



**UNIVERSITI PUTRA MALAYSIA**

**DEVELOPMENT OF MICROSATELLITE MARKERS AND  
GENETIC DIVERSITY ASSESSMENT OF KEMPAS (*KOOMPASSIA  
MALACCENSIS*) IN PENINSULAR MALAYSIA**

**LEE CHAI TING  
IB 2009 19**



**DEVELOPMENT OF MICROSATELLITE MARKERS AND  
GENETIC DIVERSITY ASSESSMENT OF KEMPAS (*KOOMPASSIA  
MALACCENSIS*) IN PENINSULAR MALAYSIA**

**By**

**LEE CHAI TING**

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia,  
in Fulfilment of the Requirements for the Degree of Doctor of Philosophy**

**December 2009**



**Specially dedicated to my beloved husband and family members,  
in loving memory of my late grandmother and uncle**



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

**DEVELOPMENT OF MICROSATELLITE MARKERS AND  
GENETIC DIVERSITY ASSESSMENT OF KEMPAS (*KOOMPASSIA  
MALACCENSIS*) IN PENINSULAR MALAYSIA**

By

**LEE CHAI TING**

**December 2009**

**Chairman: Faridah Qamaruz Zaman, PhD**

**Institute: Bioscience**

A total of 24 novel microsatellite markers have been successfully isolated and characterised in an important tropical timber species of the family Leguminosae, *Koompassia malaccensis*, locally known as kempas. The microsatellite primers were designed from a genomic library enriched for dinucleotide (CT) repeats and subsequently screened on 24 samples from a natural population. In general, these microsatellite markers are highly polymorphic (mean number of alleles per locus,  $A_a = 6.84$ ; average gene diversity,  $H_e = 0.692$ ), with two loci found to deviate significantly from Hardy-Weinberg equilibrium ( $p < 0.05$ ). The utility of these microsatellite markers were tested across 13 leguminous timber tree species and the highest transferability was found with *K. excelsa*, the only species of the same genus tested, followed by *Dialium platysepalum* of the same subtribe, Dialiinae. The amplification success appeared to be inversely associated with the phylogenetic distance, in particular up to the subtribal levels. Four of the microsatellite loci were used to study the mating system of *K. malaccensis*, based on a fruiting season at the

Semangkok Forest Reserve in year 2005. Single- and multilocus population outcrossing estimates ( $t_s$  and  $t_m$ ) were determined using the program MLTR ver 3.0. The results showed that *K. malaccensis* is predominantly outcrossing ( $t_m = 0.890$ ), with low tendency of mating between relatives [ $(t_m - t_s) = 0.027$ ]. In addition, the level of genetic diversity of *K. malaccensis* in 34 natural populations throughout Peninsular Malaysia was assessed and its distribution described. Omitting four loci due to suspected presence of null alleles and linkage disequilibrium, 20 microsatellite loci were analysed for 974 individuals. Overall, all the populations showed high levels of genetic diversity, with gene diversity ( $H_e$ ) ranging from 0.577 (Kuala Langat Selatan) to 0.787 (LenggorB) and mean  $A_a$  being 9.0. The levels of genetic diversity for the two peat swamp (PS) populations (Kuala Langat Selatan and Pekan) were significantly lower than for the non-PS populations. The estimated coefficients of population differentiation ( $F_{ST}$  and  $R_{ST}$ ) revealed that the majority of the genetic diversity resides within populations and less among populations ( $F_{ST}$ : 0.077;  $R_{ST}$ : 0.102). Results from the analysis of molecular variance (AMOVA), cluster analysis, principal component analysis (PCA) and STRUCTURE analysis consistently demonstrated that *K. malaccensis* originating from the two contrasting habitats (PS vs non-PS) were genetically distinct, supporting the ecotype hypothesis. Excluding the PS populations, the among-population component of genetic diversity was even smaller ( $F_{ST}$ : 0.028;  $R_{ST}$ : 0.023), but statistically significant. Pairwise  $F_{ST}$  values among the non-PS populations were positively correlated to geographical distance (Mantel test;  $r^2 = 0.0936$ ,  $p < 0.01$ ), indicating weak but significant isolation-by-distance. Pangkor Selatan and LenggorB were found to be relatively more divergent among the non-PS populations investigated, presumably due to genetic drift and the inclusion of freshwater swamp habitat, respectively. Significant but weaker

population genetic structure was detected among the rest of the non-PS populations surveyed, which corresponded to the topography of Peninsular Malaysia, reflecting the role of mountain ranges as geographical barriers to gene flow. The implications of the findings from this study for the genetic conservation of *K. malaccensis* are discussed and conservation strategies (both *in situ* and *ex situ*) proposed to ensure sustainable utilisation of this important timber species in Malaysia.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan ijazah Doktor Falsafah

**PEMBANGUNAN PENANDA MIKROSATELIT DNA DAN PENILAIAN  
KEPELBAGAIAN GENETIK BAGI KEMPAS (*KOOMPASSIA  
MALACCENSIS*) DI SEMENANJUNG MALAYSIA**

Oleh

**LEE CHAI TING**

**Disember 2009**

**Pengerusi: Faridah Qamaruz Zaman, PhD**

**Institut: Biosains**

Sejumlah 24 penanda mikrosatelit baru telah berjaya dipencil serta dicirikan untuk satu spesies balak tropika yang penting daripada famili Leguminosae, iaitu *Koompassia malaccensis*, yang dikenali dengan nama tempatan kempas. Pencetus mikrosatelit tersebut telah direka daripada perpustakaan genomik yang diperkayakan dengan ulangan dinukleotida (CT) dan kemudiannya disaring menggunakan 24 sampel dari satu populasi semulajadi. Secara amnya, penanda-penanda mikrosatelit yang diperolehi adalah berpolimorfik tinggi (min bilangan alel setiap lokus,  $A_a = 6.84$ ; min kepelbagaian gen,  $H_e = 0.692$ ), dengan dua lokus didapati menyimpang daripada keseimbangan Hardy-Weinberg ( $p < 0.05$ ). Kegunaan penanda-penanda molekul tersebut diuji ke atas 13 spesies balak legum dan kadar pemindahan tertinggi didapati pada *K. excelsa*, satu-satunya spesies daripada genus yang sama yang telah diuji, diikuti dengan *Dialium platysepalum* daripada subtrib yang sama, Dialiinae. Kejayaan amplifikasi didapati mempunyai perhubungan songsang dengan jarak filogenetik, terutamanya sehingga ke peringkat subtrib. Empat daripada lokus

mikrosatelit tersebut telah digunakan untuk kajian sistem kacukan *K. malaccensis*, berdasarkan suatu musim buah di Hutan Simpan Semangkok pada tahun 2005. Anggaran kadar kacukan luar populasi berdasarkan lokus tunggal serta gabungan lokus ( $t_s$  and  $t_m$ ) telah ditentukan dengan menggunakan program MLTR ver 3.0. Keputusan menunjukkan bahawa *K. malaccensis* mengamalkan kacukan luar dengan kadar yang tinggi ( $t_m = 0.890$ ) dan mempunyai kadar kacukan sesama saudara yang rendah [ $(t_m - t_s) = 0.027$ ]. Di samping itu, tahap kepelbagaian genetik *K. malaccensis* bagi 34 populasi semulajadi di Semenanjung Malaysia telah dinilai dan taburannya diterangkan. Setelah mengasingkan empat lokus yang disyaki mempunyai alel-nul serta ketidakseimbangan rangkaian, 20 lokus mikrosatelit telah dianalisis untuk 974 individu. Secara keseluruhannya, kesemua populasi menunjukkan tahap kepelbagaian genetik yang tinggi, dengan kepelbagaian gen ( $H_e$ ) menjulat daripada 0.577 (Kuala Langat Selatan) ke 0.787 (LenggorB) dan min  $A_a$  sebanyak 9.0. Tahap kepelbagaian genetik untuk kedua-dua populasi paya gambut (Kuala Langat Selatan dan Pekan) adalah lebih rendah secara statistik berbanding dengan populasi bukan-paya-gambut yang lain. Anggaran koefisien pembezaan populasi ( $F_{ST}$  dan  $R_{ST}$ ) menunjukkan bahawa majoriti kepelbagaian genetik dibahagikan di dalam populasi dan kurang di kalangan populasi ( $F_{ST}$ : 0.077;  $R_{ST}$ : 0.102). Keputusan daripada analisis varians molekular (AMOVA), analisis kelompok, analisis komponen prinsipal (PCA) serta analisis STRUCTURE secara konsisten menunjukkan bahawa *K. malaccensis* yang berasal daripada dua habitat yang berlainan (paya gambut dan bukan-paya-gambut) adalah berbeza secara genetik, iaitu menyokong hipotesis ekotip. Dengan pengecualian populasi paya gambut, komponen kepelbagaian genetik di kalangan populasi adalah jauh lebih kecil ( $F_{ST}$ : 0.028;  $R_{ST}$ : 0.023), tetapi signifikan secara statistik. Nilai  $F_{ST}$  secara berpasangan di kalangan populasi bukan-



paya-gambut berhubungan secara langsung dengan jarak geografi masing-masing (Mantel test;  $r^2 = 0.0936$ ,  $p < 0.01$ ), menggambarkan wujudnya pengasingan-oleh-jarak yang lemah tetapi signifikan. Pangkor Selatan dan LenggongB didapati lebih mencapah secara relatif di antara populasi bukan-paya-gambut yang dikaji, kemungkinan disebabkan akibat hanyutan genetik serta perangkuman habitat paya-air-tawar. Struktur genetik populasi yang lebih lemah tetapi signifikan telah dikesan di kalangan populasi bukan-paya-gambut yang lain, dan didapati berkait rapat dengan topografi Semenanjung Malaysia, menggambarkan peranan banjaran pergunungan sebagai halangan geografi terhadap aliran gen. Implikasi penemuan kajian ini terhadap pemuliharaan genetik telah dibincang dan strategi pemuliharaan (secara *in situ* dan *ex situ*) disyorkan demi memastikan penggunaan mampan spesies balak yang penting ini di Malaysia.

## ACKNOWLEDGEMENTS

*Praise be to God, Maker of all things.* I give thanks to the Almighty God for His faithfulness and amazing grace in seeing me through. It has been a challenging but enriching and memorable journey.

I would like to express my heartfelt gratitude and appreciation to my supportive and understanding supervisors, Associate Professor Dr. Faridah Qamaruz Zaman, Dr. Lee Soon Leong and Professor Dr. Siti Shapor Siraj, for their valuable guidance, insightful ideas, constructive suggestions, great patience and constant encouragement throughout the course of this study.

A note of thanks to Dr. Saw Leng Guan, Dr. William Goodwin, Dr. Iyengar Arati, Dr. Naoki Tani, Dr. Yoshihiko Tsumura, Dr. Saneyoshi Ueno and Dr. Remy Petit for sharing their expertise and sparking ideas through constructive discussions. As ultraspeed centrifuge was not available in FRIM, the CsCl purification of the genomic DNA for microsatellite development was performed with the help from Dr. Tani at the Forestry and Forest Products Research Institute (FFPRI), Tsukuba, Japan. Many thanks go to Mr. Mohd Nor Mat Isa from the Malaysia Genome Institute for his assistance in sequence analysis and to Mr. Chua Bok Hui for introducing the MISA software. I would also like to extend my sincere thanks to my lab-mates, Dr. Kevin Ng, for lending a big hand during field samplings and for sharing his technical skills generously, and Mrs. Tnah Lee Hong for helping me to familiarise with some of the software used.

I am indebted to FRIM's Director General, Dato' Dr. Abdul Latif Mohmod, and the ex-Director General, Dato' Dr. Abdul Razak Mohd Ali, for giving me the opportunity for further study. I am also thankful to the Forest Biotechnology Division Director, Dr. Norwati Muhammad, and the previous Senior Directors, Dr. Marzalina Mansor and Dr. Daniel Baskaran Krishnapillay, for they have been very supportive and understanding. My sincere appreciation is extended to my senior colleagues, Mr. Ang Khoon Cheng, Mr. Lim Seng Choon, Dr. Lillian Chua, Dr. Abdul Rahman Kassim and Dr. Ismariah Ahmad for being helpful and approachable. This research is funded by the Ministry of Science, Technology and Innovation Malaysia (IRPA Project No. 09-04-01-0098-EA001). The scholarship from the Public Services Department (JPA), the financial support from FRIM, as well as a token of graduate research grant from the ASEAN-Korea Environmental Cooperation Project (AKECOP) are gratefully acknowledged.

All field samplings were carried out with the help of the dedicated research assistants from the Genetic Laboratory, FRIM (Mr. Ramly Punyoh, Mr. Ghazali Jaafar, Mr. Yahya Marhani and Mrs. Sharifah Talib), besides Dr. Lee Soon Leong and Dr. Kevin Ng. Without their help, this study would not have been possible. It has been a challenging but rewarding experience, collecting samples from various forest reserves. I would also like to thank the Forest Department Peninsular Malaysia and the Forest Departments of the respective states for granting permission for sample collection. The kind logistic assistance of the District Forest Officers, Assistant District Forest Officers, foresters, and rangers is much appreciated. Special acknowledgement is also due to my laboratory's research assistants for their help in DNA extraction (Mrs. Mariam Din, Miss Suryani Che Semat, Mrs. Nurul Hudaini

Mamat, Mrs. Nor Salwah Abd Wahid and those mentioned above). To all the other colleagues from the Genetic Laboratory (Dr. Norwati Adnan, Dr. Norlia Basherudin, Dr. Ng Chin Hong, Dr Siti Salwana Hashim and Dr. Mohd Rosli Haron), thanks for maintaining the conducive working environment.

I am deeply grateful to my mom and dad for encouraging me to further my studies. I am thankful for all my family members and friends (Wai Mun, Phoon, How, Sik, Lee Hong, Wan Fong, Brian, Tzer Ying, etc) who cheer me up with their moral support and kept me in prayers. Finally, I want to thank my dear husband, Yik Sheng, for his love, support, understanding and encouragement. I will always cherish the joys and tears we shared in pursuing our Ph.D. at the same. His grace is sufficient for us. Hallelujah!

This thesis submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

**Faridah Qamaruz Zaman, PhD**

Associate Professor  
Institute of Bioscience  
Universiti Putra Malaysia  
(Chairman)

**Siti Shapor Siraj, PhD**

Professor  
Faculty of Agriculture  
Universiti Putra Malaysia  
(Member)

**Lee Soon Leong, PhD**

Forest Biotechnology Division  
Forest Research Institute Malaysia  
(Member)

---

**HASANAH MOHD GHAZALI, PhD.**

Professor and Dean  
School of Graduate Studies  
Universiti Putra Malaysia

Date: 17 Mac 2010

## TABLE OF CONTENTS

	<b>Page</b>
<b>DEDICATION</b>	ii
<b>ABSTRACT</b>	iii
<b>ABSTRAK</b>	vi
<b>ACKNOWLEDGEMENTS</b>	ix
<b>APPROVAL</b>	xii
<b>DECLARATION</b>	xiv
<b>LIST OF TABLES</b>	xviii
<b>LIST OF FIGURES</b>	xx
<b>LIST OF ABBREVIATIONS</b>	xxiv
<b>CHAPTER</b>	
<b>1 INTRODUCTION</b>	<b>1</b>
<b>2 LITERATURE REVIEW</b>	<b>6</b>
2.1 <i>Koompassia malaccensis</i> Maingay ex Benth.	6
2.1.1 Distribution and conservation status	6
2.1.2 Biology	7
2.1.3 Usage and economic importance	10
2.2 Genetic Diversity	11
2.2.1 Why genetic diversity matters?	11
2.2.2 Evolutionary processes that affect genetic diversity	15
2.2.3 Methods for genetic diversity assessment	18
2.2.4 Molecular markers technology	20
2.3 Applications of Molecular Markers in Conservation Genetics	22
2.3.1 Genetic diversity assessment	23
2.3.2 Resolving taxonomic uncertainties and defining management units	24
2.3.3 Detection of inter-specific hybridisation and gene introgression	25
2.3.4 Improvement of genebank management	26
2.3.5 Forensic applications in combating illegal trade of endangered species, poaching and illegal logging	27
2.3.6 Monitoring the impact of management on forest genetic resources	29
2.3.7 Other applications	30
2.4 Microsatellites	31
2.4.1 Mutational mechanisms of microsatellites	34
2.4.2 Mutation models of microsatellites	37
2.4.3 Development of microsatellites	39
2.4.4 Cross-species amplification	43
2.4.5 Limitations of microsatellites	45

<b>3</b>	<b>MATERIALS AND METHODS</b>	47
3.1	Development of Microsatellite Markers	47
3.1.1	Sample collection and DNA extraction	47
3.1.2	Isolation of microsatellite markers using enrichment Approach	48
3.1.3	Primer design	51
3.1.4	PCR amplification and fragment analysis	52
3.1.5	Primer testing and characterisation	53
3.1.6	Statistical analysis	54
3.2	Cross-species Amplification	54
3.2.1	Sample collection and DNA extraction	54
3.2.2	PCR amplification and fragment analysis	56
3.3	Mating System Study	56
3.3.1	Sample collection and DNA extraction	56
3.3.2	PCR amplification and fragment analysis	58
3.3.3	Statistical analysis	59
3.4	Population Genetic Study	59
3.4.1	Sample collection and DNA extraction	59
3.4.2	PCR amplification and fragment analysis	60
3.4.3	Statistical analysis	65
<b>4</b>	<b>RESULTS</b>	74
4.1	Development of Microsatellite Markers	74
4.1.1	Isolation of microsatellite markers using enrichment approach	74
4.1.2	Primer testing and characterisation	79
4.2	Cross-Species Amplification	82
4.3	Mating System Study	85
4.4	Population Genetic Study	87
4.4.1	Genetic diversity within and among the populations	87
4.4.2	Population differentiation	97
4.4.3	Relationships among the populations	99
4.4.4	Contributions of each population to diversity and allelic richness	112
<b>5</b>	<b>DISCUSSION</b>	114
5.1	Development of Microsatellite Markers	114
5.2	Cross-Species Amplification	120
5.3	Mating System Study	124
5.4	Population Genetic Study	126
5.4.1	Genetic diversity within the populations	128
5.4.2	Population differentiation	130
5.4.3	<i>Koompassia malaccensis</i> of the peat swamp forests as a distinct ecotype	133
5.4.4	Genetic structure of <i>K. malaccensis</i> populations from the non-peat swamp habitats	142
5.4.5	Implications for conservation	147

<b>6</b>	<b>CONCLUSION</b>	155
	<b>REFERENCES</b>	160
	<b>APPENDICES</b>	197
	<b>BIODATA OF STUDENT</b>	218
	<b>LIST OF PUBLICATIONS</b>	219





## LIST OF TABLES

Table		Page
2.1	Top five sawntimber exported in the period of January to June 2006.	11
3.1	Selected timber species (Leguminosae) tested for cross-species amplification of the microsatellite markers developed in <i>Koompassia malaccensis</i> .	55
3.2	Details of the <i>Koompassia malaccensis</i> populations investigated in this study with the respective sample sizes (n). Populations from the peat swamp forests are indicated by * whereas ** indicates the population having both peat swamp and non-peat swamp habitats in adjacent.	61
3.3	Six sets of primer combinations assigned for multiple loading.	64
4.1	Microsatellites identified from the CT enriched genomic library of <i>Koompassia malaccensis</i> .	77
4.2	Twenty-four microsatellite primer pairs derived from a CT enriched genomic library of <i>Koompassia malaccensis</i> , including locus name, Genbank Accession No., repeat motif, primer sequence, annealing temperature ( $T$ ), number of alleles ( $A$ ), allele size, observed heterozygosity ( $H_o$ ), gene diversity ( $H_e$ ) and the probability of paternity exclusion where one parent is known ( $P_e$ ). Significant departure from Hardy-Weinberg equilibrium ( $p < 0.05$ ) is indicated by (*).	81
4.3	Cross-species amplification of 24 <i>Koompassia malaccensis</i> microsatellite loci in some related timber species of the family Leguminosae. Amplification is considered positive when the primer pairs yielded specific PCR products of the expected size without multiple bands.	84
4.4	Outcrossing rates of <i>Koompassia malaccensis</i> from Semangkok Forest Reserve based on four polymorphic microsatellite loci ( <i>Kma050</i> , <i>Kma067</i> , <i>Kma147</i> and <i>Kma180</i> ).	86
4.5	Genetic diversity measures, fixation indices and number of rare alleles per individual for 34 natural populations of <i>Koompassia malaccensis</i> based on 20 microsatellite loci; standard deviations in parentheses. Population codes correspond to Table 3.2.	93
4.6	Number of private alleles at 20 microsatellite loci in 34 natural populations of <i>Koompassia malaccensis</i> .	94
4.7	Comparison of the genetic diversity between the peat swamp and	96

non-peat swamp ecotypes.

- |     |  |     |
|-----|--|-----|
| 4.8 | Pairwise $D_C$ genetic distance (below diagonal) and population differentiation $F_{ST}$ (above diagonal) among the 34 <i>Koompassia malaccensis</i> populations studied. Pop1 to Pop34 correspond to the populations listed in Table 3.2. | 98  |
| 4.9 | Results of analysis of molecular variance (AMOVA) performed by grouping the natural populations of <i>Koompassia malaccensis</i> according to ecotypes (AMOVA 1) and geographical subregions (AMOVA 2).                                    | 100 |
| 5.1 | Summary of some reported studies pertaining to the detection of genetic divergence of ecotypes in different taxa.  | 135 |

## LIST OF FIGURES

Figure		Page
2.1	<i>Koompassia malaccensis</i> : (A) mature tree; (B) young leaves from a fallen branch; (C) hermaphroditic flower; (D) fruits with papery wing; (E) flooring strip; (F) seedlings on the forest floor.	8
2.2	Model of microsatellite mutation by replication slippage (slipped strand mispairing) (source: Ellegren 2000b)	35
3.1	Schematic diagram of the microsatellite enrichment approach: selective hybridization using biotinylated oligonucleotide sequences bound to streptavidin coated magnetic beads (modified from Figure 13.7 in Peterson 2005).	49
3.2	<i>Koompassia malaccensis</i> seed samples for the study of mating system – (A) before and (B) after removal of wings; (C) seeds without wings were then soaked in water to facilitate the removal of the inner coats.	57
3.3	Map of Peninsular Malaysia showing the locations of the <i>K. malaccensis</i> populations investigated in this study. Population codes correspond to Table 3.2.	62
3.4	Some photos taken during the field trips: (A) a <i>Koompassia malaccensis</i> tree sampled from Air Cepam Forest Reserve, Perak; (B) one of the peat swamp forests at Pekan Forest Reserve, Pahang; (C) Endau Rompin State Park, Johor; (D) Compartment 32 of Ulu Sat Forest Reserve, Kelantan.	63
4.1	Single white colonies (putative colonies with recombinant plasmids) were randomly selected for plasmid extraction (e.g., as indicated with an arrow).	75
4.2	Agarose gel image showing some of the extracted plasmid DNAs of variable sizes. M1 to M5 were the concentration standards of 5, 10, 25, 50, 100 ng/μl, respectively.	76
4.3	Electrophoregrams showing microsatellite sequences from clones (A) Kma026, repeat motif: (GA) <sub>12</sub> and (B) Kma054, repeat motif: (TG) <sub>17</sub> .	78
4.4	Agarose gel image showing some of the PCR products amplified using designed primer pairs (Lanes No. 1–44). M indicates 100 bp ladder. * denotes primers which yielded good amplification products of expected sizes and without multiple bands.	80

4.5	At least one allele from the maternal genotype was inherited by the progeny arrays, indicating single-locus mode of inheritance (locus <i>Kma067</i> ).	83
4.6	Examples of gel images: (A) multiple loading set I; (B) multiple loading set V, (details of the primer combinations are given in Table 3.3).	88
4.7	Examples of some chromatograms showing alleles of variable sizes amplified from five individuals of <i>Koompassia malaccensis</i> from Sungai Lalang Forest Reserve using primer pairs (A) <i>Kma026</i> ; (B) <i>Kma141</i> and (C) <i>Kma180</i> .	89
4.8	Frequency distribution of alleles at locus <i>Kma057</i> among the <i>Koompassia malaccensis</i> populations investigated. The predominant allele for the peat swamp populations was 27 bp smaller compared with the most common allele of the non-peat swamp populations (251bp vs 278bp). Bubbles represent distinct alleles with area sizes proportional to the respective allele frequencies within each population. Population codes correspond to Table 3.2.	90
4.9	Frequency distribution of alleles at loci (A) <i>Kma163</i> and (B) <i>Kma172a</i> among the <i>Koompassia malaccensis</i> populations investigated. In both cases, the two peat swamp populations shared the predominant alleles which are of different sizes compared with the predominant/common alleles of the non-PS populations. Bubbles represent distinct alleles with area sizes proportional to the respective allele frequencies within each population. Population codes correspond to Table 3.2.	91
4.10	Neighbour-joining tree based on Chord distance showing genetic relationships among 34 populations of <i>Koompassia malaccensis</i> (PS = peat swamp; ES = East & South; NW = Northwest, NC = North & Central, SW = Southwest). Bootstrap values above 50% are given at the corresponding nodes.	101
4.11	Based on the cluster analysis, the natural populations of <i>K. malaccensis</i> throughout Peninsular Malaysia investigated in this study can be divided into peat swamp (PS) (KLSelatan: 12 and Pekan: 28; highlighted) and non-PS ecotypes (the rest of the populations, except SKarang: 8, which has both ecotypes). Excluding LenggorB: 33, the non-PS ecotype is further partitioned into four geographical subregions (NW = Northwest; SW = Southwest; NC = North & Central; ES = East & South). Population codes correspond to Table 3.2.	102
4.12A	Principal coordinate analysis (PCA 1) of the 34 <i>Koompassia malaccensis</i> populations investigated. Pink colour denotes the populations which are of peat swamp origin. Population codes correspond to Table 3.2.	104

4.12B	Principal coordinate analysis (PCA 2) of the 31 non-peat swamp <i>Koompassia malaccensis</i> populations investigated. Yellow colour denotes the populations of the Subregion Northwest (NW) as identified from the cluster analysis. Population codes correspond to Table 3.2.	105
4.12C	Principal coordinate analysis (PCA 3) of the 22 non-peat swamp <i>Koompassia malaccensis</i> populations investigated; with the exclusion of the nine populations from the Subregion NW (Northwest). The populations from the Subregions ES (East & South), NC (North & Central) and SW (Southwest) were colour coded as blue, green and orange, respectively. Population codes correspond to Table 3.2.	106
4.13	Bayesian clustering analysis of all the populations investigated to determine the optimal $K$ . (A) Mean $L(K)$ ( $\pm$ S.D.) over five runs for each $K$ . The graph exhibits the optimal number of clusters at $K = 2$ , after which $L(K)$ at larger $K$ s almost plateaus with only slight increases. (B) Rate of change of the likelihood distribution, $L'(K)$ .	108
4.14	Bayesian clustering analysis of the non-peat swamp populations to determine the optimal $K$ . (A) Mean $L(K)$ ( $\pm$ S.D.) over five runs for each $K$ . (B) $\Delta K$ calculated as $\Delta K = m( L''(K) ) / s[L(K)]$ . The modal value is $K = 8$ , inferring eight clusters.	109
4.15	STRUCTURE analysis of <i>Koompassia malaccensis</i> in Peninsular Malaysia based on microsatellite data. The width of each segment corresponds to the sample size of each population. Population codes correspond to Table 3.2. (A) All populations surveyed; the highest likelihood was found for $K = 2$ , populations of peat swamp ecotype (red; Population 12, KLSelatan and Population 28, Pekan) are distinctive from those of non-peat swamp; with Population 8 (SKarang) showing admixture genotypes; (B) Non-peat swamp populations investigated in more detail, i.e., excluding Populations 8, 12 and 28; $K = 2$ . Geographical subregions of the populations are given on top of the figure (NW = Northwest; SW = Southwest; NC = North & Central; ES = East & South).	110
4.16	Contribution of each population of <i>Koompassia malaccensis</i> to (A) total diversity, CT, and (B) allelic richness, CTR. Population codes correspond to Table 3.2.	113
5.1	(A) A <i>Koompassia malaccensis</i> tree from a peat swamp forest (Kuala Langat Selatan Forest Reserve, Selangor), exhibiting its steep, sinuous plank buttress; in comparison to <i>K. malaccensis</i> trees from the non peat swamp forests: (B) Endau Rompin State Park, Pahang; (C) Sungai Udang Forest Reserve, Melaka.	134
5.2	Topography map of Peninsular Malaysia superimposed with arbitrary boundaries of the four subregions of non-peat swamp	144

*Koompassia malaccensis* (Northwest, NS; Southwest, SW; North & Central, NC; East & South, ES) as identified from the cluster analysis (see Figure 4.10).

## LIST OF ABBREVIATIONS

AMOVA	Analysis of molecular variance
bp	Base pair
CAPS	Cleaved amplified polymorphic sequences
AFLP	Amplified fragment length polymorphism
CITES	Convention on International Trade in Endangered Species
CpDNA	Chloroplast DNA
CsCl	Caesium chloride
CTAB	Hexadecyltrimethyl-ammonium bromide
DAF	DNA amplification fingerprint
dbh	Diameter at breast height
DNA	Deoxyribonucleic acid
dNTP	2'-deoxynucleoside 5'-triphosphate
EDTA	Diaminoethanetetra-acetic acid
EMBL	European Molecular Biology Laboratory
EST	Expressed sequence tag
ESUs	Evolutionarily significant units
FAO	Food and Agriculture Organization of the United Nations
FIASCO	Fast isolation by AFLP sequences containing repeats
FR	Forest Reserve
FRIM	Forest Research Institute Malaysia
GSM	Generalized stepwise model
IAM	Infinite alleles model
IBD	Identical-by-descent