Proceeding of the International Conference on Advanced Science, Engineering and Information Technology 2011

Hotel Equatorial Bangi-Putrajaya, Malaysia, 14 - 15 January 2011

ISBN 978-983-42366-4-9

Molecular Diversification and Phylogeny of *Mangifera* (Anacardiaceae) in Indonesia and Thailand

Topik Hidayat[1,2,*], Adi Pancoro[3], Diah Kusumawaty[1], and Wichan Eiadthong[4]

[*1] *Department of Biology Education, Indonesia University of Education (UPI)*
Jalan Setiabudi 229 Bandung 40154 Indonesia
Tel.:+(62)-(0)22-2001937, E-mail: topikhidayat@upi.edu

[2] *Department of Biological Science, Faculty of Bioscience and Bioengineering, Universiti Teknologi Malaysia (UTM)*
Skudai 81310, Johor Bahru, Johor, Malaysia
Tel.:+(6) 07-5532715, E-mail: topik@fbb.utm.my

[3] *School of Life Sciences and Technology, Institute Technology of Bandung*
Jalan Ganesha 10 Bandung 40132 Indonesia
Tel.:+(62)-(0)22-2511612, E-mail: adi@sith.itb.ac.id

[4] *Department of Forest Biology, Faculty of Forestry, Kasetsart Universiy*
50 Paholyothin Road, Jattujak, Bangkok 10903 Thailand
Tel.:+(66)-2-5790176, E-mail: fforwce@ku.ac.th

***Abstract*— Phylogenetic relationships among 19 *Mangifera* L. species of Indonesia and Thailand were analyzed by comparing sequences of maturase-K gene of chloroplast genome. Phylogenetic analysis using parsimony method revealed that the gene could clasify *Mangifera* into three major groups. 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 *mat*K of two *Mangifera* species (*M. laurina* dan *M. macrocarpa*) are different between Indonesia and Thailand specimens.**

***Keywords*— *Mangifera*, matK gene, Parsimony, Phylogenetic Analysis**

I. INTRODUCTION

Genus *Mangifera* L. 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 [1]. 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, references [1], [2], [3], and [4] have revealed classification systems for the genus based upon

floral characters. However, these characters were extremely complicated in the genus and subjected to parallelism, 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 [5]. Many genes and DNA sequences have been employed in phylogenetic studies of plants. Among them, maturase-encoding gene (*mat*K) of chloroplast DNA (cpDNA) are frequently choosen by plant systematists because the region are a single copy gene and have 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. [6], [7], [8], [9], [10]). Using DNA sequences of the *mat*K gene, we have carried out phylogenetic analysis to clarify phylogenetic relationships among member of genus *Mangifera*.

II. MATERIALS AND METHOD

A total of 19 species of *Mangifera* were collected from Indonesia and Thailand, plus two species of *Bouea*. Genus *Bouea* was used as outgroup in phylogenetic analysis because based on previous research this genus was sister group to *Mangifera* [11]. Detail information about the plant can be seen in Table I.

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 II provides detail information about sequences of primer pairs.

For amplification, we used primer pairs A-D, whereas all primers were used once sequencing. Component PCR (Polymerase Chain 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 following the procedure developed by [10], which include: 1 cycle at 94^{o}C (predenaturation) for 5 minutes; 30 cycles at 94^{o}C (denaturation) for 30 second, 49^{o}C (annealing) for 30 second, and 72^{o}C (extension) for 2 minutes; and ended with 1 cycle at 72^{o}C (final extension) for 8 minutes. All amplification 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 ([12], [20], [21]) and then adjusted manually. Phylogenetic analyses based on the maximum parsimony criterion was performed using PAUP* version 4.0b10 [13]. All characters were equally weighted and unordered [14]. 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 were conducted by the bootstrap analysis [15] 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*.

TABLE I
PLANT MATERIALS

No.	Species	Origin
1	*Mangifera altissima* Blanco var bingloe	Indonesia
2	*Mangifera applanata* Kosterm.	Indonesia
3	*Mangifera foetida* Lour.	Indonesia
4	*Mangifera gedebe* Miq.	Indonesia
5	*Mangifera indica* L.	Indonesia
6	*Mangifera laurina* Bl.	Indonesia
7	*Mangifera macrocarpa* Bl.	Indonesia
8	*Mangifera odorata* Griff.	Indonesia
9	*Mangifera spp*	Indonesia
10	*Mangifera rufocostata* Kosterm.	Indonesia
11	*Mangifera similis* Auct.	Indonesia
12	*Mangifera caesia* Jack ex Wall	Indonesia
13	*Mangifera casturi* Kosterm.	Indonesia
14	*Mangifera macrocarpa* Bl.	Thailand
15	*Mangifera conchinchinensis* Englar	Thailand
16	*Mangifera flava* Evrard	Thailand
17	*Mangifera gracilipes* Hook.f.	Thailand
18	*Mangifera caloneura* Auct.	Thailand
19	*Mangifera laurina* Bl.	Thailand
20	*Bouea oppositifolia* (Roxb.) Meiss *	Indonesia
21	*Bouea macrophylla* Griff. *	Indonesia

*= Outgroup

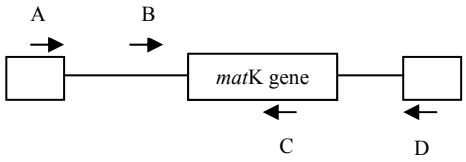

Fig. 1 Strategy of amplification and sequencing of the *mat*K gene. A=*trn*K-5F, B=TAA-09F, C=TAA- 09R, dan D=*trn*K-2R. Two internal primers were designed for this study.

TABLE II
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'

III. RESULTS AND DISCUSSION

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 as DNA sources to minimized contaminant that can inhibit amplification. DNA obtained here indicates high concentration (600 ng/uL in average) with good rasio ($\pm$ 1.750). Size and border of *mat*K gene for *Mangifera* were determined through comparative analysis in Genebank (http://www.ncbi.nlm.nih.gov/, [19]). The results indicated that size of *mat*K gene in *Mangifera* is around 1500 bp.

The first step in phylogenetic analysis is performing multiple alignment using ClustalX. The aligned *mat*K

comprised 1,601 characters. Of these, 1,429 were constant and 51 were potentially informative. Reconstruction of phylogenetic tree using PAUP resulted in 23 MPTs with a length of 121 steps, CI of 0.852, and RI of 0.739. The tree (Fig. 2) demonstrated that the genus is monophyletic and split into three major groups. Monophyletic nature of *Mangifera* is supported by character of stoma, anomositic [16].

The three major groups found in this study is not consistent with previous classification system by [17], [1], and even [11]. 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.

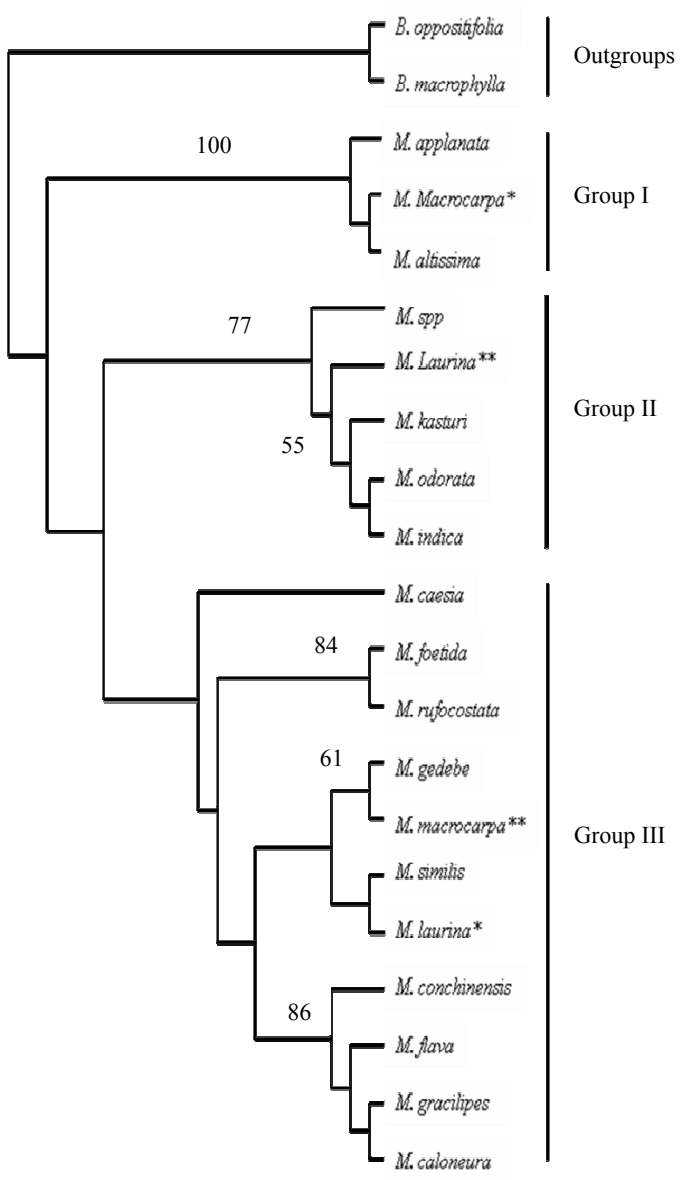

Fig. 2 One of the 23 MPTs of *Mangifera* based on *mat*K gene. Bootstrap value of >50 are shown above each branch. * = Indonesia specimen; **= Thailand specimen

However, this study has provide new information about taxonomy of *Mangifera*. As depicted in Fig. 2 *M. applanata*, *M. macrocarpa* (from Indonesia), and *M. altissima* are united (Group I), whereas *M. laurina* (form Thailand), *M. casturi, M. odorata,* and *M. indica* are closely related (Group II). Group III is housed by the rest of species. 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) is 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. All of these, of course, are related with the ability of plant to adapt to the environment change [18].

As mentioned, matK gene is highly conserved. 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 are found in other angiospermae (e.g. [8], [9], [10]). A further analysis based on the phylogenetic scheme presented here will shed more light on overlooked characters.

IV. CONCLUSION

This study demonstrated that the *mat*K can classify the *Mangifera* into three major groups. This classification system are quite different with previous system. The *mat*K gene in two species, namely *M. laurina* and *M. macrocarpa*, are different between Indonesia and Thailand specimens. Due to we found limited utility of *mat*K in *Mangifera*, it is suggested for employing another DNA region with more extensive sampling in the future.

ACKNOWLEDGMENT

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

REFERENCES

[1] A.J.G.H. Kostermans and J.M. Bompard, *The mangoes: Their botany, nomenclature, horticulture and utilization*, London: IBPGR Academic Press, 1993.

[2] G. Bentham and J.D. Hooker, "The Mangifera," *Genera Plantarum* vol. 1, pp. 314-420, 1862.

[3] L. Marchand, *Revision du Groupe des Anacardiacees*, Paris: J.B. Bailliere, 1869.

[4] L. Pierre, *Flore Forestiere de la Cochinchine*, Paris: Doin, 1897.

[5] C. Moritz and D.M. Hillis, *Molecular Systematics: context and controversies*, 2nd ed., D.M. Hillis, C. Moritz, B.K. Mable, Eds. Massachusetts: Sinauer Associates, 1996.

[6] M. Ito, A. Kamawoto, Y. Kita, T. Yukawa, and S. Kurita, "Phylogenetic relationships of Amaryllidaceae based on matK sequences data," *J Plant Res*, vol. 112, pp. 207-216, 1999.

[7] D. Ferguson and T. Sang, "Speciation through homoploid hybridization between allotetraploids in peonies (Paeonia)," *Proc Nat Acad Sci*, vol. 98, pp. 3915-3919, 2001.

[8] O. Raymond, F. Piola, and C. Sanlaville-Boisson, "Inference of reticulation in outcrossing allopolyploid taxa: caveats, likelihood and perspectives," *Trend in Eco and Evol*, vol. 17, pp. 3-6, 2002.

[9] A. Ebihara, H. Ishikawa, S. Matsumoto, S.J. Lin, K. Iwatsuki, M. Takamiya, Y. Watano, and M. Ito, "Nuclear DNA, Chloroplast DNA, and ploidy analysis clarified biological complexity of the Vandenboschia radicans Complex (Hymenophyllaceae) in Japan and adjacent areas," *Am J Bot*, vol. 92, pp. 1535-1547, 2005.

[10] T. Hidayat, T. Yukawa, and M. Ito, "Molecular phylogenetics of subtribe Aeridinae (Orchidaceae): insight from plastid *mat*K and nuclear ribosomal ITS sequences," *J Plant Res*, vol 118, pp. 271-284, 2005.

[11] K. Yonemori, C. Honsho, S. Kanzaki, W. Eidthong, and A. Sugiura, "Phylogenetic relationships of Mangifera species revealed by ITS sequences of nuclear ribosomal DNA and a possibility of their hybrid origin," *Plant Syst Evol*, vol. 231, pp. 59-75, 2002.

[12] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins, "The ClustalX windows interface: flexible strategies for multiple sequences alignment aided by quality analysis tools," *Nucl Acid Res*, vol. 24, pp. 4876-4882, 1997.

[13] D.L. Swofford, *PAUP*4.0b10. Phylogenetic analysis using parsimony (*and other methods)*, Massachusetts: Sinauer Associates, 1998.

[14] W.M. Fitch, "Toward defining the course of evolution: minimum change for a specific tree topology," *Syst Zool*, vol. 20, pp. 406-416, 1971.

[15] J. Felsenstein, "Confidence limit on phylogenies: an approach using the bootstrap," *Evolution*, vol. 39, pp. 783-791, 1985.

[16] T. Hidayat, T. Rahmi, and Kusdianti, "Diversity of stomata in genus Mangifera," Indonesia University of Education (UPI), Bandung, Tech. Rep, 2009

[17] S.K. Mukherjee, "Origin, distribution and phylogenetic affinity of the species of Mangifera L.," *J Linn Soc Bot*, vol. 55, pp. 65-83, 1953.

[18] L.T. Evans, *The physiological basis of crop yield*, London: Cambridge University Press, 1975.

[19] D.A. Benson, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, and D.L. Wheeler, "GenBank," *Nucleic Acids Res*., vol. 36, pp. D25-30, 2008.

[20] F. Jeanmougin, J.D. Thompson, M. Gouy, D.G. Higgins, and T.J. Gibson, "Multiple sequence alignment with Clustal X," *Trends Biochem Sci*., vol. 23, pp. 403-405, 1998

[21] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, and D.G. Higgins, "Clustal W and Clustal X version 2.0.," *Bioinformatics*, vol. 23, pp. 2947-2948, 2007.