UNIVERSITI PUTRA MALAYSIA

GENETIC DIVERSITY STUDY OF SELECTED SARAWAK RICE CULTIVARS USING MICROSATELLITE MARKERS

PHENE NEOH PEI NEE

GENETIC DIVERSITY STUDY OF SELECTED SARAWAK RICE CULTIVARS USING MICROSATELLITE MARKERS

By

PHENE NEOH PEI NEE

A Project Report Submitted in Partial Fulfillment of the Requirement for the Degree of Bachelor of Science Bioindustry in the Faculty of Agriculture and Food Sciences Universiti Putra Malaysia Bintulu Sarawak Campus

2009

“I dedicate this to my family for their loving support”

ABSTRACT

Sixteen selected Sarawak rice cultivars were evaluated for genetic diversity using twelve simple sequence repeat (SSR) markers. Forty-nine microsatellite markers were screened and twelve polymorphic SSR markers were selected for further analysis using polyacrylamide gel electrophoresis. A total of 34 alleles were detected by using 12 polymorphic markers. Polymorphism information content (PIC) values of the markers ranged from 0.375 (RM489) to 0.703 (RM1) with an average value of 0.562. The dendrogram (r=0.78) based on SSR polymorphism indicated great variation among the rice cultivars, with similarity coefficients ranging from 0.262 to 1.000. Genetic diversity, determined by cluster analysis, indicated two clusters at 26.21% level of similarity with additional sub-groups within each cluster. This study revealed that SSR markers are a useful tool for cultivar differentiation according to their genotypes. The information on genetic diversity of rice genotypes is essential for proper identification and selection of appropriate cultivars for rice improvement.

ABSTRAK

Enam belas varieti padi telah dinilai dengan menggunakan dua belas penanda mikrosatelit (SSR). Empat puluh sembilan penanda mikrosatelit telah disaring dan dua belas penanda mikrosatelit polimorfik dipilih untuk dianalisis dengan menggunakan gel elektrophoresis poliakrilamid. Sejumlah 34 alel telah dikesan dengan menggunakan 12 penanda mikrosatelit. Nilai maklumat polimorfik penanda mikrosatelit berada di dalam lingkungan 0.375 (RM489) sehingga 0.703 (RM1) dengan nilai purata sebanyak 0.562. Dendrogram (r=0.78) yang diperolehi berdasarkan kepolimorfikan SSR menunjukkan variasi yang besar antara varieti padi, dengan nilai kesamaan di dalam lingkungan 0.262 sehingga 1.000. Kepelbagaian diversiti yang ditentukan oleh analisis kluster menunjukkan dua kelompok besar dengan paras kesamaan sebanyak 26.21% dengan tambahan kumpulan kecil di dalamnya. Penyelidikan ini menunjukkan penanda mikrosatelit adalah sangat berguna dalam pembezaan varieti padi menurut genotip mereka. Informasi mengenai kepelbagaian genetik varieti padi adalah sangat berguna dalam pengenalpastian dan pemilihan varieti yang sesuai dalam pembaikbakaan varieti padi.

ACKNOWLEDGMENTS

This thesis owes its existence to the help, support and inspiration of many people. Firstly, I would like to express my sincere appreciation and gratitude to Mr. Wong Sie Chuong for his support and encouragement in this thesis's work. I am fortunate for having Mr. Wong as my supervisor. He has provided an optimum working environment where a lack of resources is something unimaginable due to his skills and foresight. Through his uncompromising pursuit for excellence that significantly shapes everyone at the research laboratory, he also provided extraordinary research experiences during research progress. Therefore, I am truly, deeply indebted to him and feeling gratitude alone would not even match up all his patience, guidance, encouragement and excellent advice throughout the research. Thank you.

I am also indebted to Associate Professor Dr. Rajan, who has not only been a source of enthusiasm and encouragement but also agreed to be my co-supervisor. He provided me unflinching encouragement and support in various ways whilst his truly scientist intuition inspired and enriched my growth as a student and a researcher. I am indebted to him more than he knows. I am also thankful to Dr. Yiu Pang Hung for his generous assistance all the time.

I would like to express my sincere thanks to Cik Siti Fatimah for her generous assistance in helping and assisting me during this whole year of difficulties. I am also grateful to the faculty and staff of UPMKB especially

Department of Basic Science and Engineering and Department of Crop Science for their support and assistance throughout the research.

I would also like to thank my fellow researchers (coursemates), Weng Bong Siaw Then, Lee Huei Hong, Huichi and Kiing Sie Chuong. Those long days at the laboratory will be dull without both of you. Not to be forgotten, other fellow researchers that supported and helped, thanks for all the assistance. I would like to thank Loo Li Wen for her comments and suggestions for the editing of my thesis.

I would not be here without the support from my family. Thus, I would like to grab this opportunity to express my profound gratitude to my beloved parents for their inseparable support and prayers, my sister and brother in law for being supportive and caring, aunts and uncles for their moral support and patience during my study in UPMKB. Words fail me to express my deepest appreciation and gratitude to them.

Finally, I would like to thank everybody who was important to the successful fulfillment of this thesis, as well as expressing my apology for that I could not mention personally one by one.

APPROVAL

I certify that this research project entitled "Genetic Diversity Study of Selected Sarawak Rice Cultivars Using Microsatellite Markers" has been examined and approved as a partial fulfillment of the requirement for the degree of Bachelor of Science Bioindustry in the Faculty of Agriculture and Food Sciences, Universiti Putra Malaysia Bintulu Sarawak Campus.

_____

Mr Wong Sie Chuong
Lecturer (Supervisor)
Department of Basic Science and Engineering
Faculty of Agriculture and Food Sciences
Universiti Putra Malaysia Bintulu Sarawak Campus

_____

Professor Dr. Japar Sidik Bujang
Dean
Faculty of Agriculture and Food Sciences
Universiti Putra Malaysia Bintulu Sarawak Campus

Date:

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

RFLP	Restriction fragment length polymorphism
RAPD	Random amplified polymorphic DNA
SSR	Simple sequence repeats
PIC	Polymorphism information content
PCR	Polymerase chain reaction
PAGE	Polyacrylamide gel electrophoresis
bp	base pairs

CHAPTER 1

INTRODUCTION

Rice, *Oryza sativa* L., is one of the agronomically and nutritionally important cereal crops and is the principal staple food, which feeds more than half of the global population (Saw *et al.,* 2006). It plays central role in human nutrition and culture for the past 10,000 years. It had been estimated that world rice production must increase by 30% over the next 20 years to meet projected demands from population increase and economic development (IRGSP, 2005). Malaysia also aimed for rice self-sufficiency in the future where the target for 2010 is to reach 86% self-sufficiency according to Malaysia National Food Security Policy (IRRI, 2009).

Sarawak is a state rich in biodiversity of cultivated rice varieties. Rice fields in Sarawak are very enchanting and diverse as these rich ecosystems of rice paddies exist all over the state. The choice of paddy planted usually of the farmers own traditional variety that they considered adapted well to the uncertainties of the environment. The simplest distinction between the rice economies of Sarawak with other parts of Malaysia is the dominance of subsistence-cropped and low-yielding paddy. Majority of the rural communities depended on hill paddy. Due to different or diverse environments, it supported many old traditional rice varieties that are mostly not found in most of the Asian countries. A myriad of rice varieties of different grain sizes, shapes and colours can be found in the Sarawak market. Rice

genetic diversity is important to farming communities and to agriculture in general. Rice genetic diversity is crucial because breeders strive to develop the ultimate rice variety that possesses the traits with: (i) high yield, (ii) high grain quality and (iii) resistance to pests and diseases. As such, attempts to conserve the genetic materials of rice were initiated all over the world.

Rice is the second plant genome to be completely sequenced, after the model plant species *Arabidopsis thaliana* (Mackill, 2003). Rice is selected to be one of the model for plant genomic studies because of its small genome size and relatively low amount of repetitive DNA, its diploid nature, its ease of manipulation in tissue culture and global importance as a food crop (Mackill, 2003; Xu *et al.,* 2005; Ahmadikhah *et al.,* 2007). The use of rice genetics has also become increasingly popular as the model system for the study in monocotyledonous plants. Previous studies in rice have contributed to the development of several hundred microsatellite markers and a genetic map consisting of 320 microsatellites (Chen *et al.,* 1997; Temnykh *et al.,* 2000).

The genetic maps of grass species have been constructed using a variety of marker types. Most of the older species-specific molecular maps were constructed with restriction fragment length polymorphism (RFLP) markers, but in recent times, there has been increased utilization of PCR-based markers because of accessibility and higher throughput. The RFLP approach has been used successfully to identify genetic markers in plants, including rice. However, the use of radioactive elements and specific probes for the target DNA sequences makes RFLP technique costly and tedious (Joshi and Behera, 2007). Development of Polymerase Chain Reaction

(PCR) based techniques offered a good alternative to the genetic analysis. The PCR based randomly amplified polymorphic DNA (RAPD), microsatellite and amplified fragment length polymorphism (AFLP) approach require less DNA, simple, quicker and cheaper compared to RFLP (Welsh *et al.,* 1990; Williams *et al.,* 1990). Among those markers, microsatelite markers are highly polymorphic, more reproducible, co-dominant and distributed throughout the genome and are especially suitable for evaluating genetic diversity among germplasm that is closely related such as closely related rice cultivars (Akagi *et al.*, 1997; Ni *et al.,* 2002; Mackill, 2003). Comparatively, microsatellite markers showed higher level of polymorphism compared to dominant markers (Garcia *et al.,* 2004).

Microsatellites are well known for their potentially high information content and versatility as molecular tools (Chambers and Macavoy, 2000; Pessoa-Filho *et al.,* 2007). Microsatellites are among the most variable types of DNA sequence in the genome. Their polymorphisms derive mainly from variability in length rather than in the primary sequence (Ellergen, 2004). Their nature gives them a number of advantages over other molecular markers. Multiple microsatellite alleles may be detected at a single locus and are found distributed all over the genome. They are also co-dominant and require only a small amount of DNA for screening. Microsatellites also demonstrate a high degree of transferability between species, as microsatellites PCR primers designed for one species frequently amplifies a corresponding locus in related species, making them excellent markers for comparative genomic studies (Jewell *et al.,* 2006). Microsatellites may be obtained by screening sequences in databases or by screening libraries of clones.

In the effort to meet the continuously expanding needs of rice varietal improvement, assembling, evaluation and assessment of the rice genetic resources is important. In the present study, 16 rice cultivars were analysed for genetic variation using 49 microsatellite markers distributed on 12 different chromosomes. Specifically, the objectives of this study were (1) to obtain suitable microsatellite markers to detect genetic diversity within the selected Sarawak rice cultivars; (2) to investigate genetic relationships of the studied rice cultivars; and (3) to assess the genetic diversity of the selected Sarawak rice cultivars. This knowledge will serves as the first step for proper identification and selection of parents for future rice improvements.

CHAPTER 2

LITERATURE REVIEW

2.1 Rice

Rice belongs to the genus *Oryza*, which is composed of 23 species divided into 10 genome types (Buell, 2005). Rice (*Oryza sativa*) is a major staple food and is consumed by nearly half the world's population. It accounts for more than 21% of global human per capita energy and 15% of per capita protein. In the past few decades, although rice production has doubled due to the introduction of high yielding varieties or hybrids and improved cultivation practices, it is still insufficient to cope with ever-increasing global demand, which is expected to increase at the rate of about 1% per annum (Khush, 1997). At the same time, inappropriate natural resource use, along with biotic and abiotic stress pressure, is casting a shadow on rice production (Peng *et al.*, 2004).

As one of the world's most important food crops, rice has emerged as the leading experimental model for functional and evolutionary genomics studies for cereals. Rice genome has served as a catalyst for investigations in comparative genomics, functional genomics, map-based gene cloning and molecular breeding in rice (Paterson *et al.*, 2004; Sasaki *et al.*, 2005). Accessibility to the rice genome sequence will enable identification of genes responsible for traits and alleles that

will be essential to meet the growing demands of food production in the coming years.

Synteny in the cereals had been reported previously using molecular markers (Gale and Devos, 1998) and recently at a higher resolution using sequences available from the rice, sorghum, maize, and wheat genome (Goff *et al.,* 2002; Sorrells *et al.,* 2003; Lai *et al.,* 2004). Thus, access to a complete, high-resolution rice genome sequence will facilitate research on other cereals with larger, partially sequenced genomes. Furthermore, the availability of detailed information about the rice genome will likely advance the global efforts to improve other major crops in the grass family (Barry, 2001). The increased research on rice may have direct or indirect benefits for this critical crop and its relatives, and ultimately may contribute to an improvement in food production in many world areas (Barry, 2001).

2.2 Molecular Markers

Before the introduction of molecular markers, plant breeders based their selection of breed through phenotypic agronomic traits such as yield and plant height. A molecular marker is defined as a particular segment of DNA that is representative of the differences at the genome level. However, molecular markers may or may not correlate with phenotypic expression of a trait. Compared to phenotypic based markers, molecular markers are stable and detectable in all tissues regardless of growth, differentiation or development stage and will not be affected by the

environment, pleiotropic as well as epistatic effects. Development of molecular marker technology provides plant breeders with extra tools capable of monitoring the selection through genome. With the advantages of molecular markers, it is not surprising that academician and plant breeders adopted molecular marker based genotyping. Molecular markers are also an indicator of sequence polymorphism (Edwards and Mogg, 2001). Therefore, molecular markers technologies are valuable tools in both basic and applied research, such as fingerprinting genotypes, analyzing genetic diversity and relatedness, determining variety identity, marker assisted breeding, phylogenetic analysis, map-based cloning of genes as well as for germplasm conservation in gene banks and assisting conventional breeding efforts (McCouch *et al*., 1997; Shen *et al.,* 2004; Alvarez *et al.,* 2007). There are different types of molecular markers available such as restriction fragment length polymorphism (RFLP), randomly amplified polymorphic DNA markers (RAPD) and simple sequence repeat (SSR).

RFLP is a molecular marker based on the differential hybridization of cloned DNA to DNA fragments in a sample of restriction enzyme digested DNAs. The marker is specific to a single clone or restriction enzyme combination. Major sources of RFLP markers are species-specific genomic DNA and cDNA sequences. RFLP analysis displays high efficiency for the generation of genetic maps. However, due to the time-consuming multi step protocol and the requirement of radioactivity for satisfying fragment detection, RFLPs became unfavourable. Another common molecular marker used is randomly amplified polymorphic DNA markers (RAPD) where nucleotide sequence polymorphisms in the DNA were detected by using a single primer of arbitrary nucleotide sequence (Welsh and McClelland, 1990). Most

advantageous of this molecular marker technique is the use of universal primers enabling the cost-effective accomplishment of various genetic analysed in a short period of time (Tingey and del Tufo, 1993). Unfortunately, due to frequently observed problems with reproducibility of overall RAPD profiles and specific bands which lead to inappropriate inferences, this marker became less popular. Currently, the most popular molecular marker is microsatellite markers as they have additional advantages when compared to RFLP and RAPD.

The term satellite DNA originates from the observation in the 1960s of a fraction of sheared DNA that showed a distinct buoyant density, detectable as a 'satellite' peak in density gradient centrifugation and that was subsequently identified as large centromeric tandem repeats. Typical satellite consists of very high numbers of repetitions of a basic sequence motif. Monomer sizes may range from two to several thousands of base pairs. Discovery of tandem iterations of simple sequence motifs is where the term microsatellites were coined (Ellergen, 2004). Microsatellites are characterized by a low degree of repetition at a particular locus, but microsatellites consisting of identical motifs may be found at many thousand genomic loci.

Microsatelites, also known as SSR (simple sequence repeats), are tandem repeats of short DNA motifs (1-6 bp in length) which are a simple dinucleotide, trinucleotide, tetranucleotide, pentanucleotide or hexanucleotide sequence. They occur abundantly and at random while frequently exhibit variation in the number of repeats at a locus (Jarne and Lagoda, 1996; McCouch *et al.*, 1997; Temnykh *et al.*, 2001; Ellergen, 2004; Li *et al.*, 2004; Jewell *et al.*, 2006).

Due to their abundance and inherent potential for variation, these microsatellites have become a popular type and valuable source of genetic markers (McCouch *et al.*, 1997; Cho *et al.*, 2000; Li *et al.*, 2004). The existence of tandem repeats in eukaryotic genomes was first recognized in the early 1970s, when $(TAGG)_n$ repeats were found in the satellite DNA of hermit crab (Skinner *et al.*, 1974). After that, microsatellites were studied in humans (Hamada *et al.*, 1982) and have now been found in a wide array of other eukaryotes, birds, fish, insects, yeast and several monocot and dicot plant species (Morgante and Olivieri, 1993; Cheng and Crittenden, 1994).

In plants, the presence of microsatellites was first demonstrated by RFLP fingerprinting with oligonucleotide probe. PCR-generated and locus specific plant microsatellite markers were first reported in soybean in 1992 (Akkaya *et al.*, 1992). Previous studies in rice also found microsatellites to be co-dominant, abundant and well distributed throughout the rice genome (Temnykh *et al.*, 2001; Alvarez *et al.*, 2007). McCouch *et al.* (2002) found more than 2200 microsatellite markers and mapped to specific locations in rice genome. Application of SSR in rice genetic diversity research was also studied in various countries such as in Yunnan (Tu *et al.*, 2007), Indonesia (Thomson *et al.*, 2007), Philippines (Lapitan *et al.*, 2007) and many more. In Sarawak, there were only few studies on rice research using SSR. It was only initiated in 2006, where three Bario rice cultivars were studied using 50 SSRs (Tan *et al.*, 2006) and unpublished report on genotyping of 25 Sarawak rice cultivars using 12 microsatellite markers (Valentine, 2008). SSR was found to be rapid, efficient and cost-effective, genome scanning technique that can detect polymorphisms compared to other methods

(Blair *et al.*, 1999). Microsatellite markers are cost-effective, as it requires only small amount of DNA and time. It is also efficient as it is easy to use while the results are reproducible compared to RFLP and RAPD.

CHAPTER 3

METHODOLOGY

3.1 Plant Material

Sixteen cultivars from different vicinities in Sarawak were evaluated in this study (Table 1). Ten seeds per cultivar were randomly selected as representative of the cultivars and germinated in Petri dishes and seedlings were transplanted into individual pots in the net house. About 10 to 30 mg of young leaves were harvested from four seedlings after 3 to 4 week growth and stored frozen at -80 °C. Each cultivar was harvested in four replications.

Table 1: List of selected Sarawak rice cultivars.

Sample no.	Name	Source/ Collection Site
1	Adan Sederhana	DOA
2	Bukit Wangi	Lubok Nibong
3	Buntar B	DOA
4	Empawah	DOA
5	Empawah Merah	Stumbin
6	Lasak	DOA
7	Lemak	Meradong
8	Lemak	Roban Area
9	Mamut	Stumbin
10	Mamut-2	Stumbin
11	Muyun	Tatau
12	Rotan Wangi	Lubok Nibong
13	Sampangan B	DOA
14	Sia	Lubok Nibong
15	Tulang	Tatau
16	Wangi	Stumbin

Note: DOA: Department of Agriculture, Sarawak.

3.2 Molecular Analysis

3.2.1 DNA Extraction

DNA was isolated following the procedure of Vivantis GF-1 Plant DNA Extraction Kit. Briefly, the samples were frozen with liquid nitrogen and ground into a fine powder with a mortar and pestle to homogenize the tissue samples. Buffer PL was added to the homogenized sample to lyses the sample tissues. A homogenous solution was obtained by vortexing the tube for 30 seconds, and 20 µL of Proteinase K (Vivantis) was added to the vortexed samples. The samples were incubated at 60 °C for 2 hours in shaking incubator (TS-100 Thermo Shaker, Boeco Germany) to ensure thorough digestion of the samples. After incubation, the samples were centrifuged at 16,000 x g. Supernatant containing DNA was then transferred into a new and clean microcentrifuge tube. Twenty microlitres of RNase A (Vivantis: DNase-free, 20 mg/ml) was added to digest and extract the RNA from the samples. After the removal of RNA, two volumes of Buffer PB were added and incubated for 10 minutes at 60 °C. After cooling down the samples, 200 µL of absolute ethanol was added. The mixed samples and solutions were transferred into columns and centrifuged at 10,000 x g for 1 minute. The flow through were discarded and columns were washed with wash buffer and centrifuged at 10,000 x g for 1 minute. The columns were centrifuged again at 10,000 x g for 1 minute to remove residual ethanol from the wash buffer as the ethanol can affect the quality of DNA. The columns were transferred into a clean 1.5 ml microcentrifuge tube. Preheated 100 µl elution buffer (10 mM Tris-HCl, pH 8.5) was added onto column membranes and stood still on microcentrifuge stand for 2 minutes. DNA was eluted by centrifuging at 10,000 x g for 1 minute. DNA

was pooled from the four replicates and used as a template for PCR reactions using microsatellite markers.

3.2.2 Microsatellite Markers Screening

To select suitable microsatellite markers (SSRs), 49 rice microsatellite marker primer pairs were selected from the literature (Bligh *et al.,* 1999; Ni *et al.,* 2002; Lu *et al.,* 2005; Chakravarthi and Naravaneni, 2006; Brondani *et al.,* 2006; Alvarez *et al.,* 2007; Joshi and Behera, 2007; Lapitan *et al.,* 2007; Ram *et al.,* 2007; Thomson *et al.,* 2007; Tu *et al.,* 2007; Sundaram *et al.,* 2008). Microsatellite markers were selected based on the following criteria: (i) high polymorphism information content (PIC) and (ii) uniformity in distribution across the 12 chromosomes to analyze genetic diversity. These microsatellite markers were synthesized by 1st Base Laboratories Sdn. Bhd. (Malaysia). The details of selected microsatellite markers and their positions on respective chromosomes are listed in Table 2. Primer sequences for these markers can also be found on the Gramene website (www.gramene.org).

Table 2: List of rice microsatellite markers used for screening

No	SSR Marker	Chromosome	Repeat Motif	Position (cM)	Primer Sequence (5' – 3')	Source
1	RM 1	1	$(AG)_{26}$	29.7	F: GCG AAA ACA CAA TGC AAA AA R: GCG TTG GTT GGA CCT GAC	Bligh *et al.* (1999), Ni *et al.* (2002), Lu *et al.* (2005)
2	RM226	1	$(AT)_{38}$	154.8	F: AGC TAA GGT CTG GGA GAA ACC R: AAG TAG GAT GGG GCA CAA GCT C	Ram *et al.* (2007), Sundaram *et al.* (2008)
3	RM237	1	$(CT)_{18}$	115.2	F: CAA ATC CCG ACT GCT GTC C R: TGG GAA GAG AGC ACT ACA GC	Ni *et al.* (2002), Lu *et al.* (2005), Lapitan *et al.* (2007), Thomson *et al.* (2007)
4	RM243	1	$(CT)_{18}$	57.3	F: GAT CTG CAG ACT GCA GTT GC R: AGC TGC AAC GAT GTT GTC C	Ni *et al.* (2002)
5	RM154	2	$(GA)_{21}$	4.8	F: ACC CTC TCC GCC TCG CCT CCT C R: CTC CTC CTC CTG CGA CCG CTC C	Lu *et al.* (2005), Joshi and Behera (2007), Thomson *et al.* (2007)
6	RM207	2	$(CT)_{25}$	191.2	F: CCA TTC GTG AGA AGA TCT GA R: CAC CTC ATC CTC GTA ACG CC	Ni *et al.* (2002), Brondani *et al.* (2006), Lapitan *et al.* (2007)
7	RM240	2	$(CT)_{21}$	158.0	F: CCT TAA TGG GTA GTG TGC AC R: TGT AAC CAT TCC TTC CAT CC	Ni *et al.* (2002), Lapitan *et al.* (2007), Ram *et al.* (2007)
8	RM452	2	$(GTC)_{9}$	58.4	F: CTG ATC GAG AGC GTT AAG GG R: GGG ATC AAA CCA CGT TTC TG	Lu *et al.* (2005)
9	RM22	3	$(GA)_{22}$	11.0-13.0	F: GGT TTG GGA GCC CAT AAT CT R: CTG GGC TTC TTT CAC TCG TC	Ni *et al.* (2002), Lu *et al.* (2005), Brondani *et al.* (2006), Lapitan *et al.* (2007)

Table 2 : (Continued)

No	SSR Marker	Chromosome	Repeat Motif	Position (cM)	Primer Sequence (5' – 3')	Source
10	RM155	3	$(GA)_{17}$	168.2	F: CCG TCG CCG TAG TAG AGA AG R: TCC CGG TTA TTT TAA GGC G	Bligh *et al.* (1999), Ni *et al.* (2002), Lu *et al.* (2005), Thomson *et al.* (2007)
11	RM338	3	$(CTT)_6$	108.4	F: CAC AGG AGC AGG AGA AGA GC R: GGC AAA CCG ATC ACT CAG TC	Ni *et al.* (2002), Lu *et al.* (2005), Ram *et al.* (2007)
12	RM489	3	$(ATA)_8$	29.2	F: ACT TGC GAC GAT CGG ACA CC R: TCA CCC ATG GAT GTT GTC AG	Panel of 50 rice SSR markers (Gramene website)
13	RM514	3	$(AC)_{12}$	216.4	F: AGA TTG ATC TCC CAT TCC CC R: CAC GAG CAT ATT ACT AGT GG	Lu *et al.* (2005), Thomson *et al.* (2007)
14	RM124	4	$(TC)_{10}$	150.1	F: ATC GTC TGC GTT GCG GCT GCT G R: CAT GGA TCA CCG AGC TCC CCC C	Lu *et al.* (2005), Thomson *et al.* (2007)
15	RM252	4	$(CT)_{19}$	99.9	F: TTC GCT GAC GTG ATA GGT TG R: ATG ACT TGA TCC CGA GAA CG	Brondani *et al.* (2006), Lapitan *et al.* (2007)
16	RM307	4	$(AT)_{14}$-$(GT)_{21}$	0.0	F: GTA CTA CCG ACC TAC CGT TCA C R: CTG CTA TGC ATG AAC TGC TC	Ni *et al.* (2002), Lu *et al.* (2005), Sundaram *et al.* (2008)
17	RM335	4	$(CTT)_{25}$	21.5	F: GTA CAC ACC CAC ATC GAG AAG R: GCT CTA TGC GAG TAT CCA TGG	Sundaram *et al.* (2008)
18	RM161	5	$(AG)_{20}$	96.9	F: TGC AGA TGC GAA GCG GCG CCT C R: TGT GTC ATC AGA CGG CGC TCC G	Lu *et al.* (2005), Thomson *et al.* (2007)

Table 2 : (Continued)

No	SSR Marker	Chromosome	Repeat Motif	Position (cM)	Primer Sequence (5' – 3')	Source
19	RM334	5	$(CTT)_{20}$	141.8	F: GAC TTT GAT CTT TGG TGG ACG R: TTG GAT TGT TTT GCT GGC TCG C	Ni *et al.* (2002), Lu *et al.* (2005), Sundaram *et al.* (2008)
20	RM413	5	$(AG)_{11}$	26.7	F: GGC GAT TCT TGG ATG AAG AG R: TCC CCA CCA ATC TTG TCT TC	Lu *et al.* (2005), Thomson *et al.* (2007)
21	RM421	5	$(AGAT)_6$	111.2	F: AGC TCA GGT GAA ACA TCC AC R: ATC CAG AAT CCA TTG ACC CC	Lu *et al.* (2005), Ram *et al.* (2007)
22	RM133	6	$(CT)_8$	0.0	F: TTG GAT TGT TTT GCT GGC TCG C R: GGA ACA CGG GGT CGG AAG CGA C	Lu *et al.* (2005), Thomson *et al.* (2007)
23	RM162	6	$(AC)_{20}$	108.3	F: GCC AGC AAA ACC AGG GAT CCG G R: CAA GGT CTT GTG CGG CTT GCG G	Lu *et al.* (2005), Thomson *et al.* (2007), Sundaram *et al.* (2008)
24	RM204	6	$(CT)_{44}$	25.1	F: GTG ACT GAC TTG GTC ATA GGG R: GCT AGC CAT GCT CTC GTA CC	Ni *et al.* (2002), Brondani *et al.* (2006), Ram *et al.* (2007), Sundaram *et al.* (2008)
25	RM541	6	$(TC)_{16}$	75.5	F: TAT AAC CGA CCT CAG TGC CC R: CCT TAC TCC CAT GCC ATG AG	Lu *et al.* (2005), Lapitan *et al.* (2007), Thomson *et al.* (2007)
26	RM11	7	$(GA)_{17}$	47.0	F: TCT CCT CTT CCC CCG ATC R: ATA GCG GGC GAG GCT TAG	Bligh *et al.* (1999), Lu *et al.* (2005), Brondani *et al.* (2006), Alvarez *et al.* (2007), Thomson *et al.* (2007), Tu *et al.* (2007)

Table 2 : (Continued)

No	SSR Marker	Chromosome	Repeat Motif	Position (cM)	Primer Sequence (5' – 3')	Source
27	RM118	7	$(GA)_8$	96.9	F: CCA ATC GGA GCC ACC GGA GAG C R: CAC ATC CTC CAG CGA CGC CGA G	Lu *et al.* (2005)
28	RM125	7	$(GCT)_8$	24.8	F: ATC AGC AGC CAT GGC AGC GAC C R: AGG GGA TCA TGT GCC GAA GGC C	Lu *et al.* (2005), Chakravarthi and Naravaneni (2006), Thomson *et al.* (2007)
29	RM138	7	$(GT)_{14}$	196.8	F: AGC GCA ACA ACC AAT CCA TCC G R: AAG AAG CTG CCT TTG ACG CTA TGG	Lapitan *et al.* (2007)
30	RM149	8	$(AT)_{10}$	103.7	F: GCT GAC CAA CGA ACC TAG GCC G R: GTT GGA AGC CTT TCC TCG TAA CAC G	Lu *et al.* (2005), Chakravarthi and Naravaneni (2006), Sundaram *et al.* (2008)
31	RM152	8	$(GGC)_{10}$	9.4	F: GAA ACC ACC ACA CCT CAC CG R: CCG TAG ACC TTC TTG AAG TAG	Ni *et al.* (2002), Lu *et al.* (2005)
32	RM284	8	$(GA)_8$	83.7	F: ATC TCT GAT ACT CCA TCC ATC C R: CCT GTA CGT TGA TCC GAA GC	Lu *et al.* (2005), Lapitan *et al.* (2007)
33	RM404	8	$(GA)_{33}$	60.9-69.0	F: CCA ATC ATT AAC CCC TGA GC R: GCC TTC ATG CTT CAG AAG AC	Lapitan *et al.* (2007)
34	RM105	9	$(CCT)_6$	32.1	F: GTC GTC GAC CCA TCG GAG CCA C R: TGG TCG AGG TGG GGA TCG GGT C	Ni *et al.* (2002), Lu *et al.* (2005), Thomson *et al.* (2007)

Table 2 : (Continued)

No	SSR Marker	Chromosome	Repeat Motif	Position (cM)	Primer Sequence (5' – 3')	Source
35	RM215	9	$(CT)_{16}$	99.4	F: CAA AAT GGA GCA GCA AGA GC R: TGA GCA CCT CCT TCT CTG TAG	Ni *et al.* (2002), Lu *et al.* (2005), Lapitan *et al.* (2007), Thomson *et al.* (2007), Tu *et al.* (2007), Sundaram *et al.* (2008)
36	RM278	9	$(GA)_{17}$	77.5	F: GTA GTG AGC CTA ACA ATA ATC R: TCA ACT CAG CAT CTC TGT CC	Ni *et al.* (2002)
37	RM316	9	$(GT)_8$-$(TG)_9(TTTG)_4$-$(TG)_4$	1.8	F: CTA GTT GGG CAT ACG ATG GC R: ACG CTT ATA TGT TAC GTC AAC	Lu *et al.* (2005)
38	RM171	10	$(GATG)_5$	92.8	F: AAC GCG AGG ACA CGT ACT TAC R: ACG AGA TC\AC GTA CGC CTT TG	Ni *et al.* (2002), Lu *et al.* (2005), Chakravarthi and Naravaneni (2006), Thomson *et al.* (2007)
39	RM271	10	$(GA)_{15}$	59.4	F: TCA GAT CTA CAA TTC CAT CC R: TCG GTG AGA CCT AGA GAG CC	Ni *et al.* (2002), Lu *et al.* (2005)
40	RM333	10	$(TAT)_{19}$-$(CTT)_{19}$	110.4	F: GTA CGA CTA CGA GTG TCA CCA A R: GTC TTC GCG ATC ACT CGC	Lu *et al.* (2005), Chakravarthi and Naravaneni (2006), Lapitan *et al.* (2007), Sundaram *et al.* (2008)
41	RM474	10	$(AT)_{13}$	0.0	F: AAG ATG TAC GGG TGG CAT TC R: TAT GAG CTG GTG AGC AAT GG	Lu *et al.* (2005), Thomson *et al.* (2007), Sundaram *et al.* (2008)

Table 2: (Continued)

No	SSR Marker	Chromosome	Repeat Motif	Position (cM)	Primer Sequence (5' – 3')	Source
42	RM144	11	$(ATT)_{11}$	123.2	F: TGC CCT GGC GCA AAT TTG ATC C R: GCT AGA GGA GAT CAG ATG GTA GTG CAT C	Lu *et al.* (2005), Chakravarthi and Naravaneni (2006), Thomson *et al.* (2007)
43	RM202	11	$(CT)_{30}$	54.0	F: CAG ATT GGA GAT GAA GTC CTC C R: CCA GCA AGC ATG TCA ATG TA	Bligh *et al.* (1999), Ni *et al.* (2002), Lu *et al.* (2005), Alvarez *et al.* (2007), Lapitan *et al.* (2007), Sundaram *et al.* (2008)
44	RM206	11	$(CT)_{21}$	102.9	F: CCC ATG CGT TTA ACT ATT CT R: CGT TCC ATC GAT CCG TAT GG	Lu *et al.* (2005), Chakravarthi and Naravaneni (2006), Lapitan *et al.* (2007), Ram *et al.* (2007), Sundaram *et al.* (2008)
45	RM287	11	$(GA)_{21}$	68.6	F: TTC CCT GTT AAG AGA GAA ATC R: GTG TAT TTG GTG AAA GCA AC	Ni *et al.* (2002), Lu *et al.* (2005), Thomson *et al.* (2007)
46	RM19	12	$(ATC)_{10}$	20.9	F: CAA AAA CAG AGC AGA TGA C R: CTC AAG ATG GAC GCC AAG A	Bligh *et al.* (1999), Lu *et al.* (2005), Chakravarthi and Naravaneni (2006), Thomson *et al.* (2007)

Table 2: (Continued)

No	SSR Marker	Chromosome	Repeat Motif	Position (cM)	Primer Sequence (5' – 3')	Source
47	RM235	12	$(CT)_{24}$	101.8-103.8	F: AGA AGC TAG GGC TAA CGA AC R: TCA CCT GGT CAG CCT CTT TC	Ni *et al.* (2002), Lapitan *et al.* (2007), Ram *et al.* (2007), Sundaram *et al.* (2008)
48	RM277	12	$(GA)_{11}$	57.2	F: CGG TCA AAT CAT CAC CTG AC R: CAA GGC TTG CAA GGG AAG	Thomson *et al.* (2007)
49	RM463	12	$(TTAT)_5$	75.5	F: TTC CCC TCC TTT TAT GGT GC R: TGT TCT CCT CAG TCA CTG CG	Lapitan *et al.* (2007)

Note: F: Forward primer sequence; R: Reverse primer sequence

3.2.3 Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) analysis follows the procedures recommended by the manufacturer (Vivantis). It was performed in 25 µl reaction mixture containing 2 µl DNA, 0.2 µM of each primer, 0.08 mM of each dNTP and 2.0 U Taq DNA polymerase in 1 X Vi buffer [100 mM Tris-HCl pH 9.1, 500 mM KCl, 1.5 mM $MgCl_2$, and 0.1% Triton™ X-100]. The PCR was performed using a XP Thermal Cycler Blocks (Model: TC-XP, Bioer Technology, China) according to the cycle profile: initial denaturation at 95 °C for 3 min and then 32 cycles of 30 second denaturation at 95 °C, 30 second annealing at the temperature depending on the marker used (55 °C, 61 °C, 67 °C), and 1 minute extension at 72 °C and 7 minute at 72 °C for final product extension. The amplified samples were stored at -20 °C until further use.

3.2.4 Gel Electrophoresis

3.2.4.1 Agarose Gel Electrophoresis

PCR amplified samples were subjected to electrophoresis in 2.0% agarose gels at 100 volts for 1.5 hour in 0.5x Tris-Borate-EDTA (1 M Tris-Cl, 0.5 M EDTA, Boric acid) buffer and electrophoresis apparatus (Fisher Scientific FB300). Five-microlitre aliquot of the amplified samples was combined with 1 µl loading dye. The gels were then stained in ethidium bromide (0.5 µg/ml) for 20 minutes, de-stained for 15 – 30 minutes with distilled water and observed under a UV transilluminator. The images were documented in a gel documentation system

(Model: FluorChem 5500, Alpha Innotech, USA). The sizes of the amplified bands were estimated by using AlphaEaseFC Version 3.3.2 software utility of the gel documentation system using 'Gene Ruler™ Ultra Low Range' DNA ladder (MBI Fermentas) and VC 100bp DNA ladder (Vivantis) as the size standards for measuring the molecular weights.

3.2.4.2 Microsatelite Markers Selection

Only 12 polymorphic rice microsatellite primers were selected from 49 microsatellite primers after visualization with the agarose gels. The selected SSR primers were ran on polyacrylamide gels to achieve better resolution of the bands.

3.2.4.3 Polyacrylamide Gel Electrophoresis

The PCR products from selected SSR primers were subjected to electrophoresis in 5% polyacrylamide gel in 0.5x TBE buffer at 80 volts using Mini-PROTEAN®3 (Bio-Rad, USA) where the running time depended on the size of the PCR products, from 45 to 50 minutes. The gels were stained with ethidium bromide (0.5 μg/ml) for 20 minutes. DNA bands were visualized under UV light using gel documentation system (Model: FluorChem 5500, Alpha Innotech, USA). The bands were scored for further analysis.

3.3 Data Analysis

Clear and unambiguous polymorphic products from three rounds of microsatellite analysis were scored qualitatively for presence (1) and absence (0) for each marker allele-genotype combination in Microsoft Excel 2003 file. The binary data was imported into NT Edit of 'Numerical Taxonomy and Multivariate Analysis System' (NTSYSpc) Version 2.20r N (Rohlf, 2005; Exeter Software, Setauket, USA). The (0/1) matrix was used to calculate Similarity as DICE coefficient using 'SIMQUAL' subroutine in 'SIMILARITY' routine. The resulting similarity matrix was employed to construct dendrograms using 'Sequential Agglomerative Hierarchical Nesting' (SAHN) based on 'Unweighted Pair Group Method with Arithmetic Means' (UPGMA) to infer genetic relationships and phylogeny. The pairwise similarity coefficients among varieties were quantified according to Nei (1978). Cophenetic value was generated and the Mantel matrix correspondence test (Mantel, 1967) was used to compare cophenetic with similarity matrices to define the degree of congruence in the estimation of genetic relationships by the SSR marker. The Mantel matrix correspondence test was carried out using the MXCOMP procedure in NTSYSpc Version 2.20r N.

Most informative primers were selected based on the extent of polymorphism. Botstein *et al.* (1980) originally introduced the term polymorphism information content (PIC) into human genetics. It refers to the value of a marker for detecting polymorphism within a population, depending on the number of detectable alleles and the distribution of their frequency. In the present study, polymorphism information content (PIC) was calculated by applying the formula simplified by Anderson *et al.* (1993):

$$PIC = 1 - \sum_{j=1}^{n} P^2_{ij}$$

Where P_{ij} is the frequency of j^{th} allele for marker i and the summation extends over *n* alleles.

CHAPTER 4

RESULTS

4.1 Microsatellite Markers Screening

In order to study the genetic diversity of the rice cultivars, 49 microsatellite markers were screened and analysed using agarose gel and grouped into two groups: (a) polymorphic or (b) monomorphic. The results are shown in Table 3. Four SSR markers were tested for each chromosome except chromosome 3 with five SSR markers. Twenty-eight SSR markers were found to be polymorphic. The rest (21 markers) were found to be monomorphic. In chromosome 1, RM1 showed greater variation in banding pattern compared to RM237 that showed lower variation in the banding pattern. Banding pattern analyzed in chromosome 2 found only RM240 to be polymorphic. In evaluating the cultivars at chromosome 3, microsatellite markers RM22 and RM 489 showed slight polymorphism compared to the other three SSR markers. However, RM489 showed more distinctiveness in the banding pattern compared to RM22. It was also found that chromosome 4 had the same trend as chromosome 3. However, only microsatellite marker RM252 showed polymorphism at chromosome 4.

Chromosome 5 has only one microsatellite marker (RM413) that showed polymorphism while RM334 did not show any allele at all. In the preliminary screening at chromosome 6, two microsatellite markers, RM204 and RM541

showed polymorphism. Out of the four microsatellite markers screened in chromosome 7, only one microsatellite marker (RM11) appeared to be polymorphic. Three SSR markers in chromosome 8 were found to be polymorphic with RM404 showing better difference compared to RM149 and RM152.

Polymorphism of microsatellite markers at chromosome 9 resulted in attaining one marker that showed polymorphism which is RM316. At the same time, microsatellite markers for chromosome 10 produced two SSR markers (RM271 and RM333) that showed polymorphic banding patterns. Out of the twelve chromosomes screened for polymorphic SSR markers, chromosome 11 provides more choice in selecting polymorphic markers. All four SSR markers screened in chromosome 11 showed polymorphism with RM206 showing greater competency in differentiating the cultivars. Only RM19 in chromosome 12 showed polymorphic banding pattern. For each chromosome, only one SSR marker showing polymorphism was selected to be further analyzed. The selected SSR markers were RM1, RM240, RM489, RM252, RM413, RM204, RM11, RM404, RM316, RM271, RM206 and RM19 due to their ability to differentiate the cultivars compared to the other 37 SSR markers.

Table 3: Preliminary screening of 49 SSR markers and their polymorphism

	Polymorphism	SSR Markers
1	Polymorphic	RM1, RM237, RM240, RM252, RM413, RM204, RM541, RM11, RM149, RM152, RM404, RM316, RM271, RM144, RM202, RM206, RM287, RM19, RM154, RM452, RM22, RM489, RM161, RM421, RM138, RM215, RM278, RM333
2	Monomorphic	RM226, RM243, RM207, RM55, RM338, RM514, RM124, RM307, RM335, RM133, RM118, RM125, RM284, RM105, RM171, RM474, RM235, RM277, RM463, RM334, RM162

4.2 Selected Microsatellite Markers Analysis

Microsatellite markers that were subjected to further analysis after the preliminary screening were RM1, RM240, RM489, RM252, RM413, RM204, RM11, RM404, RM316, RM271, RM206 and RM19. A total of 34 alleles were identified and assessed.

4.2.1 Microsatellite Primer RM1

Five alleles were detected using the microsatellite primer RM1 (Figure 1). Number of alleles detected per cultivar was one allele per cultivar. The molecular weights of the alleles ranged from 92 to 129 base pairs. The microsatellite primer RM1 was considered polymorphic as it can detect and help to differentiate the cultivars better.

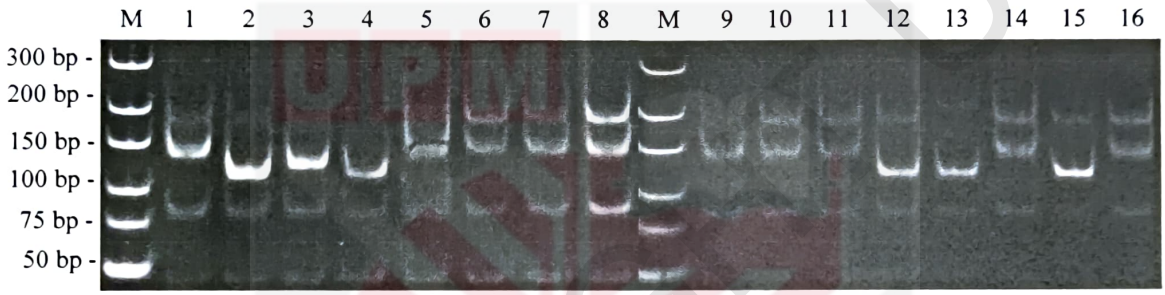
Note: 5 % polyacrylamide gel, 80 volt, 45 minutes

M:	Ultra Low Range DNA Ladder (10bp)	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 1: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM1

4.2.2 Microsatellite Primer RM240

Two alleles were detected using the microsatellite primer RM240. All the cultivars showed only one allele that was amplified (Figure 2). Their molecular weight ranged from 113 to 146 base pairs.



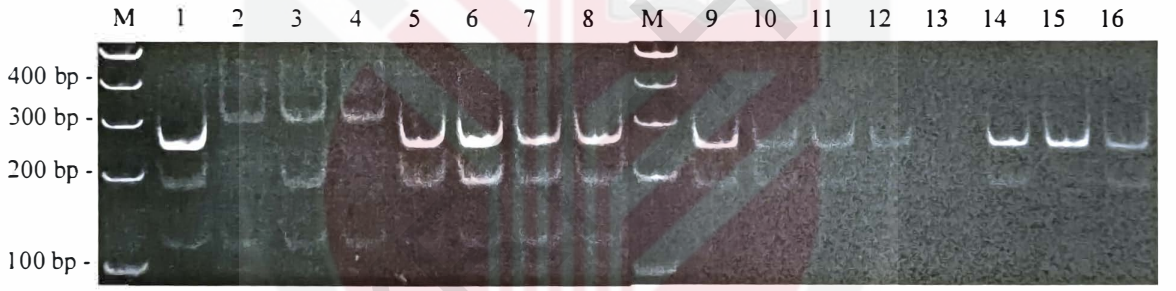
Note: 5 % polyacrylamide gel, 80 volt, 45 minutes

M:	Ultra Low Range DNA Ladder (10bp)	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 2: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM240

4.2.3 Microsatellite Primer RM489

The microsatellite primer RM489 detected two alleles at this locus (Figure 3). The number of alleles detected from all the cultivars was only one allele per cultivar and the molecular weights of the alleles ranged from 251 to 315 base pairs. The expected PCR product size was 271 base pairs. The microsatellite primer RM489 had a slight ability in distinguishing the cultivars except for cultivars, Bukit Wangi, Buntar B, Empawah and Sampangan B.



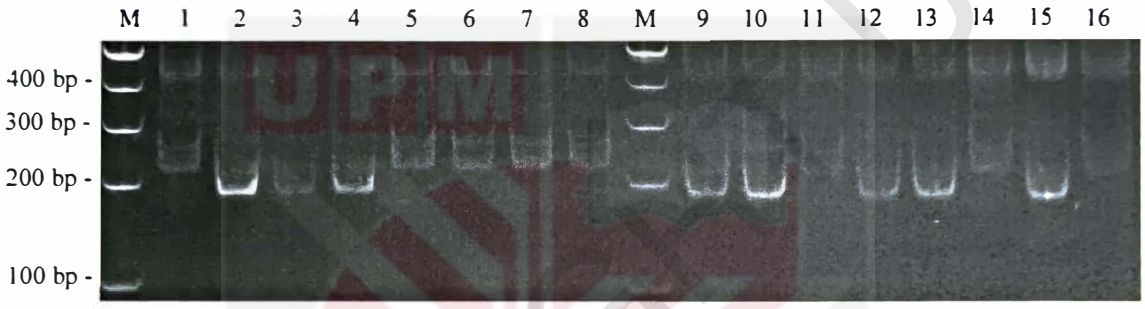
Note: 5 % polyacrylamide gel, 80 volt, 50 minutes

M:	VC 100 bp DNA Ladder	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 3: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM489

4.2.4 Microsatellite Primer RM252

Only two alleles were detected using microsatellite primer RM252. Only one allele was amplified for each of the sixteen cultivars (Figure 4). Their molecular weight ranged from 189 to 238 base pairs.



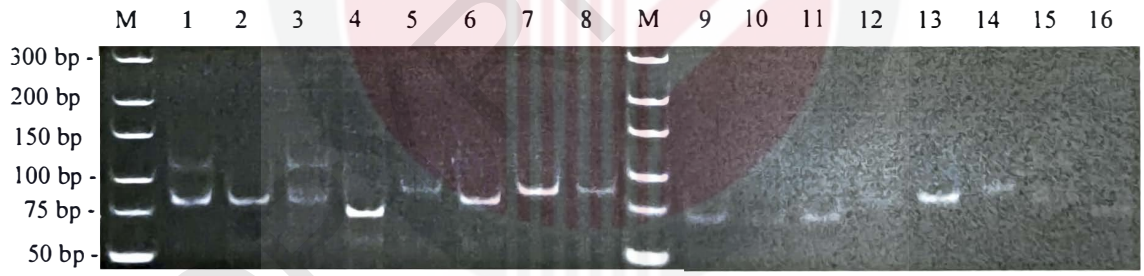
Note: 5 % polyacrylamide gel, 80 volt, 50 minutes

M:	VC 100 bp DNA Ladder	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 4: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM252

4.2.5 Microsatellite Primer RM413

Three alleles were detected using microsatellite primer RM413 (Figure 5). The number of alleles detected from all the cultivars was only one allele per cultivar and the molecular weights of the alleles ranged from 70 base pairs until 118 base pairs. The expected PCR product size was 79 base pairs. The cultivars can be grouped according to the three alleles. Adan Sederhana, Bukit Wangi, Buntar B, Lasak, Rotan Wangi, Sampangan B and Tulang were detected at the same molecular weights while Empawah, Mamut, Mamut 2, Muyun and Wangi can be grouped into one group. Empawah Merah, Lemak (Meradong and Roban Area) and Sia were found to be in another group.



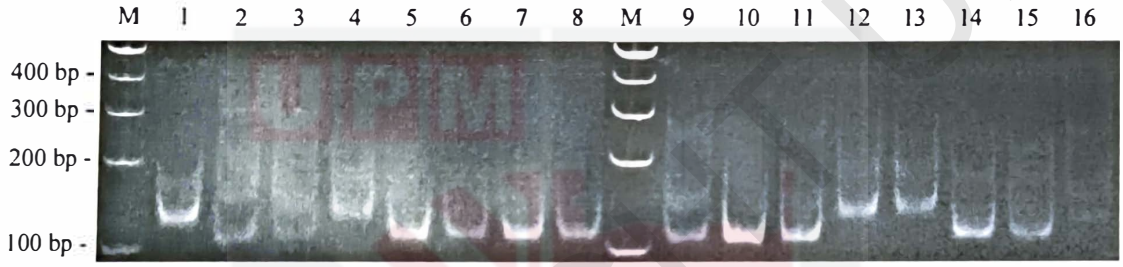
Note: 5 % polyacrylamide gel, 80 volt, 45 minutes

M:	Ultra Low Range DNA Ladder (10bp)	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 5: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM413

4.2.6 Microsatellite Primer RM204

Three alleles were detected using microsatellite primer RM204. Most of the cultivars showed only one allele that was amplified (Figure 6). Their molecular weights ranged from 109 to 156 base pairs.



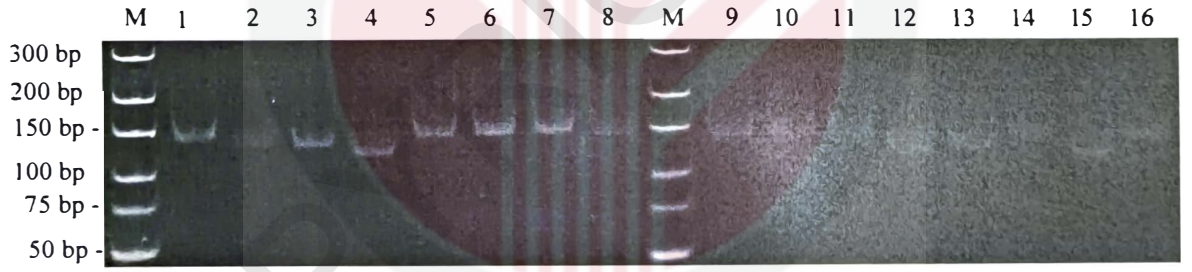
Note: 5 % polyacrylamide gel, 80 volt, 50 minutes

M:	VC 100 bp DNA Ladder	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 6: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM204

4.2.7 Microsatellite Primer RM11

Only three alleles were detected using microsatellite primer RM11 (Figure 7). The number of alleles detected from all the cultivars was only one allele per cultivar and the molecular weights of the alleles range from 127 to 154 base pairs. The expected PCR product size was at 140 base pairs. The cultivars can be grouped into three groups according to their banding pattern. The first group consists of Adan Sederhana, Bukit Wangi, Buntar B, Sia and Wangi. The second group consists of Empawah, Rotan Wangi, Sampangan B and Tulang. The last group consists of Empawah Merah, Lasak, Lemak (Meradong and Roban Area), Mamut, Mamut 2 and Muyun.



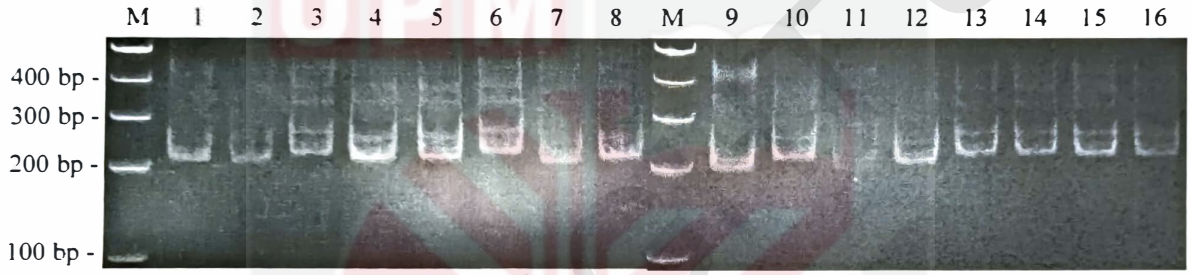
Note: 5 % polyacrylamide gel, 80 volt, 45 minutes

M:	Ultra Low Range DNA Ladder (10bp)	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 7: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM11

4.2.8 Microsatellite Primer RM404

Only three alleles were detected across the cultivars where all the cultivars showed only one allele that was amplified. The alleles could be divided into three groups using their molecular weight which ranged from 204 to 234 base pairs (Figure 8). Most of the alleles were found in between 210 to 230 base pairs.



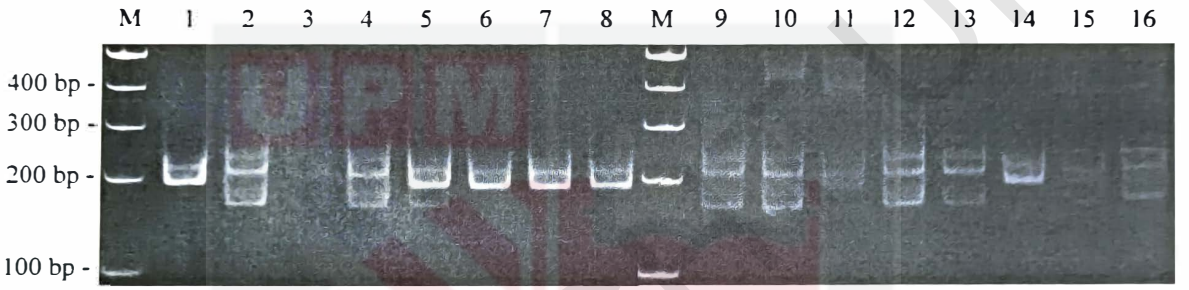
Note: 5 % polyacrylamide gel, 80 volt, 50 minutes

M:	VC 100 bp DNA Ladder	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 8: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM404

4.2.9 Microsatellite Primer RM316

Only two alleles were detected using microsatellite primer RM316 (Figure 9). The number of alleles detected from all the cultivars was only one allele per cultivar. The molecular weights ranged from 161 to 199 base pairs.



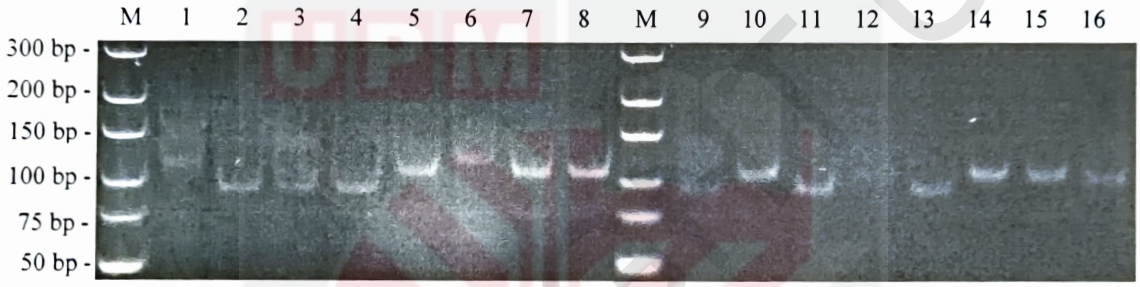
Note: 5 % polyacrylamide gel, 80 volt, 50 minutes

M:	VC 100 bp DNA Ladder	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 9: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM316

4.2.10 Microsatellite Primer RM271

Three alleles were detected using this microsatellite primer. All the cultivars showed only one amplified allele (Figure 10). Their molecular weights ranged from 94 to 122 base pairs.



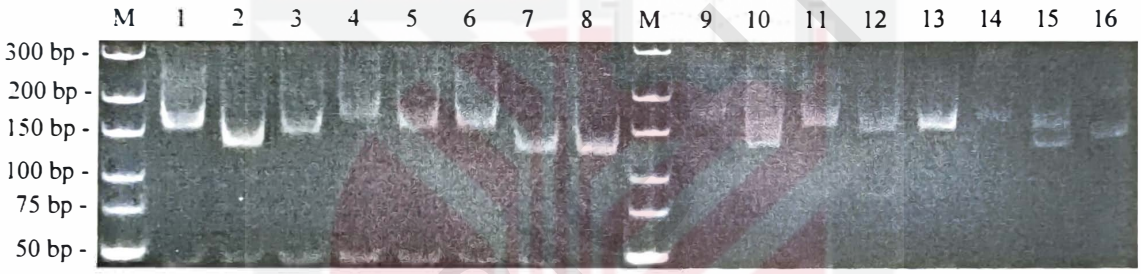
Note: 5 % polyacrylamide gel, 80 volt, 45 minutes

M:	Ultra Low Range DNA Ladder (10bp)	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 10: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM271

4.2.11 Microsatellite Primer RM206

Three alleles were detected using microsatellite primer RM206 (Figure 11). The number of alleles detected from all the cultivars was only one allele per cultivar and the molecular weights of the alleles ranged from 133 base pairs until 174 base pairs. The expected PCR product size is at 147 base pairs. Most of the alleles' sizes were found to be between 150 to 160 base pairs.



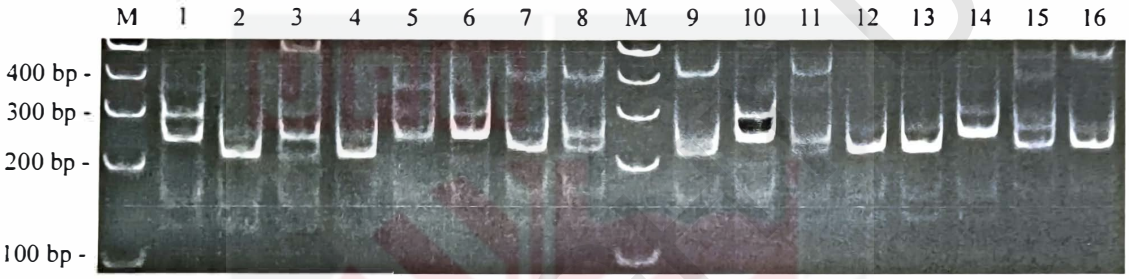
Note: 5 % polyacrylamide gel, 80 volt, 45 minutes

M:	Ultra Low Range DNA Ladder (10bp)	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 11: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM206

4.2.12 Microsatellite Primer RM19

Microsatellite primer RM 19 detected three alleles (Figure 12). The number of alleles detected per cultivar was from one to two alleles. Only Buntar B rice cultivar had two alleles while the remaining cultivars had only one alleles. The alleles' molecular weights ranged from 212 to 308 base pairs.



Note: 5 % polyacrylamide gel, 80 volt, 50 minutes

M:	VC 100 bp DNA Ladder	bp:	base pairs
1:	Adan Sederhana	9:	Mamut
2:	Bukit Wangi	10:	Mamut 2
3:	Buntar B	11:	Muyun
4:	Empawah	12:	Rotan Wangi
5:	Empawah Merah	13:	Sampangan B
6:	Lasak	14:	Sia
7:	Lemak (Meradong)	15:	Tulang
8:	Lemak (Roban Area)	16:	Wangi

Figure 12: PAGE analysis of selected Sarawak rice cultivars using microsatellite primer RM19

4.3 Allele Number and Allele Frequency

Table 4 summarizes the results obtained based on the analysis of the sixteen cultivars using twelve selected polymorphic microsatellite markers. A total of 34 alleles were detected at the twelve microsatellite loci evaluated in the sixteen cultivars. The number of alleles per microsatellite primer ranged from two to five alleles. The highest number of alleles detected in this study came from microsatellite primers RM1 with five alleles identified. Three alleles were found using microsatellite primer RM413, RM204, RM11, RM404, RM271, RM206 and RM19. Microsatellite primer RM240, RM489, RM252 and RM316 detected only two alleles each. The average number of alleles per locus was 2.83.

4.4 SSR Analysis

The polymorphism information content (PIC) values are a reflection of allele diversity and frequency among the cultivars and varies from one locus to another. The PIC values are recorded in Table 4. The PIC value ranged from 0.375 (RM489) to 0.703 (RM1). The most polymorphic microsatellite marker detected in this study was RM1 with PIC value of 0.703 and five alleles identified. The second highest was RM19 that found three alleles with the PIC value of 0.651. It was followed by RM413, RM11 and RM206 with a total of three alleles found and PIC value of 0.648 each. Meanwhile, RM404 detected three alleles with PIC value of 0.602. Microsatellite markers RM271, RM252, RM204, RM240 and RM316 had PIC values below 0.6, which are 0.594, 0.500, 0.477, 0.469, and 0.430,

respectively. The lowest PIC value of 0.375 was detected by RM489, which showed only slight polymorphism in polyacrylamide gels.

Table 4: Polymorphism information content (PIC) of the selected simple sequence repeat (SSR) markers

SSR markers	Chromosome	No. of alleles	PIC value	Repeat Motif
RM1	1	5	0.703	$(AG)_{26}$
RM240	2	2	0.469	$(CT)_{21}$
RM489	3	2	0.375	$(ATA)_8$
RM252	4	2	0.500	$(CT)_{19}$
RM413	5	3	0.648	$(AG)_{11}$
RM204	6	3	0.477	$(CT)_{44}$
RM11	7	3	0.648	$(GA)_{17}$
RM404	8	3	0.602	$(GA)_{33}$
RM316	9	2	0.430	$(GT)_8$-$(TG)_9(TTTG)_4(TG)_4$
RM271	10	3	0.594	$(GA)_{15}$
RM206	11	3	0.648	$(CT)_{21}$
RM19	12	3	0.651	$(ATC)_{10}$
Total		34		
Mean		2.83	0.562	

Note: SSR: Simple sequence repeat (microsatellite markers);
PIC: Polymorphism information content

4.5 Genetic Relationship between Rice Cultivars

Data on presence (1) or absence (0) of alleles detected across cultivars and microsatelite markers were constructed (Table 5). This data is important for generating the genetic similarity values among the rice cultivars (Table 6), and subsequent construction of the dendrogram presented in Figure 13. At a 26.21% level of similarity, the UPGMA cluster diagram showed two main clusters, A and

B with additional sub-clusters within each group. This dendrogram revealed the cultivars, which are genetically similar. Cluster A contained ten cultivars which are Adan Sederhana, Empawah Merah, Lasak, Lemak (Meradong, Roban Area), Mamut, Mamut 2, Muyun, Sia and Wangi. The similarity coefficients between any two rice cultivars in this cluster ranged from 50 to 75%. This cluster could be further divided into two groups. Group 1 clustered at a similarity coefficient of about 50%, consisted of six cultivars which were Adan Sederhana, Empawah Merah, Lasak, Lemak (Meradong, Roban Area) and Sia. Group 1 was further divided into two subgroup at similarity coefficients of 63.89%. In this subgroup, cultivars with the same designation called Lemak, but from two different areas (Meradong and Roban) were clustered as 100% similar. Empawah was also found to be 75% similar to both Lemak cultivars. From the dendrogram, Lasak was shown to be 66.67% similar to Sia. Adan Sederhana did not cluster closely with any of the cultivars in Group 1.

Group 2 was less diverse than the first group and was clustered at a similarity coefficient of about 50%, which consist of four cultivars (Mamut, Mamut 2, Muyun and Wangi). In this group, Wangi cultivar was only 50% similar to other cultivars in Group 2. At 58.33% similarity, it further divided the remaining cultivars into another smaller group. Muyun was found to be 58.33% similar to Mamut and Mamut 2. Mamut was found to be 75% similar to Mamut 2.

The second main group, cluster B was clustered at similarity coefficients of 53.21%. This cluster had six cultivars which were Bukit Wangi, Buntar B, Empawah, Sampangan B, Rotan Wangi and Tulang. The cluster could be further

subdivided into two subgroups with varying levels of similarity. Group 3 consisting of Bukit Wangi and Buntar B were clustered at a similarity coefficient of 72%. Group 4, the second subgroup, in Group B consists of Empawah, Sampangan B, Rotan Wangi and Tulang are 60.42% similar. Empawah and Sampangan B were further clustered at 66.67% while Rotan Wangi and Tulang were clustered at 83.33% similarity.

4.6 Matrix Comparison

Estimation of genetic relationship between the similarity matrix obtained by the SSR analysis and corresponding dendrogram to define the degree of congruence revealed a cophenetic correlation of 0.78. This estimation showed that the genetic relationships of the cultivars studied using the SSR markers was 78% reliable.

Table 5: Present (1) - absent (0) data according to molecular weights for selected Sarawak rice cultivars

Marker	Present (1) or Absent (0) Data															
	Rice Cultivars*															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
RM1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
RM1	0	1	0	1	0	0	0	0	1	1	0	1	0	0	1	1
RM1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
RM1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0
RM1	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0
RM240	0	1	1	1	0	0	0	0	0	0	0	1	1	0	1	0
RM240	1	0	0	0	1	1	1	1	1	1	1	0	0	1	0	1
RM489	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1	1
RM489	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0
RM252	0	1	1	1	0	0	0	0	1	1	0	1	1	0	1	0
RM252	1	0	0	0	1	1	1	1	0	0	1	0	0	1	0	1
RM413	0	0	0	1	0	0	0	0	1	1	1	0	0	0	0	1
RM413	1	1	1	0	0	1	0	0	0	0	0	1	1	0	1	0
RM413	0	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0
RM204	0	1	1	0	1	1	1	1	1	1	1	0	0	1	1	0
RM204	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
RM204	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0
RM11	0	0	0	1	0	0	0	0	0	0	0	1	1	0	1	0
RM11	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1
RM11	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0

Table 5: (Continued)

Marker	Present (1) or Absent (0) Data															
	Rice Cultivars*															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
RM404	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
RM404	0	1	0	1	0	0	0	0	1	1	0	1	1	0	1	0
RM404	0	0	0	0	1	1	1	1	0	0	1	0	0	1	0	1
RM316	0	1	1	1	1	0	0	0	1	1	1	1	1	0	1	1
RM316	1	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0
RM271	0	1	1	1	0	0	0	0	1	0	1	0	1	0	0	0
RM271	0	0	0	0	1	0	1	1	0	1	0	1	0	1	1	1
RM271	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
RM206	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1	0
RM206	1	0	1	0	1	1	0	0	0	0	0	1	1	0	0	1
RM206	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0	0
RM19	0	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0
RM19	0	0	0	0	0	0	1	1	0	0	1	1	1	0	1	1
RM19	1	0	1	0	1	1	0	0	0	1	0	0	0	1	0	0

Note: 1: Presence of alleles 0: Absence of alleles

*Rice cultivars:

1:	Adan Sederhana	6:	Lasak	11:	Muyun	16:	Wangi
2:	Bukit Wangi	7:	Lemak (Meradong)	12:	Rotan Wangi		
3:	Buntar B	8:	Lemak (Roban Area)	13:	Sampangan B		
4:	Empawah	9:	Mamut	14:	Sia		
5:	Empawah Merah	10:	Mamut 2	15:	Tulang		

Table 6: Genetic similarities between the selected Sarawak rice cultivars

Rice cultivars*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1.0000															
2	0.1667	1.0000														
3	0.4000	0.7200	1.0000													
4	0.0000	0.6667	0.4800	1.0000												
5	0.5000	0.1667	0.3200	0.0833	1.0000											
6	0.6667	0.1667	0.3200	0.0000	0.6667	1.0000										
7	0.4167	0.1667	0.0800	0.0000	0.7500	0.5833	1.0000									
8	0.4167	0.1667	0.0800	0.0000	0.7500	0.5833	1.0000	1.0000								
9	0.1667	0.5833	0.4000	0.6667	0.4167	0.3333	0.3333	0.3333	1.0000							
10	0.2500	0.5000	0.3200	0.4167	0.5833	0.4167	0.5000	0.5000	0.7500	1.0000						
11	0.2500	0.2500	0.2400	0.3333	0.5833	0.5833	0.5833	0.5833	0.6667	0.5000	1.0000					
12	0.2500	0.5000	0.4000	0.5833	0.3333	0.2500	0.2500	0.2500	0.4167	0.5000	0.2500	1.0000				
13	0.1667	0.5833	0.5600	0.6667	0.1667	0.1667	0.0833	0.0833	0.3333	0.2500	0.2500	0.7500	1.0000			

Table 6: (Continued)

Rice cultivars*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
14	0.5000	0.1667	0.24000	0.0833	0.6667	0.6667	0.6667	0.6667	0.3333	0.4167	0.5833	0.1667	0.0000	1.0000		
15	0.1667	0.6667	0.4000	0.5000	0.3333	0.2500	0.4167	0.4167	0.5000	0.6667	0.3333	0.8333	0.5833	0.2500	1.0000	
16	0.5000	0.2500	0.2400	0.2500	0.5833	0.4167	0.5000	0.5000	0.4167	0.5000	0.5833	0.5000	0.2500	0.5000	0.4167	1.0000

Note: *Rice cultivars:

1: Adan Sederhana
2: Bukit Wangi
3: Buntar B
4: Empawah
5: Empawah Merah
6: Lasak
7: Lemak (Meradong)
8: Lemak (Roban Area)
9: Mamut
10: Mamut 2
11: Muyun
12: Rotan Wangi
13: Sampangan B
14: Sia
15: Tulang
16: Wangi

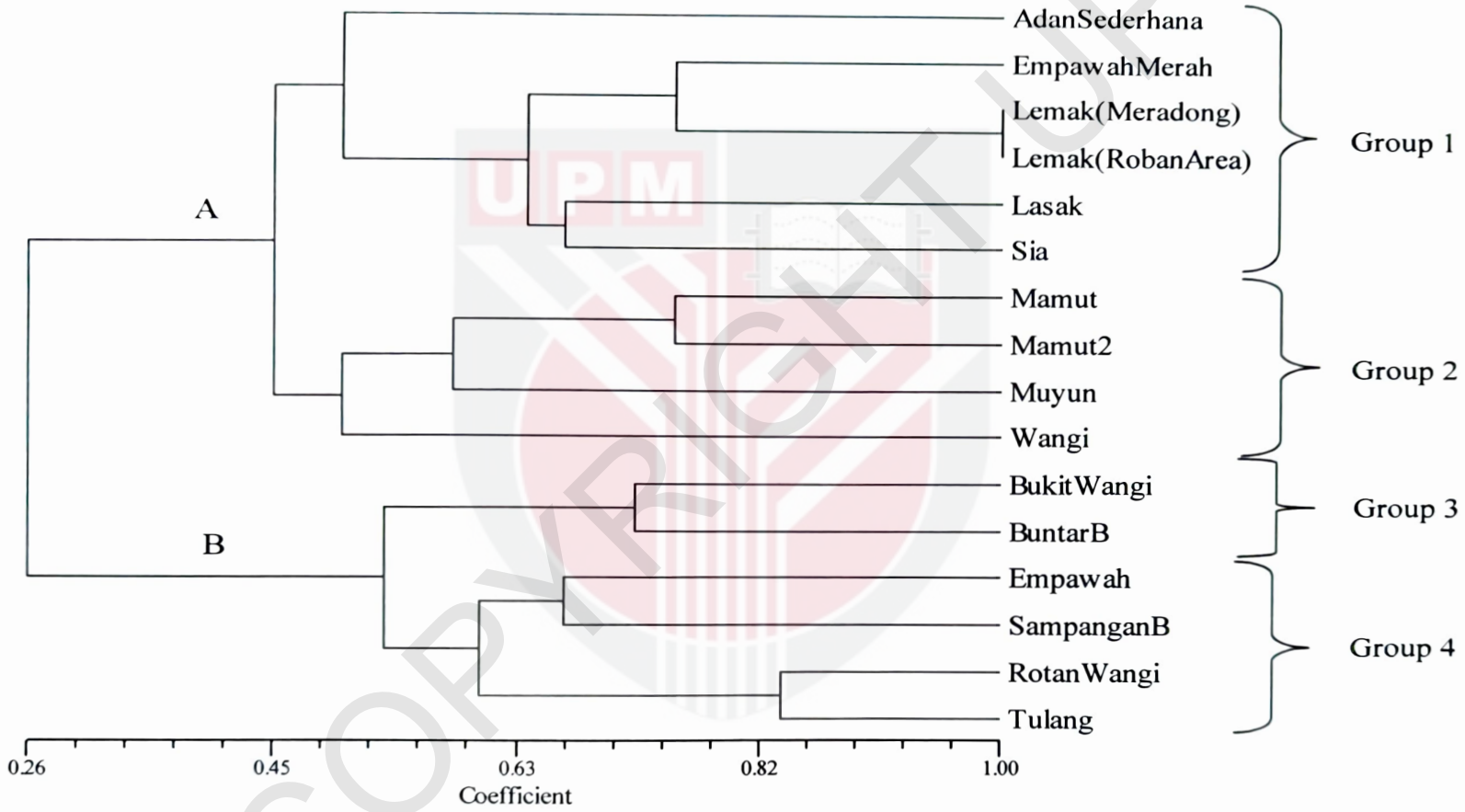


Figure 13: Genetic relationship between 16 selected Sarawak rice cultivars

CHAPTER 5

DISCUSSION

5.1 SSR Analysis

High levels of polymorphism and numbers of alleles were expected for rice microsatellite markers. Out of the 49 markers, 28 markers were found to be polymorphic and 21 markers were monomorphic. Although all the 49 microsatellite markers were selected according to high PIC values reviewed from the literature, 21 markers were found to be monomorphic markers. These monomorphic markers are probably not suitable to assess the genetic diversity of the rice cultivars in Sarawak due to different rice accession origins compared to the ones in the literature.

The 12 SSR markers detected moderate levels of polymorphism among these rice cultivars with at least two alleles identified at all loci and highest number of alleles obtained from RM1 locus where five alleles were observed. These allele numbers, as well as the average allele number per locus in this study were less than those reported by Jayamani *et al.* (2007) who observed an average of 7.7 per locus over SSR loci and as many as 16 alleles at one locus. Ni *et al.* (2002) used 111 microsatellite markers to evaluate 38 rice cultivars and found an average of 6.8 alleles, ranging from one to 17 alleles per marker. However, the number of alleles and the average number of alleles obtained in this study were higher than the

average of 2.6 alleles per locus reported by Joshi and Behera (2007). Lower allelic diversity may be caused by the cultivars' genotypes that have similar sequence at the specific locus of the chromosome (Joshi and Behera, 2007).

PIC value is a reflection of allele diversity and frequency among the varieties, but in the present study, it was not uniformly high for all the SSR loci tested. The average PIC value was 0.562 and it ranged from a low of 0.375 (RM489) to a high 0.703 (RM1). The PIC values are dependent on the genetic diversity of the accessions chosen and this study had a high proportion of closely related cultivars, which would have an effect on the PIC values. Several studies had assessed the allelic diversity of the same markers used in this study. Panaud *et al.* (1996) and Olufowote *et al.* (1997) assessed marker RM1 and both found very similar numbers of alleles and obtained slightly higher PIC values than in this study. Blair *et al.* (2002) also found PIC value of 0.88 in 59 rice cultivars using microsatellite primer RM1. As for marker RM240 in chromosome 2, the PIC value found in this study was 0.469 which was lower than the PIC value found by Lapitan *et al.* (2007) and Ram *et al.* (2007). They found PIC values of 0.750 and 0.832 in 24 and 35 rice cultivars respectively. Marker RM489 showed a low PIC value of 0.375 even though it was listed in the panel of 50 rice SSR marker recommended by the Gramene database. Although Blair *et al.* (2002) and Lapitan *et al.* (2007) found PIC values of 0.943 and 0.900 using marker RM252, the PIC value obtained in this study was much lower (0.500). A PIC value of 0.648 found with marker RM413 was lower than the value of 0.770 reported by Thomson *et al.* (2007). In chromosome 6, RM204 was assessed and the PIC value of 0.477 was obtained.

This PIC value was slightly lower than the 0.551 that was obtained by Ram *et al.* (2007).

The PIC value obtained from RM11 in the present study was 0.648, lower value than the one reported by Alvarez *et al.* (2007) and Thomson *et al.* (2007). They reported PIC values of 0.76 and 0.82, respectively. Lapitan *et al.* (2007) reported a PIC value of 0.85 when assessing RM404 across 24 rice cultivars. In this study, marker RM404 assessed gave a PIC value of 0.602, which is lower than that reported by Lapitan *et al.* (2007). Meanwhile, RM316 and RM271 gave PIC values of 0.430 and 0.594. The marker RM316 was considerably not very polymorphic in this study as it had the second lowest PIC value among the 12 SSR markers assessed. The study conducted by Lapitan *et al.* (2007) and Ram *et al.* (2007) obtained higher PIC values with marker RM206. They reported PIC values of 0.87 and 0.891 respectively, which was higher than the one obtained in this study (0.648). For RM19, the observed PIC value of 0.651 was also lower than the PIC value of 0.75 reported by Olufowote *et al.* (1997). According to De Woody *et al.* (1995), markers with PIC values 0.5 or higher are highly informative for genetic studies and are extremely useful in distinguishing the polymorphism rate of a marker at a specific locus. Therefore, markers RM1, RM252, RM413, RM204, RM11, RM404, RM271, RM206 and RM19 would be useful in the genetic studies. Out of the 12 markers, RM1 and RM19 were the better markers to assess genetic diversity of rice cultivars as they had the highest PIC values.

While genetic diversity of the accessions chosen would have an effect on the PIC values, other hypotheses including residual heterozygosity, accidental seed

mixtures, and mutation and outcrossing may have also contributed to the observed polymorphism (Jain *et al.*, 2004). Higher number of alleles (2.83) and average PIC values (0.562) according to Joshi and Behera (2007) and De Woody *et al.* (1995) detected in the present study indicates that the Sarawak rice cultivars are a good source of genetic variability to be explored.

5.2 Genetic Relationship between Rice Cultivars

The cluster analysis showed a significant genetic variation among the rice cultivars studied with genetic similarities between any two cultivars varying between 0.0800 to 1.000. Cultivars found to be poorly related were Adan Sederhana and Empawah, Adan Sederhana and Sia, and Empawah with Lasak and both of cultivars of Lemak. Two cultivars, which had 100% similarity, are Lemak from the different areas; Meradong and Roban. SSR markers detected them as cultivars with the same genotype. No other pairs of cultivars except Rotan Wangi and Tulang obtained similarity above 80%. Meanwhile, Mamut and Mamut 2 have almost similar name and they are 75% genetically similar. This indicates that they probably belong to similar genetic background, which can be explained by Brondani *et al.* (2006), where old variety after years of successive cultivation would generate populations with a genetic constitution different from original genotype due to adaptations to different environments. Tu *et al.* (2007) also remarked that variation in geographic and climate conditions, diverse farming practices and diversified utilization of rice may result in loss of identity of the varieties grown and attributes to the genetic differentiation and diversity. All this

supports the argument by Ni *et al.* (2002) that information on genetic diversity of rice for specific chromosomes are very useful in rice breeding programs especially for gene mapping and application of marker-assisted selection (MAS).

The dendrogram created using UPGMA method produced two main clusters at 26.21% similarity. The rice cultivars sampled in this study had higher genetic diversity compared to Lapitan *et al.* (2007), who assessed 24 cultivars using 56 SSR markers and clustered three main groups at 40% level of similarity. This may be contributed by geographic conditions in addition to local cultural customs. According to Lapitan *et al.* (2007), distinct geographic distribution may contribute to the genetic differentiation of the cultivars studied. Meanwhile, traditional farmers in Sarawak know the value of maintaining a high genetic diversity in fields (Teo, 2007). They practiced planting a few varieties of rice in the field to avoid major crop damage due to pests and disease. This practice helped to conserve the genetic diversity that is lost in many rice growing Asian countries (Teo, 2007). As mentioned by Deb (1999), tendency to replace traditional varieties with high yielding varieties assists in the erosion of genetic diversity. With vast genetic diversity, the knowledge and correct determination of genetic varieties is important in rice breeding programs, allowing selection of the desired rice cultivars for crossing.

5.3 Matrix Comparison

According to Silva *et al.* (2007), a cophenetic correlation coefficient of more than 0.8 indicates a dendrogram that well reflects the original data but clustering results need to be carefully checked if values obtained were lower than 0.8. Meanwhile, Rohlf (2005) stated that the cophenetic correlation could be used as a measure of degree of fit for a cluster analysis where the degree of fit can be interpreted subjectively as follows: $r > 0.9$ very good fit; $0.8 < r < 0.9$ good fit, $0.7 < r < 0.8$ poor fit, and $r < 0.7$ very poor fit. Therefore, the cophenetic correlation ($r = 0.78$) obtained between the similarity matrix and corresponding dendrogram in this study was fairly poor fit compared to 0.92 reported by Giarroco *et al.* (2007) but higher than the one reported by Ravi *et al.* (2003). This was probably due to a large number of pairwise genetic similarity coefficients with intermediate values, which allows for a number of similar variants for dendrogram branching (Maras *et al.*, 2008).

CHAPTER 6
CONCLUSION

The present study provides an overview of the genetic diversity among the selected Sarawak rice cultivars. The rice SSR markers used in genetic diversity analysis enabled grouping of cultivars according to their genotypes. Since the SSR markers are neutral and co-dominant, they are powerful tools to assess and identify the genetic variability of the closely related rice cultivars in Sarawak. At least one marker in this study uniquely distinguished all varieties except for Lemak. The cultivar Lemak from Meradong and Roban areas are believed to have a common origin.

Understanding of the rice genetic diversity is a prerequisite for future studies aimed at improving rice quality and other breeding programs. Screenings of further markers are recommended in order to develop a suite of markers that are polymorphic between closely related varieties for use in variety identification, gene tagging and plant breeding.

REFERENCES

Ahmadikhah, A., G.I. Karlov, G.H. Nematzadeh and B.K. Ghasemi. 2007. Inheritance of the fertility restoration and genotyping of rice lines (*Rf*) loci using molecular markers. *International Journal of Plant Production* **1(1):**13-21

Akagi, H., Y. Yokozeki, I. Inagaki and T. Fujimura. 1997. Highly polymorphic microsatellites of rice consist of AT repeats, and a classification of closely related cultivars with those microsatellite loci. *Theoretical and Applied Genetics* **94:**61-67

Akkaya, M.S., A.A. Bhagwat and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* **132:**1131-1139

Alvarez, A., J.L. Fuentes, V. Puldon, P.J. Gomez, L. Mora, M.C. Duque, G. Gallego and J.M. Tohme. 2007. Genetic diversity analysis of Cuban traditional rice (*Oryza sativa* L.) varieties based on microsatellite markers. *Genetics and Molecular Biology* **30(4):**1109-1117

Anderson, J.A., G. A. Churchill, J.E. Autrique, S.D. Tanksley and M.E. Sorrells. 1993. Optimizing parental selection for genetic linkage maps. *Genome* **36:**181-186

Barry, G.F. 2001. The use of the Mosanto draft rice genome sequence in research. *Plant Physiology* **12:**1164-1165

Blair, M.W., O. Panaud and S.R. McCouch. 1999. Application of ISSR amplification for the fingerprinting of rice cultivars and for the analysis of microsatellite motif frequencies in the rice genome. *Theoretical and Applied Genetics* **98:**780-792

Blair, M.W., V. Hedetale and S.R. McCouch. 2002. Fluorescent-labelled microsatellite panels useful for detecting allelic diversity in cultivated rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* **105:**449-457

Bligh, H.F.J., N.W. Blackhall, K.J. Edwards and A.M. McClung. 1999. Using amplified fragment length polymorphisms and simple sequence length polymorphisms to identify cultivars of brown and white milled rice. *Crop Science* **39:**1715-1721

Botstein, D., R.L. White, M. Skolnick and R.W. Davis. 1980. Construction of a genetic linkage in man using restriction fragment length polymorphisms. *American Society of Human Genetics* **32:**314-331

Brondani, C., T.C.C.O. Borba, P.H.N. Rangel and R.P.V. Brondani. 2006. Determination of genetic variability of traditional varieties using microsatellite markers. *Genetics and Molecular Biology* **29(4):**676-684

Buell, C.R. 2005. Sequence, annotation, and analysis of synteny between rice chromosome 3 and diverged grass species. *Genome Research* **15:** 1284-1291

Chakravarthi, B.K. and R. Naravaneni. 2006. SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa* L.). *Journal of Biotechnology* **5(9):**684-688

Chambers, G.K. and E.S. Macavoy. 2000. Microsatellites: Consensus and controversy. *Comparative Biochemistry and Physiology Part B* **126:**455-476

Chen, X., S. Temnykh, Y. Xu, Y.G. Cho and S.R. McCouch. 1997. Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* **95:**553-567

Chen, X., Y.G. Cho, and S.R. McCouch. 2002. Sequence divergence of rice microsatellites in *Oryza* and other plant species. *Molecular Genetics Genomics* **268:**331-343

Cheng, H.H. and L.B. Crittenden. 1994. Microsatellites markers for genetic mapping in the chicken. *Poultry Science* **73:**539-546

Cho, Y. G., T. Ishii, S. Temnykh, X. Chen, L. Lipovich, S.R. McCouch, W.D. Park, and N. Ayres. 2000. Diversity of microsatellites derived from genomic libraries and GeneBank sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* **100:**713- 722

Coburn, J.R., S. Temnykh and S.R. McCouch. 2002. Design and application of microsatellite marker panels for semi-automated genotyping of rice (*Oryza sativa* L.). *Crop Science* **42:**2092-2099

Deb, D. 1999. In search of the forgotten rice varieties. *Down to Earth* **8:**19990531

De Woody, J.A., R.L. Honeycutt and L.C. Skow. 1995. Microsatellite markers in white-tailed deer. *Journal of Hereditary* **86:**317-319

Edwards, K. and R. Mogg. 2001. Plant genotyping by analysis of single nucleotide polymorphisms. In *Plant Genotyping: The DNA Fingerprinting of Plants,* ed. R.J. Henry, p.1-13. Wallingford, U. K.: CABI Publishing.

Ellergen, H. 2004. Microsatellites: Simple sequences with complex evolution. *Nature Reviews: Genetics* **5:**435-445

Gale, M.D. and K.M Devos. 1998. Comparative genetics in the grasses. *Proceeding National Academy of Science USA* **95(5):**1971-1974

Gao, L.Z., C.H. Zhang and J.Z. Jia. 2005. Cross-species transferability of rice microsatellites in its wild relatives and the potential for conservation genetic studies. *Genetic Resources and Crop Evolution* **52:**931-940

Garcia, A.A.F., L.L. Benchimol, A.M.M. Barbosa, I.O. Geraldi, C.L. Souza Jr. and A.P. de Souza. 2004. Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genetics and Molecular Biology* **27(4):**579-588

Giarrocco, L.E., M.A. Marassi and G.L. Salerno. 2007. Assessment of the genetic diversity in Argentine rice cultivars with SSR markers. *Crop Science* **47:** 853-860

Goff, S.A., D. Ricke, T.H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, and H. Varma. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296(5565):**92-100

Gramene Database: Release 27 April 2008. Available at: http://www.gramene.org Last modified: Tue Aug 5 08:56:51 2008.

Hamada, H. and T. Kakunaga. 1982. Potential Z-DNA forming sequences are highly dispersed in the human genome. *Nature* **298:**296-298

IRGSP. 2005. International Rice Genome Sequencing Project: The map-based sequence of the rice genome. *Nature* **436:**793-800

IRRI. 2009. International Rice Research Institute: Malaysia: Rice self sufficiency target for next year is 86%. National Food Security Policy.

Jain, S., R.K. Jain, S.R. McCouch. 2004. Genetic analysis of Indian aromatic and quality rice (*Oryza sativa* L.) germplasm using panels of fluorescently labelled microsatellite markers. *Theoretical and Applied Genetics* **109:**965-977

Jarne, P. and P.J.L. Lagoda. 1996. Microsatellites, from molecules to populations and back. *TREE* **11(10):**424-429

Jayamani, P., S. Negrao, M. Martins, B. Macas, and M.M. Oliveira. 2007. Genetic relatedness of Portuguese rice accessions from diverse origins as assessed by microsatellite markers. *Crop Science* **47:**879-886

Jewell, E., A. Robinson, D. Savage, T. Erwin, C.G. Love, G.A.C. Lim, X. Li, J. Batley, G.C. Spangenberg and D. Edwards. 2006. SSR primer and SSR taxonomy tree: Biome SSR discovery. *Nucleic Acid Research* **34:**W656-659

Joshi, R.K. and I. Behera. 2007. Identification and differentiation of indigenous non-Basmati aromatic rice genotypes of India using microsatellite markers. Central Rice Research Institute, Cuttack, India.

Khush, G.S. 1997. Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology* **35(1-2):**25-34

Lai, J., J. Ma, Z. Swigonova, W. Ramakrishna, E. Linton, V. Liaca, B. Tanyolac, Y.J. Park, O.Y. Jeong and J.L. Bennetzen. 2004. Gene loss and movement in the maize genome. *Genome Research* **14(10A):**1924-1931

Lapitan, V.C., D.S. Brar, T. Abe and E.D. Redoña. 2007. Assessment of genetic diversity of Philippine rice cultivars carrying good quality traits using SSR markers. *Breeding Science* **5:**263-270

Li, C., Y. Zhang, K. Ying, X. Liang and B. Han. 2004. Sequence variations of simple sequence repeats on chromosome-4 in two subspecies of the Asian cultivated rice. *Theoretical and Applied Genetics* **108:**392-400

Lu, H., M.A. Redus, J.R. Coburn, N. Rutger, S.R. McCouch and T.H. Tai. 2005. Population structure and breeding patterns of 145 U. S. rice cultivars based on SSR marker analysis. *Crop Science* **45:** 66-76

Mackill, D.J. 2003. Mini review: Applications of genomics to rice breeding. *IRRN,* 9-15.

Mantel, N. 1967. The detection of disease clustering and a generalised regression approach. *Cancer Research* **27:**209-220

Maras, M., J. Sustar-vozlic, B. Javornik and V. Meglic. 2008. The efficiency of AFLP and SSR markers in genetic diversity estimation and gene pool classification of common bean (*Phaseolus vulgaris* L.). *Acta agriculturae Slovenica* **91(1)**:87-96

McCouch, S.R., X. Chen, O. Panaud, S. Temnykh, Y. Xu, Y.G. Cho, N. Huang, T. Ishii and M. Blair. 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Molecular Biology* **35**:89-99

McCouch, S.R., L. Teytelman, Y.B. Xu, K.B. Lobos, K. Clare, M. Walton, B.Y. Fu, R. Maghirang, Z.K. Li, Y.Z. Xing, Q.F. Zhang, I. Kono, M. Yano, R. Fjellstrom, G. DeClerck, D. Schneider, S. Cartinhour, D. Ware and L. Stein 2002. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Research* **9:**199-207

Morgante, M. and A.M. Olivieri. 1993. PCR-amplified microsatellites as markers in plant genetics. *Plant Journal* **1:**175-182

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89:**583-590

Ni, J., P.M. Colowit and D.J. Mackill. 2002. Evaluation of genetic diversity in rice subspecies using microsatellite markers. *Crop Science* **42:**601-607

Olufowote, J.O., Y. Xu, X. Chen, W.D. Park, H.M. Beachel, R.H. Dilday, M. Goto and S.R. McCouch. 1997. Comparative evaluation of within-cultivar variation of rice (*Oryza sativa* L.) using microsatellite and RFLP markers. *Genome* **40:**370-378

Panaud, O., X. Chen and S.R. McCouch. 1995. Frequency of microsatellite sequences in rice (*Oryza sativa* L.). *Genome Research* **38:**1170-1176.

Panaud, O., X. Chen and S.R. McCouch. 1996. Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Molecular Genetics and Genomics* **252:**597-607

Paterson, A.H, J.E. Bowers and B.A Chapman. 2004. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proceeding National Academy of Science USA* **26:**9903-9908

Peng, S., J. Huang, J.E. Sheehy, R.C. Laza, R.M. Visperas, X. Zhong, G.S. Centeno, G.S. Khush, and K.G. Cassman. 2004. Rice yields decline with higher night temperature from global warming. *Proceeding National Academy of Science USA* **101(27):**9971-9975

Pessoa-Filho, M., A. Belo, A.A.N. Alcochete, P.H.N. Rangel and M.E. Ferreira. 2007. A set of multiplex panels of microsatellite markers for rapid molecular characterization of rice accessions. *BMC Plant Biology* **7(23):**1-10

Ram, S.G., V. Thiruvengadam and K.K. Vinod. 2007. Genetic diversity among cultivars, landraces and wild relatives of rice as revealed by microsatellite markers. *Journal of Applied Genetics* **48(4):**337-345

Ravi, M., S. Geethanjali, F. Sameeyafarheen and M. Maheswaran. 2003. Molecular marker based genetic diversity analysis in rice (*Oryza sativa* L.) using RAPD and SSR markers. *Euphytica* **133:**243-252

Rohlf, F.J. 2005. NTSYS-PC: Numerical taxonomy and multivariate analysis systems, Version 2.20r N, Exeter software, Setauket, New York.

Sasaki, T., T. Matsumoto, B.A. Antonio and Y. Nagamura. 2005. From mapping to sequencing, post-sequencing and beyond. *Plant Cell Physiology* **46(1):**3-13

Saw, O.M., K. Doi, K. Aye, K. Irie and A. Yoshimura. 2006. Genetic diversity of Myanmar rice cultivars detected by DNA markers. *Faculty of Agriculture, Kyushu University* **51(2):**181-187

Shen, Y.J., H. Jiang, J.P. Jin, Z.B. Zhang, B. Xi, Y.Y. He, G. Wang, C. Wang, L. Qian, X. Li, Q.B. Yu, H.J. Liu, D.H. Chen, J.H. Gao, H. Huang, T.L. Hi, and Z.N. Yang. 2004. Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. *Plant Physiology* **135:**1198-1205

Silva, D.D.M.E., A.D da Cruz, R.P. Bastos, R.L Reis, M.P. de Campos Telles, and J.A.F. Diniz-Filho. 2007. Population structure of *Euphemphix natteri* (Amphibia, Anura, Leiuperidae) from Central Brazil. *Genetics and Molecular Biology* **30(4):**1161-1168

Skinner, D.M., W.G. Beattie and F.R. Blattner. 1974. The repeat sequence of a hermit crab satellite deoxyribonucleic acid is $(\text{-T-A-G-G-})_n$ x $(\text{-A-T-C-C-})_n$. *Biochemistry* **13:**3930-3937

Sorrells, M.E., M. La Rota, C.E. Bermudez-Kandianis, R.A. Greeene, R. Kantety, J.D. Munkvold, Miftahudin, A. Mahmoud, X. Ma and P.J. Gustafson. 2003. Comparative DNA sequence analysis of wheat and rice genomes. *Genome Research* **13(8):**1818-1827

Sundaram, R.M., B. Naveenkumar, S.K. Biradar, S.M. Balachandran, B. Mishra, M. IlyasAhmed, B.C. Viraktamath, M.S. Ramesha and N.P. Sarma. 2008. Identification of informative SSR markers capable of distinguishing hybrid rice parental lines and their utilization in seed purity assessment. *Euphytica* **163:**215-224

Tan, C.S., G.K. Teo and P.H. Tan. 2006. Microsatellite markers for differentiation of Bario rice varieties. *National Conference on Agrobiodiversity conservation and Sustainable Utilization*, 6-8 November 2006. Kuching, Sarawak.

Temnykh, S., W.D. Park, N. Ayers, S. Cartinhour, N. Hauck, L. Lipovich, Y. G. Cho, T. Ishii and S.R. McCouch. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* **100:**697-712

Temnykh, S., G. DeClerck, Lukashova, L. Lipovich, S. Cartinhour and S.R. McCouch. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Research* **11:**1441-1452

Teo, G.K. 2007. Traditional rice and their eating qualities. *Eastern Times*, Sarawak. 20 August 2007.

Thomson, M.J., E.M. Septiningsih, F. Suwardjo, T.J. Santoso, T.S. Silitonga and S.R. McCouch. 2007. Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Theoretical and Applied Genetics* **114:**559-568

Tingey, S.V. and J.P. del Tufo. 1993. Genetic analysis with random amplified polymorphic DNA markers. *Plant Physiology* **101**:349-352

Tu. M., B. Lu, Y. Zhu and Y. Wang. 2007. Abundant within-varietal genetic diversity in rice germplasm from Yunnan province of China revealed by SSR fingerprints. *Biochemistry Genetics* **45:**789-801

Valentine, S. 2008. DNA profiling of traditional rice cultivars in Sarawak using microsatellite markers. Project Report for Degree of Bachelor of Science Bioindustry. Universiti Putra Malaysia Bintulu Sarawak Campus.

Vivantis GF-1 Plant DNA extraction and Vivantis DNA amplification user's guide. Nucleic Acid Extraction Kit Handbook.

Welsh, J. and M. McClelland. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acid Research* **18:**7213-7218

Weising, K., H. Nybom, K. Wolff and G. Kahl. 2005. DNA Fingerprinting of Plants: Principles, Methods and Applications. Boca Raton: CRC Press.

Williams, J.K., A.R. Kubelik, K.J. Lival, J.A Rafalski, and V.T. Scott. 1990. DNA polymorphism amplified by arbitrary primers is useful as genetic markers. *Nucleic Acid Research* **18:**6531-6535

Xu, Y., S.R. McCouch, and Q. Zhang. 2005. How can we use genomics to improve cereals with rice as a reference genome? *Plant Molecular Biology* **59:**7-26

PUBLICATION OF THE PROJECT UNDERTAKING

This is to certify that I have no objection to publish the project entitled "Genetic Diversity of Selected Sarawak Rice Cultivars Using Microsatellite Markers" by the supervisor in a joint authorship. However, it has to be evaluated by the Faculty of Agriculture and Food Sciences, Universiti Putra Malaysia Bintulu Sarawak Campus and published in the form approved by the Faculty.

_____

Phene Neoh Pei Nee

Date : 30/04/09