IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Identification and Genomic study of the biopotent actinomycetes species Microbacterium Barkeri (LMA4): a computational approach

Krishna Kumar Das[a], Smaranika Pattnaik[*a], Santosh Kumar Behera[b]

[*a]Laboratory of Medical Microbiology, Dept. of Biotechnology and Bioinformatics, Sambalpur University, JyotiVihar, Burla, Odisha -768019, India.

[b]National Institute of Pharmaceutical Education and Research, Ahmedabad, Gujarat, India

ABSTRACT

This is important to mention that antibiotic therapy is most prioritized approach to subside bacteria generated infections. The emergence of multi drug resistance (MDR) conferred by the clinically relevant bacterial isolates is observed to be a serious threat for the developing countries like India. The bacteria being smart enough deploy more than one technique to evade the action of conventional antibiotics. Hence, this circumstance provokes for introduction of novel antibiotic molecules, its derivatives or its analogues. Therefore, this study was carried out to identify bio potent actinomycetes antibiotic producing candidates from soil habitat. Using pertinent isolation and identification methods, including 16s rRNA sequencing, a bio potent strain of *Microbacterium barkeri* (LMA4), a Torques Green pigment producing, with citrate utilization activity was identified. The microscopic features were also analysed. The 16s rRNA sequencing followed by BLASTn analysis, inferring 98.84 % of identity had substantiated the identification. The genomic analysis was done by the NCBI database from *Microbacterium sp.* to *Microbacterium barkeri* species. By purifying the predicted genes lead to the specific key role genes named RPLP, RPSL, RPSR, RPSE, RPMA, RPLB, RPSG, RPSS, RPMG and RPLQ to the gene –gene interaction studies of *M. barkeri* in their biological functions. Modeling of the genes related to the *M. barkeri* species in SWISS-MODEL database leads to predict the new chemical structures of the key genes.

Keywords: Multi drug resistance, 16srRNA sequencing, BLASTn analysis, *Microbacterium barkeri,* SWISS-MODEL.

INTRODUCTION

The genus Microbacterium was first proposed in 1919 [1]. *Microbacterium sp.* belongs to the family Microbacteriaceae, order Actinomycetales, class Actinobacteria which comprises mainly aerobic Gram positive bacteria with high G+C content and a peptidoglycan defined by a B-type cross linkage [2]. *Microbacterium barkeri* is one of the species in the genus *Microbacterium*, which belongs to the high-GC-content phylum Actinobacteria. It has been reported that *Microbacterium* strains can cause human, animal, and plant diseases. However, *Microbacterium* spp. have frequently been isolated from the soil and used as bio control agents [3]. *Microbacterium barkeri* strains are strong in plant colonization and play a very important role in biocontrol. Several *Microbacterium barkeri* strains with potential biocontrol ability from soil. *Microbacterium barkeri* may be useful to further explore the commercial potential for biocontrol in rice fields, and genome-wide sequence analysis could play an important role for genetic studies of this bacterium. *Microbacterium barkeri* and *Paenibacillus amylolyticus* were isolated and identified for the degradation of polyvinyl alcohol (PVA) contained in textile waste water. The said Actinomycetes was (*Microbacterium barkeri*) isolated from a pond side marshy soil. The soil was diluted with distilled water and made into various dilutions and incorporated onto NA plates. The mixed colonies grown on the NA plates were studied and a presumptive Actinomycetal colony was picked up and subjected to further analysis using optical microscopic and 16s r RNA sequencing. After analyzing in BLASTn window in NCBI portal, the strain (LMA4) was screened for its bacterial growth restriction activity [4-7].

MATERIALS AND METHODS

Isolation and Identification of Actinomycetes from Soil:

Soil samples were collected from different areas near University campus ponds and pond side marshy soil. Soil samples were collected from 5-25 cm depth in sterile plastic bags and transported aseptically to the Microbiology laboratory. The samples were air dried for 24-48 hr at room temperature. Isolation and enumeration of Actinomycetes were done by serial dilution and spread plate technique on the actinomycetes isolation agar (AIA) plates and the plates incubated for 4-6 days in 36-37^{o}C [5-11].

DNA sequencing of 16S rRNA gene fragment: Sequencing of 16S rRNA gene fragments of selected bacterial isolates was done from both forward and reverse directions. The sequence obtained was subjected to BLASTn search leading to identification of bacterial species [12-15].

Computational analysis (BLAST) and identification of bacterial species:

Basic Local Alignment Search Tool (BLAST) uses an algorithm for searching similarities above certain threshold between a query sequence and all other corresponding sequences present in a database [23-25].

Phylogenetic tree analysis:

Phylogenetic tree was constructed using Neighbour-Joining (NJ) method of mathematical averages (UPGMA) among 16S rRNA gene sequence and corresponding gene sequences. Bacterial isolates selected were united with quite high statistical support by the bootstrap method estimates for 100 replications and values inferred greater than 38 percent [16, 17, 22].

Gene prediction from NCBI

The National Center for Biotechnology Information (NCBI) at the National Institutes of Health was created in 1988 to develop information systems for molecular biology. In addition to maintaining the GenBank(R) (1) nucleic acid sequence database, to which data is submitted by the scientific community, NCBI provides data retrieval systems and computational resources for the analysis of GenBank data as well as a variety of other biological data. The representative genes sequence data for microbacterium were downloaded from the ftp site of the National Center for Biotechnology Information (NCBI) according to the list as of March 14, 2014 (https://www.ncbi.nlm.nih.gov/gene/?term=microbacterium). The sequence data were originally composed of 3559 genes, and the valid genes were selected using the four following criteria: (1) the scientific name is given (2) genome sequence is not fragmented; (3) gap region is small (<5% of the genome); and (4) protein-coding regions are predicted. Finally, a total of 508 genes were found as the associated genes of *Microbacterium species* [34].

Annotation and GO association of genes of *Microbacterium sp.* by DAVID:-

The functional annotation of a total of 508 genes was performed through Gene Ontology (GO) analysis which describes the functions along the three categories *viz.*, molecular functions (MF), biological processes (BP) and the cellular components (CC). The Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resources (http://david.abcc.ncifcrf.gov/) was used for GO term annotation (i.e., the common vocabulary for the functional description of genes and gene products) annotation. Finally to find the statistically significant GO terms of the genes, GO term enrichment analysis was performed. The DAVID parameters were filtered to reduce the false positives and the output was taken into account after applying multiple testing correction (p-values <0.05), fold change and False Discovery Rate (FDR). Genes from significantly enriched biological processes were termed as key genes and were used for network construction [35].

Generation of gene network and its interactions using STRING database:-

A total of 43 genes obtained from significantly enriched biological processes were termed as key genes and were used for network construction of gene-gene interaction were analyzed through STRING database.The most interacted genes are taken as the key functional genes for the *Microbacterium sp*.[36]

UniProt-(http://www.uniprot.org/):-

The Universal Protein Resource (UniProt) was a comprehensive resource for protein sequence and annotation data. The corresponding protein sequences encoded by these genes were retrieved from UniProtKB database. [37]

Retrieval of Drugs and proteins:-

The Structure Data Format (SDF) 3D/2D structure of the reported drugs were retrieved from the NCBI PubChem database (http://www.ncbi.nlm.nih.gov/ pccompound/) along with its PubChem ID, Molecular weight and Molecular formula. The compounds were converted into pdb format structure using the PyMol (academic version) tool. The structures of the corresponding proteins of reported genes were retrieved from PDB Protein Data Bank (PDB). The unknown structures were predicted using tool SWISS-MODEL web servers due to unavailability at Protein Data Bank (PDB) server [38-40].

RESULTS AND DISCUSSION

The site of collection of soil and the type of soil collected are shown in **Figure.1**. The texture of soil was under category of 'Clay soil'.

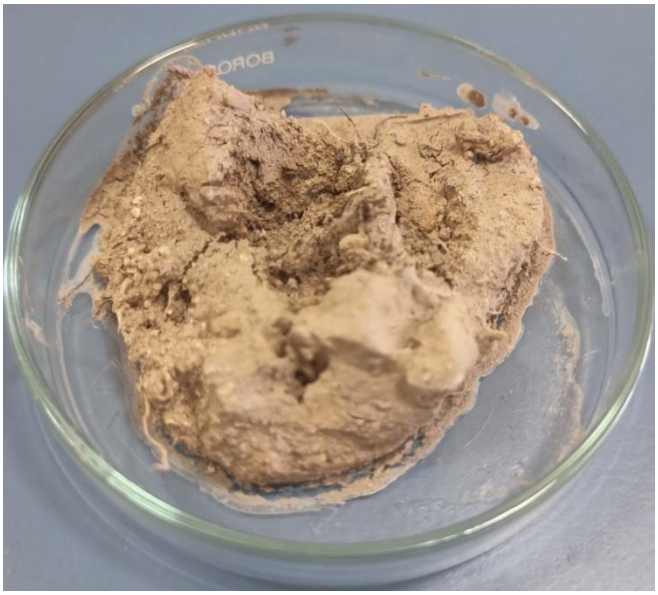


Figure.1: Depicting the type of soil collection

Growth of Actinomycetes on Agar media

The following mediums were used for the isolation of Actinomycetes Nutrient agar medium and Actinomycetes Isolation Agar media. Isolated plates were incubated at 37°C for 3-4 days for fast growing Actinomycetes. Plates were checked for the growth of typical Actinomycetes colonies up to 7 days. The Actinomycetes grown on the Actinomycetes agar plate, is given in Figure.2. The growth was affluent, diffuse, off white in colour.

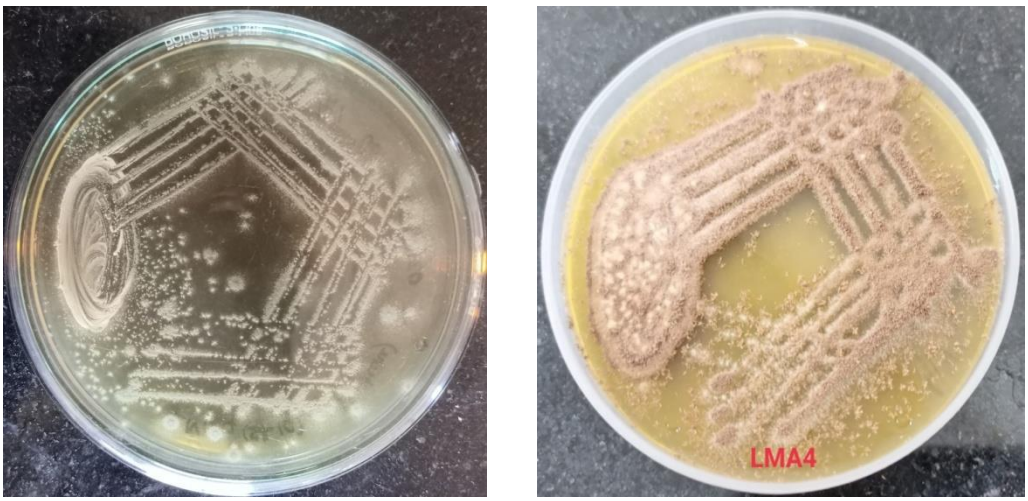


Figure.2: Displaying the growth of LMA4 from preliminary stage to proper growth on Actinomycetes agar media

A smear of culture was taken in a clean glass slide and heated gently over a flame. Then the smear was going through the gram staining process, then the slide was washed, drained, air dried, and viewed under microscope. The culture retaining the slightly purple / violet color indicated that it was Gram-positive organism with branched slender Gram positive mycelia with rounded spores, as given in **Figure.3**.

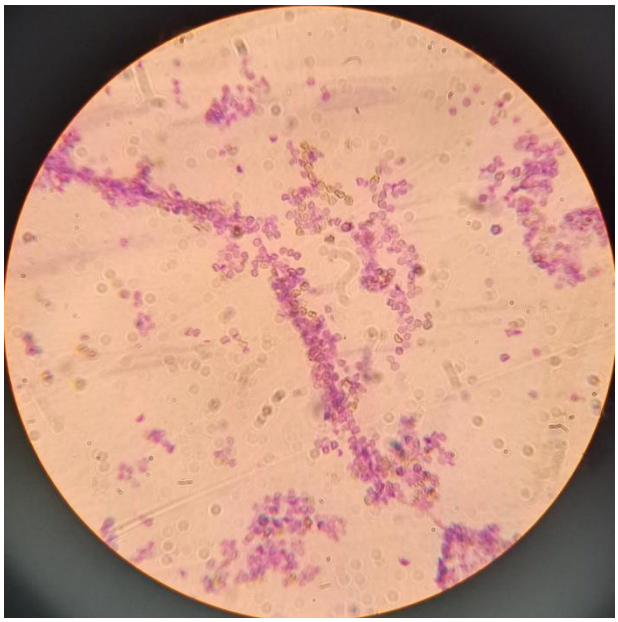


Figure.3: Showing the slender branched mycelia with rounded spores

The first amplification was to detect members of the rare actinomycete-like bacteria using KTED161F (5'ATACCGGBGMGAAAKYGYCGAC3') and GNSB 941R (5'AAACC ACACGCTCCGCT3') primers as described by Yabe *et al.* [5]. The strain which detected as rare actinomycete-like bacteria then amplified using universal primer 9F(5'-GAGTTTGATCCTGGCTCAG-3') and 1510R (5'-GGCTACCTTGTTACGA-3') as described by Nonaka *et al.*[29] to obtain the nearly-full sequence which is given below

figure.4.

>LMA4

TGTCGGTTACCCCTTGTACGACTTCACCCCAATCATCTGTCCCACCTTCGGCGGCTGGCTCCAAAAGGTTACCTCAC
CGACTTCGGGTGTTACAAACTCTCGTGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCA
TGCTGATCCGCGATTACTAGCGATTCCAGCTTCACGCAGTCGAGTTGCAGACTGCGATCCGAACTGAGAACAGATT
TGTGGGATTGGCTTAGCCTCGCGGCTTCGCTGCCCTTTGTTCTGCCCATTGTAGCACGTGTGTAGCCCAGGTCATA
AGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCACCTTAGAGTGCCCAACTGA
ATGCTGGCAACTAAGATCAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACA
ACCATGCACCACCTGTCACTCTGCCCCCGAAGGGGAAGCCCTATCTCTAGGGTTGTCAGAGGATGTCAAGACCTG
GTAAGGTTCTTCGCGTTGCTTCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTT
TCAGTCTTRCGRCCGTACTCCCCAGGCCGGGAGTGCTTAATGCGTTTGCTGCAGCACTAAAGGGCGGAAACCCTCT
AACACTTAGCAYTCATCGTTKACGGCGTGGACTACCAGGGTATCTAATCCTGTTCGCTCCCCACGCTTTCGCGCCTC
AGCGTCAGTTACAGACCAGAGAGTCGCCTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTCACCGCTACACGT
GGAATTCCACTCTCCTCTTCTGCACTCAAGTTCCCCAGTTTCCAATGACCCTCCCCGGTTGAGCCGGGGGCTTTCAC
ATCAGACTTAAGAAACCGCCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGTATTACCG
CGGCTGCTGGCACGTAGTTAGCCGTGGCTTTCTGGTTAGGTACCGTCAAGGTACCGCCCTATTCGAACGGTACTTG
TTCTTCCCTAACAACAGAGTTTTACGATCCGAAAACCTTCATCACTCACGCGGCGTTGCTCCGTCAGACTTTCGTCC
ATTGCGGAAGATTCCCTACTGCTGCCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCCGATCACCCT
CTCAGGTCGGCTACGCATCGTCGCCTTGGTGAGCCGTTACCTCACCAACTAGCTAATGCGCCGCGGGTCCATCTGT
AAGTGGTAGCTGAAAGCCACCTTTTATGATTGAACCATGCGGGTCAATCAAGCATCCGGTATTAGCCCCGGTTTCC
CGGAGTTATCCCAGTCTTACAGGCAGGTTACCCACGTGTTACTCACCCGTCCGCCGCTGACCTAAGGGAGCAAGCT
CCCGTCGGTCCGCTCGACTTGCATGTATTAGGGACGCCCCCCGCGTTCGTCCTGAGCAGAGACAAAAATCTAAAAA
AA

Figure.4: The 16s rRNA sequence of Actinomycetes strain

The sequencing of sequencing data is then trimmed and assembled manually using DNA Basser Assembler software according to Nayak *et al.,* 2009 [26]. The sequences then compared with the gene data bank at EzBio Cloud (https://www.ezbiocloud.net/identify) [27]. The sequences of rare actinomycete-like bacteria were aligned with type strains achieved from the DDJBL/EMBL/GenBank databases. Phylogenetic analysis was carried out using the Molecular Evolutionary Genetics Analysis version X software package [28]. The phylogenetic tree was arranged using the neighbor-joining method [29], maximum-likelihood [30] and maximum-parsimony [31] in MEGAX software with a bootstrap value based on 1000 replications and distance calculation according to the Kimura two-parameter method [32].

From the 16 s r RNA sequencing and subsequent BLASTn analysis, it was found that the isolated LMA4 was a strain of *Microbacterium barkeri*. Figure 4 is displaying a Phylogenetic tree, inferring about the phylogenetic linkage of *Microbacterium barkeri* (LMA4) with other similarly or identically attached actinomycetes strains16s r RNA cluster.

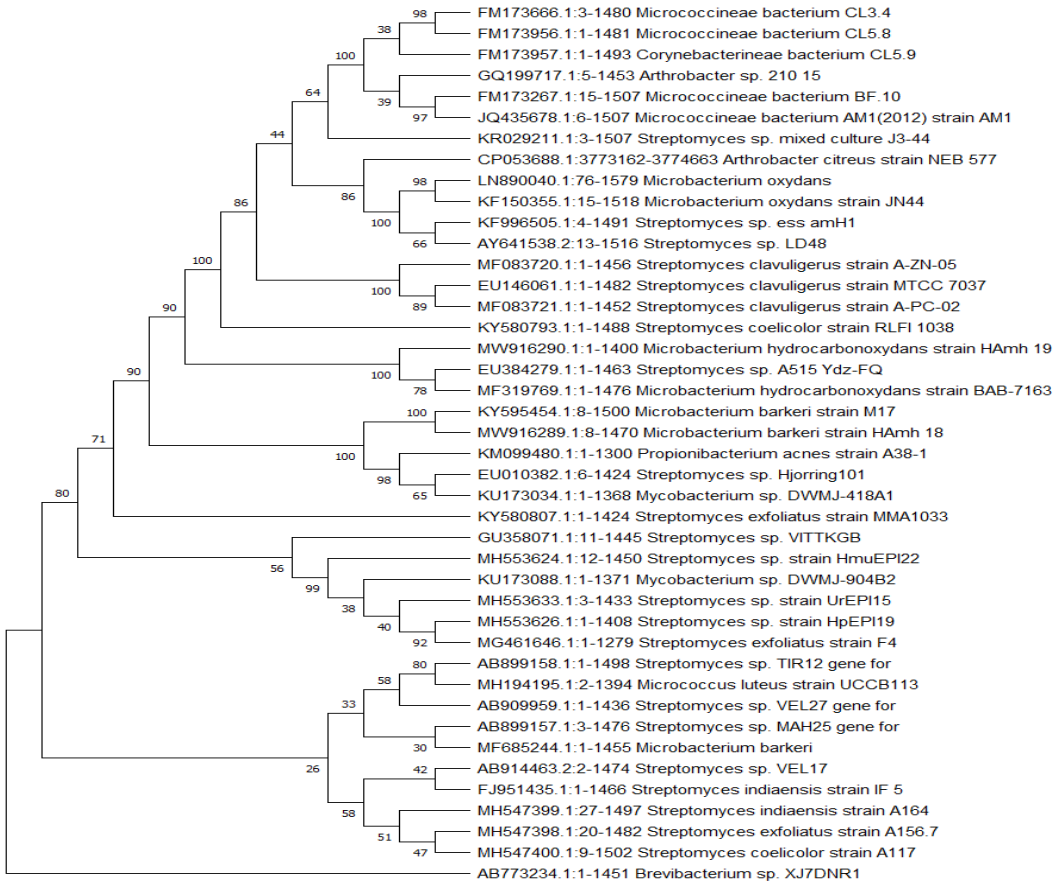


Figure.5: The Phylogenetic Tree of ***M. barkeri*** with other strains of Actinomycetes species

The genome studies of *Microbacterium sp.* with a total of 3559 genes mapped to discrete genomic locations of *Microbacterium* genome. From the bioinformatics functional enrichment analysis reported 43 genes and 5 GO terms for biological processes (BP), 63 genes and 5 GO terms for molecular functions (MF) and 61 genes and 4 GO terms for cellular components (CC). The Venn diagram of the genes related to *Microbacterium sp.* for BP, CC and MF is shown in the below **figure.6**

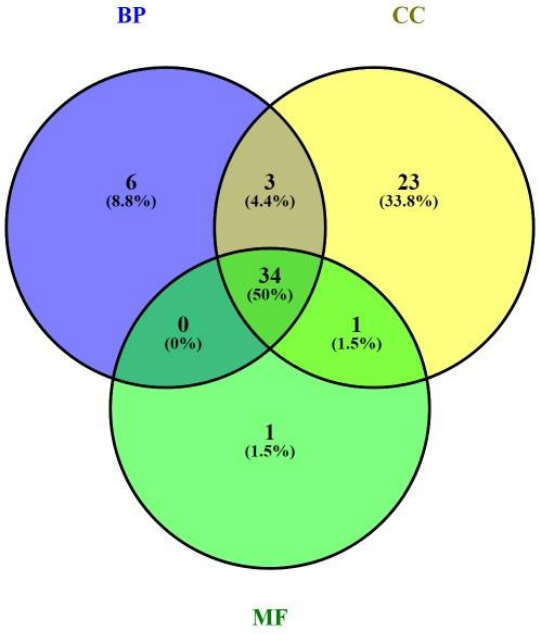


Figure.6: Venn diagram of genes characterized by BP, CC and MF category

Based on the essential role of biological processes, 43 genes and 5 GO terms obtained from significantly enriched biological processes are termed as key genes that were used for network construction of *Microbacterium Barkeri* (LMA4). The key genes were used for network construction of *Microbacterium Barkeri* (LMA4) were analyzed through STRING database. The result of the string is represented in **Figure.7**. The **LMA4** network of STRING database reported the genes namely RPLP, RPSL, RPSR, RPSE, RPMA, RPLB, RPSG, RPSS, RPMG and RPLQ at the core region of the network. These 10 genes may be said to play a key role in different biological functions of *Microbacterium Barkeri* (LMA4) as well as they can be differentially expressed in many aspects caused by the *Microbacterium sp*. specially LMA4.

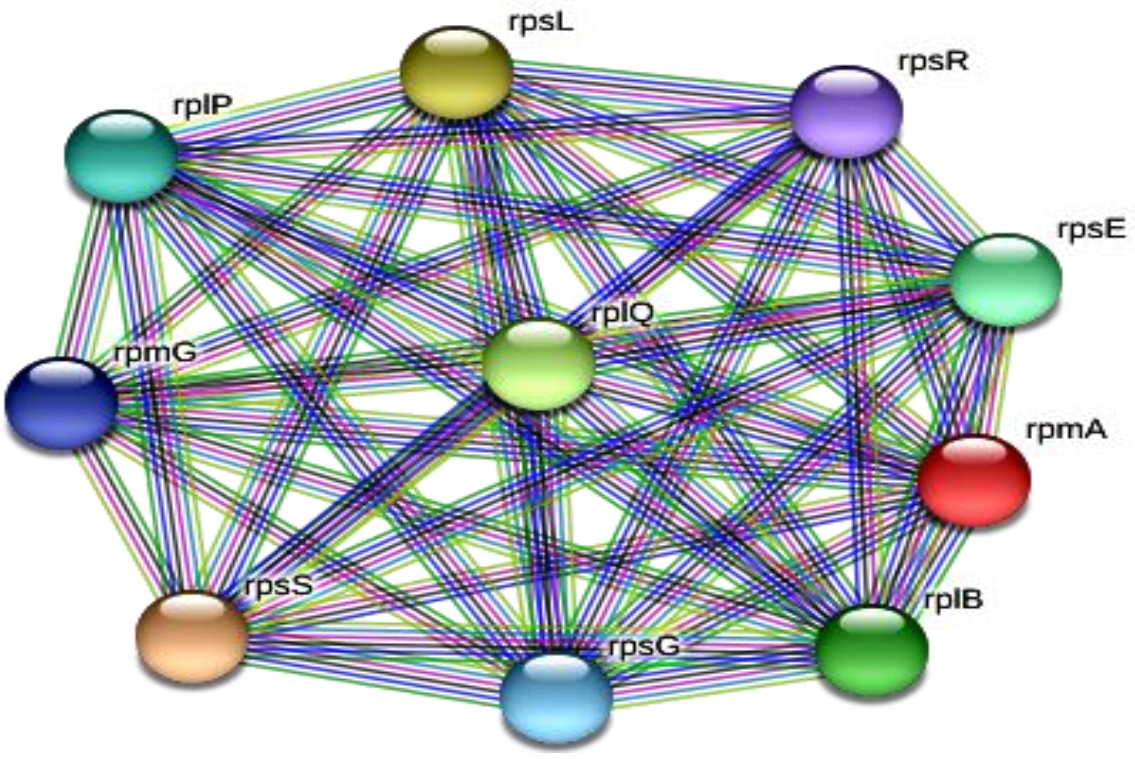


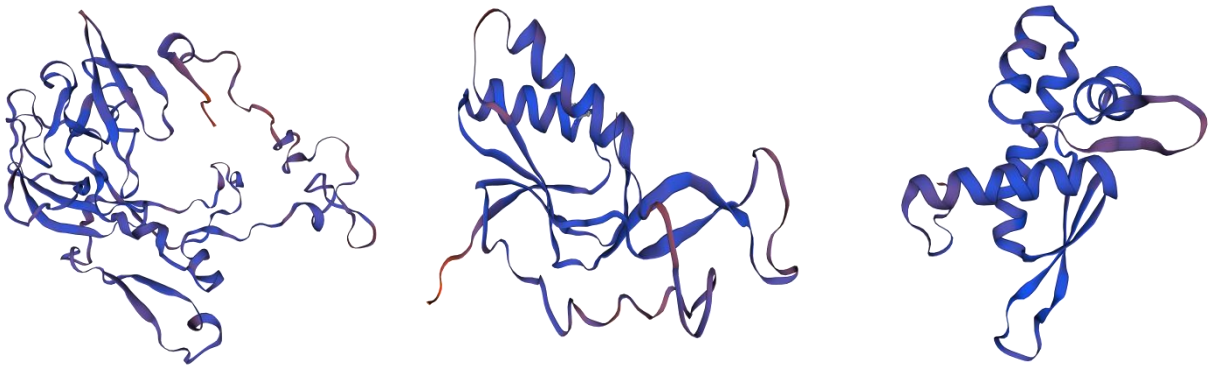
Figure.7: Associated genes interaction of *Microbacterium barkeri*

The amino acid sequences of corresponding key proteins encoded by the reported 10 genes (RPLP, RPSL, RPSR, RPSE, RPMA, RPLB, RPSG, RPSS, RPMG and RPLQ) retrieved from UniProtKB database. The 3D structure of all the 10 proteins is predicted by homology modeling and threading tool named SWISS- MODEL due to unavailability of information at Protein data Bank (PDB). Validation and quality estimation of the generated structure using SWISS- MODEL online server suggests that the quality of the generated structure is good and can be used for protein-ligand studies. Details of the template and the modeled structure is given in the below.

Table.1: Table for the Modeled structure details and the template details of interacted genes of *Microbacterium barkeri* (LMA4)

S.L NO	Protein name	UniProt IDS	Template Id	Seq Identity	Full name of the Template	Ramachandran Favoured %
1	RPLP	P14577	7aqd.1.M	100.00%	50S ribosomal protein L16	97.79%
2	RPSL	P21472	7qv2.1.g	100.00%	30S ribosomal protein S12	91.79%
3	RPSR	P21475	6ha1.1.k	100.00%	30S ribosomal protein S18	95.16%
4	RPSE	P21467	6ha1.1.7	100.00%	30S ribosomal protein S5	96.91%
5	RPMA	P05657	7aqc.1.U	100.00%	50S ribosomal protein L27	97.50%
6	RPLB	P42919	7aqd.1.D	100.00%	50S ribosomal protein L2	96.70%
7	RPSG	P21469	7qv2.1.b	100.00%	30S ribosomal protein S7	98.01%
8	RPSS	P21476	7qv2.1.n	100.00%	30S ribosomal protein S19	96.05%
9	RPMG	Q06798	7kgb.1.B	51.02%	50S ribosomal protein L33 2	95.65%
10	RPLQ	P20277	7aqd.1.N	100.00%	50S ribosomal protein L17	97.44 %

The 3D structure of the unidentified predicted gene structure was constructed and visualized by the bioinformatics visualization tool named “Discovery Studio 4.0 client”. The newly modeled structures were given in the below Figures (**Figure.8**).



RPLB **RPLP** **RPLQ**

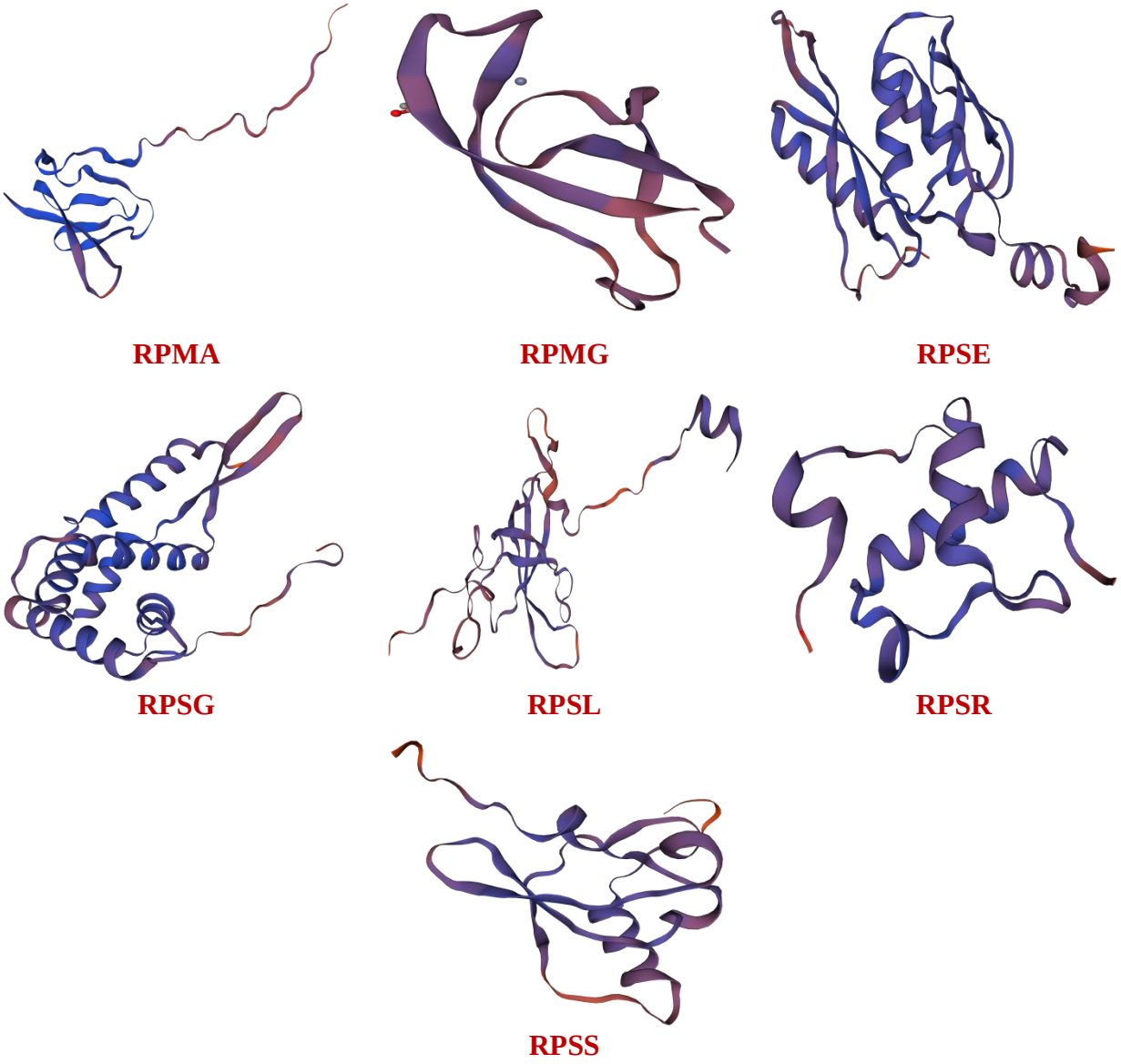


Figure.8: The modeled structure of the interacted genes of *Microbacterium barkeri* (LMA4)

The validation and quality estimation of the figures for the modeled structure was done by the Ramachandran plot analysis by the online SWISS-MODEL server. So the % of identity or favorable percentage amount was mentioned in the Table.1 and also the plots are shown in the below **Figure.9**

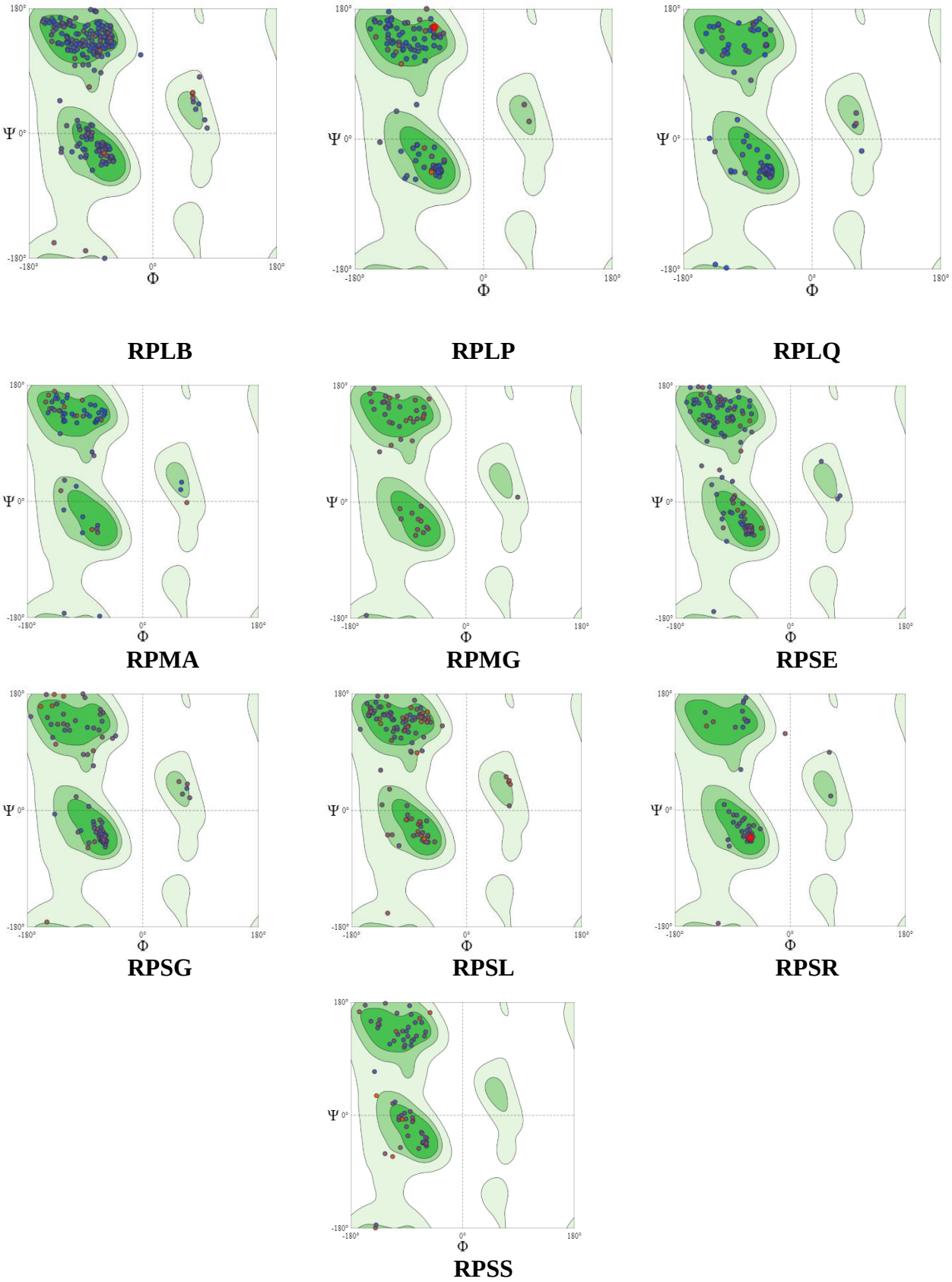


Figure.9 : Ramachandran Plot validation for modeled structure of genes predicted from the NCBI for *Microbacterium barkeri* (LMA4)

CONCLUSION

This research may pave a way for identification of bioactive interacted genes from novel strain of *Microbacterium barkeri* which will help the genomic study along with the gene-gene docking study against the pathogenic bacteria gene to produce a new antibiotics compound prediction against diseases.

Acknowledgement: This research is part of my thesis work registered under Sambalpur University.

REFERENCE

1. Priscila Gava Mazzola1,*, Angela Faustino Jozala2, LetíciaCélia de Lencastre Novaes2, Patricia Moriel1, Thereza Christina Vessoni Penna2, Minimal inhibitory concentration (MIC) determination of disinfectant and/or sterilizing agents, Brazilian Journal of Pharmaceutical Sciences, vol. 45, n. 2, abr./jun., 2009.
2. Gouse Basha Sheik[1*], Muazzam Sheriff Maqbul[2], Gokul Shankar S[3], Ranjith Ms[4],Isolation And Characterization Of Actinomycetes From Soil Of Ad-Dawadmi, Saudi Arabia And Screening Their Antibacterial Activities, *Int J Pharm PharmSci, Vol 9, Issue 10, 276-279,* DOI: http://dx.doi.org/10.22159/ijpps.2017v9i10.15402.
3. 16S rRNA gene sequencing for bacterial identification of pullulanase synthesizing thermophilic bacteria contributing to big data, IJCS 2018; 6(2): 2769-2773 © 2018 IJCS, International Journal of Chemical Studies 2018; 6(2): 2769-2773.
4. Augustine SK, Bhavsar SP and Kapadnis BP. 2005. A Non-Polyene Antifungal Antibiotic From *Streptomyces Albidoflavus PU 23. Journal of Biosciences, 30(2), 201-211.*
5. Elliah P, Ramana T, Bapi Raju KVVS, Sujatha P, Uma Sankar AM. Investigation on marine actinomycetes from the bay of Bengal near Karnataka coast of Andhra Pradesh. Asian J Microbiol Biotechnol Environ Sci 2004;6:53-6.
6. Aghamirian RM, Ghiasian AS. Isolation and characterization of medically important aerobic actinomycetes in the soil of Iran. Open Microbiol J 2009;3:53-7.
7. Reddy NG, Ramakrishna DPN, Rajagopal SV. Morphological, physiological and biochemical studies of marine Streptomyces rochei (MTCC 10109) showing antagonistic activity against human pathogenic microorganisms. Asian J Biol Sci 2011; 4:1-14.
8. Ramendra SP, Charu S, Pragati S, Ajay K. Isolation and screening of antimicrobial and extracellular pigment producing actinomycetes from Chambal Territory of Madhya Pradesh region, India. Asian J Pharm Clin Res 2016;9:157-60.
9. KUster, E., Williams, S. T. (1964): Production of hydrogen sulphide by streptomycetes and methods for its detection. Appl. Microbiol. 12,46-52.
10. Keinath, A. P., Loria, R. (1990): Melanin producing *Streptomyces* spp. respond to potato plant growth and differentially to potato cultivars. Can. J. Microbiol. 36, 279-285.

11. Shirling, E. B., Gottlieb, D. (1966): Methods for characterization of *Streptomyces* species. Int. J. Syst. Bacteriol. 16, 312-340.
12. Lane DJ (1991). 16S/23S rRNA sequencing. In: Stackebrandt et al.(eds) Nucleic acid techniques in bacterial systematic. John Wiley and Sons, New York, pp. 115-175.
13. Gordon RE, Mihm JM (1962). Identification of *Nocardia caviae* comb. nov. Ann. NY. Acad. Sci. 98:628-639.
14. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance and Maximum Parsinomy Methods. Mol. Bio. Evol. 28:2731-2739.
15. Jukes TH, Cantor CR (1969). Evolution of protein molecules. In Munro H.N (ed.) Mammalian protein metabolism, Acedemic Press, New York. pp. 21-132.
16. Saitou N, Nei M (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425.
17. Chen L, Xiong Z, Sun L, Yang J, Jin Q (2012) VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res 40: D641–645.
18. Zhou CE, Smith J, Lam M, Zemla A, Dyer MD, et al. (2007) MvirDB–a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. Nucleic Acids Res 35: D391–394.
19. Gulig PA, de Crecy-Lagard V, Wright AC, Walts B, Telonis-Scott M, et al. (2010) SOLiD sequencing of four Vibrio vulnificus genomes enables comparative genomic analysis and identification of candidate clade-specific virulence genes. BMC Genomics 11: 512.
20. Wegmann U, O'Connell-Motherway M, Zomer A, Buist G, Shearman C, et al. (2007) Complete genome sequence of the prototype lactic acid bacterium Lactococcus lactis subsp. cremoris MG1363. J Bacteriol 189: 3256–3270.
21. Altschul SF (1997) Evaluating the statistical significance of multiple distinct local alignments. In: Suhai S, editor. Theoretical and Computational Methods in Genome Research. New York: Plenum. 1–14.
22. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods," Molecul. Biol. Evol., vol. 28, pp. 2731–2739, 2011.
23. Neha Salaria, Sarika Sharma, Isolation and Screening of Actinomycetes from Soil Sample of Dal Lake (Kashmir) Against Selected Pathogens, ISSN NO: 2395-0730, Intl J Engg Sci Adv Research 2016 March; 2(1):51-54.
24. Narendra Kumar1*, Ravi Kant Singh1, Mishra S.K.1, Singh A.K.2, Pachouri U.C.3, Isolation and screening of soil *Actinomycetes* as source of antibiotics active against bacteria, International Journal of Microbiology Research, ISSN: 0975-5276, Volume 2, Issue 2, 2010, pp-12-16.
25. Nonaka L, Maruyama F, Onishi Y, Kobayashi T, Ogura Y, Hayashi T, Suzuki S, Matsuda M 2014 *Front. Microbiol*. **5** (152) 1-15

26. Nayak S N, Balaji J, Upadhyaya H D, Hash C T, Chattopadhyay P B K, Rodriquez L M, Blair M W, Baum M, McNall K, This D, Hoisington D A, and Varshney R K 2009 *Plt. Scie.***177** 460-467.
27. Yoon S, Ha S, Kwon S, Lim J, Kim Y, Seo H, and Chun J 2017 *J.Syst.Evol.Microbiol* **67** 1613-1617
28. Kumar S, Stecher G, Li M, Knyaz C, and Tamura K *MoL. Biol. Evol.* **35:**1547-49
29. Saitou N and Nei M 1987 *Mol. Biol. Evol.***4**(4):406-425.
30. Felsenstein J 1981 *J. Mol. Evol.***17** (6) 368-76
31. Fitch W M 1971 *Syst.Zool.***20** 406-416
32. Y Eshananda1, F Ningsih1,2, Y Sakai3,4, A Yokota3, S Yabe3,4, W Sjamsuridzal1,2, based on 16S ribosomal RNA gene sequences, Journal of Physics: Conference Series, 1524 (2020) 012062 , doi:10.1088/1742-6596/1524/1/012062.
33. Hadi Maleki1*, Omid Mashinchian2, Characterization of *Streptomyces* Isolates with UV, FTIR Spectroscopy and HPLC Analyses, *BioImpacts*, 2011, 1(1), 47-52
34. Wang, Z., Chen, Y. and Li, Y., 2004. A brief review of computational gene prediction methods. Genomics, proteomics & bioinformatics, 2(4), pp.216-221.
35. B.T. Sherman, M. Hao, J. Qiu, X. Jiao, M.W. Baseler, H.C. Lane, T. Imamichi and W. Chang. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Research. 23 March 2022. 50(W1):W216-W221. doi:10.1093/nar/gkac194.[PubMed].
36. Szklarczyk D*, Gable AL*, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P‡, Jensen LJ‡, von Mering C‡, The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets, Nucleic Acids Res. 2021 Jan 8;49(D1):D605-12. [PubMed]
37. Wang Y, Wang Q, Huang H, Huang W, Chen Y, McGarvey PB, Wu CH, Arighi CN, UniProt Consortium.A crowdsourcing open platform for literature curation in UniProt, Plos Biology. 19(12):e3001464 (2021)
38. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46(W1), W296-W303 (2018).
39. Yuan, S., Chan, H.S. and Hu, Z., 2017. Using PyMOL as a platform for computational drug design. Wiley Interdisciplinary Reviews: Computational Molecular Science, 7(2), p.e1298.
40. Kim S, Cheng T, He S, Thiessen PA, Li Q, Gindulyte A, Bolton EE. PubChem Protein, Gene, Pathway, and Taxonomy Data Collections: Bridging Biology and Chemistry through Target-Centric Views of PubChem Data. J Mol Biol. 2022 Jun 15; 434(11):167514. doi:10.1016/j.jmb.2022.167514.