



Molecular identification of chlorpyrifos and malathion degrading bacteria

Sumanjali Avanigadda^{1*} and Meena Vangalapati¹

¹Department of chemical engineering, Andhra University, Visakhapatnam

Article History

Received: 28-05-2022

Revised: 17-06-2022

Accepted: 29-07-2022



Abstract

Fifteen strains of bacteria were successfully isolated from agriculture soil samples contains pesticides and from soil moisture around them. The isolates were capable of utilizing chlorpyrifos (Cp) and malathion as the sole source of carbon, phosphorus and energy. High yielding isolate KSCM-08 was identified based on 16S rRNA sequencing. Genomic DNA was successfully extracted from bacterial isolate KSCM-08, and the total DNA content was determined to be $734 \pm 23 \mu\text{g/gm}$. The absorbance ratio of extracted DNA at A260/280 has been found to be 1.96 ± 0.2 . 1381 nucleotide sequence of the PCR amplified 16s rRNA gene from bacterial isolate KSCM-08. The 16s rRNA gene sequence comprises 344 bp of adenine (A), 280 bp of thymine (T), 430 bp of guanine (G), and 327 bp of cytosine (C). The G+C percentage was determined to be 54.8%. The BLAST report demonstrates the similar sequences with the KSCM-08 16s rRNA sequence, along with their percentage of identity. The BLASTn analysis of the 1381 bp length 16s rRNA gene sequence of KSCM-08 against the GenBank database revealed that KSCM-08 exhibits similarity with *Bacillus* species. The evolutionary divergence of bacterial isolate KSCM-08 with its relative members was determined. A satisfactory result was established by the use of the 16S rRNA gene as a marker to evaluate the phylogenetic relationship and the bacteria was confirmed as *Bacillus safensis*.

Keywords: Molecular characterization, sequencing, phylogenetic relationship and dendrogram

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Copyright © 2022 Author(s) retain the copyright of this article.



*Corresponding Author

Sumanjali Avanigadda

<https://doi.org/10.37022/jiaps.v7i2.339>

Production and Hosted by

www.saap.org.in

Introduction

The term pesticide covers a wide range of compounds including insecticides, fungicides, herbicides, rodenticides, molluscicides, nematocides, plant growth regulators and others [1]. Organophosphorus compounds (OP), which are a group of highly toxic agricultural chemicals widely used for plant protection, have generated a number of environmental problems such as contamination of air, water and terrestrial ecosystems, they have harmful effects on different biota, and disrupt biogeochemical cycling [2,3]. Malathion (Mt) and chlorpyrifos (Cp) as insecticides and acaricides, are the most commonly used OPs.

A reliable cost-effective technique for pesticide removal is to biodegrade the organophosphate compounds. In general, microorganisms demonstrate considerable capacity to metabolize many pesticides. They possess the unique ability to completely mineralize many aliphatic, aromatic, and heterocyclic compounds [4]. Many microorganisms can specifically hydrolyze the phosphoester bonds of OPs and thus reduce the toxicity of OP pesticides and OP chemical warfare agents (e.g. sarin). The study of Munnecke showed that the rate of enzymatic hydrolysis was two to 450 times faster than that of chemical hydrolysis, when parathion was used as a substrate. The present study aimed to characterize the chlorpyrifos and malathion degrading bacteria using 16s rRNA sequencing and BLAST methods.

Materials and Methods

Molecular identification

KSCM-08 had the highest chlorpyrifos and malathion tolerance out of the 15 active bacterial isolates [5, 6]. Hence, 16s rRNA gene sequencing was performed to describe KSCM-08 molecularly in order to identify the species.

Extraction of genomic DNA

The genomic DNA of the active bacterial isolate KSCM-08 was extracted by using CTAB lysozyme technique. To extract genomic DNA, the isolate KSCM-08 was maintained in nutrient broth for two days at 30°C under the constant agitation at 120 rpm. For PCR amplification, the strain's genomic G+C content was measured with the help of thermal denaturation technique [7].

Amplification and sequencing of 16S rRNA gene

The PCR amplification and sequencing of the 16s rRNA gene were carried out by the methodology of Li et al., (2010). The amplification of the 16s rRNA gene from the bacterial isolate KSCM-08 was performed by using two universal primers such as 27F-5' AGAGTTTGATCCTGGCTCAG-3' and 1492R 5'-CGGTTACCTTGTTTACGACTT-3' [8]. 25 µl of PCR reaction buffer comprised of 1 unit of Taq DNA polymerase, 2 µl of 10X PCR buffer with 20 mM MgSO₄, 0.2 mM dNTPs, 20 pmol of each forward and reverse primer, and 25 ng of template DNA. The following procedure was used for 16s rRNA gene amplification in a thermocycler: an initial denaturation step was performed at 94°C for 5 minutes and then PCR was adjusted to 40 cycles. In each cycle, the denaturation was conducted at 94°C for 1 minute, annealing was conducted at 48°C for 1 minute, and extension was conducted at 72°C for 2 minutes. The final elongation step was performed for 5 minutes at 72°C, followed by cooling at 4°C. The quality of the amplified PCR product was assessed by electrophoresis on a 1.5% agarose gel containing 0.5 µg/ml ethidium bromide, and photographed. Standard DNA samples with molecular weights ranging from 2 kb to 100 bp were utilized as ladder sequences to determine the amplified product molecular size. The purified PCR products were subjected to Sanger's di-deoxy sequencing in both forward and reverse directions on an ABI Prism 3700 DNA analyzer using the big dye terminator v3.1 cycle sequencing kit [9].

Sequence analysis and phylogenetic tree construction

The nucleotide composition of the 16s rRNA gene from bacterial isolate KSCM-08 was computed by the SeqstateV.1.21 server (Muller, 2005). The molecular identification and evolutionary analysis of bacterial isolate KSCM-08 were determined by constructing a phylogenetic tree with the related homologous and outer group bacterial species. The 16s rRNA gene sequence of bacterial isolate KSCM-08 was run through the NCBI server's BLASTn to identify homologous sequences or species, and 12 species were selected from the NCBI database based on their degree of similarity with the target gene sequence. The selected gene

sequences were retrieved from the nucleotide data base of GenBank for the evolutionary analysis and phylogenetic tree construction. The homology and evolutionary statistics of the KSCM-08 16s rRNA gene sequence with the selected 12 sequences were performed by multiple sequence alignment (MSA) with an advanced cluster method UPGMA, using Muscle software embedded in MEGA-X. A distance matrix was calculated to study the differences among the sequences. Based on the differences as expressed in the distance matrix, a maximum-parsimony phylogenetic tree was constructed using MEGA-X. The reliable test of the tree was done by the bootstrap method 1000 times.

Results and discussion

Molecular and Phylogenetic characterization bacterial isolate KSCM-08

Microbial community characterization is significantly more varied, with a wide range of abilities. For identifying such types of microbial communities, molecular approaches including 16s rRNA isolation and sequencing techniques have been developed. Moreover, these approaches have been proven to be effective tools for identifying bacterial species in microbiological research. The 16s rRNA method by PCR amplification was used in this research to identify the bacterial isolate KSCM-08, and the molecular characterization also helps to understand the evolutionary links between KSCM-08 and its related species. The Maximum Composite Likelihood method was applied to calculate the evolutionary divergence [10].

Total DNA quantification

The results were presented as Mean ± Standard deviation (SD) of three distinct experiments. Genomic DNA was successfully extracted from bacterial isolate KSCM-08, and the total DNA content was determined to be 734±23 µg/gm. The absorbance ratio of extracted DNA at A₂₆₀/A₂₈₀ has been found to be 1.96±0.2. The thermal denaturation midpoint (T_m) of genomic DNA from KSCM-08 was calculated as 93°C. The molar percentage of G+C content was calculated according to the formula proposed by Marmur and Doty, (1962) by using the T_m value. The G+C molar percentage of KSCM-08 genomic DNA was found to be 57.8%. The T_m of DNA is directly proportional to the molar percentage of G+C content. In the Marmur equation, 69.3 is the value derived by estimating the DNA with 100 mol% A+T intersect the temperature axis at 69.3°C, and slope of r_{bvdbdsew} regression is 0.41. The extracted genomic DNA of KSCM-08 was displayed as a conspicuous band on a 1% agarose gel. Figure 1 depicts the genomic DNA band in a 1% agarose gel.

$$\text{Mol \% G + C} = \frac{T_m - 69.3}{0.41}$$

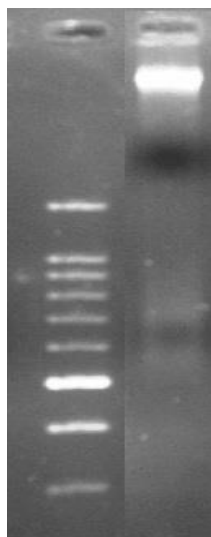


Figure: 1. Photograph showing bacterial isolate KSCM-08 genomic DNA in 1% agarose gel.

PCR amplification and sequencing of 16s rRNA gene

In the PCR experiment, genomic DNA extracted from the bacterial isolate KSCM-08 was used to amplify the 16s rRNA gene using two universal primers such as 27F-5' AGAGTTTGATCCTGGCTCAG-3' and 1492R 5'-CGGTACCTGTTTACGACTT-3'. The 16s rRNA gene has been regularly employed for phylogenetic analyses due to its high primer specificity and discriminatory strength (Schneider et al., 2004). The primers used to amplify the 16s rRNA of bacterial isolate KSCM-08 produce distinct single bands with good quality. Figure 2 shows the PCR amplified 16s rRNA gene band in a 1.5% agarose gel. In a 1.5% agarose gel, the first well contains marker DNA, while the second has an amplified gene product of about 1500bp size. The PCR amplified gene fragment was purified for further studies, including nucleotide sequencing and phylogenetic analysis.

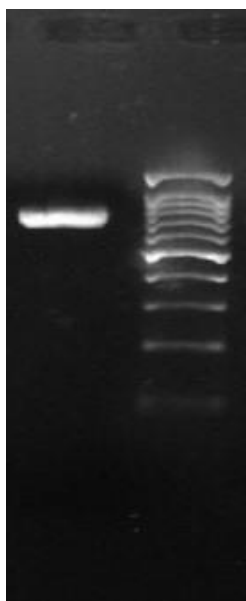


Figure: 2. Photograph showing the PCR amplified 16S rRNA gene band in 1.5% agarose gel.

Microbiology aspires to quick advancement by finding and describing microbial species in order to create an accurate phylogenetic connection. According to Salipante et al., (2014), identification of bacterial species using traditional methods is extremely challenging since certain bacteria are difficult to cultivate owing to poor growth and others are non-culturable due to intolerance. As a result, bacterial species identification by traditional methods might take days to weeks. These circumstances have resulted in the development of an alternative method for the quick and precise identification of bacterial species through the isolation and sequencing of the 16s rRNA gene (Srinivasan et al., 2015). Clegg and Zurawski, (1991) mentioned that the sequencing of PCR amplified gene products is significantly expanding the phylogenetic and evolutionary research of microbes. According to the requirements, an ideal DNA barcode should be consistently recoverable with a single primer pair, be compatible for bidirectional sequencing with little manual editing of sequences, and provide maximal species identification. The 16s rRNA gene is well described but not polymorphic in the prokaryotic genome, and it is easily retrieved using standard PCR primers. The 16s rRNA is one of the main constituents in the 30s ribosomal subunit of prokaryotes and it plays an important role in protein synthesis by stabilising the codon and anticodon pairing during translation (Wimberly et al., 2000).

The 16s rRNA gene has become an essential evolutionary indicator and it can be used to determine the three aspects of life due to its functional consistency and highly conserved existence (Woese et al., 1990). Sabat et al., (2017) stated that the 16s rRNA is a highly conserved 1.5 kilobase pair size gene that is a prime target of phylogenetic research since it is ubiquitous in all bacteria as a single or multiple copies. Woo et al., (2008) has been suggested that identification and evolutionary relationships of uncultivable bacteria were made possible with 16s rRNA sequencing. In several cases, the 16s rRNA gene sequence has been primarily utilised to differentiate the Actinomycetes and non-Actinomycetes bacteria from anaerobic gram-positive bacilli, which is typically problematic in clinical microbiology (Colmegna et al., 2003). Tazi et al., (2014) reported that the direct sequencing of 16s rRNA genes from environmental samples has become a common and effective way to analyse microbial community prevalence, organisation, and function. The current study took advantage of the 16s rRNA gene since the coding region of the 16s rRNA gene is easily amplified and sequenced and it has an influence on phylogenetic studies by giving consistent allocation of a taxon into a group and genus.

Sequence analysis

Seqstate v.1.21 server was used to determine the characteristics of sequence including nucleotide composition and GC percentage (Muller, 2005). Figure 3 shows the 1381 nucleotide sequence of the PCR amplified 16s rRNA gene from bacterial isolate KSCM-08. The 16s rRNA gene sequence comprises 344 bp of adenine (A), 280 bp of thymine (T), 430 bp of guanine (G), and 327 bp of cytosine (C). The G+C percentage was determined to be 54.8%. The nucleotide composition of the KSCM-08 16s rRNA gene is shown in table 1.

Table: 1. Nucleotide composition of KSCM-08 16s rRNA gene

S. No	Parameter	Number of nucleotides	% of nucleotides
1	Total bases	1381	--
2	Adenine	344	24
3	Thymine	280	22
4	Guanine	430	31
5	Cytosine	327	23
6	G+C	757	54.8

```

gtccccggatgtagcggcgacgggtgagtaacacgtgggtaacctgacctgtaagactgggataactccgggaaaccggagctaataccggat
agttcctgaaccgatggtcaaggatgaaagacgggttcggctgcaactacagatggaccgcggcgattagctagtgtggggtaatggct
caccaggcgacgatgcgtagccgacctgagagggtgatcgccacactgggactgagacacggcccagactcctacggggagcagcagta
gggaatcttcgcaatggacgaaagtctgacggagcaacgccgctgagtgatgaagggtttcggatcgtaaagctctgtgttagggaagaaca
agtgcgagagtaactgctgcaccttgacggtacctaaccagaaagccacggctaactacgtgccagcagcccggttaatactaggtggcaag
cgttgtccggaattattggcgtaaaaggctcgcaggcgggtttctaagtctgatgtaaagccccggctcaaccggggagggtcattgaaact
gggaaactgagtgagaagaggagtggaattccacgtgtagcgggtgaaatgcgtagagatgtggaggaaaccagtggcgaagggcact
ctctgctgtaactgacgctgaggagcgaagcgtggggagcgaacaggattagataacctggtagtcacgccgtaaacgatgagtgtaag
tgtagggggtttccgcccttagtgctgcagctaacgcactcagcctgggggagtagcggtcgcaagactgaaactaaaggaattgacg
ggggcccgcacaagcgggtggagcatgtggttaattcgaagcaacgcgaagaacctaccaggtctgacatcctctgacaacctagagatagg
gcttccctcggggacagagtgacaggtggtgcatggttgcctcagctcgtgctgagatgttgggtaagtcccgaacgagcgaaccctt
gatcttagtgcagcattcagttgggcactctaagggtactgccggtgacaaccgggaggaaggtggggatgacgtcaaatcatcatgccctta
tgacctgggctacacacgtgctacaatggacagaacaaagggtgcaagaccgaaggttagccaatccataaatctgttctcagttcggatcg
cagctgcaactcactgctgaaagctggaatcgtagtaatcgcgatcagatgccggtgaaatcgtcccggccttgacacaccgcc
gtcacaccagagagttgcaacaccgaagtcggtgaggttaaccttatggag
    
```

Figure: 3. Nucleotide sequence of KSCM-08 16s rRNA gene (1381 bp).

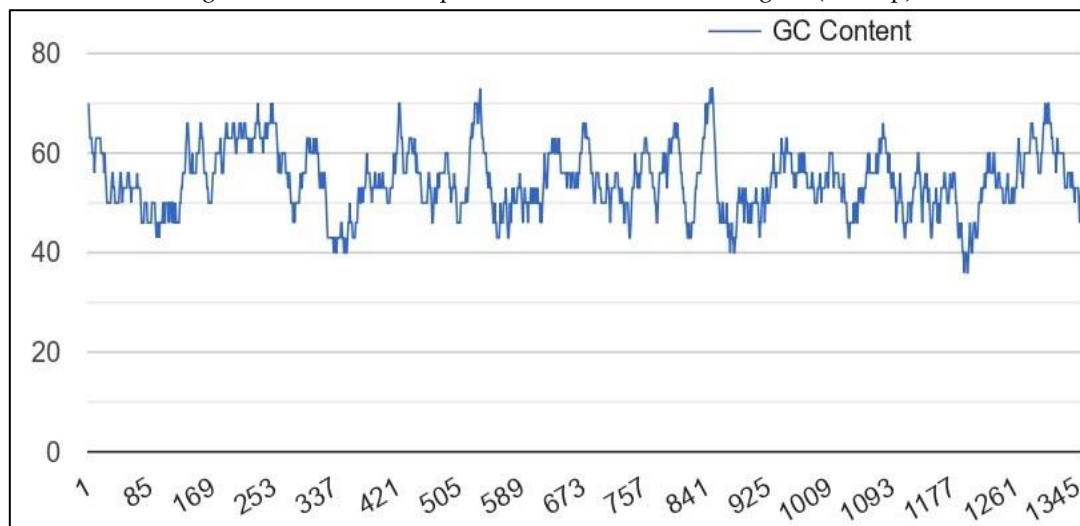


Figure: 4. GC distribution over the amplified 16s rRNA gene sequence of KSCM-08.

The GC content of each and every species is a highly dynamic characteristic. Lynch, (2007) has been stated that the GC composition of bacterial species can range from 25 to 75 percent. From the current findings, the 16s rRNA gene from bacterial isolate KSCM-08 has a relatively high GC content, which is consistent with earlier research. Hildebrand et al., (2010) found that GC content varies greatly in bacterial species, with variations ranging from 13 to 75% in different species. Although the specific causes for these variances in GC content within and between the species are unknown, Mann and Chen, (2010) predict that several factors are involved in the GC variation, which includes evolutionary developments and environmental impacts. Sinden, (1994) proposed that high GC content might be promoted by natural selection for more resilient DNA since the stacking of guanine and cytosine costs greater energy than the stacking of adenine and thymine. This is also evident with the results of Zhang and Gao, (2017) who found that the increasing percentages of amino acids translated by GC rich codons contribute positively to increased evolutionary trends while the amino acids are selectively adopted to be in GC rich species. Hence, high GC composition makes microbial genomes less common than AT nucleotides, implying that bacterial genes have a high GC content. Foerstner et al., (2005) has been reported that the microbial communities cohabiting the same habitats tend to have a similar percentage of GC regardless of taxa.

Multiple sequence alignment

The 16s rRNA gene sequence of bacterial isolate KSCM-08 was used as a query sequence for nucleotide BLAST to determine the similarities with other species' 16s rRNA gene sequences which are available in the GenBank database. The BLAST report demonstrates the similar sequences with the KSCM-08 16s rRNA sequence, along with their percentage of identity. The BLASTn analysis of the 1381 bp length 16s rRNA gene sequence of KSCM-08 against the GenBank database revealed that KSCM-08 exhibits similarity with *Bacillus* species. Figure 5 shows the result of the BLASTn search against the KSCM-08 16s rRNA gene sequence. According to the findings of the BLASTn search, the 16s rRNA gene sequence of bacterial isolate KSCM-08 has the maximum resemblance to *Bacillus pumilus*, with an identity of 100% and an E value of 0.0. From the matches, a total of 12 sequences belonging to the *Bacillus* genus were retrieved from the GenBank database and employed for multiple sequence alignment. The selected sequences for the multiple sequence alignment and phylogenetic tree construction were listed in table 2. CLUSTAL was used to produce a multiple sequence alignment of the 16s rRNA region (Higgins et al., 1992). Furthermore, multiple sequence alignment of the 16s rRNA

gene sequences revealed that the gene sequence is highly conserved across the prokaryotes.

Table: 2. List of selected species sequences used for MSA and phylogenetic study

Species	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession	Accession
<i>Bacillus pumilus</i>	255 1	255 1	100 %	0.0	100.0 0%	14 50	MT3677 13.1
<i>Bacillus pumilus</i>	255 1	255 1	100 %	0.0	100.0 0%	14 54	MT0721 59.1
<i>Bacillus safensis</i>	255 1	255 1	100 %	0.0	100.0 0%	14 51	MT0717 37.1
<i>Bacillus australimaris</i>	255 1	255 1	100 %	0.0	100.0 0%	14 89	MT0108 36.1
<i>Bacillus australimaris</i>	255 1	255 1	100 %	0.0	100.0 0%	14 74	ON3068 35.1
<i>Bacillus australimaris</i>	255 1	255 1	100 %	0.0	100.0 0%	14 74	ON3068 22.1
<i>Bacillus safensis</i>	255 1	255 1	100 %	0.0	100.0 0%	14 27	ON2871 04.1
<i>Bacillus australimaris</i>	255 1	255 1	100 %	0.0	100.0 0%	15 12	OM658 338.1
<i>Bacillus velezensis</i>	255 1	255 1	100 %	0.0	100.0 0%	14 16	OL6392 27.1
<i>Bacillus australimaris</i>	255 1	255 1	100 %	0.0	100.0 0%	15 16	MW287 242.1
<i>Bacillus safensis</i>	255 1	255 1	100 %	0.0	100.0 0%	14 24	MT7891 18.1
<i>Bacillus safensis</i>	255 1	255 1	100 %	0.0	100.0 0%	14 23	MT7890 93.1

Phylogenetic tree analysis

The selected homologous sequences of KSCM-08 16s rRNA were extracted from the GenBank database and used to construct a phylogenetic tree. The selected species employed in this study are listed in table 2. The distance matrix was used to determine the differences between the selected 12 sequences from the genus *Bacillus*. Depending on the expressed differences in the distance matrix, the maximum-parsimony phylogenetic tree was built using MEGAX and the topologies of the phylogenetic tree were evaluated using the bootstrap method with 1000 replicates for all nodes

(Felsenstein, 1985). The evolutionary divergence of bacterial isolate KSCM-08 with its relative members was determined. A satisfactory result was established by the use of the 16S rRNA gene as a marker to evaluate the phylogenetic relationship. The use of the 16S rRNA gene as a marker to analyse the phylogenetic connection provided a satisfactory result. In the phylogenetic tree, there were 2 main clades, and the first main clade consists of 6 subclades. The first subclade of main clade 1 has again consists two clades which were composed of *Bacillus australimaris* (ON306835.1), *Bacillus australimaris* (ON306822.1), *Bacillus australimaris* (MT010836.1), *Bacillus safensis* (MT0717371), and *Bacillus australimaris* (MW287242.1). The second and third subclades of main clade 1 was composed of *Bacillus velezensis* (OL639227.1) and *Bacillus australimaris* (OM658338.1) respectively. The fourth subclade of main clade 1 was composed of *Bacillus pumilus* (MT367713.1), and *Bacillus safensis* (ON287104.1). The fifth and sixth subclades of main clade 1 consists only of KSCM-08, and *Bacillus safensis* (MT789118.1) respectively. The second subclade consists of only two subclades, which were composed of only single species such as *Bacillus pumilus* (MT072159.1) and *Bacillus safensis* (MT789093.1) respectively. The phylogenetic tree indicated that the bacterial isolate KSCM-08 has a close evolutionary link to *Bacillus safensis* species. On the basis of morphological studies and molecular phylogenetics, it is revealed that the bacterial isolate KSCM-08 belongs to the *Bacillus safensis* species. The maximum-parsimony phylogenetic tree was shown in figure 5. DNA sequence analysis in several species provides valuable information about their taxonomy, gene makeup, and utilizations. DeGroot et al., (2011) reported that the generic level identification was deemed effective when a single genus was involved in all hits with maximum percent identification scores are greater than 95%. While species identification was deemed effective only when a single species was included with the highest percent identity score of greater than 95%. In the phylogenetic tree, the clades are organised mostly with the combination of several species and strains. Therefore, generating a local barcode database is necessary for a wide range of ecological applications, including the construction of community phylogenetics (Kress et al., 2009).

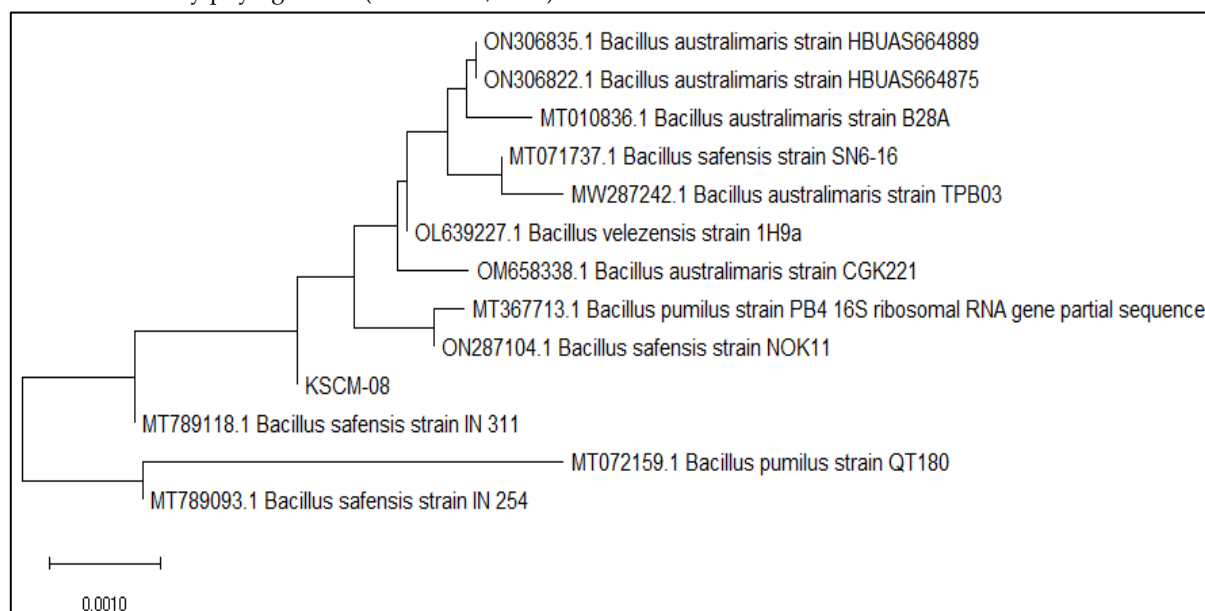


Figure: 5. Maximum Parsimony tree of bacterial isolate KSCM-08 and other relative species based on the 16S rRNA gene.

Conclusion

The results revealed that the total DNA content of KSCM-08 was determined to be 734 µg/gm, quality was found to be 1.96 and the thermal denaturation midpoint (T_m) of genomic DNA was calculated as 93°C. The 16S rRNA gene sequence comprises 344 bp of adenine (A), 280 bp of thymine (T), 430 bp of guanine (G), and 327 bp of cytosine (C). The G+C percentage was determined to be 54.8%. On the basis of morphological studies and molecular phylogenetics, it is revealed that the bacterial isolate KSCM-08 belongs to the *Bacillus safensis* species.

Funding

No funding.

Author Contribution

All authors contributed equally.

Conflict of Interest

The authors declared no conflict of interest.

References

1. Clegg MT, Zurawski G. Chloroplast DNA and the study of plant phylogeny: present status and future prospects. *Molecular systematics of plants*. 1992;1-3.
2. Colmegna I, Rodriguez-Barradas M, Young EJ, Rauch R, Clarridge J. Disseminated Actinomyces meyeri infection resembling lung cancer with brain metastases. *The American journal of the medical sciences*. 2003 Sep 1;326(3):152-5.

3. De Groot GA, During HJ, Maas JW, Schneider H, Vogel JC, Erkens RH. Use of rbcL and trnL-F as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. *PLoS one*. 2011 Jan 26;6(1):e16371.
4. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *evolution*. 1985 Jul;39(4):783-91.
5. Foerstner KU, Von Mering C, Hooper SD, Bork P. Environments shape the nucleotide composition of genomes. *EMBO reports*. 2005 Dec;6(12):1208-13.
6. Higgins DG, Bleasby AJ, Fuchs R. CLUSTAL V: improved software for multiple sequence alignment. *Bioinformatics*. 1992 Apr 1;8(2):189-91.
7. Hildebrand F, Meyer A, Eyre-Walker A. Evidence of selection upon genomic GC-content in bacteria. *PLoS genetics*. 2010 Sep 9;6(9):e1001107.
8. Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjur O, Bermingham E. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences*. 2009 Nov 3;106(44):18621-6.
9. Li M, Hong Y, Klotz MG, Gu JD. A comparison of primer sets for detecting 16S rRNA and hydrazine oxidoreductase genes of anaerobic ammonium-oxidizing bacteria in marine sediments. *Applied Microbiology and Biotechnology*. 2010 Mar;86(2):781-90.
10. Lynch M, Walsh B. *The origins of genome architecture*. Sunderland, MA: Sinauer Associates; 2007 Mar.
11. Marmur J, Doty P. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *Journal of molecular biology*. 1962 Jul 1;5(1):109-18.
12. Müller K. SeqState. *Applied bioinformatics*. 2005 Mar;4(1):65-9.
13. Sabat AJ, van Zanten E, Akkerboom V, Wisselink G, van Slochteren K, de Boer RF, Hendrix R, Friedrich AW, Rossen JW, Kooistra-Smid A. Targeted next-generation sequencing of the 16S-23S rRNA region for culture-independent bacterial identification-increased discrimination of closely related species. *Scientific reports*. 2017 Jun 13;7(1):1-2.
14. Salipante SJ, Roach DJ, Kitzman JO, Snyder MW, Stackhouse B, Butler-Wu SM, Lee C, Cookson BT, Shendure J. Large-scale genomic sequencing of extraintestinal pathogenic *Escherichia coli* strains. *Genome research*. 2015 Jan 1;25(1):119-28.
15. Schneider WL, Sherman DJ, Stone AL, Damsteegt VD, Frederick RD. Specific detection and quantification of Plum pox virus by real-time fluorescent reverse transcription-PCR. *Journal of Virological Methods*. 2004 Sep 1;120(1):97-105.
16. Sinden RR. *DNA structure and function*. Gulf Professional Publishing; 1994 Nov 24.
17. Srinivasan R, Karaoz U, Volegova M, MacKichan J, Kato-Maeda M, Miller S, Nadarajan R, Brodie EL, Lynch SV. Use of 16S rRNA gene for identification of a broad range of clinically relevant bacterial pathogens. *PLoS one*. 2015 Feb 6;10(2):e0117617.
18. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences*. 2004 Jul 27;101(30):11030-5.
19. Tazi L, Breakwell DP, Harker AR, Crandall KA. Life in extreme environments: microbial diversity in Great Salt Lake, Utah. *Extremophiles*. 2014 May;18(3):525-35.
20. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. *Journal of bacteriology*. 1991 Jan;173(2):697-703.
21. Wimberly BT, Brodersen DE, Clemons WM, Morgan-Warren RJ, Carter AP, Vonnrhein C, Hartsch T, Ramakrishnan V. Structure of the 30S ribosomal subunit. *Nature*. 2000 Sep;407(6802):327-39.
22. Woese CR, Winker S, Gutell RR. Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". *Proceedings of the National Academy of Sciences*. 1990 Nov;87(21):8467-71.
23. Woo PC, Lau SK, Teng JL, Tse H, Yuen KY. Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clinical Microbiology and Infection*. 2008 Oct 1;14(10):908-34.
24. Zhang G, Gao F. Quantitative analysis of correlation between AT and GC biases among bacterial genomes. *PLoS One*. 2017 Feb 3;12(2):e0171408.