UNIVERSITI PUTRA MALAYSIA

DEVELOPMENT OF A PCR - CYCLE SEQUENCING SYSTEM FOR ANIMAL MITOCHONDRIAL DNA ANALYSIS

LAU CHIN HOON

FSAS 1994 10

DEVELOPMENT OF A PCR - CYCLE SEQUENCING SYSTEM FOR ANIMAL MITOCHONDRIAL DNA ANALYSIS

By

LAU CHIN HOON

the degree of Master of Science in the Faculty of
Science and Environmental Studies,
Universiti Pertanian Malaysia

March 1994

Dedicated to the readers:

Be true and simple.
If not, please leave science alone.

ACKNOWLEDGEMENTS

My heartiest gratitude to:

Dr. Khatijah Yusoff, Professor Dr. Yukio Yamada and Associate Professor Dr. Tan Soon Guan for their support in this research and supervision of this thesis;

Associate Professor Dr. Norani A. Samad for access to her facility at the Virology Laboratory, and also her capacity as the Head of Department of Biochemistry and Microbiology during the period of my M.S. candidacy;

Ms Lau Pin Sian, Mr. Kok Kah Mun and Mr. Lau Lam Kwee for their company at Serdang and for their assistance which rid the worry of daily life from my mind, without whom I could not possibly be able to concentrate on my research;

Associate Professor Dr. Gan Yik Yuen for introducing population genetics of mitochondrial DNA to me;

Dr. Hiroshi Yasue and his officers at the Laboratory of Cell Genetics, National Institute of Animal Industry, Tsukuba, for providing a 2-month training in 1992; the stint was sponsored by the Society of Techno-innovation in Agriculture, Forestry and Fisheries (STAFF), Japan;

Mr. Ng Ban Kim, Mr. Tan Wen Siang and all other researchers at Lab 142 / 143 whose enchanting interest in molecular genetics has liven my life in the lab;

Mr. Julian Lim of Bio-diagnostic for regular updates on advances in molecular techniques.

This study was supported by Intensification of Research in Priority Areas (IRPA) grant no. 1-07005-041 and 1-07005-027

TABLE OF CONTENTS

		PAGE
ACKNOWLEDGEMENTS		iii
LIST OF FIGURES		vii
LIST OF PLATES		viii
LIST OF ABBREVIATIONS		ix
ABSTRACT		xi
ABSTRAK		xiii
CHAPTER		
I	INTRODUCTION	1
II	LITERATURE REVIEW	5
	Introduction	5
	Structure and Characteristics of Mitochondrial DNA	5
	Approaches in the Analysis of Mitochondrial DNA Variation	8
	Restriction Analysis	8
	Sequencing of Mitochondrial DNA	12
	Polymerase Chain Reaction	14
	Methods of DNA Sequence Analysis	16

III MATERIALS AND METHODS 22

Blood and DNA Samples 22

Chicken Livers Mitochondrial DNA . . . 22

Chicken Blood 23

Water Buffalo 23

Other Animals 24

Analysis of DNA Samples by Electrophoresis 24

Extraction of Blood Mitochondrial DNA . . . 25

Construction of Oligonucleotide Primers . . 25

#1. D-loop Primers UPM257 and UPM258 . 25

#2. Cytochrome b Primers KY1 and KY2 . 26

#3. D-loop Primers LCH2-UP and LCH4-DO 26

Polymerase Chain Reaction 28

Purification of PCR Products 29

By Column Chromatography 29

By Gel Electrophoresis 29

Selective Alcohol Precipitation . . . 30

DNA Sequencing 30

IV RESULTS and DISCUSSIONS 32

Polymerase Chain Reaction 32

D-loop Amplification 32

Cytochrome b Amplification 32

Purification and Sequencing of PCR Products 48

Purification of PCR Fragments by Low Melting Point Gel Electrophoresis and by Column Chromatoghaphy 48

Purification by Selective Ethanol Precipitation 51

General Discussion 54

Rationale in Method Development . . . 54

Polymerase Chain Reaction 55

DNA Sequence Fidelity 55

Sequence Quality 56

Sequence Output 57

Template Concentration 58

Limitation of Current Method and Recommendations for Future Practice . . . 59

V CONCLUSION 65

BIBLIOGRAPHY 68

APPENDICES 75

A. Protocol for Choma Spin Column Chromatography 76

B. Protocol for Thermal Cycle Sequencing with Labeled Incorporation 77

C. Partial Cytochrome b DNA Sequence of Bovine (cytb Bovine) and Chicken (Cytb Leghorn) 78

D. Chicken D-loop DNA Sequence 79

VITA 81

LIST OF FIGURES

Figure		Page
1	The Genomic Organisation of the Chicken MtDNA Compared with that of other Vertebrates	7
2	Comparison of DNA Sequencing Protocols between (1)Direct Sequencing of Denatured Double-Stranded DNA; (2) Sequencing of Single-Stranded Asymmetric PCR Product and (3) Cycle Sequencing	19
3	Location of D-loop and Cytochrome b PCR Primers	26
4	Alignment of Primer KY1 Leghorn Sequence with those of Human (Hum), Bovine (Bov), Frog (Xen) and Kocher (Koc)Universal Primer Showing Conservation at 3' Region . .	40

LIST OF PLATES

Plate		Page
1	The Results of PCR Amplification of Chicken D-loop from Liver MtDNA	33
2	Analysis of Chicken D-loop Amplification from Liver MtDNA	34
3	(a)Amplification of Chicken D-loop 3' Region from Blood by UPM 258/ LCH2-UP and (b)PCR Products from Extended 15 Cycles of Amplification	38
4	Amplification of D-loop 5'-Region from Blood by Primer UPM257/LCH4-DO	39
5	Amplification of Cytochrome b from Chicken Liver MtDNA Samples	41
6	Analysis of Amplified Cytochrome b from Buffalo	43
7	Analysis of Amplified Cytochrome b from Other Farm Animals	44
8	PCR of Cytochrome b from Chicken Blood	45
9	Sequence Autoradiogram of Column Purified (Panel 1 and 2) and Agarose Gel Purified (panel 3) Cytochrome b PCR Product	50
10	Comparison of Sequence Quality Obtained from DNA Purified with Chroma Spin-100 and Selective Precipitation	53

LIST OF ABBREVIATIONS

bp	base pair
cyt.b	cytochrome b
dNTP	deoxyribonucleotides
ddNTP	dideoxyribonucleotides
Kb	kilobase
min	minutes
mM	millimolar
ug	microgram
ul	microlitre
ng	nanogram
OD	optical density
PCR	Polymerase Chain Reaction
pmol	picomole
s	seconds
TBE	Tris-borate-EDTA buffer
TE	10 mM Tris.Cl, 1 mM EDTA
v/v	volume/volume
w/v	weight/volume

Abstract of thesis submitted to the Senate of Universiti Pertanian Malaysia in fulfilment of the requirements for the degree of Master of Science.

DEVELOPMENT OF A PCR - CYCLE SEQUENCING SYSTEM FOR ANIMAL MITOCHONDRIAL DNA ANALYSIS

By

LAU CHIN HOON

March, 1994

Chairperson: Khatijah Yusoff, Ph.D.

Faculty : Science and Environmental Studies

A system for the study of sequence polymorphisms in animal mitochondrial DNA was developed. This system consists of two candidate markers which are mitochondrial cytochrome b and D-loop. Six novel primers were constructed for their amplifications by polymerase chain reaction (PCR). The primers KY1 and KY2 amplified a 359-bp portion of the cytochrome b gene from chicken, water buffalo, horse, goat and sheep. Primers UPM257 and UPM258 were used to synthesize a 1.4 Kb-fragment containing the complete chicken mitochondrial D-loop. Primers LCH4-DO and LCH2-UP paired with primers UPM257 and UPM258 respectively to amplify a 355-bp 5' region and 395-bp 3' region of the D-loop. These primers were specific in their amplifications. The mitochondrial DNA could be amplified from

heterogeneous or degraded DNA preparations, as well as directly from blood incubated in proteinase K thus bypassing lengthy DNA isolation steps. Only 5 ul of blood was required for a 100-ul PCR.

For the purification of PCR products, ethanol selective precipitation was found to be both efficient and economical. Despite its lower recovery compared to the Chroma Spin column, sufficient quantity of DNA could be obtained for subsequent sequencing.

The sequence of the amplified double-stranded cytochrome b was effectively determined by high temperature cycle sequencing which does not require the production of single-stranded DNA or pre-reaction denaturation of template DNA and primer annealing. Clean sequence ladder was obtained using ^{35}S dATP incorporation.

The system developed in this research is simple, with the gene amplification and sequencing being handled by a thermal cycler. The semi-automatic nature of this system would make the production of data on mtDNA polymorphisms faster with minimum operation errors. These features are particularly useful for analyzing large numbers of samples in population studies. Moreover, a cycle-sequencing based method will undoubtedly dominate sequence analysis work in molecular biology.

Abstrak tesis yang dikemukakan kepada Senat Universiti Pertanian Malaysia bagi memenuhi keperluan untuk Ijazah Master Sains.

PERKEMBANGAN SATU SISTEM PCR - PENJUJUKAN KITARAN UNTUK ANALISIS DNA MITOKONDRIA BINATANG

Oleh

LAU CHIN HOON

March, 1994

Pengerusi: Khatijah Yusoff, Ph.D.

Fakulti : Sains dan Pengajian Alam Sekitar

Satu sistem untuk mengkaji polimorfisma jujukan dalam DNA mitokondria (mtDNA) telah dimajukan. Sistem ini mengandungi dua calon penanda iaitu sitokrom b dan gelung-D mitokondria. Enam primer baru telah direkabina dalam kajian ini bagi tindak balas rantaian polimerase ('polymerase chain reaction', PCR) gen itu. Primer KY1 dan KY2 digunakan dalam amplifikasi sebahagian (359 pasangan bes) gen sitokrom b daripada ayam, kerbau, kuda, kambing dan biri-biri. Primer UPM257 dan UPM258 digunakan untuk amplifikasi serpihan 1.4-kilobes yang mengandungi gelung-D lengkap mitokondria ayam. Primer LCH4-DO dan LCH2-UP berpasangan dengan primer UPM257 dan UPM258 masing-masing untuk melipatgandakan bahagian 5' (355 pasangan bes) dan 3' (395 pasangan bes) dari gelung-D. Kesemua primer itu adalah spesifik dalam amplifikasi. DNA mitokondria boleh

diamplifikasi daripada penyediaan DNA yang tidak tulen ataupun sempurna, di samping daripada darah yang dieramkan dalam proteinase k secara terus dan ini dapat mengelakkan langkah-langkah dalam pengasingan DNA. Hanya 5 ul darah diperlukan bagi setiap 100-ul PCR.

Untuk penulenan produk PCR, pemendakan pilihan etanol telah dikenal pasti sebagai cara yang berkesan dan ekonomik. Walaupun pemulihannya adalah lebih rendah daripada turus Chroma Spin, kuantiti DNA yang cukup boleh diperolehi untuk penjujukan selanjutnya.

Penjujukan bebenang-dua sitokrom b telah dijalankan dengan penjujukan kitaran ('cycle-sequencing') suhu tinggi yang tidak memerlukan penghasilan DNA bebenang-tunggal dan pra-tindak balas penguraian templat DNA dan pelekapan primer. Tangga jujukan yang bersih dapat diperolehi daripada penjujukan sitokrom b yang menggunakan penyatuan ^{35}S dATP.

Kaedah yang dimajukan ini adalah mudah, dengan amplifikasi dan penjujukan gen disediakan oleh pemutar haba ('thermal cycler'). Ciri separa automatik ini menjadikan penghasilan data polimorfisme mtDNA lebih cepat dengan kesilapan yang minimum. Semua ini amatlah berguna untuk analisis bilangan sampel yang banyak. Selain daripada

ini, kaedah yang berasaskan penjujukan kitaran tanpa keraguannya akan mengaruh analisis jujukan dalam kebanyakan bidang penyelidikan biologi molekul.

CHAPTER I

INTRODUCTION

A research programme entitled 'Genetic Studies of Domestic Animals' has been started at the Universiti Pertanian Malaysia. The first phase of work in the project was to identify a suitable approach to the study of DNA polymorphisms. As this line of research was new in Malaysia, the logical starting point of the research was the development of research methodology which consisted of two components. The first component was the selection of candidate gene loci that could function as markers for probing genetic differences. The second, which was the development of a technique appropriate for detecting sequence variations at these loci, constitutes the subject matter of this thesis. This Chapter gives a review on the evolution of the research direction which later led to the adoption of the current approach in genetic analysis.

The initial objective was to perform Restriction Fragment Length Polymorphisms (RFLP) analysis of poultry growth hormone and prolactin genes. The sources of both the bovine growth hormone (Gordon et al., 1983; Woychik et al., 1982) and prolactin (Cooke and Baxter, 1981) genomic and cDNA probe were located. However, two

possible pitfalls of applying these probes in the hybridization experiments were anticipated. Firstly, use of the heterologous probes (from bovine) to detect poultry DNA necessitated the lowering of hybridization stringency and this would reduce target specificity. Secondly, cross hybridization of growth hormone to chorionic somatomammotropin in human had been reported. A similar event (i.e. cross hybridization within the prolactin family of genes and with other growth hormone-like genes) could take place in poultry RFLP analysis. In both cases, there is the possibility of picking up false positives which could lead to erroneous interpretation of experimental data. Knowledge on the gene structure of these hormones and the level of local technical know-how at that time was considered to be insufficient for quality assurance in this work.

Informative restriction enzymes and ethnic specific mitochondrial DNA (mtDNA) types have been identified in human. It is thought that similar findings could be obtained in domestic animals; this being the alternative approach for detecting genetic polymorphisms. For this purpose, the research then concentrated on developing techniques for isolation of mtDNA from chicken livers and purification of the mtDNA by caesium chloride density gradient ultracentrifugation.

As the understanding and insight for the population genetics of mtDNA grew deeper, the need for an alternative strategy was gradually felt. Some information and exposure had been instrumental in the shift of research strategy away from the RFLP-based method. Firstly, using the mtDNA-RFLP method, Wakana et al (1986) found only one restriction variant from 16 chicken varieties examined. Such a low level of polymorphism was indeed unexpected for mtDNA. It is only by DNA sequencing that a more thorough screening of the genome could be made to resolve whether this low level of variation was due to a technical drawback of restriction analysis or to a result of an evolutionary bottleneck. Secondly, A PCR-based method would be a more superior way of studying genetic variation. Furthermore, the complete chicken mitochondrial sequence became available and this made possible the construction of PCR oligonucleotide primers. The then arrival of a thermal cycler at the laboratory catalyzed the adoption of PCR-sequencing strategy in the research, which formed the core of this thesis.

In analyzing mitochondrial DNA polymorphisms, DNA sequencing undoubtedly yields the finest level of resolution and suits most of the problems. However, sequencing of large numbers of samples needed for population studies can be tiresome and time-consuming. The objective of this thesis research was to alleviate these problems by improving DNA template generation and the reaction process.

To achieve this end, two current technological advances - the polymerase chain reaction (PCR) coupled to cycle or linear amplification sequencing were utilised. This research, therefore, comprised the following:

1. Design of suitable PCR and sequencing primers;
2. selection of a DNA extraction method;
3. development of a PCR condition for mtDNA amplification
4. identification of a suitable method for purifying amplified product, and
5. development of high temperature sequencing. The mitochondrial cytochrome b and D-loop were selected as markers for future studies of genetic variation in local chicken, water buffaloes and other domestic animals. Naturally, the DNA fragments from these loci were used as templates in the development of PCR and cycle sequencing.

It is anticipated that the technical foundation built up in this present work will be capable of carrying the overall research into the next phase which is the collection of data on mitochondrial DNA polymorphisms. Furthermore, the simpler, semi-automated PCR-sequencing method developed in this project would allow more person-time to be re-allocated from mechanical data generation into consolidating research planning, data analysis and application and exploration of new research area which would improve the management of local animal genetic resources.

CHAPTER II

LITERATURE REVIEW

Introduction

Animal mitochondrial DNA (mtDNA) had been studied extensively to address problems in population genetics and systematics of closely related species and individuals within a species. It had been applied to the study of matrilineal relationships in a population, the history of speciation, and geographic population structure. These studies had been thoroughly reviewed by Avise *et al.* (1986; 1987) and Moritz *et al.*(1987). In animal breeding, mtDNA has been used as a molecular marker in the study of productive traits and the degree of genetic variability (Hecht, 1990). This chapter would give an overview of the major techniques in population genetic analysis of mtDNA which was considered in the course of methodology selection. The emphasis would be on the PCR and DNA sequencing technology.

Structure and Characteristics of Mitochondrial DNA

Animal mitochondrial DNA (mtDNA) is a covalently closed circular molecule (Wallace, 1986). The complete genome in many vertebrates had been sequenced. It contains genes for 13 protein

subunits functioning in ATP production, 2 ribosomal RNAs (12s and 16s), 22 transfer RNAs and a non-coding region known as displacement-loop (D-loop) (it is common that the 'D-loop' and the 'non-coding region' are used interchangeably). The gene order is conserved in human (Anderson *et al.*, 1981), mouse (Bibb *et al.*, 1981) and bovine (Anderson *et al.*, 1982). Some rearrangements were present in chicken, quail, Guinea fowl, pheasant and turkey (Desjardins and Morais, 1990), in which the contiguous $tRNA^{Glu}$ and ND6 genes are located adjacent to the D-loop (Figure 1).

There are unique features which make mtDNA an ideal molecular system for population analysis. Firstly, its ubiquity enables comparison of data between species. Secondly, the complication of genetic recombination or rearrangement is bypassed due to its maternal inheritance. In mice, only 10^{-4} of the mtDNA is paternally inherited relative to the maternal contribution (Gyllensten *et al.*, 1991). Thirdly, the mtDNA has a base substitution rate of 1% per 10^6 years (Brown *et al.*, 1979) which is 1 to 10 times faster than in single-copy nuclear DNA (Avise *et al.*, 1987). This results in considerable sequence variation between individuals. Fourthly, the number of copies of mtDNA molecules present was estimated to be 1000 to 8000 per fibroblast (Reis and Goldstein, 1983), and 2.6×10^5 per oocyte (Michaels *et al.*, 1982) in bovine. This high copy number, along with their compartmentalized location separate from nuclear DNA, have made mtDNA relatively simple to isolate and assay.

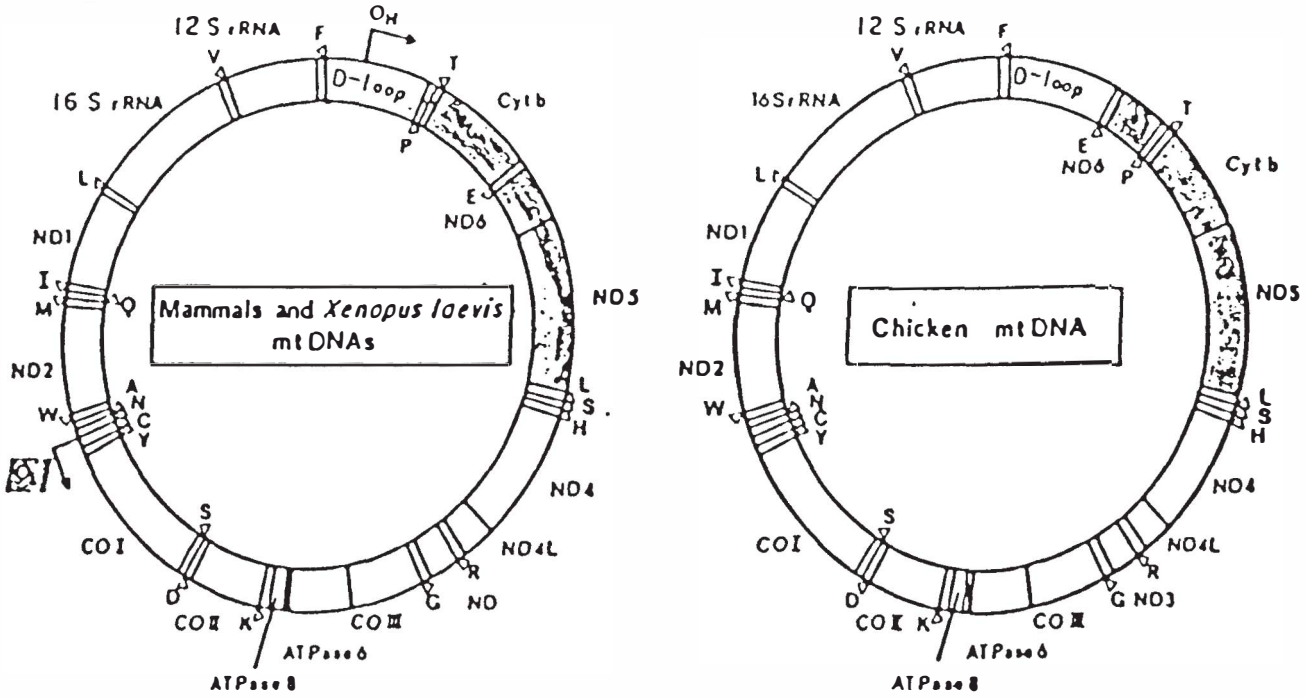

Figure 1. The Genomic Organisation of the Chicken MtDNA Compared with that of other Vertebrates (Desjardins and Morais, 1990).

Approaches in the Analysis of Mitochondrial DNA Variation

Mitochondrial DNA had been mainly studied as a whole by restriction analysis and in part by DNA sequencing. The choice of either method depends on the level of resolution required in solving the issue(s) at hand, the amount of DNA that can be extracted from the animals or biological specimens and budget expediency. Between the two approaches, the restriction analysis method is more commonly used mainly because it is technically less demanding and less costly compared to DNA sequencing.

Restriction Analysis

In the method involving restriction analysis (also known as 'restriction fragment length polymorphism', RFLP), the sequence variation is revealed by the number and size of DNA fragments produced by restriction endonuclease digestion. The variations arise from base substitution, addition or deletion. In order to be detected and become informative, the sequence variation must be located in the recognition sequence or within two cleavage sites (in the case of length mutation) of the restriction enzyme being used in the study. The restricted fragments are separated by gel electrophoresis and can be detected by blot hybridization, radioactive end-labelling or fluorescence using ethidium bromide.

In the blot hybridization method (Southern, 1975), the restricted DNA is transferred onto a membrane after electrophoresis. The immobilized DNA is then detected by a labelled probe prepared from purified mtDNA. In this method, 2-7 ug of total cellular DNA extracted from blood (Ferris *et al.*, 1981) or about 0.2 ug of platelet DNA (Denaro *et al.*, 1981) is required. Since the fragment labelling is mtDNA probe-directed, the samples need not be homogeneous. The paper of Johnson *et al.* (1983) studying divergence of human ethnic groups is an elegant classic example of applying this technique: restriction endonuclease morphs (banding patterns) were sorted into mtDNA types, and subsequently genetic distances, time of divergence and evolutionary rates were calculated. In Malaysia, the blot hybridization approach had been applied to a preliminary study of buffalo restriction cleavage patterns (Gan *et al.*, 1991).

In contrast to the tolerance to sample impurities in the hybridization method, the other two methods of detecting restriction fragments (Brown, 1980; Cann and Wilson, 1983) requires that the mtDNA be intact and of high purity. These criteria are meant to prevent false positive and high background which complicate band scoring. Choice of either one of the identification methods depends on the amount of mtDNA available. They both normally require 1 or 2 rounds of caesium chloride density gradient centrifugation. The

radioactive based methods are sensitive in that picogram quantity of DNA can be detected and restriction analysis can be carried out in a situation where the amount of DNA is limiting. The ethidium bromide staining method, on the other hand, requires at least 3-5 ng of DNA per band for detection, and as much as 8 ug of mtDNA is required for *Hae*III digestion (Horai and Matsunaga, 1986). However, due to the requirement for special containment and the health hazard, the radioisotope based methods are not always favourable. In addition, the blot hybridization has the added disadvantage of requiring too many steps and too many days for DNA detection.

Abundant DNA had been isolated from liver or brain in cow (Hauswirth *et al.*, 1984), heart or pectoral muscle from birds (Ball *et al.*, 1988) and whole bodies of fish (Bermingham *et al.*, 1991). In addition, 200 to 300 ug closed circular mtDNA had been recovered from individual placenta (Horai *et al.*, 1984). The availability of mtDNA isolated from some of these sources makes feasible the use of the direct fluorescent detection method which is fast, inexpensive, and justifies the labour and expense in DNA extraction and purification. As the experimental procedure is simple, confirmation or duplication of results or trial of new restriction enzymes can be easily done long after the original experiment. Use of the other methods require more effort in setting-up and material preparation.