Available online freely at www.isisn.org

Bioscience Research

Print ISSN: 1811-9506 Online ISSN: 2218-3973

Journal by Innovative Scientific Information & Services Network

RESEARCH ARTICLE BIOSCIENCE RESEARCH, 2021 18(2): 1277-1283. **OPEN ACCESS**

Comparison of three DNA barcode loci for distinguishing Hoan-ngoc (*Pseuderanthemum palatiferum*) from its relative

Viet The HO* and Son Tuyen BUI

Ho Chi Minh City University of Food Industry, 140 Le Trong Tan, Tan Phu district, Ho Chi Minh City, **Vietnam**

***Correspondence:** thehv@hufi.edu.vn **Received 13-01-2021, Revised: 15-03-2021, Accepted: 20-03-2021 e-Published: 10-04-2021**

Hoan-ngoc is a highly important herbal plant in Vietnam. To improve the effectiveness of conservation and development of this plant, an accurate method for classification and identification is needed. In this study, *matK*, *rbcL*, and *trnH-psbA* barcodes were utilized to evaluate genetic relatedness among 13 hoang-ngoc accession collected throughout Vietnam and distinguish hoan-ngoc from its close relative plant (*Clinacanthus nutans* or "Xuong-khi" in Vietnamese). The results revealed that three are close genetic relatedness among hoan-ngoc samples. The *matK* and *trnH-psbA* loci reveal potential to differentiate hoang-ngoc from *Clinacanthus nutans.* The findings of this project could provide valuable information that is necessary for classification, plant origin identification, breeding, and conservation program of hoan-ngoc in Vietnam.

Keywords: DNA barcode, hoan-ngoc; *matK*, *rbcL*, *trnH-psbA*.

INTRODUCTION

Hoan-ngoc (*Pseuderanthemum palatiferum*) is a native plant in Vietnam which is commonly used as traditional medicine not only for this country but for neighbor countries (Sittisart et al., 2016). This herbal plant was first identified in 1990 in Cuc Phuong National Park in northern Vietnam provinces (Nguyen and Eun, 2013). The plant was also found popularly distributed in nearby areas such as the Red River Delta, Thai Binh, Nam Dinh, Hung Yen and Hai Duong provinces. Nowadays, the planted area has been replicated in the southern provinces of Vietnam. In addition, hoan-ngoc is also found in some countries such as Thailand, China, Cambodia, India, Laos, Malaysia and Myanmar. Hoan-ngoc is well known for its medicinal value which contains activities to treat diarrhea, hypertension, rheumatoid arthritis, and pharyngitis (Padee, 2009). The leaves have high antibacterial properties, can inhibit the growth of gram-negative, gram-positive bacteria, and both mold and yeast and high ability to protect liver cells (Huynh et al. 2008).

As an important herbal plant in Vietnam, hoan-ngoc is commonly mixed with other inexpensive medicinal plants with similar morphological characteristics, especially *Clinacanthus nutans*. The traditional way is to distinguish based on morphological differences in the plant. However, the distinction of these ingredients is not effective since ordinary ingredients reach the users' hands that have been either preliminary processed or processed into products.

In 2003, DNA barcode was proposed as new criteria to identify existing and unknown species by comparing with Genbank's DNA sequence. DNA barcode plays a role as a tool in taxonomy and is being successfully exploited in assessing biodiversity, determining the purity of herbal samples. DNA barcode are being popular use to manage the adulteration of herbs and plant

identification. In plants, the barcode DNA regions used are usually the sequences of the chloroplast genome and the nuclear genome. Different gene regions have been utilized as barcodes such as nuclear ribosomal internal transcribed space (ITS), rbulose-1,4-bisphospate carboxylase larger subunit gene (*rbcL)*, ribosomal rNA maturase gene (*matK)*, *atpF-atpH*, *psbK-psbI* and *trnH-psbA* intergentic spacers. Among these, *matK*, *rbcL* and *trnH-psbA* have been extensively utilized for classification of different plant species such as *Olea europaea* (Gismondi et al. 2012); *Plectranthus hadiensis* (Amarasinghe et al. 2013); Cymodocea (Bchir et al. 2019). This study was aimed to assess the efficiency of three barcode loci consisting of *rbcL*, *matK* and *psbA-trnH* to examine genetic variation among different hoan-ngoc accession and to identify its effectiveness in distinguishing hoan-ngoc from *Clinacanthus nutans*.

MATERIALS AND METHODS

Total of 13 hoan-ngoc genotypes and 1 *Clinacanthus nutans* (XK) genotype were collected from different provinces in Vietnam and presented in Table 1 and Figure 1. Leaf samples were dried and stored in silica gel until use.

DNA was extracted with CTAB method (Cetyl Trimethyl Ammonium Bromide) as described by Doyle and Doyle (1987). PCR reactions were amplified using the composition as follows: 7.5 µL 2X Mytaq Mix (Bioline, UK), 20 ng DNA, 0.2 µM primer (either *matK* 390F: 5'-CGATCTATTCATTCAATATTTC-3'; and 1326R: 5'- TCTAGCACACGAAAGTCGAAGT-3' (Cuenoud et al. 2002); *rbcL*: cF: 5'-TGAAAACGTGAATTCCCAACCGTTTATGCG-3'; cR: 5'- GCAGCAGCTAGTTCCGGGCTCCA-3' (Hasebe et al. 1994) or *trnH-psbA* (trnH-psbA-F: 5'- GTTATGCATGAACGTAATGCTC-3'; and trnH-psbA-R: 5'- CGCGCATGGTGGATTCACAATCC-3' (Ferri et al. 2009), and PCR water (Sigma-Aldrich, USA) to final volume of 15 µL. The PCR reaction conditions were as follows: initial denaturation at 94 °C for 5 minutes; then 30 cycles of 30 seconds at 94 ^{o}C, 30 seconds at 52 ^{o}C (for *matK* and *trnH-psbA* loci) or 58 ^{o}C (for *rbcL* locus), and 1 minute at 72 ^{o}C. Finally, addition 5 minutes was continued at 72 ^{o}C to complete the reaction. All reactions were carried out in SureCycler 8800 Thermal Cycler (Agilent, USA). The PCR products were electrophorized on 1% agarose gel using 1 kb DNA marker (Bioline, UK) to confirm the amplification length. The PCR products were then purified by ISOLATE II PCR and Gel Kit (Bioline, UK) and sequenced by Sanger method at PhusaBiochem (Can Tho, Vietnam). The obtained electropherograms were edited using FinchTV (Digital World Biology Products, USA). Only the sequences with scores more than 20 PHRED score were considered for further analysis. Sequences were trimmed at both ends of the alignment in order to avoid too many missing data at the ends. The obtained sequences were submitted to GenBank (NCBI, USA) and are publicly accessible under the accession numbers listed in Table 1

For species identification, the DNA sequences were identified by using Basic Local Alignment Search Tools (BLAST) at http://blast.ncbi.nlm.gov using default parameters. The identification was deemed correct if the highest identity percentage of searched sequences was derived from expected species or genus. On the other hand, the identification was considered ambiguous when the highest identity percentage of searched sequences was not derived from expected species or genus or family (Nio et al., 2018). DNA sequences were then aligned with the ClustalW algorithm, implemented in MEGA7 package (Tamura et al. 2013), using the default parameters. Evolutionary divergence for each data set and pattern of nucleotide substitution were performed by same software. Evolutionary trees were constructed based on two methods consisting of Maximum Likelihood (ML) and Neighbour Joining (NJ) representing for discrete character methods and distance methods, respectively. The reliability of phylogenetic analysis was validated by 1000 bootstrap replicates. Bootstrap support (BS) was categorized as strong (> 85%), moderate (70%-85%), weak (50%-69%) or poor (< 50%) (Kress et al. 2002). In order to estimate species resolution for a given barcode locus, we considered the species were resolved if conspecific individual grouped into one monophyletic branch in the phylogenetic tree with strong bootstrap support. In converse, if conspecific individuals were separated in paraphyletic branches, then the species was considered as identification failure (Sikdar et al. 2018).

Table 1: List of samples used in this study and the corresponding accession numbers of DNA sequences in three different barcode loci.

No.	Sample ID	Species	*matK*	*rbcL*	*trnH-psbA*
1	BT01	*Pseuderanthemum palatiferum*	MW460671	MW460684	MW460697
2	TG01	*Pseuderanthemum palatiferum*	MW460672	MW460685	MW460698
3	LA01	*Pseuderanthemum palatiferum*	MW460673	MW460686	MW460699
4	SG01	*Pseuderanthemum palatiferum*	MW460674	MW460687	MW460700
5	SG02	*Pseuderanthemum palatiferum*	MW460675	MW460688	MW460701
6	SG03	*Pseuderanthemum palatiferum*	MW460676	MW460689	MW460702
7	SG04	*Pseuderanthemum palatiferum*	MW460677	MW460690	MW460703
8	TN01	*Pseuderanthemum palatiferum*	MW460678	MW460691	MW460704
9	BP01	*Pseuderanthemum palatiferum*	MW460679	MW460692	MW460705
10	HN01	*Pseuderanthemum palatiferum*	MW460680	MW460693	MW460706
11	HN02	*Pseuderanthemum palatiferum*	MW460681	MW460694	MW460707
12	PT01	*Pseuderanthemum palatiferum*	MW460682	MW460695	MW460708
13	TH01	*Pseuderanthemum palatiferum*	MW460683	MW460696	MW460709
14	XK	*Clinacanthus nutans*	MW460887	MW460888	MW460889

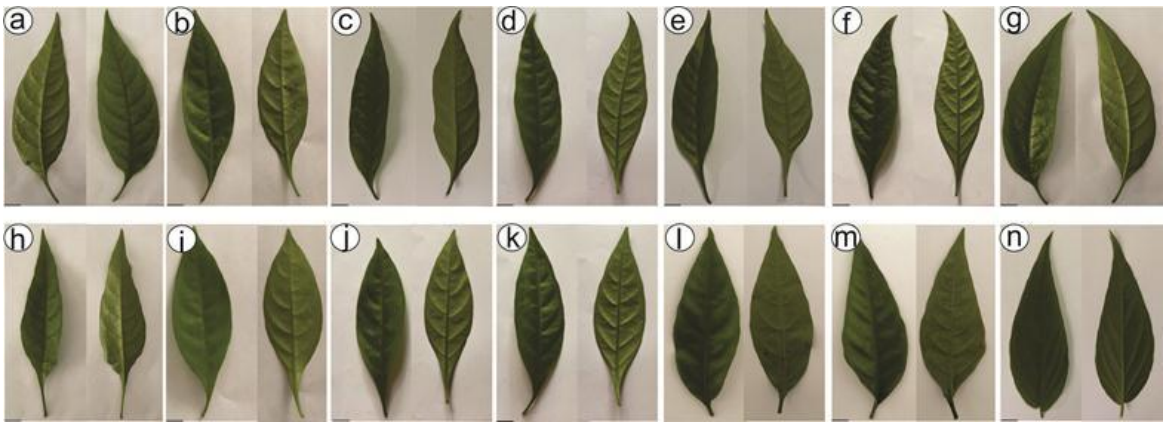

Figure 1. Leaf morphology of collected samples (a: BT01; b: TG01; c: LA01; d: SG01; e: SG02; f: SG03; g: SG04; h: TN01; i: BP01; j: HN01; k: HN02; l: PT01; m: TH01; n: XK (*Clinacanthus nutans)*). Bar = 1cm.

RESULTS AND DISCUSSION

Species identification

In this study, all three genes were amplified and sequenced successfully. For homologous identification, only sequences showing a minimum of 80% percentage identity were considered. Using BLAST, sequences from three barcode loci show inconsistent results as described in Table 2. For all hoan-ngoc accessions, the species identification by BLAST is not reliable, for *matK* studied sample was incorrectly classified as *Pseuderanthemum carruthersii* and *Cosmianthemum viriduliflorum* with 8 and 5 accessions, respectively. Whereas, BLAST with *rbcL* revealed all accessions belonging to *Ruttya fruiticosa* or *Pseuderanthemum atropurpureum* when *trnH-psbA* was used. We found that only *Clinacanthus nutans* was correctly identified by all loci. The low accuracy of BLAST could be resulted from their database small size and insufficient completeness. Consequently, the missing species in the database cannot be identified and the method may assign the query sequence to an incorrect species (Parmentier et al. 2013).

Phylogenetic analysis

Firstly, phylogenetic trees were constructed by single maker. The results show the high similarity between two employed methods namely ML and NJ (Figure 2). For *matK* and *trnH-psbA* loci,*Clinacanthus nutans* used as out group is separated in distinct branch indicating the effectiveness of these two loci in differentiating hoan-ngoc and *Clinacanthus nutans.* The evolution of *matK* region is considered as the fastest in plastid genome and the sequence is highly similar to COI sequence in animal which is commonly used as key barcode region in animal identification. A study on Casuarinaceae found that *matK* gene gave higher resolution than *rbcL* (Sogo et al., 2001). A research group in Vietnam also reported that *matK* region is more reliable marker than *rbcL* on *Hopea chinesis* (Trang et al, 2015). On another hand, *trnH-psbA* is also previously considered as a high potential locus since it contains short sequence which facilitate the amplification and sequence more straightforward. Previous study on Combretaceae family also indicated that *trnH-psbA* is better than *rbcL* and *matK* (Gere et al., 2013). Li and colleagues used several barcode to evaluate *Ficus* in China, they found that the success rate of species discrimination reduce from *trnH-psbA*, *matK* and *rbcL* (Li et al., 2012). Whereas in *rbcL* locus, this sequences is more similar to that of HN02, PT01 and TH01 and gathered into a group (Figure 2b and 2e). A mega study in Atlantic forest also reported that *rbcL* was not able to identify any plant species (Vivas et al., 2014). This could be due to the low mutation rate of *rbcL* locus in the comparison to that of *matK* and *trnH-psbA (*Kellogg and Juliano, 1997*).*

Combination of multi locus barcodes could improve the classification of species and several studies have proven this idea. In our study, the combination of two barcode regions show the variation in species power resolution compared to the single use. The combination of *matK* and *trnH-psbA* showing best discrimination capacity since *Clinacanthus nutans* (XK) is classified in single branch (Figure 3b, 3e), whereas the combination of either *matK* + *rbcL* or *rbcL* + *trnH-psbA* revealed that *Clinacanthus nutans* is grouped to TH01, HN02 and PT01 accessions (Figure 3a, 3c, 3d, 3f). The combination of *rbcL* to either *matK* or *trnH-psbA* also did not increase the discrimination power of forest trees in early reported study (Vivas et al. 2014). Thus, the proposed combination from *matK* and *rbcL* of The Consortium for the Barcode of Life (CBOL) is not applicable for all plant species.

Although previous study shows the superior of combining these three markers in species classification such as Combretaceace (Gere et al. 2013); Actinidia (Weihong et al. 2018); the phylogeny combination gives result similarly to the combination of *matK*+*rbcL* and *rbcL*+ *trnH-psbA* show (Figure 4). In in study, XK is not distinguished from three hoan-ngoc accessions namely NH02, PT01, and TH01. The lower species discrimination when combining three loci was also previously reported in forest trees (Vivas et al. 2014).

The result shows that both ML and NL are more effective in species classification of hoan-ngoc. Similar result was reported from Yang and colleagues (2012) when studied *Calamus* species in China. Although both ML and NJ are commonly used in phylogenetic analysis, NJ can be easily performed in a short time with personal computer while ML is considered as professional method in phylogenetic analysis, since it could consider the possibility for all events happening simultaneously and produced the best tree, supported at higher probability in the comparison to other methods (Felsenstein, 1981).

Table 2: Species identification by using BLAST with three DNA barcode loci.

No	Sample	BLAST with *matK*	BLAST with *rbcL*	BLAST with *trnH-psbA*
1	BP01	*Pseuderanthemum carruthersii*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
2	TN01	*Pseuderanthemum carruthersii*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
3	TH01	*Pseuderanthemum carruthersii*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
4	HN02	*Pseuderanthemum carruthersii*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
5	HN01	*Pseuderanthemum carruthersii*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
6	SG04	*Pseuderanthemum carruthersii*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
7	SG03	*Pseuderanthemum carruthersii*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
8	PT01	*Pseuderanthemum carruthersii*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
9	LA01	*Cosmianthemum viriduliflorum*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
10	SG02	*Cosmianthemum viriduliflorum*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
11	SG01	*Cosmianthemum viriduliflorum*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
12	BT01	*Cosmianthemum viriduliflorum*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
13	TG01	*Cosmianthemum viriduliflorum*	*Ruttya fruiticosa*	*Pseuderanthemum atropurpureum*
14	XK	*Clinacanthus nutans*	*Clinacanthus nutans*	*Clinacanthus nutans*

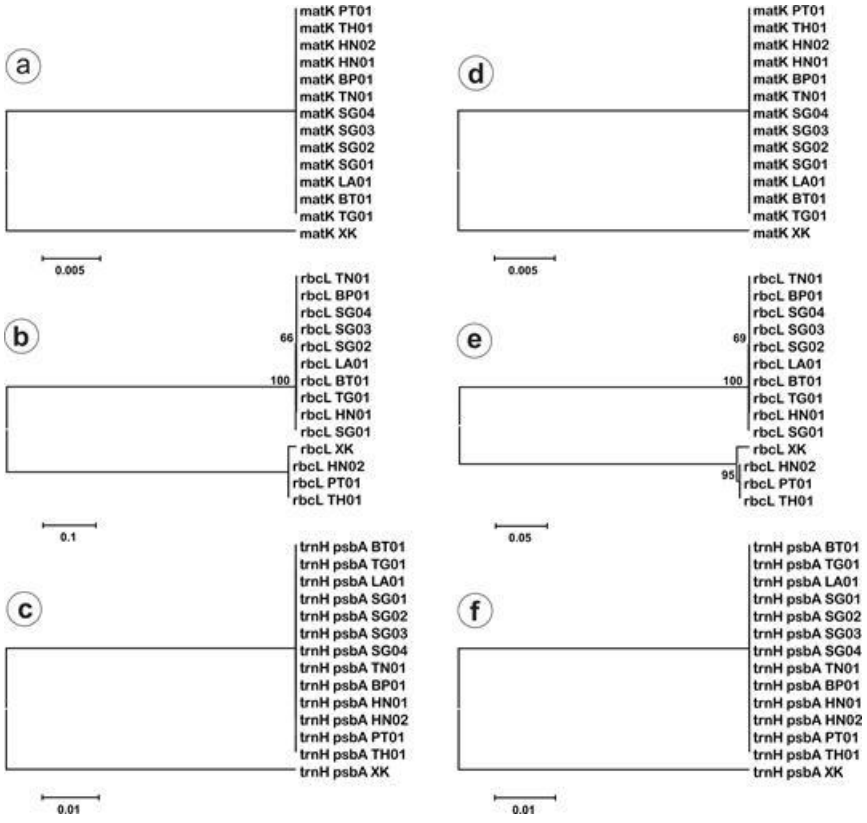

Figure 2: Phylogenetic tree developed by Maximum Likelihood (a,b,c) and Neighbour Joining (d,e,f) methods using single barcode locus.

Figure 3: Phylogenetic tree developed by Maximum Likelihood (a,b,c) and Neighbour Joining (d,e,f) methods using combination of two barcode loci.

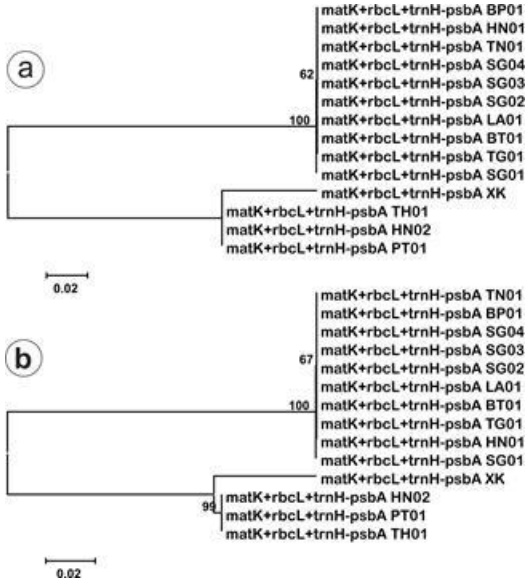

Figure 4: Phylogenetic tree developed by Maximum Likelihood (a) and Neighbour Joining (b) methods using combination of three barcode loci.

CONCLUSION

The obtained show the low genetic variation of collected hoan-ngoc accession within three DNA barcode loci consisting of *matK*, *rbcL* and *trnH-psbA*. *matK* and *trnH-psbA* barcode loci could be used as potential tool for hoan-ngoc identification, while *rbcL* is not effective to identify this herbal tree. The range of future studies should include more different relative species of hoan-ngoc which could find the higher reliable combined loci for identifying this herbal plant.

CONFLICT OF INTEREST

The authors declared that present study was performed in absence of any conflict of interest.

ACKNOWLEGEMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors would like to thank Ho Chi Minh University of Food Industry-Vietnam for supporting laboratory and research equipment.

AUTHOR CONTRIBUTIONS

Conceptualization: VTH; Methodology: HVT; Project administration: Supervision: VTH; Validation; Laboratory work: HVT, STB; Writing - original draft: VTH, STB; Writing - review and editing: VTH, STB. All authors read and approved the final manuscript.

REFERENCES

Amarasinghe APPR, Siriwardhane DAS, Samarasekera RR, Weerasena OVDSJ, 2013. Development of DNA Barcodes and Optimization of DNA Isolation and PCR Protocols for a Medicinal Plant *Plectranthus hadiensis*. Int J Sci Res 4(6): 2852-2856.

Bchir R, Djellouli AS, Zitouna N, Aurelle D, Pergent G, Pergent-Martini C, Langar H, 2019. Morphology and genetic studies of *Cymodocea* seagrass genus in Tunisian coasts. Phyton 88(2): 171-184.

Cuenoud P, Savolainen V, Chatrou LW, Powell M,

Grayer RJ, Chase MW, 2020. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbcL*, *atpB*, and *matK* DNA sequences. Am J Bot 89: 132-144.

Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical bulletin 19: 11-15.

Felsenstein J, 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J Mol Evol 17(6):368-76.

Ferri G, Alu M, Corradini B, Beduschi G, 2009. Forensic botany: species identification of botanical trace evidence using a multigene barcoding approach. Int J Legal Med 123: 395-401.

Gere J, Yessoufou K, Daru BH, Mankg LT, Maurin O, van der Bank M, 2013. Incorporating *trnH-psbA* to the core DNA barcodes improves significantly species discrimination within southern African Combretaceae. Zookyes 365: 127-147.

Gismondi A, Rolfo MF, Leonardi D, Rickards O, Canini A, 2012. Identification of ancient *Olea europaea* L. and *Cornus mas* L. seeds by DNA barcoding. C R Biol 335: 472-479.

Hasebe M, Omori T, Nakazawa M, Sano T, Kato M, Iwatsuki K, 1994. *rbcL* gene sequences provide evidence for the evolutionary lineages of leptosporangiate ferns. Proc Natl Acad Sci USA 91:5730–5734.

Huynh KD, 2008. Isolating some chemical ingredients in *Pseuderanthemum palatiferum* leaves. Sci J Can Tho Uni 9: 232-240.

Kellogg EA, Juliano ND, 1997. The structure and function of RuBisCO and their Implications for Systematic Studies. American Journal of Botany 84: 413–28.

Kress WJ, MacKechnie LM, Wiliams KJ, 2002. The phylogeny and a new classification of the gingers (Zingiberaceae): Evidence from molecular data. Am J Bot 89(10): 1682-1696.

Li HQ, Chen JY, Wang S, Xiong SZ, 2012. Evaluation of six candidate DNA barcoding loci in Ficus (Moraceae) of China. Mol Ecol Resour 12: 783-790.

Li Q, Wang X, Wang J, Su N, Zhang L, Ma Y, Chang Z, Zhao L, Potter D, 2019. Efficient ientification of Pulsatilla (Ranunculaceae) uing DNA barcodes and micro-morphological characters. Front Plant Sci 10:1196.

Nguyen QV, Eun JB, 2013. Antimicrobial activity of some Vietnamese medicinal plants extracts. Med Plant Res 7(35): 2597-2605.

Nio SA, Kolondam BJ, Talei TK, 2018. Evaluation of *matK* and *rbcL* genes as markers in DNA barcoding of *Codiaeum variegatum* (L.) Blume. Biosci Res 15(1): 192-198.

Padee P, 2009. Current information of medicinal plants: Pseuderanthemum palatiferum (Nees) Radlk. J Health Sci 18(1): 131–138.

Parmentier I, Duminil J, Kuzmina M, Philippe M, Thomas DW, Kenfack D, Chuyong GB, Cruaud C, Hardy OJ, 2013. How effective are DNA barcodes in the identification of African rainforest trees? PloS One 8(4):e54921.

Sikdar S, Tiwari S, Thakur VV, Sapre S, 2018. An *in silico* approach for evaluation of rbcL and matK loci for DNA barcoding of Fabaceae family. Int J Chem Stud 6(6): 2446-2451.

Sittisart P, Chitsomboon B, Kaminski NE, 2016. *Pseuderanthemum palatiferum* leaf extract inhibits the proinflammatory cytokines, TNF-a and IL-6 expression in LPS-activated macrophages. Food Chem Toxicol 97: 11-22.

Sogo A, Setoguchi H, Noguchi J, Jaffre T, Tobe H, 2001. Molecular Phylogeny of Casuarinaceae based on *rbcL* and *matK* Gene Sequences. Int J Plant Res 114: 459-464.

Tamura K, Stecher G, Peteron D, Filpski A, Kumar S, 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol 30(12): 2725–2729.

Trang NTP, Duc NM, Sinh NV, Triest L, 2015. Application of DNA barcoding markers to the identification of Hopea species. Genet Mol Res 14(3): 9181-9190.

Vivas CV, Moraes RCS, Alves-Araujo AA, Alves M, Mariano-Neto E, van den Berge C, Gaiotto FA, 2014. DNA barcoding in Atlantic Forest plants: What is the best marker for Sapotaceae species identification? Genet Mol Biol 37(4): 662-670.

Weihong B, Li D, Li X, 2018. DNA barcoding of Actinidia (Actinidiaceae) using internal transcribed spacer, *matK*, *rbcL* and *trnH-psbA*, and its taxonomic implication. N Z J Bot 4: 360-371.

Yang HQ, Dong YR, Gu ZJ, Lang N, Yang JB, 2012. A preliminary assessment of *matK*, *rbcL* and *trnH–psbA* as DNA barcodes for Calamus (Arecaceae) species in China with a note on ITS. Ann Bot Fenn 49: 319–330