PLANT SCIENCE TODAY
Vol 13(1): 1-10
https://doi.org/10.14719/pst.8889

eISSN 2348-1900

RESEARCH ARTICLE

Molecular identification of ethnomedicinal *Calotropis* using DNA barcoding

Ayan Kumar Naskar[1], Sk Md Abu Imam Saadi[2], Mehebub Sarwar Hossain[2] & Amal Kumar Mondal[3*]

[1]Wildlife Institute of India, Chandrabani 248 001, Uttarakhand, India

[2]Molecular Plant Taxonomy Laboratory, Department of Biological Sciences, Aliah University, Kolkata 700 160, West Bengal, India

[3]Plant Taxonomy, Biosystematics, and Molecular Taxonomy Laboratory, UGC-DRS-SAP Supported Department, Department of Botany & Forestry, Vidyasagar University, Midnapore 721 102, West Bengal, India

*Correspondence email - akmondal@mail.vidyasagar.ac.in

Received: 15 April 2025; Accepted: 14 December 2025; Available online: Version 1.0: 11 March 2026; Version 2.0: 19 March 2026

Cite this article: Ayan KN, Sk Md AIS, Mehebub SH, Amal KM. Molecular identification of ethnomedicinal *Calotropis* using DNA barcoding. Plant Science Today. 2026; 13(1): 1-10. https://doi.org/10.14719/pst.8889

Abstract

The presence of high demand for the ethno-medicinally important plant *Calotropis* R.Br. (Apocynaceae Juss.) in the trade market makes substitution or adulteration likely. To address this problem of accurate species identification within this genus, DNA barcoding methods are employed. We investigated the species discriminating power of the recommended barcode loci (rbcL, matK, trnL-F and ITS) and their combinations using distance-based (inter- and intra-specific distances), similarity-based (BM and BCM) and phylogeny-based analyses of available *Calotropis* species. In the present study, the BLAST identity rate is high for the recommended barcode region rbcL (99.66–100 %), followed by matK (99.57–100 %), trnL-F (98.12–100 %) and ITS (97.50–100 %). A notable difference was found between inter- and intraspecific distances in all the selected genes except rbcL. The BM and BCM approaches revealed the highest rate of correct identification with ITS (60 %) as a single gene and the combination with ITS (50 %) as a double gene. It is further confirmed that only the ITS single gene successfully separated the *Calotropis* species in phylogenetic analysis, whereas the other single locus and double locus showed some ambiguity in discriminating the species properly. Therefore, we suggest that the nrITS gene is the most suitable barcode for differentiating *Calotropis* species.

Keywords: Apocynaceae; *Calotropis* sp.; DNA barcoding; ethnomedicinal; molecular authentication

Introduction

India has one of the richest diversities of medicinal herbs and has used these herbs since ancient times (1) for healthcare. Approximately 2400 medicinal plants and 6000 higher plants are utilized in different folk healthcare systems, as listed in the Indian System of Medicine. About 9500 registered and many unregistered cottage-level industries supply medicinal herbs for the preparation of raw herbal products (2). Rarely do 10 % of the supplied medicinal plants come from cultivated sources, while the remaining are sourced from the wild (2, 3). The herbal drugs have high demand in the international trade market due to the inexpensive, more effective, readily accessibility and harmless. However, the herbal sector confronts the substitution and adulteration of medicinal plants with closely related species (4, 5). The reasons for this may include misidentification, scarcity of the correct herb, etc. (6, 7). Impurities in herbal remedies or incorrect plant identification can have detrimental effects on one's health and, in cases of negligence, can be fatal (8). To prevent adulteration, people must be educated on the right way of identification of these herbal raw materials.

The genus *Calotropis* R.Br. popularly referred to as "Akanda" consists of terrestrial, woody shrub or tree falling within the sub-family Asclepioideae under the milkweed family Apocynaceae Juss. Robert Brown first described it on 3rd April 1810. Its native range is tropical and sub-tropical Old World. *Calotropis* consists of three species viz. *Calotropis gigantea* (L.) W.T.Aiton, *Calotropis procera* (Aiton) W.T.Aiton and *Calotropis acia* Buch.-Ham. Hamilton, Francis Buchanan first described the *C. acia* in "Transactions of the Linnean Society of London (1824)". After that, no information is available about that species. Leprosy, tumors, wounds, dropsy, worms, digestive disorders, asthma and skin conditions are among the ailments for which *C. procera* is employed in ethnomedicine. The root is used to treat any toxic stinging or dyspepsia (9). Whereas, for leucorrhoea, anti-inflammatory (10), worms, anti-diarrhoeal (11), *C. gigantea* is employed. Leaf blades are utilized for epilepsy, jaundice, elephantiasis and rheumatic discomfort. Latex is utilized for toothaches, bronchitis, asthma, leprosy (9, 13) and scorpion stings (12).

Even though *Calotropis* have numerous observable characteristics like root, leaves, stem, flower, fruit and inflorescence structure. Besides that, both the *Calotropis* species are very similar in appearance which is one of the reasons of adulteration or substitution. DNA barcoding has become a crucial tool for species differentiation within the genus in a recent year. The CBOL evaluated seven chloroplast genomic areas in the kingdom of plants and suggested a set of markers like matK and rbcL as plant barcodes for efficient identification. However, rbcL has a poor species resolution despite having high universality, matK provides

greater resolution at the less universality. A combination of these two markers can maximize species discrimination. However, their ability to discriminate the species is limited (14, 15). To achieve the highest rate of identification, even among closely related species, the China Plant BOL Group (16) recommended that, using the nuclear ITS with the combination of matK+rbcL as a best suitable plant barcode (17). This investigation evaluates the potentiality of the recommended barcode markers for authentication of ethnomedicinal plants with special reference to *Calotropis* sp.

Materials and Methods

Taxon sampling and ethics statement

At least three accessions of each specimen were collected from natural distributional areas of Eastern India (Table 1) which cover the sequence variability in the chosen barcode regions. This specimen vigorously growing in South-West Bengal. The voucher specimen has been deposited in the Herbarium section (Accession number: VU/AYAN/015.1-015.3 and VU/AYAN/016.1-016.3) of Vidyasagar University and Ulluberia Botanical Institute Herbarium (UBIH00000210-215). All herbarium procedures were carried out following the established guidelines (18). In accordance with the recommendations of the China Plant BOL Group, we used the standard DNA barcode markers (rbcL, matK, trnL-F and ITS) for sequencing and the resulting sequences were subsequently deposited in the NCBI GenBank database. Finally, create a brief gene library (Table 2) with the laboratory-derived sequences.

DNA isolation, amplification and sequencing

Genomic DNA was isolated at the Rajiv Gandhi Centre for Biotechnology (Thiruvananthapuram) using the Sanger sequencing platform. Silica-dried samples were utilized for DNA extraction with the NucleoSpin® Plant II Kit (Macherey-Nagel), following the manufacturer's protocol. After isolation, the DNA quality was checked by the agarose gel electrophoresis (0.8 %) and the DNA profile was visualized on a UV trans illuminator. The recommended targeted genes viz. rbcL, matK, trnL-trnF & ITS (Table 3) were chosen for this study. PCR amplification was performed in a 20 µL reaction volume which contained 1X Phire PCR buffer (containing 1.5 mM $MgCl_2$), 0.2 mM each dNTPs (dATP, dGTP, dCTP and dTTP), 1 µL DNA, 0.2 µL Phire Hotstart II DNA polymerase enzyme, 0.1 mg/mL BSA and 3 % DMSO, 0.5M Betaine, 5 pmol of forward and reverses primers. The PCR amplification was executed in a PCR thermal cycler (GeneAmp PCR System 9700, Applied Biosystems). After successful amplification, the purification was done by the treatment of ExoSAP-IT (GE Healthcare) for the removal of unwanted primers and dNTPs. DNA sequencing was done using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA) following the protocol of manufacturer. After cleaning up, the air-dried product was sequenced in a DNA sequencer (ABI 3500 DNA Analyser, Applied Biosystems) according to the protocol. The sequence data were investigated using Sequence Scanner Software v.1 (Applied Biosystems).

Table 1. Selected samples from different districts of West Bengal

Sl No.	Specimens	Collection areas		
		1st	2nd	3rd
1.	*Calotropis gigantea* (L.) W.T.Aiton	Arambag, Hooghly	Mukutmanipur, Bankura	Sarbari, Purulia
2.	*Calotropis procera* (Aiton) W.T.Aiton	Sarbari, Purulia	Panchet, Jharkhand	Pakhanna, Bankura

Table 2. List of accessions which are used in this study uploaded and downloaded from NCBI

Sl. No.	Species name	Chloroplast marker			Nuclear marker
		Codding		Non-codding	Non-codding
		matK	rbcL	trnL-F	ITS
1.	*Calotropis gigantea* (L.) W.T.Aiton	OR900760	OR795026	PP066974	OR973574
		OR900761	OR859991	PP066975	OR973575
		OR900762	OR859990	PP066976	OR973576
2.	*Calotropis procera* (Aiton) W.T.Aiton	OR921066	OR859992	PP066977	OR973784
		OR921067	OR859989	PP066978	OR973785
		OR921068	OR900759	PP066979	OR973786
3.	*Vincetoxicum atratum* (Bunge) C.Morren & Decne.	MN317494	MN163308	AJ410270	MH808514
		LC625498	LC625487	HE793772	MH808513
		LC625499	LC625488	MF400813	MG818132
4.	*Nerium oleander* L.	KU556667	MW960591	DQ221167	MH548391
		MF447465	MW960592	FJ490765	MH548390
		MF476855	MW960606	MT078031	MT106635

Table 3. Selected primer for amplification of target region of the genomic DNA

Target	Primer Name	Direction	Sequence (5′ → 3′)
RBCL	RBCL-AF	Forward	ATG TCA CCA CAA ACA GAG ACT AAA GC
	RBCL-724R	Reverse	TCG CAT GTA CCT GCA GTA GC
MATK	KIMF	Forward	CGT ACA GTA CTT TTG TGT TTA CGA G
	KIMR	Reverse	ACC CAG TCC ATC TGG AAA TCT TGG TTC
	MATK-XF	Forward	TAA TTT ACG ATC AAT TCA TTC
	MATK-NR1	Reverse	ACA AGA AAG GCG AAG TAT
TRNLF	trnLF-F-A50272	Forward	TCC TCC GCT TAT TGA TAT GC
	trnLF-E-B49873	Reverse	GGA AGT AAA AGT CGT AAC AAG G
ITS	ITS-5F	Forward	GGA AGT AAA AGT CGT AAC AAG G
	ITS-4R	Reverse	TCC TCC GCT TAT TGA TAT GC
	ITS-F2	Forward	GAT TGA ATG ATC CGG TGA AG
	ITS-R2	Reverse	CTC GCC GTT ACT AGG GGA AT

Sequence analysis

The Forward and reverse raw sequence histograms were inspected by Bioedit sequence alignment editor ver.7.2.5 (19) for trimming the poor quality of 5⍰ & 3⍰ sequence ends and the existing primer sequences. After trimming, the sequences retained at least 60 % of the original read data and were processed to meet the minimum average quality score of Q20 using Finch TV ver.1.4.0. The edited final sequences were applied for BLASTn analysis against the online nucleotide database and identify the samples. The MUSCLE algorithm in MEGA ver.11.0 (20) was used to multiple-align of each discovered barcode sequences. The fundamental sequence statistics, such as conserved sites, variable sites and parsimony informative sites, were computed using this software. Using the K2P model (21) in MEGA, the genetic distance was calculated in two ways: intra-specific variation (Average intraspecific distance, Theta (θ), Maximum intraspecific distance) and inter-specific divergence (Average interspecific distance, Theta prime (θ'), Minimum interspecific distance). To determine the proposed barcode's capacity for discrimination, the DNA barcode gap was computed using automatic barcode gap discovery (22). Here, the Best Match (BM) and Best Close Match (BCM) parameters were utilized by the TaxonDNA or SpeciesIdentifier, ver.1.8 software to evaluate the effectiveness of this possible barcode for species determination (23). The phylogenetic tree analysis was then constructed using the consensus sequences of either solo or combination barcode loci. Here, phylogeny creation in MEGA is done by the Neighbor-Joining (NJ) approach using the best-fitting substitution model utilizing the jmodel test (24) ver.2.1.10. Maximum likelihood (ML) analysis was performed in RAxML software (25). The beast tree was constructed in Beast software (26). We use the Gamma model with MCMC length of chain 100000. The thread pool size was two and the posterior probability limit was 0.5. Posterior probability 1 means 100 % accuracy.

Data availability

All the isolated DNA sequences collected from different accessions of *Calotropis* were deposited into the NCBI (https://www.ncbi.nlm.nih.gov/genbank) and are publicly accessible under the gene accession numbers listed in Table 2.

Results

The BLAST identity

Each concatenated sequence's BLASTn search was effective in identifying the sample down to the genus level. However, rbcL had the highest rate of identification at the species level, followed by matK, ITS and trnL-F (Table 4). The sequence statistics, including conserved sites, variable sites and parsimony informative sites were presented in Table 5. The length of the matrix including the aligned single locus genes ITS, rbcL, matK and trnL-F is 478 bp, 441 bp, 478 bp and 369 bp, respectively. Highest percentage of conserved sites was observed by the rbcL region (96.82 %), followed by trnL-F (71 %), matK (68.41 %) and ITS (54.39 %). Whereas ITS demonstrates the most variable site (44.35 %), then matK (31.38 %), trnL-F (24.39 %) and rbcL (2.94 %). The matK has the most parsimony informative sites, followed by ITS, trnL-F and rbcL. ITS+matK has the highest variable site and parsimony informative site, while the combined data set of trnL-F+rbcL has the highest conserved site.

Table 4. BLAST identity of the collected accessions, up to species level

Gene name	Taxa name	Sequence length	Max score	Total score	Query cover	E value	Percentage of identity	Accession length	Match sequence accession no.
matK	*Calotropis gigantea I*	554	1016	1016	100 %	0.0	99.82 %	1051	KX911177.1
	Calotropis gigantea II	747	1330	1330	100 %	0.0	98.93 %	1051	KX911177.1
	Calotropis gigantea III	682	1249	1249	100 %	0.0	99.71 %	1051	KX911177.1
	Calotropis procera I	701	1279	1279	100 %	0.0	99.57 %	1536	KT344854.1
	Calotropis procera II	641	1184	1184	100 %	0.0	100 %	823	MF694833.1
	Calotropis procera III	750	1378	1378	100 %	0.0	99.87 %	849	MK125101.1
rbcL	*Calotropis gigantea I*	548	1003	1003	100 %	0.0	99.64 %	1236	KX910823.1
	Calotropis gigantea II	576	1050	1050	100 %	0.0	99.48 %	1236	KX910823.1
	Calotropis gigantea III	441	815	815	100 %	0.0	100 %	1236	KX910823.1
	Calotropis procera I	649	1199	1199	100 %	0.0	100 %	1437	AJ419736.1
	Calotropis procera II	649	1195	1195	100 %	0.0	99.85 %	701	MZ291553.1
	Calotropis procera III	664	1221	1221	100 %	0.0	99.85 %	1437	AJ419736.1
trnL-F	*Calotropis gigantea I*	373	662	662	100 %	0.0	98.12 %	165928	NC_041431.1
	Calotropis gigantea II	870	1526	1526	100 %	0.0	99.77 %	165928	NC_041431.1
	Calotropis gigantea III	862	1552	1552	100 %	0.0	100 %	165928	NC_041431.1
	Calotropis procera I	860	1507	1507	100 %	0.0	98.14 %	875	HE805509.1
	Calotropis procera II	866	1552	1552	98 %	0.0	99.30 %	875	HE805509.1
	Calotropis procera III	869	1559	1559	98 %	0.0	99.53 %	875	HE805509.1
ITS	*Calotropis gigantea I*	767	771	771	54 %	0.0	100 %	432	OP627197.1
	Calotropis gigantea III	769	771	771	54 %	0.0	100 %	432	OP627197.1
	Calotropis procera I	519	941	941	100 %	0.0	99.23 %	670	MW412686.1
	Calotropis procera II	470	865	865	100 %	0.0	99.79 %	670	MW412686.1
	Calotropis procera III	360	610	610	100 %	6e-170	97.50 %	710	KR149556.1

Table 5. General characters of single locus and combined locus

Locus	Sequence length (bp)	Alignment length (bp)	Conserve site (%)	Variable site (%)	Parsimony informative site (%)
rbcL	441–730	441	427 (96.82)	13 (2.94)	11 (2.49)
matK	513–858	478	327 (68.41)	150 (31.38)	144 (30.12)
trnL-F	371–966	369	262 (71)	90 (24.39)	19 (5.14)
ITS	360–769	478	260 (54.39)	212 (44.35)	96 (20.08)
rbcL+matK	–	919	754 (82.04)	163 (17.73)	155 (16.86)
rbcL+trnL-F	–	810	689 (85.06)	103 (12.71)	30 (3.70)
matK+trnL-F	–	847	589 (69.53)	240 (28.33)	163 (19.24)
rbcL+ITS	–	919	687 (74.75)	225 (24.48)	107 (11.64)
matK+ITS	–	956	587 (61.40)	362 (37.86)	240 (25.10)
trnL-F+ITS	–	847	522 (61.62)	302 (35.65)	115 (13.57)
rbcL+matK+trnL-F+ITS	–	1766	1276 (72.25)	465 (26.33)	270 (15.28)

Distance-based analysis

There is diverse variance in each of the chosen markers and their combinations to allow *Calotropis* sp. to be identified. In the single gene rbcL have very low variation, no species level resolution due to completely overlap of inter- and intraspecific distance (0.0007 ± 0.0011 and 0.0015 ± 0.0017), even with SD value. The theta prime value ($\theta' = 0.125$) is also low which is confirm that the limited sequence polymorphism. The moderate variation or partial discrimination is found in matK (0.1849 ± 0.2773 and 0.0934 ± 0.2154) and trnL-F (0.1757 ± 0.2829 and 0.0976 ± 0.2279) genes although SD values are greater than the mean value, indicating wide overlap. Highest genetic variation but overlap remains in ITS gene (0.1621 ± 0.2431 and 0.1089 ± 0.1802) and the SD value also strongly support this but it has highest theta prime value (0.006) among all individual markers. The combined matK+ITS marker shows the best single two gene barcode due to smaller overlap compared to single marker and θ′ and θ values align with moderate variability. In four genes super matrix, genetic distances are lower overall due to conserved chloroplast genes dominating the signal. The most interspecific distance and the least intraspecific divergence are necessary for an effective DNA barcode marker. Therefore, all the chosen markers in the current study aside from rbcL have notable variations between intraspecific divergence and interspecific distance. There is no such distinction between intraspecific and interspecific genetic distance in the rbcL (Table 6).

DNA barcode gap analysis

A barcode gap is considered to exist when the minimum interspecific genetic distance is greater than the maximum intraspecific genetic distance. Under such conditions, species can be reliably and unambiguously identified based on their DNA sequence variation (27). However, in the present study, none of the selected barcode markers demonstrated a distinct barcode gap. Instead, the minimum interspecific divergence was found to be lower than the maximum intraspecific divergence, resulting in an overlap between the two distance distributions. Due to absence of barcode gap, ABGD analysis was performed for deeper evaluation of species delimitation. The result clearly illustrated in Fig. 1, which are indicating that the absence of a clean separation between species except ITS gene.

Despite this overlap, the analysis still revealed meaningful genetic differences. The mean interspecific distances were consistently higher than the mean intraspecific distances across all single-locus markers (except rbcL), as well as the multilocus combinations (matK+ITS and rbcL+matK+trnL-F+ITS). These differences, although insufficient to create a distinct barcode gap, still reflect notable divergence between the two species and contribute to their molecular delimitation.

Similarity-based analysis

Using TaxonDNA software, the BM and BCM characteristics were utilized to assess the proposed barcode's effectiveness for species determination. In TaxonDNA software different threshold values are used like 1 %, 3 % and 4 %. For our study, we used only 3 % threshold values. In a single locus, ITS shows good results (60 %) in Best Match (BM) and 60 % for Best Close Match (BMC) analysis. For double locus marker matK+ITS, rbcL+ITS and trnL-F+ITS show 50 % correct identification of species in BM and BMC. In quadruple, locus matK+rbcL+trnL-F+ITS shows 50 % and 33.33 % correct sequence similarity identification for BM and BMC respectively (Table 7). So, the ITS is the potential marker for discriminative the *Calotropis* sp. and then trnL-F.

Phylogeny-based analysis

Using their best-fitting substitution model, Neighbor-Joining trees (Fig. 2A–D) were built based on the single locus. All the outgroup taxa (species of *Vincetoxicum* and *Nerium*) are separated from the in -group taxa and interestingly all the *Calotropis* accessions are developing a monophyletic clade by the single locus gene ITS and trnL-F. In matK-based phylogeny, all five accessions of *Calotropis* are forming a monophyletic clade but a single accession goes to outgroup taxa. Whereas, in rbcL-based phylogeny, *Nerium* species developed outgroup clade, but *Vincetoxicum* species showed paraphyletic relation with all the *Calotropis* species. In double locus, (matK+ITS) based phylogeny shows (Fig. 3) the same result as shown by the matK gene. The maximum likelihood tree (Fig. 4) and beast tree (Fig. 5) were constructed in RAxML and BEAST ver.1.10.4 software based on the quadruple locus (matK+rbcL+trnL-F+ITS). Both the trees cannot significantly discriminate the in-group taxa from out-group taxa.

Table 6. Interspecific and intraspecific distances of single locus and combined locus

Distance parameters	Average interspecific distance	Theta prime (θ')	Minimum interspecific distance	Average intraspecific distance	Theta (θ)	Maximum intraspecific distance
rbcL	0.0007 ± 0.0011	0.125	0.0000	0.0015 ± 0.0017	0.0007	0.0022
matK	0.1849 ± 0.2773	0	0.0000	0.0934 ± 0.2154	0.1849	0.5547
trnL-F	0.1757 ± 0.2829	0	0.0000	0.0976 ± 0.2279	0.1952	0.5856
ITS	0.1621 ± 0.2431	0.006	0.0000	0.1089 ± 0.1802	0.1712	0.4863
matK+ITS	0.1556 ± 0.1816	0.1556	0.0000	0.0843 ± 0.1062	0.1457	0.2776
rbcL+matK+trnLF+ITS	0.0918 ± 0.1093	0.0919	0.0000	0.0505 ± 0.0720	0.0892	0.2916

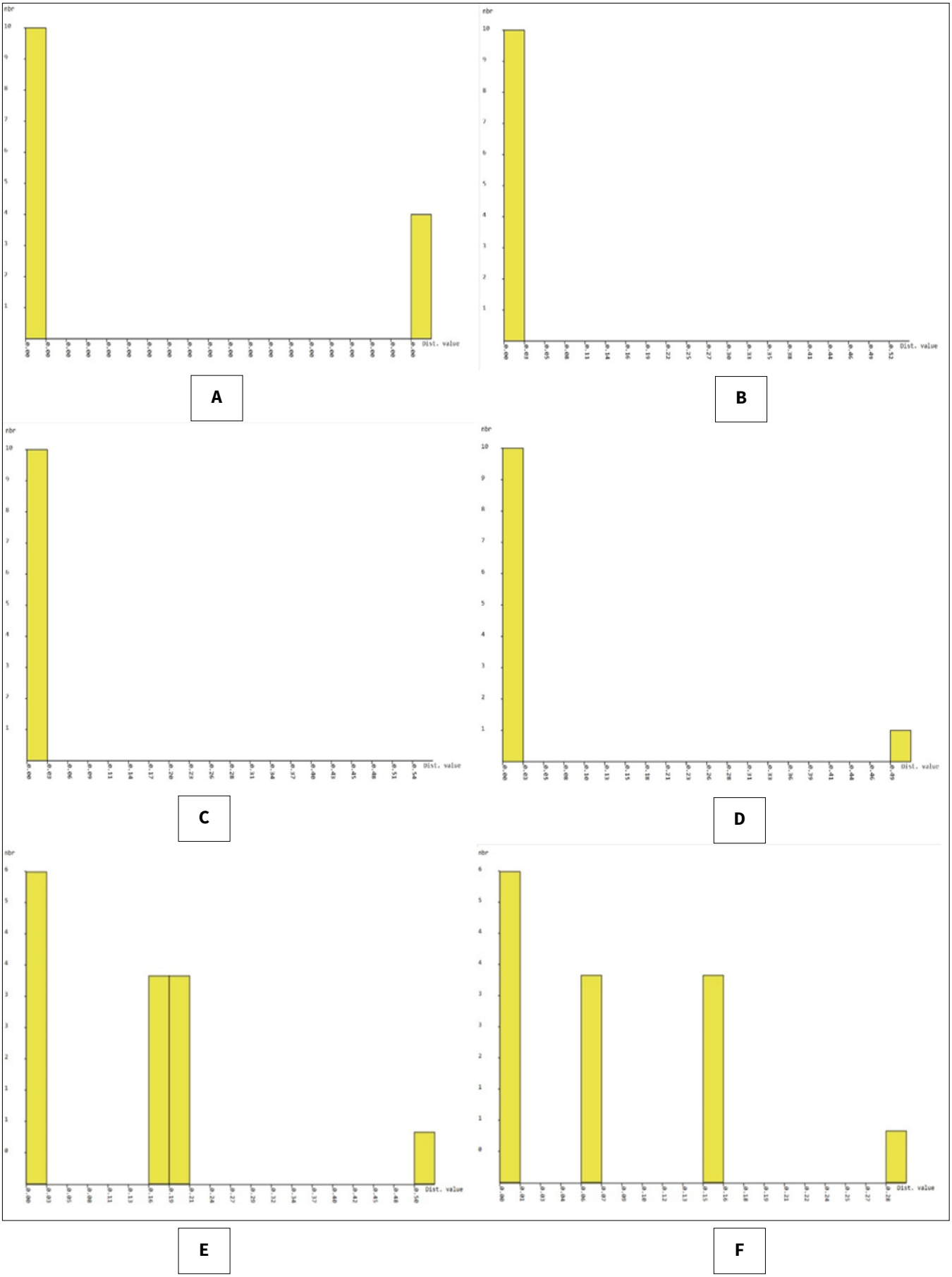

Fig. 1. ABGD barcode gap: (A) rbcL; (B) matK; (C) trnL-F; (D) ITS; (E) matK+ITS; (F) matK+rbcL+ trnL-F+ITS.

Table 7. Best match and best close match analysis in TaxonDNA

Gene loci name	Loci type	Best Match (BM)			Best Close Match (BCM)			
		Correct	Ambiguous	Incorrect	Correct	Ambiguous	Incorrect	Sequences without any match closer than 3.0 %
matK	Single locus	0(0 %)	6(100 %)	0(0 %)	0(0 %)	5(83.33 %)	0(0 %)	1(16.66 %)
rbcL		0(0 %)	6(100 %)	0(0 %)	0(0 %)	6(100 %)	0(0 %)	0(0 %)
trnL-F		1(16.66 %)	4(66.66 %)	1(16.66 %)	1(16.66 %)	4(66.66 %)	0(0 %)	1(10.66 %)
ITS		3(60 %)	0(0 %)	2(40 %)	3(60 %)	0(0 %)	2(40 %)	0(0 %)
matK+rbcL	Double locus	0(0 %)	6(100 %)	0(0 %)	0(0 %)	5(83.33 %)	0(0 %)	1(16.66 %)
matK+trnL-F		1(16.66 %)	4(66.66 %)	1(16.66 %)	1(16.66 %)	4(66.66 %)	0(0 %)	1(16.66 %)
matK+ITS		3(50 %)	1(16.66 %)	2(33.33 %)	2(33.33 %)	0(0 %)	2(33.33 %)	2(33.33 %)
rbcL+trnL-F		1(16.66 %)	4(66.66 %)	1(16.66 %)	1(16.66 %)	4(66.66 %)	0(0 %)	1(16.66 %)
rbcL+ITS		3(50 %)	0(0 %)	3(50 %)	3(50 %)	0(0 %)	2(33.33 %)	1(16.66 %)
trnL-F+ITS		3(50 %)	0(0 %)	3(50 %)	2(33.33 %)	0(0 %)	2(33.33 %)	2(33.33 %)
matK+rbcL+trnL-F+ITS	Quadruple locus	3(50 %)	0(0 %)	3(50 %)	2(33.33 %)	0(0 %)	2(33.33 %)	2(33.33 %)

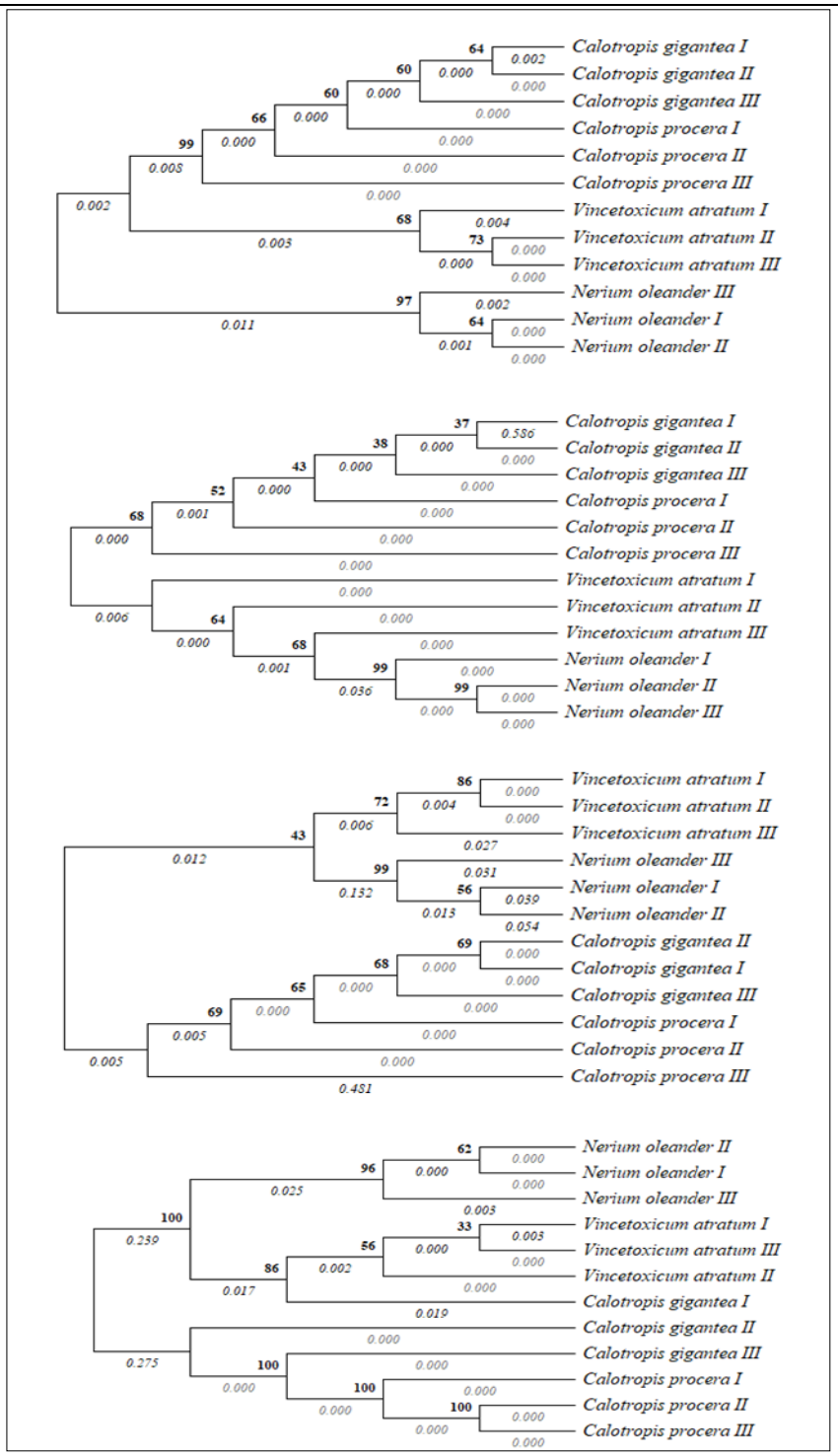

Fig. 2. Neighbor-Joining trees. NJ trees were constructed using MEGA-XI based on best-fitted substitution model: (A) rbcL; (B) trnL-F; (C) ITS; (D) matK. ≥40 % bootstrap supported value was present above the branches.

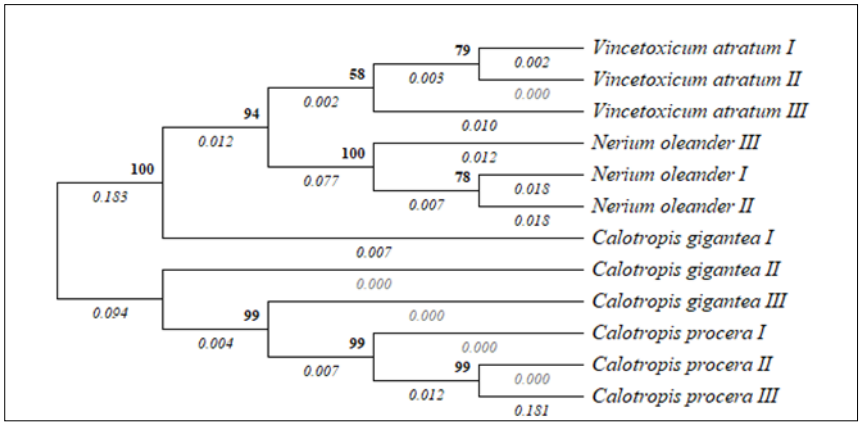

Fig. 3. Neighbor-Joining trees (NJ) of matK+ITS. ≥40 % bootstrap supported value was present above the branches.

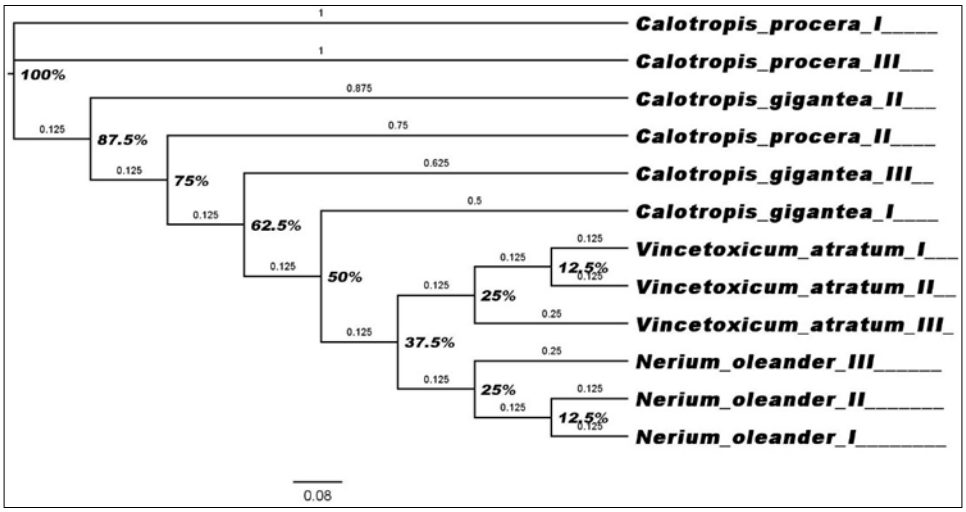

Fig. 4. Maximum likelihood (ML) tree of matK+rbcL+trnL-F+ITS in RAxML.

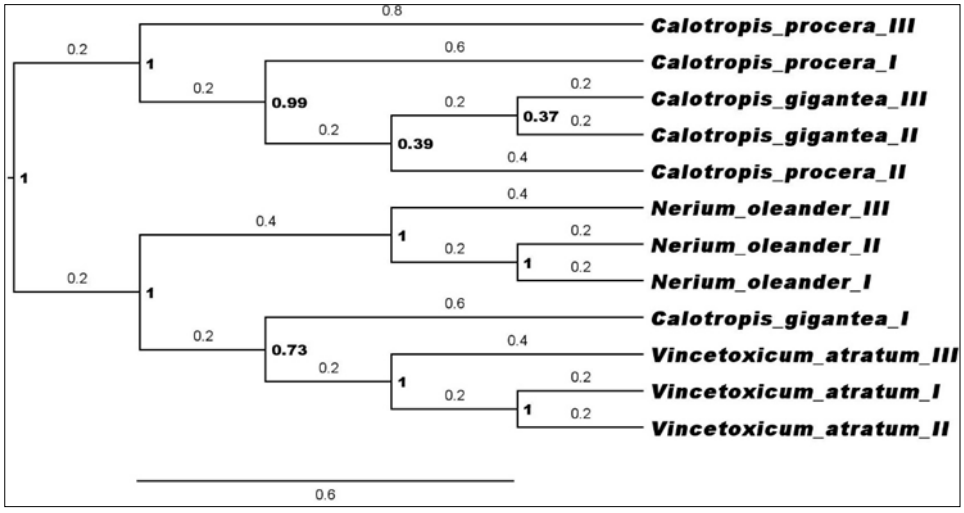

Fig. 5. Beast tree of matK+rbcL+trnL-F+ITS concatenated sequence with posterior probability value (in bold) in each node. Posterior probability 1 means 100 % accuracy so it is a good node supporting value.

Discussion

An effective barcode should be simple to amplify, sequence and resolve with excellent species discrimination and identification (28, 29). The current study uses the four widely recognized universal barcode markers, which are reasonably priced and have an appropriate length (30). All the synthetic sequences of *Calotropis* accessions and repository sequences (out-group taxa) from the gene bank were analysed here. After building a brief gene library using the chosen single locus and the combination locus, barcode analysis was carried out. For species discrimination in the Apocynaceae family, it was recommended that, the ITS2 is the core barcode and psbA-trnH is the supplementary barcode among the universal barcode locus (31). In 2009, CBOL suggested that the single gene matK and the matK+rbcL combination serve as the universal barcode for all terrestrial plants. Even though rbcL is widely used as a universal barcode identifier, its interspecific distances are less than its intraspecific divergence. But the species distinction ability of this rbcL locus is insignificant in the present investigation due to highly conserved sequences (96.82 %), showing almost no variation in allied species of *Calotropis* and it was also supported in previous study (32) on this family Apocynaceae and other angiosperm families (33, 34). The matK and trnL-F, has a moderate variation between inter and intra-specific distance value. The matK gene have the large standard deviations show that

variability is inconsistent across samples, leading to poor phylogenetic separation in gene trees. Like matK, trnL-F has no clear barcode gap and displays large amounts of background noise. ITS has the high variability and showing small barcode gap. This gene alone improves species differentiation rather than the other markers but does not fully resolve the two species.

However, the matK is the primary DNA barcode and trnL-F is a supplementary barcode marker for the Apocynaceae family described in previous research (32, 35) which is not supported by the present work. Whereas ITS2 is a suitable barcode marker for this family proposed in previous studies (31, 36) which may be partially support by the current work. Low species variation was indicated here by the lack of a barcode gap or overlap between intra-specific and inter-specific distances among the single gene (except ITS gene) and combined loci. The BM and BCM statistics parameters have been utilized for further species assignments to evaluate the success rate of species identification. These analyses revealed that none of the chloroplast loci (rbcL, matK, trnL-F) were able to correctly identify the two *Calotropis* species, with all sequences classified as ambiguous or incorrectly assigned. The nuclear ITS region showed the highest discriminatory power, correctly identifying 60 % of the sequences; however, 40 % of the identifications were incorrect, indicating substantial haplotype sharing and insufficient interspecific divergence. Two-locus and four -locus combinations did not improve assignment accuracy, as the inclusion of highly conserved chloroplast regions diluted the resolving power of ITS. Overall, these results suggest that the two *Calotropis* species exhibit low genetic differentiation across standard barcode loci, likely due to recent divergence, incomplete lineage sorting and potential chloroplast capture.

Finally, the phylogenetic tree supports the above-mentioned statistical values of the distance-based and similarity-based calculation. Only the single locus ITS has the power to discriminate the *Calotropis* species from the out-group taxa. The rbcL-based NJ phylogram shows, all the *Nerium* species (belongs to Apocynoideae subfamily) developing out-group and all the *Vincetoxicum* species (belongs to Tylophorinae subtribe) and *Calotropis* species (belongs to Asclipiadinae subtribe) developing single group as they are belonging to the same tribe. Although *Vincetoxicum* forms a separate clade from the *Calotropis* clade due to different subtribes. In trnL-F-based NJ phylogram, all the *Vincetoxicum* and *Nerium* species developed the outgroup and separated from all the *Calotropis* species which is insignificant. Whereas in a matK-based tree, all the *Calotropis* species separated well from the other taxa but a single accession comes with out-group taxa. In the matK+ITS and quadruple locus-based phylogeny the species is not separated well. Although the findings of the present study align with those reported in earlier research (39), our work expands upon their conclusions by incorporating a more rigorous and stepwise barcode evaluation framework. The earlier study relied primarily on BLAST identity scores and multiple sequence alignments of laboratory-generated sequences, which provide useful preliminary insights but are not, on their own, sufficient for a comprehensive assessment of DNA barcode performance. In contrast, the current investigation applies a detailed and methodologically robust set of analytical approaches, thereby offering a more reliable and well-supported evaluation of species discrimination within the group.

Conclusion

Based on the evaluation exercise of multiple DNA barcode markers, rbcL was found to be the most conserved region, while ITS exhibited the highest sequence variability among all the loci examined. A clear barcode gap was observed only for ITS, reflecting its comparatively higher genetic divergence. This greater variability also contributed to ITS achieving the highest correct identification rates in both the BM and BCM analyses, as well as showing a notable difference in theta prime (θ′) values in distance-based evaluations. The phylogenetic analysis further supported the effectiveness of ITS, this marker consistently separated the target group (*Calotropis*) from the selected outgroups and recovered a monophyletic clade for all *Calotropis* samples. Although ITS did not fully resolve the two species (*C. gigantea* and *C. procera*) into distinct clades, it provided moderate intraspecific-level resolution, supported by bootstrap values above 65 %. In summary, among the recommended plant barcode markers, ITS is the most suitable and informative DNA barcode for the genus *Calotropis*. However, additional experimental work beyond the standard barcode regions may be necessary to achieve complete species-level resolution. Integrating morphological traits with molecular data is strongly recommended for accurate species delimitation. This approach is especially valuable in sectors dealing with raw plant materials, such as the pharmaceutical and herbal medicine industries, where precise identification is essential. Misidentification or the use of adulterated plant material can compromise drug efficacy and, in some cases, lead to serious health risks. Therefore, correct identification and proper formulation are critical to ensuring the safety and success of herbal medicinal products.

Acknowledgements

We would like to express our thanks to Dr. Ritesh Kumar Choudhary and his scholars, Agarkar Research Institute, Pune, Maharashtra for their continuous guidance. We also acknowledge the Dr. Saptarshi Biswas, Rajiv Gandhi Centre for Biotechnology (RGCB), Thiruvananthapuram, Kerala for supporting us with DNA sequencing service. This work cannot be possible without the wholehearted help of Mr. Dheeman Mondal, research scholar, Department of Botany, Vidyasagar University. The authors are also thankful to Dr. Avishek Bhattacharjee Scientist-E & Dr. Ranjith Layola M R, Botanist of BSI, Kolkata for his support and guidance during this work.

Authors' contributions

AKN contributed to conceptualization, visualization, investigation, methodology, data curation, formal analysis and drafting of the manuscript. SMAIS and Hossain, MSH contributed to review and editing and validation. AKM contributed to supervision. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interests to declare.

Ethical issues: None

References

1. Valiathan MS. Ayurveda: Putting the house in order. Curr Sci. 2006;90:5–6.
2. Ved DK, Goraya G. Demand and supply of medicinal plants in India. New Delhi: NMPB; Bangalore: FRLHT; 2007.
3. Seethapathy GS, Ganesh D, Santhosh Kumar JU, Senthilkumar U, Newmaster SG, Ragupathy S, et al. Assessing product adulteration in natural health products for laxative yielding plants, *Cassia*, *Senna* and *Chamaecrista* in southern India using DNA barcoding. Int J Legal Med. 2015;29:693–700. https://doi.org/10.1007/s00414-014-1120-z
4. Prakash O, Jyoti Kumar A, Kumar P, Manna NK. Adulteration and substitution in Indian medicinal plants: an overview. J Med Plants Stud. 2013;1:127–32.
5. Sagar PK. Adulteration and substitution in endangered ASU medicinal plants of India: a review. Int J Med Aromat Plants. 2014;4:56–73.
6. Keshari P. Controversy, adulteration and substitution: burning problems in Ayurveda practices. In: El-Shemy HA, editor. Pharmacognosy – medicinal plants. London: IntechOpen; 2021. p. 1–12. https://doi.org/10.5772/intechopen.98220
7. Prasanth M, Anvar K. An ethnobotanical survey on adulterants of medicinal plants used by traditional practitioners of Palakkad district, Kerala, India. Int J Res Pharm Chem. 2019;9:78–84. https://doi.org/10.33289/IJRPC.9.3.2019.937
8. Ladani MR, Parabia FM. DNA barcoding: an identification of medicinal plants, databases and a promising future. Quest. 2014;2:12–7.
9. Naskar AK, Bhunia AK, Mondal AK. A survey on ethnomedicinal plants used by forest-dependent communities of the southwestern part of West Bengal, India. J Tradit Folk Pract. 2021;9:1–34. https://doi.org/10.25173/jtfp.2
10. Das S, Das S, Das M, Basu S. Evaluation of anti-inflammatory effect of *Calotropis gigantea* and *Tridax procumbens* on Wistar albino rats. J Pharm Sci Res. 2009;1:123–6.
11. Chitme H, Chandra R, Kaushik S. Studies on anti-diarrhoeal activity of *Calotropis gigantea* R Br in experimental animals. J Pharm Pharm Sci. 2004;7:70–5.
12. Rahaman C, Karmakar S. Ethnomedicine of Santal tribe living around Susunia hill of Bankura district, West Bengal, India: the quantitative approach. J Appl Pharm Sci. 2015;5:127–36. https://doi.org/10.7324/JAPS.2015.50219
13. Chaudhury S, Singh H, Rahaman CH. Ethnomedicinal uses of plants by the Lodhas tribal group of West Bengal, India. J Tradit Folk Pract. 2018;6:32. https://doi.org/10.25173/jtfp.106
14. Cameron KM, Chase MW, Whitten WM, Kores PJ, Jarrell DC, Albert VA, et al. A phylogenetic analysis of the Orchidaceae: evidence from rbcL nucleotide sequences. Am J Bot. 1999;86:208–24. https://doi.org/10.2307/2656938
15. Singh HK, Parveen I, Raghuvanshi S, Babbar SB. The loci recommended as universal barcodes for plants on the basis of floristic studies may not work with congeneric species as exemplified by DNA barcoding of *Dendrobium* species. BMC Res Notes. 2012;5:42. https://doi.org/10.1186/1756-0500-5-42
16. Li DZ, Gao LM, Li HT, Wang H, Ge XJ, Liu JQ, et al. Comparative analysis of a large dataset indicates that ITS should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci USA. 2011;108:19641–6. https://doi.org/10.1073/pnas.1104551108
17. Techen N, Parveen I, Pan Z, Khan AI. DNA barcoding of medicinal plant material for identification. Curr Opin Biotechnol. 2014;25:103–10. https://doi.org/10.1016/j.copbio.2013.09.010
18. Jai SK, Rao RR. A handbook of field and herbarium methods. New Delhi, Today and Tomorrow's printers and Publishers; 1977.
19. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 1999;41:95–8.
20. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2011;38:3022–7. https://doi.org/10.1093/molbev/msab120
21. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 1980;16:111–20. https://doi.org/10.1007/BF01731581
22. Puillandre N, Lambert A, Brouillet S, Achaz G. ABGD: automatic barcode gap discovery for primary species delimitation. Mol Ecol. 2012;21:1864–77. https://doi.org/10.1111/j.1365-294X.2011.05239.x
23. Meier R, Shiyang K, Vaidya G, Ng PKL. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst Biol. 2006;55:715–28. https://doi.org/10.1080/10635150600969864
24. Darriba D, Taboada G, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012;9:772. https://doi.org/10.1038/nmeth.2109
25. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006;22:2688–90. https://doi.org/10.1093/bioinformatics/btl446
26. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol. 2012;29:1969–73. https://doi.org/10.1093/molbev/mss075
27. Quanxi M, Xiaolu C, Li X, Yue L, Yanyan S, Yuqiao G, et al. DNA barcode for identifying Folium *Artemisiae argyi* from counterfeits. Biol Pharm Bull. 2016;39:1531–7. https://doi.org/10.1248/bpb.b16-00336
28. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. PLoS One. 2011;6:e19254. https://doi.org/10.1371/journal.pone.0019254
29. Yang JB, Wang YP, Möller M, Gao LM, Wu D. Applying plant DNA barcodes to identify species of *Parnassia* (Parnassiaceae). Mol Ecol Resour. 2012;12:267–75. https://doi.org/10.1111/j.1755-0998.2011.03095.x
30. Kress WJ. Plant DNA barcodes: applications today and in the future. J Syst Evol. 2017;55:291–307. https://doi.org/10.1111/jse.12254
31. Lv YN, Yang CY, Shi LC, Zhang ZL, Xu AS, Zhang LX, et al. Identification of medicinal plants within the Apocynaceae family using ITS2 and psbA-trnH barcodes. Chin J Nat Med. 2020;18:594–605. https://doi.org/10.1016/S1875-5364(20)30071-6
32. Cabelin VLD, Alejandro GJD. Efficiency of matK, rbcL, trnH-psbA and trnL-F to authenticate Philippine ethnomedicinal Apocynaceae through DNA barcoding. Pharmacogn Mag. 2016;12:S384–8. https://doi.org/10.4103/0973-1296.185780
33. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci U S A. 2005;102:8369–74. https://doi.org/10.1073/pnas.0503123102
34. Steven GN, Subramanyam R. Testing plant barcoding in a sister species complex of pantropical *Acacia* (Mimosoideae, Fabaceae). Mol Ecol Resour. 2009;9:172–80. https://doi.org/10.1111/j.1755-0998.2009.02642.x
35. Mahadani P, Sharma GD, Ghosh SK. Identification of ethnomedicinal plants (Rauvolfioideae: Apocynaceae) through DNA barcoding from northeast India. Pharmacogn Mag. 2013;9:255–63. https://doi.org/10.4103/0973-1296.113284
36. Selvaraj D, Sarma RK, Shanmughanandhan D, Srinivasan R, Ramalingam S. Evaluation of DNA barcode candidates for the discrimination of the large plant family Apocynaceae. Plant Syst Evol. 2015;301:1263–73. https://doi.org/10.1007/s00606-014-1149-y

37. Sidhu MC, Kumari A, Jain B, Kaur S, Kamra A, Rai J. Morphological and DNA barcoding-based identification of *Calotropis procera* and *Calotropis gigantea*. Biol Bull Russ Acad Sci. 2023;50:474–7. https://doi.org/10.1134/S1062359022602968

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.