VOLUME 12 ISSUE 1 2026

ISSN 2454 – 3055

INTERNATIONAL JOURNAL OF ZOOLOGICAL INVESTIGATIONS

Forum for Biological and Environmental Sciences

Published by Saran Publications, India

International Journal of Zoological Investigations Vol. 12, No. 1, 21-34 (2026)

International Journal of Zoological Investigations

Contents available at Journals Home Page: www.ijzi.net

Editor-in-Chief: Prof. Ajai Kumar Srivastav

Published by: Saran Publications, Gorakhpur, India

ISSN: 2454-3055

Medicinal Plant *Crotalaria pallida*: Molecular Sequencing and Characterization of its Seed Extract

Farook M.A.[1]*, Muhammed Shariq K.[2], Zakirullah S.[1], Waheed D.[1] and Rohini J.[1]

[1]Department of Biotechnology, Islamiah College (Autonomous) (Affiliated to Thiruvalluvar University, Serkadu, Vellore 632115), Vaniyambadi 635 75, Tamil Nadu, India

[2]Department of Biochemistry, Islamiah College (Autonomous) (Affiliated to Thiruvalluvar University, Serkadu, Vellore 632115), Vaniyambadi 635 752, Tamil Nadu, India

**Corresponding Author*

Received: 14th July, 2025; **Accepted:** 22nd September, 2025; **Published online:** 4th January, 2026

https://doi.org/10.33745/ijzi.2026.v12i01.003

Abstract: Medicinal plants, long revered for their therapeutic properties, continue to play a vital role in traditional medicine, particularly in developing countries where accessibility and affordability are crucial. *Crotalaria pallida*, a species within the Fabaceae family, is traditionally used in various indigenous healing practices. This study integrates morphological, molecular, and phytochemical approaches to identify and characterize the medicinal properties of *C. pallida* seeds. DNA barcoding using the chloroplast rbcL gene was employed for molecular identification, with successful amplification and sequencing confirming the species as *Crotalaria pallida*. The sequence was submitted to GenBank (Accession No: OR387359), marking the first molecular documentation of this species from the Vellore district. A neighbor-joining phylogenetic tree further supported its taxonomic placement. Fourier Transform Infrared Spectroscopy (FTIR) analysis revealed the presence of several functional groups indicative of bioactive secondary metabolites, such as alcohols, phenols, alkanes, and amines. Additionally, Gas Chromatography-Mass Spectrometry (GC-MS) analysis of the ethanolic seed extract identified fifteen phyto-compounds, including 1,9-Cyclohexadecadiene, Trichloromethane-D, and Beta-Sitosterol, all of which possess known pharmacological potential. The combination of traditional and modern analytical methods offers a comprehensive understanding of *C. pallida*'s medicinal relevance. This multi-faceted approach highlights the plant's potential as a source of novel bioactive compounds and validates its traditional use. Furthermore, the study emphasizes the importance of DNA barcoding as a reliable tool for authenticating medicinal plants, especially in the absence of comprehensive reference databases. The integration of molecular data, chemical profiling, and traditional knowledge sets the foundation for future pharmacological and therapeutic explorations of *C. pallida* and other underutilized medicinal species.

Keywords: *Crotalaria pallida*, Medicinal plants, DNA barcoding, rbcL gene, FTIR spectroscopy, GC-MS analysis, Molecular identification, Phylogenetic tree

Citation: Farook M.A., Muhammed Shariq K., Zakirullah S., Waheed D. and Rohini J.: Medicinal plant *Crotalaria pallida*: Molecular sequencing and characterization of its seed extract. Intern. J. Zool. Invest. 12(1): 21-34, 2026.

https://doi.org/10.33745/ijzi.2026.v12i01.003

Introduction

Medicinal plants, also referred to as herbal medicines or phytotherapeutic agents, have long served as a cornerstone of traditional healthcare systems across the globe. Utilized since ancient times, these botanical resources have been employed for the prevention and treatment of various ailments, promotion of healing, and enhancement of overall well-being. Their therapeutic potential is largely attributed to the presence of bioactive phytochemicals such as alkaloids, flavonoids, terpenoids, and phenolic compounds. Owing to their accessibility, affordability, and perceived safety, herbal remedies remain particularly valuable in rural and underdeveloped regions, where modern health-care infrastructure may be limited (Rashedul *et al.*, 2010). India, in particular, stands out as a biodiversity-rich nation, home to approximately 45,000 plant species of which 5,150 are endemic contributing to its longstanding heritage of plant-based medicinal practices (India State of Forest Report, 2009). Despite the increasing availability of conventional medicine, a significant portion of the global population continues to rely on herbal treatments, often due to their alignment with cultural traditions and a holistic approach to healing. Unlike modern pharmaceuticals that typically target specific symptoms, herbal medicines are believed to restore systemic balance, making them a preferred alternative in many communities. *Crotalaria*, a diverse genus within the Fabaceae family, encompasses over 600 species of herbs, shrubs, and small trees predominantly found in tropical and subtropical regions. Known for their bright yellow, legume-like flowers and nitrogen-fixing capabilities, several species of *Crotalaria* are agriculturally important for improving soil fertility and are also used as fodder crops or green manure. Importantly, various *Crotalaria* species hold ethnomedicinal significance, having been traditionally used to treat a range of ailments. Advancements in plant molecular biology, particularly in DNA barcoding and sequencing, have facilitated the accurate identification and taxonomic classification of plant species. DNA barcoding, which involves sequencing short, standardized genetic markers, such as the chloroplast-encoded rbcL gene, has become a reliable technique for species authentication (Hebert *et al.*, 2003). Such genomic tools are crucial in differentiating morphologically similar species and in conserving genetic resources (Caicedo *et al.*, 2005). In parallel, spectroscopic and chromatographic techniques have proven invaluable in phytochemical characterization. Fourier Transform Infrared Spectroscopy (FTIR) allows for the identification of functional groups in plant metabolites through infrared absorption patterns, providing insights into both primary and secondary metabolic profiles (Surewicz *et al.*, 1993). Similarly, Gas Chromatography-Mass Spectrometry (GC-MS) is widely applied in phytochemical analysis due to its sensitivity, accuracy, and efficiency in detecting and quantifying volatile and semi-volatile compounds. This dual analytical approach enables comprehensive profiling of plant extracts and supports the identification of bioactive constituents with therapeutic potential (Uma *et al.*, 2009). The present study explores the pharmacological and biochemical properties of *Crotalaria pallida,* a traditionally used medicinal plant, through molecular identification, FTIR, and GC-MS analyses. The investigation also aims to assess its anticancer activity, contributing to the growing body of evidence supporting the therapeutic applications of underexplored plant species.

Materials and Methods

Collection of Plant Sample

As part of the study, frequent field surveys were carried out between May 2019 and August 2021. The excursion was taken regularly, ranging from once a month to thrice a month. The plant samples were gathered in the village of Bathalapalli in the Koundinya Forest, Tamil Nadu, India. As part of the investigation, plant samples representing various types were collected from the study

region. Following the collection of plant samples from the study area, identification was carried out.

Authentication of Plant Sample

The plant materials were confirmed by taxonomists, Dr. S. Soosairaj (Department of Botany, St. Joseph's College,Tiruchirapalli, Tamil Nadu) and Dr. G. Gnanasekaran (Department of Botany, Madras Christian College, Chennai, Tamil Nadu). The sample was identified and authenticated as *Crotalaria pallida* Aiton (Fabaceae) and given the Accession number 3053, dated 08.02.2022. The plant samples were allowed to dry in the shade, powdered and then extracted for further analysis.

Molecular Identification-DNA Sequencing

Genomic DNA isolation

DNA extraction from *C. pallida* plant samples utilized the EXpure Plant DNA isolation kit, an innovation by Bogar Bio Bee Stores Pvt. Ltd. The process began with lysis and homogenization, involving the pulverization of approximately 50 mg of the plant sample into a powdered form. This powdered sample was then mixed with 500 µl of lysis buffer in a 2 ml micro centrifuge tube. Cellular lysis was achieved through repeated pipetting, followed by the addition of 4 µl of RNAse and 500 µl of neutralization buffer. Subsequent vortexing and a 30 min incubation at 65°C in a water bath ensured efficient cell lysis while minimizing DNA shearing through careful inversion. Centrifugation at 10,000 rpm for 10 min followed, enabling the separation of a viscous supernatant from the cellular debris without disrupting the pellet. To further purify the DNA, 600 µl of Chloroform Isoamyl Alcohol was vigorously mixed by hand, and another round of centrifugation at 10,000 rpm for 10 min was carried out. The resultant aqueous phase, approximately 600 µl, was cautiously transferred to a fresh 2 ml microcentrifuge tube. 600 µl of binding buffer was added, well mixed using a pipette, and incubated at room temperature for 5 min to establish DNA binding. Then, 600 µl of this combination was added to a spin column that was positioned within a 2 ml collecting tube. The flow-through was removed after 2 min of 10,000 rpm centrifugation. The remaining 600 µl of lysate was transferred after reassembling the spin column and collecting tube, followed by another 2 min centrifugation at 10,000 rpm, and flow-through removal.

The purification process continued with two successive washes. After adding 500 µl of washing buffer I, the spin column underwent 2 min of 10,000 rpm centrifugation. The flow-through was thrown away. The spin column was then put back together, 500 µl of washing buffer II was added, and a 2 min centrifugation at 10,000 rpm was performed, followed by flow-through disposal. A final step involved a 5 min dry spin at 10,000 rpm. The spin column was then transferred to a sterile 1.5 ml microcentrifuge tube for elution. For elution, 100 µl of Elution buffer was carefully added to the center of the spin column to avoid touching the filtrate. Incubation of the tubes at room temperature for 2 min was followed by centrifugation at 10,000 rpm for an additional 2 min. The resulting buffer in the microcentrifuge tube contained the purified DNA, with DNA concentrations subsequently quantified using the Qubit 3.0 instrument.

PCR Protocol

Polymerase Chain Reaction (PCR) is a molecular biology technique that employs specialized primers to selectively amplify specific DNA sequences, whether they originate from cloned or genomic DNA sources. This amplification process relies on the activity of a unique enzyme known as DNA polymerase. PCR harnesses the catalytic power of DNA polymerase to orchestrate the synthesis of DNA strands using deoxynucleotide substrates. The template for this synthesis is a single-stranded DNA molecule. When a custom-designed oligonucleotide primer is annealed to this single-stranded DNA template (Table 1), DNA polymerase exhibits its enzymatic activity by adding individual nucleotide units to the 3' end of the custom-designed oligonucleotide primer. This addition occurs when the primer is aligned and

Table 1: Primer details

Primer Name	Sequence Details	Number of Bases
RBCLAF	5'ATGTCACCACAAACAGAGACTAAAGC3'	26
RBCLAR	5'GTAAAATCAAGTCCACCRCG3'	20

Table 2: PCR Conditions

STAGES	TEMPERATURE	TIME	No. of Cycles
Initial Denaturation	95˚C	2 min	1
Denaturation	95˚C	30 sec	
Annealing	55˚C	30 sec	30
Extension	72˚C	2 min	
Final extension	72˚C	10 min	1
Hold	4˚C	∞	∞

bound to a longer template DNA strand that shares sequence complementarity with the oligonucleotide primer. Consequently, when a synthetic oligonucleotide primer anneals to a single-stranded template DNA molecule containing a region that matches its sequence, DNA polymerase can effectively utilize the oligonucleotide primer as a starting point or primer. It then proceeds to extend the primer's 3' end, ultimately generating an elongated region of double-stranded DNA. This newly synthesized double-stranded DNA region faithfully reflects the sequence of the original template, thereby enabling precise DNA amplification and replication through repeated PCR cycles.

Primer details

Mix 5 µl of isolated DNA with 1.5 µl each of Forward Primer and Reverse Primer, along with 5 µl of deionized water, within a total volume of 25 µl. Subsequently, execute the PCR procedure using the subsequent thermal cycling parameters (Table 1).

Denaturation

The DNA template undergoes a thermal treatment at 95°C (Table 2), a temperature at which the weak hydrogen bonds that stabilize the DNA double helix are disrupted. This thermal denaturation process results in the separation of the DNA strands, effectively generating single-stranded DNA molecules.

Annealing

The mixture is subsequently cooled to a temperature typically in the range of 55°C (Table 2). This controlled cooling step facilitates the binding or annealing of primers to their complementary sequences present in the template DNA.

Extension

Following the annealing step, the reaction temperature is raised to 72°C (Table 2), which is the optimal temperature for DNA polymerase activity. DNA polymerase, functioning at this temperature, extends the primers by adding nucleotides sequentially onto the primers. This process utilizes the target DNA as a template, resulting in the synthesis of complementary DNA strands.

Purification of PCR Production

To eliminate unincorporated PCR primers and dNTPs from the PCR products, a Montage PCR Clean-up kit from Millipore was employed. Subsequently, the cleaned PCR product underwent

sequencing using specific primers. The sequencing reactions were conducted using ABI PRISM® BigDyeTM Terminator Cycle Sequencing Kits, which incorporate AmpliTaq® DNA polymerase (FS enzyme) by Applied Biosystems.

Sequencing protocol

Single-pass sequencing was performed for each DNA template, employing a set of universal 16S rRNA primers. The fluorescent-labeled fragments generated during the sequencing process were purified using an ethanol precipitation protocol to remove any unincorporated terminator molecules. Following purification, the samples were reconstituted in distilled water and then subjected to electrophoresis using an ABI 3730xl sequencer, which is a sequencing instrument manufactured by Applied Biosystems.

Bioinformatics protocol

The sequence underwent a similarity search using the NCBI BLAST tool, and subsequent phylogenetic analysis involved comparing the query sequence with closely related sequences obtained from the BLAST results. This analysis was followed by multiple sequence alignment. Multiple sequence alignments were conducted using the MUSCLE 3.7 software (Edgar, 2004). The resulting aligned sequences were refined through the application of G blocks 0.91b, which effectively removed poorly aligned positions and divergent regions, thereby eliminating alignment noise (Talavera and Castresana, 2007). Subsequently, phylogenetic analysis was performed using PhyML 3.0 with the HKY85 substitution model.PhyML, known for its efficiency, was demonstrated to be as accurate as other existing phylogenetic analysis programs when tested with simulated data, and it exhibited a one-order-of-magnitude increase in speed. To visualize the phylogenetic tree, the Tree Dyn 198.3 program was employed (Dereeper *et al.*, 2008).

Fourier Transform Infrared Spectroscopy (*FTIR*)

Fourier Transform Infrared Spectroscopy is performed for the identification and description of chemical compounds of *C. pallida* through their interaction with infrared radiation. Infrared (IR) radiation exists in the electromagnetic spectrum between visible light and microwave radiation. As IR radiation traverses a sample, it engages with the molecular vibrations of the chemical bonds within that sample. Each type of chemical bond exhibits distinctive vibrational frequencies, thereby creating characteristic absorption patterns in the infrared region. This technique enables researchers to gain valuable insights into the composition and structure of the analyzed substances.

Gas Chromatography-Mass Spectrometry (*GCMS*)

Gas Chromatography-Mass Spectrometry (GC-MS) is an analytical technique utilized for the separation, documentation, and quantification of components within a mixture of *C. pallida*. The process begins with Gas Chromatography (GC), where the mixture is separated into individual components using an inert phase and a traveling phase. The inert phase is typically a coated solid or a liquid on a solid support, while the mobile phase is a carrier gas like helium or nitrogen. Injecting the extract of *C. pallida* into the GC system allows it to traverse the column, where the individual components interact differently with the inert phase based on their chemical properties, leading to their separation. Subsequently, Mass Spectrometry (MS) comes into play as the second step. It involves the detection and identification of the separated components from the GC centered on their mass-to-charge ratio. As the components elute from the GC column, they enter the mass spectrometer and undergo ionization, forming charged particles or ions. These ions are then accelerated and sorted permitting their mass-to-charge ratio to spend electric and/or magnetic fields. The divided ions are finally detected and their abundance is recorded, generating a mass spectrum that acts as a unique fingerprint for the components present in the mixture. The analysis was performed using the Clarus 680 Gas Chromatograph (GC) equipped with a fused silica column, packed with Elite-5MS (consisting of 5% biphenyl and 95% dimethylpolysiloxane) measuring 30 meters in length, with an inner

>0123_616_005_PCR_ICBTKMS_1_RBCLA_F_D04.ab1

CTACCCGGGAATTCTCACCGGGGTTCCCCCACACCTAGGGATACTAGTAATTCCAAGAGGGAAAAAGAGGTACTCGTT
GATTTTTCCCTGAAAACGTGATTGAAGAGAGGCGGTATGTTTCCAACTGTGATATAAACAGTAGGACACTCCCTGTCT
CCCGGACCCAAGGTGTGAAAAGCTTTGTACTCAAATACCTTTTTCTCCAATAGAAAGTACTTTGGTCGAAAAAGAACC
TTTCTTTAAAAAGGGCTACGGGGATAAACTCCATACCCCTTTTTTTTTTTTCTTCTCCCAGCAACGGGCTCAACGAGAA
ACCATAGTCCTTTGAAACCATCAAACACTCGTGAGCCCCATCGGCCCCCCCGGGTGTCCATTGTACAGGAAAAAATTCG
GCAGTTTCCTTCTTCACCTGGTTTTTCAGGGGGGACCTCCGGGGGTGGGGGAGGTAACTCGAAAAGCTGGGAAAAAAA
ATCAGCCTTCTTTTGGTTTTCATACCCGGGAGAAAAAAAAGCAATTTTATAATCCTTTAACAACCGGCCTTTGAACCCC
ACCCTTGGTTTTTTTCTCTGGTTTGGGGGGGGAAAAAAAAAA

Fig. 1: rbcL gene sequences (Forward direction) in FASTA format.

diameter of 0.25 mm and a film thickness of 250 µm. Helium was used as the carrier gas, maintaining a constant flow rate of 1 ml/min during the chromatographic run. The injector temperature was set to 260°C, and for each analysis, a 1 µl sample extract was introduced into the device. The temperature profile of the oven followed these parameters: it began at 60°C for 2 min, then ramped up to 300°C at a rate of 10°C per minute. Subsequently, the temperature was held at 300°C for 6 min. The conditions for the mass detector were as follows: the transmission line temperature was set at 240°C, the ion source temperature at 240°C, and the ionization mode used was electron effect at 70 eV. The scan time was set to 0.2 sec, covering a scan interval of 0.1 sec, with fragments detected in the mass range from 40 to 600 Da. To identify the individual components, their spectra were related to the catalog of known component spectra stored in the GC-MS NIST (2008) library.

Results and Discussion

Molecular Genomic Sequence

Whole genomic DNA isolation

Genomic DNA was effectively extracted from *C. pallida* plant seeds using the EXpure Plant DNA isolation kit (a product of Bogar Bio Bee Stores Pvt. Ltd.). Quality assessment was performed through agarose gel electrophoresis, revealing a substantial yield of genomic DNA under the Gel Documentation system. This kit, known for its ability to produce large quantities of high-quality DNA, was employed. While some RNA contamination and DNA degradation were observed, the DNA's quality remained sufficient for the subsequent PCR amplification of the rbcL gene, a widely-used standard barcode for species identification.

The amplification process utilized the forward primer RBCLAF and the reverse primer RBCLAR. The resulting PCR product was analyzed on a 1.2% Agarose gel, with a 1Kb DNA ladder serving as a reference to determine the approximate size of the rbcL gene. The PCR amplification of the rbcL gene produced a fragment measuring 592 base pairs in forward shown in Figure 1 with its chromatogram documented as Figure 2 and the PCR amplification of the rbcL gene produced a 577 bp fragment in reverse shown in Figure 3 and its chromatogram was recorded and presented in Figure 4.

DNA sequencing analysis and PCR amplification of rbcL gene

The purified PCR product was subjected to Sanger sequencing, yielding a high-quality chromatogram (Figs 2, 4). Analyzing the rbcL sequence is a crucial step in uncovering potential cryptic species. To achieve this, a National Center for Biotechnology (BLAST) search was conducted to identify regions of sequence similarity between the rbcL gene sequence of *Crotalaria* species and sequences within the database.

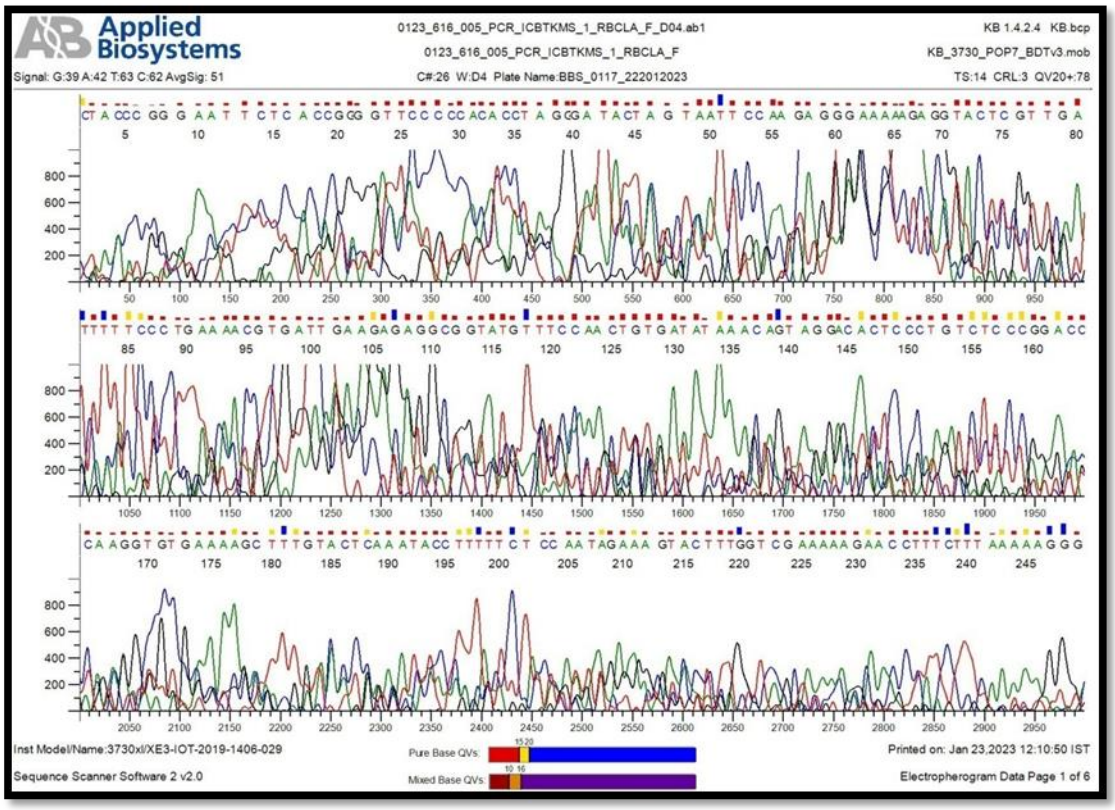

Fig. 2: Chromatogram of forward sequence of the rbcL gene of the study plant.

>0123_616_006_PCR_ICBTKMS_1_RBCLA_R_D03.ab1

ACCGGGTTCTACCCGTAATCTCTTAGCGGATAACCCCAATTCAGGTTTAATAGTACATCCCAATAGCGGACGGCCAT
ACTTGTTCATTTATCTCTTTCAACTTGGATGCCGTGAGGCGGACCTTGGAAAGTTTTAATATAAGCAGGAGGGATTC
GCAAATCTTCCACACGTAAAGCGCGCATAGCTTTGAACCCAAATACATTACCTACAATAGAAGGAAACATGTTACTA
ACAGAACCTTCTTCAAAAAGGTCTACGGGATAAGCTACATAAGCAATATATTGATTTTCTTCTCCAGCAACGGGCTC
AAGGGGATAGCATCGTCCTTTGTAACGATCAAGACTGGAAAGCCCATCGGTCCACACAGTTGTCCATGTACCAGTAG
AAAATTCGGCAGCTACCGCAGCACCTGCTTCTTCAGGCGGAACTCCGGGTTGAGGAGTTACTCGGAATGCTGCTAAA
ATATCAGTATTTTTGGTTTCATACTCAGGAGTATAATAAGTCAATTTATAATCTTTAACACCGGCTTTGAACCCAAC
ACTTGCTTTAGTCTCGGGGTGGGGGGGACAAAAAAACA

Fig. 3: rbcL gene sequences (Reverse direction) in FASTA format.

Homology search

The BLAST search found that the species (ICBTKMS1) matched *Crotalaria pallida* voucher Zhu S.S.194 (Accession No: MH050000.1) and *Crotalaria pallida* voucher 4331(JRAU) (Accession No: JQ067595.1) with Percentage identify (PI) = 97.087% and E value 0 (Fig. 5).

The obtained sequences were subsequently submitted to GenBank, resulting in the acquisition of accession numbers for the respective strains, with ICBTKMS1 assigned GenBank accession number OR387359. Remarkably, the ribulose-1,5-bisphosphate carboxylase large subunit (rbcL) gene sequence from the Vellore District species is now documented in GenBank for the first time. The outcomes of the BLAST analyses played a pivotal role in confirming the taxonomic identification of the presumed plant species in this investigation. The sequence analysis of the collected plant unequivocally identified it as

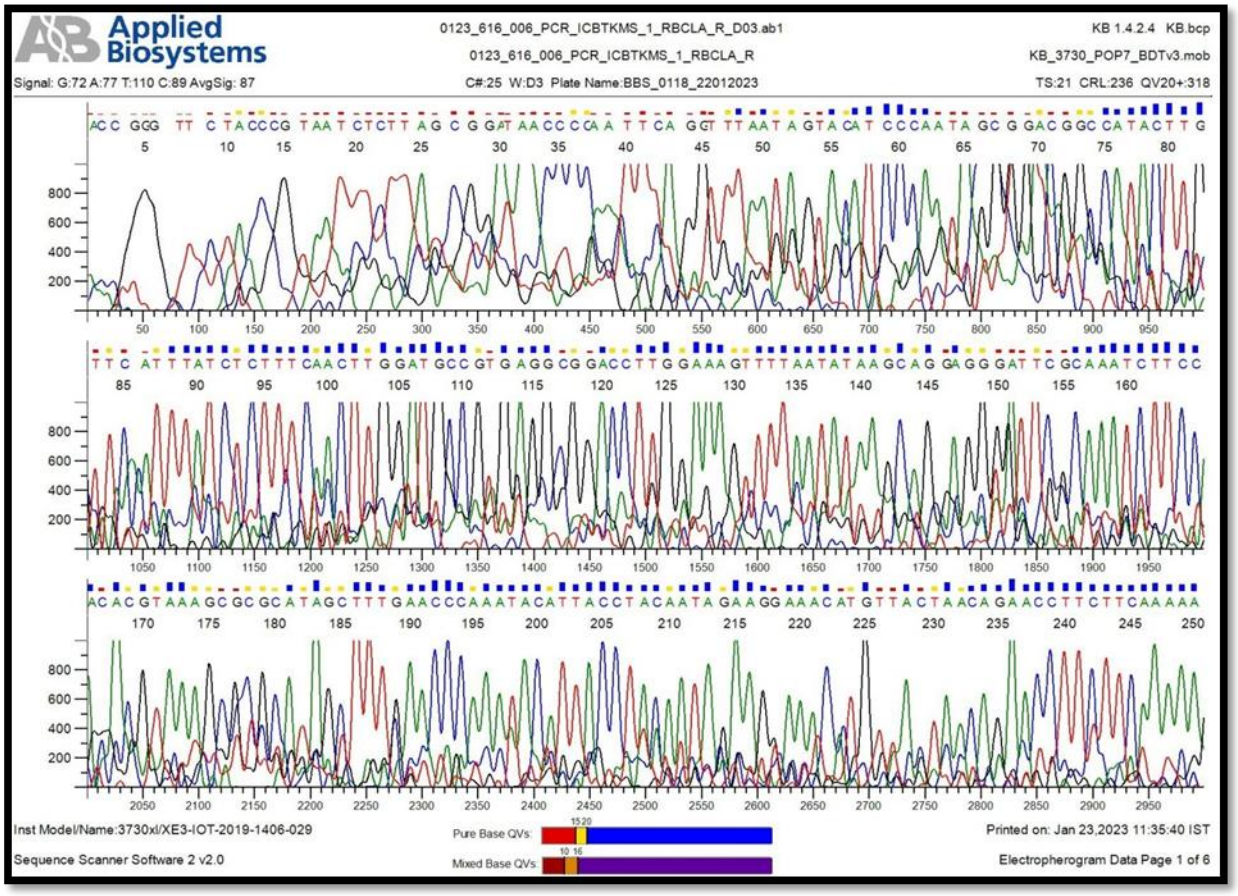

Fig. 4: Chromatogram of a reverse sequence of the rbcL gene of the study plant.

Crotalaria pallida. Further, the BLAST search highlighted that this plant exhibits the highest sequence similarity with the rbcL sequences of *Crotalaria pallida*.

Neighbor-joining phylogenetic tree

A neighbor-joining phylogenetic tree was constructed using the rbcL sequence of *C. pallida* in conjunction with other rbcL sequences available in GenBank. The tree was computed using the Kimura-2 parameter method and involved 1000 bootstrap replications to assess the robustness of the branching patterns. Bootstrap values greater than 25 are displayed at each branching point, indicating the degree of support for those specific branches. The scale bar on the tree represents the estimated genetic distance (Fig. 6). It is important to note that gaps were treated as missing data and were excluded from the dataset during the tree construction process to ensure accuracy (Tamura *et al.* 2007).

Accurate identification of medicinal plants is of growing importance for ensuring their safe application. Molecular identification, also known as DNA barcoding, is an established technique that relies on specific, standardized DNA regions for species-level identification. By comparing the distinctive barcode sequences of known species to reference sequences kept in public databases, primarily serves the goal of identifying known species. Additionally, DNA barcoding can aid in the discovery of new species. Compared to traditional methods, DNA barcoding offers several advantages, including speed, objectivity, and accuracy in species identification. Consequently, it has gained widespread recognition as a valuable tool for species identification and has proven to be effective, particularly in the context of identifying medicinal plants. In this study, both conventional and molecular methods were employed to identify

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
Crotalaria pallida voucher Zhu S.S.194 ribulose-1,5-bisphosphate carboxylase large subunit (rbcL) gene, partial cds; chloroplast	906	906	94%	0.0	97.08%	MH050000.1
Crotalaria lanceolata voucher FLAS:Whitten 3975 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	906	906	94%	0.0	97.08%	KY627420.1
Crotalaria lanceolata voucher Rockinger 20143 (M) ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	906	906	94%	0.0	97.08%	KX083396.1
Crotalaria pallida chloroplast, complete genome	906	906	94%	0.0	97.08%	NC_053562.1
Crotalaria pallida voucher 4331 (JRAU) ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; plastid	904	904	94%	0.0	97.07%	JQ067595.1
Crotalaria pallida var. obovata voucher FLAS:Whitten 3984 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast	902	902	94%	0.0	96.90%	KY627429.1

Fig. 5: Homology search result of a reverse sequence of the study plant.

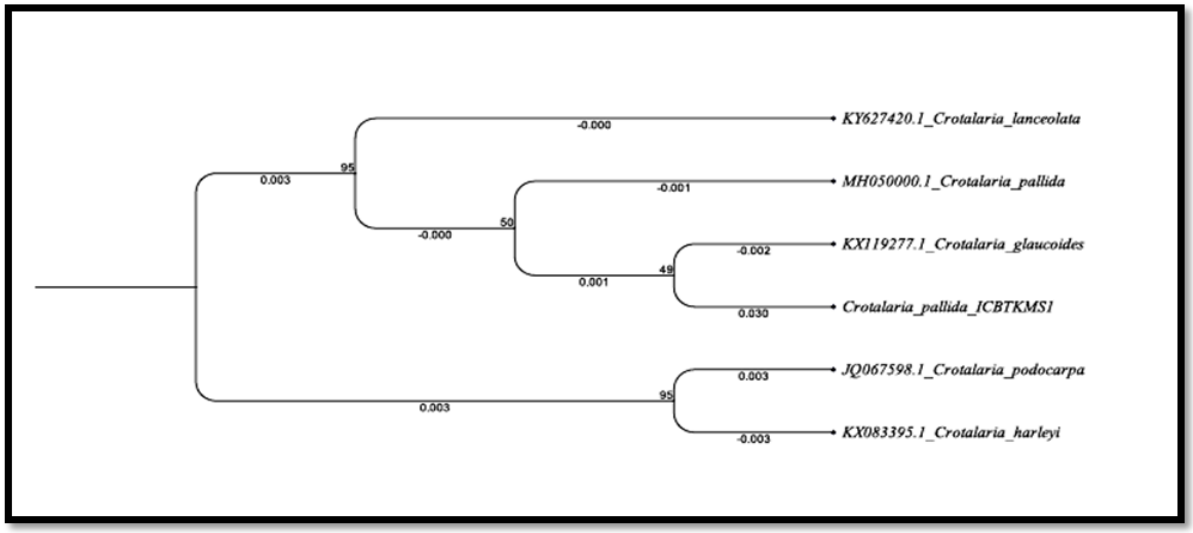

Fig. 6: Neighbor-joining tree of five rbcL sequences of the genus *Crotalaria*.

the plant species under investigation, allowing for a comprehensive and reliable assessment.

The sequences were submitted to GenBank, and an accession number was obtained for the strain ICBTKMS1, specifically GenBank accession number OR387359. Notably, the rbcL gene sequence from the Vellore species, as mentioned above, is the first-ever report of this particular species in GenBank. The outcomes of the BLAST analyses played a pivotal role in confirming the taxonomic identification of the presumed plant species in this study.The species from Vellore District was identified as *Crotalaria pallida* based on the highest sequence similarity with multiple other sequences of the same species obtained from GenBank. The sequence similarities observed ranged between 96% and 97%. To visualize the phylogenetic relationships of *Crotalaria pallida*,

a neighbor-joining phylogenetic tree was constructed. This tree helps depict the evolutionary relationships among different genetic sequences and provides insights into how closely related *Crotalaria pallida* is to other related species.

The present study underscores the utility of molecular identification through DNA bar coding as a valuable tool for the identification of medicinal plants. However, it is essential to recognize that the accuracy of species assignments using DNA barcoding is contingent upon the comprehensiveness of the sequence reference database. Currently, there is no exhaustive database available. In many instances, the identification results obtained through DNA barcoding were inconsistent when compared with putative species names derived from herbal pharmacopeia. Agreement at the species level between herbal pharmacopeia and the integrative barcoding approach was observed in 71% of cases. However, in 24% of cases, the use of herbal pharmacopeia for identification led to incorrect results. The problem of inadequate reference database coverage might be problematic when working with fully unidentified samples from a diverse range of species. In such scenarios, the most effective approach to address this problem is an integrative one. This approach involves the use of multiple sources of evidence, including multi-marker DNA barcoding, morphological characteristics, distribution data, information from pharmacopeias, and relevant literature. By combining these various sources of information, a more reliable and accurate identification can be achieved, particularly when dealing with complex or poorly documented species.

The chloroplast genomes of *Crotalaria pallida* were found to be 152,658 base pairs (bp) long, exhibiting the typical quadripartite structure characteristic of chloroplasts. This structure included a pair of inverted repeat regions (IRa and IRb) each with a length of 25,489 bp, a large single-copy region (LSC) spanning 83,652 bp, and a small single-copy region (SSC) covering 18,028 bp. The overall GC content of the genome was 36.7%, with the IR regions having a higher GC content of 42.9% compared to the LSC (34.3%) and SSC (30.6%) regions. In total, 111 genes were successfully annotated within the chloroplast genome, comprising 77 coding protein genes, 30 tRNA genes, and four rRNA genes (Nong Zhou *et al.*, 2021). In this study, the plant species were identified by integrating morphology, distribution, pharmacopeias, literature and sequencing the chloroplast gene rbcL.

Fourier Transform Infrared Spectroscopy (*FTIR*)

The infrared spectrum displays distinctive peaks corresponding to various functional groups existing in the extract of *C. pallida*. The peak observed at 3293 cm^{-1} indicates the presence of alcohols and phenols, characterized by the O-H stretching associated with hydrogen bonding. The peak detected at 2915 cm^{-1} is endorsed to alkanes, with the peak arising from the C-H stretching vibration. Furthermore, the peak at 1634 cm^{-1} signifies the presence of alkenes, characterized by the extending vibration of C=C bonds. A peak at 1531 cm^{-1} suggests the occurrence of cyclic alkenes, reflecting C=C stretching vibrations.

The occurrence of a peak at 1388 cm^{-1} corresponds to Alkane C-H bending, demonstrating the bending vibration of alkane carbon-hydrogen bonds. The peak located at 1228 cm^{-1} suggests the occurrence of amines, as it corresponds to C-H bending in amine functional groups. Similarly, another peak at 1028 cm^{-1} is associated with amines, reflecting C-H bending vibrations in this functional group. Finally, the presence of an alkyl halide is confirmed by a peak at 511 cm^{-1}, corresponding to the extending vibration of the C-Br bond. Overall, these spectral features provide valuable insights into the chemical composition and functional groups present in the *C. pallida* ethanolic seed sample depicted in Table 3 and Figure 7.

The Fourier Transform Infrared (FT-IR) spectrum serves as a pivotal analytical tool for confirming the presence of specific functional

Table 3: FT-IR Analysis of *Crotalaria pallida* seed extract

S. No.	Peak at absorbance cm^{-1}	Corresponding chemical group	Structure
1	3293	Alcohol	O-H(stretch, H-bonded)
2	2915	alkanes	(C–H stretch)
3	1634	alkenes	(C=C stretch)
4	1531	cyclic alkene	(C=C stretching)
5	1388	alkane	(C-H bending)
6	1228	amines	(C–H bend)
7	1028	amines	(C–H bend)
8	511	Alkyl Halide	(C-Br stretch)

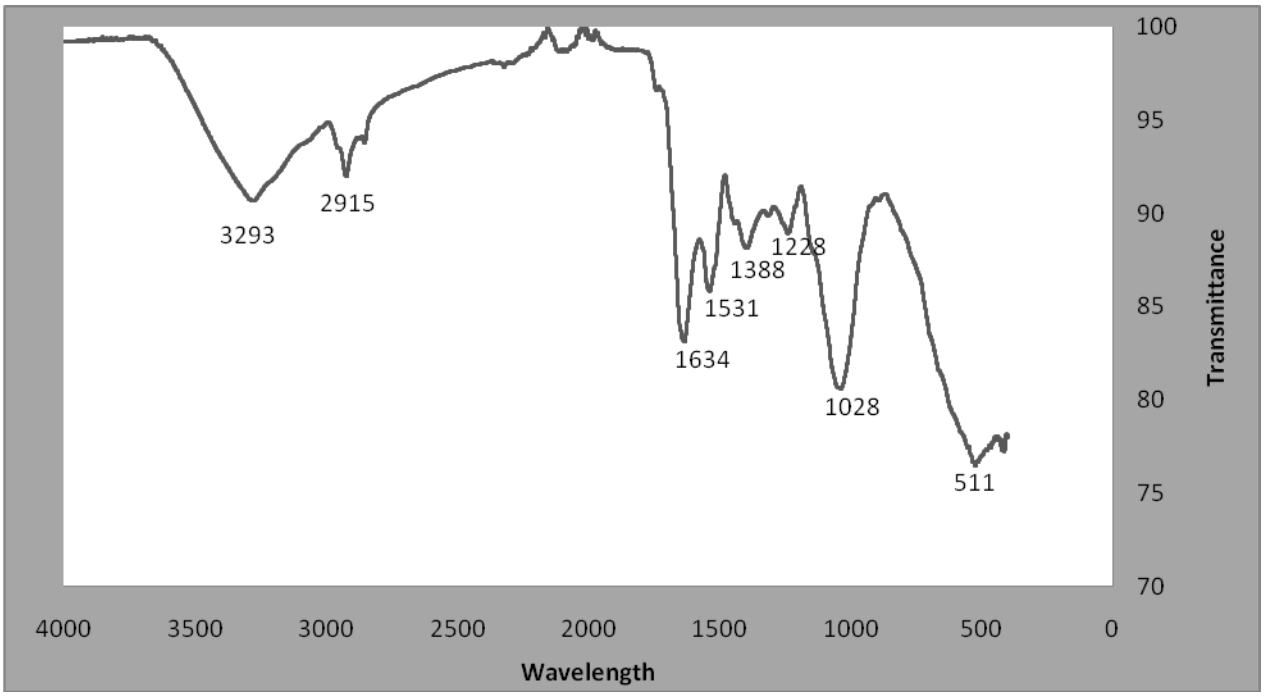

Fig. 7: FT-IR Analysis of *Crotalaria pallida* Seed extract.

constituents within provided samples and extracts. This technique enables the differentiation of authentic medicinal materials from adulterants and facilitates the assessment of the quality attributes of medicinal substances. In numerous studies, researchers have effectively employed FT-IR spectroscopy to discern between closely related botanical species and other organic entities. This application highlights the versatility and reliability of FT-IR spectroscopy as an essential resource for both qualitative and comparative analysis within the realm of natural products and materials authentication. Utilizing the FT-IR spectrum, the identification of active component functional groups was accomplished by analyzing peak values within the infrared radiation range. A diverse array of functional groups was discerned, including hydroxyl, alkyl, carboxylic, aromatic alcohols, esters, carboxylic acids and anhydrides, aromatic ethers, tertiary amines, aliphatic fluoro,

Table 4: GC-MS spectral analysis of ethanolic extract of *C. pallida* seed

S. No.	RetentionTime	Molecule Name	Formula	Molecular Weight	Area %
1.	1.259	3-Ethyl-1H-pyrrole	C_6H_9N	95	17.351
2.	1.394	1,1-Dipropylhydrazine	$C_6H_{16}N_2$	116	5.204
3.	1.469	Trichloro methane-D,	CCl_3D	119	27.660
4.	16.940	Hexahydro-1,3-benzodioxol-2-one	$C_7H_{10}O_3$	142	0.926
5.	17.000	4-Hydroxybutyl acrylate	$C_7H_{12}O_3$	269	1.700
6.	17.150	Methyl 2,4-dimethylpentanoate	$C_8H_{16}O_2$	144	0.873
7.	18.605	1,9-Cyclohexadecadiene	$C_{16}H_{28}$	220	32.113
8.	18.941	1,3(Z),13-Tetradecatriene	$C_{14}H_{24}$	192	4.050
9.	19.176	6-Heptenyl acetate	$C_9H_{16}O_2$	156	0.905
10.	20.056	1H-Pyrrole-2,5-dione	$C_4H_3NO_2$	198	1.198
11.	21.172	2-Methylheptanal	$C_8H_{16}O$	128	1.807
12.	27.114	26-Hydroxycholesterol	$C_{27}H_{46}O_2$	402	1.105
13.	27.359	10-Dodecenol	$C_{12}H_{24}O$	276	1.239
14.	27.999	Beta-Sitosterol	$C_{29}H_{50}O$	414	2.441
15.	28.489	Indolizine, 1-acetyl-2-methyl-6-nitro-	$C_{11}H_{10}N_2O_3$	305	1.427

and aliphatic compounds (Paulraj *et al.*, 2011). Notably, these compounds are classified as secondary plant metabolites, a categorization supported by the elucidations of researchers. This comprehensive analysis underscores the invaluable role of FT-IR spectroscopy in characterizing the chemical composition of the aerial parts *C. longipes* and highlighting their secondary metabolite diversity (Maobe and Nyarango, 2013).

Gas Chromatography-Mass Spectrometry (*GCMS*)

Medicinal flora serves as a reservoir for novel pharmaceutical agents, with many contemporary therapeutic agents being derived indirectly from these botanical sources. GC–MS chromatogram of the ethanolic *C. pallida* seed extract indicated fifteen peaks indicating the occurrence of 15 chemical components. On assessment of the mass spectra of the compounds with the NIST library, the fifteen compounds were characterized and recognized as 1-1,9-Cyclohexadecadiene(32.29%), Trichloro methane-D(27.66%) and 3-Ethyl-1H-pyrrole (17.35%), 1,1-Dipropylhydrazine (5.204%), 1,3(Z),13-Tetradecatriene (4.050), Beta-Sitosterol (2.441%), 2-Methylheptanal (1.807%), 4-Hydroxybutyl acrylate (1.700%), Indolizine, 1-acetyl-2-methyl-6-nitro- (1.427%), 10-Dodecenol (1.239%), 1H-Pyrrole-2,5-dione (1.198%), 26-Hydroxycholesterol (1.105%), Hexahydro-1,3-benzodioxol-2-one (0.926%), 6-Heptenyl acetate (0.905%) and Methyl 2,4-dimethylpentanoate (0.873%). Consequently, this ongoing investigation aims to ascertain the bioactive compounds inherent within the ethanolic extract of *Crotalaria pallida* seeds, leveraging Gas Chromatography and Mass Spectroscopy techniques. Table 4 delineates the active constituents, together with their respective retention times (RT), molecular weights (MW), molecular formulae and concentration (peak area %). The dataset unveils the occurrence of 15 bioactive compounds within the ethanolic extract of *C. pallida* seeds. The mass spectra of these recognized compounds from *C. pallida* seeds are visually depicted in Figure 8.

GC-MS investigation of an ethanolic seed extract derived from *Crotalaria pallida* with the documentation of a diverse range of bioactive composites validates the plant's historical practice by traditional therapists for a variety of medicinal purposes for which it has been used for centuries.

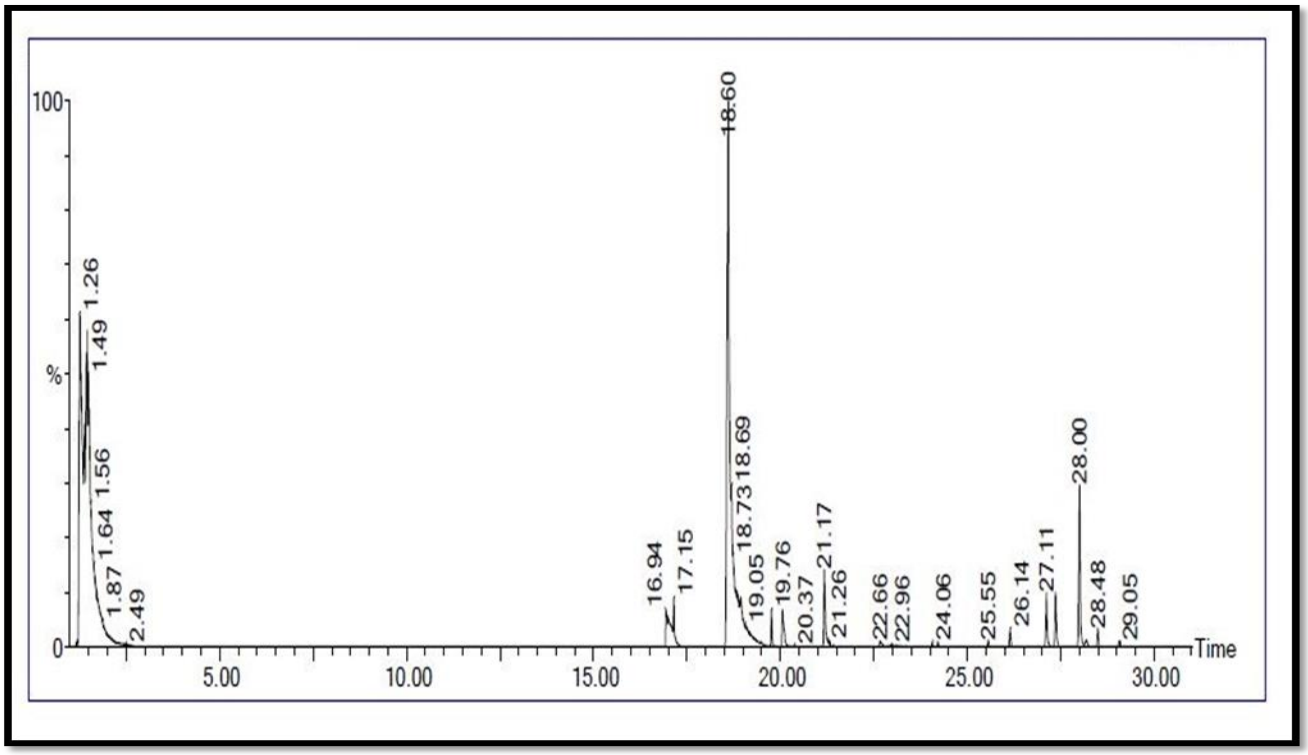

Fig. 8: GC-MS Peak level of the ethanol extract of *C. pallida* seeds chromatogram graph.

While there are still plenty of challenges to overcome, the segregation of distinct phyto-chemical components and subsequent testing of their biological activity do hold the promise of yielding substantial results. As a result of this research endeavor, a novel realm of inquiry will be initiated that focuses on the individual components and their inherent pharmacological efficacy within the solution. The GC-MS spectral analysis definitively validated the existence of fifteen prominent constituents within the sample, each characterized by specific retention times: 1.259, 1.394, 1.469, 16.940, 17.000, 17.150, 18.605, 18.941, 19.176, 20.056, 21.172, 27.114, 27.359, 27.999, and 28.489 correspondingly. The findings from the GC-MS investigation of the ethyl acetate extract of the entire *C. pallida* plants are presented, along with the potential medicinal properties associated with each identified compound. The identification of these metabolites was carried out by relating their retention times and disintegration patterns with mass spectra stored in the NIST spectral library within the GC-MS software (version 1.10 beta, Shimadzu). Additionally, the potential pharmaceutical roles of each of these bioactive molecules were cross-referenced with information from Dr. Duke's Phytochemical and Ethnobotanical Database, hosted by the National Agriculture Library, USA (Hassan Mohammad *et al.*, 2021).

Conclusion

This comprehensive study highlights the successful genomic identification and phytochemical profiling of *Crotalaria pallida*, reinforcing its potential as a valuable medicinal plant. Through effective genomic DNA extraction, PCR amplification, and rbcL gene sequencing, the species was accurately confirmed via BLAST analysis and phylogenetic tree construction. The rbcL sequence submission to GenBank as OR387359 marks a significant contribution, particularly as the first documentation of this species from the Vellore District. Furthermore, the FTIR and GC-MS analyses revealed the presence of diverse bioactive compounds such as 1,9-Cyclohexadecadiene, 3-Ethyl-1H-pyrrole, and Beta-Sitosterol, all of which are known for pharmacological relevance. These findings substantiate the plant's traditional therapeutic

use, offering scientific validation through molecular and chemical evidence. Importantly, this study underscores the significance of an integrative approach combining molecular taxonomy, spectral analysis, and phytochemical evaluation for the identification and authentication of medicinal plants. Such methodologies pave the way for the discovery of novel bioactive agents, encouraging further pharmacological investigations into *C. pallida* and its individual constituents for therapeutic applications.

Ethical Statement

No animal or microbes have been used or sacrificed for this study, hence Ethical Approval not required.

Author Contributions

Each author has contributed equally to the study's planning, analysis, writing, and editing. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

The authors extend sincere thanks to the Management and the Principal, Islamiah College (Autonomous), Vaniyambadi for providing the facilities to carry out this research work.

Conflict of Interest

The authors declare no conflicts of interest.

References

Caicedo AL and Purugganan MD. (2005) Comparative plant genomics. Frontiers and prospects. Plant Physiol. 138: 545-547.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM and Gascuel O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36(Web Server issue): W465-9.

Edgar RC. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5): 1792-1797.

Hassan Mohammad M, Kanagasabai V, Nandini MS, Prabhu K, Rao MRK, Kalaivannan J and Janaki CS. (2021) The GcMs analysis of ethyl acetate extract of one herbal plant, *Crotalaria pallida.* Nat Volatiles Essent Oils 8(4): 6791-6801.

Hebert PDN, Ratnasingham S and deWaard JR. (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Lond B. 270: S96-S99.

India State of Forest Report (2009) Himachal Pradesh.

Maobe MAG and Nyarango RM. (2013) Fourier Transformer Infra-Red Spectrophotometer analysis of *Warburgia ugandensis* medicinal herb used for the treatment of diabetes, malaria and pneumonia in Kisii region, Southwest Kenya. Global J Pharmacol. 7: 61-68.

Nong Zhou, Hai-Ling L, You Z and Dong-Qin G. (2021) The complete chloroplast genome sequences and phylogenetic analysis of *Crotalaria pallida* (Leguminosae). Mitochondrial DNA B Resour. 6(3): 1231-1232.

Paulraj K, Irudayaraj V, Johnson M and Patric Raja D. (2011) Phtochemical and anti-bacterial activity of epidermal glands extract of *Christella parasitica* (L.) H. Lev. Asian Pac J Trop Biomed. 1: 8-11.

Rashedul IM, Ahamed R, Rahman MO, Akbar MA, Alamin M and Alam KD. (2010) In vitro antimicrobial activities of four medicinally important plants in Bangladesh. Eur J Sci Res. 39: 199-206.

Surewicz WK, Mantsch HH and Chapman D. (1993) Determination of protein secondary structure by fourier transform infrared spectroscopy: a critical assessment. Biochem. 32: 389-393.

Tamura K, Dudley J, Nei M and Kumar S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24(8):1596-1599.

Talavera G and Castresana J. (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biol. 56: 564-577.

Uma B, Prabhakar K, Rajendran S and Sarayu LY. (2009) Studies on GC/MS spectroscopic analysis of some bioactive antimicrobial compounds from *Cinnamomum zeylanicum*. J Med Plants 8(31): 125-131.