Molecular identification of chlorpyrifos and malathion degrading bacteria
Sumanjali Avanigadda* and Meena Vengalapati
1Department of chemical engineering, Andhra University, Visakhapatnam

### Article History

<table>
<thead>
<tr>
<th>Received: 28-05-2022</th>
<th>Revised: 17-06-2022</th>
<th>Accepted: 29-07-2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>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±23 µg/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 <em>Bacillus</em> 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 <em>Bacillus safensis</em>. <strong>Keywords:</strong> Molecular characterization, sequencing, phylogenetic relationship and dendrogram.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*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, nematicides, 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'-CGGTACCTTGTTACGACTT-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 terminatorv3.1 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 A260/280 has been found to be 1.96±0.2. The thermal denaturation midpoint (Tm) 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 Tm value. The G+C molar percentage of KSCM-08 genomic DNA was found to be 57.8%. The Tm 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 rbvdbdsewsegregation 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.

\[
\text{Mol} \% G + C = \frac{Tm - 69.3}{0.41}
\]
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>
<thead>
<tr>
<th>S. No</th>
<th>Parameter</th>
<th>Number of nucleotides</th>
<th>% of nucleotides</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Total bases</td>
<td>1381</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>Adenine</td>
<td>344</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>Thymine</td>
<td>280</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>Guanine</td>
<td>430</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>Cytosine</td>
<td>327</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>G+C</td>
<td>757</td>
<td>54.8</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Species</th>
<th>Max Score</th>
<th>Total Score</th>
<th>Query Cover</th>
<th>E value</th>
<th>Per. Iden t</th>
<th>Ac. Len</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacillus pumilus</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>MT3677</td>
</tr>
<tr>
<td>Bacillus pumilus</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>MT0721</td>
</tr>
<tr>
<td>Bacillus safensis</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>MT0717</td>
</tr>
<tr>
<td>Bacillus australi maris</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>MT0108</td>
</tr>
<tr>
<td>Bacillus australi maris</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>ON3068</td>
</tr>
<tr>
<td>Bacillus australi maris</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>ON3068</td>
</tr>
<tr>
<td>Bacillus safensis</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>ON2871</td>
</tr>
<tr>
<td>Bacillus australi maris</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>ON6392</td>
</tr>
<tr>
<td>Bacillus velezens is</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>OL2421</td>
</tr>
<tr>
<td>Bacillus australi maris</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>MW287</td>
</tr>
<tr>
<td>Bacillus safensis</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>MT7891</td>
</tr>
<tr>
<td>Bacillus safensis</td>
<td>255</td>
<td>1</td>
<td>100%</td>
<td>0.0</td>
<td>100.0</td>
<td>14</td>
<td>MT7890</td>
</tr>
</tbody>
</table>

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.
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 (MT071737), and Bacillus australimaris (MW287242.1). The second and third subclades of main clade 1 was composed of Bacillus velezensis (OL639227.1) and Bacillus australimaris (OM658388.1) 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 (Tm) 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.

Author Contribution
All authors contributed equally.

Conflict of Interest
The authors declared no conflict of interest.

References


