RESEARCH ARTICLE

SABRAO Journal of Breeding and Genetics
55 (1) 97-106, 2023
http://doi.org/10.54910/sabrao2023.55.1.9
http://sabraojournal.org/
pISSN 1029-7073; eISSN 2224-8978

THE ABILITY OF *matK* AND *trnL-trnL-trnF INTERGENIC SPACER* TO DISCERN CERTAIN SPECIES ACCESSIONS OF THE FAMILIES SOLANACEAE AND FABACEAE

HERMAN[1*], H. AL-KHAIRI[1], A.R. WANSYAH[1], ASMANIA[1], R.I. UTAMI[1], D.N. ANZAELINA[1], Z. OKTAVIANO[1], W. LESTARI[1], ADIWIRMAN[2], and D.I. ROSLIM[1]

[1]Department of Biology, Faculty of Mathematics and Natural Sciences, University of Riau, Riau, Indonesia
[2]Department of Agrotechnology, Faculty of Agriculture, University of Riau, Indonesia
*Corresponding author's emails: herman@lecturer.unri.ac.id, hermansyahdan@ymail.com
Email addresses of co-authors: hapiz12al@gmail.com, aldyriauwansyah08@gmail.com, asmania5156@student.unri.ac.id, rahma.indah0260@student.unri.ac.id, destrya.nadya0219@student.unri.ac.id, zahara.oktaviano4182@student.unri.ac.id, wahyu.lestari@lecturer.unri.ac.id, adiwirman@gmail.com, dewiindriyaniroslim@gmail.com

SUMMARY

Indonesia, as a mega-biodiversity country, has many kinds of tropical fruits. Solanaceae and Fabaceae families contain plant species that produce fruits and vegetables for human food, such as 'rimbang' (*Solanum torvum*), 'terung asam' (*S. ferox*), 'kabau' (*Archidendron microcarpum*), 'jengkol' (*A. pauciflorum*), and 'petai' (*Parkia speciosa*). This study aimed to determine the ability of DNA barcodes (*matK* and *trnL-trnL-trnF intergenic spacer* [*IGS*]) in distinguishing some of those tropical fruit plants. Plant materials came from Riau Province, Indonesia. Methods developed include total DNA extraction from fresh leaves and PCR using the primer pairs to amplify *matK* and *trnL-trnL-trnF IGS*. The PCR products underwent sequencing, with the DNA sequences analyzed using MEGA6. In this study, the obtained *matK* and *trnL-trnL-trnF IGS* sequences showed lengths ranging from 800 bp to 892 bp and 872 bp to 1113 bp, respectively. The *matK* and *trnL-trnL-trnF IGS* sequences of *S. ferox*, *A. microcarpum*, and *A. pauciflorum* provided the first sequences reported in those species. Several nucleotide variations in *matK* sequences emerged among the studied accessions caused by substitution mutation, including transversion and transition. No indel (insertion-deletion) mutation in the studied *matK* sequences occurred. Indel can also cause variations in *trnL-trnL-trnF IGS* sequences among the studied accessions aside from substitution mutation. For species *S. torvum*, *S. ferox,* and *P. speciosa*, there came about some decisive nucleotides based on the *matK* sequences. Moreover, for species *S. torvum*, *S. ferox*, *A. microcarpum*, *A. pauciflorum,* and *P. speciose,* there occurred few critical nucleotides based on *trnL-trnL-trnF IGS*. Also, there was a critical nucleotide for large-fruited 'jengkol' (one of the populations in *A. pauciflorum*) based on *trnL-trnL-trnF IGS*. However, both DNA barcodes could not differentiate the two types of *S. ferox,* i.e., 'terung asam' and 'terung bulu.' Both DNA barcodes showed to differentiate some species members of both families Solanaceae and Fabaceae into separate clusters. Moreover, compared with *matK*, the *trnL-trnL-trnF IGS* indicated uniqueness in discriminating populations within species.

Keywords: Solanaceae, Fabaceae, species, population types, DNA barcodes, *matK, trnL-trnL-trnF intergenic spacer*

Key findings: DNA barcode sequences of *matK* and *trnL-trnL-trnF IGS* revealed the high variations caused by mutations, namely, insertion and deletion in *trnL-trnL-trnF IGS* and substitution in *matK*. These mutations could distinguish the species into separate groups belonging to the families Solanaceae and Fabaceae. DNA barcode *trnL-trnL-trnF IGS* revealed more promising for discriminating populations within species.

Citation: Herman, Al-Khairi H, Wansyah AR, Asmania, Utami RI, Anzaelina DN, Oktaviano Z, Lestari W, Adiwirman, Roslim DI (2023). The ability of *matK* and *trnL-trnL-trnF intergenic spacer* to discern certain species accessions of the families Solanaceae and Fabaceae. *SABRAO J. Breed. Genet.* 55(1): 97-106. http://doi.org/10.54910/sabrao2023.55.1.9.

Communicating Editor: Dr. Himmah Rustiami

Manuscript received: July 10, 2022; Accepted: December 29, 2022.

INTRODUCTION

Indonesia is a mega-biodiversity country that has many tropical fruits. Some types of the Solanaceae and Fabaceae families are valuable for the Indonesian community. The Solanaceae family consists of 90 genera and 2600 species. This family usually refers to a group of eggplants ('terung-terungan') with unique characteristics, such as, bushes, shrubs, and trees having single or compound leaves, curved leaf edges, trumpet-shaped flowers, and box fruit (single real dry fruit having many seeds) or buni fruit (a fruit containing a thin outer layer and a thick soft and juicy inner layer). Some species within this family commonly served as main food ingredients, such as potato (*Solanum tuberosum* L.); vegetables, such as tomato (*S. lycopersicum* L.), eggplant (*S. melongena* L.), 'ranti' or 'leunca' (*S. nigrum* L.), 'rimbang' or 'takokak' (*S. torvum* L.), 'terung asam' (*S. ferox* L.), red chili (*Capsicum annuum* L.), cayenne pepper (*C. frutescens* L., *C. baccatum* L.); industrial raw material, such as tobacco (*Nicotiana tabacum* L. and *N. rustica* L.); and weeds, such as 'kecubung' (*Datura metel* L.) (Barboza *et al*., 2016).

Meanwhile, the Fabaceae family classifies into three subfamilies, i.e., Faboideae (or Papilionoideae, a butterfly flowering-like plant), Caesalpinioideae, and Mimosoideae. The subfamily of Mimosoideae is a group of 'petai-petaian' consisting of 3,270 species (LPWG, 2013) of distinctive characteristics, with the flowers arranged on a flower base (head) in the form of a sphere causing the flowers to look like a hairy ball. Some species in this family are 'jengkol' (*Archidendron pauciflorum*), 'kabau' (*A. microcarpum*), 'saga pohon' (*Adenanthera pavonina*), 'jeungjing' (*Paraserianthes falcataria*), 'lamtoro' (*Leucaena glauca*), 'putri malu' (*Mimosa pudica*), 'petai' (*Parkia speciosa*), and 'trembesi' or '*ki hujan*' (*Albizia saman*) (Lim, 2012; Harvey-Brown, 2019).

Furthermore, the species *S. ferox* has several populations, and two of these are very important, i.e., one produces large and non-hairy fruits called 'terung asam,' and another produces smaller and thick hairy fruits known as 'terung bulu.' Similarly, several populations also appear within species *A. pauciflorum*, having a population that produces large fruits named large-fruited 'jengkol' and another population that yields small fruits called small-fruited 'jengkol' (Figure 1). This study will carry out DNA barcode analysis to determine whether different populations based on the characteristics of the fruit are also genetically distinct.

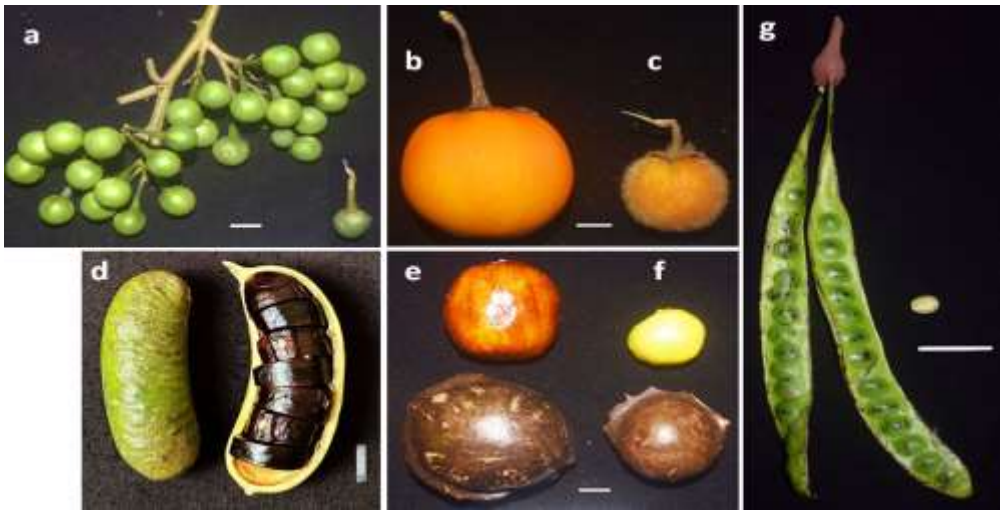

Figure 1. Fruits studied. a) 'rimbang,' b) 'terung asam,' c) 'terung bulu,' d) 'kabau,' e) large fruit 'jengkol', f) small fruit 'jengkol,' and g) 'petai.' Bar = 1 cm and 5 cm in petai capture.

DNA barcoding is a technique developed by Hebert *et al*. (2003) for identifying organisms using a short piece of DNA (DNA barcode). The difference between population and organism could be determined using this technique (Roslim *et al*., 2021). The DNA barcodes applicable in plants include coding regions like *rbcL* and *matK* and noncoding regions. The latter covers the intron of a gene, like *rps16* intron, *rpl16* intron, *trnL* intron, and the space between the two genes (called an intergenic spacer, *IGS*), like *trnL-trnF IGS*, *ITS*, and *ndhC-trnV IGS* (Shaw *et al*., 2007). The *matK* and *trnL-trnF IGS* are two DNA barcodes having the highest variation and can distinguish plants into species and sub-species levels (Sarkinen *et al*., 2013; Buerki *et al*., 2016).

The average *matK* full length in angiosperms is approximately 1500 bp (Kar *et al*., 2015). The *matK* is a gene-coding maturase protein that functions to splice pre-mRNA (Mustafa *et al*., 2018) and has a higher mutation rate than the *rbcL* gene, therefore often used in plant DNA barcoding (Roslim, 2017). Some examples include DNA barcoding analysis in *Metroxylon sagu* and other palm species (Abbas *et al*., 2020), rice (Anggraini *et al*., 2020), dipterocarps from Sumatra (Moura *et al*., 2019), and Cucurbitaceae (Ho and Nguyen, 2020).

In contrast to *matK*, *trnL-trnL-trnF IGS* is a non-coding region covering the intron of the *trnL* gene and spacer between *trnL* and *trnF* genes (intergenic spacer, *IGS*). For DNA barcoding analysis, it is better to use both non-coding regions because of their highest mutation rates (Taberlet *et al*., 1991; Shaw *et al*., 2007).

Unfortunately, DNA barcode sequences of several species members of families Solanaceae (such as *S. ferox*) and Fabaceae (like *A. microcarpum* and *A. pauciflorum*) are unavailable in the public database. This situation potentially complicates and aborts the identification using DNA barcoding since the DNA barcoding analysis needs the availability of abundant DNA sequences in the public database (Roslim, 2018). Therefore, this study aims to analyze the DNA barcodes *matK* and *trnL-trnL-trnF IGS* in several species of Solanaceae and Fabaceae. This study carries out the objectives, i.e., a) to determine the ability of both DNA barcodes to discriminate populations and species within families Solanaceae and Fabaceae, and b) to build their DNA barcode sequences database for supporting DNA barcoding analysis.

MATERIALS AND METHODS

Material and procedure

The study used the plant families Solanaceae and Fabaceae (Table 1). The primer pair used to amplify *matK* was *matK*-413f-1: 5'-TAA TTT ACR ATC AAT TCA TTC AAT ATT TCC-3' and *matK*-1227r-3: 5'-GAR GAT CCR CTR TRA TAA TGA AAA AGA TTT-3' (Heckenhauer *et al.*, 2016). However, for amplifying the *trnL-trnL-trnF IGS*, the primer pair was B49317_F2: 5'-CGA AAT CGG TAG ACG CTA CG-3' and A50272_R3: 5'-ATT TGA ACT GGT GAC ACG AG-3' (Taberlet *et al*., 1991).

Total DNA extraction used the Genomic DNA Mini Kit Plant (Geneaid) by adopting Roslim *et al*. (2021). Then, amplifying total DNA solutions used the primer pairs of *matK* and *trnL-trnL-trnF IGS*. The PCR components adopted Roslim (2017) method, with modification of using 0.2 mM dNTPs. The PCR proceeded according to manufacturer Thermo Scientific, with the following conditions: pre-PCR at 95 ºC for 3 min, followed by 35 cycles consisting of three steps, such as denaturation at 95 ºC for 45 s, annealing at 47.5 ºC for *matK*, and 49.2 ºC for *trnL-trnL-trnF IGS* for 45 s, and elongation at 72 ºC for 90 s. Then, pasca-PCR continued at 72 ºC for 10 min.

The use of electrophoresis, according to Roslim (2017), checked the amplicons. The study sent 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 for sequencing in 1st Base Malaysia.

Table 1. Family, sample names, local names, sample codes, and district of origins.

No.	Family	Sample Names	Local Names	District Origins
1	Solanaceae	*Solanum torvum*	Rimbang	Pekanbaru
2		*Solanum ferox*	Terung asam	Kampar
3		*Solanum ferox*	Terung bulu	Rokan Hulu
4	Fabaceae	*Archidendron microcarpum*	Kabau	Pelalawan
5		*Archidendron pauciflorum*	Large fruit jengkol	Kampar
6		*Archidendron pauciflorum*	Small fruit jengkol	Kampar
7		*Parkia speciosa*	Petai	Kampar

Data analysis

The alignment of DNA sequences, sequenced with the forward and reverse primers, used BioEdit version 7.0.0 software to obtain the correct sequences (Hall, 1999). The analysis of correct DNA sequences used MEGA version 6.06 software to determine the nucleotide differences and construction of dendrograms (Tamura *et al*., 2013). Outgroups used in this study were *Adenanthera pavonina* (GU135053.1 for *matK* and AF278486.1 for *trnL-trnL-trnF IGS*) and *Nicotiana tabacum* (KJ652184.1 for *matK* and AP019625.1 for *trnL-trnL-trnF IGS*), with their DNA sequences obtained from GenBank.

RESULTS

Length and nucleotide content

The DNA barcode sequences of *matK* and *trnL-trnL-trnF IGS* obtained in this study have already been registered in the GenBank (Table 2). The shortest *matK* sequence showed in 'terung asam' (800 bp) and the longest in 'rimbang' (892 bp). About the *trnL-trnL-trnF IGS* sequence, the shortest occurred in 'jengkols' (872 bp), whereas the longest one also in 'rimbang' (1113 bp). The average *trnL-trnL-trnF IGS* sequences size in the family Solanaceae (1063 bp–1113 bp) was longer than the one in the family Fabaceae (872 bp–1003 bp) (Table 2).

The contents of A+T and G+C in both studied DNA barcodes were relatively similar between the families of Solanaceae and Fabaceae. The *matK* and *trnL-trnL-trnF IGS* sequences were (A+T)-rich, such as 66%–68% and 65%–67%, respectively (Table 3). These results confirmed that the nucleotide content does not differentiate the family Solanaceae from Fabaceae.

Nucleotide variation and phylogenetic tree based on *matK* sequence

Several nucleotide variations in *matK* sequences appeared among the studied accessions that might have caused by substitutional mutation, including transversion and transition. There was no insertion-deletion (indel) mutation in the studied *matK*

Table 2. Size of *matK* and *trnL-trnL-trnF IGS* sequences.

No.	Species and local name	*matK*			*trnL-trnL-trnF IGS*		
		Sample codes	Size (bp)	Registration number	Sample codes	Size (bp)	Registration number
	Solanum torvum						
1	Rimbang 1	DIR065	892	ON101685	DIR092	1113	ON054977
2	Rimbang 2	DIR066	892	ON101686	DIR093	1113	ON054978
3	Rimbang 3	DIR067	892	ON101687	DIR094	1113	ON054979
	Solanum ferox						
4	Terung asam 1	DIR062	800	ON101682	DIR083	1063	ON054968
5	Terung asam 2	DIR063	800	ON101683	DIR084	1063	ON054969
6	Terung asam 3	DIR064	800	ON101684	DIR085	1063	ON054970
7	Terung bulu 1	DIR080	831	ON101700	DIR086	1063	ON054971
8	Terung bulu 2	DIR081	831	ON101701	DIR087	1063	ON054972
9	Terung bulu 3	DIR08	831	ON101702	-	-	-
The average in Solanaceae			841			1082	
	Archidendron microcarpum						
10	Kabau 3	DIR068	865	ON101688	DIR095	873	ON054980
11	Kabau 4	DIR069	865	ON101689	-	-	-
12	Kabau 6	DIR070	865	ON101690	DIR096	873	ON054981
	Archidendron pauciflorum						
13	Large fruit jengkol 1	DIR071	835	ON101691	DIR088	872	ON054973
14	Large fruit jengkol 2	DIR072	835	ON101692	DIR089	872	ON054974
15	Large fruit jengkol 3	DIR073	835	ON101693	-	-	-
16	Small fruit jengkol 1	DIR077	870	ON101697	DIR090	872	ON054975
17	Small fruit jengkol 2	DIR078	870	ON101698	-	-	-
18	Small fruit jengkol 3	DIR079	870	ON101699	DIR091	872	ON054976
	Parkia speciosa						
19	Petai 2	DIR074	865	ON101694	DIR097	1003	ON054982
20	Petai 4	DIR075	865	ON101695	-	-	-
21	Petai 6	DIR076	865	ON101696	DIR098	1000	ON054983
The average in Fabaceae			859			905	

Table 3. The A+T and G+C content (%) of *matK* and *trnL-trnL-trnF intergenic spacer* in the families of Solanaceae and Fabaceae.

No.	Accessions	*matK* (%)		*trnL-trnL-trnF IGS* (%)	
		A+T	G+C	A+T	G+C
1	Rimbang 1	67	33	65	35
2	Rimbang 2	67	33	65	35
3	Rimbang 3	67	33	65	35
4	Terung asam 1	67	33	67	33
5	Terung asam 2	67	33	67	33
6	Terung asam 3	67	33	67	33
7	Terung bulu 1	67	33	67	33
8	Terung bulu 2	67	33	67	33
9	Terung bulu 3	67	33	-	-
10	Kabau 3	68	32	66	34
11	Kabau 4	68	32	-	-
12	Kabau 6	68	32	66	34
13	Large fruit jengkol 1	67	33	66	34
14	Large fruit jengkol 2	67	33	66	34
15	Large fruit jengkol 3	67	33	-	-
16	Small fruit jengkol 1	67	33	66	34
17	Small fruit jengkol 2	67	33	-	-
18	Small fruit jengkol 3	67	33	66	34
19	Petai 2	68	32	66	34
20	Petai 4	68	32	-	-
21	Petai 6	67	33	66	34
22	*Nicotiana tabacum*	66	34	65	35
23	*Adenanthera pavonina*	67	33	65	35
Average		67	33	66	34

sequences. A few vital nucleotides showed in all the studied accessions, except in 'kabau.' For instance, four decisive nucleotides emerged in 'rimbang' and 'terung asam,' three in 'jengkol,' and two in 'petai.' However, no differences occurred between *matK* sequences for species *S. ferox* populations, i.e., 'terung asam' and 'terung bulu.' The same also holds true for *matK* sequences within populations of 'jengkol' species, namely, large-fruited 'jengkol' and small-fruited 'jengkol' (Table 4).

The phylogenetic tree showed that species of the family Solanaceae formed one cluster and separated from the family Fabaceae. Furthermore, every individual in the same species was grouped in the same cluster but separately from the clusters of the other species (Figure 2).

Nucleotide variation and phylogenetic tree based on *trnL- trnL-trnF IGS*

More nucleotide variations in *trnL-trnL-trnF IGS* sequences surfaced among the studied accessions than in *matK* sequences. Compared with *matK*, the variations in *trnL-trnL-trnF IGS* sequences resulted from substitution mutation and indel. Moreover, a few critical nucleotides also existed in all the studied species, such as 11 significant nucleotides in 'rimbang,' 21 in 'terung asam,' two in 'kabau,' three in 'jengkol,' and two decisive nucleotides in 'petai.' For *S. ferox* populations, i.e., 'terung asam' and 'terung bulu,' the *trnL-trnL-trnF IGS* sequences were the same. However, the *trnL-trnL-trnF IGS* sequence of *A. pauciflorum* populations, namely, the large-fruited 'jengkol' population, differed from small-fruited ones, signed by one vital nucleotide for large-fruited 'jengkol' (Table 5).

Similar to the *matK* sequence, the phylogenetic tree based on the *trnL-trnL-trnF IGS* sequence showed that species from the studied families Solanaceae and Fabaceae formed clusters that separated from each other. Every individual in each studied species formed a cluster according to each species that separated from the others. Large-fruited 'jengkol' formed a group that separated from small-fruited 'jengkol' (Figure 3).

DISCUSSION

Intentionally, the development of the DNA barcoding technique aimed to assist in identifying organisms. The process uses a DNA short piece, approximately 500–1500 bp. The advantages of this technique indicate that a) it is expected that anyone who is a non-taxonomist can perform identification of the different organisms and b) identification can

Table 4. Nucleotide variations in *matK* sequences.

Accessions	Nucleotide number (vertically)													
	1	3	3	3	3	3	3	3	4	4	6	7	7	
	9	0	1	3	5	6	7	9	2	7	1	1	5	
	7	4	3	7	7	1	0	4	2	6	0	0	5	
Solanum torvum_rimbang DIR065	C	C	A	G	T	T	T	A	G	A	G	G	C	
Solanum torvum_rimbang DIR066	.	.	.	.	.	.	.	.	.	.	.	.	.	
Solanum torvum_rimbang DIR067	.	.	.	.	.	.	.	.	.	.	.	.	.	
Solanum ferox_terong asam DIR062	A	T	C	A	G	.	.	G	C	.	A	.	.	
Solanum ferox_terong asam DIR063	A	T	C	A	G	.	.	G	C	.	A	.	.	
Solanum ferox_terong asam DIR064	A	T	C	A	G	.	.	G	C	.	A	.	.	
Solanum ferox_terong bulu DIR080	A	T	C	A	G	.	.	G	C	.	A	.	.	
Solanum ferox_terong bulu DIR081	A	T	C	A	G	.	.	G	C	.	A	.	.	
Solanum ferox_terong bulu DIR082	A	T	C	A	G	.	.	G	C	.	A	.	.	
Archidendron microcarpum_kabau DIR068	.	.	.	A	.	.	.	G	C	.	A	.	.	
Archidendron microcarpum_kabau DIR069	.	.	.	A	.	.	.	G	C	.	A	.	.	
Archidendron microcarpum_kabau DIR070	.	.	.	A	.	.	.	G	C	.	A	.	.	
Archidendron pauciflorum_large fruit jengkol DIR071	.	.	.	A	.	G	G	G	C	.	A	T	.	
Archidendron pauciflorum_large fruit jengkol DIR072	.	.	.	A	.	G	G	G	C	.	A	T	.	
Archidendron pauciflorum_large fruit jengkol DIR073	.	.	.	A	.	G	G	G	C	.	A	T	.	
Archidendron pauciflorum_small fruit jengkol DIR077	.	.	.	A	.	G	G	G	C	.	A	T	.	
Archidendron pauciflorum_small fruit jengkol DIR078	.	.	.	A	.	G	G	G	C	.	A	T	.	
Archidendron pauciflorum_small fruit jengkol DIR079	.	.	.	A	.	G	G	G	C	.	A	T	.	
Parkia speciosa_petai DIR074	.	.	G	A	.	.	.	G	C	C	A	.	T	
Parkia speciosa_petai DIR075	.	.	G	A	.	.	.	G	C	C	A	.	T	
Parkia speciosa_petai DIR076	.	.	G	A	.	.	.	G	C	C	A	.	T	
Nicotiana tabacum	.	.	.	A	.	.	.	G	C	.	A	.	.	
Adenanthera pavonina	.	.	G	A	.	.	.	G	C	.	A	.	.	

Dots (.) dots indicate that the nucleotide in a particular position was the same as the one of the *S. torvum*_rimbang DIR065 sequence.
Critical nucleotides for *Solanum torvum* (blue), *Solanum ferox* (yellow), *Archidendron pauciflorum* (brown), and *Parkia speciosa* (green).

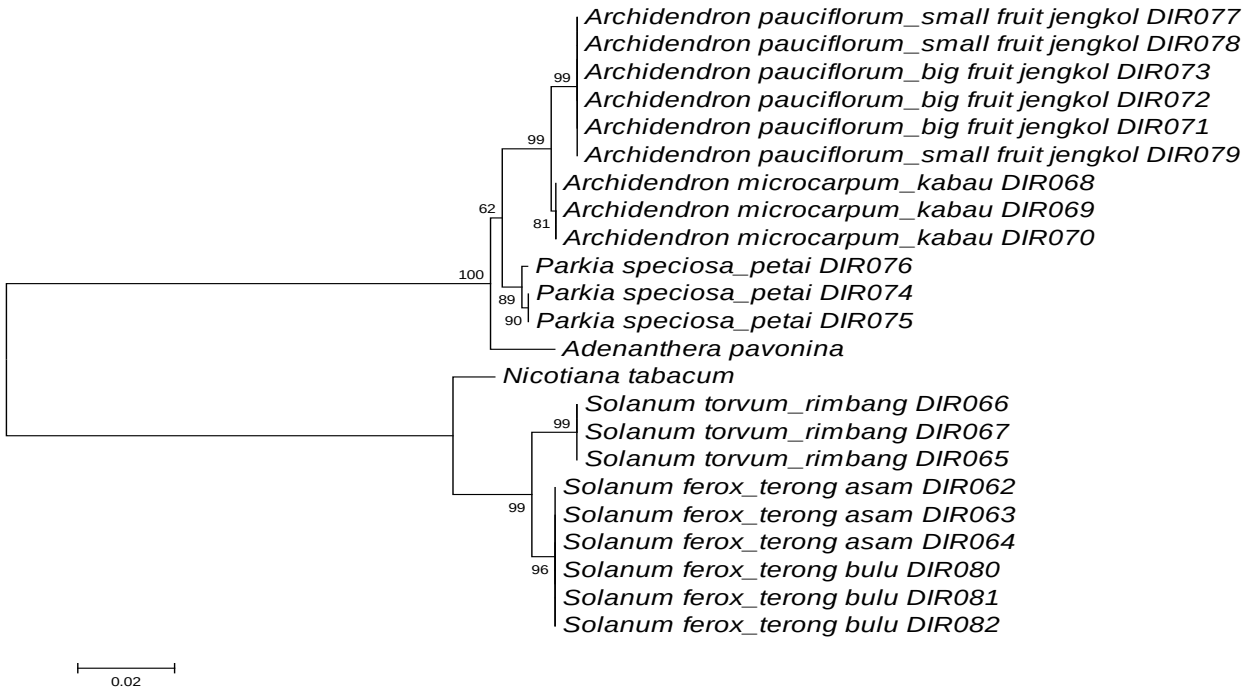

Figure 2. Dendrogram constructed based on *matK* sequences using Neighbor-Joining method with 1000 bootstraps.

Table 5. Nucleotide variations in *trnL-trnL-trnF sequence*.

Accessions	Nucleotide number (vertically)																																				
	113	156	372	492	561	583	648	693	728	734	827	850	891	925	936	997	1001	1003	1018	1024	1027	1029	1030	1035	1042	1052	1060	1061	1062	1063	1064	1068	1071	1072	1073	1077	1078	1080
Solanum torvum_rimbang DIR092	C	-	-	G	G	-	-	C	C	-	T	-	C	G	A	C	A	A	G	-	-	-	-	C	A	C	A	A	A	G	G	C	A	A	T	C	A	G
Solanum torvum_rimbang DIR093	.	-	-	.	.	-	-	.	.	-	.	-	.	.	.	.	.	.	.	-	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Solanum torvum_rimbang DIR094	.	-	-	.	.	-	-	.	.	-	.	-	.	.	.	.	.	.	.	-	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Solanum ferox_terong asam DIR083	G	-	-	.	T	-	T	T	T	T	.	-	T	T	.	.	.	.	T	T	G	T	A	A	.	T	T	T	G	A	C	A	C	C	C	T	C	A
Solanum ferox_terong asam DIR084	G	-	-	.	T	-	T	T	T	T	.	-	T	T	.	.	.	.	T	T	G	T	A	A	.	T	T	T	G	A	C	A	C	C	C	T	C	A
Solanum ferox_terong asam DIR085	G	-	-	.	T	-	T	T	T	T	.	-	T	T	.	.	.	.	T	T	G	T	A	A	.	T	T	T	G	A	C	A	C	C	C	T	C	A
Solanum ferox_terong bulu DIR086	G	-	-	.	T	-	T	T	T	T	.	-	T	T	.	.	.	.	T	T	G	T	A	A	.	T	T	T	G	A	C	A	C	C	C	T	C	A
Solanum ferox_terong bulu DIR087	G	-	-	.	T	-	T	T	T	T	.	-	T	T	.	.	.	.	T	T	G	T	A	A	.	T	T	T	G	A	C	A	C	C	C	T	C	A
Archidendron microcarpum_kabau DIR095	G	A	T	T	T	-	A	-	-	-	.	-	-	A	C	.	-	T	A	A	C	A	G	.	G	A	C	G	.	A	A	.	G	.	.	T	C	.
Archidendron microcarpum_kabau DIR096	G	A	T	T	T	-	A	-	-	-	.	-	-	A	C	.	-	T	A	A	C	A	G	.	G	A	C	G	.	A	A	.	G	.	.	T	C	.
Archidendron pauciflorum_large fruit jengkol DIR088	G	-	-	.	T	-	A	-	-	-	A	-	-	A	T	T	G	C	A	A	C	A	G	.	.	A	C	G	.	A	A	.	G	.	.	T	C	.
Archidendron pauciflorum_large fruit jengkol DIR089	G	-	-	.	T	-	A	-	-	-	A	-	-	A	T	T	G	C	A	A	C	A	G	.	.	A	C	G	.	A	A	.	G	.	.	T	C	.
Archidendron pauciflorum_small fruit jengkol DIR090	G	-	-	.	T	-	A	-	-	-	.	-	-	A	T	T	G	C	A	A	C	A	G	.	.	A	C	G	.	A	A	.	G	.	.	T	C	.
Archidendron pauciflorum_small fruit jengkol DIR091	G	-	-	.	T	-	A	-	-	-	.	-	-	A	T	T	G	C	A	A	C	A	G	.	.	A	C	G	.	A	A	.	G	.	.	T	C	.
Parkia speciosa_petai DIR097	G	-	-	.	T	A	A	G	T	G	.	A	-	A	T	.	-	-	A	A	C	A	G	.	.	A	C	G	.	A	A	.	G	.	.	T	C	.
Parkia speciosa_petai DIR098	G	-	-	.	T	A	A	G	T	G	.	A	-	A	T	.	-	-	A	A	C	A	G	.	.	A	C	G	.	A	A	.	G	.	.	T	C	.
Nicotiana tabacum	G	-	-	.	T	-	-	.	-	-	.	-	.	T	.	.	T	.	A	-	-	-	-	.	.	.	-	-	-	-	-	-	-	-	-	-	-	-
Adenanthera pavonina	G	-	-	.	T	-	A	G	T	G	.	G	-	T	T	.	-	-	A	A	C	A	G	.	.	A	C	G	.	A	T	.	G	.	.	T	C	.

Dots (.) dots indicate that the nucleotide in a particular position was the same as the one of the *S. torvum*_rimbang DIR065 sequence.
Critical nucleotides for *Solanum torvum* (blue), *Solanum ferox* (yellow), *Archidendron microcarpum* (red), *A. pauciflorum* (brown), and *Parkia speciosa* (green).

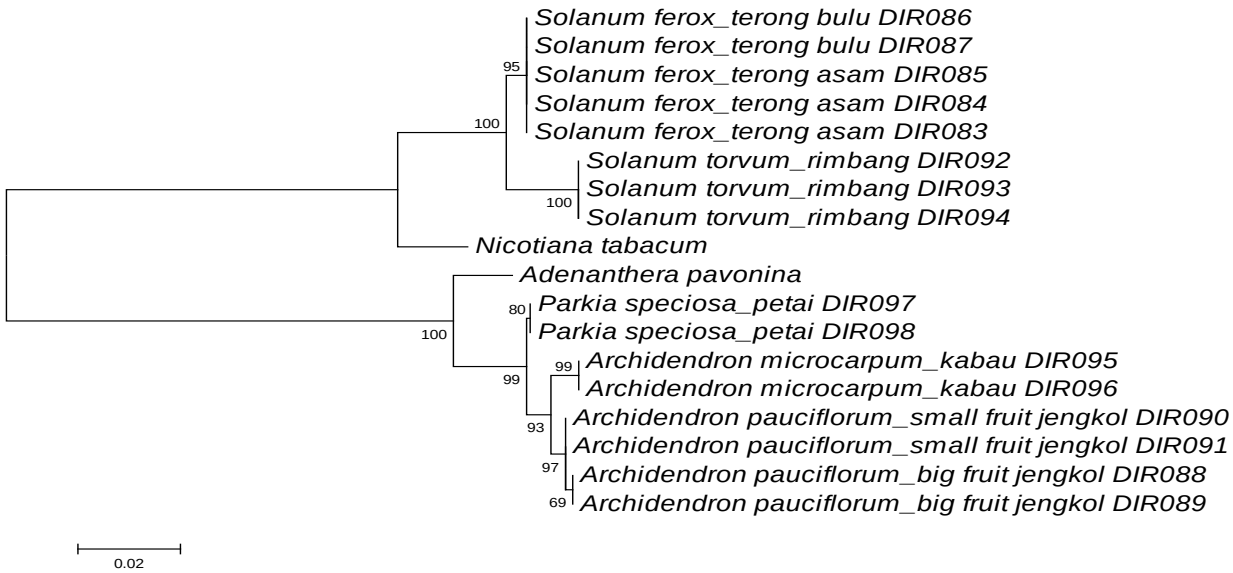

Figure 3. Dendrogram constructed based on *trnL-trnL-trnF intergenic spacer* sequences using the Neighbor-Joining method with 1000 bootstraps.

use specimens with various unfavorable conditions, such as, incomplete, damaged, un- or underdeveloped, fresh, and preserved specimens (Hebert *et al*. 2003). The DNA barcoding analysis authenticated that the levels of divergence among the individuals of the same species appeared lower than in the individuals of different species (Shen *et al*., 2013). Furthermore, identification is essential in various breeding methods, seed production, registration, and trade. The authenticity of bioproducts aims to protect consumers from product substitution and contamination (Balachandran *et al*., 2015; Zhang *et al*., 2021).

The DNA barcoding analysis among the species of families Solanaceae and Fabaceae using *matK* and *trnL-trnF IGS* sequences revealed that the length of *matK* sequences between the species of families in Solanaceae and Fabaceae varied. However, the length cannot determine the differences in both families. In contrast, the length of the *trnL-trnL-trnF IGS* sequences differs between both families (Table 2). Results further showed that the *trnL-trnL-trnF IGS* sequence could better determine the differences between both families. The *trnL-trnL-trnF IGS* sequence size difference in both families might have occurred given that the *trnL-trnL-trnF IGS* sequence has indel mutation, whereas the *matK* sequence does not. According to Sehn (2015), the insertion and deletion in DNA with a length of less than 1 kb caused the indel mutation. Many deletions in one taxon can decrease the DNA size, while many insertions in others can increase the DNA size.

Moreover, the content of A+T and G+C between *matK* and *trnL-trnF IGS* sequences appeared relatively similar (Table 3). The results showed that the average sequence size could determine the difference between the families of Solanaceae and Fabaceae instead of nucleotide content.

The phylogenetic tree analysis showed that both studied sequences could distinguish individuals into their species group separately from other species groups. The trees could also differentiate families of Solanaceae and Fabaceae separately from each other (Figures 2 and 3). Those clusterings resulted from variations of the nucleotide sequence in between species studied. Furthermore, the differences in the studied species got signed by decisive nucleotides (Tables 4 and 5). The critical nucleotide is a nucleotide that constitutes the characteristic of accession in the species. More significantly, the crucial nucleotides found in the *trnL-trnL-trnF IGS* sequence can differentiate up to sub-species levels, i.e., distinguish the *A. pauciflorum* populations into large-fruited 'jengkol' and small-fruited 'jengkol' (Table 5). The phylogenetic tree also showed that large-fruited 'jengkol' forms a cluster separated from small-fruited 'jengkol' (Figure 3). This study validated that the barcoding DNA technique using *matK* and *trnL-trnL-trnF IGS* can better

identify plant species. Moreover, the phylogenetic tree will cluster individuals in the same species into one cluster separate from other individuals from other species.

CONCLUSIONS

The DNA barcodes of *matK* and *trnL-trnL-trnF IGS,* having noncoding regions, revealed the high mutation rates and can distinguish the species and family into separate clusters. The *trnL-trnL-trnF IGS* can discern the studied accessions of families Solanaceae and Fabaceae based on its length (i.e., 1063–1113 bp in the Solanaceae that was longer than in the Fabaceae (872–1003 bp)). Populations in *S. ferox* (such as 'terung asam' and 'terung bulu') are the same genetically, but in *A. pauciflorum* (such as large-fruited 'jengkol' and small-fruited 'jengkol') are different, based on *trnL-trnL-trnF IGS*. The *trnL-trnL-trnF IGS* proved more reliable than *matK* in differentiating the populations within species. However, it needs further investigation using a large population to determine the ability of *trnL-trnL-trnF IGS* to distinguish populations in a species to compare with *matK*.

ACKNOWLEDGMENTS

This research received full sponsorship from a grant from DIPA University of Riau, The Republic of Indonesia 2022, with contract number 1346/un19.5.1.3/PT.01.03/2022.

REFERENCES

Abbas B, Kabes RJ, Mawikere NL, Ruimassa RMR, Aprianto RM (2020). DNA barcode of *Metroxylon sagu* and other palm species using *mat*K gene. *Biodiversitas* 21(9): 4047-4057.

Anggraini NB, Sholihah A, Khasna EN, Retnaningtyas RW, Suharti, Listyorini D (2020). Genetic relationship between local rice varieties based on *matK* and *rbcL* genes. *AIP Conf. Proceed.* 2260: 020020. https://doi.org/10.1063/5.0015756.

Balachandran KRS, Mohanasundaram S, Ramalingam S (2015). DNA barcoding: A genomic-based tool for authentication of phytomedicines and its products. *Botanics: Targets and Therapy* 2015: 5.

Barboza GE, Hunziker AT, Bernardello G, Cocucci AA, Moscone AE, Carrizo C, Garci´A, Fuentes V, Dillon MO, Bittrich V, Cosa MT, Subils R, Romanutti A, Arroyo S, Anton (2016). Solanaceae. In: JW Kadereit and V Bittrich (eds.), Flowering Plants. Eudicots, The Families and Genera of Vascular Plants. Aquifoliales, Boraginales, Bruniales, Dipsacales, Escalloniales, Garryales, Paracryphiales, Solanales (except Convolvulaceae), Icacinaceae, Metteniusaceae, Vahliaceae. Springer International Publishing Switzerland. Vol. 14, pp. 295-357.

Buerki S, Gallaher T, Booth T, Brewer G, Forest F, Pereira JT, Callmander MW (2016). Biogeography and evolution of the screw-pine genus *Benstonea* Callm. & Buerki (Pandanaceae). *Candollea* 71(2): 217-229.

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

Harvey-Brown Y (2019). *Parkia speciosa.* IUCN Red List of Threatened Species*. 2019:* e.T153891869A153917800. doi*:*10.2305/IUCN.UK.2019-3.

Hebert PDN, Cywinska A, Ball SL, de-Waard JR (2003). Biological identifications through DNA barcodes. *Proceed. Royal Soc. London B*, 270: 313-321.

Heckenhauer J, Barfuss MHJ, Samuel R (2016). Universal multiplexable *matK* primers for DNA barcoding of angiosperms. *App. in Plant Sci.* 4(6): 1500137.

Ho VT, Nguyen MP (2020). An in silico approach for evaluation of *rbc*L and *mat*K loci for DNA barcoding of the Cucurbitaceae family. *Biodiversitas* 21(8): 3879-3885.

Kar P, Goyal A, Sen A (2015). Maturase K-gene in Plant DNA Barcoding and Phylogenetics. Lambert Academic Publishing, Saarbrucken, Germany.

Lim TK (2012). *Archidendron jiringa*. Edible Medicinal and Non-Medicinal Plants. Springer Netherlands. pp. 544-548.

LPWG (2013). Towards a new classification system for legume: Progress report from the International Legume Conference. *South Afr. J. Bot.* 89: 3-9.

Moura CCdM, Brambach F, Bado KJH, Krutovsky KV, Kreft H, Tjitrosoedirdjo SS, Siregar IZ, Gailing O (2019). Integrating DNA barcoding and traditional taxonomy for the identification of dipterocarps in remnant lowland forests of Sumatra. *Plants* 8: 461; doi:10.3390/plants8110461.

Mustafa KM, Ewadh MJ, Al-Shuhaib MBS, Hasan HG (2018). The in silico prediction of the chloroplast maturase K gene polymorphism in several barley varieties. *Agric.* (*Pol'nohospodarstvo*) 64(1): 3-16.

Roslim DI (2017). Identification of Pandan plant (*Benstonea* sp) from Riau, Indonesia using three DNA barcodes. *SABRAO J. Breed. Genet.* 49(4): 346-360.

Roslim DI (2018). Pandan (*Pandanus* sp), Rotan (*Calamus* sp), and Rengas (*Gluta* sp) from Kajuik Lake, Riau Province, Indonesia. *Brazilian Arch. Biol. Technol.* 61: e18160419.

Roslim DI, Putra YO, Dewi YM, Bago Y, Sitohang H, Herman, Fitmawati, Sofiyanti N (2021). First record of five DNA barcodes on Nothospecies Cocor Bebek (*Kalanchoe* × *laetivirens*). *SABRAO J. Breed. Genet.* 53(2): 263-377.

Sarkinen T, Bohs L, Olmstead RG, Knapp S (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): A dated 1000-tip tree. *BMC Evol. Biol.* 13: 214.

Shaw J, Lickey EB, Schilling EE, Small RL (2007). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *Am. J. Bot.* 94(3): 275-288.

Sehn JK (2015). Chapter 9: Insertions and Deletions (Indels). in: S Kulkarni and J Pfeifer (eds.), Clinical Genomics. Academic Press. pp. 129-150.

Shen Y-Y, Chen X, Murphy RW (2013). Assessing DNA barcoding as a tool for species identification and data quality control. *PLoS ONE* 8(2): e57125. doi:10.1371/journal.pone.0057125.

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 1105-1109.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729.

Zhang W, Sun Y, Liu J, Xu C, Zou X, Chen X, Liu Y, Wu P, Yang X, Zhou S (2021). DNA barcoding of *Oryza*: Conventional, specific, and super barcodes. *Plant Mol. Biol.* 105: 215-228.