DEVELOPMENT OF MOLECULAR MARKERS FOR THE CONSERVATION OF *Phalaenopsis gigantea*, *Paphiopedilum rothschildianum* AND OTHER ENDANGERED ENDEMIC ORCHIDS OF SABAH, MALAYSIA

KENNETH FRANCIS RODRIGUES

THESIS SUBMITTED IN FULFILMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH

BIOTECHNOLOGY RESEARCH INSTITUTE
UNIVERSITY MALAYSIA SABAH
2008
JUDUL: DEVELOPMENT OF MOLECULAR MARKERS FOR THE CONSERVATION OF *Phalaenopsis gigantea*, *Paphiopedilum rothschildianum* AND OTHER ENDANGERED ENDEMIC ORCHIDS OF SABAH, MALAYSIA.

IJAHAZ: DOKTOR FALSAFAH (MOLECULAR GENETICS)

SESJI PENG AJIAN: 2004 - 2008

Saya, KENNETH FRANCIS RODRIGUES mengaku membenarkan tesis sarjana ini disimpan di perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:

1. Tesis adalah hak milik Universiti Malaysia Sabah.
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. TIDAK TERHAD

Disahkan oleh

[K. Rodriguez]

Penulis: KENNETH FRANCIS RODRIGUES

TANDATANGAN PUSTAKAWAN

[Signature]

Penyelia: Prof. Madya. Dr. Vijay Kumar

Penyelia bersama: Dr. Zaleha Abdul Aziz

Penyelia bersama: Prof. Madya. Datin Dr. Mariam Abdul Latip

Alamat:
Institut Penyelidikan Bioteknologi
Universiti Malaysia Sabah
Locked Bag 2073
88999 Kota Kinabalu Sabah

Tarikh: 2008
DECLARATION

I declare that this dissertation is the result of my own independent work and original writing except for quotations, excerpts and references, which have been duly acknowledged.

17 November 2008

Kenneth Francis Rodrigues
PS04-012-004(A)
CERTIFICATION

NAME : KENNETH FRANCIS RODRIGUES
MATRIC NO. : PS04 – 012 – 004(A)
TITLE : DEVELOPMENT AND APPLICATION OF MOLECULAR MARKERS FOR THE CONSERVATION OF Phalaenopsis gigantea, Paphiopedilum rothschildianum, AND OTHER ENDANGERED ENDEMIC ORCHIDS OF SABAH, MALAYSIA
DEGREE : DOCTOR OF PHILOSOPHY (MOLECULAR GENETICS)
VIVA DATE : 17 NOVEMBER 2008

DECLARED BY

1. SUPERVISOR
   Assoc. Prof. Dr. Vijay Kumar

2. CO-SUPERVISOR
   Dr. Zaleha Abdul Aziz

3. CO-SUPERVISOR
   Assoc. Prof. Dr. Datin Mariam Abdul Latip
ACKNOWLEDGEMENTS

All human achievements and endeavors are the Manifestations of One Infinite Source.

This thesis has been made possible through the unqualified support, assistance and encouragement from the following Institutions and individuals to whom all credit is due:

The Ministry of Science Technology & Innovation, Government of Malaysia, The Sabah Parks Authority, Postgraduate Centre, Universiti Malaysia Sabah, Office for Research and Innovation, Universiti Malaysia Sabah.

Professor Datin Dr. Ann Anton, Director, Biotechnology Research Institute.

My supervisors: Associate Professor Dr. Vijay Kumar, Dr. Zaleha Abdul Aziz, Associate Professor Datin Dr. Mariam Abdul Latip.

Associate Professor Dr. C. M. Wong, Mr. Johan Alexander, Mr. Adrian Ng, Mr. M. J. Kinsuat, Mr. Mohammad Awang Sagaf Bin Abu Bakar, Ms. Shuhadah Mustafa, Ms. Sharon Lau, Ms. Chelven Lim Ai Chen, Mr. Gordon J. Thomas, Mr. Tam Heng Keat, Ms. Grace Joy Chin, Ms. Elaine Remi, Mr. Ramlan Madin, Ms. Vidalita Maikin, Mr. Richard Dalis, Ms. Norazmayati Kassim, Mr. Zulkeflii Muhamad, Mr. Rudi Boliku, Associate professor Dr. Miroslav Radjovic, Dr. Nandkumar Kamat, Goa University, Mr. Patrick San Francisco. My wife Paula Shirlia SidRek for her constant support, encouragement and boundless optimism, and to my family.

Kenneth Francis Rodrigues
17 November 2008
ABSTRACT

Development of molecular markers for the conservation of *Phalaenopsis gigantea*, *Paphiopedilum rothschildianum* and other endangered endemic orchids of Sabah, Malaysia

*Phalaenopsis gigantea*, *Paphiopedilum rothschildianum*, *Paphiopedilum dayanum* and *Paphiopedilum lowii* are endangered orchids endemic to Sabah, Malaysia, which are currently accorded protection under the provisions of the CITES. The molecular characterization of these species is the most significant step in the development of a conservation and management strategy for wild orchid populations. This constituted the basis for this investigation, which was directed towards elucidating the population genetic structure of the endangered orchids using microsatellite loci. Eighty-three individuals representing three populations of *P. rothschildianum* derived from Bukit Ampuan, Melangkap and Telupid were sampled. DNA was extracted and authenticated by amplification of the ITS1 and ITS2 rDNA intergenic spacer region followed by sequencing and comparative analysis using the blastn program. Concurrently, nineteen individuals representing the extant population of *P. gigantea* were collected from the Tawau Hills Park. DNA was extracted and authenticated using the chloroplast DNA *trnl-trnF* intergenic region. One individual each from *P. dayanum* and *P. lowii* was sampled from Kinabalu National Park. A genomic library enriched for microsatellite loci from these four orchids was constructed using the 5’ anchored PCR technique, followed by ligation of the PCR amplicons onto a TOPO TA pCR 2.1 plasmid, transformation of constructs into competent TOP 10F’ E. coli, screening for positive transformants, plasmid extraction, purification and DNA sequencing. A total of 95 sequences containing 212 microsatellite loci and cryptic simple repeats were isolated and deposited at the NCBI GenBank. Specific primer pairs were designed to amplify the microsatellite loci, and applied to characterize the population genetic structure. In *P. gigantea*, 30 polymorphic primer pairs defined 78 alleles. The averages of the observed and expected heterozygosity were 0.3544 to 0.4910. The $F_{ST}$ value ranged from 0.1174 to 1.000 with an average of 0.6294, indicating a high level of genetic variability within the mixed population. In the case study of *P. rothschildianum*, 24 of the 30 primer pairs exhibited polymorphism. The averages of the observed and expected heterozygosity were 0.3800 and 0.4533 respectively. The mean $F_{ST}$ value was 0.5098 indicating genetic diversity within the total population, however the mean $F_{IS}$ value was 0.8766, implying that there is a deficiency of level of heterozygosity. A test for cross amplification was conducted to determine the degree of genomic similarity within the genus *Paphiopedilum* indicated that very little homology exists between genomes of the three species examined. This is in accordance with karyotype analysis data, which indicates that the species in island ecosystems of which one representative is Borneo, have evolved into distinct species as a result of speciation events involving reduction in chromosome numbers. The study has concluded that both the endangered species being investigated exhibit a reduced genetic diversity and warrant categorization as ESUs for the purpose of conservation and has recommended a strategy for the conservation and maintenance of current diversity levels by complying with the scientific breeding strategies that are delineated within the contents of this study. The focus of this investigation has been to develop a strategic approach for the conservation of the endangered endemic species of Sabah, Malaysia with the objective of developing a scientific and pragmatic approach for the conservation of a diverse range of species in addition to orchids.
Developing a statutory framework for conservation involves the consolidation of legislative guidelines and scientific data. This comprehensive study will form the basis of future research investigations into the wide range of genetically diverse Malaysian endemic species.
ABSTRAK

Phalaenopsis gigantea, Paphiopedilum rothschildianum, Paphiopedilum dayanum dan Paphiopedilum lowii adalah spesis orkid terancam yang endemik kepada negeri Sabah, Malaysia, yang pada masa kini di bawah pemuliharaan CITES. Pencirian molecular spesies endemik yang menghadapi ancaman kepupusan ini telah membawa kepada langkah pertama dalam perkembangan terhadap pemuliharaan dan strategi pengendalian bagi populasi liar. Struktur populasi diukur dengan menggunakan penanda molecular berdasarkan kepada loci mikrosatelit. Lapan puluh tiga individu mewakili tiga populasi P. rothschildianum yang diperolehi daripada Bukit Ampuan, Melangkap dan Telupid telah disampel. DNA diekstrak dan disahkan dengan mengamplifikasi kasikan ITS1 dan ITS2 rDNA intergenic spacer region yang diikuti oleh penjukฐาน DNA dan analisis perbandingan menggunakan program blastn. Sembilan belas individu yang mewakili populasi sedia ada daripada spesis P. gigantea telah disampel daripada Taman Bukit Tawau, di mana DNA telah diekstrak dan disahkan dengan menggunakan DNA kloroplas trnL–trnF intergenic region. Satu daripada setiap individu P. dayanum dan P. Lowii telah disampel dari Taman Negara Kinabalu. Suatu perpustakaan genomik yang diperkayakan untuk loci mikrosatelit daripada keempat-empat orkid ini telah dibina menggunakan teknik tindakbalas rantaian polimeras (PCR) 5’ anchored, dan diikuti dengan pencantuman amplikon PCR ke dalam plasmid TOPO TA pCR 2.1, ditransformasikan ke sel kompeten TOP 10F'. coli, penyaringan untuk transforman, kloning, pengekstrakkan plasmid, penulenan dan penjukฐาน DNA. Sebanyak 95 sequences yang mengandungi 212 microsatellite loci dan kriptik 'simple repeats' telah diasingkan dan disimpan di dalam NCBI GenBank. Pasangan primer yang spesifik telah direka untuk mengamplifikasi kasikan loci mikrosatelit, dan digunakan untuk pencirian struktur genetik populasi. Dalam P. gigantea, 30 pasang primer polimorfi memberikan 78 allele. Heterozigositi yang telah diperhatikan dan dianggarkan adalah dari 0.3544 hingga 0.4910. Nilai FST yang berada dalam linkungan 0.1174 hingga 1.000 dengan nilai purata sebanyak 0.6294 menunjukkan perubahan tahap genetik yang tinggi dalam populasi campuran. Dalam kes kajian P.rothschildianum, 24 daripada 30 pasangan primer memaparkan ciri polimorfisme. Heterozigositi yang telah diperolehi hasil daripada pemerhatian adalah sebanyak 0.3800 dan 0.4533. Nilai min FST sebanyak 0.5098 menunjukkan terdapatnya kepelbagaian genetik dalam jumlah keseluruhan populasi, Walau bagaimanapun, nilai min FST sebanyak 0.8766 menunjukkan bahawa terdapat sedikit kepelbagaian genetik dalam populasi. Satu ujian bagi amplifikasi bersilang yang telah dijalankan untuk menentukan tahap persamaan genomik dalam genus Paphiopedilum menunjukkan bahawa terdapat sedikit homologi yang wujud di antara ketiga-tiga genom spesis yang telah dikaji. Ini adalah berdasarkan kepada analisis data kariotip yang menunjukkan bahawa spesis dalam ekosistem pulau yang mana diwakili oleh Borneo ini telah berkembang kepada spesis yang unik hasil daripada pembentukan spesis baru yang melibatkan pengurangan dalam bilangan kromosom. Kajian ini telah menunjukkan bahawa spesies terancam yang dikaji telah menunjukkan pengurangan kepelbagaian genetik dan pencirian ESUs yang diperlukan bagi tujuan pemuliharaan. Strategi bagi pemuliharaan dan pengekalan tahap kepelbagaian sedia ada ini telah dicadangkan dengan mematuhi strategi pembiakan saintifik sepertimana yang telah dinyatakan dalam kandungan kajian ini.
CHAPTER 1: INTRODUCTION

CHAPTER 2: LITERATURE REVIEW

2.1 Overview of Global Biodiversity
2.2 Status of Global and Malaysian Forests
2.3 Global Conservation Efforts: IUCN and CITES
2.4 Orchid diversity
  2.4.1 Molecular Markers in Orchid Phylogenetics
  2.4.2 DNA Barcoding of the Family Orchidaceae
  2.4.3 Orchid Diversity of Malaysia with Reference to Sabah
2.5 Measures of Biological Diversity
  2.5.1 The ESU: Its Significance and Application
  2.5.2 The Management Unit
2.6 Molecular Conservation
2.7 Types of Molecular Markers
  2.7.1 Criteria for Molecular Marker Selection
  2.7.2 Microsatellites: an Overview
  2.7.3 Isolation of Microsatellite Loci
2.8 The Hardy-Weinberg Principle
2.9 Fixation Index
2.10 Linkage Disequilibrium
2.11 Genetic Distance Measures

CHAPTER 3: IDENTIFICATION OF Phalaenopsis gigantea, Paphiopedilum rothschildianum, Paphiopedilum dayanum AND Paphiopedilum lowii AND ISOLATION OF MICROSATELLITES

3.1 Introduction
3.2 Materials and Methods
  3.2.1 Sampling and DNA Extraction
  3.2.2 DNA Authentication of Paphiopedilum spp. and Phalaenopsis gigantea
  3.2.3 5' Anchored Polymerase Chain Reaction
  3.2.4 Cloning of PCR products
  3.2.5 Plasmid Miniprep
  3.2.6 DNA Sequencing and Data Analysis
3.2.7 Primer Design 74

3.3 Results
3.3.1 Sampling and DNA Extraction 75
3.3.2 Amplification of Nuclear rDNA ITS1 and ITS 2 Region in the Genus Paphiopedilum. 76
3.3.3 Characterization of the trnL – trn F Intergenic region in P. gigantea 85
3.3.4 Isolation of Microsatellites Using the 5’ Anchored PCR Method 88
3.3.5 Cloning of PCR Products 94
3.3.6 Alkaline Lysis Miniprep 97
3.3.7 Sequencing of Clones Containing Inserts 98
3.3.8 Design of Specific Primers 115

3.4 Discussion
3.4.1 Sampling and DNA Extraction 127
3.4.2 DNA Authentication of Paphiopedilum Species 129
3.4.3 DNA Authentication of Phalaenopsis gigantea 131
3.4.4 Isolation of Microsatellites Using the 5’ Anchored PCR Method 132
3.4.5 Cloning of PCR Products 135
3.4.6 Alkaline Lysis Miniprep 140
3.4.7 Mining of Sequences for Microsatellite loci 141
3.4.8 Primer Design 147
3.4.9 Conclusions 148

CHAPTER 4: MOLECULAR CHARACTERIZATION OF Phalaenopsis gigantea

4.1 Introduction 149
4.2 Materials and Method 154
4.2.1 Sample Collection 156
4.2.2 Primer Optimization and Scoring of PCR amplicons 154
4.2.3 Data Analysis 157
4.2.4 Heterozygosity 157
4.2.5 F-Statistics 158
4.2.6 Gene Flow, Linkage Disequilibrium and Neutrality 159
4.2.7 Genetic Distance and Dendrograms 160
4.2.8 Bootstrapping and Consensus Trees 160
4.2.9 Population Bottleneck 161
4.2.10 Breeding Program 162
4.2.11 Characterization of the Chalcone Synthase Gene 162
4.2.12 Chloroplast DNA ycf2 Locus 163

4.3 Results
4.3.1 Sample Collection and DNA Extraction 164
4.3.2 Primer Optimization 164
4.3.3 PCR Amplification of Polymorphic Microsatellite Loci 165
4.3.4 Heterozygosity and Effective Number of Alleles 169
4.3.5 Hardy Weinberg Equilibrium, F-statistics 169
4.3.6 Test for Neutrality and Linkage Disequilibrium 176
4.3.7 Population Bottleneck 176
4.3.8 Genetic Distance 181
4.3.9 Dendrogram: NJ and UPGMA 181
4.3.10 Consensus Tree 181
LIST OF TABLES

Table 3.1.1  Summary of loci used for phylogenetic classification within the family Orchidaceae indicating organelle, locus, expected sizes (bp), degree of divergence observed and primary reference.  58

Table 3.1.2  An overview of the data currently available in literature revealed sequences containing microsatellite loci isolated from orchids.  60

Table 3.2.1  Record of samples collected for identification and isolation of microsatellite loci, indicating species name, accession number and location of the plant.  62

Table 3.2.2  PCR Primers used for identification and authentication of specimens collected for microsatellite isolation.  66

Table 3.2.3  Complete list of the 5’ anchored degenerate PCR primers used for the isolation of microsatellite loci indicating 5’ anchored PCR primer utilized, primer sequence, annealing temperature range, target species.  69

Table 3.2.4  Optimization of PCR conditions indicating thermal cycler profiles tested for optimum yield and resolution of PCR amplicons.  70

Table 3.3.1  Sequence of the ITS1 and ITS2 regions derived from P. lowii (PL01), P. rothschildianum (TP1), (MK1), (BA1) and P. dayanum (PD01).  77

Table 3.3.2  Sequence alignment of P. rothschildianum, ITS1 and 2 intergenic spacer region for populations derived from (BA1), (MK1) and (TP1) using CLUSTALW alignment.  80

Table 3.3.3  Blastn sequence alignment result for P. rothschildianum ITS 1 and ITS2 Intergenic spacer region.  81

Table 3.3.4  Blastn sequence alignment result for P. lowii (PL01) ITS 1 and ITS2 Intergenic spacer region.  82

Table 3.3.5  Blastn sequence alignment result for P. dayanum (PD01) ITS 1 and ITS2 Intergenic spacer region.  83
Table 3.3.6  Table indicating degree of similarity between samples collected in the wild and GenBank sequence data. Summary of BlastN sequence alignment result for *P. rothschildianum*, *P. lowii* and *P. dayanum* ITS1 and ITS2 sequence indicating species, GenBank accession number, GenBank accession number of closest possible match, score (bits), E value, identities and gaps.

Table 3.3.7 Chloroplast *tml*-tmF intergenic spacer region sequence derived from *Phalaenopsis gigantea* (PGTW 3644).

Table 3.3.8 Blastn sequence alignment result for Chloroplast *tml*-F intergenic spacer region sequence derived from *Phalaenopsis gigantea* (PGTW 3644).

Table 3.3.9 Table indicating degree of similarity between samples collected in the wild and GenBank sequence data.

Table 3.3.10 Optimum PCR conditions indicating primer concentrations for each of the degenerate primers tested.

Table 3.3.11 Optimum PCR conditions indicating thermal cycler profiles tested for optimum yield and resolution of PCR amplicons.

Table 3.3.12 Total number of transformed *E. coli* containing clones from *Phalaenopsis gigantea*.

Table 3.3.13 Total number of transformed *E. coli* containing clones from *Paphiopedilum rothschildianum*.

Table 3.3.14 Total number of transformed *E. coli* containing clones from *Paphiopedilum dayanum*.

Table 3.3.15 Total number of transformed *E. coli* containing clones from *Paphiopedilum lowii*.

Table 3.3.16 Summary of clones containing inserts that were sequenced in *Phalaenopsis gigantea*, *Paphiopedilum rothschildianum*, *Paphiopedilum lowii* and *Paphiopedilum dayanum*.

Table 3.3.17 Di- and Trinucleotide microsatellite loci isolated from *Phalaenopsis gigantea* indicating repeat motif targeted by degenerate primer, clone, 5' and 3' terminal repeat, and internal repeat sequences, clone size and NCBI GenBank Accession numbers.

Table 3.3.18 Microsatellite loci isolated from *Phalaenopsis gigantea* indicating repeat motif category and total number.
Summary of microsatellite loci isolated from *Paphiopedilum dayanum* indicating loci targeted by degenerate primer, clone number, microsatellite loci at 5' and 3' terminal ends, internal repeats, clone size (bp) and NCBI GenBank Accession number.

Table 3.3.20
Microsatellite loci isolated from *Paphiopedilum dayanum* indicating repeat motif category and total number.

Table 3.3.21
Summary of microsatellite loci isolated from *Paphiopedilum lowii*, indicating loci targeted by degenerate primer, clone number, microsatellite loci at 5' and 3' terminal ends, internal repeats, clone size (bp) and NCBI GenBank Accession number.

Table 3.3.22
Microsatellite loci isolated from *Paphiopedilum lowii* indicating repeat motif category and total number.

Table 3.3.23
Di-, Tri- and Tetranucleotide microsatellite loci isolated from *Paphiopedilum rothschildianum* indicating repeat motif targeted by degenerate primer, clone, 5' and 3' terminal repeat, and internal repeat sequences, clone size and NCBI GenBank Accession numbers.

Table 3.3.24
Microsatellite loci isolated from *Paphiopedilum rothschildianum* indicating repeat motif category and total number.

Table 3.3.25
Summary of primers that were designed to amplify loci containing microsatellite repeat motifs and cryptic simple repeats in *P. gigantea*, *P. rothschildianum*, *P. lowii* and *P. dayanum*.

Table 3.3.26
Summary of primers designed to characterize microsatellite loci in *P. gigantea*, indicating degenerate primer used, clone, locus, primer sequence, repeat motif characterized and expected size.

Table 3.3.27
Summary of primers designed to characterize microsatellite loci in *P. dayanum*, indicating degenerate primer used, clone, locus, primer sequence, repeat motif characterized and expected size.

Table 3.3.28
Summary of primers designed to characterize microsatellite loci in *P. lowii*, indicating degenerate primer used, clone, locus, primer sequence, repeat motif characterized and expected size.
#### Table 3.3.29
Summary of primers designed to characterize microsatellite loci in *P. rothschildianum*, indicating degenerate primer used, clone, locus, primer sequence, repeat motif characterized and expected size.

#### Table 4.1.1
Binomial classification of *Phalaenopsis gigantea*.

#### Table 4.2.1
*Phalaenopsis gigantea* samples obtained from Tawau Hills Park, indicating accession numbers.

#### Table 4.2.2
NCBI GenBank Accession derived for Chalcone Synthase gene derived from species belonging to the genus *Phalaenopsis* and *Oncidium*.

#### Table 4.2.3
Primers designed for amplification of *Phalaenopsis gigantea* Chalcone synthase gene and intron region.

#### Table 4.2.4
Primers designed for amplification of *Phalaenopsis gigantea* chloroplast DNA *ycf2* region.

#### Table 4.3.1
Polymorphic microsatellite and cryptic simple repeat sequence primer pairs, locus name, repeat motif, primer sequences (5'-3'), expected size (bp), specific annealing temperature (°C), number of alleles (bp), observed heterozygosity (H₀), expected heterozygosity (Hₑ), and GenBank Accession no. and sample size.

#### Table 4.3.2
Chi-squared and likelihood ratio test for Hardy-Weinberg equilibrium for all populations.

#### Table 4.3.3
Table indicating locus name, sample size, observed number of alleles (nₐ) and effective number of alleles (nₑ) *(Kimura and Crow, 1964)* across each of the thirty polymorphic loci tested.

#### Table 4.3.4
Table indicating locus name, sample size, F<sub>ST</sub> and F<sub>IT</sub> at each of the thirty loci tested.

#### Table 4.3.5
Overall Ewens–Watterson test for neutrality.

#### Table 4.3.6
Detection of population bottleneck within the population, assuming the IAM model, indicating locus name, sample size, number of alleles, Heterozygosity, Equilibrium heterozygosity, Standard deviation (SD), DH / SD and Probability.

#### Table 4.3.7
Nei's original measure of genetic distance (Nei, 1978) between individuals of *P. gigantea* sampled at Tawau Hills Park.
Combinations of breeding pairs exhibiting a genetic distance (Nel, 1978) of 0.5 and above, indicating possible combinations of breeding pairs. Shaded blocks represent combinations, which have a low genetic distance (GD < 0.5) which may result in F1 progeny with low genetic diversity levels. (ACN: Accession number).

Schematic representation of Neis’ Genetic Distance as functions of frequency distribution, 75 combinations of genotypes exhibit a genetic distance in excess of 0.5 and are an indication of the possible breeding pairs for improvement of gene diversity levels in the F1 generation.

CLUSTAL W alignment of Chalcone synthase gene derived from *Phalaenopsis* Hybrid (AY954515) and *Phalaenopsis gigantea* PGTW3644 (EF139435).

Translated blastx results for *P. gigantea*, CHS gene indicating GenBank Accession number, source, score in bits and e value.

Translated blastx alignment of Chalcone Synthase gene derived from *Phalaenopsis* Hybrid (PHCHS) and *Phalaenopsis gigantea* (PGCHS).

Blastn alignment results of Chloroplast DNA Ycf2 gene derived from *Phalaenopsis gigantea* (PGYCF) and *Phalaenopsis aphrodite* (PAYCF).

Binomial classification of *Paphiopedilum rothschildianum*.

*Paphiopedilum rothschildianum* samples obtained from the Sabah Parks Authority *exsitu* conservation facility at Kinabalu National Park indicating population, assigned sample number and total number of samples collected.

Polymorphic microsatellite and cryptic simple repeat sequence primer pairs, locus name, repeat motif, primer sequences (5'-3'), size range (bp), specific annealing temperature (°C), number of alleles, observed heterozygosity (H₀), expected heterozygosity (Hₑ), and GenBank Accession no. Sample size, N = 83.

Table indicating locus name, sample size, observed number of alleles (n₀) and effective number of alleles (nₑ) *(Kimura and Crow, 1964) across each of the 24 polymorphic loci tested.

Chi-squared and likelihood ratio tests for Hardy-Weinberg equilibrium for all populations indicating locus, number of samples (n), χ² and G² values and the respective probability (P) values.
<p>| Table 5.3.4 | Table indicating locus name, sample size $F_{ST}$ and $F_{IS}$ at each of the 24 loci tested. | 232 |
| Table 5.3.5 | Overall Ewens–Watterson test for neutrality. | 233 |
| Table 5.3.6 | Genetic distance between populations. | 234 |
| Table 6.2.1 | Primers for the amplification of chloroplast DNA loci indicating the locus and the primer sequence. | 255 |
| Table 6.2.2 | Restriction digestion mixture for RFLP of PCR products indicating component, concentration and final volume. | 256 |
| Table 6.3.1 | PCR amplicons of the three specific chloroplast DNA loci, PSIA1, PSIA2 and PSIID2 indicating amplicon size in base pairs for each of the three species of <em>Paphiopedilums</em>. | 258 |
| Table 6.3.2 | Restriction enzyme digestion of PCR amplicons of the three specific chloroplast DNA loci, PSIA1, PSIA2 and PSIID2, using the restriction enzyme <em>Taq</em> I (Fermentas) indicating restriction fragment size in base pairs for <em>P. dayanum</em>, <em>P. lowii</em> and <em>P. rothschildianum</em>. | 259 |
| Table 6.3.3 | Cross amplification of microsatellite loci from <em>Paphiopedilum dayanum</em> indicating locus, primer sequences, size(s) of PCR amplicons in base pairs in target species and GenBank accession number. | 261 |
| Table 6.3.4 | Cross amplification of microsatellite loci from <em>Paphiopedilum lowii</em> indicating locus, primer sequences, size(s) of PCR amplicons in base pairs in target species and GenBank accession number. | 262 |
| Table 6.3.5 | Cross amplification of microsatellite loci from <em>Paphiopedilum rothschildianum</em> indicating locus, primer sequences, size(s) of PCR amplicons in base pairs in target species and GenBank accession number. | 263 |</p>
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1</td>
<td>Systematic representation of the Chloroplast DNA genome of <em>Phalaenopsis aphrodite</em> subspecies <em>formosana</em> showing the Large Single Copy (LSC), Small Single Copy (SSC) and inverted repeat regions (IRA /IRB).</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Schematic representation of rDNA repeat region showing the 5S, 18S, 5.8S and 26S ribosomal subunit coding regions.</td>
</tr>
<tr>
<td>Figure 3.2.1</td>
<td>Map indicating locations of the original sampling sites which are located at Melangkap, Bukit Ampuan and Telupid.</td>
</tr>
<tr>
<td>Figure 3.3.1</td>
<td>Electrophoresis of DNA extracts on a 1.2% TBE Agarose gel.</td>
</tr>
<tr>
<td>Figure 3.3.2</td>
<td>Agarose gel electrophoresis of PCR amplification products generated using the primers 17 SE and 26 SE.</td>
</tr>
<tr>
<td>Figure 3.3.3</td>
<td>Agarose gel electrophoresis of PCR amplicons generated by amplification of DNA with the primers <em>trnLF</em>and <em>trnFR</em>.</td>
</tr>
<tr>
<td>Figure 3.3.4</td>
<td>PCR amplification of DNA with 5’ anchored PCR Primers under conditions of low stringency.</td>
</tr>
<tr>
<td>Figure 3.3.5</td>
<td>PCR amplification of DNA with 5’ anchored PCR Primers under conditions of low annealing temperature stringency.</td>
</tr>
<tr>
<td>Figure 3.3.6</td>
<td>PCR amplification of DNA with Primer PCT7 under conditions of high stringency.</td>
</tr>
<tr>
<td>Figure 3.3.7</td>
<td>PCR amplification of <em>P. rothschildianum</em> DNA with degenerate primers under conditions of high stringency.</td>
</tr>
<tr>
<td>Figure 3.3.8</td>
<td>PCR amplification of <em>P. rothschildianum</em> (BA1) DNA with degenerate primer (PCT04) using Taq Polymerase from two manufacturers.</td>
</tr>
<tr>
<td>Figure 3.3.9</td>
<td>Screening for positive transformants using on LB agar incorporated with 50 µg / ml Ampicillin and 40 µg/ ml X-Gal.</td>
</tr>
<tr>
<td>Figure 3.3.10</td>
<td>Purified plasmids resolved on a 1.2% TBE Agarose Gels.</td>
</tr>
<tr>
<td>Figure 3.3.11</td>
<td>Restriction Enzyme digests of plasmids resolved on a 1.2% TBE.</td>
</tr>
<tr>
<td>Figure 4.1.1</td>
<td><em>Phalaenopsis gigantea</em> inflorescence in flowering stage at Royal Botanical Garden, Kew.</td>
</tr>
<tr>
<td>Figure 4.2.1</td>
<td><em>Phalaenopsis gigantea</em> inflorescence.</td>
</tr>
</tbody>
</table>
Figure 4.2.2  Schematic representation of scoring methodology. 157

Figure 4.3.1  Analysis of *P. gigantea* genomic DNA using 1% Agarose gel. 164

Figure 4.3.2  Gel electrophoresis of PCR amplicons representing microsatellite locus PGAC05. 166

Figure 4.3.3  Gel electrophoresis of PCR amplicons representing microsatellite locus PGCT35. 166

Figure 4.3.4  Gel electrophoresis of PCR amplicons representing microsatellite locus PGAC04B. 167

Figure 4.3.5  Gel electrophoresis of PCR amplicons representing microsatellite locus PGV01. 167

Figure 4.3.6  Gel electrophoresis of PCR amplicons representing microsatellite locus PGJ45. 168

Figure 4.3.7  Gel electrophoresis of PCR amplicons representing microsatellite locus PGAC36. 168

Figure 4.3.8  Graphical representation of Expected Heterozygosity (*H*<sub>e</sub>) and Expected equilibrium gene diversity (*H*<sub>eq</sub>) at each of the thirty loci tested. 178

Figure 4.3.9  Phylogenetic tree reconstructed by using UPGMA from the GD matrix (Nei, 1978) for the population of *P. gigantea* from Tawau Hills Park. 184

Figure 4.3.10  Graphical representation of Genetic Distance (Nei, 1978) using the Neighbor Joining method. 185

Figure 4.3.11  Consensus tree generated after 1000 iterations using PHYLIP. 186

Figure 4.3.12  Schematic representation of Neis' Genetic Distance as functions of frequency distribution, 75 combinations of genotypes exhibit a genetic distance in excess of 0.5 and are an indication of the possible breeding pairs for improvement of gene diversity levels in the F1 generation. 187

Figure 5.2.1  *Paphiopedilum rothschildianum* plants flowering at the ex-situ conservation facility of the Sabah Parks Authority located at Kinabalu National Park. 217

Figure 5.3.1  Gel electrophoresis of PCR amplicons representing microsatellite locus PRA12 depicting four alleles of this heterozygous locus. 221
Figure 5.3.2 Gel electrophoresis of PCR amplicons representing microsatellite locus PRC16 depicting two alleles of this heterozygous microsatellite locus. 222

Figure 5.3.3 Gel electrophoresis of PCR amplicons representing microsatellite locus microsatellite locus PRE22A depicting the single allele of this homozygous, monomorphic, tetranucleotide locus. 222

Figure 5.3.4 Gel electrophoresis of PCR amplicons representing microsatellite locus microsatellite locus PRC15 depicting the three alleles amplified in this locus. 223

Figure 5.3.5 Gel electrophoresis of PCR amplicons representing microsatellite locus microsatellite locus PR704A depicting the two alleles amplified in this locus. 223

Figure 5.3.6 Gel electrophoresis of PCR amplicons representing microsatellite locus microsatellite locus PRC13 depicting the two alleles amplified in this locus. 224

Figure 5.3.7 Gel electrophoresis of PCR amplicons representing microsatellite locus microsatellite locus PRA17 depicting the three alleles amplified in this locus. 224

Figure 5.3.8 Gel electrophoresis of PCR amplicons representing microsatellite locus microsatellite locus PRC09 depicting the three alleles amplified in this locus. 225

Figure 5.3.9 Neighbor Joining Tree for populations of *P. rothschildianum* constructed using pairwise genetic distance data. 234

Figure 5.3.10 UPGMA dendrogram showing relative rates of evolution between populations of *P. rothschildianum* computed using the genetic distance data. 235

Figure 5.3.11 UPGMA tree for the population of Bukit Ampuan. 236

Figure 5.3.12 UPGMA tree for the population of Melangkap. 237

Figure 5.3.13 UPGMA tree for the population of Telupid. 238

Figure 5.3.14 Consensus UPGMA tree generated after 1000 iterations for the overall population using PHYLP. 239

Figure 5.3.15 Consensus Neighbor Joining tree generated after 1000 iterations for the overall population. 240

Figure 5.3.16 Minimum Evolution tree generated after for the overall population using MEGA 4.0. 241
Figure 6.2.1 *Paphiopedilum lowii* specimen flowering at the Sabah Parks Authority *ex-situ* conservation facility at Kinabalu national park, Sabah, Malaysia. 254

Figure 6.2.2 *Paphiopedilum dayanum* specimen flowering at the Sabah Parks Authority *ex-situ* conservation facility at Kinabalu national park, Sabah, Malaysia. 254

Figure 6.3.1 PCR-RFLP of chloroplast DNA loci PSIA1, PSIA2 and PSIID2 derived from *Paphiopedilum dayanum*, *Paphiopedilum lowii* and *Paphiopedilum rothschildianum* using the restriction enzyme *Taq* I. 259

Figure 6.3.2 Gel electrophoresis of loci derived from *P. lowii*. 264

Figure 6.3.3 Gel electrophoresis of loci derived from *P. dayanum*. 264
LIST OF SYMBOLS AND ABBREVIATIONS

% Percent
: Ratio
= Equal
°C Degree of Celsius
α Alpha
A Adenine
Acc. Accession number
AFLP Amplified Fragment Length Polymorphism
β Beta
blastn Basic local alignment search tool for nucleotide
bp Base pair
CaCl₂ Calcium chloride
C Cytosine
cm Centimeter
DNA Deoxyribonucleic acid
DNase Deoxyribonuclease
CTAB Cetyl-trimethly ammonium bromide
dATP deoxyadenosine-5'-triphosphate
dCTP deoxycytidine-5'-triphosphate
dGTP deoxyguanosine-5'-triphosphate
dTTP deoxythymidine-5'-triphosphate
dNTP deoxynucleoside-5'-triphosphate
ddH₂O Double distilled water
E. coli Escherichia coli
EDTA Ethylene-diamine-trichloro-acetic-acid
EtBr Ethidium Bromide
g Gram
G Guanine
hr Hour
IPTG Isopropyl -1-thio-β-D-galactoside
Kb Kilo base
KCl Potassium chloride
L Litre
LB Luria-Bertani
\( \lambda \) Lambda
mM Milimolar
MgCl\(_2\) Magnesium chloride
\( \mu \) Micro
M Molar
min Minutes
\( \mu \text{L} \) Microlitre
\( \mu \text{g} \) Microgram
ng/uL Nanogram per microlitre
\( \mu \text{g/L} \) Microgram per litre
NaOH Sodium hydrochloride
p Pico
PCR Polymerase Chain Reaction
PAGE Polyacrylamide Gel Electrophoresis
RFLP Restriction Fragment Length Polymorphism
RAPD Random Amplified Polymorphic DNA
RNA Ribonucleic acid
RNase Ribonuclease
rpm Revolutions per minute
s Seconds
SDS Sodium dodecyl sulphate
SSR Simple sequence repeat
SNP Single Nucleotide Polymorphism
TBE Tris- Boric-EDTA
TAE Tris- Acetic Acid - EDTA
TE Tris-EDTA
T Thymine
UV Ultra violet
U Unit
V Volt
REFERENCES


