

**DEVELOPMENT AND CHARACTERIZATION
OF ANONYMOUS NUCLEAR MARKERS TO
DETERMINE GENETIC VARIATION IN
TWO GROUPERS SPECIES AND THEIR
HYBRIDS**

PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH



AHMAD ZAIDI BIN TANI

UMS
UNIVERSITI MALAYSIA SABAH

**BIOTECHNOLOGY RESEARCH INSTITUTE
UNIVERSITI MALAYSIA SABAH
2015**

**DEVELOPMENT AND CHARACTERIZATION
OF ANONYMOUS NUCLEAR MARKERS TO
DETERMINE GENETIC VARIATION IN
TWO GROUPERS SPECIES AND THEIR
HYBRIDS**

PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH



AHMAD ZAIDI BIN TANI

UMS
UNIVERSITI MALAYSIA SABAH

**THESIS SUBMITTED IN PARTIAL
FULFILLMENT FOR THE DEGREE OF
MASTER OF SCIENCE**

**BIOTECHNOLOGY RESEARCH INSTITUTE
UNIVERSITI MALAYSIA SABAH
2015**

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS

JUDUL: DEVELOPMENT AND CHARACTERIZATION OF ANONYMOUS NUCLEAR MARKERS TO DETERMINE GENETIC VARIATION IN TWO GROUPERS SPECIES AND THEIR HYBRIDS

IJAZAH: SARJANA SAINS dalam bidang BIOTEKNOLOGI

Saya AHMAD ZAIDI TANI, Sesi Pengajian 2011-2015, mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis ini 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. Sila tandakan (/)

SULIT (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)


TERHAD (Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

TIDAK TERHAD

Disahkan oleh,

NURULAIN BINTI ISMAIL
LIBRARIAN

UNIVERSITI MALAYSIA SABAH

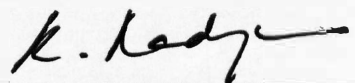


(Tandatangan Penulis)



(Tandatangan Pustakawan)

Tarikh: 31 Ogos 2015

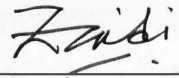


(DR. KENNETH FRANCIS RODRIGUES)
Penyelia

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, excerpts, equations, summaries and references, which have been duly acknowledged.

27 August 2015



Ahmad Zaidi Bin Tani
PB20118049



UMS
UNIVERSITI MALAYSIA SABAH

CERTIFICATION

NAME : **AHMAD ZAIDI BIN TANI**

MATRIC NO : **PB20118049**

TITLE : **DEVELOPMENT AND CHARACTERIZATION OF ANONYMOUS
NUCLEAR MARKERS TO DETERMINE GENETIC VARIATION
IN TWO GROUPERS SPECIES AND THEIR HYBRIDS**

DEGREE : **MASTER OF SCIENCE (BIOTECHNOLOGY)**

VIVA DATE : **23 APRIL 2015**

CERTIFIED BY

1. SUPERVISOR

Dr. Kenneth Francis Rodrigues

Signature



UMS
UNIVERSITI MALAYSIA SABAH

K. Rodrigues

ACKNOWLEDGEMENT

I would like to express my gratitude to Allah S.A.W for granting me the opportunity for doing my Master in Science (Biotechnology). I am deeply indebted to my supervisor Dr. Kenneth Francis Rodrigues for entrusting this research to me and supervising me throughout the research. Thank you for stimulating suggestion and encouragement throughout the time of my research and thesis writing. To my former colleagues (Noor Hydayaty, Ong Wen Dee, Mohd Faiz and Senty), lab assistants (Vidarita, Mony and Azian) lecturers (Assoc. Prof. Dr. Vijay Kumar) and management team (Mdm Noor Asha); thank you for all of their help, support and interest. To my lovely family; my parents En Tani Bin Duet and Pn Betty Chin, my sisters and brother, Ibrahim, Zaiton and Normala, thank you for giving me your full support, love and understanding all this while till the completion of my study and this thesis.

Ahmad Zaidi Tani

23 April 2015



UMS
UNIVERSITI MALAYSIA SABAH

ABSTRACT

Grouper (Genus: *Epinephelus*) represent one of the high value fisheries resource in Malaysia and several inter-specific hybrids have been developed by the fish breeders. DNA based markers are ideal elucidation of the fundamental associated with inheritance and linkage in determinate hybrids. First generation markers (Microsatellite, RAPD, ISSR etc.) have limitations when applied to characterize intraspecific hybrids. This study set forth to overcome this limitation by developing the second generation of non-coding non informative anonymous nuclear markers (ncni-ANM) for subsequent genetic characterization of F1 grouper hybrids. Small-insert genomic DNA library of *Epinephelus lanceolatus* (EL) and *Epinephelus fuscoguttatus* (EF) were constructed via a shotgun approach following which they were annotated, subjected to a similarity searches using the blastn algorithms and deposited at the GenBank. Annotated genomic sequences of *E. lanceolatus* and *E. fuscoguttatus* were then used for designation of anonymous nuclear markers (ANM) and further categorized into being non-coding non-informative ANM (ncni-ANM), non-coding informative ANM (nci-ANM) or coding informative ANM (ci-ANM). Intraspecific DNA fingerprinting analysis was carried out using five *E. lanceolatus* and 12 *E. fuscoguttatus* broodstock, scored as binary matrix and calculated for genetic distance (GD), genetic similarity (GS) and cophenetic correlation index (CCC) via Jaccard's and Dice's coefficient. An unweighted pair-group method with arithmetic mean (UPGMA) phenogram was constructed based on the GD and CCC values to verify the genetic relationship among the grouper broodstock operational taxonomic units (OTUs). Subsequently, thirty two half-sib F1 hybrids obtained by crossing of female *E. fuscoguttatus* and male *E. lanceolatus* (EFEL) brooders were characterized using 34 male *E. lanceolatus* and 32 *E. fuscoguttatus* ncni-ANM in order to evaluate Mendelian segregation pattern based on genetic inheritance and segregation distortion analyses. A library of 127 genomic sequences was developed of which seven *E. lanceolatus* and 16 *E. fuscoguttatus* sequences could be mapped back to the *Dicentrarchus labrax* genome. Locus specific markers were developed using 60 *E. lanceolatus* and 67 *E. fuscoguttatus* genomic sequences. A total of 46 *E. lanceolatus* ANM were designed containing 34 ncni-ANM, 4 nci-ANM and 8 ci-ANM. While as in total of 42 *E. fuscoguttatus* ANM were designed containing 32 ncni-ANM, 3 nci-ANM and 7 ci-ANM. Polymerase chain reaction (PCR) amplification patterns of ncni-ANM in both *E. lanceolatus* and *E. fuscoguttatus* broodstocks revealed 11 *E. lanceolatus* and 10 *E. fuscoguttatus* ncni-ANM were reproducible. Among these 13 *E. lanceolatus* and *E. fuscoguttatus* ncni-ANM loci ranged in size from 153 to 177 bp and 157 to 624 bp respectively revealing variation across all OTUs through visualization of agarose gel electrophoresis patterns and binary matrix scoring. Clustering comparison based on CCC goodness-of-fit showed both species samples were best presented using Jaccard's coefficient. The phenogram constructed for the five *E. lanceolatus* OTUs showed two clusters with polyphyletic groups. The first cluster (C1) consisted of OTUs EL6924, EL5477 and EL140E with monophyletic group while as the second cluster (C2) consisted of OTUs EL4E4D and EL174A with

paraphyletic grouping. Four clusters of polyphyletic groups were revealed when a phenogram was constructed based on the 12 *E. fuscoguttatus* OTUs. The first and second clusters (C1 and C2) consisted of OTUs EF02X, EF002, EF009 and EF03X; and OTUs EF001, EF006, EF005 and EF006 respectively which showed monophyletic grouping. The third and fourth clusters (C3 and C4), consisted of OTUs EF007 and EF0010; and OTU EF003 which showed paraphyletic grouping. Genetic distances (≤ 0.500) of all five *E. lanceolatus* OTUs (EL6924, EL140E, EL434D and EL174A) and four *E. fuscoguttatus* OTUs (EF02X, EF003, EF005 and EF008) ranged from 0.182-0.500 and 0.364-0.500 respectively; implying that high dissimilarities among OTUs were well characterized based on UPGMA. Applying the same nci-ANM, 34 *E. lanceolatus* and 32 *E. fuscoguttatus* markers showed acceptable monohybrid segregation patterns in which 10 and 24 *E. lanceolatus*; 11 and 22 *E. fuscoguttatus* were grouped as 1:1 and 3:1 segregation pattern ratios respectively, thus proving the accordance of nci-ANM to Mendelian segregation patterns. Analysis of segregation distortion analysis (χ^2) proved that among the reported segregation loci, only six *E. lanceolatus* and 12 *E. fuscoguttatus* nci-ANM were in accordance ($\chi^2 < 3.841$) with Mendelian law while the remaining were classified as distorted loci. This is the first genomic library and the first report that affirmed the use of nci-ANM markers in DNA fingerprinting and segregation inheritance study. Thus, the preliminary isolation of nci-ANM of groupers in the study of groupers genotyping and inheritance is revealed and be made feasible and benefit the Malaysian aquaculture genetics research and industry.



UMS
UNIVERSITI MALAYSIA SABAH

ABSTRAK

PEMBANGUNAN DAN PENCIRIAN PENANDA NUKLEAR UNTUK MENENTUKAN VARIASI GENETIK DALAM DUA SPESIES KERAPU DAN HIBRID MEREKA

Kerapu (Genus: Epinephelus) merupakan salah satu daripada sumber perikanan bernilai tinggi di Malaysia, dimana terdapat beberapa kacukan kerapu khas yang telah dihasilkan dan diusahakan oleh penternak ikan. Penanda DNA merupakan penjelasan yang ideal bagi asas berkaitan dengan perwarisan dan hubungan di dalam hibrid; namun penanda DNA generasi pertama (Mikrosatelit, RAPD, ISSR dll) mempunyai had apabila digunakan untuk mencirikan kacukan intraspecies. Kajian ini ditetapkan untuk mengatasi had tersebut dengan membangunkan generasi kedua penanda DNA non-coding non informative anonymous nuclear marker (ncni-ANM) untuk digunakan di dalam pencirian genetik kerapu hibrid F1. Perpustakaan DNA genom skala kecil Epinephelus lanceolatus (EL) dan Epinephelus fuscoguttatus (EF) telah dibina melalui pendekatan berikut shotgun carian persamaan jujukan menggunakan algoritma blastn dan seterusnya diserahkan kepada GenBank. Beranotasikan jujukan genom E. lanceolatus dan E. fuscoguttatus, penanda ANM telah berjaya dikategorikan kepada non-coding non-informative ANM (ncni-ANM), non-coding informative ANM (nci-ANM) atau coding informative ANM (ci-ANM). Analisis intraspecies cap jari DNA telah dijalankan menggunakan lima induk E. lanceolatus dan 12 induk E. fuscoguttatus, dijaringkan sebagai matriks binari dan jarak genetik (GD), persamaan genetik (GS) dan indeks korelasi cophenetic (CCC) dikira menggunakan pekali Jaccard dan Dice. Kaedah pasangan-kumpulan unweighted dengan aritmetik min (UPGMA) phenogram telah dibina berdasarkan nilai GD dan CCC bagi mengesahkan hubungan genetik di antara induk ikan kerapu. Sebanyak, 32 hibrid F1 separuh keluarga yang dihasilkan melalui pembiakan bersilang diantara E. fuscoguttatus betina dan E. lanceolatus jantan (EFEL) telah dicirikan menggunakan 34 penanda jantan E. lanceolatus dan 32 penanda betina E. fuscoguttatus ncni-ANM bagi menganalisa corak pengasingan berasaskan prinsip Mendel dan analisis herotan pengasingan. Perpustakaan 127 jujukan genom telah dibangunkan di mana tujuh jujukan E. lanceolatus dan 16 jujukan E. fuscoguttatus berjaya dipetakan semula kepada genom Dicentrarchus labrax. Penanda lokus khas telah dihasilkan berdasarkan 60 jujukan E. lanceolatus dan 67 jujukan E. fuscoguttatus. Sebanyak 46 penanda ANM E.lanceolatus telah direka mengandungi 34 ncni-ANM, 4 nci-ANM dan 8 ci-ANM. Manakala 42 penanda ANM E.fuscoguttatus telah Berjaya direka mengandungi 32 ncni-ANM, 3 IKN-ANM dan 7 ci-ANM. Corak amplifikasi ncni-ANM di dalam kedua-dua induk ikan E. lanceolatus dan E. fuscoguttatus mendedahkan 11 penanda E. lanceolatus dan 10 penanda E. fuscoguttatus ncni-ANM menunjukkan penghasilan semula. Antaranya 13 lokus E. lanceolatus dan E. fuscoguttatus ncni-ANM di antara saiz 153-177 bp dan 157-624 bp menunjukkan variasi melalui gel agarose elektroforesis dan jaringan matriks binari. Pengelompokan perbandingan berdasarkan nilai CCC, menunjukkan kedua-

dua spesies terbaik dikelaskan dengan menggunakan pekali Jaccard. Phenogram untuk lima induk *E. lanceolatus* membentuk dua kelompok dengan kumpulan polyfili. Kelompok pertama (C1) terdiri daripada EL6924, EL5477 dan EL140E iaitu kumpulan monofili manakala kelompok kedua (C2) terdiri daripada EL4E4D dan EL174A iaitu kumpulan parafili. Empat kelompok kumpulan polyfili telah dibina berdasarkan 12 induk *E. fuscoguttatus*. Kelompok pertama dan kedua (C1 dan C2) terdiri daripada EF02X, EF002, EF009 dan EF03X; dan EF001, EF006, EF005 dan EF006 dimana masing-masing yang menunjukkan kumpulan monofili. Manakala kelompok ketiga dan keempat (C3 dan C4), yang terdiri daripada EF007 dan EF0010; dan EF003 menunjukkan kumpulan parafili. Jarak genetik (≤ 0.500) di antara lima *E. lanceolatus* (EL6924, EL140E, EL434D dan EL174A) dan empat *E. fuscoguttatus* (EF02X, EF003, EF005 dan EF008) adalah di antara 0.182-0.500 dan 0.364-0.500; dimana ketidaksamaan yang tinggi telah berjaya dicirikan berdasarkan UPGMA. Selain itu, 34 dan 32 penanda *E. lanceolatus* dan *E. fuscoguttatus* juga menunjukkan corak pengasingan monomorfik di mana 10 dan 24 penanda *E. lanceolatus*; dan 11 dan 22 penanda *E. fuscoguttatus* telah dikelaskan sebagai nisbah perwarisan 1:1 dan 3:1, sekali gus membuktikan kecekapan nci-ANM dalam menganalisa corak perwarisan Mendel. Analisis herotan pengasingan (χ^2) membuktikan bahawa di antara pengasingan lokus yang dilaporkan, hanya enam lokus *E. lanceolatus* dan 12 lokus *E. fuscoguttatus* nci-ANM diterima ($\chi^2 < 3,841$) berdasarkan prinsip Mendel manakala bakinya dikelaskan sebagai lokus tidak berdasarkan prinsip Mendel; sekaligus menunjukkan pengasingan sekata. Ini merupakan perpustakaan genom dan laporan pertama yang mengesahkan fungsi penanda generasi kedua nci-ANM di dalam analisis cap jari DNA dan kajian pengasingan warisan. Oleh itu, kepentingan nci-ANM dalam genotyping dan kajian warisan kerapu telah dinyatakan dan dilaksanakan dimana memberikan manfaat kepada penyelidikan dan industri genetik akuakultur di Malaysia.

TABLE OF CONTENT

TITLE	Page
DECLARATION	ii
CERTIFICATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
<i>ABSTRAK</i>	vii
LIST OF CONTENTS	ix
LIST OF TABLES	xii
LIST OF FIGURES	xv
LIST OF ABBREVIATIONS	xviii
LIST OF SYMBOL	xix
LIST OF APPENDIX	xx
CHAPTER 1: INTRODUCTION	1
1.1 Research Background	1
CHAPTER 2: LITERATURE REVIEW	3
2.1 Groupers and Taxon	3
2.1.1 <i>Epinephelus lanceolatus</i> (Bloch, 1790)	4
2.1.2 <i>Epinephelus fuscoguttatus</i> (Forsskal, 1775)	6
2.2 Breeding and Status of Groupers Culture in Southeast Asia	8
2.2.1 Problematic Issues in Breeding and Productions of Groupers	9
2.3 Fish Cross-Breeding and Hybridization	13
2.4 Genomic DNA Libraries	15
2.4.1 Construction and Screening of Genomic DNA Libraries	15
2.4.2 Genomic Library Construction Approaches	17
2.5 DNA Marker Technologies	18

CHAPTER 3: METHODOLOGY	20
3.1 Construction of Small Insert Genomic DNA Library	20
3.1.1 Sample Collection	20
3.1.2 Isolation of Total Genomic DNA	21
3.1.3 Preparation pUC19 Cloning Vector	22
3.1.4 Quantification and Determination of Genomic and Plasmid Purity and Concentration	23
3.1.5 Enzymatic Restriction of Genomic and Plasmid DNA	24
3.1.6 Genomic DNA Purification (QIAquick PCR Purification Kit)	26
3.1.7 Plasmid DNA Purification (QIAquick Gel Extraction Kit)	26
3.1.8 Cloning of 14 sets of Double Digestion Products	28
a. Preparation of One-shot CaCl ₂ Chemically Competent <i>E. coli</i> TOP10 Cells	28
b. Ligation	29
c. Transformation	29
d. Analysis of Recombinant Clones	30
e. Plasmid DNA Extraction and Glycerol Stocks Preparation	31
3.1.9 Single Pass Plasmid Sequencing, Vector Trimming and Genbank Submission	32
3.1.10 Bioinformatics Analysis	33
a. Repeat Masking Analysis	33
b. Similarities Search Analysis (BLAST)	33
3.2 Development of Anonymous Nuclear Markers (ANM)	34
3.2.1 Primer Designation and Selection	34
3.2.2 Primer Optimization	35
3.2.3 Primer-Amplicon Anonymity Analysis	37
a. Anonymity and Primer Classification Analysis	37
b. Primer Nomenclature	38
3.3 Molecular Characterization and Intra-Species Analysis Groupers	39
3.3.1 Fish Sampling and Genomic DNA Extraction	39
3.3.2 PCR Amplification and the Resolution of ncni-ANM	39
3.3.3 Phenetic Analysis	40
3.4 Mendelian Inheritance Patterns in F1 Hybrid Grouper <i>Epinephelus fuscoguttatus</i> X <i>Epinephelus lanceolatus</i>	42
3.4.1 Generation and Sampling of F1 Hybrid Groupers from Known Families	42
3.4.2 Isolation and Quantification of Total Genomic DNA	42
3.4.3 Amplification of Non-Coding Non-Informative Anonymous Nuclear Markers	43
3.4.4 Band scoring and Segregation Distortion Analysis	43
CHAPTER 4: RESULT AND DISCUSSIONS	46
4.1 Construction of Small Insert Genomic DNA Library	46

4.1.1	Fish Sampling, Total Genomic DNA Extraction and pUC19 Cloning	46
4.1.2	Enzymatic Restriction of Genomic DNA and pUC19	49
4.1.3	Cloning, Transformation and the Analysis of Recombinant Clones	50
4.1.4	Plasmid Minipreparation, Single Pass Plasmid Sequencing , Genbank Submission and Insert Size Clustering	54
4.1.5	Bioinformatics Analysis	59
a.	Repeat Masking Analysis	59
b.	Characterization of Genomic Sequences from Groupers DNA Library	63
i.	Nucleotide Similarity Analysis (blastN)	63
ii.	Translated Nucleotide Similarity Analysis (blastX)	66
4.2	Development of Anonymous Nuclear Markers (ANM)	78
4.2.1	Primer Designation and Selection	78
4.2.2	Primer Optimization	87
4.2.3	Primer-Amplicon Anonymity Analysis	94
4.3	Application of Non-coding Non-informative Anonymous Nuclear Markers (ncni-ANM) in Molecular Characterization of Intra-species Groupers	101
4.3.1	Generation of ncni-ANM and Identification of Strain-Specific Diagnostic Markers	101
4.3.2	Cluster Analysis	106
4.4	Application of Non-coding Non-informative Anonymous Nuclear Markers (ANM) in Mendelian Inheritance of F1 Hybrid Grouper	117
4.4.1	Amplification of Anonymous Nuclear Markers (ANM)	117
4.4.2	Genetic Inheritance Analysis	119
4.4.3	Segregation Distortion	127
	CHAPTER 5: CONCLUSION	133
5.1	Research Summary	133
	REFERENCES	136
	APPENDIX	151

LIST OF TABLES

		Page
Table 2.1:	Summary of Grouper Breeding and Status in Six Other Country in Southeast Asia (Pomeroy <i>et al.</i> , 2002) (Cont.)	11
Table 2.2:	Examples of Inter-specific Hybrid Groupers (Cont.)	14
Table 2.3:	Comparison between Genomic and cDNA Libraries (Ausubel <i>et al.</i> , 1991)	16
Table 2.4:	Genomic Library Construction	17
Table 3.1:	Fourteen Sets of Restriction Enzymes, Recommended Buffer and Thermal Inactivation (Thermo Scientific, USA)	26
Table 3.2:	Similarity Hits Clusters	34
Table 3.3:	General Primer Designation Parameters	35
Table 3.4:	Example of Grouper Primer Nomenclature	38
Table 3.5:	List of Non-coding Non-Informative Anonymous Markers (ncni-ANM) used in the PCR Amplification	39
Table 3.6:	Groupers F1 Hybrid and Parental Information	42
Table 3.7:	F1 hybrid grouper of <i>E. fuscoguttatus</i> X <i>E. lanceolatus</i> analysis (EFEL)	43
Table 3.8:	Segregation Calculation, Grouping and Ratio (Lallias <i>et al.</i> , 2007)	44
Table 3.9:	χ^2 Decision Rule Interpretation (Lallias <i>et al.</i> , 2007)	45
Table 4.1:	Example of genomic DNA and pUC19 cloning vector purity and concentration quantification	47
Table 4.2:	Number of Restriction Enzyme (RE) Combination Sets with Successful and Unsuccessful Transformation	51
Table 4.3:	Example of Acceptable and Edited Plasmid Sequence of <i>E. lanceolatus</i> Genomic Sequence EL101 in FASTA Format	55
Table 4.4:	Genbank Submission Accession No.	55

Table 4.5:	Total Number of Plasmids Based on Species and RE Combination Sets	58
Table 4.6:	Summarization of Repeat Masking Analysis of Two Groupers Genomic Libraries	60
Table 4.7:	Simple Repeats Motifs	60
Table 4.8:	Distribution and occurrences of masked simple repeats motifs	59
Table 4.9:	Distribution of low complexity motifs	61
Table 4.10:	BlastN Analysis of 60 Genomic sequences of <i>E. lanceolatus</i> (Cont.)	69
Table 4.11:	BlastN Analysis of 67 Genomic sequences of <i>E. fuscoguttatus</i> (Cont.)	71
Table 4.12:	BlastN Identities and Hits Distribution	64
Table 4.13:	BlastN Significant Hits Distribution based on E-value	64
Table 4.14:	BlastX Analysis of 60 Genomic sequences of <i>E. lanceolatus</i> (Cont.)	75
Table 4.15:	BlastX Analysis of 67 Genomic sequences of <i>E. fuscoguttatus</i> (Cont.)	76
Table 4.16:	BlastX Identities and Hits Distribution	66
Table 4.17:	BlastX Significant Hits Distribution based on E-value	66
Table 4.18:	Total Number of Successful and Unsuccessful Generated Primers	78
Table 4.19:	<i>E. lanceolatus</i> Primer List and Descriptions (Cont.)	79
Table 4.20:	<i>E. fuscoguttatus</i> Primer List and Descriptions (Cont.)	82
Table 4.21A:	Total Number of Successful and Unsuccessful Optimized Primers	87
Table 4.21B:	<i>Epinephelus lanceolatus</i> Primer-Amplicons Anonymity Analysis (Cont.)	97

Table 4.22:	<i>Epinephelus fuscoguttatus</i> Primer-Amplicons Anonymity Analysis (Cont.)	99
Table 4.23:	Generation of 11 of Highly Variable <i>E. lanceolatus</i> ncnI-ANM	104
Table 4.24:	Generation of 10 Highly Variable <i>E. fuscoguttatus</i> ncnI-ANM	105
Table 4.25:	Similarity and Distance Matrix Computed with Jaccard's Coefficient	106
Table 4.26:	Similarity and Distance Matrix Computed with Dice's Coefficient	106
Table 4.27:	Cophenetic Correlation Index Value Based on Jaccard's and Dice' Coefficient for <i>E. lanceolatus</i> Phenogram	107
Table 4.28:	Similarity and Distance Matrix Computed with Jaccard's Coefficient for 12 OTUs of <i>Epinephelus Fuscoguttatus</i>	112
Table 4.29:	Similarity and Distance Matrix Computed with Dice's Coefficient for 12 OTUs of <i>Epinephelus fuscoguttatus</i>	113
Table 4.30:	Cophenetic Correlation Index Value Based on Jaccard's and Dice' Coefficient for and <i>E. fuscoguttatus</i>	108
Table 4.31:	Amplification of 34 <i>Epinephelus lanceolatus</i> ncnI-ANM Primers Representing Male EL6924 Markers against Female parent of <i>Epinephelus fucoguttatus</i> EF02X and 32 F1 Hybrid of EFEL (Sample 1-32) (Cont.)	123
Table 4.32:	Amplification of 32 <i>Epinephelus fuscoguttatus</i> ncnI-ANM Primers Representing Female EF02X Markers against Male parent of <i>Epinephelus lanceolatus</i> EL6924 and 32 F1 Hybrid of EFEL (Sample 1-32) (Cont.)	125
Table 4.33:	Segregation Distortion Analysis of 34 Locus of <i>Epinephelus lanceolatus</i> EL6924	128
Table 4.34:	Segregation Distortion Analysis of 32 Locus of <i>Epinephelus fuscoguttatus</i> EF02X	129
Table 4.35:	Segregation distortion Analysis for 1:1 and 3:1 for Mapping 32 F1 Hybrid Progenies	127

Table 4.22:	<i>Epinephelus fuscoguttatus</i> Primer-Amplicons Anonymity Analysis (Cont.)	99
Table 4.23:	Generation of 11 of Highly Variable <i>E. lanceolatus</i> ncnI-ANM	104
Table 4.24:	Generation of 10 Highly Variable <i>E. fuscoguttatus</i> ncnI-ANM	105
Table 4.25:	Similarity and Distance Matrix Computed with Jaccard's Coefficient	106
Table 4.26:	Similarity and Distance Matrix Computed with Dice's Coefficient	106
Table 4.27:	Cophenetic Correlation Index Value Based on Jaccard's and Dice' Coefficient for <i>E. lanceolatus</i> Phenogram	107
Table 4.28:	Similarity and Distance Matrix Computed with Jaccard's Coefficient for 12 OTUs of <i>Epinephelus Fuscoguttatus</i>	112
Table 4.29:	Similarity and Distance Matrix Computed with Dice's Coefficient for 12 OTUs of <i>Epinephelus fuscoguttatus</i>	113
Table 4.30:	Cophenetic Correlation Index Value Based on Jaccard's and Dice' Coefficient for and <i>E. fuscoguttatus</i>	108
Table 4.31:	Amplification of 34 <i>Epinephelus lanceolatus</i> ncnI-ANM Primers Representing Male EL6924 Markers against Female parent of <i>Epinephelus fuscoguttatus</i> EF02X and 32 F1 Hybrid of EFEL (Sample 1-32) (Cont.)	123
Table 4.32:	Amplification of 32 <i>Epinephelus fuscoguttatus</i> ncnI-ANM Primers Representing Female EF02X Markers against Male parent of <i>Epinephelus lanceolatus</i> EL6924 and 32 F1 Hybrid of EFEL (Sample 1-32) (Cont.)	125
Table 4.33:	Segregation Distortion Analysis of 34 Locus of <i>Epinephelus lanceolatus</i> EL6924	128
Table 4.34:	Segregation Distortion Analysis of 32 Locus of <i>Epinephelus fuscoguttatus</i> EF02X	129
Table 4.35:	Segregation distortion Analysis for 1:1 and 3:1 for Mapping 32 F1 Hybrid Progenies	127

LIST OF FIGURES

	Page
Figure 2.1: <i>E. lanceolatus</i> morphological differences A: Adult Stages; B: Juvenile Stages (Heemstra and Randall, 1993).	6
Figure 2.2: <i>E. fuscoguttatus</i> Adult Morphology (Heemstra and Randall, 1993).	7
Figure 3.1: A) pUC19 Cloning Vector-Circularized (Thermo Scientific, USA); B) pUC19 Multiple cloning sites (Thermo Scientific, USA).	25
Figure 3.2: Range of Genomic Sequence Subjected to Anonymity Analysis	37
Figure 4.1: Example of the analysis of grouper genomic DNA on 1% agarose gel. Lane M: Lambda DNA/ <i>Hind III</i> marker (Thermo Scientific, USA); Lane 1: Genomic DNA of <i>E. lanceolatus</i> isolate EL6924; Lane 2: Genomic DNA of <i>E. fuscoguttatus</i> isolates EF02X.	48
Figure 4.2: Example of successfully digested total genomic DNA, digested pUC19 plasmid and undigested pUC19 plasmid on 1% agarose gel. Lane M: 1kb DNA Ladder (Thermo Scientific, USA); Lane 1: Digested Genomic DNA of <i>E. fuscoguttatus</i> using <i>BamHI</i> and <i>XbaI</i> ; Lane 2: Undigested pUC19; Lane 3: Digested pUC19.	50
Figure 4.3: Example of colony PCR for <i>E. lanceolatus</i> transformant. Lane M: 1 kb DNA marker (Thermo Scientific); Lane 1-16 : Colony PCR products.	52
Figure 4.4: Example of plasmid minipreparation of <i>E. fuscoguttatus</i> clones using Thermo Scientific Plasmid Miniprep Kit. Lane M: 1 kb DNA Ladder (Thermo Scientific, USA); Lane 1-11: Plasmid minipreparation products. Red Box: Example of Supercoiled Plasmid.	54
Figure 4.5: Example of Good Sequencing of <i>E. fuscoguttatus</i> Plasmid EFJ002 clone Sequencing Result	55
Figure 4.6: Distribution of 127 Groupers Genomic Sequences based on Fragment Size (bp)	56
Figure 4.7: Distribution of BlastN Hits Description	65
Figure 4.8: Distribution of BlastX Hits Description.	67
Figure 4.9: Successfully optimized <i>E. lanceolatus</i> primers.	89

Lane 100bp: 100 bp DNA ladder (Thermo Scientific, USA);
 Lane 1-16, 18-26, 28-38: *E. lanceolatus* primers
 (Refer Table 4.20 for primer descriptions) (Cont.).

- Figure 4.10: Successfully optimized *E. fuscoguttatus* primers. 91
 Lane 100bp: 100 bp DNA ladder (Thermo Scientific, USA);
 Lane 1-15, 16-27, 28-42: *E. fuscoguttatus*
 primers (Refer Table 4.21 for primer descriptions).
- Figure 4.11: Example of Primer Designation and Repeat Identification 9
 (RepeatMasker)
- Figure 4.12: Example of PCR Amplification for Four *Epinephelus* 102
lanceolatus OTUs. Lane M: 100 bp DNA Ladder
 (Thermo Scientific, USA); Lane 1: EL4E4D; Lane 2:
 EL140E; Lane 3: EL174A; Lane 4: EL5477. Sub-Figure 1
 (A-Primer ELJPK001: No Variation; B-ELJPK002: No
 Variation); Sub-Figure 2 (B-ELJK01: Variation Observed)
- Figure 4.13: UPGMA Phenogram showing relationship among 5 OTUs 109
 of *Epinephelus lanceolatus*. Phenogram are based on
 genetic distance value calculated by Jaccard's and Dice's.
 (C=Cluster).
- Figure 4.14: UPGMA Phenogram showing relationship among 12 OTUs 110
 of *Epinephelus fuscoguttatus*. Phenogram are based on
 genetic distance value calculated by Jaccard's. (C=Cluster).
- Figure 4.15: UPGMA Phenogram showing relationship among 12 OTUs 111
 of *Epinephelus fuscoguttatus*. Phenogram are based on
 genetic distance value calculated by Dice's. (C=Cluster).
- Figure 4.16 (a): Example of consistent PCR amplification of nci-ANM: 117
 ELJPK002 against EFEL01-EFEL16. Lane 100bp: 100bp
 DNA Ladder (Thermo Scientific, USA); Lane 1-16:
 F1 hybrid isolate EFEL01-EFEL16
- Figure 4.16 (b): Example of consistent PCR amplification of nci-ANM: 117
 ELJPK009 against EFEL01-EFEL16. Lane 100bp: 100bp
 DNA Ladder (Thermo Scientific, USA); Lane 1-16: F1
 hybrid isolate EFEL01-EFEL16.
- Figure 4.17 (a): Example of PCR amplification of nci-ANM markers 118
 against EFEL04. Lane 100bp: 100bp DNA Ladder
 (Thermo Scientific, USA); Lane 28: Marker EFPHI003;
 Lane 29: Marker EFPHI004; Lane 30: Marker
 EFPHI006; Lane 31: Marker EFBP001; Lane 32:
 Marker EFEH001; Lane 34: Marker EFEX002;
 Lane 35: Marker EFEX003; Lane 36: Marker EFEX004;

Lane 38: Marker EFHX003; Lane 39: Marker EFXB001;
Lane 40: Marker EFXB003A; Lane 41: Marker EFXB003B;
Lane 42: Marker EFXP001; Lane 33 and 37 indicates
ci-ANM.

Figure 4.15 (b): Example of no PCR amplification of nci-ANM:
EFP007 against EFEL01-EFEL16. Lane 100bp: 100bp
DNA Ladder (Thermo Scientific, USA); Lane 1-16: F1
hybrid isolate EFEL01-EFEL16

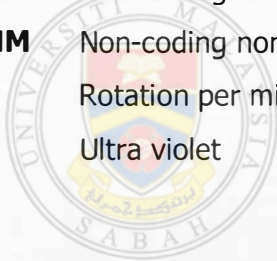
118



UMS
UNIVERSITI MALAYSIA SABAH

LIST OF ABBREVIATIONS

μL	Microliter
ng	Nanogram
bp	Base Pair
cm	Centimetre
dNTP	Deoxyribonucleotide triphosphate
g	Graviti
min	Minute
mL	Mililitre
mM	Milimolar
MgCl₂	Magnesium chloride
U	Unit
ci-ANM	Coding infomrative-anonymous nuclear markers
nci-ANM	Non-coding informative-anonymous nuclear markers
ncni-ANM	Non-coding non-informative-anonymous nuclear markers
rpm	Rotation per minute
UV	Ultra violet



UMS
UNIVERSITI MALAYSIA SABAH

LIST OF SYMBOL

°C	Degree Celsius
%	Percentage



UMS
UNIVERSITI MALAYSIA SABAH

LIST OF APPENDIX

	Page
Appendix A: List of groupers samples collection from Borneo Marine Research Institute (BMRI) Hatchery, Universiti Malaysia Sabah (UMS) and their standard characteristics measurements of BWt-body weight (kg), TL-total length (cm), SL-standard length (cm), HL-head length (cm), BH-body height (cm), BWd-body width (cm), M-Male, F-Female and U-Unknown sexes.	151
Appendix B: A representation of sequences from single pass sequencing of two different species of groupers clones	153
Appendix C: RepeatMasking full course analysis for two different species of groupers genomic DNA library sequences	159
Appendix D: Intra-species amplification	166
Appendix E: F1 Hybrid Grouper amplification	167



UMS
UNIVERSITI MALAYSIA SABAH

CHAPTER 1

INTRODUCTION

1.1 Research Background

Groupers are a carnivorous marine fish that inhabit the tropical and subtropical ocean worldwide. Groupers are farmed globally and represent high-value fisheries resources. Malaysia is one of the seven countries in Southeast Asia that produce high quantities of grouper breed. However, because of the high demand, small numbers of hatcheries produce and problematic issues in the production of groupers fry, the numbers of groupers production are tremendously decreasing.

The Hatchery of Borneo Marine Research Institute (BMRI), Universiti Malaysia Sabah (UMS) have conducted and successfully develop and produce approximately 10 different breeds of F1 hybrid groupers from a cross between various types of groupers from the genus of *Epinephelus*. The successful production of F1 hybrid crosses solve the aquaculture production. However, does not improve any of genetics study respectively.

During the past decade, genetic molecular markers (First generation markers) have been applied for various types of application in genetic study. The most common genetic markers used in the groupers study are microsatellite markers which used in the study of population genetics as well as mitochondrial DNA and various nuclear DNA markers which used in the study of strain and species identification. However for the past few years researcher have found that the competence of the existing genetic molecular markers is decreasing (Thomson *et al.*, 2010).