; Database

Introduction

Probiotics and healthy microflora play an impotent role in human health, thus research in those two microbial groups is essential. Beneficial microbes means "let good microbes work for us in different fields, get their benefits and take a rest" [1,2]. Additionally, "let them fight for us against bad microbes and be safe". How do humans benefit from microbes? How do they defend pathogens in our body? Such questions were answered in previous studies [1]. Other questions can be raised based on our opportunistic use for microbes. How could we discover more helpful microbes by using the information we have about them? Bacitracin producing microbes are good candidates for such research and will be addressed in this study.

Many researches are proving fact that our microflora supports our health. Intestinal microflora had metabolic functions, such as fermenting indigestible dietary residues and endogenous mucus, saving of energy, production of vitamin K, and absorption of ions [3]. Probiotics have roles in epithelial cell proliferation and differentiation, and the development and the homeostasis of the immune system [4-6]. Probiotics are not an invention but existed in our traditional foods such as beverages, salty fishes, yogurt, different types of cheeses and so on since olden times [7,8]. Probiotics can be defined as living microorganisms administered in an adequate number that continue to exist in the intestinal bionetwork, to perform a health positive effect [9]. Probiotics as a term was first used by Lilly and Stillwell (1965) to describe the `substances secreted by one microorganism that stimulate the growth of another' [10]. Parker (1974), proposed that Probiotics are `organisms which contribute to intestinal microbial balance' [11].

Nowadays information and database banks such as, amino acid sequence of protein is essential for any biologist. Specific functions such as, Bacitracin production could be used to identify microbes has not been identified yet as friendly microflora, to be investigated and added to the information bank. Bacitracin is an antibiotic produced by some bacterial species mostly Bacillus [5,12,13]. It is a mixture of cyclic peptides. Its mode of action is based on its ability to interfere with the gram positive bacterial cell wall and peptidoglycan synthesis. It has some side effects and is able to cause kidney damage if used internally. Bacitracin is nephrotoxic. However, it has excellent superficial activity against pathogenic bacteria. It acts as a dermatological irritants and has healing activity [14,15]. It is used as ointment topically to treat eye and skin infections. Non -ointment formula is available for eye

Citation: Amara AA and Sharaf MM. The Amino Acids' Profiles for Searching Proteins Similarity: Bacitracin Synthase as a Model. Austin J Proteomics Bioinform & Genomics. 2017; 4(1): 1020.

Table 1: The different sequences (100) used in this study and the microbe names.

***Bacillus cereus* VD014**	(gb|EJR21234.1), (emb|SCC54161.1), (emb|SCC54179.1), (ref|WP_002107052.1), (ref|WP_002107057.1), (ref|WP_002164124.1), (ref|WP_016118418.1) (ref|WP_016118420.1), (ref|WP_033693014.1), (ref|WP_033693015.1), (ref|WP_048529241.1) (ref|WP_048529245.1), (ref|WP_048535930.1), (ref|WP_048535935.1), (ref|WP_052767462.1) (ref|WP_074594549.1), (ref|WP_074594551.1) and (ref|WP_080345776.1)
***Bacillus licheniformis* LMG 7559**	(gb|KUL05994.1), (gb|KND05497.1), (ref|WP_043927787.1) (ref|WP_059231853.1), (ref|WP_075213156.1), (ref|WP_075213158.1), (ref|WP_081049785.1) (ref|WP_081376395.1), (sp|O68006.1)and (sp|O68008.1)
Bacillus methylotrophicus	(gb|KJR70738.1), (ref|WP_006636543.1) and (ref|WP_006637435.1)
Bacillus sonorensis	(ref|WP_029419408.1)
***Bacillus sp*. CPSM8**	(ref|WP_049936870.1)
***Bacillus sp*. GO-13**	(ref|WP_065894575.1), (ref|WP_065894577.1)
***Bacillus sp*. KJ-16**	(gb|OMI07176.1), (gb|OPF70356.1), (gb|OPF70757.1), (ref|WP_020452077.1) (ref|WP_020452079.1), (ref|WP_035338352.1), (ref|WP_057957861.1), (ref|WP_057957862.1) (ref|WP_065643262.1), (ref|WP_065643264.1), (ref|WP_075752472.1), (ref|WP_080131435.1) (ref|WP_080131437.1)
***Bacillus sp*. MSP5.4**	(ref|WP_039072838.1)
***Bacillus sp*. N35-10-2**	(ref|WP_071758549.1)
***Bacillus sp*. NRRL B-41282**	(ref|WP_076797459.1)
***Bacillus sp*. NRRL B-41294**	(ref|WP_076758518.1)
***Bacillus sp*. NRRL B-41327**	(ref|WP_076793436.1)
***Bacillus sp*. SB47(**	ref|WP_026579340.1)
***Bacillus sp*. TH008**	(gb|KKB72147.1)
***Bacillus subtilis* group**	(ref|WP_043054383.1), (ref|WP_043054385.1)
Bacillus subtilis	(gb|ABB80123.1), (gb|ABB80125.1)
***Bacillus thuringiensis* serovar huazhongensis BGSC 4BD1**	(gb|EEM80701.1)
Bacillus thuringiensis	(ref|WP_000023653.1), (ref|WP_033703281.1), (ref|WP_061884453.1), (ref|WP_061884455.1)
Bacillus	(ref|WP_026579338.1)
Brevibacillus laterosporus	(ref|WP_003336314.1), (ref|WP_003339532.1), (ref|WP_018669922.1), (ref|WP_022584750.1), (ref|WP_031415226.1), (ref|WP_051870643.1), (ref|WP_068792335.1)
***Brevibacillus sp*. SK.DU.10**	(ref|WP_064017180.1), (ref|WP_064018610.1)
Paenibacillus alvei	(ref|WP_005546175.1), (ref|WP_021252556.1), (ref|WP_021259807.1)
Paenibacillus durus	(ref|WP_042206282.1)
Paenibacillus elgii	(gb|AFJ14794.1), (ref|WP_063177908.1)
Paenibacillus kribbensis	(ref|WP_068501588.1)
Paenibacillus pabuli	(ref|WP_062318729.1)
Paenibacillus polymyxa	(ref|WP_013310169.1), (ref|WP_034095029.1), (ref|WP_069291848.1)
***Paenibacillus sp*. 181mfcol5.1**	(emb|SEJ02277.1)
***Paenibacillus sp*. Cl6col**	(emb|SDG14332.1)
***Paenibacillus sp*. FSL R5-0765**	(ref|WP_076217152.1),
***Paenibacillus sp*. FSL R7-0273**	(ref|WP_052421498.1)
***Paenibacillus sp*. IHBB 10380**	(ref|WP_044876767.1), (ref|WP_052675792.1)
***Paenibacillus sp*. MSt1**	(gb|KEQ22994.1), (gb|KEQ27596.1), (ref|WP_036675334.1), (ref|WP_051775118.1)
***Paenibacillus sp*. NAIST15-1**	(dbj|GAV14649.1), (ref|WP_083545290.1)
***Paenibacillus sp*. OSY-SE**	(gb|AGM16412.1)
***Paenibacillus sp*. UNC217MF**	(ref|WP_036667063.1)
Paenibacillus	(ref|WP_024633946.1)
Unknown	(lcl|Query_57839)

infections. Bactoprenol phosphate Bacitracin interferes with the dephosphorylation of C55- isoprenyl pyrophosphate, also known as bactoprenol, a membrane carrier molecule that transports the building blocks of the peptidoglycan bacterial cell wall outside of the

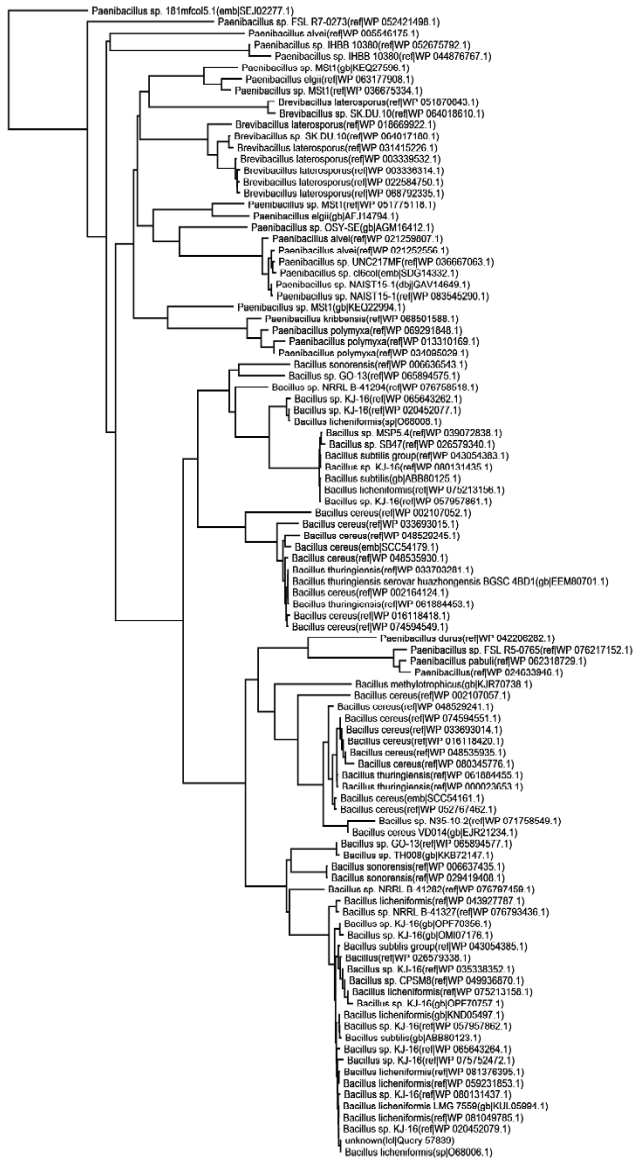

Figure 1: Phylogenic tree from Blast protein database for the 100 aligned sequences which show homology to the Bacitracin.

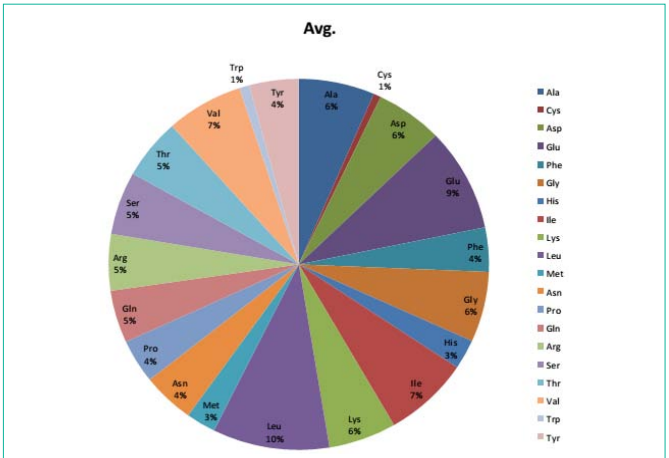

Figure 2: Amino acids of the 100 used sequence as %.

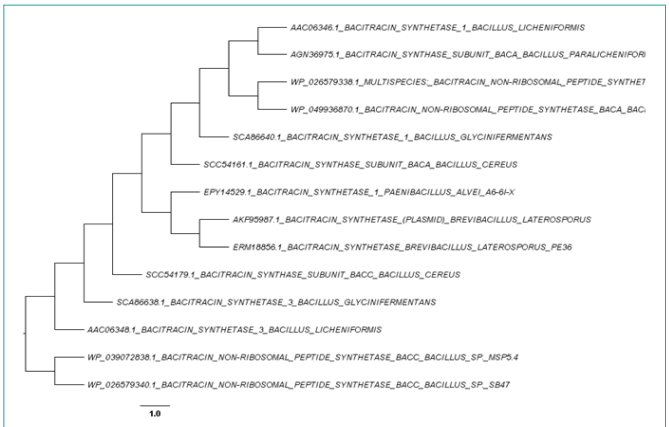

Figure 3: The phylogenic tree of the selected 14 Bacitracin synthases sequences.

inner membrane [16,17]. Bacitracin is synthesized via what is called Non-ribosomal peptide synthetases (Nrpss), meaning that ribosomes are not directly involved in its synthesis.

This study aims to encourage mini-local databases linked with the global database for proteins or protein products useful for humans. Bacitracin, which silently defends many microbes trying to colonize our body every nano-second, was used as a model for our hypothesis.

Material and Methods

Bacitracin protein start sequence

Bacitracin sequence is the starting point of this study and is represented by AAC06346.1 BACITRACIN SYNTHETASE 1 [BACILLUS LICHENIFORMIS]. The sequence was obtained from the www.ncbi.nlm.nih.gov nucleotide database and was transferred to the protein blast search. This database search for proteins using a protein query (Blast.ncbi.nlm.nih.gov/Blast.cgi) [18].

Searching protein database for similarity

The Blast protein database searches for similar sequences; the search was adjusted for 100 aa sequences results. The obtained data was saved as FASTA format and a tree was generated from the tree generation option using the ID and the species names in the Blast search. The generated files were saved for further analysis.

The FASTA file adjustment

The obtained FASTA file was adjusted manually to remove any of the long header of the source data of the obtained amino acid sequences. Simply the sequence ID, the strain and the protein names were kept. After finishing the removal of any excess data, the file was saved as FASTA format for further investigation.

Protein sequences alignment

The 100 obtained aa sequences were subjected to alignment using ClustalW option in the BioEdit ver 7.2.3. [19] as well as Mega 6 software [20] and the sequences investigated on the GeneDoc 2.7. [21]. Bioedit ver 7.2.3. was used to convert the FASTA text file to a true FASTA format which is readable by Mega 6. Mega 6 was used for the alignment of protein sequences, visualizing the sequences

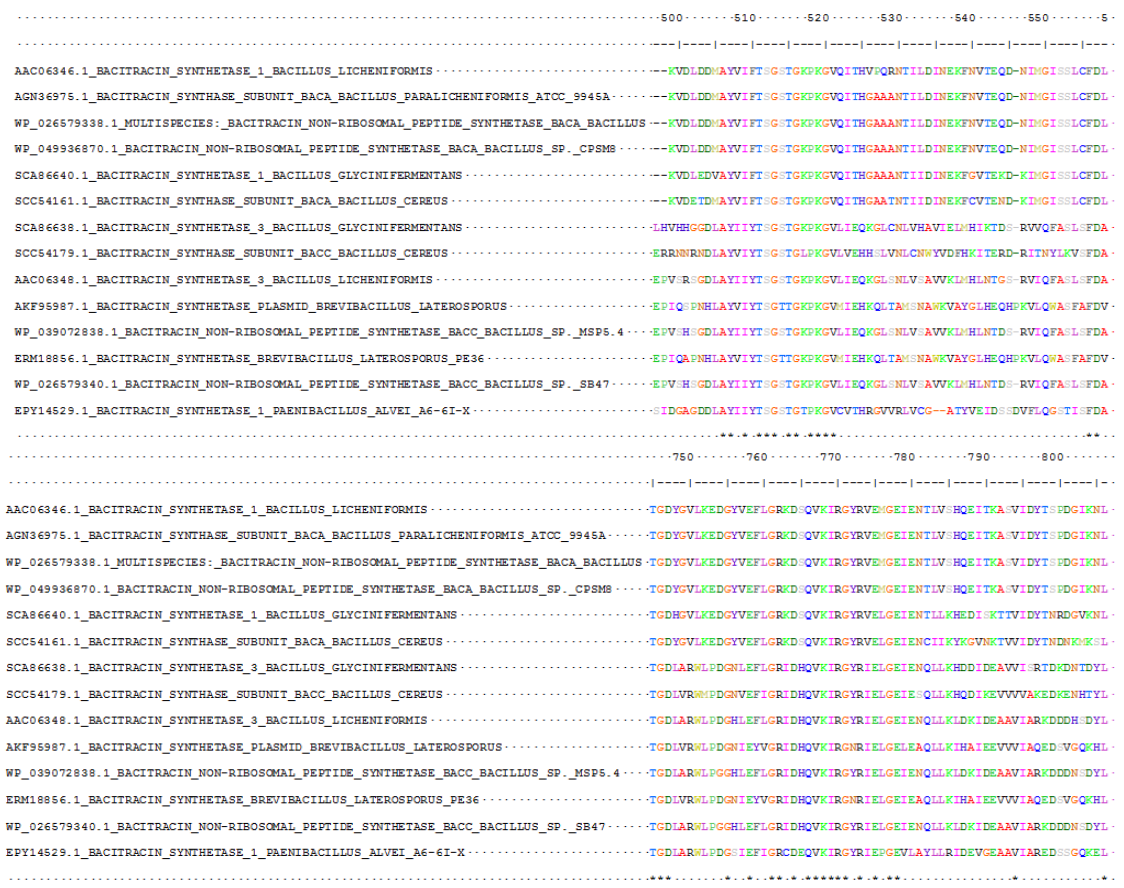

Figure 4: Part of the 14 Bacitracin synthases sequences alignment.

and for removing gaps and long/short odd sequences. The alignment and shorten 94 aa sequences were obtained (992 amino acids as recorded by Mega 6). And after the first adjustment the sequences were realigned using the MUSCLE alignment option in Mega 6 [20]. After being sure about the alignment homogeneity and identity, 14 aa Bacitracin synthase sequences were selected out of the 94, and a separate alignment and phylogenic tree generation step for them was generated.

The consensus sequence of the 14 aa sequences alignment plus the EPY14529.1 sequence was merged to construct a hypothetical protein. The constructed hypothetical protein was then inserted in the 94 aa selected and shortened sequences which include EPY14529.1. The 95 aa sequences were aligned again and the file was saved. A tree was generated for the 95 aa sequences through from the tree option in the phylogeny option in Mega 6 using construct/test Maximum Parsimony tree (s). The obtained tree was then saved as Newick file. FigTree software [22] as well as Mega 6 [20] were used to collect and visualize data from the obtained trees (even not represented).

Analysis for the amino acid ratio in the protein sequences

The amino acid composition option in Mega 6 software was used to generate an amino acid profile for the 94 selected sequences plus the generate hypothetical sequence. The obtained data was saved as Excel file for further analysis.

Cluster analysis for the amino acid % profile

The amino acid % profile for the 95 aa sequences Excel file was adjusted and only the names of the sequences , the amino acid symbols were kept. The file was then opened by Past 3 software and the data was readjusted to be suitable for the statistical analysis. Neighbor joining clustering with similarity index (Euclidean) was used. And the result was obtained as image file.

Results and Discussion

This study is aims to high lighten the importance of establishing new local and global database for the amino acids profile %. Bacitracin synthase was used as a model. Bacitracin synthase is a protein involved in the Bacitracin antibiotic production. It is well known that Bacitracin is produced by our skin microflora. Bacillus sp. Are responsible for that. By searching for Bacitracin synthases similarity, using Blast protein database against a protein sequence, the result included different Bacillus sp. (Table 1 and Figure 1 (100 aa sequences)). Their amino acids % profile obtained from Mega 6 was summarized in Figure 2. The species were clustered using the phylogenic tree option in the blast database to visualize the different proteins, which have homology with Bacitracin synthase sequence. The sequences were subjected to alignment, adjustment, shortening and readjustment as described in the above part. The sequences were reduced to 94 and their length was shortened to the best alignment

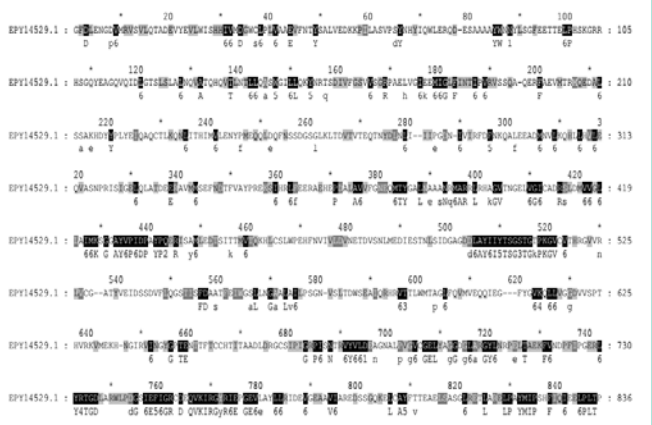

Figure 5: Figure show both of the consensus sequence and the EPY14529.1 which were used in the generation of the hypothetical sequence.

part. The sequences were realigned and a phylogenic tree was obtained. From the 94 aa sequences 14 Bacitracin synthase sequences were obtained, alignment and their phylogenic tree was generated Figure 3 and 4.

The evolutionary history was inferred using the Maximum Parsimony method (Figure 3). The most parsimonious tree with a length of 2237 is shown. The consistency index is 0.918641 (0.916667), the retention index is 0.941741 (0.941741), and the composite index is 0.865122 (0.863263) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm (pg. 126 in ref. [23]) with search level 0, in which the initial trees were obtained by the random addition of sequences (10 replicates). The analysis involved 14 amino acid sequences. All positions containing gaps and missing data were eliminated.

There were a total of 882 positions in the final dataset. Evolutionary analyses were conducted in MEGA 6 [20].

The full name of the sequences can be found in Table 2 and Figure 3 (95 aa sequences with full description). The consensus of the aligned sequences and the EPY14529.1 were used to generate a hypothetical protein as in Figure 6.

The sequence of the generated hypothetical protein is: GFDLENGPVMRVSVLQTADEVYEVLWISHHIVM DGWSLPLVAAEVFNTYSALVEDKKPILASVPDYNHYIQ WLERQDESAAAAYWNNYLSGFEETTELPHSKGRR HSGQYEAGQVQIDLGTSLSLALNQVATQHQV TLNTLLQASWGILLQKYNRQSDIVFGSVVSGRPAELVGIK EMIGLFINTIPVRVSSQAQERFAEVMTRMQEDALS SAKHDYYPLYEIQAQCTLKQNLITHIMVLENYFMEQQ LEQFNSSDGSGLKLTDVTVTEQTNYDLNLIIIPGDNIVIRFDF NKQAFEEADMNVLKQHLLHVLEQVASNPRISIGELQLATD EERAVMMSEFNDTFVAYPREKSIHRLFEERAEHEPDALAV VFGNEQMTYGALNAASNQMARRLRHKGVTNGELV GICADRSLDMVVGLLAIMKSGGAYVPIDPAYPQERISAM LEDTSIKTMVTQKHLCSLWPEHFVIVLDVNETDVSNLMEDI ESTNLSIDGAGDDLAYIIYTSGSTGKPKGVCVTHRG VVRLVCGATYVEIDSSDVFLQGSTISFDAATFEIWGAL LNGAALVILPSGNVSLTDWSEAIQRHRVTTLWMTPGLFQ VMVEQQIEGFYGVKQLLVGGDVVSPTHVRKVMEKHNGIR VINGYGPTENTTFTCCHTITAADLDRGCSIPIGRPISNTRV YVLDENGNALPVGVCGELYAGGDGLARGYLNRPELTAEKFVN DPFIPGERLYRTGDLARWLPDGSIEFIGRCDEQVKIRGYRIEP GEVEAYLLRIDEVGEAAVIAREDSSGQKELCAYFVTEAE LSASGLRETLARELPAYMIPSHFIQIEELPLTP

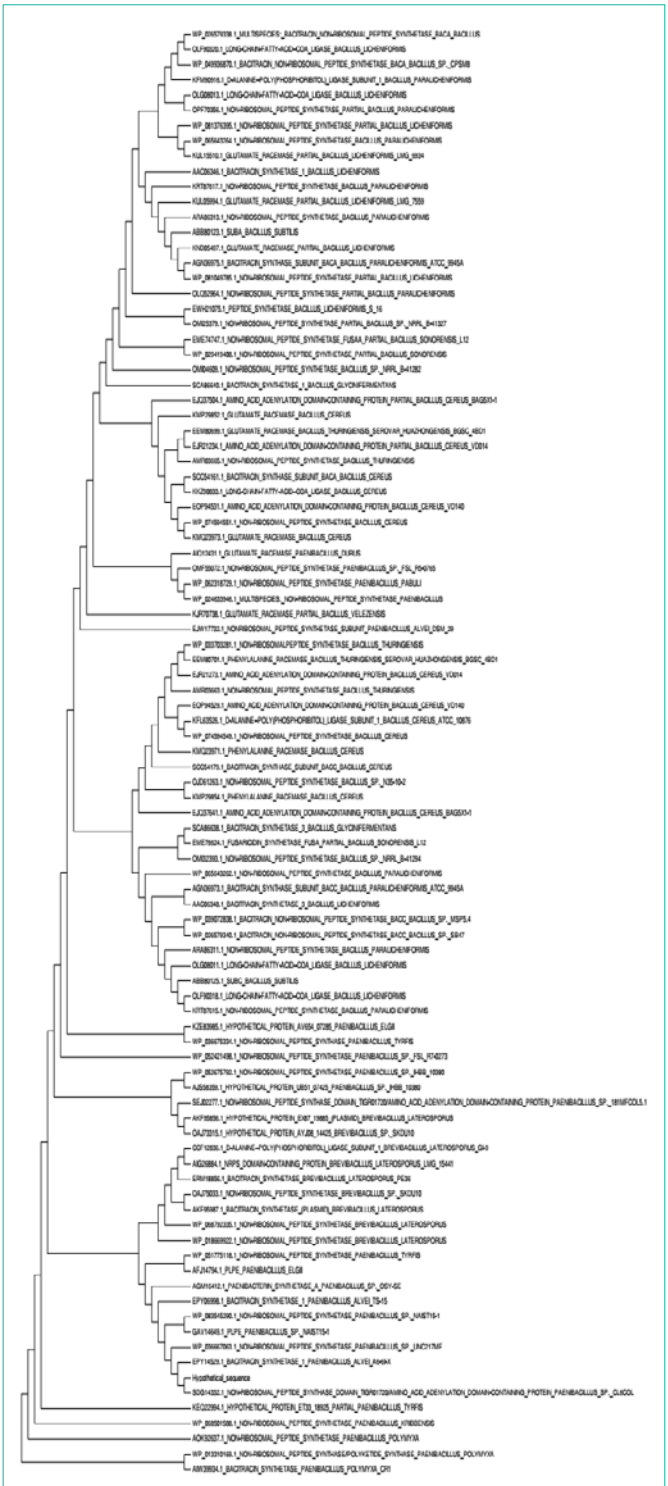

Figure 6: Phylogenic tree from Mega 6 for the 95 aligned sequences.

Table 2: The amino acids profile of the Hypothetical sequence.

	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Total
Hypothetical sequence	7,4	1,3	5,3	8,1	3,2	7,4	2,5	6,5	2,8	9,8	2,5	4,4	3,9	4,7	4,9	6,3	5,7	8,4	1,2	3,6	836

Table 3: The amino acids profiles of the used 95 sequences including the hypothetical sequence.

Sequence description	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr
AAC06346.1_BACITRACIN_SYNTHETASE_1_BACILLUS_LICHENIFORMIS	5,2	0,5	7,4	6,8	4,7	5,6	3,2	8,3	7,2	8,6	2,4	4,7	4,1	4,4	3,3	6,7	4,6	6,7	1,2	4,6
AGN36975.1_BACITRACIN_SYNTHASE_SUBUNIT_BACA_BACILLUS_PARALICHENIFORMIS_ATCC_9945A	5,6	0,5	7,4	6,8	4,7	5,7	3,2	8,3	7,2	8,6	2,4	4,7	4,0	4,3	3,2	6,7	4,6	6,5	1,2	4,6
ARA86313.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_PARALICHENIFORMIS	5,6	0,5	7,4	6,8	4,7	5,7	3,2	8,3	7,2	8,6	2,4	4,7	4,0	4,3	3,2	6,7	4,6	6,5	1,2	4,6
WP_065643264.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_PARALICHENIFORMIS	5,3	0,5	7,5	6,8	4,7	5,7	3,2	8,3	7,0	8,6	2,4	4,6	4,0	4,3	3,2	6,9	4,7	6,7	1,2	4,6
KRT87617.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_PARALICHENIFORMIS	5,6	0,5	7,4	6,8	4,7	5,7	3,2	8,3	7,2	8,6	2,4	4,7	4,0	4,3	3,2	6,7	4,6	6,5	1,2	4,6
ABB80123.1_SUBA_BACILLUS_SUBTILIS	5,6	0,5	7,3	6,7	4,6	5,9	3,3	8,3	7,2	8,6	2,4	4,7	4,0	4,3	3,2	6,7	4,6	6,5	1,2	4,6
OLG08013.1_LONG-CHAIN-FATTY-ACID--COA_LIGASE_BACILLUS_LICHENIFORMIS	5,6	0,5	7,5	6,5	4,7	5,7	3,4	8,4	7,3	8,6	2,1	4,6	4,0	3,8	3,2	6,9	4,7	6,8	1,2	4,6
KFM90916.1_D-ALANINE--POLY(PHOSPHORIBITOL)_LIGASE_SUBUNIT_1_BACILLUS_PARALICHENIFORMIS	5,6	0,5	7,5	6,5	4,7	5,7	3,3	8,4	7,3	8,6	2,1	4,6	4,0	3,8	3,2	6,9	4,7	6,8	1,2	4,7
WP_026579338.1_MULTISPECIES:_BACITRACIN_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACA_BACILLUS	5,6	0,5	7,5	6,5	4,7	5,7	3,3	8,4	7,3	8,6	2,1	4,6	4,0	3,8	3,2	6,9	4,7	6,8	1,2	4,7
OLF90320.1_LONG-CHAIN-FATTY-ACID--COA_LIGASE_BACILLUS_LICHENIFORMIS	5,6	0,5	7,5	6,5	4,7	5,7	3,3	8,4	7,3	8,6	2,1	4,6	4,0	3,8	3,2	6,9	4,7	6,8	1,2	4,7
EWH21075.1_PEPTIDE_SYNTHETASE_BACILLUS_LICHENIFORMIS_S_16	5,2	0,7	7,8	6,4	4,7	5,7	2,8	8,5	7,2	9,0	2,3	4,7	3,9	4,1	3,5	6,5	4,5	6,6	1,2	4,6
KND05497.1_GLUTAMATE_RACEMASE_PARTIAL_BACILLUS_LICHENIFORMIS	5,6	0,5	7,4	6,8	4,7	5,7	3,2	8,3	7,2	8,6	2,4	4,7	4,0	4,3	3,2	6,7	4,6	6,5	1,2	4,6
OLQ52964.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PARTIAL_BACILLUS_PARALICHENIFORMIS	5,3	0,5	7,5	6,9	4,7	5,8	3,2	8,2	7,1	8,8	2,4	4,6	4,0	4,1	3,3	6,7	4,7	6,4	1,2	4,6
WP_049936870.1_BACITRACIN_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACA_BACILLUS_SP._CPSM8	5,6	0,5	7,5	6,5	4,7	5,7	3,3	8,4	7,3	8,6	2,1	4,6	4,0	3,8	3,2	6,9	4,7	6,8	1,2	4,7
OMI25379.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PARTIAL_BACILLUS_SP._NRRL_B-41327	5,2	0,5	7,9	6,4	4,7	5,8	2,9	8,5	7,1	9,1	2,2	4,7	3,9	4,1	3,3	6,5	4,7	6,7	1,2	4,6
WP_081376395.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PARTIAL_BACILLUS_LICHENIFORMIS	5,3	0,5	7,5	6,8	4,7	5,7	3,2	8,3	7,0	8,6	2,4	4,6	4,0	4,3	3,2	6,9	4,7	6,7	1,2	4,6
KUL05994.1_GLUTAMATE_RACEMASE_PARTIAL_BACILLUS_LICHENIFORMIS_LMG_7559	5,6	0,5	7,4	6,8	4,7	5,7	3,2	8,3	7,2	8,6	2,4	4,7	4,0	4,3	3,2	6,7	4,6	6,5	1,2	4,6
WP_081049785.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PARTIAL_BACILLUS_LICHENIFORMIS	5,6	0,5	7,4	6,8	4,7	5,7	3,2	8,3	7,2	8,6	2,4	4,7	4,0	4,3	3,2	6,7	4,6	6,5	1,2	4,6
OMI04609.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_SP._NRRL_B-41282	5,0	0,7	6,9	7,9	5,0	5,6	2,4	9,4	8,0	8,9	2,1	5,2	4,1	4,1	3,2	5,6	4,4	5,9	1,3	4,5
KUL15510.1_GLUTAMATE_RACEMASE_PARTIAL_BACILLUS_LICHENIFORMIS_LMG_6934	5,3	0,5	7,5	6,8	4,7	5,7	3,2	8,3	7,0	8,6	2,4	4,6	4,0	4,3	3,2	6,9	4,7	6,7	1,2	4,6
SCA86640.1_BACITRACIN_SYNTHETASE_1_BACILLUS_GLYCINIFERMENTANS	5,2	0,5	7,2	7,3	4,6	5,8	2,5	8,7	8,8	8,9	2,0	4,6	3,7	4,1	3,2	5,0	5,0	6,7	1,3	4,8
EME74747.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_FUSAA_PARTIAL_BACILLUS_SONORENSIS_L12	4,5	0,8	7,2	7,1	4,8	5,7	2,4	9,1	7,9	8,6	2,1	5,8	3,8	3,9	3,2	6,1	4,8	6,4	1,3	4,7
WP_029419408.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PARTIAL_BACILLUS_SONORENSIS	4,5	0,9	7,2	7,1	4,8	5,7	2,4	9,1	7,9	8,6	2,1	5,8	3,8	3,9	3,2	6,1	4,8	6,4	1,2	4,7
WP_074594551.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_CEREUS	4,4	1,5	7,4	7,1	4,6	5,0	1,2	10,1	9,4	8,1	1,7	6,8	3,5	3,4	2,5	5,7	4,6	6,8	1,2	5,1
KMQ23973.1_GLUTAMATE_RACEMASE_BACILLUS_CEREUS	4,4	1,5	7,3	7,1	4,6	5,0	1,2	10,3	9,4	8,0	1,7	6,9	3,5	3,4	2,5	5,7	4,6	6,8	1,2	5,1
EOP94531.1_AMINO_ACID_ADENYLATION_DOMAIN-CONTAINING_PROTEIN_BACILLUS_CEREUS_VD140	4,4	1,5	7,3	7,2	4,6	5,0	1,2	10,1	9,3	8,1	1,7	6,9	3,5	3,4	2,5	5,7	4,6	6,8	1,2	5,1
SCC54161.1_BACITRACIN_SYNTHASE_SUBUNIT_BACA_BACILLUS_CEREUS	4,4	1,4	7,5	7,2	4,6	4,9	1,2	10,0	9,4	8,2	1,7	6,5	3,5	3,4	2,6	5,7	4,6	6,8	1,3	5,1
KMP29852.1_GLUTAMATE_RACEMASE_BACILLUS_CEREUS	4,4	1,5	7,2	7,3	4,6	4,9	1,3	9,9	9,3	8,3	1,7	7,0	3,6	3,3	2,5	5,6	4,6	6,8	1,2	5,1
KKZ90833.1_LONG-CHAIN-FATTY-ACID--COA_LIGASE_BACILLUS_CEREUS	4,4	1,4	7,4	7,2	4,6	4,9	1,2	10,0	9,4	8,2	1,7	6,7	3,5	3,4	2,6	5,7	4,6	6,8	1,3	5,1
AMR03665.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_THURINGIENSIS	4,3	1,5	7,3	7,3	4,6	4,9	1,2	10,0	9,2	8,2	1,7	6,9	3,5	3,3	2,7	5,7	4,7	6,8	1,2	5,1
EEM80699.1_GLUTAMATE_RACEMASE_BACILLUS_THURINGIENSIS_SEROVAR_HUAZHONGENSIS_BGSC_4BD1	4,3	1,5	7,2	7,3	4,6	4,9	1,2	10,0	9,2	8,2	1,7	7,0	3,5	3,3	2,7	5,7	4,7	6,8	1,2	5,1
OJD61263.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_SP._N35-10-2	3,5	1,0	6,3	9,0	3,5	4,8	2,6	7,4	7,5	10,5	2,8	7,0	3,6	3,5	4,3	5,6	5,3	6,9	0,5	4,4
AIQ12431.1_GLUTAMATE_RACEMASE_PAENIBACILLUS_DURUS	5,8	1,1	5,9	6,9	4,9	5,1	2,5	8,3	6,5	9,4	2,4	5,1	4,0	4,4	4,2	6,8	4,5	6,4	1,3	4,6
WP_062318729.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_PABULI	5,6	1,3	6,9	6,4	4,9	5,2	2,1	6,6	5,2	9,8	2,6	4,8	3,7	5,9	4,5	6,4	4,2	8,0	1,3	4,6

WP_024633946.1_MULTISPECIES:_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS	5,3	1,3	6,8	6,3	4,8	5,3	2,1	6,9	5,2	9,7	2,6	4,6	3,8	6,1	4,5	6,5	4,1	8,2	1,3	4,6
OMF55072.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_SP._FSL_R5-0765	5,6	1,3	6,8	6,1	4,9	5,3	2,1	6,4	5,6	9,8	2,5	4,8	3,8	6,0	4,2	6,5	4,4	8,1	1,3	4,6
EJQ37504.1_AMINO_ACID_ADENYLATION_DOMAIN-CONTAINING_PROTEIN_PARTIAL_BACILLUS_CEREUS_BAG5X1-1	4,5	1,4	7,2	7,1	4,9	4,9	1,4	9,5	9,1	8,3	1,7	7,1	3,6	3,1	2,8	5,8	4,0	7,4	1,2	5,0
OPF70356.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PARTIAL_BACILLUS_PARALICHENIFORMIS	5,6	0,5	7,5	6,5	4,7	5,7	3,4	8,4	7,3	8,6	2,1	4,6	4,0	3,8	3,2	6,9	4,7	6,8	1,2	4,6
KZE83985.1_HYPOTHETICAL_PROTEIN_AV654_07285_PAENIBACILLUS_ELGII	8,5	0,6	4,8	9,0	3,8	7,6	2,6	5,8	4,0	10,1	2,4	2,6	5,8	4,7	6,5	4,1	5,4	6,9	1,0	3,8
WP_036675334.1_NON-RIBOSOMAL_PEPTIDE_SYNTHASE_PAENIBACILLUS_TYRFIS	8,5	0,6	4,7	9,3	4,0	7,7	2,6	6,2	3,7	9,8	2,5	2,7	5,8	4,8	6,2	4,0	5,7	6,5	1,0	3,8
WP_052675792.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_SP._IHBB_10380	6,1	0,4	3,7	10,9	3,2	8,5	2,0	6,1	4,8	8,8	2,8	3,5	3,7	5,7	5,3	4,8	4,6	8,7	1,3	5,0
SCA86638.1_BACITRACIN_SYNTHETASE_3_BACILLUS_GLYCINIFERMENTANS	6,7	0,9	7,7	7,8	3,9	6,1	2,8	7,0	7,6	10,7	2,3	3,8	4,4	3,5	4,7	4,6	5,3	6,0	0,7	3,6
OMI32393.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_SP._NRRL_B-41294	6,7	0,9	6,7	8,6	3,9	5,4	3,2	6,5	7,4	10,7	2,3	3,9	4,7	4,0	4,5	4,9	5,2	6,0	0,8	3,7
SCC54179.1_BACITRACIN_SYNTHASE_SUBUNIT_BACC_BACILLUS_CEREUS	3,5	0,9	6,3	8,8	3,5	4,8	2,6	7,4	7,8	10,3	2,9	6,6	3,5	3,5	4,0	5,9	5,3	7,4	0,5	4,4
EJR21273.1_AMINO_ACID_ADENYLATION_DOMAIN-CONTAINING_PROTEIN_BACILLUS_CEREUS_VD014	3,6	0,9	6,2	9,0	3,6	5,0	2,4	7,6	7,4	9,9	2,8	6,9	3,6	3,4	4,3	5,6	5,1	7,5	0,5	4,6
WP_033703281.1_NON-RIBOSOMALPEPTIDE_SYNTHETASE_BACILLUS_THURINGIENSIS	3,6	0,9	6,2	9,0	3,6	5,0	2,4	7,6	7,4	9,9	2,8	6,9	3,6	3,4	4,3	5,6	5,1	7,5	0,5	4,6
AMR03663.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_THURINGIENSIS	3,6	0,9	6,2	9,0	3,6	5,0	2,4	7,9	7,4	9,9	2,8	6,9	3,6	3,4	4,3	5,6	5,0	7,4	0,5	4,6
EJR21234.1_AMINO_ACID_ADENYLATION_DOMAIN-CONTAINING_PROTEIN_PARTIAL_BACILLUS_CEREUS_VD014	4,3	1,5	7,2	7,3	4,6	4,9	1,2	10,0	9,2	8,2	1,7	7,0	3,5	3,3	2,7	5,7	4,7	6,8	1,2	5,1
KFL63526.1_D-ALANINE--POLY(PHOSPHORIBITOL)_LIGASE_SUBUNIT_1_BACILLUS_CEREUS_ATCC_10876	3,6	0,9	6,2	8,8	3,6	5,0	2,4	7,8	7,4	9,9	2,8	7,0	3,6	3,4	4,3	5,6	5,1	7,4	0,5	4,6
EOP94529.1_AMINO_ACID_ADENYLATION_DOMAIN-CONTAINING_PROTEIN_BACILLUS_CEREUS_VD140	3,6	0,9	6,2	8,8	3,7	5,0	2,4	7,8	7,4	9,9	2,8	7,0	3,6	3,4	4,3	5,6	5,1	7,3	0,5	4,6
EEM80701.1_PHENYLALANINE_RACEMASE_BACILLUS_THURINGIENSIS_SEROVAR_HUAZHONGENSIS_BGSC_4BD1	3,6	0,9	6,2	9,0	3,6	5,0	2,4	7,6	7,4	9,9	2,8	6,9	3,6	3,4	4,3	5,6	5,1	7,5	0,5	4,6
AJS58359.1_HYPOTHETICAL_PROTEIN_UB51_07425_PAENIBACILLUS_SP._IHBB_10380	5,8	0,6	4,9	9,6	4,2	6,5	2,8	6,8	5,1	9,3	2,2	4,5	3,7	2,8	5,0	6,3	5,6	9,1	1,2	3,9
AGN36973.1_BACITRACIN_SYNTHASE_SUBUNIT_BACC_BACILLUS_PARALICHENIFORMIS_ATCC_9945A	7,7	0,8	7,6	7,6	3,9	6,2	2,5	6,6	7,2	10,9	2,4	2,9	4,4	3,6	5,3	5,3	5,2	5,4	0,8	3,6
WP_074594549.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_CEREUS	3,6	0,9	6,2	8,8	3,6	5,0	2,4	7,8	7,4	10,0	2,8	7,0	3,6	3,4	4,1	5,6	5,1	7,4	0,5	4,6
KMQ23971.1_PHENYLALANINE_RACEMASE_BACILLUS_CEREUS	3,6	0,9	6,1	9,1	3,6	4,8	2,4	7,8	7,5	9,9	2,7	6,9	3,6	3,4	4,3	5,6	5,2	7,5	0,5	4,6
AAC06348.1_BACITRACIN_SYNTHETASE_3_BACILLUS_LICHENIFORMIS	7,7	0,8	7,6	7,6	3,9	6,2	2,5	6,6	7,2	10,9	2,4	2,9	4,4	3,6	5,3	5,3	5,2	5,4	0,8	3,6
OAJ75033.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BREVIBACILLUS_SP._SKDU10	6,5	0,9	5,7	8,2	3,1	5,7	2,9	7,4	3,3	10,0	3,0	4,0	4,7	6,1	4,6	6,3	4,4	8,2	1,0	4,1
OLF90318.1_LONG-CHAIN-FATTY-ACID--COA_LIGASE_BACILLUS_LICHENIFORMIS	7,6	0,8	7,6	7,6	4,2	6,3	2,5	6,6	7,0	11,2	2,4	3,2	4,4	3,4	5,4	5,0	5,1	5,4	0,8	3,5
OLG08011.1_LONG-CHAIN-FATTY-ACID--COA_LIGASE_BACILLUS_LICHENIFORMIS	7,7	0,8	7,5	7,6	4,2	6,3	2,5	6,6	7,1	11,2	2,3	3,3	4,5	3,4	5,4	5,0	5,0	5,3	0,8	3,5
WP_018669922.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BREVIBACILLUS_LATEROSPORUS	7,0	0,8	5,5	8,2	3,1	6,2	2,6	7,1	3,6	10,1	2,7	4,1	4,5	6,2	4,5	5,9	5,2	7,7	1,0	4,2
ARA86311.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_PARALICHENIFORMIS	7,7	0,8	7,5	7,6	4,2	6,3	2,5	6,6	7,1	11,2	2,3	3,3	4,5	3,4	5,4	5,0	5,0	5,3	0,8	3,5
ABB80125.1_SUBC_BACILLUS_SUBTILIS	7,6	0,8	7,6	7,6	4,2	6,3	2,5	6,6	7,0	11,2	2,4	3,2	4,5	3,4	5,4	4,9	5,1	5,4	0,8	3,5
KRT87615.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_PARALICHENIFORMIS	7,6	0,8	7,6	7,6	4,2	6,3	2,5	6,6	7,0	11,2	2,4	3,2	4,4	3,4	5,4	5,0	5,1	5,4	0,8	3,5
WP_065643262.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACILLUS_PARALICHENIFORMIS	7,7	0,8	7,5	7,8	3,8	6,2	2,5	6,6	7,1	11,0	2,3	3,0	4,2	3,4	5,7	5,2	5,2	5,7	0,8	3,5
AKF95987.1_BACITRACIN_SYNTHETASE_(PLASMID)_BREVIBACILLUS_LATEROSPORUS	6,3	0,9	5,6	8,2	3,1	5,8	2,9	7,4	3,4	10,2	2,9	3,9	4,7	6,1	4,5	6,3	4,5	8,3	1,0	4,1
WP_039072838.1_BACITRACIN_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACC_BACILLUS_SP._MSP5.4	7,7	0,8	7,6	7,6	4,2	6,3	2,6	6,5	7,1	11,1	2,3	3,2	4,5	3,4	5,4	5,0	5,1	5,3	0,8	3,5
ERM18856.1_BACITRACIN_SYNTHETASE_BREVIBACILLUS_LATEROSPORUS_PE36	6,7	0,9	5,5	8,2	3,1	5,8	2,7	7,3	3,4	10,0	2,9	4,0	4,7	6,2	4,7	6,0	4,5	8,5	1,0	4,1
EJQ37641.1_AMINO_ACID_ADENYLATION_DOMAIN-CONTAINING_PROTEIN_BACILLUS_CEREUS_BAG5X1-1	4,7	0,7	6,2	9,5	3,8	5,1	2,3	8,1	8,8	10,7	2,8	6,1	3,7	2,9	3,5	5,2	5,0	6,1	0,5	4,0
CCF12836.1_D-ALANINE--POLY(PHOSPHORIBITOL)_LIGASE_SUBUNIT_1_BREVIBACILLUS_LATEROSPORUS_GI-9	6,5	0,9	5,6	8,2	3,0	5,8	2,6	7,4	3,4	10,0	2,9	3,8	4,7	6,3	4,7	6,0	4,6	8,5	1,0	4,2
WP_026579340.1_BACITRACIN_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BACC_BACILLUS_SP._SB47	7,7	0,8	7,6	7,6	4,2	6,3	2,6	6,5	7,1	11,1	2,3	3,2	4,5	3,4	5,4	5,0	5,1	5,3	0,8	3,5
WP_051775118.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_TYRFIS	9,5	0,5	5,0	8,5	3,8	6,9	2,6	5,1	3,8	9,4	3,1	3,2	4,8	5,1	5,0	5,4	4,9	7,9	1,1	4,3

WP_068792335.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_BREVIBACILLUS_LATEROSPORUS	6,5	0,9	5,5	8,2	3,1	5,8	2,6	7,2	3,4	10,1	2,9	4,1	4,7	6,3	4,6	6,0	4,6	8,5	1,0	4,1
AFJ14794.1_PLPE_PAENIBACILLUS_ELGII	9,7	0,5	4,9	8,7	3,7	6,7	2,6	5,3	3,6	9,8	3,2	3,5	4,5	5,1	5,1	5,2	5,2	7,1	1,1	4,4
AIG26884.1_NRPS_DOMAIN-CONTAINING_PROTEIN_BREVIBACILLUS_LATEROSPORUS_LMG_15441	6,4	0,9	5,5	8,3	3,1	5,8	2,6	7,3	3,4	10,0	2,8	3,9	4,7	6,3	4,7	6,1	4,6	8,6	1,0	4,1
KMP29854.1_PHENYLALANINE_RACEMASE_BACILLUS_CEREUS	3,5	1,0	6,2	9,1	3,5	4,8	2,6	7,2	7,5	10,4	3,1	7,0	3,6	3,5	4,3	5,6	5,3	7,0	0,5	4,4
AGM16412.1_PAENIBACTERIN_SYNTHETASE_A_PAENIBACILLUS_SP._OSY-SE	9,0	0,9	5,3	8,2	2,4	7,4	2,7	6,2	3,4	10,5	2,6	3,1	4,4	5,2	4,8	4,9	6,0	8,2	1,2	3,5
EJW17732.1_NONRIBOSOMAL_PEPTIDE_SYNTHETASE_SUBUNIT_PAENIBACILLUS_ALVEI_DSM_29	8,9	0,8	5,4	6,6	4,0	7,7	3,6	5,1	2,7	11,8	2,8	2,7	5,4	4,3	5,1	5,4	5,5	7,6	0,9	3,9
Hypothetical_sequence	7,4	1,3	5,3	8,1	3,2	7,4	2,5	6,5	2,8	9,8	2,5	4,4	3,9	4,7	4,9	6,3	5,7	8,4	1,2	3,6
KEQ22994.1_HYPOTHETICAL_PROTEIN_ET33_18925_PARTIAL_PAENIBACILLUS_TYRFIS	9,9	0,6	4,5	9,5	3,4	7,1	2,0	5,5	3,0	9,5	2,3	3,0	4,9	5,4	6,1	4,3	5,2	8,2	1,5	4,3
WP_083545290.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_SP._NAIST15-1	7,6	1,2	5,3	7,9	2,8	7,7	2,5	6,1	2,4	10,6	2,7	3,9	4,1	4,6	5,3	6,3	6,0	8,6	1,2	3,3
GAV14649.1_PLPE_PAENIBACILLUS_SP._NAIST15-1	7,6	1,2	5,3	7,9	2,8	7,7	2,5	6,1	2,4	10,6	2,7	3,9	4,1	4,6	5,3	6,3	6,0	8,6	1,2	3,3
WP_036667063.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_SP._UNC217MF	7,6	1,3	5,3	7,9	3,0	7,5	2,6	6,3	2,5	10,5	2,6	4,1	4,1	4,6	5,1	6,2	6,1	8,3	1,2	3,3
AKF95856.1_HYPOTHETICAL_PROTEIN_EX87_19885_(PLASMID)_BREVIBACILLUS_LATEROSPORUS	5,7	0,6	4,1	8,4	3,2	6,4	2,8	7,0	4,4	11,1	2,5	3,5	4,6	5,9	6,4	5,8	4,8	7,1	1,4	4,2
EPY14529.1_BACITRACIN_SYNTHETASE_1_PAENIBACILLUS_ALVEI_A6-6I-X	7,6	1,3	5,3	7,9	3,0	7,5	2,6	6,3	2,5	10,5	2,6	4,2	4,0	4,6	5,1	6,2	6,1	8,3	1,2	3,3
OAJ73315.1_HYPOTHETICAL_PROTEIN_AYJ08_14425_BREVIBACILLUS_SP._SKDU10	5,8	0,6	4,0	8,4	3,3	6,7	2,9	7,1	4,5	11,1	2,4	3,4	4,6	5,9	6,0	6,0	4,7	7,0	1,4	4,2
KJR70738.1_GLUTAMATE_RACEMASE_PARTIAL_BACILLUS_VELEZENSIS	5,3	1,2	7,0	7,5	5,1	4,9	1,6	9,4	8,2	9,4	2,3	5,9	3,4	3,4	3,2	7,0	4,0	5,6	1,2	4,4
SDG14332.1_NON-RIBOSOMAL_PEPTIDE_SYNTHASE_DOMAIN_TIGR01720/AMINO_ACID_ADENYLATION_DOMAIN-CONTAINING_PROTEIN_PAENIBACILLUS_SP._CL6COL	7,6	1,3	5,3	7,9	3,0	7,5	2,6	6,3	2,5	10,5	2,6	4,2	4,0	4,6	5,1	6,2	6,1	8,3	1,2	3,3
AOK92637.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_POLYMYXA	7,6	0,5	4,5	9,1	3,1	6,7	2,7	5,7	3,1	9,8	2,4	3,3	4,6	6,6	4,7	5,3	6,0	9,0	1,4	3,9
EPY06998.1_BACITRACIN_SYNTHETASE_1_PAENIBACILLUS_ALVEI_TS-15	7,9	1,2	5,4	7,8	3,0	7,4	2,4	6,3	2,5	10,7	2,6	4,2	4,2	4,6	5,2	6,1	5,7	8,2	1,3	3,3
WP_013310169.1_NON-RIBOSOMAL_PEPTIDE_SYNTHASE/POLYKETIDE_SYNTHASE_PAENIBACILLUS_POLYMYXA	7,7	0,5	4,9	8,8	3,1	6,7	2,7	6,0	3,1	9,3	2,6	3,2	4,8	6,3	4,9	5,7	5,5	9,0	1,4	3,9
AIW39934.1_BACITRACIN_SYNTHETASE_PAENIBACILLUS_POLYMYXA_CR1	7,4	0,4	5,1	8,9	3,1	7,2	2,7	6,1	3,0	9,7	2,2	2,8	4,9	6,2	4,8	5,7	5,8	8,6	1,5	3,9
WP_068501588.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_KRIBBENSIS	8,2	0,5	5,0	8,7	3,1	7,3	2,1	6,3	3,3	9,4	2,5	3,1	4,6	6,2	5,2	5,2	5,3	8,7	1,4	4,0
WP_052421498.1_NON-RIBOSOMAL_PEPTIDE_SYNTHETASE_PAENIBACILLUS_SP._FSL_R7-0273	9,2	1,1	4,4	9,0	3,0	8,3	2,2	4,9	3,2	11,2	2,5	3,4	4,6	5,1	5,7	4,1	5,3	7,5	1,4	4,2
SEJ02277.1_NON-RIBOSOMAL_PEPTIDE_SYNTHASE_DOMAIN_TIGR01720/AMINO_ACID_ADENYLATION_DOMAIN-CONTAINING_PROTEIN_PAENIBACILLUS_SP._181MFCOL5.1	6,1	0,5	5,1	7,4	3,5	7,0	2,8	8,2	5,0	10,7	1,9	3,4	4,6	5,6	4,7	6,1	5,8	6,3	1,3	3,9
EME76624.1_FUSARICIDIN_SYNTHETASE_FUSA_PARTIAL_BACILLUS_SONORENSIS_L12	7,5	0,9	7,1	7,8	3,9	5,8	2,8	7,7	6,9	10,1	2,6	3,6	4,5	4,0	4,7	4,9	5,2	5,7	0,7	3,7

The generated hypothetical protein, its amino acid profile and the 94 aa sequences were inserted with their amino acid profiles (Table 3).

Both, of the 95 sequences and their amino acids profiles were subjected to alignment. Trees were generated using Mega 4 and data clustering using Past statistic software (Figure 5).

The tree was obtained for the 95 sequences from Mega 6 (Figure6), where the evolutionary history was inferred using the Maximum Parsimony method. The most parsimonious tree with length = 6040 is shown. The consistency index is 0,603477 (0,601298), the retention index is 0,915820 (0,915820), and the composite index is 0,552676 (0,550681) for all sites and parsimony-informative sites (in parentheses). The MP tree was obtained using the SPR algorithm (pg. 126 in ref. [23]), with search level of 0, in which the initial trees were obtained by the random addition of sequences (10 replicates). The analysis involved 95 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 761 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 [20].

Both trees, either obtained from the 95 aa sequences using Mega 6 or obtained from clustering the profiles of the amino acid sequence % (Figure 7), show that the hypothetical sequence was clustered with the same group in both cases (Figure 6 and 7). The hypothetical protein was clustered in both with EPY06998.1; GAV14649.1; WP_083545290.1; EPY14529.1. That proves the possibility of using the amino acids % profile in identifying unknown protein. Or at least decrease the investigation area, where all of the obtained species were related to Bacillus. That additionally highlights the possibility of building microbial database by searching a specific protein. But one should find a specific protein and avoid common proteins involved in the general metabolic pathways which can be found nearly in each microbial strain.

Moreover, the amino acid profile could give some information

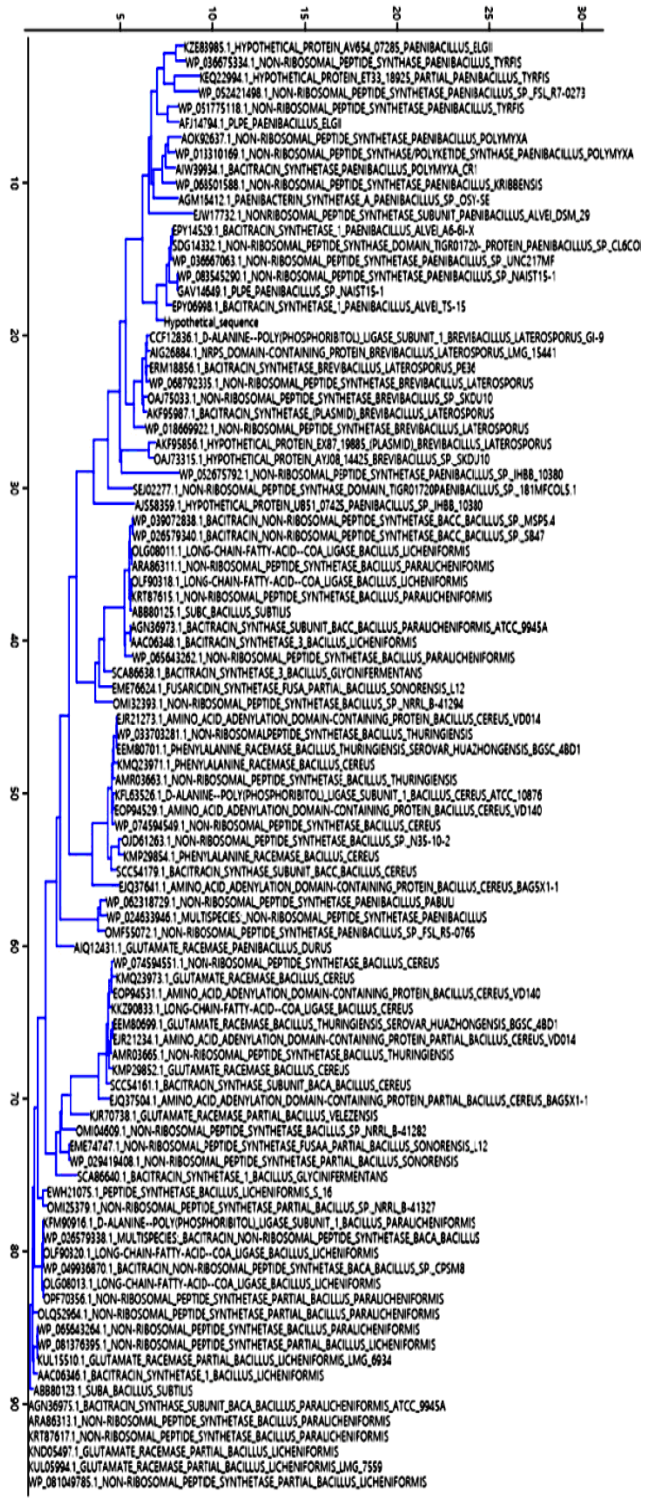

Figure 7: Cluster result of the amino acids % profile of the used 95 sequences (including the hypothetical sequencing).

about the protein structure particularly those amino acids involved in catalysis. The amino acid % profile proves identity for the Bacitracin synthases. In the field of the biotechnology, the availability to express genes in different species would be of great importance.

This study recommends using the amino acids % profile in searching for proteins. And it is suggested to conduct more research concerning the amino acids % profile of other specific proteins.

Conclusion

The main results obtained from this study could be concluded in the following points.

1. Amino acid % profiles could be used to identify unknown proteins. The study proves its sensitivity after being compared with the aligned sequences.

2. Specific proteins could build better database for certain microbial groups such as Bacillus sp which are able to produce Bacitracin.

3. Specific proteins could enable chances to select and use better heterologous expression alternative microbe(s).

4. The Bioinformatics could be implemented in new fields in biotechnology and vise verse.

Acknowledgement

The author acknowledges the City of Scientific Research and Technological Applications for supporting this work.

References

1. Amara AA. Shibi A. Role of Probiotics in health improvement, infection control and disease treatment and management. Saudi Pharma J. 2015; 23: 107-114.

2. Siriken B, Bavaram I, Onol AG. Effects of probiotics: alone and in a mixture of Biosacc plus Zinc Bacitracin on the caecal microflora of Japanese quail. Res Vet Sci. 2003; 75: 9-14.

3. O'Sullivan MG, Thornton G, O'Sullivan GC, Collins JK. Probiotic bacteria: myth or reality. Trends Food Sci Technol 1992; 309-314.

4. Cammarota M, De Rosa M, Stellavato A, Lamberti M, Marzaioli I, Giuliano M. *In vitro* evaluation of Lactobacillus plantarum DSMZ 12028 as a probiotic: emphasis on innate immunity. Int J Food Microbiol. 2009; 135: 90-98.

5. Rocha B, Correia M, Barbosa R, Laranjinha J. A dietary-driven redox modulation of gut microbiome-host interactions: the rescue of epithelial barrier and mucus production during dysbiosis by dietary nitrate. Free Radic Biol Med. 2014; 75: S36-37.

6. Moreau MC, Raibaud P, Muller MC. [Relationship between the development of the intestinal IgA immune system and the establishment of microbial flora in the digestive tract of young holoxenic mice]. Ann Immunol (Paris). 1982; 133D: 29-39.

7. Amara AAK. In: Amara, A. (Ed.), Toward Healthy Genes. Schüling Verlage, Germany 2012.

8. Amara AA. The inevitability of balanced lives: genes–foods –action-interactions. IIOBJ 2013; 4: 1.

9. Gismondo MR, Drago L, Lombardi A. Review of Probiotics available to modify gastrointestinal flora. Int. J. Antimicrob. Agents. 1999; 12: 287-292.

10. Lilly DM, Stillwell RH. Probiotics: Growth-promoting factors produced by microorganisms. Science 1965; 147: 747-748.

11. Parker RB. Probiotics, the other half of the antibiotic story. Anim. Nutr. Health 1974; 29: 4.

12. Collinder E, Cardona ME, Berge GN, Norin E, Stern S, Midtvedt T. Influence of zinc bacitracin and Bacillus licheniformis on microbial intestinal functions in weaned piglets. Vet Res Commun. 2003; 27: 513-526.

13. Adami A, Cavazzoni V. Occurrence of selected bacterial groups in the faeces of piglets fed with Bacillus coagulans as probiotic. J Basic Microbiol. 1999;

39: 3-9.

14. Spann CT, SC; Weinberg, JM. Topical antimicrobial agents in dermatology. Dis Mon. 2004; 50: 407-421.

15. Trookman NR, RL; Weber, T. Treatment of minor wounds from dermatologic procedures: A comparison of three topical wound care ointments using a laser wound model. J Am Acad Dermatol. 2011; 64: S8-15.

16. Karala AR, Ruddock LW. Bacitracin is not a specific inhibitor of protein disulfide isomerase. FEBS J. 2010; 277: 2454-2462.

17. Weston BS, Wahab NA, Roberts T, Mason RM. Bacitracin inhibits fibronectin matrix assembly by mesangial cells in high glucose. Kidney Int. 2001; 60: 1756-1764.

18. Madden T. The BLAST Sequence Analysis Tool. 2002 Oct 9 [Updated 2003 Aug 13]. In: McEntyre J, Ostell J, editors. The NCBI Handbook [Internet]. Bethesda (MD): National Center for Biotechnology Information (US) 2002; Chapter 16.

19. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999; 41: 95-98.

20. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013; 30: 2725-2729.

21. Nicholas KB, Nicholas HB Jr. Gene Doc: a tool for editing and annotatong multiple sequence alignments. Distributed by the author. 1997.

22. Drummond AH, J. Lemey, P. de Oliveira, T. Pybus, O. Shapiro, B. Suchard, M. Tree Figure Drawing Tool Version 1.4.2. Andrew Rambaut Institute of Evolutionary Biology, University of Edinburgh.

23. Nei M, Sudhir K. Molecular Evolution and Phylogenetics. Oxford University Press, New York: 2000.

Austin J Proteomics Bioinform & Genomics - Volume 4 Issue 1 - 2017
ISSN : 2471-0423 | www.austinpublishinggroup.com

Citation: Amara AA and Sharaf MM. The Amino Acids' Profiles for Searching Proteins Similarity: Bacitracin Synthase as a Model. Austin J Proteomics Bioinform & Genomics. 2017; 4(1): 1020.