Genetic Identification and Phylogeny of *Clerodendrum Phlomidis* Growing in Egypt Using Some DNA Barcodes

Ateto A.A¹, Sayed A.I.H¹, EL-Shaer H.F¹, Mahdy E.M.B²

¹Agricultural Botany (Genetics) Department, Faculty of Agriculture, Al-Azhar University, Cairo11884, Egypt.
²National Gene Bank (NGB), Agricultural Research Center (ARC), 9 Gamma St., 12619, Giza, Egypt.

**ABSTRACT**

*Clerodendrum* is one of the Lamiaceae genera, which has significance in taxonomical position. *Clerodendrum phlomidis* is an important wild species in Egypt and plays a main role in ecological, biodiversity, and taxonomical approaches. In the current study, we conducted this research work to investigate the identification of *C. phlomidis* collected from the North Coastal region of Egypt using three DNA barcode sequences viz., *rbcL*, ITS, and *rpoC1* genes. The investigation was conducted in the corporation between the Agric. Botany-Genetics Dept., Fac. of Agric, Al-Azhar Univ. and National Gene Bank (NGB), Agricultural Research Center (ARC), Egypt. Our findings revealed that accessions KF199889 (98.7%) and ON010669.1 (98.6%) were identical to the taxon of *C. phlomidis* by the ITS and *rbcL* sequences. On the other hand, the sequence of *C. phlomidis* revealed closer to the taxa of *C. buchneri*, *C. cephalanthum*, *C. thomsoniae*, and *C. rotundifolium* by ITS sequences, *C. indicum*, *C. thomsoniae*, *C. longiflorum*, *C. infortunatum*, *C. floribundum*, *C. cephalanthum*, *C. buchneri*, and *C. splendens* by *rbcL*, and *C. japonicum*, *C. thomsoniae*, *C. cyrtophyllum*, and *C. trichotomum* by *rpoC1* sequences. It can be recommended that these partial sequences are accurate genes for identifying and verifying the species of *Clerodendrum phlomidis*.

**KEYWORDS:** *Clerodendrum phlomidis*, ITS, *rbcL*, *rpoC1*, Molecular Identification, DNA barcoding, BLAST

1. INTRODUCTION

The Lamiaceae family, commonly known as the mint family, is a diverse and economically significant plant family with a global distribution. With over 200 genera and around 7,000 species, this family includes many well-known herbs and ornamental plants, contributing to various aspects...
of human life. Lamiaceae family is globally significant due to its economic, culinary, medicinal, and ecological roles. Its members play a vital part in human culture and are integral to various industries, making the family a subject of interest for researchers, horticulturists, and conservationists worldwide (Carović-Stanko et al., 2016). Clerodendrum (Lamiaceae) has about 580 species distributed over the world. Some species viz., C. indicum, C. phlomidis, C. serratum, and C. trichotomum used for the treatment of various diseases. Clerodendrum phlomidis (syn. Volkameria multiflorum) is an erect or scrambling shrub and is commonly found in Egypt. This taxon has significance in ecological, taxonomical, and ethnobotanical approaches, as well as its use as a remedy for various diseases such as diabetes, hypertension, headaches, wounds, snakebites and toothaches. C. phlomidis identification is needed to reduce substitution and the possibility of contamination. (Adamu et al., 2020)( Newmaster et al., 2013). Using molecular identification such as DNA barcoding for medicinal plants could be very challenging at the same time in terms of generating barcode data and analyzing these data to stand on the discrimination power (Cowan and Fay 2012). A barcode is a short fragment sequence of a gene, which has been agreed upon and used to identify species based on the references contained in the BLAST or BOLD DNA sequence database (Kress et al., 2009). The selection of an accurate barcode is very challenging to distinguish very closely related or newly developed species. Species cannot be identified using DNA barcodes if the variation within the barcode between species is low or has a history of hybridization (Kress and Erickson 2008). Development of accurate DNA barcoding sequence is necessary for authentication and identification of plant species (CBOL 2009) The chloroplast genes rbcL (Ribulose -1,5 bisphosphate carboxylase/oxygenase large subunit), rpoC1 (RNA polymerase subunit C1) and nuclear region ITS (Internal transcribed spacer) consider as a most DNA barcoding genes used in the phylogeny of plants (Hollingsworth et al., 2016). Our study aimed to evaluate rbcL, rpoC1 and ITS DNA barcode sequences to provide recommendations about the identification efficiency of Clerodendrum phlomidis.

2. MATERIAL AND METHODS

2.1. Taxon collection and authentication:

The taxon of C. phlomidis (syn. Volkameria multiflorum) was collected from the North Coast Region, Egypt. The species was identified and verified under the authority of the National Gene Bank (NGB), Agricultural Research Center (ARC), Giza, Egypt. Morphological identification was taken at the field during the collection mission, and herbarium voucher specimens at the NGB Herbarium, Taxonomy Dept., NGB. Observations followed the terminology by (El-Gazzar et al., 2019).

2.2. DNA extraction and purification:

The genomic DNA was extracted from 5g of fresh tissues using the CTAB-modified method (Elsherbey, 2016). The extracted DNA was qualified and electrophoresed (5µl of DNA sample/100ml TE buffer) using 1.2% agarose gel and visualized by UV light. The extracted DNA was measured using the UV spectrophotometer (BIO-RAD- SmartSpec Plus spectrophotometer) between a range from 260nm to 280nm (Elsherbey, 2016).

2.3. PCR and Amplification of Target Genes:

The PCR amplification used three universal genes viz., ITS, rbcL, and rpoC1 were chosen as presented in Table (1). The PCR reaction was performed with approximately 50 µl reaction mixture (1x Flexi buffer, 50ng DNA template, 2.5mM MgCl 2, 10uM dNTPs, 0.4uM of each primer, and 1U Pro mega© Green Go Taq ™ enzyme). The PCR was optimized at 94°C for 5 min for an initial denaturation cycle, followed by 40 cycles of 94°C for 30 sec for the denaturation step, 50°C for 1 min for the annealing step, and 72°C for 30 sec for the elongation step, and 72°C for 7 min for the final extensions. The PCR product was run on 1.5% agarose gel, visualized under a trans illuminator UV light, photographed and recorded via the gel documentation system.
Table 1. Primer sequences and PCR conditions based on ITS, rbcL and rpoC1 barcodes are under study.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer and Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITS</td>
<td>Forward 5'-ACGTCCCTGCCCCTTTGTACACA-3' (22 bp)</td>
</tr>
<tr>
<td></td>
<td>Reverse 5'-GCGGTACTTGTGCTATCGGT-3' (21 bp)</td>
</tr>
<tr>
<td>rbcL</td>
<td>Forward 5'ATGTCAACCACAAAACAGAGACTAAAGC-3' (26 bp)</td>
</tr>
<tr>
<td></td>
<td>Reverse 5'-GTAAAATCAAGTCCACCRCG-3' (20 bp)</td>
</tr>
<tr>
<td>rpoC1</td>
<td>Forward 5'-GGCAAAGAGGGAAGATTTGG-3' (20 bp)</td>
</tr>
<tr>
<td></td>
<td>Reverse 5'-CCATAAGCATATCTTGAGTTGG-3' (22 bp)</td>
</tr>
<tr>
<td>PCR conditions</td>
<td>F. 95°C 10 min (94°C 30s, 55°C 30s, 72°C 1min) x 35 cycles</td>
</tr>
<tr>
<td></td>
<td>R. 72°C 10 min</td>
</tr>
</tbody>
</table>

2.4 Sequencing of PCR products:

The Purified PCR products were sequenced according to the principles investigated by (Sanger et al., 1977). Sequence Analysis Software recorded and collected sequenced data. The sequence data was performed with the same amplification DNA barcode primers.

2.5 Assignment of taxa:

BLAST software was applied to all recorded sequences of Clerodendrum phlomidis using NCBI online databases. The hits with maximal percent identity scores > 90% were recorded as successful when all involved a single genus.

2.6 Data analysis:

2.6.1 Genetic parameters based on type of barcode:

Nucleotide bases and their percentages with gap % based on ITS, rbcL and rpoC1 sequences of Clerodendrum phlomidis were estimated after sequencing.

2.6.2 Pairwise Alignment:

A pairwise alignment was carried out individually based on ITS, rbcL and rpoC1 genes marker on NCBI database using BLASTN, some random sequences are selected, from BLASTN results ranging from 90 – 100% identity with the query sequence were performed to alignment on GenBank using ClustalW algorithm. The Sequence was assigned to taxon by comparing it with the sequences in the GenBank database. The E-value with lower percentage was more similar query sequence to the reference in the database. Reference sequences with their accession numbers were downloaded from GenBank (Janzen, 2019). The identified sequences have been submitted in GenBank.

2.6.3 Phylogenetic relationship examination:

Phylogenetic relationship clusters were reconstructed using NCBI with neighbor-joining software based on the aligned sequences. Based on the single locus alignment, Phylogenetic clusters were reconstructed (Tamura et al., 2018).

2.6.4 Predicate amino acid of DNA sequence:

Translation and estimation of amino acid compositions based on ITS, rbcL, and rpoC1 nucleotide sequences carried out by computer-assisted generated reading frames using online translation software (www.web.expasy.org/translate) according to (Gasteiger et al., 2005).

3 RESULTS

3.1 Morphological identification:

Clerodendrum phlomidis were collected from the North Coast Region, Egypt. The plant material was morphologically identified, authenticated and data-passported under the authority of the National Gene Bank, Egypt, based on the rules of gene banks shown in Fig (1).
3.2. Taxa Identification:

The obtained data from ITS, rbcL and rpoC1 sequences of C. phlomidis revealed a total length of 714, 1968, and 509 bp, respectively. The amount of each base present within the nucleotide sequence with their respective percentages is presented in Table (2). The query of obtained three partial sequences was blasted at online GenBank databases individually. Over hundred reference sequences showed significant alignment with percent identity between 95 – 100%. The accessions KF199889 (98.7%) and ON010669.1 (98.6%) were revealed by ITS and rbcL sequences identical to the query taxon of C. phlomidis, as shown in Table (3&4). On the other hand, the query sequence revealed similar to the taxa of C. buchneri, C. cephalanthum, C. thomsoniae, and C. rotundifoliuma by ITS sequences Table (3). C. indicum, C. thomsoniae, C. longiflorum, C. infortunatum, C. floribundum, C. cephalanthum, C. buchneri, and C. splendens by rbcL Table (4) C. japonicum, C. thomsoniae, C. cyrtophyllum and C. trichotomum by rpoC1 sequences Table (5). These taxa have a high identity to the query of C. phlomidis which range from 93% to 98% in ITS, from 98% to 99% in rbcL, and from 98.7 to 98.9% in rpoC1. In the case of the ITS sequence (714 bp), the closest species was C. phlomidis with an identity percentage of 98.74%. In the rbcL sequence (1968 bp), the closest species was C. indicum with an identity percentage of 98.89%. Finally, in the rpoC1 sequence (509 bp), the closest species was C. japonicum with an identity percentage of 98.96%.

<table>
<thead>
<tr>
<th>NO.</th>
<th>Barcode Name</th>
<th>Seq. Length</th>
<th>A</th>
<th>T</th>
<th>C</th>
<th>G</th>
<th>Gap %</th>
<th>GC</th>
<th>GC %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ITS</td>
<td>714</td>
<td>166</td>
<td>140</td>
<td>200</td>
<td>208</td>
<td>50.35%</td>
<td>408</td>
<td>57.14%</td>
</tr>
<tr>
<td>2</td>
<td>rbcL</td>
<td>1968</td>
<td>516</td>
<td>359</td>
<td>370</td>
<td>723</td>
<td>5.84%</td>
<td>1093</td>
<td>55.54%</td>
</tr>
<tr>
<td>3</td>
<td>rpoC1</td>
<td>509</td>
<td>132</td>
<td>161</td>
<td>99</td>
<td>117</td>
<td>70.56%</td>
<td>216</td>
<td>42.44%</td>
</tr>
</tbody>
</table>
Table 3. Description of the Identity percentage between C. phlomidis and the queries based on the partial ITS region against NCBI database.

<table>
<thead>
<tr>
<th>No.</th>
<th>Scientific Name</th>
<th>Length (bp)</th>
<th>Query Cover %</th>
<th>E-value</th>
<th>Identity %</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Clerodendrum phlomidis</td>
<td>655</td>
<td>88%</td>
<td>0.0</td>
<td>98.74%</td>
<td>KF199889</td>
</tr>
<tr>
<td>2</td>
<td>Clerodendrum buchneri</td>
<td>674</td>
<td>92%</td>
<td>0.0</td>
<td>95.65%</td>
<td>KT728417</td>
</tr>
<tr>
<td>3</td>
<td>Clerodendrum cephalanthum</td>
<td>686</td>
<td>90%</td>
<td>0.0</td>
<td>95.83%</td>
<td>JN575348</td>
</tr>
<tr>
<td>4</td>
<td>Clerodendrum rotundifolium</td>
<td>691</td>
<td>92%</td>
<td>0.0</td>
<td>93.24%</td>
<td>U77766</td>
</tr>
<tr>
<td>5</td>
<td>Clerodendrum thomsonae</td>
<td>697</td>
<td>92%</td>
<td>0.0</td>
<td>93.40%</td>
<td>U77745</td>
</tr>
<tr>
<td>6</td>
<td>Clerodendrum thomsonae</td>
<td>689</td>
<td>92%</td>
<td>0.0</td>
<td>93.24%</td>
<td>U77743</td>
</tr>
<tr>
<td>7</td>
<td>Clerodendrum thomsonae</td>
<td>596</td>
<td>83%</td>
<td>0.0</td>
<td>96.48%</td>
<td>AF477778</td>
</tr>
<tr>
<td>8</td>
<td>Clerodendrum thomsonae</td>
<td>492</td>
<td>55%</td>
<td>7e-172</td>
<td>94.95%</td>
<td>KC896701</td>
</tr>
<tr>
<td>9</td>
<td>Clerodendrum thomsonae</td>
<td>485</td>
<td>53%</td>
<td>3e-165</td>
<td>94.56%</td>
<td>KC896700</td>
</tr>
</tbody>
</table>

Table 4. Description of the Identity percentage between C. phlomidis and the queries based on the rbcL gene against NCBI database.

<table>
<thead>
<tr>
<th>No.</th>
<th>Scientific Name</th>
<th>Length (bp)</th>
<th>Query Cover %</th>
<th>E-value</th>
<th>Identity %</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Clerodendrum indicum</td>
<td>694 C. DNA</td>
<td>27%</td>
<td>0.0</td>
<td>98.89%</td>
<td>MK331785.1</td>
</tr>
<tr>
<td>2</td>
<td>Clerodendrum phlomidis</td>
<td>562 C. DNA</td>
<td>26%</td>
<td>0.0</td>
<td>98.62%</td>
<td>ON010669.1</td>
</tr>
<tr>
<td>3</td>
<td>Clerodendrum thomsoniae</td>
<td>706 C. DNA</td>
<td>27%</td>
<td>0.0</td>
<td>98.71%</td>
<td>MK331803.1</td>
</tr>
<tr>
<td>4</td>
<td>Clerodendrum longiflorum</td>
<td>573 C. DNA</td>
<td>27%</td>
<td>0.0</td>
<td>98.71%</td>
<td>KU564780.1</td>
</tr>
<tr>
<td>5</td>
<td>Clerodendrum infortunatum</td>
<td>536 C. DNA</td>
<td>27%</td>
<td>0.0</td>
<td>98.69%</td>
<td>JQ724863.1</td>
</tr>
<tr>
<td>6</td>
<td>Clerodendrum infortunatum</td>
<td>536 C. DNA</td>
<td>27%</td>
<td>0.0</td>
<td>98.69%</td>
<td>JQ724864.1</td>
</tr>
<tr>
<td>7</td>
<td>Clerodendrum floribundum</td>
<td>566 C. DNA</td>
<td>27%</td>
<td>0.0</td>
<td>98.87%</td>
<td>KF496552.1</td>
</tr>
<tr>
<td>8</td>
<td>Clerodendrum cephalanthum</td>
<td>552 C. DNA</td>
<td>26%</td>
<td>0.0</td>
<td>99.05%</td>
<td>KU568049.1</td>
</tr>
<tr>
<td>9</td>
<td>Clerodendrum buchneri</td>
<td>552 C. DNA</td>
<td>27%</td>
<td>0.0</td>
<td>98.89%</td>
<td>KU568050.1</td>
</tr>
<tr>
<td>10</td>
<td>Clerodendrum splendens</td>
<td>552 C. DNA</td>
<td>27%</td>
<td>0.0</td>
<td>98.89%</td>
<td>KX783849.1</td>
</tr>
</tbody>
</table>

Table 5. Description of the Identity percentage between C. phlomidis and the queries based on the rpoC1 gene against NCBI database.

<table>
<thead>
<tr>
<th>No.</th>
<th>Scientific Name</th>
<th>Length (bp)</th>
<th>Query Cover %</th>
<th>E-value</th>
<th>Identity %</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Clerodendrum japonicum</td>
<td>152,217 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.96%</td>
<td>MW181770.1</td>
</tr>
<tr>
<td>2</td>
<td>Clerodendrum japonicum</td>
<td>152,279 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.96%</td>
<td>MT473745.1</td>
</tr>
<tr>
<td>3</td>
<td>Clerodendrum japonicum</td>
<td>152,215 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.96%</td>
<td>MW307827.1</td>
</tr>
<tr>
<td>4</td>
<td>Clerodendrum japonicum</td>
<td>152,171 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.96%</td>
<td>NC_056260.1</td>
</tr>
<tr>
<td>5</td>
<td>Clerodendrum thomsoniae</td>
<td>151,053 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.75%</td>
<td>OM912812.1</td>
</tr>
<tr>
<td>6</td>
<td>Clerodendrum thomsoniae</td>
<td>151,053 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.75%</td>
<td>NC_064126.1</td>
</tr>
<tr>
<td>7</td>
<td>Clerodendrum thomsoniae</td>
<td>151,053 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.75%</td>
<td>OM912812.1</td>
</tr>
<tr>
<td>8</td>
<td>Clerodendrum cyrtophyllum</td>
<td>152,004 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.75%</td>
<td>MW858153.1</td>
</tr>
<tr>
<td>9</td>
<td>Clerodendrum cyrtophyllum</td>
<td>152,004 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.75%</td>
<td>MW858153.1</td>
</tr>
<tr>
<td>10</td>
<td>Clerodendrum trichotomum</td>
<td>151,693 C. DNA</td>
<td>94%</td>
<td>0.0</td>
<td>98.75%</td>
<td>MT473746.1</td>
</tr>
</tbody>
</table>
3.3. Phylogenetic analysis:

The molecular phylogenetic trees are represented in a linear form using neighbor-joining methods based on the BLAST-NCBI Genbank database. The phylogenetic clusters were reconstructed for each sequence based on closest species with the estimated substitution matrix shown in Figures (2,3&4). In the ITS sequences, the species divided into two main groups. The first group (left cluster) includes *C. rotundifolium* only; while, the second group (right cluster) is divided into more clusters. It is notably that the taxa of *C. buchneri* and *C. cephalanthum* fall into a cluster Fig (2). In the rbcL sequences, the species divided into two main groups. The first group (left cluster) includes *C. phlomoidis* (ON010669.1) and our taxa (*C. phlomoidis*) under study, while, the second group (right cluster) is divided into more clusters, one of these clusters contain more than 5 accessions of *Volkameria inermis*, this species is another name of *C. phlomoidis* Fig (3). In the rpoC1 sequences, the species divided into two main groups. The first group (left cluster) includes only our taxa (*C. phlomoidis*) under study, while, the second group (right cluster) is divided into more clusters, three accessions named *C. thomsoniae* fall into a cluster and four accessions of *C. japonicum* fall into another cluster Fig (4).

Figure 2. Distance tree of results between the queries and *C. phlomoidis* based on the partial ITS region against NCBI database.
Figure 3. Distance tree of results between the queries and *C. phlomidis* based on the *rbcL* gene against NCBI database.

Figure 4. Distance tree of results between the queries and *C. phlomidis* based on the *rpoC1* gene against NCBI database.
3.4. Sequence Alignment:

The alignment is done with the closest taxa derived by three genes viz., ITS, rpoc1 and rbcL sequences. The ITS sequence revealed the identical taxon to *C. phlomidis* (KF199889.1) with a similarity of 98.74%; where the difference in alignment scored 8 nucleotides with a gap of 4 nucleotides, as shown in Fig (5). In the rbcL results, the closest taxon recorded with *Clerodendrum indicum* (MK331785.1) with a similarity of 98.89%. The variance scored 5 nucleotides without gap, as presented in Fig (7). Sequence alignment (SA) is a method for detection variations and similarities between DNA nucleotide sequences. It is important analysis in bioinformatics for characterize functional or structural relation between sequences and is effective for biodiversity analysis (Zutau, 2017).

---

**Figure 5.** Alignment between the highest identities query and *C. phlomidis* based on partial ITS region against NCBI database.

**Figure 6.** Alignment between the highest identities query and *C. phlomidis* based on rbcL gene against NCBI database.
3.5. In silico translation of DNA sequence:

*In silico* translation of the ITS, rbcL, and rpoC1 barcode nucleotide sequences carried out into three reading frames (Protein sequences). Figures (8, 9 & 10) displayed Frames 1 - 3 (5'-3'). Nucleotide sequences in DNA or RNA can be divided into a collection of sequential, non-overlapping triplets called reading frames (RFs), which translated to stop codons or amino acids (Pearson et al., 1997). DNA sequences designated as open reading frames (ORFs) in the genome tend to be translated into amino acids (Yazhini, 2018). Previously, only large ORFs (> 300 nucleotides) were used to be protein-coding areas, which are regarded to be genes, but the discovery of short ORFs (less than 100 nucleotides) led to discovered many translating functional small ORFs and presence of their stable peptide chain. As a result, we checked the small ORFs contained within our sequences, given that some exons are extremely short. Consequently, Finding ORFs in higher eukaryotic genome sequences is significantly more challenging than in prokaryotic genomes, despite the exons of protein-encoding genes must be ORFs (Parker, 2001).
Table 1. Summary of primers used for amplification of examined markers

<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer 1</th>
<th>Primer 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITS</td>
<td>F: 5′-TCCATTAACCAAGTGTCATAT-3′</td>
<td>R: 5′-TCATATTGAGGCTACCTGATAC-3′</td>
</tr>
<tr>
<td>rbcL</td>
<td>F: 5′-GAGATTTTGAAATATCCAGAATG-3′</td>
<td>R: 5′-TCATATTGAGGCTACCTGATAC-3′</td>
</tr>
<tr>
<td>rpoC1</td>
<td>F: 5′-GAGATTTTGAAATATCCAGAATG-3′</td>
<td>R: 5′-TCATATTGAGGCTACCTGATAC-3′</td>
</tr>
</tbody>
</table>

Figure 9. rbcL nucleotide sequence of C. phlomidis translated to 3 reading frames using Expasy online translation software.

Figure 10. rpoC1 nucleotide sequence of C. phlomidis translated to 3 reading frames using Expasy online translation software.

4. DISCUSSION

Plant DNA barcoding works best when coding and noncoding genetic markers are combined. The plastid-conserved rbcL gene and the more variable rpoC1 gene are the most commonly investigated markers in several studies. Recently, the ITS region has also shown itself to be an effective variable marker (Kress, 2017). Therefore, the purpose of this study was to determine for the first time how efficiently these markers worked as DNA barcoding for the rare Egyptian plant Clerodendrum phlomidis. Except for the pairs used for the rbcL gene, which may require more specificity for successful amplification, the primers used in the amplification of the examined markers under study were sufficient, although it is not particularly useful in identifying plant species on their own, we have other markers to corroborate the identification of plant genus, so this is not much of a barrier in this case (Kang et al., 2017). The primers utilized for amplification in each case were also used to partially sequence the remaining markers, which included the ITS region, rbcL and rpoC1 genes. The identification of the plant species C. phlomidis has been confirmed by the results. The ITS region, rbcL and rpoC1 genes have been demonstrated in numerous studies to be highly effective genetic markers for identifying Clerodendrum species, making them promising candidates for plant DNA barcoding (Han et al., 2016)(Yu et al., 2021). Numerous Clerodendrum species were discovered to be comparable to the rbcL and rpoC1 gene sequences, which...
demonstrated efficacy in identifying the species when both pairs of primers were used, although the closest was C. indicum and C. japonicum respectively. Using phylogenetic analysis, one may determine potential evolutionary paths for the nucleotide family. By representing the sequences as outer branches of a tree, the evolutionary ties between them are illustrated. The degree to which various sequences are related is then reflected in the branching relationships on the inner portion of the tree. (Mount, 2001). If two DNA molecules from different organisms have identical sequences, it is likely that previous generations inherited these DNA sequences (Erickson et al., 2008). Egyptian Clerodendrum phlomidis was identified and documented here in this study with three DNA barcode sequence for the first time. The taxon was matched to neighboring taxa in each barcode sequence and the phylogenetic clusters were reconstructed. Additionally, by aligning between C. phlomidis and C. indicum sequences based on rbcL gene and another aligning between C. phlomidis and C. japonicum sequences based on rpoC1 gene obtained a single sequence with high identity, the identification process produced identical findings, supporting the sequence's utility as a differentiating marker of the C. phlomidis plant. Regarding the ITS region, differences in the relatedness of Clerodendrum species to C. phlomidis were found in the sequences produced using the ITS primers and their alignment. This provided yet another argument in favor of the ability of species to discriminate in this area. The two distinct sets of primers allowed for the successful sequencing of the ITS region. Typically, even with good amplification, sequencing the ITS region can provide some challenges (Yu et al., 2021)(Wang et al., 2016).

5. CONCLUSION

It can be said that the usefulness of these partial sequences viz., ITS, rbcL and rpoC1 genes are reliable markers for C. phlomidis DNA barcoding and identification, especially two genes of rbcL and ITS because of their linkage to C. phlomidoids. Even if the result of rpoC1 was not particularly encouraging, more specifically designed primers can produce better results. Researchers can use a standard DNA region as a DNA barcoding to identify unknown species. The results of this research proved that it is possible to ITS, rbcL, and rpoC1 genes in C. phlomidis successfully amplified and sequenced. Furthermore, the ITS region, rbcL and rpoC1 gene sequences have a 98–100% identity to taxa from the Lamiaceae family, according to the results of BLASTN analysis, single sequence alignment, and phylogenetic analysis.

6. REFERENCES


Zutau W (2017). Kmer-Based Sequence Representations for Fast Retrieval and Comparison. Faculty of Science and Engineering, Queensland University of Technology. 1-175.
الملخص العربي

تعريف الجيني و شجرة القرابة الوراثية لنبات Clerodendrum phlomidis في مصر باستخدام تقنية ترميز الحمض النووي (الباركود)

عبدالرحمن أحمد عطيتو ١، عبدالهادي إبرهيم حسن سيد ١، حسام فوزي الشاعر ١ و إيهاب معوض بدر مهدي ٢

قسم النبات الزراعي (الوراثة)，كلية الزراعة بالقاهرة، جامعة الأزهر
قسم الأصول الوراثية النباتية، البنك القومي للجينات، مركز البحوث الزراعية

Clerodendrum phlomidis هو أحد أجناس Lamiaceae، التي لها أهمية من الناحية التصنيفية. يعد Clerodendrum phlomidis من أنواع البرية المهمة في مصر ويلعب دورًا رئيسيًا في النهج البيئي والتنوع البيولوجي والتصنيفي. في الدراسة الحالية، أجرينا هذا البحث البحت هوية C. phlomidis التي تم جمعها من المنطقة الساحلية الشمالية في مصر باستخدام ثلاثة تسلسلات باركود (DNA) وهي جينات rbcL و ITS و rpoC1. تم إجراء هذا البحث بالشراكة بين قسم الوراثة بكلية الزراعة جامعة الأزهر وبنك الجينات القومي (NGB) بمركز البحوث الزراعية (ARC) بمصر. كشفت النتائج التي توصلنا إليها أن المدخلات KF199888 (98.7 %) و ON010669.1 (98.6 %) كانت متطابقة مع تصنيف Clerodendrum phlomidis و ITS و rbcL و ITS و C. indicum و C. thomsoniae و C. splendens و C. floribundum و C. japonicum و C. thomsoniae و C. cyrtophyllum و C. trichotomum و C. phlomidis. يمكن التوصية بأن تكون هذه التسلسلات الجزئية جيدة دقيقة لتحديد وتعريف أنواع C. phlomidis.