Jurnal Teknologi

Full paper

Development *mat*K Gene as DNA Barcode to Assess Evolutionary Relationship of Important Tropical Forest Tree Genus *Mangifera* (Anacardiaceae) in Indonesia and Thailand

Topik Hidayat[a,b*], Adi Pancoro[c], Diah Kusumawaty[b], Wichan Eiadthong[d]

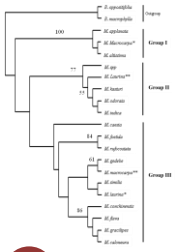
[a]*Department of Biological Science, Faculty of Bioscience and Bioengineering, Universiti Teknologi Malaysia (UTM), Malaysia*
[b]*Department of Biology Education, Indonesia University of Education (UPI), Bandung, Indonesia*
[c]*School of Life Sciences and Technology, Institute Technology of Bandung, Indonesia*
[d]*Department of Forest Biology, Faculty of Forestry, Kasetsart Universiy, Bangkok, Thailand*

*Corresponding author: topik@fbb.utm.my

Article history

Received :1 August 2012
Received in revised form :7 Sept. 2012
Accepted :1 October 2012

Graphical abstract



Abstract

MaturaseK (matK) gene of chloroplast DNA has been served as appropriate candidate to be a DNA barcode in angiosperms. Using this DNA marker, 19 species of genus *Mangifera*, one of ecologically important crop, collected from Indonesia and Thailand were analyzed. Phylogenetic analysis using parsimony method revealed that the gene could clasify *Mangifera* into three major groups, namely group I, II, and III. Moreover, the *mat*K barcode can identify *Mangifera* species that is originated from Thailand. Although this classification system is different with the previous system, it can provide a new information about *Mangifera* taxonomy. Results further exhibited that DNA sequences of the matK of two *Mangifera* species (*M. laurina* dan *M. macrocarpa*) are different between Indonesia and Thailand specimens.

Keywords: DNA barcode; *Mangifera*; *mat*K gene; parsimony; phylogenetic analysis

1.0 INTRODUCTION

The genus *Mangifera* L., one of the most important plant groups in deciduous forest and wets tropical rain forest including mountain forest, is one of the largest genera in family Anacardiaceae to which approximately 69 species have already described. The genus is mostly distributed in the tropical parts of Asia (India, Burma, Sri Lanka, Thailand, South Tropical China, Malaysia, Indonesia, Papua New Guinea, the Philippines, the Solomon Islands) but also extend to the Pacific Islands (Kostermans and Bompard, 1993). In spite of their economical importance, phylogenetic relationships among species within the genus have been poorly understood due to their extremely complicated vegetative and reproductive organs.

Previously, Bentham and Hooker (1862), Marchand (1869), Pierre (1897), and Kostermans and Bompard (1993) have revealed classification systems for the genus based upon floral characters. However, these characters were extremely complicated in the genus and subjected to parallelism (Yonemori *et al*., 2002), suggesting many taxonomic and phylogenetic problems still remain unresolved. Given the shortcomings of these characters, data obtained from nucleotide substitutions of appropriate molecules are preferable for clarifying phylogenetic relationships (e.g., Moritz and Hillis, 1996).

Methods for identifying species or group organisms by using short DNA sequences, known as ‘‘DNA barcodes,’’ have been proposed and initiated to facilitate biodiversity studies, identify juveniles, associate sexes, and enhance forensic analyses. It is proposed that maturaseK gene of chloroplast genome as potentially usable DNA regions for applying barcoding to flowering plants (Kress *et al*., 2005). The matK gene is frequently choosen by plant systematists because the region is a single copy gene and has enough variable sites of nucleotide substitution. Recently, the *mat*K gene has been widely used in phylogenetic inferences of various groups of plant (e.g. Ito *et al*., 1999; Ferguson and Sang, 2001; Raymond *et al*., 2002; Ebihara *et al*., 2005; Hidayat *et al*., 2005).

Using DNA sequences of the *mat*K gene, we have carried out phylogenetic analysis to clarify phylogenetic relationships among member of genus *Mangifera*. An understanding of the evolutionary relationships in this group may contribute in the field of plant systematics or ecology.

2.0 MATERIALS AND METHOD

A total of 19 species of *Mangifera* were collected from Indonesia and Thailand, plus two species of *Bouea*. Two members of genus *Bouea* (M9 and M13) were used as outgroup in phylogenetic analysis because based on previous research this genus was sister group to *Mangifera* (Yonemori *et al*., 2002). Detail information about the plant is summarized in Table 1.

DNA genome was extracted from fresh materials (young leaf or flower) or in the form of silica gel material using QIAGEN *Dneasy Mini Plant Kit* with slight modification. Amplification was conducted using four primers as seen in Fig. 1. Table 2 provides detail information about sequences of primer pairs.

Table 1 Plant materials, their geographic origins and codes used in this study

No.	Species	Origin	Code
1.	*Mangifera altissima* Blanco var bingloe	Indonesia	M18
2.	*Mangifera applanata* Kosterm.	Indonesia	M14
3.	*Mangifera foetida* Lour.	Indonesia	M17
4.	*Mangifera gedebe* Miq.	Indonesia	M10
5.	*Mangifera indica* L.	Indonesia	M11
6.	*Mangifera laurina* Bl.	Indonesia	M7
7.	*Mangifera macrocarpa* Bl.	Indonesia	M3
8.	*Mangifera odorata* Griff.	Indonesia	M16
9.	*Mangifera spp*	Indonesia	M12
10.	*Mangifera rufocostata* Kosterm.	Indonesia	M8
11.	*Mangifera similis* Auct.	Indonesia	M2
12.	*Mangifera caesia* Jack ex Wall	Indonesia	M5
13.	*Mangifera casturi* Kosterm.	Indonesia	M15
14.	*Mangifera macrocarpa* Bl.	Thailand	S1
15.	*Mangifera conchinchinensis* Englar	Thailand	S6
16.	*Mangifera flava* Evrard	Thailand	S3
17.	*Mangifera gracilipes* Hook.f.	Thailand	S2
18.	*Mangifera caloneura* Auct.	Thailand	S5
19.	*Mangifera laurina* Bl.	Thailand	S7
20.	*Bouea oppositifolia* (Roxb.) Meiss	Indonesia	M13
21.	*Bouea macrophylla* Griff.	Indonesia	M9

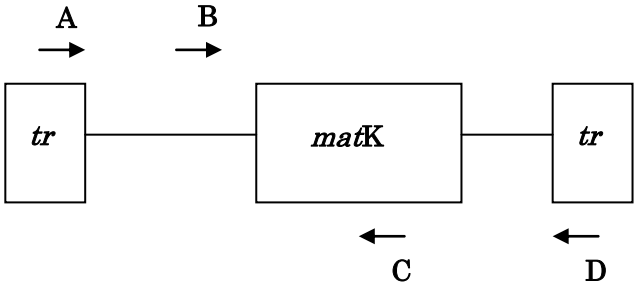


Figure 1 Strategy of amplification and sequencing of the *mat*K gene. A=*trn*K-5F, B=TAA-09F, C=TAA- 09R, dan D=*trn*K-2R. Two internal primers (B and C) were designed for this study

Table 2 Primers used in this study

Name	Sequences
trnK-5F	5' TGGGTTGCTAACTCATGG 3'
trnK-2R	5' AACTAGTCGGATGGAGTAG 3'
TAA-09F	5'GGTTTTCCCATGAGTAGATTATCG 3'
TAA-09R	5' CGAAGTAGACGAAGCTCTTGG 3'

For amplification, we used primer pairs A and D, whereas all primers were used once sequencing. PCR (Polymerase Chain Reaction) reaction included buffer PCR (1x), $MgCl_2$ (2-3mM), primers (@ 0,5 mM), enzyme Taq polymerase (1 U/uL), dNTPs Mix (1,6 mM), and DNA template (100-150 ng/uL). PCR was conducted according to Hidayat *et al*. (2005). PCR cycles include 1 cycle at 94°C (predenaturation) for 5 minutes; 30 cycles at 94°C (denaturation) for 30 second, 49°C (annealing) for 30 second, and 72°C (extension) for 2 minutes; and ended with 1 cycle at 72°C (final extension) for 8 minutes. PCR products were cloned into pGEM-T Easy (Promega) before sending them to Macrogen (Korea) for sequencing.

DNA sequences obtained from the *mat*K gene were aligned with Clustal X (Thompson *et al*., 1997) and then adjusted manually. Phylogenetic analyses based on the maximum parsimony criterion was performed using PAUP* version 4.0b10 (Swofford, 1998). All characters were equally weighted and unordered (Fitch, 1971). All the data sets were analysed by the heuristic search method with tree bisection-reconnection (TBR) branch swapping and the MULTREES option ON, ten replications of random addition sequences with the stepwise addition option, and all most parsimonious trees (MPTs) were saved. Evaluation of internal support of clades was conducted by the bootstrap analysis (Felsenstein, 1985) utilizing 1,000 replicates with TBR branch swapping and the MULTREES option OFF. Number of steps, consistency indices (CI) and retention indices (RI) were calculated on one of the MPTs in each analysis with the TREE SCORES command in PAUP*.

3.0 RESULTS

DNA extraction can be done using various type of DNA sources such as leaf, stem, flower, and seed. In this research, young leaf was used for DNA extraction to minimize contamination that can inhibit PCR amplification. High level of concentration (600 ng/uL in average) with good rasio ($\pm$ 1.750) was obtained. Size and border of matK gene for *Mangifera* were determined through comparative analysis in genebank (www.ncbi.nlm.nig.gov). The results indicated that size of matK gene in *Mangifera* is about 1500 bp.

Multiple alignment analysis was performed by using ClustalX (Thompson *et al*., 1997). The aligned *mat*K comprised 1,601 characters. Of these, 1,429 were constant and 51 were potentially informative. Reconstruction of phylogenetic tree (Fig. 2) using PAUP resulted in 23 MPTs with a length of 121 steps, CI of 0.852, and RI of 0.739.

4.0 DISCUSSION

The tree (Fig. 2) demonstrated that the genus was monophyletic and split into three major groups. Monophyletic nature of

Mangifera was supported by character of stoma, anomositic (Hidayat, unpublished data).

The three major groups found in this study is not consistent with previous classification system by Mukherjee (1953), Kostermans and Bompard (1993) based upon morphological characters, and even Yonemori *et al*. (2002) on the basis of DNA sequences of internal transcribed spacer (ITS) region. Number of plant materials used in this study is likely to be insufficient (only 19 from 69 recognized species). Further phylogenetic analysis therefore is desired using more extensive sampling.

However, this study has provided new information about taxonomy of *Mangifera*. As depicted in Fig. 2 *M. applanata*, *M. macrocarpa* (from Indonesia), and *M. altissima* were united (Group I), whereas *M. laurina* (form Thailand), *M. casturi, M. odorata,* and *M. indica* were closely related (Group II). Group III was housed by the rest of species. And *Mangifera* species which is originated from Thailand were placed within Group III (see Fig. 2). Unfortunately, no single synapomorphic character is found to support each group.

Moreover, this research has revealed that there are variation of *mat*K in *M. laurina* and *M. macrocarpa* which come from Indonesia and Thailand. As seen in Fig. 2, *M. laurina* (from Thailand) was separated from that of Indonesia (Group III; Thailand specimen in Group II). Similar situation has been found in *M. macrocarpa*: Thailand in Group III and Indonesia in Group I. Different nature between these two countries has driven the mutation in *mat*K, but this does not lead to shift the morphology. These are related with the ability of plant to adapt to the environment changes (Evans, 1975).

As mentioned, *mat*K gene is highly conserved (e.g. Ebihara *et al*., 2005; Hidayat *et al*., 2005). Mutation rate in this kind of gene is very slow. This is reflected by the small number of informative characters (only 51 from a total 1,601 characters) to build the tree. As consequence, bootstrap value in most branches of the tree are less than 50. Similar condition was found in other angiosperms (e.g. Raymond *et al*., 2002; Ebihara *et al*., 2005; Hidayat *et al*., 2005). A further analysis based on the phylogenetic scheme presented here will shed more light on overlooked characters.

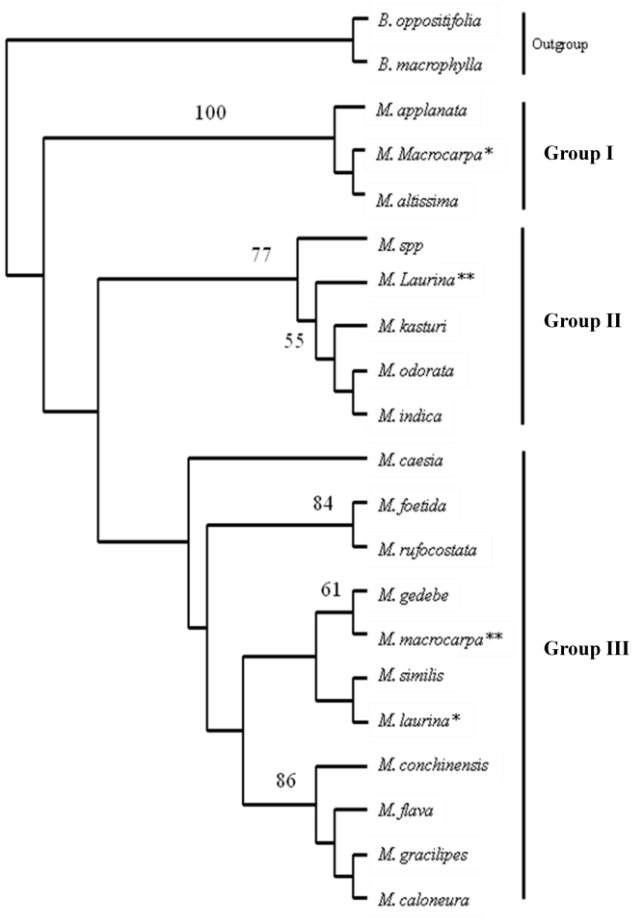


Figure 2 One of the 23 MPTs of *Mangifera* based on *mat*K gene. Bootstrap value of >50 are shown above each branch. * = Indonesia specimen; **= Thailand specimen. Species inside the box are originated from Thailand

■5.0 CONCLUSION

This study demonstrated that the *mat*K as a barcode classified the *Mangifera* into three major groups. Further the *mat*K barcode identified *Mangifera* species that is originated from Thailand. The *mat*K gene in two species, namely *M. laurina* and *M. macrocarpa*, were different between Indonesia and Thailand specimens. Due to we found limited utility of matK in *Mangifera*, it is suggested for employing another DNA region with more extensive sampling in the future.

Acknowledgement

We gratefully acknowledge Nisa, Puri, and Asri for their kind assistance during the completion of the study. We would like to thank Campbell Webb for fruitful discussion during preparation of this paper.

References

[1] Bantham, G. and Hooker, J. D. 1862. *Genera Plantarum* 1: 420.
[2] Ebihara, A., Ishikawa, H., Matsumoto, S., Lin, S. J., Iwatsuki, K., Takamiya, M., Watano, Y. and Ito, M. 2005. Nuclear DNA, Chloroplast DNA, and Ploidy Analysis Clarified Biological Complexity of the Vandenboschia Radicans Complex (Hymenophyllaceae) in Japan and adjacent areas. Am J Bot 92:1535-1547.
[3] Evans, L. T. 1975. *The Physiological Basis of Crop Yield.* London, UK: Cambridge University Press.
[4] Felsenstein, J. 1985. Confidence Limit on Phylogenies: An Approach Using the Bootstrap. *Evolution*. 39: 783–791
[5] Ferguson, D. and Sang, T. 2001. Speciation through Homoploid Hybridization Between Allotetraploids in Peonies (Paeonia). *Proc Nat Acad Sci*. 98: 3915–3919.
[6] Fitch, W. M. 1971. Toward Defining the Course of Evolution: Minimum Change for a Specific Tree Topology. *Syst Zool*. 20: 406–416.
[7] Hidayat, T., Yukawa, T. and Ito, M. 2005. Molecular Phylogenetics of Subtribe Aeridinae (Orchidaceae): Insight from Plastid Matk and Nuclear Ribosomal ITS Sequences. *J Plant Res.* 118: 271–284.
[8] Ito, M., Kamawoto, A., Kita, Y., Yukawa, T. and Kurita, S. 1999. Phylogenetic relationships of Amaryllidaceae based on matK sequences data. *J Plant Res*. 112: 207–216.

[9] Kostermans, A. J. G . H. and Bompard, J. M. 1993. *The Mangoes: Their Botany, Nomenclature, Horticulture and Utilization*. London: IBPGR Academic Press.
[10] Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A. and Janzen, D. A. 2005. Use of DNA Barcodes to Identify Flowering Plants. *PNAS.* 102: 8369–8374
[11] Marchand, L. 1869. Revision du Groupe des Anacardiacees. J.B. Bailliere, Paris.
[12] Moritz, C. and Hillis, D. M. 1996. Molecular Systematics: Context And Controversies. In: Hillis DM, Moritz C, Mable BK (Eds) Molecular Systematic. 2nd ed. Sinauer Associates, Sunderland, MA.
[13] Mukherjee, S. K. 1953. *Origin*, Distribution and Phylogenetic Affinity of the Species of Mangifera L. *J Linn Soc*. 55: 65–83.
[14] Pierre, L. 1897. *Flore Forestiere de la Cochinchine*. Doin, Paris.
[15] Raymond, O., Piola, F. and Sanlaville-Boisson, C. 2002. Inference of Reticulation in Outcrossing Allopolyploid Taxa: Caveats, Likelihood And Perspectives. *Trend in Eco and Evol*. 17: 3–6.
[16] Swofford, D. L. 1998. PAUP*4.0b10. Phylogenetic Analysis Using Parsimony (*And Other Methods). Version 4. Sinauer Associates, Sunderland, Massachussets.
[17] Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. 1997. The Clusstalx Windows Interface: Flexible Strategies For Multiple Sequences Alignment Aided by Quality Analysis Tools. *Nucl Acid Res.* 24: 4876–4882.
[18] Yonemori, K., Honsho, C., Kanzaki, S., Eidthong, W. and Sugiura, A. 2002. Phylogenetic Relationships of Mangifera Species Revealed by ITS Sequences of Nuclear Ribosomal DNA and a Possibility of Their Hybrid Origin. *Plant Syst Evol*. 231: 59–75.