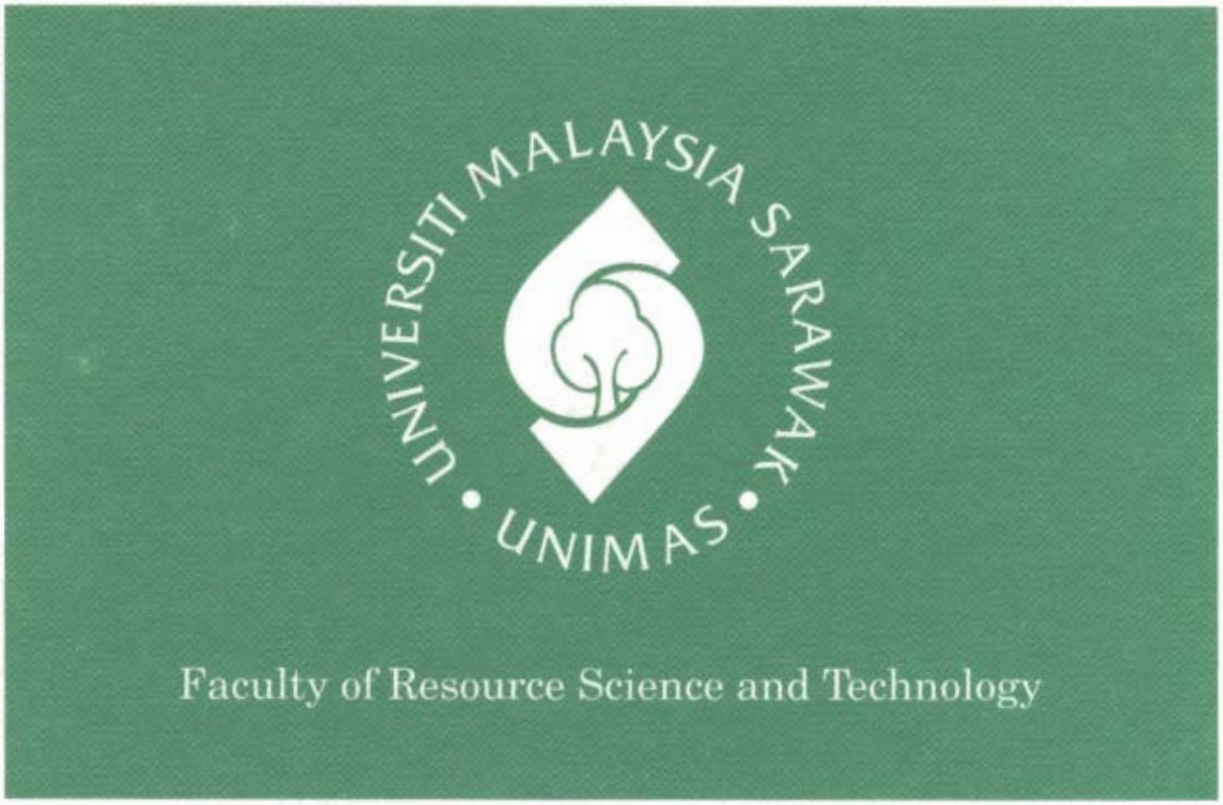

IDENTIFICATION OF GENOMIC MARKERS FOR BELIAN USING RAPD

Tan Sia Hong

Bachelor of Science with Honours
(Resource Biotechnology)
2005

IDENTIFICATION OF GENOMIC MARKERS FOR BELIAN USING RAPD

TAN SIA HONG

This project is submitted in partial fulfillment of the requirements for the degree
of
Bachelor of Science with Honours
(Resource Biotechnology)

Faculty of Resource Science and Technology
UNIVERSITI MALAYSIA SARAWAK

2005

ACKNOWLEDGEMENTS

This project was performed in partial fulfillment for the Degree of Bachelor of Science with Honors. I would like to express my deepest gratitude to the Faculty of Resource science and Technology, UNIMAS, Sarawak and my respectful supervisor, Dr. Hairul Azman Roslan and my co-supervisor, Assoc. Prof Dr. Isa Ipor for their guidance, support, encouragement and valuable discussion. Thanks to Dr. Ho Wei Seng for his advise and comments toward my project. I am also grateful to Assoc. Prof Dr. Sim Soon Liang for so kindly shares her experience and support in various ways that helps me to complete this project. Also thanks to Plant Genomic Lab for provided the NTSYSpc software for data analysis in this study.

Besides that, I would like to thank all our lab assistants especially Encik Azis Ajim, all master students especially Lau Ee Tiing, and all my fellow coursemates for helping me during the process of completing this final year project. Not forgotten my family, thanks for their support, love and encouragement.

IDENTIFICATION OF GENOMIC MARKERS FOR BELIAN USING RAPD

Tan Sia Hong

Program Resource Biotechnology
Faculty Resource Science and Technology
University Malaysia Sarawak

ABSTRACT

Belian refer to the most valuable hardwood timber of Borneo which is produced from two different species of two monotypic genera namely *Eusideroxylon zwageri* and *Potoxylon melagangai*. Both genera are belong to the family Lauraceae from order Laurales. These two species of belian are very closely related and difficult to distinguish morphologically. Random Amplified Polymorphic DNA (RAPD) has been used to identify the genomic markers for both Belian species. Genomic DNA for both species was successfully isolated using SDS (Sodium Dodecyl Sulfate) method with modification. The DNA samples were screened for RAPD markers using 20 decamer random primers. Only one primer namely OPD18 was found to be suitable for generating good and consistent RAPD profile for both species. The RAPD profile is reproducible and the bands produced ranged from 500bp to 2000bp. Dendogam constructed was clearly divided the two species of belian into two cluster Based on the RAPD profile, there are differences in banding pattern for both species. This study shows that RAPD is useful in identify genomic markers for closely related species.

Key words: Genomic markers, *Eusideroxylon zwageri*, *Potoxylon melagangai*, RAPD.

ABSTRAK

Belian merujuk kepada kayu balak keras yang berharga dari Borneo dan dihasilkan oleh dua spesies iaitu Eusideroxylon zwageri dan Potoxylon melagangai yang datang dari dua genera yang monotipik. Kedua-dua genera ini adalah tergolong dalam famili Lauraceae dari order Laurales. Kedua-dua spesies ini adalah sangat serupa dan sukar untuk dibezakan secara morfologi. Teknik "Random Amplified Polymorphic DNA" (RAPD) telah digunakan untuk mengenalpasti penanda genomik untuk kedua-dua spesies Belian itu. Genomik DNA untuk Belian telah berjaya diekstrak dengan menggunakan kaedah SDS (Sodium Dodecyl Sulfate) yang telah diubahsuai. Sample DNA telah diskrin dengan menggunakan 20 primer rawak untuk mencari penanda genomik. Hanya satu primer iaitu OPD18 didapati sesuai untuk menghasilkan profil RAPD yang bagus dan konsisten untuk kedua-dua spesies. Profil RAPD ini dapat diulang semula dengan keputusan yang konsisten dan saiz jalur yang dihasilkan adalah dalam lingkungan 500bp ke 2000bp. Dendogram yang dihasilkan telah membahagikan kedua-dua spesies belian kepada dua kelompok yang berasingan. Berdasarkan profil RAPD yang dihasilkan, terdapat corak jalur yang berbeza untuk kedua-dua spesies belian itu. Kajian ini menunjukkan bahawa RAPD adalah amat berguna dalam mengenalpasti penanda genomik bagi spesies yang sangat rapat.

Kata kunci: Penanda genomik, Eusideroxylon zwageri, Potoxylon melagangai, RAPD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Belian refer to the very hard and durable timber produced by two species of tree. Belian is also call ironwood because of its hardness and durability. Belian tree is native to Borneo and Sumatera. This study cover two species of belian namely *Eusideroxylon zwageri* and *Potoxylon melagangai* (this species is formerly referred as *Eusideroxylon melagangai*). These two species are monotypic species come from a 'hard' family, the Lauraceae which contains many valuable timber species. *E. zwageri* is locally known as belian (Brunei, Indonesia, Sabah and Sarawak), tambulian (Sabah and Philippines), and tebelian (Kalimantan) (Kostermans *et al.*, 1994). *P. melagangai* also has different local names such as malangangai (Brunei), legangai (Dusun), tebelian kebuau, belian kapok, belian kebuau (Iban) and belian malagangai (Dayak) (Teo, 1998).

Morphological and anatomical characteristics have been used to identify the two belian species. Both species are very closely related and some people think that they are the same. *Potoxylon melagangai* is formerly known as *Eusideroxylon melagangai* and considered as second species of genus *Eusideroxylon*. This situation is due to similarities in their fruits and leaf. This is very confusing during identification in field (Kostermans *et al.*, 1994 and Teo, 1998). Later, Kosterman separated *E. melagangai* to a monotypic genus namely Potoxylon and given new name, *Potoxylon melagangai* Kosterm (Kostermans *et al.*, 1994). The general morphology features of *E. zw*ageri is shows in Figure 1.1.

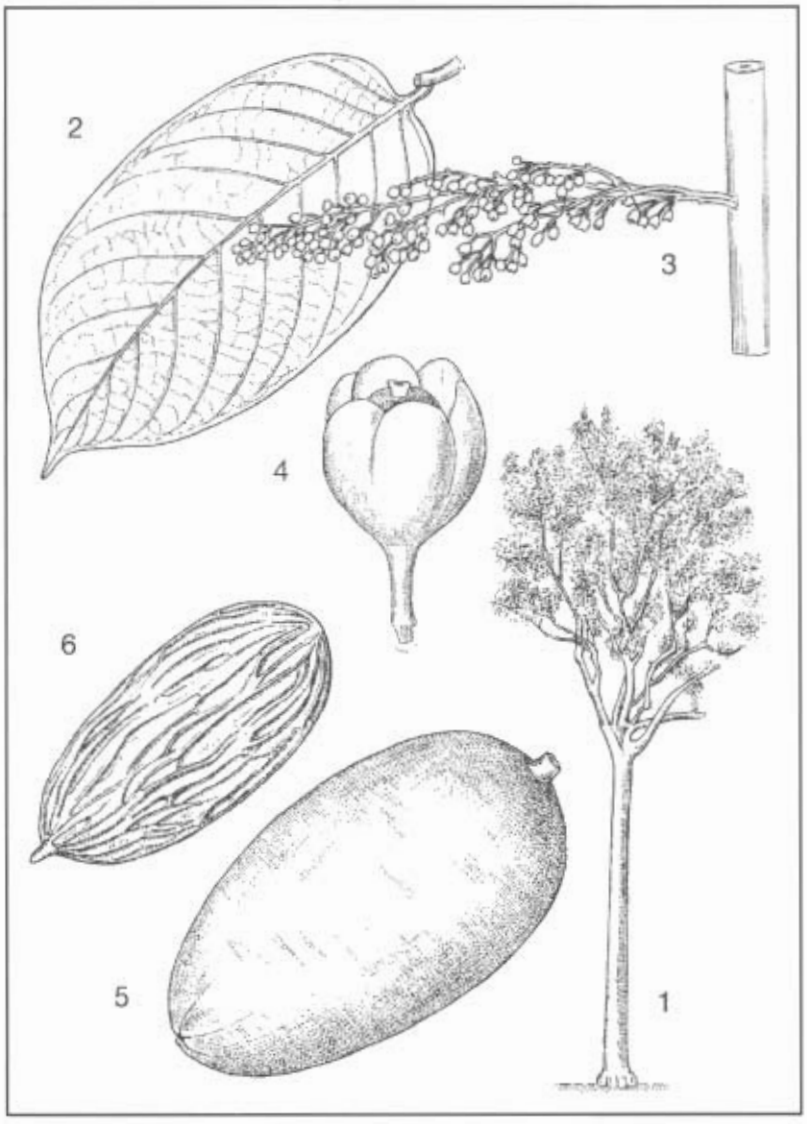

Figure 1.1: *Eusideroxylon zwageri*: 1, tree habit; 2, leaf; 3, inflorescence; 4, flower; 5, fruit; 6, seed. (Adapted from Plant Resources of South-East Asia 5 (1). Timber trees: major commercial timbers)

There is not much information about these two species especially in molecular aspect. Previous work by Ariffin (2003) in protein and polymerase chain reaction (PCR) analysis of belian but no substantial result was obtained from the protein and PCR analysis. Another study by Saban (2002) had analyzed genetic variation for belian using Random Amplified Polymorphic DNA (RAPD). His work was focus on genetic

variation among individual within *Eusideroxylon zwageri* population. Unfortunately, there was also no result obtained from his study due to no PCR product.

Traditionally, morphological characteristics have been used for the assessment of genetic variation but these characteristics are often controlled by multiple genes and subject to varying degree of environmental modification and interaction (Liu *et al.*, 1993). Therefore, differences between closely related species are not always absolute. Since the two species of belian are difficult to distinguish morphologically, molecular approach will be developed to analyze its genetic variation.

The assessment of genetic variation is a major concern of plant breeders and population genetists. This is important at several levels. First, the ability to distinguish members of different species is critical in controlling the material entering a breeding program and in population genetic analyses. Second, the ability to identify or "fingerprint" different genotypes is important in breeding programs that rely on clonal propagation in testing or production and in population genetic analyses of naturally clonal species. Finally, an estimate of the amount of variation within a species is useful for predicting potential genetic gain in a breeding program and in testing population genetic hypotheses (Liu *et al.*, 1993).

Nowadays, advances in techniques for DNA analysis and subsequent data analysis have greatly increased the ability for understanding the genetic relationships among organism at molecular level by developing molecular markers (Stiles *et al.*, 1993). *P. melagangai* is very closely related to *E. zwageri* and some authors consider them to be cogeneric (Teo, 1998). RAPD are efficient molecular markers that can be used to

estimate genetic relationship among genotypes from any organism (Thormann *et al.*, 1994) therefore can be applied to the two species of belian.

Random Amplified Polymorphic DNA (RAPD) has been developed in which DNA is amplified by the polymerase chain reaction using arbitrary short primer (Williams *et al.*, 1990). Depending on the nucleotide sequence of the primer used and the source of the template DNA, different sequence may be amplified in different organisms to get a genome specific "DNA fingerprint". RAPD data will be obtained by staining agarose electrophoresis gel containing fragments that synthesized using the automated technology of PCR (Stiles *et al.*, 1993). The polymorphisms between individuals result from sequence differences in one or both of the primer binding sites, and are visible as presence or absence of a particular RAPD band (Rafalski *et al.*, 1994). RAPD have been successfully used for genetic fingerprinting for dipterocarps (Rath *et al.*, 1998), olive (Germas *et al.*, 2000), cashew (Dhanaraj *et al.*, 2002), guava (Prakash *et al.*, 2002) and mango (Karihaloo *et al.*, 2003).

In this study, RAPD was carried out in attempt to identify the genomic markers and analyze the genetic variation for the two species of belian.

CHAPTER 2

LITERATURE REVIEW

2.1 Belian

2.1.1 Taxonomy of belian

Order: Laurales

Family: Lauraceae

Genera/Species: *Eusideroxylon zwageri*,

Potoxylon melagangai

Eusideroxylon zwageri and *Potoxylon melagangai* are the scientific names for two species of belian. Both belong to two monotypic genera namely *Eusideroxylon* and *Potoxylon*. The family of both genera is Lauraceae from order of Laurales. Lauraceae is a very large family of 35 genera with 2500 species of tree and shrubs. This family can be found throughout the tropics and subtropics. Although consisting of many genera, the family Lauraceae is remarkably uniform in the structure of the fruit and seed, and also in seedling morphology (Ng, 1991).

2.1.2 Variations in belian

According to Browne (1995), *P. melagangai* and *E. zwageri* are difficult to distinguish because the morphological characteristics are almost the same. A distinct morphological difference between these species is the distribution of stomata. *P.*

melagangai has stomata on both side of the leaf (amphistomataus) while *E. zwageri* only has stomata under the leaf (hypostomataus) (Awee, 2002).

Another two distinct differences between these two species are the wood density and colour. Wood density of *P. melagangai* is 525-920kg/m^3 at 15% moisture content (Teo, 1998) while *E. zwageri* is 880-1190kg/m^3 at 15% moisture content (Kostermans *et al.*, 1994). The heartwood's colour for *E. zwageri* is yellowish-brown to reddish-brown when freshly cut, but become silvery brown, dark brown or almost black on exposure; the sapwood is sharply differentiated from the heartwood and bright yellow when freshly cut, darkening yellowish-brown on exposure (Kostermans *et al.*, 1994). For *P. melagangai*, the heartwood is brown with a distinct reddish tinge, sharply differentiated from the yellowish sapwood when freshly cut (Teo, 1998).

2.1.3 Distribution of belian

According to Teo (1998), *P. melagangai* is a monotypic species occurs to all part of Borneo. This species is characteristic of primary lowland, evergreen forest, mainly mixed with dipterocarp forest. It is especially found on sandy-clayey alluvial soil in riverine forest and adjacent hills including limestone hills, up to 300m altitudes. It is prefers well drained but moist soils and usually found scattered or is rarely locally abundant (Teo, 1998).

According to Kostermans *et al* (1994), *E. zwageri* is also monotypic and occurs in eastern and southern Sumatra, Bangka, Belitung, Borneo and the Sulu archipelago and Palawan (Philippines). It is a constituent of primary or old secondary tropical rain

forest. It prefers well-drained soils in valleys or on hillsides or even low ridges when soil moisture is sufficient. It is found from sea level up to 500m altitudes. This species generally found on clay-loam soils or on sandy silt-loam soils, but large specimens have also been found on limestone. *E. zwageri* occurs scattered or gregarious and is often the dominant canopy species. Sometimes its forms almost pure stand or in mixed dipterocarp forest and have been found associated with *Koompassia*, *Shorea* and *Intsia* species.

2.1.4 Uses of belian

Belian is one of the heaviest and most durable timbers of South-East Asia. Besides its hardness and durability, Belian wood is resistant to termite attack (Supriana, 1988) and wood-rotting fungi (Kostermans *et al.*, 1994). Belian grow very slowly, a tree of 100cm diameter may have taken more than 200 years to form (Carter, 1984). *E. zwageri* is more prefer to be used in marine constructions such as pilings, wharfs, dams and ship due to its density and durability. Others less important uses of this species such as in making furniture, chopsticks, blowpipes and survey pegs. Both Belian species can be used for heavy construction, for pepper posts and shingles (Kostermans *et al.*, 1994 and Teo, 1998). In area where *E. zwageri* is not readily available, *P. melagangai* will be a popular substitute. Both are primarily used locally and seldom traded commercially.

According to Kostermans *et al.* (1994), the fruits of *E. zwageri* are poisonous and pulverized fruits have been used medicinally to heal swelling.

2.2 Random Amplified Polymorphic DNA (RAPD)

2.2.1 RAPD

The genetic characterization of individuals within a species or between species and the relationship between them has been determined by a combination of morphological and agronomic traits or by biochemical test such as isoenzyme analysis (Waugh, 1997). However, these methods are subject to environmental influences and their overall effectiveness in estimating genetic relationship has been subjected to some debates (Waugh, 1997). Nowadays, it is widely accepted that information generated from DNA based polymorphisms provides the best estimation of genetic diversity.

Random Amplified Polymorphic DNA (RAPD) is a PCR based technique for identifying genetic variation. It involves the use of a single arbitrary primer in a PCR reaction, resulting in the amplification of many discrete DNA products. The technique was developed independently by two different laboratories (Williams *et al.*, 1990; Welsh and McClelland, 1990) and called as RAPD and AP-PCR (Arbitrary primed PCR) respectively. This procedure detects nucleotide sequence polymorphisms in a DNA amplification-based assay using only a single primer of arbitrary nucleotide sequence. Polymorphisms between individuals are detected as differences between the patterns of DNA fragment amplified from the different DNA using given primers. According to Williams *et al.* (1990), there are some sources of polymorphisms such as deletions of a priming site, insertions that render priming sites too distance to support amplification or insertions that change the size of a DNA segment without preventing its amplification.

Standard PCR reaction using two specific oligonucleotides, which can prime the amplification of the target DNA, sequence selectively (Bova and Micheli, 1997) but RAPD utilized a single primer binds to the genomic DNA at two different sites on opposite strands of the DNA template. If these priming sites are within an amplifiable distance of each other, a discrete DNA product is produced through thermocyclic amplification. The polymorphisms between individuals result from sequence differences in one or both of the primer binding sites, and are visible as the presence or absence of a particular RAPD band (Rafalski *et al.*, 1994).

The primer for RAPD analysis is designed with an arbitrary sequence that is made up of 10 bases. The target sequences to be amplified are unknown. In the other words, the primers can be designed without any information of the nucleotide sequence. Usually arbitrary oligonucleotide primers that do not contain any palindromic sequences and have a minimum of 40% G+C content (generally 50%-80%) will be used in this analysis (William *et al.*, 1990).

In RAPD, a target DNA sequence is exponentially amplified with the help of arbitrary primers, a thermostable DNA polymerase, dNTPs, magnesium chloride and reaction buffer. The reaction involves repeated cycles, each consisting of a denaturation, an annealing and an extension step. In the first step the DNA is made single stranded by raising the temperature to 94°C (denaturation). In the second step, lowering of the temperature to about 40°C to 65°C results in annealing of the primer to their target sequences on the template DNA (annealing step). In the third cycle, temperature is chosen where the activity of the thermostable Taq DNA polymerase is optimal, usually 72°C. The polymerase now extends the 3' ends of the DNA-primer hybrids

towards the other primer-binding site. Since this happens at both primer-annealing sites on both the DNA strands, the target fragment is completely replicated. Repeating these three step cycles 40 to 50 times results in the exponential amplification of the target between the 5' ends of the two primer binding sites. Amplification products are separated by gel electrophoresis and visualized by ethidium bromide staining.

Reproducibility of RAPD has received particular concern. Quality and quantity are two parameters of template DNA that served as the factors mainly affecting RAPD reproducibility, thus controlling both factors is required in order to ensure this. Therefore, the choice of an appropriate DNA extraction protocol is highly recommended. Presence of contaminants could interfere the amplification process. DNA template has to be accurately quantified because template concentration also a crucial parameter to obtain reproducible result, too high or too low template concentrations may lead to unreliable pattern. Apart from that, some primer/template combinations unable to give reproducible pattern when amplification is perform under standard conditions (Waugh, 1997). So, the optimization of the RAPD-PCR reaction parameters is crucial.

2.2.2 RAPD related techniques

AP-PCR (arbitrary primed-PCR) and DAF (DNA amplification fingerprinting) are the two other PCR based related techniques for DNA genotyping using arbitrary primers. RAPD uses single arbitrary primers of 9 or 10 nucleotides in length, low stringency annealing conditions, separation on agarose gels and detection by ethidium bromide staining, whereas DAF (Caetano-Anolles *et al.*, 1991) requires single/ multiple

arbitrary primers that greater than four nucleotides in length, separation by polyacrylamide gel electrophoresis and detection by silver staining. AP-PCR (Welsh and McClelland, 1990) uses single arbitrary primers with 20 to 34 nucleotides in length, low stringency annealing followed by high stringency annealing, separation by polyacrylamide gel electrophoresis and detection by autoradiography. The resolution obtained is low (up to 10 products), moderate (3-20products) and high (up to 100 products) with RAPD, AP-PCR and DAF respectively.

2.2.3 Applications of RAPD

The area of research that has shown the most growth with respect to the use of RAPD technology is that of population genetics. RAPD markers have been used to create fingerprints for the study of individual identity and taxonomic relationship in both eukaryotic and prokaryotic organisms. RAPD can detect polymorphism in closely related organisms such as different population of single species or individuals within a population. Therefore, this technique provides a powerful tool for gene mapping, marker-assisted selection in breeding programs, population and pedigree analysis, phylogenetic studies, and individual and strain identification (Bova and Micheli, 1997).

RAPD markers are being used effectively to assess the amount of genetic diversity in germplasm collections especially crops such as coffee (*Coffea arabica* L.) (Carvalho *et al.*, 2003), pea (*Pisum sativum* L.) (Chegharmiza *et al.*, 2002), pepper (*Piper nigrum* L.) (Archak *et al.*, 2001), coconut (*Cocos nucifera* L.) (Jayadev *et al.*, 2004) and mints (*Mentha spicata* L.) (Bahl *et al.*, 2002). Besides these crops, there are also

RAPD studies in tree species such as in *Lansium domesticum* Corr. (Meliaceae) (Clyde *et al.*, 2000), *Licuala glabra* (Palmae) (Kumar *et al.*, 1999), *Leucadendron elimense* (Proteaceae) (Brown and Tansley, 2000) and *Pilgerodendron uviferum* (Cupressaceae) (Allnutt *et al.*, 2003).

RAPD has been used widely in the assessment of genetic relationship in plants. Since not much molecular work has been done on both Belian species, the reviews of some research below shows the usefulness of RAPD in the study of plant which is in the same family with Belian such as Avocado, others forest tree (Dipterocarps), and example of identification of closely related species, *Ixora.*

Avocado (*Persea Americana* Mill) is in the genus *Persea* of the *Lauraceae* family. Bangerth *et al.* (1998) had study 17 accession of avocado using RAPD and 16 accessions can be differentiated with only one primer. They successfully demonstrated a relatively small but carefully selected set of avocado accessions which show the suitability of RAPD markers for the assessment of genetic relationship of avocado (Bangerth *et al.*, 1998).

Rajaseger *et al.* (1997) also had used RAPD for analysis of genetic diversity among *Ixora* cultivars (Rubiaceae). The identification and classification of species and cultivars in this genus has been very difficult due to the existence of natural hybrids among species that cannot be ruled out. The taxa have been morphologically distinguished using the nature of the leaves and floral parts. By using RAPD, they show the successful analysis for cultivars identification in *Ixora*, a group of tropical ornamental plants.

The family Dipterocarpaceae is the predominant family of timber trees in the rainforest of the Malaysian region. Rath *et al.* (1997) have been using RAPD to perform phylogenetic analysis of dipterocarps. The phylogenetic relationships among 12 species from three genera (*Shorea, Hopea and Anisoptera*) of Dipterocarpaceae were studied. Identification of forest trees, especially dipterocarps is not an easy task. Standard morphological markers such as height of the tree, shape of the canopy, and leaf shape and size are ill-suited to Dipterocarpaceae because of the sheer size, similar shapes and near-identical leaf morphologies exhibited by many members of this family. Due to these difficulties in identification, RAPD technique was successfully applied to study the relationships between species with in a particular genus, and populations within a species of dipterocarps (Rath *et al.*, 1997).

CHAPTER 3

MATERIALS AND METHODS

3.1 Plant Materials

Young and healthy leaves (reddish green) were collected from both belian species (*E. zwageri* and *P. melagangai*) that are grown in UNIMAS Green House. Young leaves of *E. zwageri* were priory destarched by covering the leaf with black plastic for 2 days before plucking the leaf for DNA extraction. Fresh leaf samples were surface sterile by washing with 70% ethanol then rinsed with distilled water before DNA extraction.

3.2 Methods

3.2.1 Genomic DNA Extraction

3.2.1.1 CTAB (Cetyltrimethylammoniumbromide) Method

Genomic DNA of the two species of Belian was isolated using method reported by Doyle & Doyle, (1987) with slight modification by Gillies and Bush. Young leaves were cut into fine pieces with a sterile scalpel and the veins will be discarded. About 0.1g of the cut pieces were weight and place into a mortar. The leaf was frozen in liquid nitrogen rapidly and grounded to a powder form with a pestle as the liquid nitrogen boiled off. A little more liquid nitrogen will be added if necessary to keep the powder from thawing while grinding. 100µl of CTAB extraction buffer (refer Appendix IV) was added to the tissue and continue grinding until slurry was formed. A further 900µl of CTAB extraction buffer was added into the mortar. The tissue

together with the extraction buffer was transfer into a 1.5ml eppendorf tube. The content was mixed gently and incubate in a water bath at 65°C for 1 hour. After that the tube was removed from the water bath and allowed to cool to room temperature for approximately 5 minutes. Then, 400µl of "wet" chloroform (CIA) was added and mixed gently to a single phase. The mixture was centrifuged at 13000rpm for 5 minutes at 4°C. The aqueous layer was removed and transferred to a clean Eppendorf tube. 600µl of ice-cold isopropanol was added and mixed gently to precipitate the DNA. The mixture was left to stand overnight at -20°C.

The next day, the tube was taken out from the freezer and centrifuged at 13000rpm for 2 minutes at 4°C to pellet the DNA. The supernatant was removed and 1ml of wash buffer was added. The tube was agitated to dislodge the pellet and then leave to stand at room temperature for 30 minutes. The DNA was pelleted for second time at 13000rpm for 2 minutes at 4°C. The supernatant was discarded and the pellet was air-dried for approximately 15 minutes. The DNA pellet was then dissolved in 100µl TE buffer and stored at 4°C until required.

3.2.1.2 SDS (Sodium Dodecyl Sulfate) Method

Genomic DNA of the two species of Belian was isolated using method reported by Kikuchi *et al.*, (1998) with modification by Sim (personal communication). Young leaves were collected and washed with 70% ethanol then rinsed with distilled water. About 0.1g of fresh leaves was ground in liquid nitrogen until fine powder. The ground leaf powder was transferred to an Eppendorf tube containing 600µl lysis buffer (refer Appendix IV). Solution was mixed well by inversion. 200µl of 5M