RESEARCH ARTICLE

SABRAO Journal of Breeding and Genetics
54 (5) 1066-1076, 2022
http://doi.org/10.54910/sabrao2022.54.5.9
http://sabraojournal.org/
pISSN 1029-7073; eISSN 2224-8978

DNA BARCODING OF *CANANGA ODORATA* (LAM. HOOK.F. & THOMSON) USING CHLOROPLAST GENES IN NORTH SUMATRA, INDONESIA

E. PRASETYA[1*], LAZUARDI[1], A. HASAIRIN[1], Y. RACHMAWATI[2], M. JANNAH[3], and T. HARSONO[1]

[1]Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Negeri Medan, Medan, Indonesia
[2]Department of Biology, Faculty of Sciences and Technology, Islamic State University, Surabaya, Indonesia
[3]Department of Biology, Faculty of Sciences and Technology, Islamic University of Assyafiiyah, East Jakarta, Indonesia
*Corresponding author's email: ekoprasetya.biologi@gmail.com
Email addresses of co-authors: drslazuardi@gmail.com, asharhasairin@unimed.ac.id, yuanitarhartono@uinsby.ac.id, mifta.frozi@gmail.com, triharsonounimed@gmail.com

SUMMARY

Cananga odorata (Lam.) Hook.f. & Thomson, a member of the Annonaceae family, has become widespread from the Indo-Malayan region to northern Australia and Malaysia. This plant has economic value because it produces essential oils used in the perfume and food industries. No research existed yet related to DNA barcoding to see intraspecific variations in *C. odorata*. This study aimed to analyze the potential of chloroplast genes *maturase* K (*mat*K)-*trn*K, ribulose 1,5-biphosphate carboxylase (*rbc*L), and *trn*L-F intergenic spacer as DNA barcodes on *C. odorata* from North Sumatra, Indonesia. The research sample obtained from five areas included Tanjung Balai, Langkat Regency (Tangkahan Conservation Forest), Medan, Tanjung Morawa (Deli Serdang Regency), and Sukarasa (Deli Serdang Regency). Phylogenetic analysis using the three markers showed that *C. odorata* was monophyletic grouped from a common ancestor. The results of the phylogenetic tree construction also showed that the three markers used were able to group each genus in the Annonaceae family and form separate branches from the outgroup (*Magnolia*, *Sarcandra*, and *Litsea*). The barcoding gap on the genetic distance could not be observed in this study, although phylogenetic results showed that the three markers were discriminatory enough to separate species. The results showed that the *mat*K-*trn*K, *rbc*L, and *trn*L-F markers could be effectively used as DNA barcodes to identify *C. odorata*.

Keywords: *Cananga odorata*, DNA barcoding, chloroplast genes, North Sumatra

Key findings: There are no barcoding gaps in the *mat*k-*trn*K, *rbc*L, and *trn*L-F markers, although phylogenetic results show that three barcode markers are quite discriminatory to separate the Annonaceae species. DNA barcode *mat*K-*trn*K, *rbc*L, and *trn*L-F revealed as efficient tools to identify *C. odorata*.

Communicating Editor: Prof. Dr. Bambang Sapta Purwoko

Manuscript received: August 4, 2022; Accepted: November 29, 2022.

To cite this manuscript: Prasetya E, Lazuardi, Hasairin A, Rachmawati Y, Jannah M, Harsono T (2022). DNA barcoding of *Cananga odorata* (Lam. Hook.F. & Thomson) using chloroplast genes in North Sumatra, Indonesia. *SABRAO J. Breed. Genet.* 54(5): 1066-1076. http://doi.org/10.54910/sabrao2022.54.5.9.

INTRODUCTION

Extensive functional convergence in Annonaceae flowers with widespread homoplasy in many characters (Saunders, 2009) becomes an issue in identifying species in this family. Recent molecular phylogenetic studies on this family have generated a broad consensus by producing many changes in generic delimitation involving the incorporation of (Zhou *et al.*, 2009) or separation of (Mols *et al.*, 2008) the existing genus. Genus *Cananga* is a member of the Annonaceae family reported to have early-divergent lineage (Doyle *et al.*, 2004; Richardson *et al.*, 2004; Couvreur *et al.*, 2008). *C. odorata* is a member of the *Cananga* genus, which is still rarely studied regarding molecular markers for the plant's identification.

A tall tree, *C. odorata,* grows 10–30 m in height, with very fragrant flowers and greenish yellow when ripe (Mallavarapu *et al.*, 2016). *C. odorata* generally blooms all year round (Manner and Elevitz, 2006), peaking in April to June and October to December (De Bontin, 2006). *C. odorata* is a tropical tree native to Indonesia and Malaysia but was naturalized by cultivation in several other countries, such as the Pacific Islands, Northern Australia, Thailand, Vietnam, and Madagascar (Manner and Elevitz, 2006). *C. odorata* produces essential oils of economic value in the perfume and food industries (Mallavarapu *et al.*, 2016). In Indonesia, this plant closely relates to various traditional activities of the Malay tribe.

One of the best approaches to unequivocally identify species uses DNA barcoding techniques. DNA barcoding aims to identify species by utilizing specific DNA short sequences that allow identifying species and discrimination to distinguish between species. DNA barcoding allows taxonomists to quickly sort specimens and highlight changing or different taxa that might lead to new species (Ali *et al.*, 2014). Currently, the International Barcode of Life (iBOL) carries out an international consortium to maintain DNA barcode references (Ratnasingham and Hebert, 2007). The Consortium for the Barcode of Life (CBOL) recommends markers of two combination loci chloroplast— ribulose-1, 5-bisphosphate carboxylase/oxygenase large subunit gene (*rbc*L) and maturase K gene (*mat*K)—as barcodes (Hollingsworth *et al.*, 2009). Several cases showed that the markers *rbc*L and *mat*K were unable to discriminate species in closely related taxa (Zhang *et al.*, 2012; Ali *et al.*, 2014) or divergent species (Van-Velzen *et al.*, 2012). Therefore, the use of additional sequences, such as, the *trn*L-F intergenic spacer, is recommended for identification purposes (Hollingsworth *et al.*, 2009; Taberlet *et al.*, 2007, 1991).

Research related to molecular markers of the Annonaceae family showed that *C. odorata* nests in *Cyathocalyx,* although only with 66% bootstrap support (Surveswaran *et al.*, 2010). Another study showed that *C. odorata* and *Cyathocalyx* form a polytomy with a combination of *rbc*L and *trn*L-F markers (Richardson *et al.*, 2004; Saunders *et al.*, 2011). The *trn*L-*trn*F intergenic spacer region revealed intergeneric relationships by showing a slight homoplasy in the Annonaceae family, however, the combination of *trn*L-F and *rbc*L resulted in higher resolution results with higher bootstrap values, as well (Mols *et al.*, 2004).

As a plant closely related to local culture and has economic value, this plant needs conservation for further use in the future. Currently, no research on file on DNA barcoding to see an intraspecies variation in *C. odorata* exists. This study analyzed and evaluated the potential of *rbc*L, *mat*K-*trn*K, and *trn*L-F markers as DNA barcodes in *C. odorata* plants from North Sumatra, Indonesia. This research expects to provide important information about the use of DNA barcoding in *C. odorata,* in particular, and in the Annonaceae family, in general. The results of this study can provide important information regarding the conservation efforts of *C. odorata*.

MATERIALS AND METHODS

Sample collection and DNA extraction

The obtained fresh *C. odorata* leaves came from various regions in the province of North Sumatra, Indonesia, namely, Tanjung Balai, Langkat Regency (Tangkahan Conservation Forest), Medan, Tanjung Morawa (Deli Serdang Regency), and Sukarasa (Deli Serdang Regency) (Figure 1). Storing the fresh leaf samples obtained from the field received temperature at -20 °C for long-term storage. Performing the DNA isolation was performed using the GeneJet Plant Genomic DNA Purification Kit (Thermo Fisher Scientific, Waltham, MA, USA) followed the manufacturer's recommended protocol. Using gel electrophoresis and gel documentation (Bio-Rad Laboratories, Hercules, CA, USA) ensured the purity of the DNA isolates.

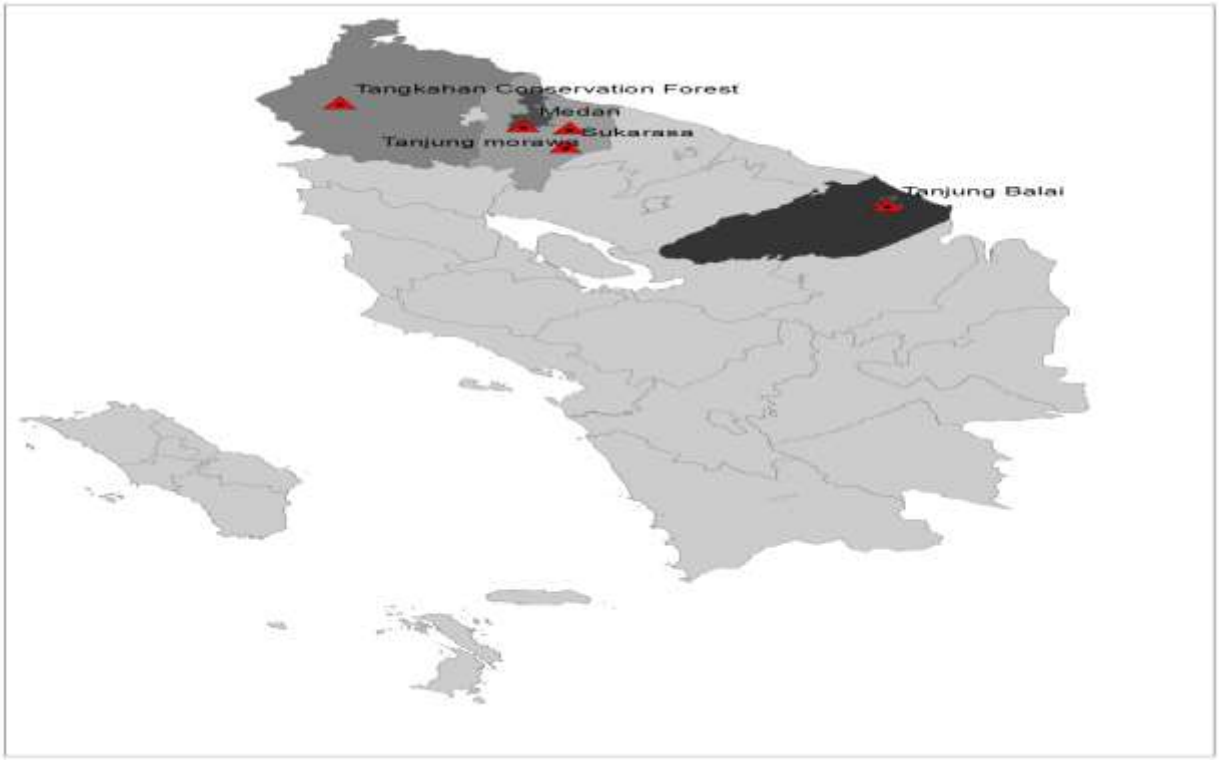


Figure 1. Sampling locations in North Sumatra, Indonesia at five points in four regencies. (*Source*: Google Map and processed using ArcGis software).

Amplification and sequencing

Polymerase chain reaction (PCR) proceeded with a total volume of 25 µL (2.5 µL of DNA template; 2.5 µL of forward primer, 2.5 µL of reverse primer; 5 µL of distilled water, and 12.5 µL of PCR mix [MyTaq HS Red Mix, Bioline, USA]) with a template DNA concentration of 50 ng. Primers for *mat*K-*trn*K sequences are *Mat*K-6F (5'-TGG GTT GCT AAC TCA ATG G-3') and *Mat*K-1R (5'-AAC TAG TCG GAT GGA GTA G-3') (Johnson and Soltis, 1995); primers for *rbc*L sequences are *rbc*L a_f (5'- ATG TCA CCA CAA ACA GAG ACT AAA GC-3') and *rbc*L a_rev (5'-GTA AAA TCA AGT CCA CCR CG-3') (Costion *et al.*, 2011), and primers for *trn*L-F sequences are *trn*L-F F (5-GGT TCA AGT TCT ATC CCC CC-3') and *trn*L-F R (5'-ATT TGA GAC ACG AG ACT GGT-3') (Taberlet *et al.*, 1991). The PCR product visualization used 1% agarose gel. A repeat of PCR occurs if the DNA band is not visible. Results from PCR products that show positive (visible DNA bands) will proceed to the sequencing process at the First Base DNA Sequencing Service in Singapore. The sequencing method used was Sanger Bi-directional PCR product sequencing.

Statistical analysis

The sequencing results analysis used the Bioedit application (Hall, 1999) to determine the consensus sequence. The research used a tree-based method in studying the DNA barcodes to analyze the resolution efficiency of the three markers and their combinations. The phylogenetic tree reconstruction used the Maximum Likelihood (ML) method using the MEGA 11 application (Tamura *et al.*, 2021). A marker rates discriminatory if it can group the same species in one branch (Hollingsworth *et al.*, 2009). ML analysis used the Kimura-2-parameter model method with 1000 bootstrap replicates (Kimura, 1980). Pairwise K2P (Kimura 2-parameter) distances for all three DNA regions were calculated in MEGA 11 to evaluate intraspecific and interspecific divergence in *C. odorata*. The use of average interspecific distance and the smallest interspecific distance represented the interspecific divergence (Meyer and Paulay, 2005). This study also used several DNA sequences of close relatives of *C. odorata* taken from the GeneBank National Center for Biotechnology Information (NCBI) as an outgroup.

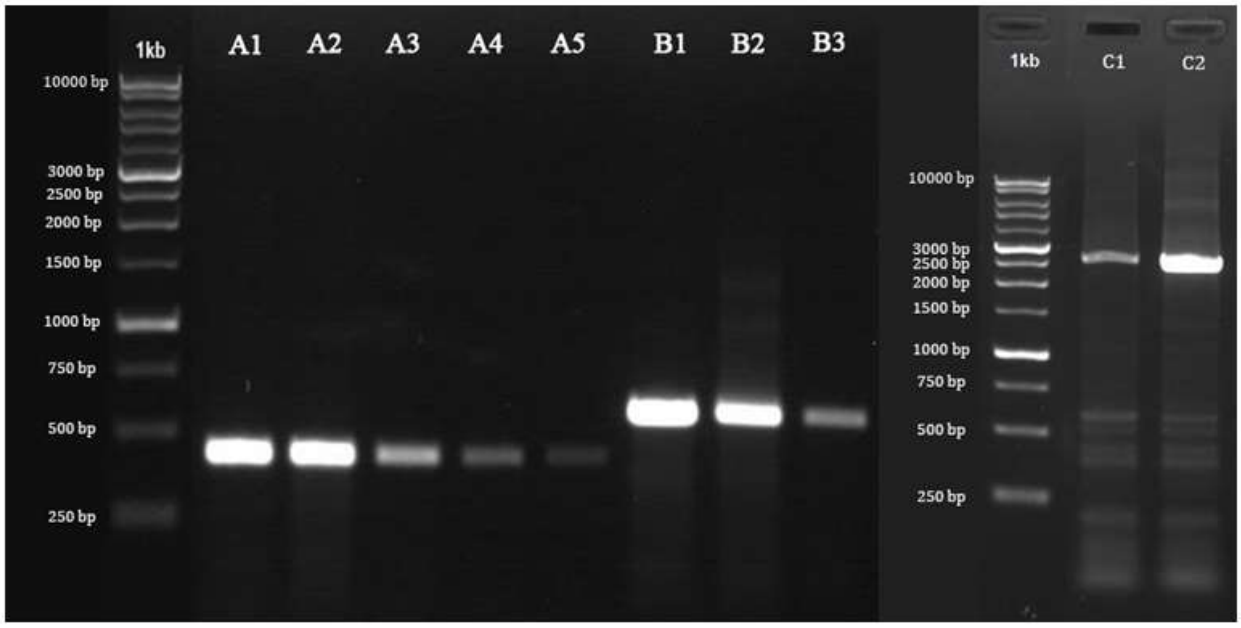


Figure 2. Visualization of amplification results with three markers (*rbc*L, *trn*L-F, and *mat*K-*trn*K) with 1% agarose gel. A1: Deli Serdang, *trn*L-F; A2: Langkat, *trn*L-F; A3: Sukarasa, *trn*L-F; A4: Tanjung Balai, *trn*L-F; A5: Medan, *trn*L-F; B1: Deli Serdang, *rbc*L; B2: Langkat, *rbc*L; B3: Tanjung Balai, *rbc*L; C1: Deli Serdang, *mat*K-*trn*K; C2: Langkat, *mat*K-*trn*K.

RESULTS AND DISCUSSION

Amplifying three markers used in the study of DNA chloroplast *C. odorata* showed varying results. A total of two samples using the *mat*K-*trn*K intergenic spacer, three samples using the *rbc*L marker, and five samples using the *trn*L-F marker achieved successful amplification. The results of PCR visualization using agarose gel appears in Figure 2.

The results of sample sequencing that show positive outcomes indicate variations in the level of success. A total of two samples with the *mat*K-*trn*K marker (Langkat and Deli Serdang), three samples with the *rbc*L marker (Langkat, Deli Serdang, and Tanjung Balai), and five samples with the *trn*L-F marker (Langkat, Deli Serdang, Tanjung Balai, Sukarasa, Medan) attained successful sequencing. The success of sequencing on flowering plants in Sumatra, Indonesia, using *rbc*L and *mat*K-*trn*K markers showed 94.7% and 65.8%, respectively (Amandita *et al.*, 2019). The success of sequencing using the *trn*L-F marker in five species of the Annonaceae family showed 100% success (Lestari and Azrianingsih, 2019). The sequencing results showed that the *rbc*L marker in *C. odorata* had the highest intraspecific homology (99.8%) compared with the *mat*K-*trn*K (99.4%) and *trn*L-F (99.3%) markers.

The nucleotide composition of amplicons of *C. odorata* showed variations in AT and GC content. GC content is one of the first metrics to be measured in genomic character analysis. A phylogenetic tree rich in GC content produces contradictory trees compared with rich AT content, which leads to tree construction bias (Romiguier and Roux, 2017). In the *rbc*L marker amplicons, the nucleotide compositions ranged from 55.0%–55.5% (AT) and 44.5%–45% (GC), with lengths ranging from 576–587 bp. In the *mat*K-*trn*K marker amplicons, the nucleotide compositions ranged from 63% (AT) and 37% (GC), with lengths ranging from 2453–2468 bp, whereas the *trn*L-F marker amplicons had nucleotide compositions ranging from 62.4%–62.6% (AT) and 37.3%–37.6% (GC), with a length ranging from 418–429 bp. Overall, all barcode areas used have a higher average AT content than the GC content. The percentage of GC content in the Annonaceae family (*Fissistigma, Mitrephora, Popowia, Anaxagorea, Annona, Artabotrys, Cananga,* and *Goniothalamus*) on the *trn*L marker ranged from 38.2%–36.3% (Lestari and Azrianingsih, 2019). In *Uvaria macrophylla* (Annonaceae), the overall GC content in the chloroplast

Table 1. Nucleotide composition, GC content, and AT content with three markers (*rbc*L, *trn*L-F, and *mat*K-*trn*K) in *C. odorata.*

Sample	Barcode	Composition (%)				Content (%)		Total (bp)
		A	C	G	T	A/T	G/C	
C. odorata Langkat	*rbc*L	27.6	21.4	23.6	27.4	55.0	45.0	576
C. odorata Deli Serdang	*rbc*L	27.6	21.1	23.3	27.8	55.5	44.5	587
C. odorata Tanjung Balai	*rbc*L	27.7	21.3	23.6	27.4	55.1	44.9	577
C. odorata Deli Serdang	*trn*L-F	29.6	20.7	16.9	32.9	62.4	37.6	426
C. odorata Langkat	*trn*L-F	29.2	20.3	17.0	33.3	62.6	37.4	418
C. odorata Sukarasa	*trn*L-F	29.5	20.6	16.9	32.8	62.4	37.6	427
C. odorata Tanjung Balai	*trn*L-F	29.7	30.6	16.9	32.8	62.5	37.5	427
C. odorata Medan	*trn*L-F	29.8	20.5	16.8	32.9	62.7	37.3	429
C. odorata Langkat	*mat*K-*trn*K	31.0	18.2	18.8	32.0	63.0	37.0	2468
C. odorata Deli Serdang	*mat*K-*trn*K	30.9	18.3	18.7	32.1	63.0	37.0	2453

Table 2. Characteristic features of *rbc*L, *mat*K-*trn*K, and *trn*L-F region among *C. odorata* species.

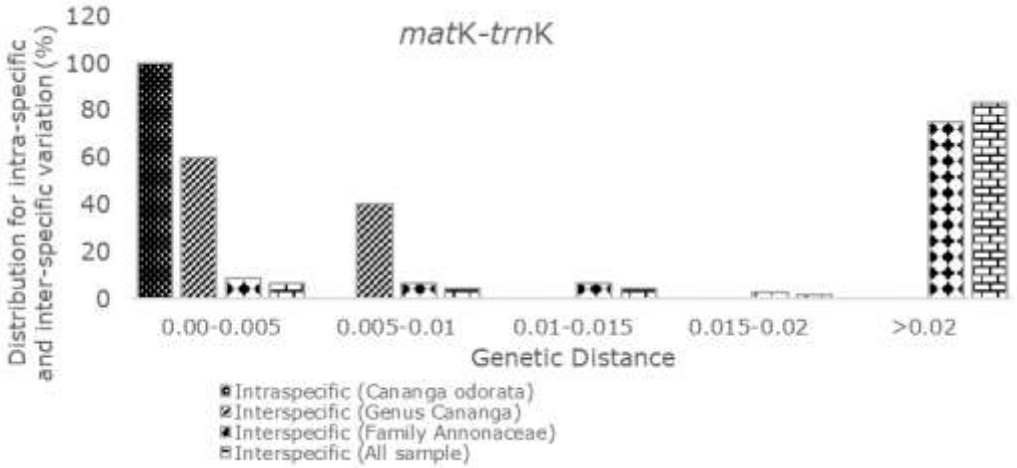
Locus	*rbc*L	*mat*K-*trn*K	*trn*L-F
Total number of sites	1434	2469	891
Conserved sites	1420 (99.02%)	2443 (98.95%)	425 (47.70%)
Variable sites	7 (0.49%)	9 (0.36%)	2 (0.22%)
Parsimony Informative sites	3 (0.21%)	1 (0.04%)	1 (0.11%)
Singleton sites	3 (0.21%)	1 (0.04%)	1 (0.11%)

genome was 38.7% (Fitmawati *et al*., 2018; Wu *et al.*, 2019). Table 1 presents the nitrogen base frequencies of the three markers in *C. odorata* .

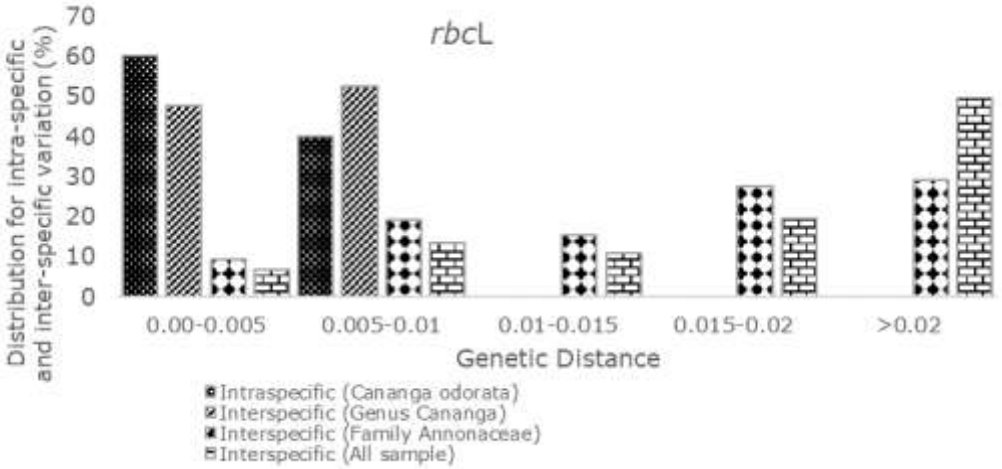
Previous studies reported that the sequencing lengths of the *rbc*L, *mat*K-*trn*K and *trn*L-F regions in the genera *Cananga*, *Cyathocalyx*, and *Drepananthus* were 1239, 737, and 1069 bp, respectively (Surveswaran *et al.*, 2010). The length of the barcode region in each species indicates a variation (Bousquet *et al.*, 1992; Hilu and Liang, 1997; Pirie *et al.*, 2007). The *mat*K sequence, which is encoded by the *trn*K intron, is common in land plants and is the only *maturase*-like gene in plant plastids (Chuang and Hu, 2004). The *mat*K in the *trn*K intron also serves as a standard marker for DNA barcode land plants promoted by the Barcode of Life's Plant Working Group (CBOL-PWG) (Cheng and Houston, 2021; Hollingsworth *et al.*, 2009). The length of the *mat*K-*trn*K sequence varied in various plant species, such as 2478 bp (*Lycopodiella cernua*), 2407 bp (*Selaginella doederleinii*), and 2322 bp (*Ophioglossum petiolatum*) (Chuang and Hu, 2004).

All samples from forward and reverse sequencing showed good quality results on the *mat*K-*trn*K, *rbc*L, and *trn*L-F markers. The sequence variability of the three barcode regions in *C. odorata* species reflects in Table 2. The length of the alignment results sequence was 1434 bp for *rbc*L, 2469 bp for *mat*K-*trn*K, and 891 bp for *trn*L-F. The variable sites in *rbc*L (0.49%) are the highest after *mat*K-*trn*K (0.36%) and *trn*L-F (0.22%). The highest level of parsimony informative sites resulted in the *rbc*L (0.21%) after *trn*L-F (0.11%) and *mat*K-*trn*K marker (0.04%). Conserved site levels of barcode markers in *C. odorata* species ranged, namely, *rbc*L (99.02%), *mat*K-*trn*K (98.95%), and *trn*L-F (47.70%). It indicates that the *rbc*L marker is more suitable for barcoding in *C. odorata* but may not necessarily be compatible with other species. Variable character ingroup in the *Disepalum* genus (Annonaceae) was 1.9% with the *mat*K marker and 3.3% with the *trn*L-F marker (Li *et al.*, 2015). In the genus *Fitzalania* and *Meiogyne* (Annonaceae), the character ingroup variables on the *mat*K, *rbc*L, and *trn*L-F markers were 6.6%, 2.2%, and 3.3%, respectively (Xue *et al.*, 2014).

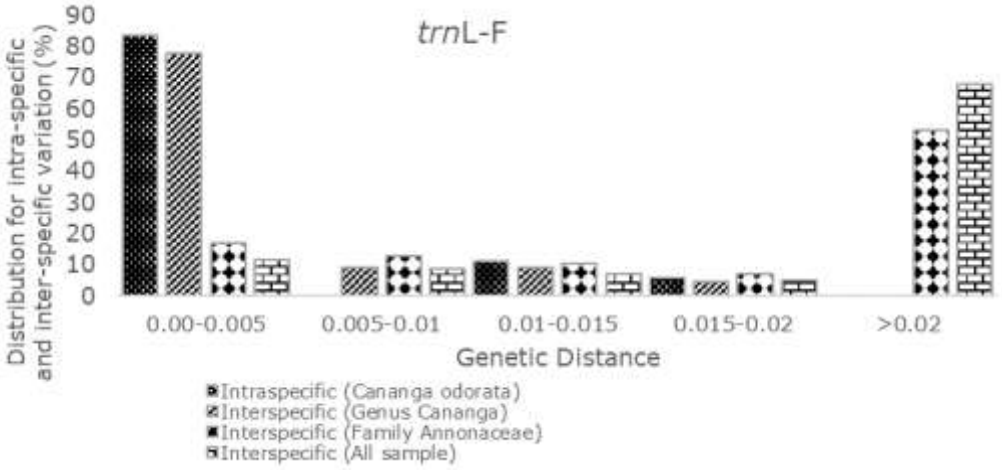
The analysis of estimates of the degree of variation in the three barcode regions used pairwise distance analysis using MEGA 11. The frequency distributions of interspecific and intraspecific divergences show in Figure 3. The markers *rbc*L, *mat*K-*trn*K, and *trn*L-F present lower intraspecific variation compared with the interspecific variation. The frequency of distribution of intraspecific variation was not more than 0.01, except for the *trn*L-F marker, which showed considerable variation. The



(A)

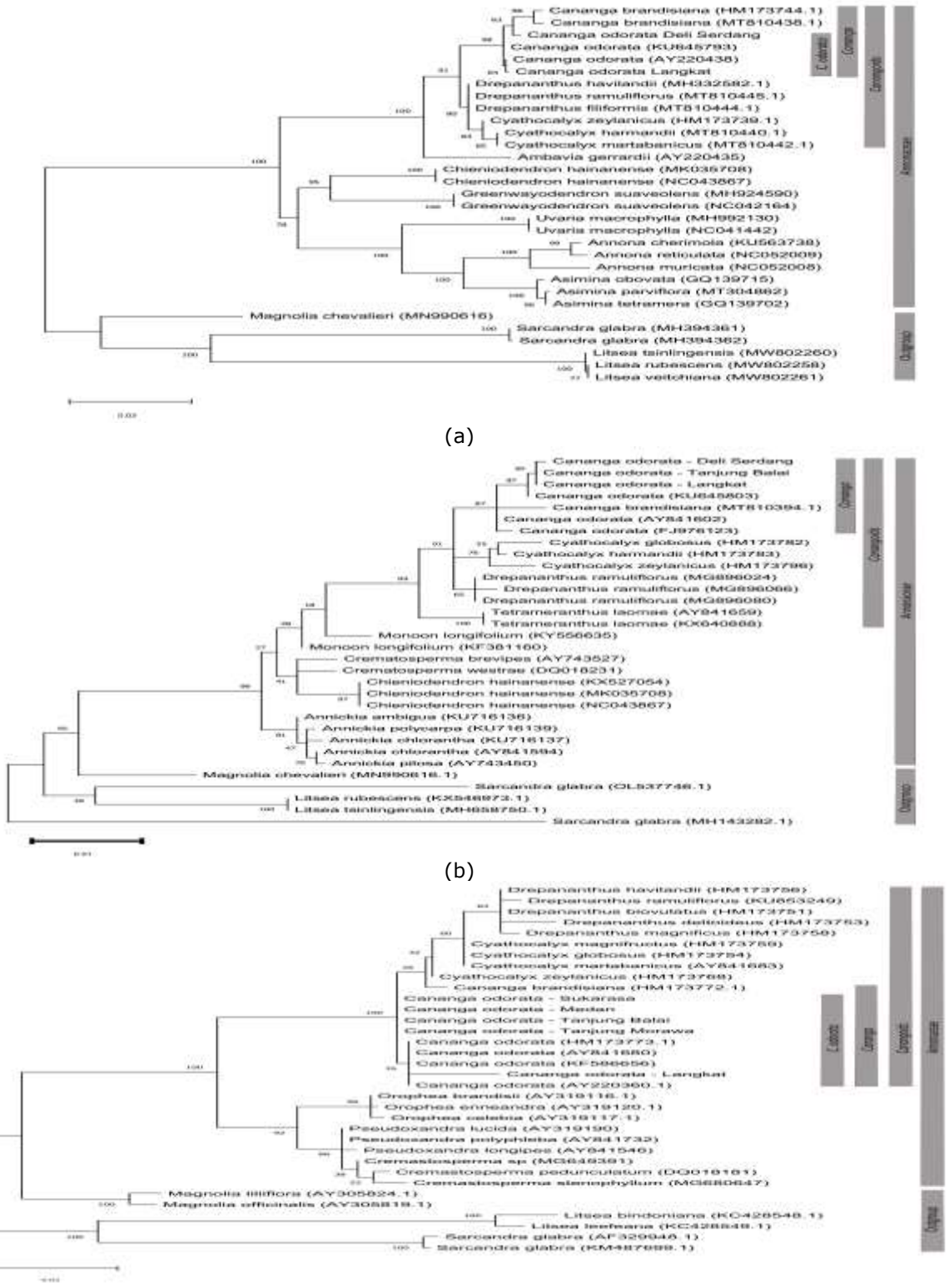


(B)



(C)

Figure 3. Relative distribution of interspecific and intraspecific variation using *mat*K-*trn*K (A), *rbc*L (B), and *trn*L-F (C) barcode.



(a)

(b)

(c)

Figure 4. Maximum-Likelihood tree for *Cananga odorata*, closely related genera, and outgroup (genus *Litsea, Magnolia,* and *Sarcandra*) using *mat*K-*trn*K (a), *rbc*L (b), and *trn*L-F (c) barcode. Reconstruction of Maximum-Likelihood methods was based on the Kimura-2-parameter model. Branching was analyzed with a bootstrap value of 1000×.

interspecific variation was also quite high in all three markers, with most reaching values greater than 0.02. The frequency of the inter-specific distribution was higher in the overall sample analysis, and there was still an overlap with the intra-specific distance even though the distribution was small. The three barcode regions, namely *mat*K-*trn*K (Figure 3A), *rbc*L (Figure 3B), and *trn*L-F (Figure 3C) did not show a clearly observable barcoding gap. However, the high intraspecies distribution at a low percentage of variable sites indicates that the three markers are suitable in DNA barcodes in identifying *C. odorata* species and species in the Annonaceae family. In addition, although the presence of barcoding gaps is common in animals and has resulted in a high level of species discrimination, this is not always the case in plants (Jaén-Molina *et al.*, 2015; Fazekas *et al.*, 2009).

In DNA barcode studies of certain plant taxa, the distribution of intraspecific and interspecific distances were relatively separate (Fu *et al.*, 2011; de Vere *et al.*, 2012; Armenise *et al.*, 2012; Zhang *et al.*, 2013; Ghahramanzadeh *et al.*, 2013; Leavitt *et al.*, 2014). But also, in some studies, barcoding gaps do not result in species discrimination (Yang *et al.*, 2012; Ashfaq *et al.*, 2013). Generally, the higher the overlap between intraspecific and interspecific distances, the less ability to distinguish species (Fazekas *et al.*, 2009).

The ideal DNA barcode determination may result from the presence of a barcoding gap that arises because of the natural gap between the highest intraspecific value and the lowest interspecific value (Meyer and Paulay, 2005). This gap is also known as the global barcoding gap (Chapple and Ritchie, 2013). This gap presence indicates that there is a limit to variation that occurs in all species in the genus so that the species threshold can be adjusted (Rajaram *et al.*, 2019). The high level of overlap at this threshold indicates insufficient variation by barcode region (Rajaram *et al.*, 2019). In addition, the geographical distribution of the sample is also an important part of the analysis of plant DNA barcodes (Kim *et al.*, 2014).

Databases related to DNA barcoding in *C. odorata* plants in the BOLD System using *mat*K markers have only one sample with a homology level of 99.85%, while using *rbc*L markers have eight samples with homology rates ranging from 99.64%–100.00%. BOLD System does not support DNA barcoding using *trn*L-F markers, so no analysis can proceed further. The species discriminatory ability of the three barcoding markers analysis used a phylogenetic tree with the Maximum Likelihood method (Figure 4). The three markers used in this study showed a high level of success in discrimination at the species level, except for the *trn*L-F marker, which was able to classify the genus *Cananga* but could not separate groups of *C. odorata* species. All markers were able to group the Annonaceae family in the same group and separated from the outgroup (*Magnolia*, *Sarcandra*, *Litsea*). The species *C. odorata* formed monophyly on the phylogenetic tree based on the markers *mat*K-*trn*K, *rbc*L, and *trn*L-F. The monophyletic group shows that there is a very close relationship and is assumed to carry the same genetic pattern (Assis and Rieppel, 2011).

In the phylogenetic analysis of the three markers used, *C. odorata* always grouped with the genera *Drepananthus* and *Cyathocalyx*. It also confirms that the genera *Cananga, Drepananthus,* and *Cyathocalyx* are more closely related and clustered in the clade *'canangoid'*. (Surveswaran *et al.*, 2010). Relationships within the *'canangoid'* clade have been controversial since *Cyathocalyx's* discovery as a sister group of *Drepananthus* (Guo *et al.*, 2017), with *Cananga* and *Drepananthus* included in the sister genera (Xue *et al.*, 2018). *Cananga, Drepananthus,* and *Cyathocalyx* also clustered on the parsimony tree based on the psbA-trnH spacer, *trn*L-F, *rbc*L, and *mat*K sequences (Surveswaran *et al.*, 2010). In addition, phylogenetic analysis of the Ambavioideae subfamily (Annonaceae) based on eight plastid DNA regions (*mat*K, ndhF, and *rbc*L exons; *trn*L intron; atpB-*rbc*L, psbA-trnH, *trn*L-*trn*F, and trnS trnG intergenic spacers) grouped *Cananthus*, *Cyathocalyx*, and *Lettowianthus* into the same group based on the maximum likelihood method phylogenetic tree analysis (Chaowasku, 2020). Based on the analysis results, the markers *rbc*L, *mat*K-*trn*K, and *trn*L-F have the potential as DNA barcodes in *C. odorata* plants native to North Sumatra, Indonesia. This research can provide information about DNA barcoding use, especially in the *C. odorata* species, and the Annonaceae family, as a whole.

CONCLUSIONS

Based on the results of the study using DNA barcoding in *C. odorata*, it has an intraspecific homology level of 99.8% (*rbc*L), 99.4% (*mat*K-*trn*K), and 99.3% (*trn*L-F). Conserved site levels of barcode markers in *C. odorata*

ranged, namely, *rbc*L (99.02%), *mat*K-*trn*K (98.95%), and *trn*L-F (47.70%). The barcoding gap analysis shows the absence of barcoding gaps, although, in the phylogenetic analysis, the third barcode DNA marker can group species and genera in the Annonaceae family. DNA barcodes *mat*K, *rbc*L, and *trn*L-F efficiently identified *C. odorata* species and species in other Annonaceae families.

ACKNOWLEDGMENTS

This study received support from the Penelitian Dasar scheme of the Research Institutions and Community Service, Universitas Negeri Medan, in 2022.

REFERENCES

Ali MA, Gyulai G, Hidvégi N, Kerti B, Al Hemaid FMA, Pandey AK, Lee J (2014). The changing epitome of species identification - DNA barcoding. *Saudi J. Biol. Sci.* 21(3): 204-231.

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.

Armenise L, Simeone MC, Piredda R, Schirone B (2012). Validation of DNA barcoding as an efficient tool for taxon identification and detection of species diversity in Italian conifers. *Eur. J. For. Res.* 131(5): 1337-1353.

Ashfaq M, Asif M, Anjum ZI, Zafar Y (2013). Evaluating the capacity of plant DNA barcodes to discriminate species of cotton (Gossypium: Malvaceae). *Mol. Ecol. Resour.* 13(4): 573-582.

Assis LCS, Rieppel O (2011). Are monophyly and synapomorphy the same or different? Revisiting the role of morphology in phylogenetics. *Cladistics* 27(1): 94-102.

Bousquet J, Strauss SH, Doerksen AH, Price RA (1992). Extensive variation in evolutionary rate of rbcL gene sequences among seed plants. *Proc. Natl. Acad. Sci.* 89(16): 7844-7848.

Chaowasku T (2020). Toward a phylogenetic reclassification of the subfamily Ambavioideae (Annonaceae): Establishment of a new subfamily and a new tribe. *Acta Bot. Brasilica* 34(3): 522-529.

Chapple DG, Ritchie PA (2013). A retrospective approach to testing the DNA barcoding method. *PLoS One* 8(11): e77882.

Cheng YC, Houston R (2021). Evaluation of the *trn*K-*mat*K-*trn*K, *ycf3*, and *acc*D-*psal* chloroplast regions to differentiate crop type and biogeographical origin of *Cannabis sativa*. *Int. J. Legal Med.* 135(4): 1235-1244.

Chuang S, Hu J (2004). The evolution of chloroplast *mat*K genes, including identification of new homologues from *Ophioglossum petiolatum* and two Lycophytes. 49(1): 273-287.

Costion C, Ford A, Cross H, Crayn D, Harrington M, Lowe A (2011). Plant DNA barcodes can accurately estimate species richness in poorly known floras. *PLoS One* 6(11): e26841.

Couvreur TLP, Richardson JE, Sosef MSM, Erkens RHJ, Chatrou LW (2008). Evolution of syncarpy and other morphological characters in African Annonaceae: A posterior mapping approach. *Mol. Phylogenet. Evol.* 47(1): 302-318.

De Bontin O (2006). Ylang-ylang oil production in Madagascar and the Comoros. *Proceedings of IFEAT Conference on "The Industry in Sub-Saharan Africa and the Indian Ocean Islands."* Cape Town South Africa, pp.123-135.

De Vere N, Rich TCG, Ford CR, Trinder SA, Long C, Moore CW, Satterthwaite D, Davies H, Allainguillaume J, Ronca S, Tatarinova T, Garbett H, Walker K, Wilkinson MJ (2012). DNA barcoding the native flowering plants and conifers of Wales. *PLoS One* 7(6): e37945.

Doyle JA, Sauquet H, Scharaschkin T, Le-Thomas A (2004). Phylogeny, molecular and fossil dating, and biogeographic history of Annonaceae and Myristicaceae (Magnoliales). *Int. J. Plant Sci.* 165(S4): S55-S67.

Fazekas AJ, Kesanakurti PR, Burgess KS, Percy DM, Graham SW, Barrett SCH, Newmaster SG, Hajibabaei M, Husband BC (2009). Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Mol. Ecol. Resour.* 9: 130-139.

Fitmawati, Hayati I, Mahatma R, Suzanti F (2018). Phylogenetic study of Mangifera from Sumatra, Indonesia using nuclear and chloroplast DNA sequences. *SABRAO J. Breed. Genet.* 50(3): 295-312.

Fu YM, Jiang WM, Fu CX (2011). Identification of species within *Tetrastigma* (Miq.) Planch. (Vitaceae) based on DNA barcoding techniques. *J. Syst. Evol.* 49(3): 237-245.

Ghahramanzadeh R, Esselink G, Kodde LP, Duistermaat H, Valkenburg JLCH, Marashi SH, Smulders MJM, Wiel CCM (2013). Efficient distinction of invasive aquatic plant species from non-invasive related species using DNA barcoding. *Mol. Ecol. Resour.* 13(1): 21-31.

Guo X, Tang CC, Thomas DC, Couvreur TLP, Saunders RMK (2017). A mega-phylogeny of the Annonaceae: Taxonomic placement of five enigmatic genera and support for a new tribe, Phoenicantheae. *Sci. Rep.* 7(1): 7323.

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

Hilu KW, Liang G (1997). The matK gene: Sequence variation and application in plant systematics. *Am. J. Bot.* 84(6): 830-839.

Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS, Erickson DL, Fazekas AJ, Graham SW, James KE, Kim KJ, Kress WJ, Schneider H, van AlphenStahl J, Barrett SCH, van-den-Berg C, Bogarin D, Burgess KS, Cameron KM, Carine M, Chacón J, Clark A, Clarkson JJ, Conrad F, Devey DS, Ford CS, Hedderson TAJ, Hollingsworth ML, Husband BC, Kelly LJ, Kesanakurti PR, Kim JS, Kim Y-D, Lahaye R, Lee H-L, Long DG, Madriñán S, Maurin O, Meusnier I, Newmaster SG, Park CW, Percy DM, Petersen G, Richardson JE, Salazar GA, Savolainen V, Seberg O, Wilkinson MJ, Yi D-K, Little DP (2009). A DNA barcode for land plants. *Proc. Natl. Acad. Sci.* 106(31): 12794-12797.

Jaén-Molina R, Marrero-Rodríguez Á, Reyes-Betancort JA, Santos-Guerra A, Naranjo-Suárez J, Caujapé-Castells J (2015). Molecular taxonomic identification in the absence of a 'barcoding gap': A test with the endemic flora of the Canarian oceanic hotspot. *Mol. Ecol. Resour.* 15(1): 42-56.

Johnson LA, Soltis DE (1995). Phylogenetic inference in Saxifragaceae sensu stricto and Gilia (Polemoniaceae) using *mat*K sequences. *Ann. Missouri Bot. Gard.* 82(2): 149.

Kim HM, Oh SH, Bhandari GS, Kim CS, Park CW (2014). DNA barcoding of Orchidaceae in Korea. *Mol. Ecol. Resour.* 14(3): 499-507.

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16(2): 111-120.

Leavitt SD, Esslinger TL, Hansen ES, Divakar PK, Crespo A, Loomis BF, Lumbsch HT (2014). DNA barcoding of brown Parmeliae (Parmeliaceae) species: A molecular approach for accurate specimen identification, emphasizing species in Greenland. *Org. Divers. Evol.* 14(1): 11-20.

Lestari DA, Azrianingsih R (2019). Short Communication: Species identity and taxonomical position of selected species of Annonaceae based on *trn*L molecular marker. *Biodiversitas J. Biol. Divers.* 20(4): 1012-1019.

Li PS, Thomas DC, Saunders RMK (2015). Phylogenetic reconstruction, morphological diversification and generic delimitation of *Disepalum* (Annonaceae). *PLoS One* 10(12): e0143481.

Mallavarapu GR, Gurudutt KN, Syamasundar KV (2016). Ylang-ylang (*Cananga odorata*) oils. V.R. Preedy (ed.), *Essential oils in food preservation, flavor and safety*. London: Elsevier, pp. 865-873.

Manner HI, Elevitz CR (2006). *Cananga odorata* (ylang ylang). *Species Profiles for Pacific Islands Agroforestry*. Honolulu, Hawai: Permanent Agriculture Resources, pp. 1-10.

Meyer CP, Paulay G (2005). DNA Barcoding: Error rates based on comprehensive sampling. *PLoS Biol.* 3(12): e422.

Mols JB, Gravendeel B, Chatrou LW, Pirie MD, Bygrave PC, Chase MW, Keßler PJA (2004). Identifying clades in Asian Annonaceae: Monophyletic genera in the polyphyletic Miliuseae. *Am. J. Bot.* 91(4): 590-600.

Mols JB, Keßler PJA, Rogstad SH, Saunders RMK (2008). Reassignment of six *Polyalthia* species to the new genus *Maasia* (Annonaceae): Molecular and morphological congruence. *Syst. Bot.* 33(3): 490-494.

Pirie MD, Vargas MPB, Botermans M, Bakker FT, Chatrou LW (2007). Ancient paralogy in the cpDNA *trn*L-F region in Annonaceae: Implications for plant molecular systematics. *Am. J. Bot.* 94(6): 1003-1016.

Rajaram MC, Yong CSY, Gansau JA, Go R (2019). DNA barcoding of endangered *Paphiopedilum* species (Orchidaceae) of Peninsular Malaysia. *Phytotaxa* 387(2): 94-104.

Ratnasingham S, Hebert PDN (2007). BARCODING: bold: The barcode of life data system (http://www.barcodinglife.org). *Mol. Ecol. Notes* 7(3): 355-364.

Richardson JE, Chatrou LW, Mols JB, Erkens RHJ, Pirie MD (2004). Historical biogeography of two cosmopolitan families of flowering plants: Annonaceae and Rhamnaceae. *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.* 359(1450): 1495-1508.

Romiguier J, Roux C (2017). Analytical biases associated with GC-content in molecular evolution. *Front. Genet.* 8:16.

Saunders RMK (2009). Floral evolution in the Annonaceae: Hypotheses of homeotic mutations and functional convergence. *Biol. Rev.* 85: 571-591.

Saunders RMK, Su YCF, Xue B (2011). Phylogenetic affinities of *Polyalthia* species (Annonaceae) with columellar-sulcate pollen: Enlarging the Madagascan endemic genus *Fenerivia*. *Taxon* 60(5): 1407-1416.

Surveswaran S, Wang RJ, Su YCF, Saunders RMK (2010). Generic delimitation and historical biogeography in the early-divergent "ambavioid" lineage of Annonaceae: *Cananga*, *Cyathocalyx* and *Drepananthus*. *Taxon* 59(6): 1721-1734.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E (2007). Power and limitations of the chloroplast *trn*L (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 35(3): e14-e14.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* 17(5): 1105-1109.

Tamura K, Stecher G, Kumar S (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* 38(7): 3022-3027.

Van-Velzen R, Weitschek E, Felici G, Bakker FT (2012). DNA barcoding of recently diverged species: Relative performance of matching methods. *PLoS One* 7(1): e30490.

Wu Q, Liu X, Chen Y, Zhang D, Dai S (2019). The complete chloroplast genome of *Uvaria macrophylla* Roxb. (Annonaceae). *Mitochondrial DNA Part B* 4(2): 3012-3013.

Xue B, Tan Y-H, Thomas DC, Chaowasku T, Hou X-L, Saunders RMK (2018). A new Annonaceae genus, *Wuodendron*, provides support for a post-boreotropical origin of the Asian-Neotropical disjunction in the tribe Miliuseae. *Taxon* 67(2): 250-266.

Xue B, Thomas DC, Chaowasku T, Johnson DM, Saunders RMK (2014). Molecular phylogenetic support for the taxonomic merger of *Fitzalania* and *Meiogyne* (Annonaceae): New nomenclatural combinations under the conserved name *Meiogyne*. *Syst. Bot.* 39(2): 396-404.

Yang J, Wang Y, Möller M, Gao L, Wu D (2012). Applying plant DNA barcodes to identify species of *Parnassia* (Parnassiaceae). *Mol. Ecol. Resour.* 12(2): 267-275.

Zhang C, Wang F, Yan H, Hao G, Hu C, Ge X (2012). Testing DNA barcoding in closely related groups of *Lysimachia* L. (Myrsinaceae). *Mol. Ecol. Resour.* 12(1): 98-108.

Zhang W, Fan X, Zhu S, Zhao H, Fu L (2013). Species-specific identification from incomplete sampling: Applying DNA barcodes to monitoring invasive Solanum plants. *PLoS One* 8(2): e55927.

Zhou L, Su YCF, Saunders RMK (2009). Molecular phylogenetic support for a broader delimitation of *Uvaria* (Annonaceae), inclusive of *Anomianthus*, *Cyathostemma*, *Ellipeia*, *Ellipeiopsis* and *Rauwenhoffia*. *Syst. Biodivers.* 7(3): 249-258.