Phylogenetic analysis of Dipterocarpaceae in Ketambe Research Station, Gunung Leuser National Park (Sumatra, Indonesia) based on \textit{rbcL} and \textit{matK} genes

ESSY HARNELLY$^{1,*}$, ZAIRIN THOMY$^2$, NIR FATHIYA$^{3,**}$

Program in Magister Biology, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala. Jl. Tgk Chik Pante Kulu No. 5, Kopelma Darussalam, Syiah Kuala, Banda Aceh 23111, Aceh, Indonesia. Tel.: +62-651-8012505, *email: eharnelly@gmail.com, **nirfathiya@gmail.com

Abstract. Harnelly E, Thomy Z, Fathiya N. 2018. Phylogenetic analysis of Dipterocarpaceae in Ketambe Research Station, Gunung Leuser National Park (Sumatra, Indonesia) based on \textit{rbcL} and \textit{matK} genes. Biodiversitas 19: 1074-1080. Gunung Leuser National Park has several Research Stations. Ketambe is one of the Research Stations which is located in Aceh Tenggara district, Aceh province. One of the timber plants family which is abundant in Leuser Mountain is Dipterocarpaceae. However, the data of species and relationship of Dipterocarpaceae in Ketambe Research Station is lack and not available. The aim of the study is to analyze the phylogenetic of Dipterocarpaceae based on \textit{rbcL} and \textit{matK} gene in Ketambe Research Station. This research was carried out from July 2015 to August 2016 in Ketambe Research Station and Forestry and Forest Genetics Laboratory of Molecular, Bogor Agricultural University. The method used quadrat sampling technique with purposive sampling and experimental laboratory consisting of DNA extraction, PCR, electrophoresis, and sequencing. The data analysis was done using BioEdit and MEGA6. The results showed that based on morphological identification, there were five Dipterocarpaceae species found namely: \textit{Parashorea lucida}, \textit{Shorea parvifolia}, \textit{Shorea lepidota}, \textit{Shorea johorensis}, and \textit{Hopea dryobalanoides}. The phylogenetic tree based on \textit{rbcL} gene showed that there were two monophyletic groups, the first group was \textit{S. johorensis}, \textit{S. lepidota}, and \textit{H. dryobalanoides}; and the second group consisted of \textit{S. parvifolia} and \textit{P. lucida}. The phylogenetic tree reconstruction based on \textit{matK} gene showed that \textit{Shorea parvifolia} and \textit{S. johorensis} were separated in two different monophyletic groups.

Keywords: Dipterocarpaceae, Ketambe Research Station, \textit{matK} gene, \textit{rbcL} gene

INTRODUCTION

Ketambe Research Station is one of Research Stations in Gunung Leuser National Park in northern Sumatra, Indonesia with an area of 450 ha. Administratively, it locates in Aceh Tenggara District. According to Rijksen (1978), Ketambe Research Station is a lowland rainforest which has a complete stratification and rarely dominated by one plant species so that it creates a diverse and complex ecosystem. This forest consists of 332 trees species from 179 genera and 68 families.

Based on data from Bappenas (1993), one of the plants in Gunung Leuser National Park is Dipterocarpaceae family. Dipterocarpaceae is a pantropical plant widely used in the field of timber and commercially valuable. According to Alrasyid et al. (1991), the most Dipterocarpaceae exist in wet climates, high humidity, and grow at an altitude of 0-800 meters above sea level with rainfall 2,000 mm/year.

Dipterocarpaceae belong to the most important plant species in the tropical rainforest, both ecologically and economically. Ecologically, Dipterocarpaceae is a key element of various lowland forests. This family is generally large trees which reach the top of the canopy and emergent trees (Newman, et al. 1999). Currently, the dipterocarps predominate in the international tropical timber market, because it has the best wood quality. In addition, the non-timber products of dipterocarp are used by some wildlife in the forest for their survival (Panayotou and Ashton 1992).

Data of relationship Dipterocarpaceae species in Ketambe Research Station is not available. Phylogenetic analysis is one of the most commonly used methods in systematic to understand the diversity of living things through reconstruction of relationship. Along with the rapid advancement of molecular biology, data of DNA have been used in many phylogenetic studies to get more accurate information (Hidayat and Pancoro 2008).

The basic idea of DNA sequences utilization in phylogenetic studies is that occur change of nucleotide base over time. Therefore, it can be estimated the rate of evolution and reconstructed evolutionary relationships between one group of organisms with another. The purpose of phylogenetic is to construct the relationships between organisms and to estimate the differences that occur from one ancestor to the offspring (Dharmayanti 2011).

One source of DNA characters for Phylogenetic analysis is the gene from the chloroplast genome (cpDNA), \textit{rbcL} and \textit{matK} genes. The \textit{rbcL} gene is a gene encoding a large subunit of ribulose 1,5 bisphosphate carboxylase (Rubisco or RuBPCase) which is important for photosynthesis. The sequence of \textit{rbcL} gene data is extensively used in the reconstruction of the whole seed plants phylogeny because it has a fairly conservative level of evolution (Doebley et al. 1990). While the \textit{matK} gene is more variable than the \textit{rbcL} gene and useful for tree-level phylogenetic analyses.
a gene encoding the maturase enzyme subunit K. The length of \textit{matK} gene region can produce approximately 1500 bp (base pair). The \textit{matK} gene is used in many phylogenetic studies because the accuracy is more specific at the species level. In plant systematics, \textit{matK} appears as a valuable gene because it has a high phylogenetic signal than another gene (Muller et al. 2006).

Data collection on the diversity of plant species is important for research and practical purposes in the future as well as to offset the loss rate of biodiversity. Therefore, further research is needed to get accurate data diversity of Dipterocarpaceae in Ketambe Research Station.

MATERIALS AND METHODS

Study area

The research was conducted in Ketambe Research Station, Gunung Leuser National Park, Aceh Tenggara, Aceh Province, Indonesia (Figure 1) and Forestry and Forest Genetic Laboratory of Molecular, Bogor Agricultural University, West Java, Indonesia. The research was carried out from July 2015 to August 2016.

Samples collection

Samples of Dipterocarpaceae were collected from Ketambe Research Station, Gunung Leuser National Park, Aceh Tenggara (Figure 1). Samples were collected using Quadrat Sampling Technique. Intake of vegetation data by sampling plot was done by purposive sampling. There were 25 plots, with the area of each plot 20 m x 20 m (Indriyanto 2006). The total area of the plots was 1 ha (10,000 m²). Samples consisted of leaves from sampling stage, pole stage, or tree of Dipterocarpaceae.

Three individuals per species were collected represent Ketambe Research Station location. The three individual samples had the same ID number, and they were numbered individually. Three sets of specimen leaves were collected from each individual sample: (i) two sets of leaves for the herbarium (leaves must have important taxonomic characters such as leaf tip, leaf surface, stipule, ptiole/leaf stalk). Specimens for herbarium were put on paper sheets and moistened with 70% alcohol. The specimens were dried and glued on herbarium paper pairs. Specimens were labeled information such as ID numbers, collector name, collection date, and taxonomy. Identification of samples was conducted using Dipterocarpaceae identification book; and (ii) a set of leaves for DNA extraction (soft, fresh, and young leaf tissue). Specimen for DNA extraction: leaf sheets were cleaned with the dry cloth. Specimens were placed into an existing sac containing another sac filled with silica gel (ratio of silica gel 5-10: 1). The ID number of the specimens were written on the outside of the bag using a permanent marker. All packets/bags of the specimen were stored in containers.

Figure 1. Map of research sites in Ketambe Research Station, Gunung Leuser National Park, Aceh Tenggara, Indonesia
Table 1. Primer data in this study

<table>
<thead>
<tr>
<th>DNA region</th>
<th>Primer</th>
<th>Sequence (5'—3')</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>rbcL</td>
<td>rbcL F</td>
<td>ATGTCACCCAACAACAGAGACTAAA</td>
<td>Kress and Erickson (2007)</td>
</tr>
<tr>
<td></td>
<td>rbcL R</td>
<td>GFAAACATCAATCCACCCACCG</td>
<td>Kress and Erickson (2007)</td>
</tr>
<tr>
<td>matK</td>
<td>matK 472F</td>
<td>CCCRTCATCTGGAAATCTTGGTC</td>
<td>Yu et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>matK 1248R</td>
<td>GCCTRATATAAGGGAGATTGCTG</td>
<td>Yu et al. (2011)</td>
</tr>
</tbody>
</table>

Molecular analysis

DNA extraction

DNA extraction was carried out using Cetyltrimethyl Ammonium Bromide (CTAB) method developed by Doyle and Doyle (1987). Young leave of 200 mg was ground in a mortar with liquid nitrogen. The leaf powder was put into two mL tubes, 500 extraction buffer solutions and 100 μL polyvinylpyrrolidone (PVP) solution were added. The mixture was vortex and then incubated in a water heater (water bath) for 60 minutes at a temperature of 65°C and every 15 minutes once reversed. After cooling to room temperature for 15 minutes, the mixture was added with 500 μL chloroform-isoamyl alcohol (24:1) and centrifuged for 10 minutes at 10,000 rpm. The upper layer (water phase/supernatant) was separated from the organic phase by using the micropipette into the new tube. Chloroform-isoamyl alcohol was added twice. The Supernatant was added with 500 μL cold isopropanol and NaCl of 300 μL. Samples were incubated overnight in the freezer. The precipitation result was centrifuged at 10,000 rpm for 10 minutes. The DNA pellet was washed twice using 96% ethanol of 300 μL and dried in a desiccator for 15 minutes. The DNA was stored at -20°C in the freezer.

Polymerase Chain Reaction (PCR)

The final conditions of each PCR reaction were 16 μL consisted of 8 μL Green GoTaq DNA polymerase, 2 μL Nuclease-Free Water, 2 μL primer (forward), 2 μL primer (reverse), and 2 μL diluted DNA (1 μL pure DNA: 99 μL aquabides). The primers were rbcL and matK (Table 1). The temperature setting of the thermocycler was begun with the initial denaturation at 95°C for 4 minutes and then proceeds in 35 cycles with denaturation at 94°C for 30 seconds, annealing at 58°C (for rbcL) and temperature 56 ºC (for matK) for 1 minute, extension at 72°C for 1 minute and rest at 4°C (Kristina et al. 2007).

Gel electrophoresis

The result of PCR (DNA) was visualized by 1% agarose gel electrophoresis. Gel electrophoresis procedure consists of making agarose gel, sample loading into the gel, running electrophoresis, and observation of electrophoresis running with UV transilluminator. In the process of making 1% agarose, 0.66 grams agarose was weighed and mixed with 33 mL TAE 1x (for mold 17-25 wells) in an Erlenmeyer. The mixture was put into the microwave and boiled for 2 minutes. The agarose gel was dyed with 1 μL gel red dye and poured into an electrophoresis mold. The hardened gel was inserted into electrophoresis chamber and added with TAE running buffer. The DNA samples and DNA ladder were separately mixed with loading buffer (6 μL + 2 μL). Mixing was done on the parafilm, and each mixture was put into well. Gel electrophoresis was done by using electric current with 100-volt voltage for 45 minutes. For the electrophoresis of PCR products (amplicons), 3 μL of 1 kb DNA ladder was placed at the first well hole, and 3 μL of the amplicones were added into the next wells. The process of running electrophoresis was 30 minutes with a voltage of 100 volts. The DNA profiles were observed using UV transilluminator.

Sequencing

The nucleotide sequence of the amplicon was identified using the Sanger method carried out by 1st BASE Sequencing INT in Malaysia. The sequencing process was done twice with different directions (forward and reverse). The sequencing data were used for the construction of phylogenetic trees.

Data analysis

The result of sequencing was analyzed with the following stages: (i) Sequence alignment using Bioedit program (Hall 1999), (ii) The result of sequence alignment was used to develop phylogenetic tree by Neighbor-Joining (NJ) method with MEGA (Molecular Evolutionary Genetics Analysis) version 6 (Tamura et al. 2011), (iii) The reliable test of the tree was done by the bootstrap method 1000 times.

RESULTS AND DISCUSSION

DNA extraction

DNA extraction is a method of separating DNA from other cell components. The extraction of DNA Dipterocarpaceae was performed to obtain DNA from the genome total of Dipterocarpaceae which were used as DNA template for PCR amplification process. The DNA extraction used CTAB buffer solution as cell wall degradation because it has advantages instead of other methods, i.e., it is easy to do, the possibility of DNA degrading enzymes is smaller than other methods, and can produce a large quantity of DNA (Demeke et al. 2009). The quality of genomic DNA was observed from the purity of the extraction result using agarose gel electrophoresis (Figure 2).
Based on the visualization of electrophoresis results (Figure 2), there were DNA bands of the fifteen DNA samples. The DNA bands on samples 1, 2, 3, 4, 5, and 6 look quite thick than other samples. The thinness of the DNA band indicates the low of DNA concentration. This is due to that incomplete lysis in DNA extraction process, or some supernatant containing DNA genomic was wasted (Restu et al. 2012). In addition, all DNA bands were smear, showing the poor DNA quality. This might be due to the incomplete lysis process like the absence of RNase A and proteinase K. The addition of RNase A in purification and proteinase K after incubation resulted in the lower smear intensity of the DNA (Utami 2012). According to Tiwari et al. (2012), the CTAB extraction method could be modified, by adding more concentrations of NaCl, Ethylene Diamine Tetra Acetic acid (EDTA), and mercaptoethanol to enhance the CTAB extraction and purification activity. Additional quantities of NaCl and mercaptoethanol enhanced the DNA extraction and proteins degradation respectively, while an increased concentration of EDTA protected DNA. They also increased the water bathing time and temperature for effective extraction.

### Amplification of the rbcL and matK genes

Amplification of the rbcL and matK genes was performed to multiply DNA sequence from DNA genome of Dipterocarpaceae by targeting the rbcL and matK genes. The amplification stage consists of 35 PCR cycles with amplification products $2^n$ (n = 35) DNA strands. The result of electrophoresis is presented in Figure 3 and Figure 4. The DNA bands of the 15 samples were between the 500-750bp (base pair) (Figure 3). All of the samples showed the good quality with thick DNA bands indicating that the processes of rbcL gene amplification of 15 samples were successfully done. According to Kress and Erickson (2007), a pair of primer is a universal primer for recognizing regions in the rbcL gene of Angiospermae plant and this primer show a high level of universality in land plants.

Amplification of the matK gene showed that from the total of 15 samples, only 8 samples exhibited the existence of DNA bands, i.e., samples 3, 4, 5, 6, 9, 10, 11, and 12 (Figure 4). It might be due to the low amplification and sequencing rate of the matK sequence affecting the PCR amplification results (Yu et al. 2011). In addition, the universality of matK primers was reported to be low in some studies (Sass et al. 2007; Fazekas et al. 2008; Ford et al. 2009; Kress et al. 2009). The DNA sample sizes were between 500-750 bp.

### Phylogenetic tree reconstruction

**Phylogenetic analysis of rbcL gene**

The construction of phylogenetic tree was conducted using MEGA 6 program with Neighbor-Joining (NJ) method. The construction of phylogenetic tree aims to determine the relationship of among several
Dipterocarpaceae species. In this study, the phylogenetic tree was statistically tested using the bootstrap method of 1000 repetitions (Salemi and Vandame 2003). According to Claverie and Notredame (2007), bootstrap method is a randomization method of characters into new data sets with the same number of characters as the initial data set and the phylogenetic tree. In this study, only 11 amplicons were successfully sequenced, while the rest showed poor sequencing results.

The Dipterocarpaceae samples in this study were divided into two groups (clade) (Figure 5). The first group has bootstrap values 63, consisted of *Shorea johorensis*, *Shorea lepidota*, and *Hopea dryobalanoides*, and the second group has bootstrap values 95 consisted of *Parashorea lucida* and *Shorea parvifolia*. Each group formed a monophyletic group. A group of species is monophyletic if all of the species present in the branches come from one common ancestor (Campbel et al. 2003).

In this phylogenetic tree, *S. johorensis* has a closer relationship with *S. lepidota* than *S. parvifolia*. This is in accordance with the research by Cao et al. (2009), *S. johorensis* formed a separate group with *S. parvifolia*. In addition, chloroplast DNA analysis by Tsumura et al. (2011), also explained that *S. johorensis* and *S. parvifolia* each formed the monophyletic group.

Two samples of *S. lepidota* (sample 1 and 3) were in a different branch with sample 2. However, the three samples of *S. lepidota* were still in a monophyletic group. According to Olivar et al. (2014), *rbcL* has low interspecific but exceedingly high intraspecific divergence. Hence, *rbcL* was not able to group similar species in these samples. In addition, Li (2011) stated that plastid DNA barcoding loci might not be able to discriminate between closely related species within a genus. They posited that the low performance of *rbcL* in this parameter is attributed to the low mutation rate of bases in plastid DNA markers.

*Shorea parvifolia* showed the closer relationship with *P. lucida*. This is in accordance with a study by Gamage et al. (2006) and Indrioko (2005), chloroplast DNA analysis explained that *Parashorea* is the same group with *Shorea*. In addition, molecular data from Tsumura et al. (1996), Kajita et al. (1998), and Kamiya et al. (2005) also explained that *Parashorea* is relatively close to several species of *Shorea*. It is clear in this study that *Parashorea* is in the same group with *Shorea* not with *Hopea*. But, the opposite result was reported by Kamiya et al. (2005), *P. lucida* and *S. parvifolia* were separate and not belong to a monophyletic group.

The interesting result in this study, *H. dryobalanoides* was grouped with *S. johorensis* and *S. lepidota*. Molecular analysis using *trnL-trnF* and ITS regions by Yulita et al. (2005) observed that the molecular analysis has not been able to separate *Shorea* from *Hopea* into different monophyletic groups. Therefore, it is strongly suspected that the *Hopea* group has the same common ancestor with *Shorea*. Phylogenetic analysis based on the *trnL-trnF*, *trnL*, and *matK* from some Dipterocarpoidea subfamily species showed that *Hopea* formed a monophyletic group with several genera of *Shorea* (Gamage et al. 2006). In addition, *PgC* (Kamiya et al. 2005) and *rbcL* (Dayanandan et al. 1999) tree topologies also showed that *Hopea* is in a monophyletic with several species in the *Shorea* tribe.

Flower morphology data from the *Hopea* genus and several *Shorea* species showed similarities, i.e., the flower has the corolla with urceolate shape and stamen with an additional acicular linkage (Dayanandan et al. 1999). In addition, they also have different unique morphological characters (Kamiya et al. 2005). Therefore, it is assumed that the Dipterocarpoidea species have not yet shown differences in molecular genus levels even though they have evolved into species with several different morphological characters (Gamage et al. 2006).

*Monotes katangensis* (sequence from GeneBank database) is an outgroup because it has a distant relationship with Dipterocarpaceae research samples. According to Hidayat and Pancoro (2008), in the analysis of phylogenetics, outgroup lead to the polarization of characters or characteristics, namely apomorphic and plesiomorphic characters. Apomorphic characters are the changed and derived characters which were found in the ingroup (species which were studied), whereas the plesiomorphic character is the primitive character was found in the outgroup. The synapomorphic character is a derived character in the monophyletic group.

**Phylogenetic analysis of matK gene**

The reconstruction of phylogenetic trees based on the *matK* gene is presented in Figure 6. In this study, only 4 amplicons were successfully sequenced, while the rest showed poor sequencing results. There were two monophyletic groups namely group 1 and group 2 with the support of a very large (100) bootstrap value. The high bootstrap value indicates the more stable grouping in the phylogenetic tree. *S. johorensis* and *S. parvifolia* each formed a separate monophyletic group. This is in accordance with *rbcL* gene analysis. In addition, Tsumura et al. (2011) and Cao et al. (2009) explained that *S. johorensis* and *S. parvifolia* were not in a monophyletic group. *Monotes madagascariensis* (sequence from GeneBank database) was an outgroup because it has a distant relationship with Dipterocarpaceae research samples.

**Figure 5.** The phylogenetic tree of Dipterocarpaceae based on *rbcL* gene using the neighbor-joining method
In conclusion, reconstruction of the phylogenetic tree based on the rbcL gene showed that there were two monophyletic groups, the first group consisted of *S. johorensis*, *lepidota Shorea*, and *H. dryobalanoides* and the second group consisted of *S. parvifolia* and *P. lucida*. In addition, the reconstruction of the phylogenetic tree based on the matK gene showed that *S. johorensis* and *S. parvifolia* were separated in two different monophyletic groups.

**ACKNOWLEDGEMENTS**

This project was funded by Directorate General of Higher Education (DIKTI), Ministry of Research, Technology and Higher Education, Indonesia through the fundamental scheme. The authors also thank Laboratory of Forest Genetics and Molecular Forestry, Bogor Agricultural University, Gunung Leuser National Park Office, and Leuser Conservation Forum which have been very helpful in conducting this research.

**REFERENCES**


Doebely J, Durbin M, Goldberg EM, Clegg MT, Ma DP. 1990. Evolutionary analysis of the large subunit of carboxylase (rbcL) nucleotide sequence among the grasses (Gramineae). Evolution 44: 1097-1108.


