



IDENTIFICATION OF PANDAN PLANT (*Benstonea* sp) FROM RIAU, INDONESIA USING THREE DNA BARCODES

DEWI INDRIYANI ROSLIM¹

¹Genetic Laboratory, Department of Biology, Faculty of Mathematics and Natural Sciences, University of Riau, Binawidya Campus, Jl HR Soebrantas Km 12.5, Panam, Pekanbaru, Riau, Indonesia
*Corresponding author's email: dewiindriyaniroslim@gmail.com

SUMMARY

Pandan from Riau is one of the important plants in Kajuik Lake located in Langgam, Riau Province of Indonesia, although its scientific name has not been recognized. This study reports the use of three DNA barcodes: *matK*, *rbcL*, and *trnL-trnF* intergenic spacer; to determine the Pandan's taxonomic status. The methods included DNA isolation, PCR, electrophoresis, and sequencing. The software BLASTn, BioEdit, and MEGA were used to analyze the data. The *matK*, *rbcL*, and *trnL-trnF* intergenic spacer sequences obtained were 639 bp, 539 bp, and 1014 bp in size, respectively. The results showed that although the identification had already been performed using two standard DNA barcodes sequences for plants, i.e. the *matK* and *rbcL*, and also the *trnL-trnF* intergenic spacer sequence which was commonly used as a DNA barcode in Pandanaceae and abundantly available in GenBank, none of them had 100% similarity to Pandan from Riau. In addition, the dendrogram generated from those sequences showed that Pandan from Riau had the closest relationship with a few species of *Benstonea* rather than *Pandanus*, *Martellidendron*, and *Freycinetia*. It can be concluded that the scientific name of Pandan from Riau can only be determined up to the genus level, i.e. *Benstonea* sp. This suggests that Pandan from Riau is an identified species but its DNA barcode sequences are not present in the GenBank database. It can also be understood that this plant is a new species of *Benstonea* which has never been identified and reported.

Key words: *Benstonea* sp, DNA barcode, Langgam, *matK*, *rbcL*, *trnL-trnF* intergenic spacer.

Key findings: As a flood-adapted plant, Pandan from Riau may provide many submergence tolerance genes from its genome and the genes could be used in crop breeding through genetic engineering.

Manuscript received: May 7, 2017; Decision on manuscript: August 17, 2017; Manuscript accepted: September 22, 2017
© Society for the Advancement of Breeding Research in Asia and Oceania (SABRAO) 2017

Communicating Editor: Naqib Ullah Khan

INTRODUCTION

Pandan from Riau is one of the plants growing in and around Kajuik Lake located in Langgam, Pelalawan Regency, Riau Province, Indonesia. This plant, along with others, has an important role for all animals in Kajuik Lake ecosystem (Elvyra and Yus, 2010; Roslim, 2017; Roslim *et al.*, 2016a, 2016b). Pandan from Riau has been

identified as *Pandanus* sp (a member of Pandanaceae family) based on its morphological character without being furnished by female and male flowers (Elvyra and Yus, 2010). The identification was then resumed in 2016 using DNA barcode, namely internal transcribed spacer (ITS). Unfortunately, the taxonomic status still could not be determined up to the species level due to the limited number of ITS

sequence database of *Pandanus* in GenBank. There is only one datum of ITS available in GenBank updated on April 24th of 2016 (Roslim, 2017).

Prior to the discovery of molecular identification technique, people have identified plants based on their morphological characters and it had to be supported by complete organs such as: stem, leaves, flower, and fruit (Sofiyanti *et al.*, 2016; Harsono *et al.*, 2016; Roslim *et al.*, 2016c). Moreover, if plants are identified from a genus with a lot of morphological similarities, it will be difficult to distinguish them as different species. Another constraint is that the morphological characteristic-based identification requires a specific expertise in the taxonomic field. The DNA barcoding technique using short DNA sequences may be to assist and facilitate the identification of organisms, including plants (Hebert *et al.*, 2003).

One advantage offered by the DNA barcoding technique is that the identification can be performed without the complete organs and anyone could perform it, regardless of their expertise in the field of taxonomy (Jarman and Elliott, 2000; Hebert *et al.*, 2003; Stoeckle, 2003; Ali *et al.*, 2014). In practice, the convenience offered by the identification technique should be accompanied with a complete database of DNA barcodes and other related data, such as morphology, and other phenotypes. It means that both identification techniques, conventional and molecular, complement each other. Moreover, the DNA barcoding technique used to determine the taxonomic status of an organism cannot replace the conventional one (Will and Rubinoff, 2004; DeSalle, 2006; Roslim, 2017; Roslim *et al.*, 2016b, 2016c).

Since the discovery of the DNA barcoding technique, some of the DNA barcodes have been developed and deposited in GenBank. The two standardised DNA barcodes for plant identification are *matK* and *rbcL* genes. These genes are selected based on some considerations, such as the ability to be recovered, the quality of sequence, and the ability to discriminate species (CBOL Plant Working Group, 2009). Both genes are part of the plant chloroplast genome.

The *matK* encodes the maturase enzyme

and has higher substitution mutation rate than the *rbcL*. Consequently, the *matK* variation between species is higher than *rbcL*. Thus, the *matK* is often used in plant phylogenetic studies and plant molecular identification. The *rbcL* gene encodes the large subunit of the ribulose biphosphate carboxylase enzyme that plays an important role in photosynthesis. The *rbcL* sequence is more conservative than the *matK* sequence (Fazekas *et al.*, 2008; Lahaye *et al.*, 2008; Patwardhan *et al.*, 2014; Guo *et al.*, 2016).

Another DNA barcode, which is also frequently used in plant molecular identification in Pandanaceae family, is the *trnL-trnF* intergenic spacer (Buerki *et al.*, 2012; Gallaher *et al.*, 2015). It is a region in the plant chloroplast genome located in between the genes encoding tRNA Leu (UAA) and tRNA Phenylalanine (GAA). Unlike the *matK* and *rbcL* which is the coding region, the *trnL-trnF* intergenic spacer is a non coding region that has the highest variation compared to *matK* and *rbcL* (Kress *et al.*, 2005).

This study reports the use of three plant DNA barcodes (*matK*, *rbcL*, and *trnL-trnF* intergenic spacer) to determine the taxonomic status of Pandan from Riau.

MATERIALS AND METHODS

Plant materials

The plant materials used in this study were Pandan plants (six plants) collected from Kajuik Lake located in Langgam, Pelalawan Regency, Riau Province, Indonesia. However, the DNA was extracted from only one plant. The primer pairs for the *matK* and *rbcL* used in this study were designed based on the DNA sequences that were available in the GenBank database. Moreover, the primer pairs for *trnL* exon and *trnL-trnF* intergenic spacer were designed on the basis of Taberlet *et al.* (1991) suggestions (Table 1). Sequences of *rbcL*, *matK*, and *trnL-trnF* intergenic spacer from some accessions used to create phylogenetic trees were derived from the GenBank.

Table 1. Primers for amplification of *matK*, *rbcL*, and *trnL-trnF* intergenic spacer regions.

Primers	5'-----3'	Annealing Temperatures (°C)	Regions
P- <i>matK</i> -F	GCGATTCTTTCTCCATG	47.0	<i>maturase K</i>
P- <i>matK</i> -R	GGTCCAGATCGGCTTACTA		
P- <i>rbcL</i> -F1	ACTCCTGAATACGAAACCA	52.5	<i>ribulose biphosphate carboxylase</i>
P- <i>rbcL</i> -R1	TCACCTGTTTCGGCCTGCGCT		
B49317_F2	CGAAATCGGTAGACGCTACG	51.8	<i>trnL-trnF</i> exon
A49855_R2	GGGGATAGAGGGACTTGAAC		
B49873_F3	GGTTCAAGTCCCTCTATCCC	50.6	<i>trnL-trnF</i> Intergenic Spacer
A50272_R3	ATTGAACTGGTGACACGAG		

Total DNA isolation

The total DNA was extracted from fresh leaves using DNeasy plant mini kit (Qiagen). The DNA was then migrated on 1.2% agarose gel in 1X TBE buffer (Tris-Borate-EDTA pH 8.0) at 65 volts for 30 minutes to predict the quality and quantity of the DNA. The band was recorded using Olympus SP-500 UZ camera.

The DNA amplification using PCR (polymerase chain reaction) technique

The total DNA was amplified in 50 µl PCR reaction with the following components: 1X PCR buffer (plus Mg²⁺), 0.1 mM dNTPs, 2.4 µM primer forward, 2.4 µM primer reverse, 2 U enzim Dream Taq DNA polymerase (Thermo Scientific), and 1 ng DNA total, and water (Porebski *et al.*, 1997). The PCR analysis was conducted with the following conditions: 5 minutes at 94 °C for 1 cycle, followed by 45 seconds at 94 °C, 45 seconds at annealing temperature (Table 1), and 1 minute at 72 °C for 35 cycles. The PCR process was ended with 1 cycle of post-PCR for 10 minutes at 72 °C.

Electrophoresis

The PCR products were migrated using electrophoresis on 1.2% agarose gels in 1X TBE buffer at 65 volts for 1 hour. After that, the bands were stained using 5 µg/ml ethidium bromide solution, then visualized on the UV lamp transilluminator (WiseUv WUV-M20,

Daihan Scientific), and then documented using a digital camera (Olympus SP-500 UZ).

PCR purification and sequencing

The PCR products were then sent to PT Genetika Science located in Jakarta to be purified and sequenced at 1st Base Malaysia. Sequencing was performed using the PCR primer pairs.

Data analysis

The forward and reverse sequences were aligned using BioEdit version 7.0.0 software (Hall, 1999) to get the sequence of each region (*matK*, *rbcL*, and *trnL-trnF* intergenic spacer). After that, the sequences were analyzed using BLASTn program (Basic Local Alignment Search Tool) at <http://www.ncbi.nlm.nih.gov/BLAST> (Altschul *et al.*, 1997) to search the similarity between the three sequences and the sequences deposited in the GenBank database. Phylogenetic tree or dendrogram was then created using MEGA version 6.06 software (Build#: 6140226) (Molecular Evolutionary Genetics Analysis) (Tamura *et al.*, 2013). The trees were reconstructed by nucleotide sequences using Kimura 2-parameter model and neighbor joining method with 1000 bootstrap. *Stemona tuberosa*, which is a member of Stemonaceae family and Pandanales ordo, was used as an outgroup. The sequences of *matK*, *rbcL*, and *trnL-trnF* intergenic spacer of *S. tuberosa* were obtained from the GenBank database.

RESULTS

The Total DNA of Pandan from Riau was thick enough and was used as a PCR template. The obtained PCR products of *matK*, *rbcL*, *trnL* exon, and *trnL-trnF* intergenic spacer regions were approximately 650 bp, 550 bp, 570 bp, and 450 bp in size, respectively (Figure 1).

The *matK* sequence analysis of Pandan from Riau

It was obtained that the *matK* sequence of Pandan from Riau was 639 bp in size and had been registered in GenBank with the accession number KY503024 (Figure 2). The alignment of the sequence using BLASTn analysis had also been performed and showed that Pandan from Riau had high similarity (ident = 99%) to several species of *Pandanus* and *Martellidendron*. This was supported by query cover value reaching 100%, E-value = 0.0, high max score and total score values (Table 2).

The analysis of *matK* sequences of Pandan from Riau, several species of Pandanaceae, and an outgroup (*S. tuberosa*) showed that there were some variations or differences in nucleotides (Table 3). The length of the *matK* sequence of Pandan from Riau was 639 bp and there were 59 nucleotide differences between Pandan from Riau and accessions analyzed. Nucleotide number 599 was a critical nucleotide for the identification of Pandan from Riau. At that position, Pandan from Riau had Timin (T) while all other accessions analyzed had Cytosine (C) (Table 3).

The *matK* nucleotide sequence-based phylogenetic analysis showed that Pandan from Riau was in group I together with *P. affinis*, *P. tsaratananensis*, *P. maromokotrensis*, and *P. gibbsianus*. Based on the database in GenBank, *P. gibbsianus* and *P. affinis* have synonyms, namely *Benstonea gibbsianus* and *B. affinis*, respectively. Consequently, this clustering showed that group I actually consisted of mixture species of *Benstonea* and *Pandanus*. Moreover, Pandan from Riau had closest relationship with *P. affinis* syn. *B. affinis* rather than other *Pandanus* species. In addition, group II comprised two species of *Martellidendron* such as *M. karaka* and *M. kariangense*, and

group III was only composed of the outgroup, i.e. *S. tuberosa* (Figure 3).

The *rbcL* sequence analysis of Pandan from Riau

The *rbcL* sequence of Pandan from Riau obtained in this study was 539 bp in size and was already registered in GenBank with the accession number KY503025 (Figure 4). The BLASTn analysis was also performed and it showed that Pandan from Riau had high similarity (ident = 99%) to several species from *Pandanus* and *Martellidendron* genus. This result was supported by query cover value reaching 100%, E-value = 0.0, and high values of max score and total score (Table 4).

The analysis of the *rbcL* sequences of Pandan from Riau, several species of Pandanaceae, and *S. tuberosa* as an outgroup showed that there were variations or differences on nucleotide. The length of the *rbcL* sequence of Pandan from Riau analyzed was 539 bp and the sequence contained 23 nucleotide differences. Nucleotide number 536 was a critical nucleotide for the identification of Pandan from Riau. At that position, Pandan from Riau had Cytosine (C) while *P. vandermeeschii*, *P. copelandii*, and *S. tuberosa* had Guanine (G) and *P. tectorius* and *M. Karaka* had Adenine (A) (Table 5). These results also showed that the *rbcL* genes were more conservative and less varied than *matK* gene.

The dendrogram, which was constructed based on the *rbcL* nucleotide sequences, showed that Pandan from Riau formed the same cluster (group I) as three species of *Pandanus*, namely *P. vandermeeschii*, *P. copelandii*, and *P. tectorius*. Based on the GenBank database, *Pandanus copelandii* has a synonymous name, i.e. *B. copelandii*. In other words, these results also showed that group I comprised mixture species of *Pandanus* and *Benstonea*. Furthermore, group II only comprised one species of *Martellidendron*, *M. Karaka*, and group III was only composed of *S. tuberosa* as an outgroup (Figure 5).

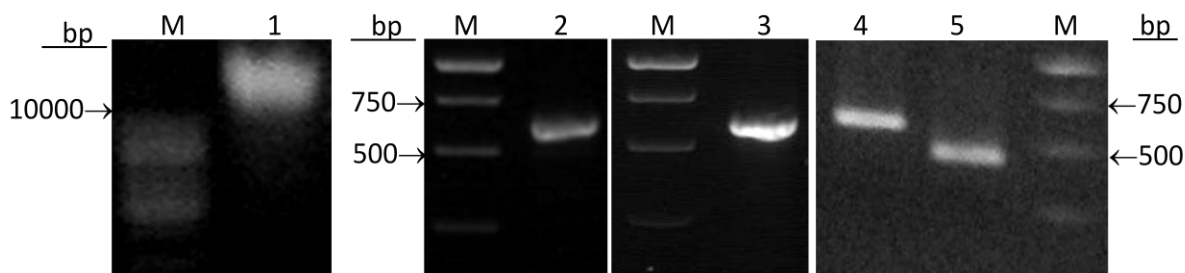


Figure 1. The total DNA (1) and DNA fragments of *matK* (2), *rbcL* (3), the *trnL* exon (4), and *trnLtrnF* intergenic spacer (5) of Pandan from Riau that were migrated on 1.2% agarose gel in 1X TBE buffer. (M) 1 kb DNA Ladder (Thermo Scientific).

```
>KY503024 | Pandan from Riau maturase K (matK) gene, partial
sequence; chloroplast
TCTCCATGAATATCATAATTGGAGTAGTTTCATTACTCTGAAGAAATCTATTTACGGTTT
TTCAAAAAGAAAATAAAAAGACTATTTTCGATTCCCTATATAATTCTTATGTATCTGAATGCGA
ATTTGTATTAGTTTTTTTTTTTGTAAACAATCTTCTTATTTACGATCAACATCTTCTGGAAC
CTTTCTTGAGCGAACACATTTCTATGGAAAAATAGAACATCTTAATCCTATAGTAGTGTG
TCGTAATTATTTTCCGAAGAACCCTTTGTTCTTCAAGGATCCTTTCATGCATTATGTTTCG
ATATCAAGGAAAAGCAATTCTGGCTTCAAAGGAACTCATCTTCTGATGGAAAAATGGAG
ATGTCACCTTGTCAATTTCTGGCAATATTATTTTCACTTTTGGTCTCAACCGTACAGGAT
TCGTATAAACCATTATCAAACCATCTTTCTATTTTCTGGGTTATCTTTTAAGTCTACT
AATAAATCCTTCGTCGGTAAGGAATCAAATGCTAGAAAATTCATTCTAATAGATACTGT
TACTAAAAAATTCGATCCCATAGTCCCAATTATTCCTCTTATTGGATCATTGTCTAAAGT
TAAGTTTTGTACCGTATCAGGGCATCCTAGTAGTAAGCC
```

Figure 2. The *matK* sequence of Pandan from Riau.

Table 2. The alignment analysis using BLASTn on the *matK* sequence of Pandan from Riau.

Description	Max score	Total score	Query cover	E value	Ident	Accession
<i>Pandanus gibbsianus</i>	1144	1144	100%	0.0	99%	JX286821.1
<i>Pandanus affinis</i>	1135	1135	100%	0.0	99%	JX286801.1
<i>Martellidendron karaka</i>	1135	1135	100%	0.0	99%	JX286739.1
<i>Martellidendron kariangense</i>	1135	1135	100%	0.0	99%	JX286737.1
<i>Pandanus tsaratananensis</i>	1130	1130	100%	0.0	99%	JX286742.1
<i>Pandanus maromokotrensis</i>	1130	1130	100%	0.0	99%	JX286743.1


```
>KY503025 | Pandan from Riau ribulose-1,5-bisphosphate
carboxylase large subunit (rbcL) gene, partial sequence;
chloroplast
CCCGGAGTCCGCCTGAAGAAGCAGGGGCAGCGGTAGCTGCCGAATCTTCTACTGGTACA
TGGACAACCTGTGTGGACTGATGGACTTACCAGTCTTGATCGTTACAAAGGACGATGCTAC
CACATAGAGGCCGTTGTTGGGGAGGACAATCAATATATTGCTTATGTAGCTTATCCTTTA
GACCTTTTTGAAGAAGGTTCCGTTACTAACATGTTTACTTCCATCGTAGGTAATGTATTT
GGTTTCAAAGCCCTACGAGCTCTACGTCTGGAGGATTTGCGAATTCCTCCTGCTTATTCC
AAAACTTTTCAAGGCCCGCCTCATGGCATCCAAGTTGAAAGAGATAAATTGAACAAGTAT
GGTCGTCCCTTATTGGGATGTACTATTAACCAAAAATTGGGATTATCTGCAAAGAACTAC
GGTAGGGCGGTTTATGAATGTCTACGCGGTGGACTTGATTTTACCAAGGATGATGAAAAC
GTGAACTCACACCATTATGCGTTGGAGAGACCGTTTCGTATTTTGTGCCGAAGCCCT
```

Figure 4. The *rbcL* sequence of Pandan from Riau.

Table 4. The alignment analysis using BLASTn on the *rbcL* sequence of Pandan from Riau.

Description	Max score	Total score	Query cover	E value	Ident	Accession
<i>Pandanus vandermeeschii</i>	964	964	100%	0.0	99%	JX903247.1
<i>Pandanus tectorius</i>	964	964	100%	0.0	99%	JN407337.1
<i>Pandanus copelandii</i>	964	964	100%	0.0	99%	AY465701.1
<i>Martellidendron karaka</i>	955	955	100%	0.0	99%	FN870868.1

Table 5. Nucleotide differences on *rbcL* sequence.

Species	Nucleotide number*																						
	1	1	1	1	1	2	2	2	2	3	3	3	3	3	4	4	4	4	5	5			
	4	6	2	2	3	4	4	0	0	2	7	9	0	0	1	2	5	2	2	4	4	1	3
	1	6	5	9	0	3	6	0	9	4	6	3	2	8	4	6	9	5	8	6	9	9	6
<i>Pandan from Riau</i>	C	A	A	G	C	G	C	C	C	T	T	A	T	C	C	T	G	G	C	T	G	C	
<i>Pandanus vandermeeschii</i>	C	G
<i>Pandanus tectorius</i>	C	A
<i>Pandanus copelandii</i>	C	G
<i>Martellidendron karaka</i>	.	.	.	C	C	.	.	A	A
<i>Stemona tuberosa</i>	T	T	C	A	G	A	A	T	T	T	C	C	G	.	A	T	C	A	C	G	G	T	G

*The numbers arranged vertically show nucleotide positions referring to Pandan from Riau sequence. Dots (.) indicate that the nucleotide on particular position was the same as the one of Pandan from Riau sequence. Nucleotides in the box and bold are the critical nucleotides for the identification of Pandan from Riau.

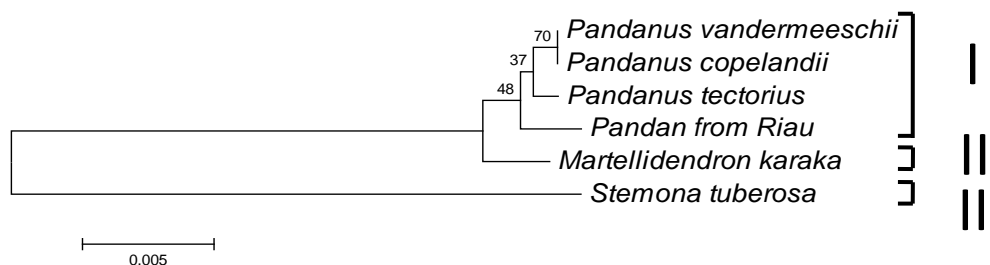


Figure 5. Dendrogram constructed based on the *rbcL* sequences using Neighbor Joining method with 1000 bootstraps.

The intergenic spacer *trnL-trnF* sequence analysis of Pandan from Riau

The sequences of *trnL* exon (570 bp) and *trnL-trnF* intergenic spacer (466 bp) had been obtained and made a contig sizing the 1014 bp. The contig of *trnL-trnF* intergenic spacer had been registered in GenBank with the accession number KY649618 (Figure 6). The BLASTn analysis was performed and it showed that Pandan from Riau had higher similarity to species from genus *Benstonea* and *Martellidendron* (ident = 99%) compared to species from genus *Pandanus* (ident = 96%-97%) and *Freycinetia* (ident = 97%). This was supported by high values of max score, total score, and query cover (91%), as well as E-value = 0.0 (Table 6).

The analysis of *trnL-trnF* intergenic spacer sequences of Pandan from Riau, several species of Pandanaceae, and *S. tuberosa* as an outgroup showed that there were more nucleotide variations or differences compared to *matK* and *rbcL* sequences. A total of 195 nucleotide differences were observed. Nucleotides at positions 854 and 855 were critical nucleotides for the identification of Pandan from Riau. At that position, Pandan from Riau had Timin (T) and Cytosine (C) respectively, while others had Adenine (A) at both position (Table 7).

The dendrogram that was created based on the *trnL-trnF* intergenic spacer sequences showed that there were five groups. The first group (I) comprised of Pandan from Riau and three species of *Benstonea*. The second group (II) only consisted of *M. hornei*. The third group (III) comprised four species of *Pandanus*. The fourth (IV) and fifth (V) groups were composed of *Freycinetia multiflora* and *S. tuberosa*, respectively (Figure 7). These results demonstrated that Pandan from Riau was more closely related to species of *Benstonea* than the other three genus (*Pandanus*, *Martellidendron*, and *Freycinetia*) of Pandanaceae based on the *trnL-trnF* intergenic spacer sequence.

General description of Pandan from Riau

Pandan from Riau is an acaulescent shrub that grows under large trees such as *Syzygium sp*, *Planchonia valida*, and *Ochreanauclea maingayi* trees at the edge of the lake. Plant height reaches 62 cm with one clump of leaves bearing a single fruit. The leaves have prickles at their margin, their main vein on the lower part of leaf surface, and the tip of main leaf vein on the upper part of leaf surface. The width of the leaf is 1.5 cm and the length of the leaf can reach up to 43.5 cm. The ripe fruit has the diameter of 14 cm and length of 5 cm. There are approximately 430 seeds in one fruit with the length of fruit stem is 8 cm (Figure 8).

DISCUSSION

Identity or taxonomic status of an organism is crucial for studies in the biological field. The taxonomic status can be determined based on the morphological and agronomical characteristics, protein or biochemical, and DNA. The identification of an organism identity using short DNA sequence, which is located in the nuclear and plastid genomes, is called DNA barcoding technique, and this is applicable for plant molecular identification (White *et al.*, 1990; Herbert *et al.*, 2003; Kress *et al.*, 2005; Fazekas *et al.*, 2008; Lahaye *et al.*, 2008; Buerki *et al.*, 2012; Patwardhan *et al.*, 2014; Ali *et al.*, 2014; Gallaher *et al.*, 2015; Guo *et al.*, 2016). This technique is crucial for plants and is required to help identify plant specimens which, when identification process is conducted, is not accompanied by complete plant parts such as leaf, stem, flower, and fruit. The plant specimen is a member of a genus with very extensive morphological variation such as *Pandanus* (Stone, 1974, 1983, 1993; Buerki *et al.*, 2012).

Pandanus is a genus which belongs to the Pandanaceae family composed of monocotyledonous and dioecious plants. Prior to 2012, the Pandanaceae family was divided into four genus with 700 species which are: *Pandanus* Parkison, *Freycinetia* Gaudich, *Martellidendron* (Pic. Serm.) Callm & Chassot,


```

>KY649618 | Pandan from Riau trnL-Leu (trnL) gene, partial
sequence; trnL-trnF intergenic spacer, complete sequence;
and tRNA-Phe (trnF) gene, partial sequence; chloroplast.
GGTGACGCTACGGACTTGATTGGATTGAGCCTTGGTATGGAAACCTGCTAAGTGGTAACT
TCCAAATTCAGAGAAACCCTGGAATTA AAAATGGGCAATCCTGAGCCAAATCTTTATTTT
GAGAAAACAAAACAAGGGTTTATAAAACCAGAATCAAAAAAAGGATAGGTGCAGAGACTC
AATGGAAGCTGTTCTAACGAATGGAGTTGATTGCGTTACGTTGGTAGCTGGAATCCCTCT
ATCGAAATTACAGAAGGGATTGACCTATATTATATACCTGTACGTATACATACTAACATA
GCAAACGATTAATCACGACCAAATCCATATATATATGAATATGAAAAAATTCAGAATTAT
TGTGAATCCATTTCAATCGAAATCGAAGTTGAAGGAAGAATTGAATATTCAGTGATCAA
TCATTCAATTCAGAGCCGATCTTTTGAAAAACTGATTAATCGGACGAGAATAAAGAGAGA
GTCCCGTTCTACATGTCAATACCAACAACAATGAAATTTATAGTAAGAGGAAAATCCGTC
GACTTTAGAAATCGTGAGGGTTCAAGTCCCTCTATCCCCAATAAAAAGACCATTTTACCT
CCTAACTAGTTATTCTCTTTTTTTTTCATCAGCGCTTTAAACTTCACTATCTTTTTTCATTC
ACTCTACTCTTTCAAAAACAGATCCGAACAGAAATCTTTGGATCTTATCCTAAGTCTTTT
GGATAGATACGATACCCGTACCCGTACAAATAAACATATATGGAAAAGGAATTTCTATTA
TTGAATCATTACAGTCCATATCATTATCTTTACACTTACAAAGATAGTCTTCTTTTGA
AGATCTTCAAAAATTTGGAGACTGGTTAAGATTTATTAATTTAATACTTTTTTAGTTCCTT
TAATTGACATAGAATAGATAACAAGTACTCTACTCGGATGATGCGCGGGAATGGTCGGGA
TAGCTCAGTTGGTAGAGCAGAGGACTGAAAATCCTCGTGTCCACGATTCAAATA

```

Figure 6. The *trnL-trnF* intergenic spacer sequence of Pandan from Riau.

Table 6. The alignment analysis using BLASTn on the *trnL-trnF* intergenic spacer sequence of Pandan.

Description	Max score	Total score	Query cover	E value	Ident	Accession
<i>Benstonea monticola</i>	1654	1654	91%	0.0	99%	KJ681562.1
<i>Benstonea lauterbachii</i>	1645	1645	91%	0.0	99%	KJ681568.1
<i>Benstonea affinis</i>	1636	1636	91%	0.0	99%	KJ681518.1
<i>Martellidendron hornei</i>	1631	1631	91%	0.0	99%	KJ681504.1
<i>Pandanus tectorius</i>	1613	1613	94%	0.0	97%	FJ194471.1
<i>Pandanus pygmaeus</i>	1564	1564	91%	0.0	97%	KJ681527.1
<i>Pandanus furcatus</i>	1544	1544	91%	0.0	96%	KJ681539.1
<i>Pandanus pancheri</i>	1544	1544	91%	0.0	96%	KJ681508.1
<i>Freycinetia multiflora</i>	1537	1537	91%	0.0	97%	KJ681552.1

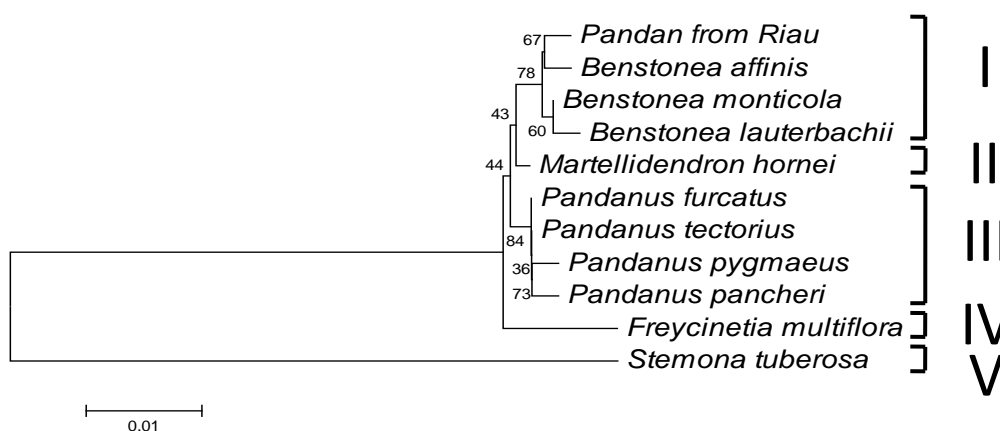


Figure 7. Dendrogram generated from the *trnL-trnF* intergenic spacer sequences using Neighbor Joining method with 1000 bootstraps.



Figure 8. Pandan from Riau plant habit. (a) an acaulescent shrub habit with a terminal fruit, (b) zoom of a terminal syncarp, (c) zoom of a single drupe freely from syncarp, (d) a terminal female flower, and (e) a terminal male flower.

and *Sararanga* Helms (Stone, 1983; Stone *et al.*, 1998; Huynh, 2001; Callmender *et al.*, 2003; Keim, 2009). *Pandanus* is dominated by trees and shrubs and has wide geographic distribution, i.e. from India to South Pacific. Meanwhile, Borneo and Peninsular Malaysia are the diversity center of this genus. *Pandanus* is then divided into four sections, sect. *Acrostigma* (Kruz) B.C. Stone, sect. *Epiphytica* Martelli, sect. *Fusifforma* H. St. John, and sect. *Pseudoacrostigma* B.C. Stone (Stone, 1978, 1983, 1993; Callmender, 2000; Keim, 2009).

Recently, an update on phylogenetic relationship between species and genus has been performed in the Pandanaceae family using three DNA chloroplast sequences, including *matK*, *trnQ-rps16*, and *trnL-trnF*. It is concluded that there are five clades in Pandanaceae family. The five clades are *Freycinetia*, *Sararanga*, *Martellidendron*, *Pandanus* sect. *Acrostigma*, and the rest of *Pandanus*. Such analysis was done using molecular techniques, as morphological analysis often caused error in placing species into a genus or clade (Buerki *et al.*, 2012).

Buerki *et al.* (2012) stated that morphologically, *Pandanus* sect. *Acrostigma* has its own flowers characteristics which are different from the three sections of *Pandanus*, and this difference leads *Pandanus* sect. *Acrostigma* to further become the fifth new genus in Pandanaceae family and named as *Benstonea* Callm. & Buerki (Callmender *et al.*, 2012). At this time, 60 species are recorded in *Benstonea* (originates from Benjamin Stone's name) and 21 of them have synonyms (Callmender *et al.*, 2012).

Previously, Pandan from Riau had been characterized morphologically and categorized into *Pandanus*. However, the species' name has not been determined due to the absence of flower during identification (Elvyra and Yus, 2010). Therefore, DNA barcoding technique was applied in this study to determine the scientific name of Pandan from Riau using three plant DNA barcodes. Two of them, *matK* and *trnL-trnF* intergenic spacer regions, were commonly applied in Pandanaceae family identification (Buerki *et al.*, 2012, 2016; Callmender *et al.*, 2012, 2013) and there were abundance of data in the GenBank database. The

166 sequence of *matK*, 54 of *rbcL*, and 320 of *trnL-trnF* intergenic spacer of *Pandanus* were already available in the GenBank database (updated 26 January 2017), but none of them has 100% similarity to the Pandan from Riau sequences.

Four nucleotides, such as nucleotide number 599 on *matK* sequence, number 536 on *rbcL*, number 854 and 855 on *trnL-trnF* intergenic spacer, were stated as critical nucleotides for the identification of Pandan from Riau. It was because at those nucleotide positions, Pandan from Riau had different nucleotides compared to other accessions that were analyzed. These results support the idea that DNA barcoding technique can help in confirming the taxonomic status of an organism based on the nucleotide differences on certain position or called critical nucleotide.

This study also showed several pieces of evidence of the limitations of *matK* and *rbcL* sequences as molecular identification tools for the Pandanaceae family. Firstly, both sequences were more conservative compared to *trnL-trnF* intergenic spacer sequence. However, the *rbcL* sequence was the most conservative one. Secondly, the *matK* and *rbcL* sequences cannot discriminate species in *Benstonea* and *Pandanus* genus. This can be seen on the dendrograms generated from those sequences. On the contrary, the *trnL-trnF* intergenic spacer sequence can differentiate and cluster the species of *Benstonea* and *Pandanus* genus separately (Figure 7). Therefore, the *trnL-trnF* intergenic spacer sequence was often used to identify and analyze the phylogenetic relationship, evolution, and distribution on the species of Pandanaceae family (Buerki *et al.*, 2012, 2016; Callmender *et al.*, 2012).

This study had succeeded in determining the taxonomic status of Pandan from Riau up to the genus level (*Benstonea* sp). It also suggested that the Pandan from Riau might be an identified species, although its DNA barcode sequences were not present in the GenBank database or this plant is a new species of *Benstonea* which has never been identified and reported.

CONCLUSION

The combination of DNA barcode sequences can be used to identify plants, particularly for specimens that are members of a genus with wide morphological variation such as the Pandanaceae. The taxonomic status of Pandan from Riau stipulates that it is a member of *Benstonea* by using the three DNA barcodes. It is suggested that Pandan from Riau is an identified species, although its DNA barcode sequences were not present in the GenBank database. It can also be understood that this plant is a new species of *Benstonea* which is never been identified and reported.

ACKNOWLEDGEMENTS

This research was funded by a grant from the Ministry of Research, Technology, and Higher Education, Republic of Indonesia (2017) under 'Hibah Penelitian Fundamental' on the behalf of Dr. Dewi Indriyani Roslim. I would like to thank Putri Nurkhairani, S.Si. and Siti Khumairoh, S.,Si. for their technical assistance in the laboratory.

REFERENCES

- Ali MA, Gyulai G, Hidve'gi N, Kerti B, Al Hemaïd FMA, Pandey AK, Lee J (2014). The changing epitome of species identification – DNA barcoding. *Saudi J. Biol. Sci.* 21: 204-231.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25: 3389-3402.
- Buerki S, Callmander MW, Devey DS, Chappell L, Gallaher T, Munzinger J, Haevermans T, Forest F (2012). Straightening out the screw-pines: A first step in understanding phylogenetic relationships within Pandanaceae. *Taxon* 61(5): 1010-1020.
- 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.
- Callmander MW (2000). *Pandanus* subg. *Martellidendron* (Pandanaceae) part 1: New findings on *Pandanus hornei* Balf. f. (sect. *Seychellea*) from the Seychelles. *Webbia* 55: 317-329.
- Callmander MW, Chassot P, K pfer P, Lowry II PP (2003). Recognition of *Martellidendron*, a new genus of Pandanaceae, and its biogeographic implications. *Taxon* 52: 747-762.
- Callmander MW, Lowry II, PP, Forest F, Devey, Beentje H, Buerki S (2012). *Benstonea* Callm. & Buerki (Pandanaceae): characterization, circumscription, and distribution of a new genus of screw-pines, with a synopsis of accepted species. *Candollea* 67: 323-345.
- Callmander MW, Booth T, Beentje H, Buerki S (2013). Update on the systematics of *Benstonea* (Pandanaceae): when a visionary taxonomist foresees phylogenetic relationships. *Phytotaxa* 112: 57-60.
- CBOL Plant Working Group (2009). A DNA barcode for land plants. *Proceed. Natl. Acad. Sci. USA.* 106: 12794-12797.
- DeSalle R (2006). Species discovery versus species identification in DNA barcoding efforts: response to Rubinoff. *Conserv. Biol.* 20(5): 1545-1547.
- Elvyra R, Yus Y (2010). Ikan Lais dan Sungai Papan Banjar di Provinsi Riau. UR Press, Pekanbaru.
- Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* 3: e2802.
- Gallaher T, Callmander MW, Buerki S, Keeley SC (2015). A long distance dispersal hypothesis for the Pandanaceae and the origins of the *Pandanus tectorius* complex. *Mol. Phylogenet. Evol.* 83: 20-32.
- Guo Y-Y, Huang L-Q, Liu Z-J, Wang X-Q (2016). Promise and Challenge of DNA Barcoding in *Venus Slipper* (Paphiopedilum). *PLoS ONE* 11(1): e0146880.
- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acid S.* 41: 95-98.
- Harsono T, Pasaribu N, Sobir, Fitmawati (2016). Diversity of gandaria (*Bouea*) based on morphological characters in Indonesia. *SABRAO J. Breed. Genet.* 48(4): 504-517.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. London B.* 270: 313-321.

- Huynh KL (2001). Contribution to the flower structure of *Sararanga* (Pandaceae). *Bot. J. Linn Soc.* 136: 239-245.
- Jarman SN, Elliott NG (2000). DNA evidence for morphological and cryptic Cenozoic speciations in the Anaspididae, living fossils' from the Triassic. *J. Evol. Biol.* 13: 624-633.
- Keim AP (2009). Pandaceae of the Island of Yapen, Papua (West New Guinea), Indonesia, with their nomenclature and notes on the rediscovery of *Sararanga sinuosa*, and several new species and records. *Blumea* 54: 255-266.
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH (2005). Use of DNA barcodes to identify flowering plants. *Proceed. Natl. Acad. Sci. USA*. 102: 8369-8374
- Lahaye R, Bank MV, Bogarin D, Warner J, Pupulin F (2008). DNA barcoding the floras of biodiversity hotspots. *Proceed. Natl. Acad. Sci. USA* 105: 2923-2928.
- Patwardhan A, Samit R, Amit R (2014). Molecular markers in phylogenetic studies-A Review. *J. Phylogen. Evol. Biol.* 2: 131.
- Porebski S, Bailey LG, Baum BR (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Rept.* 15(1): 8-15.
- Roslim DI (2017). Pandan (*Pandanus* sp), Rotan (*Calamus* sp), and Rengas (*Gluta* sp) from Kajuik Lake, Riau Province, Indonesia. *Brazilian Arch. Biol. Technol.* 60: e17160419.
- Roslim DI, Khumairoh S and Herman (2016a). Confirmation of tuntun angin (*Elaeocarpus floribundus*) taxonomic status using *matK* and ITS sequences. *Biosaintifika: J. Biol. Biol. Edu.* 8(3): 392-399.
- Roslim DI, Nurkhairani P, Herman, Elvyra R (2016b). Identification of durik-durik plant (*Syzygium* sp) using the *psbA-trnH* intergenic spacer and ITS regions. *TPGM* 3: 11-16.
- Roslim DI, Herman, Sofyanti N, Chaniago M, Restiani R, Novita L (2016c). Characteristics of 22 cassavas (*Manihot esculenta* L., Crantz) from Riau Province, Indonesia. *SABRAO J. Breed. Genet.* 48(2): 110-119.
- Sofyanti N, Mat-Salleh K, Mahmud K, Mazlan NZ, Hasein MRA, Burslem DFRP (2016). *Rafflesia parvimaclulata* (Rafflesiaceae), a new species of *Rafflesia* from Peninsular Malaysia. *Phytotaxa* 253(3): 207-213.
- Stoeckle M (2003). Taxonomy, DNA, and the bar code of life. *BioSci.* 53:9.
- Stone BC (1974). Toward an improved infra-generic classification in *Pandanus* (Pandaceae). *Botanische Jahrbücher für Systematik* 94: 459-540.
- Stone BC (1978). Revisio Pandacearum, Part I. *Pandanus* subgenera Coronata and Acrostigma. *Flora Malesaina precursores. Fordia incredibilis* Whitmore, 23.
- Stone BC (1983). Some new and critical *Pandanus* species of subgenus *Acrostigma*, I. Supplement to revisio pandacearum. *Gardens' Bulletin Singapore* 36: 205-212.
- Stone BC (1993). Studies in Malesian Pandaceae, 21. The genus *Pandanus* in Borneo. *Sandakanian* 2: 3-84.
- Stone BC, Huynh KL, Poppendieck HH (1998). Pandaceae, in: Kubitzki, K. (Eds.), The families and genera of flowering plants. Vol. 3. Springer, Berlin, pp. 397-404.
- Tamura K, Stecher G, Peterson D, Filipksi A, Kumar S (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30: 2725-2729.
- 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.
- White TJ, Bruns T, Lee S, Taylor J (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. in: PCR Protocols: A guide to methods and applications. Academic Press, Inc.
- Will KW, Rubinoff D (2004). Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* 20: 47-55.