RESEARCH ARTICLE

SABRAO Journal of Breeding and Genetics
57 (3) 1070-1080, 2025
http://doi.org/10.54910/sabrao2025.57.3.19
http://sabraojournal.org/
pISSN 1029-7073; eISSN 2224-8978

EVALUATION OF THE MAMAN PLANT FROM RIAU WITH *RBCL, ITS*, AND THEIR COMBINATION AS DNA BARCODES

D.I. ROSLIM[1*], M. FAUZAN[1], M.E. TAUFIK[1], H. AL KHAIRI[1,2], HERMAN[1], A. SAFARRIDA[3], W. LESTARI[1], N. SOFIYANTI[1], A. EFFENDI[4], L. NOVITA[5], R. RENINTA[5], and A.A.F. ALTUHAISH[6]

[1]Department of Biology, Faculty of Mathematics and Natural Sciences, University of Riau, Riau, Indonesia
[2]Department of Tropical Biology, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta, Indonesia
[3]Center for Applied Microbiology Research, National Research and Innovation Agency, Cibinong, Indonesia
[4]Department of Agrotechnology, Faculty of Agriculture, University of Riau, Riau, Indonesia
[5]Research Center for Applied Botany, BRIN, Cibinong, Indonesia
[6]Department of Biology, Faculty of Science, University of Aden, Khormaksar, Yemen
*Corresponding author's email: dewiindriyaniroslim@gmail.com
Email addresses of co-authors: hermansyahdan@ymail.com, fauzanmuhamad389@gmail.com, ezat2001taufik@gmail.com, anna003@brin.go.id, hapiz12al@gmail.com, wahyulestari1965@yahoo.com, nery_yusuf@yahoo.com, arman.effendi60@gmail.com, lind003@brin.go.id; rikania.reninta@brin.go.id, adeeltuh@gmail.com

SUMMARY

Maman is one of the food plants growing in Riau Province, Indonesia. It has the names joruk maman and African cabbage, typically eaten as a fermented vegetable. Its verification as *Cleome gynandra* L. synonym *Gynandropsis gynandra* by the *matK* and *trnL-trnL-trnF intergenic spacer* (*IGS*) used the DNA barcode sequences. The latest research aimed at molecular authentication of maman based on *rbcL* and *ITS* sequences and the evaluation of locus combinations (*rbcL+ITS* and *rbcL+ITS+matK+trnL-trnL-trnF IGS*). The DNA extraction utilized a modified CTAB method and phylogenetic analysis using MEGA11. The *rbcL* and *ITS* sequences were 839 and 672 bp in length, respectively. The results revealed the maman plant from Riau was distinct as *C. gynandra* / *G. gynandra* based on single- and multi-locus analyses. The results suggested a unique nucleotide was evident in the *rbcL* sequence in the maman from Riau. All the single and multilocus can be applicable as DNA barcodes for *C. gynandra* and *G. gynandra*.

Keywords: Maman plant from Riau, *Cleome gynandra, Gynandropsis gynandra, ITS, rbcL*, DNA barcode, phylogenetic analysis

Communicating Editor: Dr. Gwen Iris Descalsota-Empleo

Manuscript received: September 16, 2024; Accepted: November 24, 2024.

Citation: Roslim DI, Fauzan M, Taufik ME, Al Khairi H, Herman, Safarrida A, Lestari W, Sofiyanti N, a. Effendi A, Novita L, Reninta R, Altuhaish AAF (2025 Evaluation of the maman plant from Riau with *rbcL, its*, and their combination as dna barcodes. *SABRAO J. Breed. Genet.* 57(3): 1070-1080. http://doi.org/10.54910/sabrao2025.57.3.19.

Key findings: DNA barcoding is a useful technique for molecular verification in crop plants. The presented results could prove vital for researchers in recognizing maman plants (*C. gynandra* and *G. gynandra*) from Riau in the future.

INTRODUCTION

Maman plant is an annual (perennial) herb with a vertical stem that can grow to heights of 0.5 to 1.5 m, with green to purple branches (Figure 1). The leaves are long, green to old green, and the entire leaf edges are flat. The flower type was determinate inflorescence, placed on the ends of the rod (terminalis) and the leaf bone (lateralis). Each flower has four white crowns and six long filaments. It is a silicious fruit type, linear or roll-shaped, yellow when ripe, and brown when peeled.

The maman plant from Riau (*Cleome gynandra* L., synonymous with *Gynandropsis gynandra*) and African cabbage became basic components of the traditional food of the Malay Rokan community in Riau Province, Indonesia. The Joruk Maman is a fermented processed food using maman's leaves (Restusari *et al*., 2023). Maman is also a potential natural source of antioxidants, as the DPPH (IC_{50}) test revealed its leaves contain antioxidant activity (40.36 µg/ml) (Chandradevan *et al*., 2020). Previously, the maman, as verified by *matK* and *trnL-trnL-trnF IGS* using DNA barcode sequences, received the scientific name *C. gynandra* L. synonym *G. gynandra* (Roslim *et al.,* 2023).

The most promising method for achieving the objectives of quick and precise species identification and sustainable biological resource use could be the DNA barcoding (Hollingsworth *et al.,* 2016). The DNA barcodes are the standard DNA sequences of 400–800 bp that can help classify unknown organisms (Nehal *et al*., 2021). It is also a most accurate way to identify unnamed plants, as carried out by various researchers worldwide (Letsiou *et al*., 2024). The capacity to verify a species at every stage of its life history has made the DNA barcoding method more attractive. The development in computational and sequencing technology has made high-throughput barcoding applications possible, and has become a universal tool for hypothesis testing (Gostel and Kress, 2022).

Some regions of the chloroplast DNA (cpDNA) are valuable for DNA barcoding and phylogenetic analysis, including the *trnL-trnL-trnF* intergenic spacer gene (IGS), situated between the trnL (UAA) and trnF (GAA) 3'exon genes (Sevindik *et al.,* 2019). The *matK* gene encodes a component of the maturase K enzyme, implicated in RNA splicing (Barthet *et al.*, 2020). In addition to *matK* and *trnL-trnL-trnF IGS*, two other DNA barcode loci, namely, *rbcL* and *ITS*, can also be effective for DNA barcoding analysis in plants. The *rbcL* gene is easily amplified and prevalent in various plant species. The location of the locus is in the chloroplast genome (cpDNA) and a coding region (Tnah *et al*., 2019). The *Internal Transcribed Spacer* (*ITS*) is a non-coding region in the structural ribosomal RNA on the common precursor transcript in the nuclear genome. The *ITS* (*ITS1* and *ITS2*) of ribosomal DNA (rDNA) plays a significant role in species identification and phylogenetic analyses (Rathnayake *et al.,* 2023). On the other hand, the *rbcL* encodes the large subunit of the ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) enzyme (Rahman and Ali 2020).

In plants, the DNA barcoding analysis will be superior if used in more than one locus (Letsiou *et al*., 2024). However, some researchers employed a combination of DNA barcodes (multilocus) to classify the medicinal plants because single-locus DNA barcode sequences cannot always provide adequate information for low-level identification (Yu *et al*., 2021). Various combinations of DNA barcodes exist, as suggested by researchers in past studies, including *matK* + *rbcL*, *matK* + *atpF-atpH IGS* + *trnH-psbA IGS*, *matK* + *atpF-atpH IGS* + *psbK-psbI IGS*, ITS + *rbcL* + *matK* + *psbA-trnH*, ITS2 + *psbA-trnH IGS, psbA-trnH* + ITS, *psbA-trnH* + ITS + *rbcL* (Amandita *et al.,* 2019; CBOL

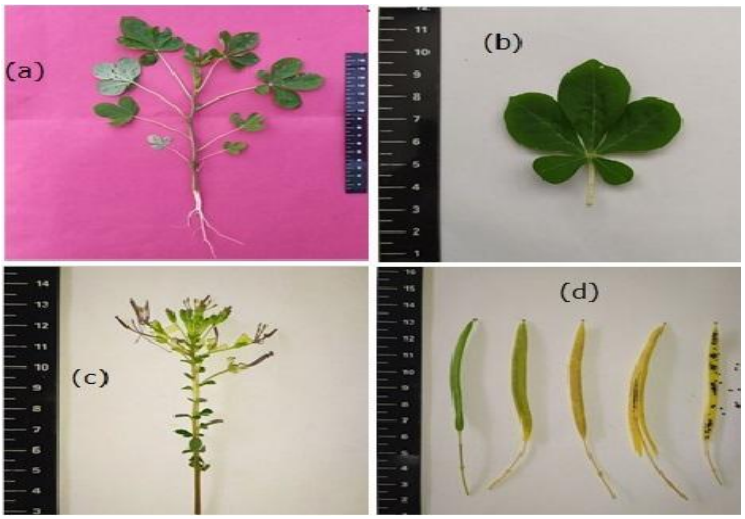

Figure 1. The morphology of maman accessions (*Cleome gynandra* L.) from Serombau Indah village, Riau Province, Indonesia. a) Plant's stature, b) leaf, c) inflorescence, and d) fruits.

Plant Working Group, 2009; Dong *et al.,* 2015; Guo *et al.,* 2018; He *et al.,* 2019). The *matK* provides greater resolution but poorer universality, while the *rbcL* delivers higher universality but lower species resolution. The greatest species discrimination came from this combination (Letsiou *et al*., 2024). Nuclear *ITS,* as recommended by the China Plant BOL Group to achieve the best identification rates even among closely related species, can also add non-coding regions to this combination (Li *et al*., 2011). These four candidate plant DNA barcoding regions (*rbcL*, *matK*, *trnL-trnL-trnF IGS*, and *ITS*) have also been applicable in several other types of plants, for example, in genera of the Gramineae family by Wang *et al.* (2022). Another example was also an undertaking by Roslim (2019) on durik-durik plants (*Syzygium* sp) using *matK, rbcL,* and *trnL-trnL-trnF IGS*.

Therefore, the presented study aimed to evaluate the *rbcL* gene and *ITS* region for the molecular identification of maman accessions and investigate the genetic relationship among the maman species from Riau with other accessions of the same genus. This study also evaluated several combinations of loci to identify the maman species from Riau, such as *rbcL+ITS* and *matK+rbcL+ITS+trnL-trnL-trnF IGS*. The DNA barcode sequence analysis hopes to enrich the molecular databases of maman accessions at GenBank.

MATERIALS AND METHODS

Material and procedure

Three specimens of maman accessions used in this study were a collection from cultivated lands of the Serombau Indah Village (latitude: 1°18′08″N, longitude: 100°40′33″E), Riau Province, Indonesia. The total gDNA, extracted from a dried leaf, used the modified CTAB method as described by Sari and Murti (2015). By adding the loading dye to the DNA mixture (1:5), the extracted gDNA's visualization employed electrophoresis (*Mupid-exU*) before storing at -20 °C. The PCR components used have the composition as follows: 1X PCR buffer (plus Mg^{2+}), 0.2 mM dNTPs, 0.8 µM forward primer, 0.8 µM reverse primer, 2 U/µl *DreamTaq* DNA polymerase (*Thermo Scientific*), 100 ng/µl the total gDNA template, and nuclease-free water. Using the primer pair of rbcLb_SF (5'-AGACCTTTTTGAAGAAGGTTCTGT-3') and rbcLb_SR (5'-TCGGTCAGAGCAGGCATATGCCA-3') served to amplify the *rbcL* gene (Dong *et al*., 2014). Meanwhile, applying the primer pairs of FP_ITS5_F (5'-

GAAAGTAAAAGTCGTAACAAGG-3') and FP_ITS4_R (5'-TCCTCCGCTTATTGATATGC-3') amplified the *ITS* region (Schoch *et al*., 2012). The conditions for PCR amplification were as follows: initial denaturation at 94 °C for 5 min; then, 35 cycles for 45 s at 94 °C (denaturation), 45 s at 56.6 °C for *rbcL* and 47 °C for *ITS* (annealing), and 1 min at 72 °C (extension). Finally, an additional 10 min ensued at 72 °C to complete the reaction.

The PCR products visualization continued on a 1.5% agarose gel using a 1 kb DNA ladder (*Thermo Scientific*) to confirm the amplicon length. The quantity measurement of PCR products, including the purity and concentration obtained, used a spectrophotometer (*NanoDrop ND-1000*) at wavelengths of 260 and 280 nm. Then, sending the single, correct, and thick bands of amplicon (40 µl), forward primer (30 µl), and reverse primer (30 µl) in separate tubes to PT Genetika Science, Indonesia (1st *BASE Malaysia*) proceeded for assessment.

Data analysis

In ensuring the generated DNA barcodes' accuracy, the alignment of DNA sequences of forward and reverse primers used the BioEdit version 7.0.0 program (Hall, 1999). Both *rbcL* and *ITS* had the BLASTn application on the https://blast.ncbi.nlm.nih.gov/website, utilized to determine the sequences under study were similar to those in the GenBank database. Additionally, the BOLD program's application (at https://v3.boldsystems.org/) was simultaneous with analogous sequence checks, particularly for *rbcL.* The top 10–12 accessions, used for the next analysis, helped determine the nucleotide variations among the studied maman accessions and construct the phylogenetic tree. The management and analysis of DNA sequences ran with MESQUITE (Maddison and Maddison, 2023). Multiple alignments engaged ClustalW (Thompson *et al*., 1994). Phylogenetic analysis ensued by using MEGA 11 to construct the Neighbor-Joining (NJ) tree (Tamura *et al*., 2021). The accuracy of the trees' assessment used techniques, such as the bootstrap re-sampling approach with 1000 repeats, to ensure the common appearance of branches most likely to reflect the genuine evolutionary relationship. Multi-locus analysis proceeded based on maman accessions in the family Cleomaceae with the whole cpDNA genome. The DNA sequences of *matK* and *trnL-trnL-trnF IGS* of maman came from previous studies with the same accession voucher (Roslim *et al*., 2023). Alignment and concatenated sequence of all markers occurred with MEGA11.

RESULTS

Total gDNA and amplicons profile

The total gDNA's successful isolation and measurement provided more than 10000 bp with thick and bright bands. These results incurred nanodrops to reveal the highest concentration values of 1004 ng/µl with a purity of 1.90 at A260/280. The DNA amplicons obtained were approximately ±800 bp for *rbcL* and ±700 bp for *ITS* (Figure 2).

Molecular characterization of maman based on *rbcL* sequence

For the potential study, the collection of plant samples transpired at the same time and same location. The resulting sequence was three sequences of *rbcL* with a length of 839 bp and no nucleotide differences. The said sequence has a previous insertion into Genbank with accession numbers OR898307, OR898308, and OR898309. The top 10 accessions appearing in the BLASTn analysis will then proceed for phylogenetic tree analysis (Table 1). The maman from the Riau *rbcL* sequence had similarity values and a percentage identity of 98.45% to 99.88%. The highest value emerged with *Gynandropsis gynandra* (99.88%) and the lowest value with *Tarenaya aculeata* (98.45%), while the other eight accessions had similarity values of 98.69% and 98.57%. The highest similarity score (99.88%) came from the BOLD identification system on *G. gynandra* (Table 1).

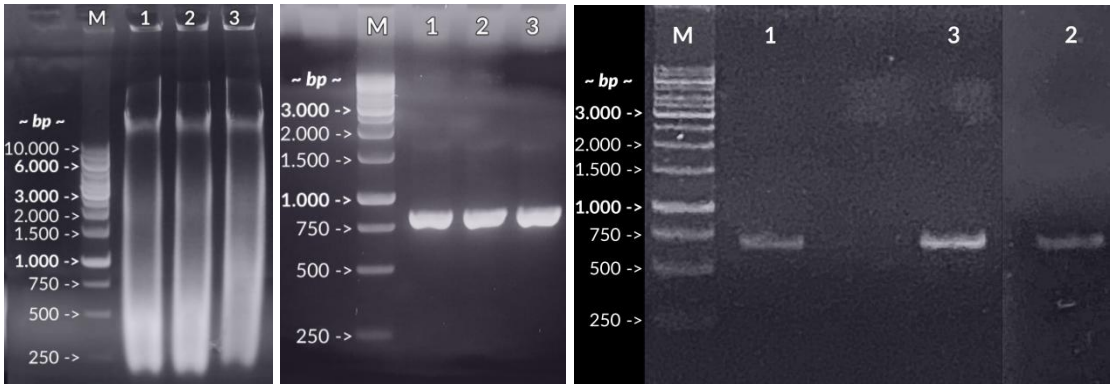

Figure 2. Profile of (a) the total DNA, (b) *rbcL*, and (c) *ITS* amplicon bands of maman from Riau. (M = 1 kb DNA Ladder from Thermo Scientific and 1, 2, 3 = biological replicate samples).

Table 1. The BLASTn analysis of maman species from Riau *rbcL* sequence.

Species Name	Query Cover (%)	E-Value	Identity (%)	Accession
Gynandropsis gynandra	100	0.0	99.88	MH767477.1
Corynandra viscosa	100	0.0	98.69	MH767481.1
Podandrogyne matthewsii	100	0.0	98.69	KU739633.1
Coalisina paradoxa	100	0.0	98.69	NC_066812.1
Tarenaya hassleriana	100	0.0	98.57	NC_034364.1
Tarenaya spinosa	100	0.0	98.57	KU739634.1
Tarenaya rosea	100	0.0	98.57	KU739624.1
Cleome arabica	100	0.0	98.57	KU739616.1
Podandrogyne chiriquensis	100	0.0	98.57	AY483269.1
Tarenaya aculeata	100	0.0	98.45	KU739631.1

Overall, 811 invariable (monomorphic) sites and 28 variable (polymorphic) sites and variations of the nucleotides existed. Of the 28 polymorphic sites, eight were singleton variable sites, and 20 others were parsimony informative sites. Nucleotide diversity (per site) was Pi = 0.01079, and the average nucleotide difference was k = 9.05128. Fu's Fs statistic was -0.108 and Tajima's D was 0.01374, with a statistical significance value of $P \leq 0.10$. The one critical nucleotide was evident in sequence number 10, which only exists on maman accessions from Riau, Indonesia. However, five other critical nucleotides were notable for the species *G. gynandra* syn. *Cleome gynandra,* in general. The nucleotides were in sequence numbers 166, 202, 454, 466, and 790 (Table 2).

The largest composite likelihood reached estimation for the nucleotide substitution pattern. Utilizing the model of Tamura *et al*. (2021), the study evaluated the possibility of replacement. For pyrimidines, the transition/transversion rate ratio was k2 = 2.841, while for purines, it was k1 = 3.245. R = 1.497 represents the transition/transversion bias as a whole. The nucleotide frequency for each was A = 27.44%, T/U = 29.38%, C = 19.09%, and G = 24.09%. The overall genetic range among the analyzed accessions relied on the *rcbL* sequence (0.000–0.016), with the lowest genetic distance value (0.001) achieved among the maman accessions from Riau and *G. gynandra*. However, among the maman accessions from Riau with *T. aculeata*, the highest genetic distance value was 0.016.

Table 2. Nucleotide variation in *rbcL* sequence.

	Nucleotide number (vertically)																										
No				1	2	2	3	3	3	3	4	4	4	4	4	4	4	5	6	6	6	6	7	7	7	8
.	1	2	6	6	0	3	2	4	5	7	9	0	2	4	5	6	8	9	7	0	5	7	8	9	3	6	9	1
	0	3	1	6	2	8	9	7	6	9	4	6	7	1	4	6	7	9	5	7	0	3	5	5	4	0	0	1
1	T	T	C	A	G	A	T	C	G	C	G	G	C	G	C	T	A	C	A	C	A	T	A	T	A	C	G	G
2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
5	G	.	T	G	T	.	G	.	.	.	.	A	.	T	G	C	.	.	C	.	.	.	.	.	.	.	A	.
6	G	C	.	G	T	.	.	.	.	.	T	.	T	.	G	C	.	T	.	.	.	.	.	C	.	.	A	.
7	G	.	T	G	T	G	.	.	.	.	.	.	.	T	G	C	.	.	.	T	.	.	.	.	.	.	A	A
8	G	.	.	G	T	.	.	A	.	.	.	.	.	.	G	C	G	.	.	.	G	G	G	.	.	T	A	.
9	G	.	.	G	T	.	.	A	.	.	.	.	.	.	G	C	G	.	.	.	G	G	G	.	.	T	A	.
10	G	.	.	G	T	.	.	A	.	A	.	.	.	.	G	C	G	.	.	.	G	G	G	.	.	.	A	.
11	G	.	.	G	T	.	.	A	.	A	.	.	.	.	G	C	G	.	.	.	G	G	G	.	.	.	A	.
12	G	C	.	G	T	.	.	.	.	.	T	.	T	.	G	C	.	T	.	.	G	.	.	C	.	.	A	.
13	G	.	.	G	T	.	.	A	T	.	.	.	.	.	G	C	G	.	.	.	G	G	G	.	G	.	A	.

Note: (1-3) Maman from Riau, 4) *Gynandropsis gynandra*, 5) *Corynandra viscosa,* 6) *Podandrogyne matthewsii*, 7) *Coalisina paradoxa*, 8) *Tarenaya hassleriana*, 9) *Tarenaya spinosa,* 10) *Tarenaya rosea*, 11) *Cleome arabica*, 12) *Podandrogyne chiriquensis,* 13) *Tarenaya aculeata*. Green highlight = critical nucleotide for maman plants from Riau, yellow highlight = critical nucleotides for *G. gynandra–C. gynandra*.

Molecular characterization of maman based on *ITS* sequence

The same conditions and maman accessions used for *rbcL* analysis underwent the *ITS* analysis. The analysis also obtained three sequences with a length of 672 bp, and with no nucleotide differences among the maman accessions. The said sequence had also reached insertion into the GenBank with the accession numbers of PP959341, PP959342, and PP959343. Collecting the BLASTn analysis results for this sequence also proceeded on 10 maman accessions with similarity values and percentage identity ranging from 100% to 85.44%. Perfect similarities (100%) were evident in *G. gynandra* and *C. gynandra*, while the lowest values observed resulted in *C. briquetii* (Table 3).

The results showed 263 sites with alignment gaps, 351 monomorphic sites, and 130 polymorphic sites. The polymorphic sites were 55 singleton variable sites and 75 parsimony informative sites. The value of nucleotide diversity (per site) was Pi = 0.10182, and the average number of nucleotide differences was k = 48.97436. Fu's Fs statistic got 7.102 points and Tajima's D = -0.55716, with a statistical significance of P ≤ 0.10. Besides, in general, the 20 critical nucleotides were prominent for *G. gynandra* syn. *C. gynandra*; however, none of them were specific to maman accessions from Riau, Indonesia (Table 4).

The specific model used helped assess the likelihood of substitution (Tamura *et al*., 2021). The transition/transversion rate ratio for pyrimidines was k2 = 3.455, whereas for purines the value was k1 = 1.396. The entire transition/transversion bias has the representation of R = 1.173. The nucleotide frequencies for each base were A = 23.11%, T/U = 20.72%, C = 28.78%, and G = 27.39%.

On the *ITS* sequence, the genetic distance ranges from 0.00 to 0.22. The lowest value appeared for all the maman accessions from Riau on its intra-species, as well as, with *G. gynandra* and *C. gynandra*. Moreover, an obtained value of 0.00 for *Corynandra viscosa* occurred with *C. viscosa*. However, the highest score (0.18) manifested for maman accessions from Riau was with *Peritoma lutea*.

Table 3. The BLASTn analysis of maman species from Riau *ITS* sequence.

Species Name	Query Cover %)	E-Value	Identity (%)	Accession
Gynandropsis gynandra	100	0.0	100	MH768084.1
Cleome gynandra	98	0.0	100	HM044252.1
Corynandra viscosa	96	1e-178	87.18	ON908405.1
Cleome viscosa	95	8e-175	87.02	KF923186.1
Cleome cleomoides	93	1e-177	87.84	KF923166.1
Cleome iberidella	93	2e-160	85.77	KF923174.1
Cleome foliosa	88	2e-171	86.92	KF923173.1
Peritoma lutea	93	8e-170	87.11	KF923177.1
Cleome briquetii	95	4e-158	85.44	KF923165.1
Cleome densifolia	95	1e-163	86.10	KF923168.1

Table 4. Twenty critical nucleotides in *ITS* sequence.

No.	Nucleotide number (vertically)																			
	2	2	2	2	2	2	2	2	2	4	4	5	5	5	6	6	6	6	6	6
	0	3	3	4	5	8	8	8	9	8	9	0	7	7	3	4	5	5	5	7
	5	0	4	5	3	0	1	4	0	0	8	9	2	5	2	6	1	2	3	8
1	T	A	T	C	T	G	T	G	T	T	C	T	T	T	A	G	A	C	T	G
2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
6	G	G	C	G	C	C	G	T	C	C	T	A	G	C	G	T	G	A	C	A
7	G	G	C	G	C	C	G	T	C	C	T	A	G	C	G	T	G	A	C	A
8	G	G	C	G	C	C	G	T	C	C	T	A	G	C	G	C	G	A	C	A
9	G	G	C	G	C	T	G	T	C	C	T	A	A	C	G	C	T	G	C	A
10	G	G	C	G	C	C	G	T	C	C	T	A	G	C	G	C	G	A	C	T
11	G	G	C	G	C	T	G	T	C	C	A	A	G	C	G	C	G	A	C	A
12	G	G	C	A	C	T	G	T	C	C	T	A	A	C	G	C	C	G	C	T
13	G	G	C	G	C	T	G	T	C	C	T	A	A	C	G	T	T	G	C	T

Note: 1-3) Maman species from Riau, 4) *Gynandropsis gynandra*, 5) *Cleome gynandra* 6) *Corynandra viscosa*, 7) *Cleome viscosa*, 8) *Cleome cleomoides*, 9) *Cleome iberidella,* 10) *Cleome foliosa*, 11) *Peritoma lutea*, 12) *Cleome briquetii,* 13) *Cleome densifolia*.

Single and multilocus analysis for phylogenetic tree

The results of the phylogenetic tree obtained from single-locus showed the maman accessions from Riau formed groups of branches with their neighbors. In addition to the accessions themselves, the tree based on *rbcL* also grouped with *G. gynandra*. As for the tree based on *ITS*, this group also had the following: *G. gynandra* and *C. gynandra* (Figure 3). The results of the phylogenetic tree depended on multilocus analysis, such as, combinations of *rbcL+ITS* and *rbcL+ITS+matK+trnL-trnF IGS* (Figure 4). The whole produced branch indicated maman accessions from Riau grouped with *G. gynandra*.

DISCUSSION

The single and multilocus analyses indicated maman accessions from Riau confirmed the species *C. gynandra* or *G. gynandra*. This statement also received support from the BLASTn analyses, which have two confirmations on different web platforms (NCBI and BOLD), showing the identity results in 99.88%–100%. Nonetheless, the presented study further indicated maman accessions from Riau had a critical nucleotide. This critical

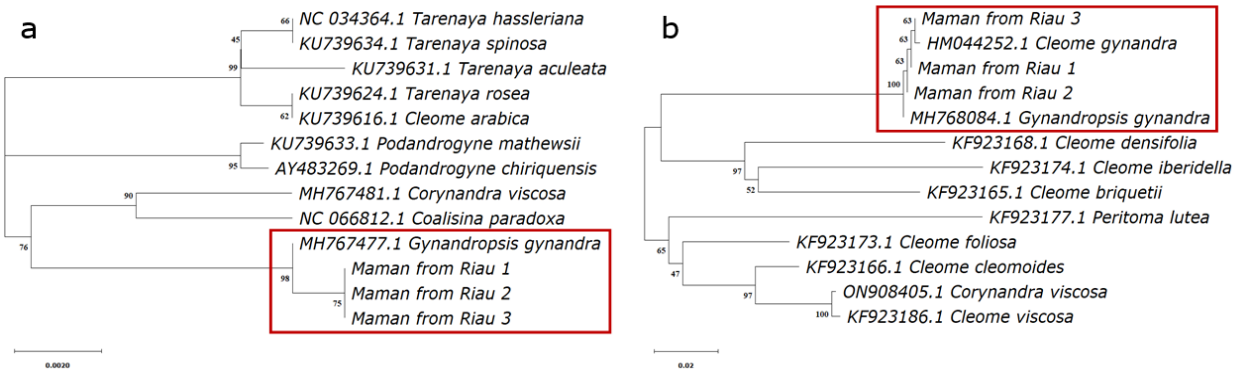

Figure 3. The single-locus phylogenetic tree of maman species from Riau and other species from NCBI using the Neighbor-joining method with 1000× bootstraps. a) *rbcL*, and b) *ITS*.

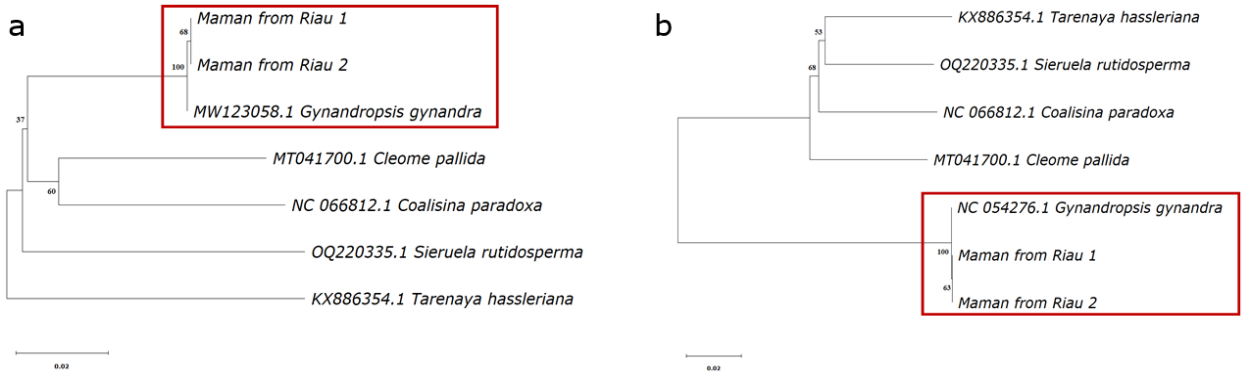

Figure 4. The multi-locus phylogenetic tree of maman species from Riau and other species from NCBI using the Neighbor-joining method with 1000× bootstraps. a) *rbcL+ITS*, and b) *rbcL+ITS+matK+trnL-trnF IGS*.

nucleotide was in order number 10 in the *rbcL* sequence, namely, T for maman accession from Riau, and G for other accessions, and was even for *G. gynandra* in Xisha, China (Li *et al*., 2018).

In the existing study, the species of *G. gynandra* and *C. gynandra,* collected from the NCBI, originated from different geographical regions. *Gynandropsis gynandra* originated from tropical islands in China, and *C. gynandra* was the accession sent to museums outside of Indonesia (Feodorova *et al*., 2010). Although, with the same species, the maman accessions from the tropical regions of Riau, Indonesia, with different climatic and environmental conditions, encouraged them to mutate and adapt to the Riau environment, having the greatest peat land area (4.04 million hectares) in Sumatera (Juniyanti and Situmorang 2023). Peatland forests, mainly lowland forests, primarily spread in Sumatra and Borneo Islands, Indonesia, are witnessing land cover changes due to the vast plantation expansion (Kuswanda *et al.,* 2022). These findings will be very useful for researchers in identifying maman accessions in the future.

Furthermore, the *ITS* sequence in this study was unable to discriminate the maman accessions from Riau and its fellow species, as showing 100% identity results, with no genetic distance. It also did not show the critical nucleotide specific to maman accessions of Riau. Therefore, it was understandable that the *rbcL* sequence can be more capable of distinguishing the maman accessions from Riau and its fellow species from a molecular

perspective and is recommended for use as a DNA barcode for identification.

However, the *ITS* can strongly discriminate the *C. gynandra* syn. *G. gynandra* and other species. Therefore, the presented study has proven that the *ITS* sequence has as many as 20 critical nucleotides for this species. These results significantly differed from the *rbcL,* which has only five critical nucleotides for this species. The *rbcL* sequence emerged rich in A+T nucleotides (56.83%), whereas the *ITS* sequence was rich in G+C nucleotides (56.17%). A rich region in the A+T sequence will be more likely to denature and undergo replication stress; thus, its structure becomes unstable compared with the rich G+C sequence with stronger hydrogen bonds to guanine and cytosine (Vinogradov and Anatskaya, 2017; Zhang *et al*., 2022). According to Lynch *et al*. (2016), the rich A+T sequences will have a high level of genetic diversity due to the accumulation of sequence variations from the highest mutation.

The single-locus phylogenetic tree also supports the statement that *rbcL* was more suitable as a DNA barcode than *ITS* on maman accessions from Riau. On the phylogenetic tree based on the *ITS* sequence, the maman accession numbered 3 got separated from the accessions 1 and 2 and was seemingly identical to the accessions of *C. gynandra* and *G. gynandra*. These results enunciated the inability of the *ITS* to discriminate the maman accessions from Riau and its fellow species. It was different from the tree based on *rbcL*, which has already grouped the third maman accession on the same branch among its neighbors and did not separate.

The most closely related species of maman accessions from Riau based on BLASTn analysis were *P. matthewsii* based on *rbcL* (Table 2) and *C. viscosa* based on *ITS* (Table 3). The results of these two markers were also unalike, caused by incomplete sequence data for each sequence. However, solving this problem can continue by using multilocus analysis to obtain the data on each species' proximity represented by more than one sequence and get the properly related species. Regarding the multilocus study, *C. pallida* became the closest accession with *C. gynandra* syn. *G. Gynandra* (Figure 4). Moreover, the multilocus analysis of the combination of *rbcL+ITS* and *rbcL+ITS+matK+trnL-trnL-trnF IGS* gave the same results, namely, both could group maman from Riau in one cluster with *G. gynandra* and separate it from other studied accessions.

Other research results also used a combination method of several markers, such as that conducted by Wang *et al*. (2022), which found combining *ITS*, *matK*, *rbcL,* and *trnL-F* produced a much higher identification effect than utilizing a single fragment. This means the identification process in different types of plants has a different combination of markers to obtain clear identification results, but the combination of using various markers does provide a clearer resolution. In the future, we recommend using multilocus for identification of plant types. The study also suggests researching more the standard for similarity value for annual plants because no findings exist for this. Authors hypothesize that since annual plants have a faster life cycle, they will accumulate more mutations due to their adaptation to the environment than the longer-living plants.

The identification research using this method faces limitations due to high costs, time requirements of multilocus sequencing, complex data analysis needed to differentiate it from closely related species, and dependency on adequate genomic reference data for accurate results. But, besides all that, the precise identification of *Maman* (*Cleome gynandra*) through optimized genetic sequences enhances its potential in agriculture and medicinal research, supporting biodiversity conservation and promoting its effective use as a valuable crop and medicinal plant.

CONCLUSIONS

Maman accessions from Riau reached validation as a species of *C. gynandra* or *G. gynandra*. However, based on the *rbcL* sequence, maman in this study is another variety of this species because it has one critical nucleotide only found in the maman sample from Riau. The *rbcL* sequence proved

more capable of distinguishing the maman accessions from Riau than the *ITS,* and emerged highly suggested serving as a DNA barcode for this plant. The phylogenetic analysis with multilocus appeared very helpful in obtaining more accurate and descriptive results.

ACKNOWLEDGMENTS

Authors thank the Directorate of Research, Technology, and Public Service; the Directorate General of Higher Education, Research, and Technology; and the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, for funding through Regular Fundamental Research with the contract number 20601/UN19.5.1.3/AL.04/2024. They also thank LPPM University of Riau for the facilities. Likewise, thanks to the Research Centre for Botanical Application, the Research Organization for Life Science and Environment, and the National Research and Innovation Agency, Indonesia (BRIN) for the facilities.

REFERENCES

Amandita FY, Rembold K, Vornam B, Rahayu S, Siregar IZ, Kreft H, Finkeldey R (2019). DNA barcoding of flowering plants in Sumatra, Indonesia. *Ecol Evol.* 9(4): 1858–1868.

Barthet MM, Pierpont CL, Tavernier E-K (2020). Unraveling the role of the enigmatic MatK maturase in chloroplast group IIA intron excision. *Plant Direct.* 4(3) Article e00208, 10.1002/pld3.208.

CBOL Plant Working Group (2009). A DNA barcode for land plants. *Proceed. Natl. Acad. Sci., USA.* 106: 12794–12797.

Chandradevan M, Simoh S, Mediani A, Ismail NH, Ismail IS, Abas F (2020). UHPLC-ESI-Orbitrap-MS analysis of biologically active extracts from *Gynura procumbens* (Lour.) Merr. and *Cleome gynandra* L. leaves. *Evid-Based Compl. Alt.* 2020(5): 1–14. https://doi.org/10.1155/2020/3238561.

Dong W, Cheng TLC, Xu C, Long P, Chen C, Zhou S (2014). Discriminating plants using the DNA barcode *rbcLb*: An appraisal based on a large data set. *Mol. Ecol. Resour*. *14*(2): 336–343. https://doi.org/10.1111/1755-0998.12185.

Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, Cheng T, Guo J, Zhou S (2015). *ycf1* the most promising plastid DNA barcode of land plants. *Sci Rep*. 5:8348.

Feodorova TA, Voznesenskaya EV, Edwards GE, Roalson EH (2010). Biogeographic patterns of diversification and the origins of C4 in *Cleome* (Cleomaceae). *Syst. Bot*. 35(4): 811–826. https://doi.org/10.1600/036364410X539880.

Gostel MR, Kress WJ (2022). The expanding role of DNA barcodes: Indispensable tools for ecology, evolution, and conservation. *Diversity* 14(3): 213. https://doi.org/10.3390/d14030213.

Guo M, Xu Y, Ren L, He S, Pang X (2018). A systematic study on DNA barcoding of medicinally important genus *Epimedium* L. (Berberidaceae). *Genes (Basel).* 9:623.

Hall TA (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids*. 41: 95–98.

He T, Jiao L, Yu M, Guo J, Jiang X, Yin Y (2019). DNA barcoding authentication for the wood of eight endangered Dalbergia timber species using machine learning approaches. *Holzforschung*. 73:277–85.

Hollingsworth PM, Li DZ, van der Bank M, Twyford AD (2016). Telling plant species apart with DNA: From barcodes to genomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150338. doi: 10.1098/rstb.2015.0338.

Juniyanti L, Situmorang ROP (2023). What causes deforestation and land cover change in Riau Province, Indonesia. *Forest Policy and Economics.* 153(2023):102999 https://doi.org/10.1016/j.forpol.2023.102999.

Kuswanda W, Garsetiasih R, Gunawan H, Situmorang ROP, Hutapea FJ, Kwatrina RT, Karlina E, Atmoko T, Zahrah M, Takandjandji M, Gunaryadi D (2022). Can humans and elephants coexist? A review of the conflict on Sumatra Island, Indonesia. *Diversity*. 14 (2022), 10.3390/d14060420.

Letsiou S, Madesis P, Vasdekis E, Montemurro C, Grigoriou ME, Skavdis G, Moussis V, Koutelidakis AE, Tzakos AG (2024). DNA barcoding as a plant identification method. *Appl. Sci.* (*Switzerland*). 14(4): 1415. https://doi.org/10.3390/app14041415.

Li DZ, Gao LM, Li HT, Wang H, Ge XJ, Liu JQ, Chen ZD, Zhou SL, Chen SL, Yang JB, Fu CX, Zeng CX, Yan HF, Zhu YJ, Sun YS, Chen SY, Zhao L, Wang K, Yang T, Duan GW (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (*ITS*) should be incorporated into the core barcode for seed plants. *Proc. Natl. Acad.*

Sci. USA. 108(49): 19641–19646. https://doi.org/10.1073/pnas.1104551108.

Li S, Qian X, Zheng Z, Shi M, Chang X, Li X, Liu J, Tu T, Zhang D (2018). DNA barcoding the flowering plants from the tropical coral islands of Xisha (China). *Ecol. Evol*. 8(21): 10587–10593. https://doi.org/10.1002/ece3.4545.

Lynch M, Ackerman MS, Gout J, Long H, Sung W, Thomas WK, Foster PL (2016). Genetic drift, selection and the evolution of the mutation rate. *Nat. Rev. Genet*. *17*: 704–714.

Maddison WP, Maddison DR (2023). Mesquite: A modular system for evolutionary analysis. Version 3.81. Available online: http://www.mesquiteproject.org.

Nehal N, Choudhary B, Nagpure A, Gupta RK (2021). DNA barcoding: A modern age tool for detection of adulteration in food. *Crit. Rev. Biotechnol.* 41(5): 767–791. https://doi.org/10.1080/07388551.2021.1874279.

Rahman MSA, Ali R (2020). Application of *rbcL* in DNA barcoding of *Phaleria macrocarpa* (Scheff.) Boerl. and its processed commercial products. *Mater. Today: Proc.* 31(2020), pp. A58–A62, 10.1016/j.matpr.2020.12.131.

Rathnayake RAS, Wedage WMM, Muthukumarana LS, De Silva BGDNK (2023). Genetic diversity, phylogenetic and phylogeographic analysis of *Anopheles culicifacies* species complex using ITS2 and *COI* sequences. *PLoS ONE.* 18(8): e0290178. https://doi.org/10.1371/journal.pone.0290178.

Restusari L, Muharni, Fitri, Sjahjadi NV (2023). Activities of fermented maman (*Cleome gynandra* L) on blood sugar level of hyperglycemic white rats. *Local. Ther*. 02(01): 1–8.

Roslim DI (2019). Analysis of *matK, rbcL* and *trnL-trnF* Intergenic Spacer Sequences on Durik-Durik (*Syzygium* sp). *J. Phys.: Conf. Ser*. 1351(2019) 012023 doi:10.1088/1742-6596/1351/1/012023.

Roslim DI, Herman H, Putri SR, Furqoni AT, Budiono DYF, Altuhaish AAF, Lestari W (2023). Verification of maman (*Cleome gynandra* (L.) Briq.) from Riau based on *matK* and *trnL-trnL-trnF* intergenic spacer. *Biosaintifika*. 15(3): 296–305. https://doi.org/10.15294/biosaintifika.v15i3.46937.

Sari VK, Murti RH (2015). An effective method for DNA extraction of mature leaf of sapodilla (*Manilkara zapota* (L.) van Royen). *Agrivita* 37(1): 18–23. https://doi.org/10.17503/agrivita-2015-37-1-p018-023.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. *Proc. Natl. Acad. Sci. USA*. 109(16): 6241–6246. https://doi.org/10.1073/pnas.1117018109.

Sevindik E, Murathan ZT, Filiz S, Yalçin K (2019). Molecular characterization based on chloroplast (*trnl-F*) DNA sequence of the apple genotypes in Ardahan/Turkey. *Bangladesh Journal of Botany*. 48(4), 1099–1106. https://doi.org/10.3329/bjb.v48i4.49058.

Tamura K, Stecher G, Kumar S (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol*. 38(7): 3022–3027. https://doi.org/10.1093/molbev/msab120.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 22(22): 4673–4680.

Tnah LH, Lee SL, Tan AL, Lee CT, Ng KKS, Ng CH, Nurul FZ (2019). DNA barcode database of common herbal plants in the tropics: A resource for herbal product authentication. *Food Control* 95(2019): 318–326. https://doi.org/10.1016/j.foodcont.2018.08.022.

Vinogradov AE, Anatskaya OV (2017). DNA helix: The importance of being AT-rich. *Mamm. Genome* 28(9–10): 455–464. https://doi.org/10.1007/s00335-017-9713-8.

Wang J, Yan Z, Zhong P, Shen Z, Yang G, Ma L (2022). Screening of universal DNA barcodes for identifying grass species of Gramineae. *Front. Plant Sci.* 13(2022): https://doi.org/10.3389/fpls.2022.998863.

Yu J, Wu X, Liu C, Newmaster S, Ragupathy S, Kress WJ (2021). Progress in the use of DNA barcodes in the identification and classification of medicinal plants. *Ecotox. Environ. Safe*. 208(2021): 111691. https://doi.org/10.1016/j.ecoenv.2020.111691.

Zhang Y, Wu L, Wang Z, Wang J, Roychoudhury S, Tomasik B, Wu G, Wang G, Rao X, Zhou R (2022). Replication stress: A review of novel targets to enhance radiosensitivity-from bench to clinic. *Front. Oncol.* 12: 838637. https://doi.org/10.3389/fonc.2022.838637.