

**FORENSIC IDENTIFICATION OF  
AVIAN SPECIES USING  
MITOCHONDRIAL LOCI**

By

**Sansook Boonseub**

A thesis in fulfilment of the requirements  
for the degree of

**Doctor of Philosophy (Forensic Genetics)**

School of Biological Sciences  
Flinders University

October 2011

## **CERTIFICATE OF AUTHORSHIP OF THESIS**

I further certify that to the best of my knowledge, the thesis contains no material previously published or written by another person except where due reference is made in the text of the thesis.

The material in the thesis has not been the basis of an award or any other degree or diploma except where due reference is made in the text of the thesis.

**Sansook boonseub**

October 2011

## **ACKNOWLEDGEMENTS**

Firstly, I would like to thank you Prof. Dr. Adrian Linacre, my supervisor who always gave me advice and encouragement at all time of my study. I would never be successful in my study without his support.

I would like to thank the Royal Thai Police and the Royal Thai Government who gave me full scholarship for doing my Ph.D. I could not be here to study in the UK and Australia without the funding support from Thailand.

I would like to thank all my colleagues here at the forensic laboratory (Prof. Linacre's lab) in Flinders University, for willing to help whenever I needed any assistance for anything.

Lastly, I would like to thank my parents and my brother who are always beside me when I feel sad, depressed or experience any trouble. They are the first ones who are always ready to hold out their hands for help, unconditionally. Thank you all very much.

## ABSTRACT

Mitochondrial DNA (mtDNA) loci are used routinely for species testing in mammalian species. This project examines their use in avian identification for forensic purposes, particularly as certain avian species are the subject of an illegal trade. Comparison of amino acid sequences for a range of mitochondrial genes taken from avian species indicated that along with two commonly used gene loci (cyt *b* and COI), two members of the NADH dehydrogenase family (ND2 and ND5) showed greater variation.

Alignment of entire avian mitochondrial sequences illustrated the potential for using such large sequences, but comparison of individual genes using 102 avian species further supported ND2 and ND5 as having greater interspecies variation and less intra-specific variation. Support for these two loci was further provided when looking at closely related avian species. As each gene is relatively large, sections of these loci were selected for ability to identify and distinguish closely related avian species and reconstruct accurate phylogenetic trees. A 452 bp section at the 5' terminus from both ND2 and ND5, at base positions 58-509 and 101-552 respectively, were superior compared to the other mitochondrial loci in species identification, including closely related members of the Fringillidae, Psittacidae and Cacatuidae families and in accurate phylogenetic tree reconstruction.

DNA extracts were obtained from individual barbs and calamus of feathers. Samples included both fresh from known species, feathers from unknown species collected in Adelaide, and archived the museum samples. Successful amplification of sections of the ND2 (561 bp) and ND5 genes (921 bp) was achieved. Sequence data comparison of the PCR products confirmed accurate species identification.

# TABLE OF CONTENTS

CERTIFICATE OF AUTHORSHIP OF THESIS .....	ii
ACKNOWLEDGEMENTS .....	iii
ABSTRACT .....	iv
TABLE OF CONTENTS .....	vi
LIST OF FIGURES .....	viii
LIST OF TABLES .....	xxi
LIST OF ABBREVIATION .....	xxv
CHAPTER 1 INTRODUCTION.....	1-58
General Introduction.....	1
Types of wildlife crime .....	2
Legislation covering wildlife protection .....	2
Background to Forensic Science .....	5
Avian taxonomy .....	15
Speciation and Species concept.....	17
Avian mitochondrial DNA (mtDNA).....	21
Generation of Genetic Variation .....	28
Molecular clocks .....	30
Mitochondrial DNA Barcoding genes.....	32
Phylogenetic tree .....	33
The Aim of this Study .....	37
Chapter 1 References.....	38
CHAPTER 2 MATERIALS AND METHODS .....	59-100
Sequence Data ... ..	59
Mitochondrial protein sequence analysis .....	68
Primer design.....	68
Sample collection.....	75
DNA extraction .....	83
DNA amplification.....	87
Gel Electrophoresis .....	89
PCR product purification by using Gel extraction kit.....	90
DNA Sequencing.....	91

Avian mitochondrial DNA analysis .....	92
Chapter 2 References.....	99
CHAPTER 3 RESULTS AND DISCUSSIONS .....	101-185
Mitochondrial protein sequences analysis.....	101
Mitochondrial DNA sequences analysis .....	121
Species identification and phylogenetic tree reconstruction in the members of the same genera .....	136
Methods and models comparison for phylogenetic tree reconstruction of the parrot and cockatoo families .....	165
Inter- and intra-species variations in closely relative species .....	170
Chapter 3 References.....	181
CHAPTER 4 IDENTIFICATION OF PROTECTED AVIAN SPECIES USING A SINGLE BARB .....	186-199
Chapter 4 References.....	199
CHAPTER 5 CONCLUSION .....	200-211
Summary .....	200
Applications.....	202
Validation .....	204
Dissemination.....	205
The technologies for species identification.....	208
Future developments .....	210
Chapter 5 References.....	211
APPENDICES .....	212-438
APPENDIX A .....	212-217
APPENDIX B.....	218-225
APPENDIX C.....	226-346
APPENDIX D .....	347-416
APPENDIX E.....	417-438

## LIST OF FIGURES

**Figure 1.1a:** shows the first ever DNA fingerprint on the left, produced in Alec Jeffreys laboratory on the 10<sup>th</sup> September 1984. It shows the banding pattern from a mother, child, and father using a multilocus probe. The image on the right shows the multilocus fingerprint from a mother, a tested child (TC), as the UK Immigration Officials did not believe that the boy was the biological son of the mother, and three of the mother's undisputed children (UD 1-3): the results demonstrated that the tested child was indeed the biological son of the woman and was therefore allowed to stay in the UK. The image on the left was supplied by Adrian Linacre and that on the right from reference [66].....7

**Figure 1.1b:** Figure 1.1b: Schematic representation of RFLP analysis; genomic DNA from sample 1 (S1) and sample 2 (S2) were cut by restriction enzymes which generated different length of restriction products. They were separated on agarose gel and transferred to membrane. Visualization was performed using hybridization with labelled probes then exposed to x-ray film. There are 2 types of RFLP profiling; Multi-locus (MLPs) profiling (using multi-locus probes) and Single locus (SLPs) profiling (using single locus probes).  
.....8

**Figure 1.2:** Schematic representation of RAPD analysis; genomic DNA from sample 1 (S1) and sample 2 (S2) were amplified using arbitrary single short primers. The primers are illustrated as blue arrows. These are typically 10 nucleotides in length. The different lengths of amplicons were separated by agarose gel electrophoresis (2a). The example species identification using RAPD profiling shows in (2b), the figure was taken from James Lee's thesis [66]. Lane L is 100 bp ladder. Lane 1 to 10 are from genus *Paneolus* which 1-4 are from the same species and lane 11-20 are from genus *Psilocybe* which 11 and 12 are from the same species and lane 19 and 20 are from the same species. The result shows that the sample from the same species shows very similar RAPD profile but none of them show 100% match therefore impossible to reproduce the same RAPD profile from this technique even the sample was taken from 1 individual.....11

**Figure 1.3:** Schematic representation of AFLP analysis; genomic DNA from sample 1 (S1) and sample 2 (S2) were double-digested with *EcoRI* and *MseI*. These generated different length of restriction products with 'sticky ends' generated by the enzymes to which adapters could be ligated; with these adapters being of a known sequence and can act as a template in PCR. Pre-amplification, single nucleotide selective amplification was performed using one base extending into the unknown sequence of interest; this reduces the number of PCR products produced. Example species identification using AFLP profiling shows in (b), the figure was taken from James Lee's thesis [66]. Lane M is 100 bp ladder. Lanes 1 to 17 are unknown samples. Lane 1-3 are from the same species, 10-12 are from the same species and 18-20 are reference samples from *Volvariella volvacea*, *Agaricus bisporus* and *Lentinus edodes*, respectively. The result shows that the sample from the same species shows very similar AFLP profile but none of them show 100% match



therefore impossible to reproduce the same AFLP profile from this technique even the sample was taken from 1 individual.....12

**Figure 1.4:** Avian taxonomy [139, 140] .....16

**Figure 1.5:** Organization of the mammalian and avian mitochondrial genomes. The D loop is normally viewed at the top with the numbered bases going clockwise. The order of the 37 genes is similar between mammalian and avian mtDNA but note that *cyt b* and ND 6 are transversed. ....27

**Figure 1.6:** Figure 1.6: Showing rooted tree (a) and unrooted tree (b).....33

**Figure 1.7:** Showing parts of a phylogenetic tree. The red arrows indicate the parts of the tree. Taxon A and B are sister groups. Taxon C is the outgroup of taxon A and B. The blue box is a clade. Root is a common ancestor of all taxa on the tree. A clade is a group of taxa and their common ancestor. Internal nodes or hypothetical taxonomic units (HTUs) represent the hypothetical ancestor. External nodes or operational taxonomic units (OTUs) are the set of data under the comparison study. Branch represents the relationship between the taxa. Branch length shows the number of changes within the branch. Sister groups represent closest relatives. Outgroup represents a taxon outside the group of study.....34

**Figure 2.1:** A map of the overlapping sites of the expected PCR products generated from the ND2 primers.....71

**Figure 2.2:** A map of the overlapping sites of the expected PCR products generated from the ND5 primers.. ....73

**Figure 2.3:** A map of the overlapping sites of the expected PCR products from the ND5 primers for the complete ND5 gene sequencing of the parrot and cockatoo species.....75

**Figure 2.4:** showing Google map positions of the River Torrens, Victoria Square and the Westfield Marion Shopping Centre, Adelaide, South Australia.....82

**Figure 2.5:** showing the feathers from Major Mitchell's Cockatoo, the blue arrow indicates the calamus part of the feather.....86

**Figure 2.6:** showing feather and barb samples from the smaller size bird: superb parrot (*Polytelis swainsonii*) and the larger size bird: short-billed black cockatoo (*Calyptorhynchus latirostris*). The red arrows indicate a barb from each species.....87

**Figure 2.7:** showing p-distance calculation using MEGA program. The orange arrows indicate the three steps to compute the overall mean value.. .....98

**Figure 3.1:** Protein alignment of COI gene in 33 different avian species using MEGA 4 program. The different colours are used to indicate the different groups of amino acid based on their side-chain properties; where polar is green, non-polar is red, acidic is blue and basic is grey. ....107

**Figure 3.2:** Protein alignment of *cyt b* gene in 33 different avian species using MEGA 4 program. The different colours are used to indicate the different groups of amino acid based on their side-chain properties; where polar is green, non-polar is red, acidic is blue and basic is grey ..... 109

**Figure 3.3:** Protein alignment of ND2 gene in 33 different avian species using MEGA 4 program. The different colours are used to indicate the different groups of amino acid based on their side-chain properties; where polar is green, non-polar is red, acidic is blue and basic is grey ..... 111

**Figure 3.4:** Protein alignment of ND5 gene in 33 different avian species using MEGA 4 program. The different colours are used to indicate the different groups of amino acid based on their side-chain properties; where polar is green, non-polar is red, acidic is blue and basic is grey ..... 115

**Figure 3.5:** The partial sequences of the COI, *cyt b*, ND2 and ND5 that have been used for avian species identification and evolutionary studies. The numbers indicate the base positions according to the mtDNA sequence of *Gallus gallus* (AY235571). Single letters indicate the tRNA gene for each amino acid: H is tRNA-His, S is tRNA-Ser, L is tRNA-Leu, T is tRNA-Thr, P is tRNA-Pro, I is tRNA-Ile, Q is tRNA-Gln, M is tRNA-Met, W is tRNA-Trp, A is tRNA-Ala, N is tRNA-Asn, C is tRNA-Cys, Y is tRNA-Tyr, and D is tRNA-Asp.....117

**Figure 3.6:** Genetic variation at the protein level of the COI (a), *cyt b* (b), ND2 (c) and ND5 (d). The X-axis represents an amino acid position along the four loci and the Y-axis represents the percentage of homology between amino acid of each locus based on 33 avian species, 33 genera, 29 Families and 13 Orders.....118

**Figure 3.7:** Diagram for the human cyt *b* peptide (NC\_012920) [42]. The red arrows indicate the parts of the peptide that fold outside an inner and outer mitochondrial membrane which are expected exhibit greater variation than the conserved membrane bound amino acids. The numbers shows the position of the amino acids based on the human sequence.....120

**Figure 3.8:** Phylogenetic tree reconstruction of the complete mitochondrial genome sequences. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analysed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 102 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 9166 positions in the final dataset. The Galliformes, Anseriformes, Struthioniformes, Tinamiformes, Falconiformes and Passeriformes are clustered together as expected if the genetic data of the whole mitochondrial genome matches the current taxonomic groups at the taxonomic level of Order. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....122

**Figure 3.9:** The percentage of nucleotide similarity for each of the thirteen avian mitochondrial genes, two rRNA genes and the 22 tRNA gene sequences based on 102 avian species. The similarity is provided next to each locus. The size of each gene is provided. The tRNA loci are coloured in yellow (except the tRNA-Glu is coloured in pink), loci with similarities of between 81 – 90 % are coloured in green, those loci coloured in pink have similarities of between 71 – 80%, and the D-loop is coloured in blue to note the nucleotide similarity of 64.6%.....124

**Figure 3.10:** Phylogenetic tree reconstruction of the complete gene sequences of the COI gene. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 102 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1238 positions in the final dataset. The Galliformes and Gruiformes are clustered together as expected if the genetic data of the complete gene sequences of the COI gene matches the current taxonomic groups at the taxonomic level of Order. The green

dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct.....127

**Figure 3.11:** Phylogenetic tree reconstruction of the complete gene sequences of the *cyt b* gene. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 102 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 770 positions in the final dataset. The Galliformes is clustered together as expected if the genetic data of the complete gene sequences of the *cyt b* gene matches the current taxonomic groups at the taxonomic level of Order. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....128

**Figure 3.12:** Phylogenetic tree reconstruction of the complete gene sequences of the ND2 gene. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 102 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 460 positions in the final dataset. All Orders are not clustered together as expected if the genetic data of the complete gene sequences of the ND2 gene. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....129

**Figure 3.13:** Phylogenetic tree reconstruction of the complete gene sequences of the ND4 gene. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 102

nucleotide sequences. Codon positions included were 1st, +2nd, +3rd, +Non-coding. All positions containing gaps and missing data were eliminated. There were a total of 652 positions in the final dataset. The Galliformes are clustered together as expected if the genetic data of the complete gene sequences of the ND4 gene matches the current taxonomic groups at the taxonomic level of Order. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....130

**Figure 3.14:** Phylogenetic tree reconstruction of the complete gene sequences of the ND5 gene. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 102 nucleotide sequences. Codon positions included were 1st, +2nd, +3rd, +Non-coding. All positions containing gaps and missing data were eliminated. There were a total of 908 positions in the final dataset. The Galliformes are clustered together as expected if the genetic data of the complete gene sequences of the ND5 gene matches the current taxonomic groups at the taxonomic level of Order. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....131

**Figure 3.15:** The gene loci chosen were divided into 100 bp and 450 bp segments from which the sequences were extracted and used in alignments using the MEGA 5.....132

**Figure 3.16a:** Phylogenetic tree reconstruction of the 450 bp segment from the ND2 gene at base positions 1-450. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history of the 1-450 ND2 fragment was inferred using the Neighbor-Joining method . The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 104 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 450 positions in the final dataset. The Galliformes, Anseriformes, Struthioniformes and Tinamiformes are clustered together as expected if the genetic data of the complete gene sequences of the ND2 gene matches the current taxonomic groups at the taxonomic level of Order. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....133

**Figure 3.16b:** Phylogenetic tree reconstruction of the 450 bp segment from the ND5 gene at base positions 101-550 bp. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history of the 101-550 bp ND5 fragment was inferred using the Neighbor-Joining method . The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 102 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 450 positions in the final dataset. The Galliformes and Anseriformes are clustered together as expected if the genetic data of the complete gene sequences of the ND5 gene matches the current taxonomic groups at the taxonomic level of Order. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....134

**Figure 3.17:** Higher avian taxonomy including subclass Neognathae and subclass Paleognathae [57]. All avian species were split into a group that includes all the flightless birds (emus and ostriches etc) and all other birds that evolved flight. The gene sequences of both the ND2 and ND5 loci were able to split the avian species at this higher taxonomic level as well as at lower taxonomic levels. In this study, the 19 Orders were classified as show in this diagram.....135

**Figure 3.18:** Phylogenetic tree reconstruction of the partial sequences at 5' terminus (57-508, 452 bp) of the ND2 gene of 10 different species from *Carduelis* and *Fringilla* genera. The out group was a moa. Evolutionary analyses were conducted in MEGA5 [43].The evolutionary history was inferred using the Neighbor-Joining method [44] . The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 44 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 452 positions in the final dataset. The two genera are clustered together as expected if the genetic data of the 452 bp at 5' end of the ND2 gene matches the current taxonomic groups at the taxonomic level of genus. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....145

**Figure 3.19:** Phylogenetic tree reconstruction of the partial sequences at 5' terminus (136-639, 504 bp) of the *cyt b* gene from 10 different species. The out group was a moa. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was

inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 39 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 503 positions in the final dataset. The two genera are clustered together as expected if the genetic data of the 504 bp at 5' end of the *cyt b* gene matches the current taxonomic groups at the taxonomic level of genus. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....146

**Figure 3.20:** Phylogenetic tree reconstruction of the partial sequences at 5' terminus (98-708, 611 bp) of the COI gene from 10 different species. The out group was a moa. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 81 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 610 positions in the final dataset. The two genera are clustered together as expected if the genetic data of the 611 bp at 5' end of the COI gene matches the current taxonomic groups at the taxonomic level of genus. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....147

**Figure 3.21:** Phylogenetic tree reconstruction of the partial sequences at 5' terminus (58-509, 452 bp) of the ND2 gene of 14 different species from Psittacidae and Cakatuidae families. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 15 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 449 positions in the final dataset. The two Families are clustered together as expected if the genetic data of the 452 bp at 5' end of the ND2 gene matches the current taxonomic groups

at the taxonomic level of family. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....160

**Figure 3.22:** Phylogenetic tree reconstruction of the partial sequences at 5' terminus (101-552, 452 bp) of the ND5 gene of 14 different species from Psittacidae and Cacatuidae families. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 15 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 448 positions in the final dataset. The two Families are not clustered together as expected if the genetic data of the 452 bp at 5' end of the ND5 gene matches the current taxonomic groups at the taxonomic level of family. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....161

**Figure 3.23:** Phylogenetic tree reconstruction of the partial sequences at 5' terminus (39-760, 722bp) of the COI gene of 5 different species. According to there is no sequences at 5' end of this gene from Psittacidae family submitted on the database so that the tree was reconstruct from Cacatuidae family. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 6 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 722 positions in the final dataset. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....162

**Figure 3.24:** Phylogenetic tree reconstruction of the partial sequences at 3' terminus (741-1272, 532 bp) of the COI gene of 10 different species from Psittacidae and Cacatuidae families. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap



replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 10 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 523 positions in the final dataset. The two Families are clustered together as expected if the genetic data of the 532 bp at 5' end of the COI gene matches the current taxonomic groups at the taxonomic level of family. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....163

**Figure 3.25:** Phylogenetic tree reconstruction of the partial sequences at 3' terminus (686-860, 175 bp) of the *cyt b* gene of 11 different species from Psittacidae and Cacatuidae families. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 12 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 175 positions in the final dataset. The two Families are clustered together as expected if the genetic data of the 175 bp at 5' end of the *cyt b* gene matches the current taxonomic groups at the taxonomic level of family. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....164

**Figure 3.26:** The phylogenetic tree was reconstructed using partial sequences at 5' terminus (57-508, 452 bp) of the ND5 gene of 15 species of parrots and 8 species of cockatoos. The out group was a moa. Evolutionary analyses were conducted in MEGA5 [43]. The evolutionary history was inferred using the Neighbor-Joining method [44]. The bootstrap consensus tree inferred from 1000 replicates [45] is taken to represent the evolutionary history of the taxa analyzed [45]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [45]. The evolutionary distances were computed using the Kimura 2-parameter method [46] and are in the units of the number of base substitutions per site. The analysis involved 43 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 452 positions in the final dataset. The green dots indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....168

**Figure 3.27:** The phylogenetic trees generated using a range of parrot and cockatoo species using the partial sequences of the ND5 gene at 5' terminus created from the MrBayes program. The out-group was a moa. The pink branches indicate the bootstrap value for a given interior branch is 95% or higher, then the topology at that branch is considered correct [47].....169

**Figure 3.28:** Inter- and intra-species of the Fringillidae family (finches) at genus and species taxonomic level using partial uences of the ND2 gene at base positions 58-509, 452 bp. The numbers in red indicates the numbers of sequences used in this study.**Figure 3.29:** Inter- and intra-species of the Fringillidae family (finches) at genus and species taxonomic level using partial sequences of the *cyt b* gene at base positions 140-620, 481 bp. The numbers in red indicates the numbers of sequences used in this study.**Figure 3.30:** Inter- and intra-species of the Fringillidae family (finches) at genus and species taxonomic level using partial sequences of the COI gene at base positions 81-705, 625 bp. The numbers in red indicates the numbers of sequences were used in this study.**Figure 3.31:** Inter- and intra-species of the Psittacidae family (parrots) and Cacatuidae family (cockatoos) at genus and species taxonomic level using partial sequences of the ND2 gene at base positions 58-509, 452 bp. The numbers in red indicates the numbers of sequences used in this study.....171

**Figure 3.29:** Inter- and intra-species of the Psittacidae family (parrots) and Cacatuidae family (cockatoos) at genus and species taxonomic level using partial sequences of the ND5 gene at base positions 101-552, 452 bp. The numbers in red indicates the numbers of sequences used in this study.....172

**Figure 3.30:** Inter- and intra-species of the Psittacidae family (parrots) and Cacatuidae family (cockatoos) at genus and species taxonomic level using partial sequences of the COI gene at base positions 81-702, 622 bp. The numbers in red indicates the numbers of sequences used in this study.....173

**Figure 3.31:** Inter- and intra-species of the Psittacidae family (parrots) and Cacatuidae family (cockatoos) at genus and species taxonomic level using partial sequences of the COI gene at base positions 751-1256, 506 bp. The numbers in red indicates the numbers of sequences used in this study.....174

**Figure 3.32:** Inter- and intra-species of the Psittacidae family (parrots) and Cacatuidae family (cockatoos) at genus and species taxonomic level using partial sequences of the *cyt b* gene at base positions 300-751, 452 bp. The numbers in red indicates the numbers of sequences used in this study.....175

**Figure 3.33:** Inter- and intra-species of the Psittacidae family (parrots) and Cacatuidae family (cockatoos) at genus and species taxonomic level using partial sequences of the COI

gene at base positions 81-702, 622 bp. The numbers in red indicates the numbers of sequences used in this study.....176

**Figure 3.34:** Inter- and intra-species of the Psittacidae family (parrots) and Cacatuidae family (cockatoos) at genus and species taxonomic level using partial sequences of the COI gene at base positions 751-1256, 506 bp. The numbers in red indicates the numbers of sequences used in this study.....177

**Figure 3.35:** Inter- and intra-species of the Psittacidae family (parrots) and Cacatuidae family (cockatoos) at genus and species taxonomic level using partial sequences of the cyt *b* gene at base positions 300-751, 452 bp. The numbers in red indicates the numbers of sequences used in this study.....178

**Figure 4.1:** showing the calamus section of the bird feather.....188

**Figure 4.2:** showing amplification of sections of the avian mitochondrial ND2 and ND5 genes from 2 barbs and 5 barbs of the Short-billed Black Cockatoo (*Calyptorhynchus latirostris*) under various annealing temperature of 50°C, 52°C, 54°C, 57°C and 60°C. The amplification from the ND2 gene producing a product of 561 bp is in Figure 4.2(a) and M is a 100 bp marker, N is a negative control, P is a positive control from chicken muscle. The amplification from the ND5 gene producing a product of 921 bp is in Figure 4.2(b) and M is a 1 kb marker, N is a negative control, P is a positive control from chicken muscle.....189

**Figure 4.3:** showing an example of a single barb and the feather from which it was taken. The larger feather and barb (bottom) is from the Short-billed Black Cockatoo (*Calyptorhynchus latirostris*), the Regent Parrot (*Polytelis anthopeplus*) is shown top right and cockatiel (*Nymphicus hollandicus*) is shown top left.....190

**Figure 4.4:** showing amplification of sections of the avian mitochondrial ND2 and ND5 genes from the Short-billed Black Cockatoo (*Calyptorhynchus latirostris*). Lanes 1-8 are amplification from the ND2 gene producing a product of 561 bp and are, left to right, 100 bp marker, negative control, positive control from chicken muscle, 1 barb, 2 barbs, 5 barbs, 10 barbs and 20 barbs. Lanes 9 – 16 are amplifications from the ND5 gene producing a product of 921 bp and are left to right, 1 kb marker, negative control, positive control from chicken muscle, 1 barb, 2 barbs, 5 barbs, 10 barbs and 20 barbs.....191

**Figure 4.5:** showing amplification of sections of the avian mitochondrial ND2 and ND5 genes from the Superb Parrot (*Polytelis swainsonii*). Lanes 1-6 are amplification from the ND5 gene producing a product of 921 bp and are, left to right, 1 kp marker, negative control, positive control from chicken muscle, 10 barb, 40 barbs and 80 barbs.....192

**Figure 4.6a:** showing an example of the comparison sequence data from one barb after amplification of ND2. The complete 561 bp fragment was sequenced and found to match that of the Short billed Black Cockatoo (*Calyptorhynchus latirostris*) (accession number JF414243) with a similarity of 99%.....194

**Figure 4.6b:** showing an example of the comparison sequence data from two barbs after amplification of DNA of a 921 bp fragment of ND5. A partial sequence of 404 bp was found to match the Short-billed Black Cockatoo (*Calyptorhynchus latirostris*) (accession number JF414243) with a similarity of 99% .....195

**Figure 4.7:** showing an example of the comparison sequence data from two barbs after amplification of DNA of a 921 bp fragment of ND5. A partial sequence of 475 bp from red-winged Parrot (*Aprosmictus erythropterus*) was found to match the kakapo (*Strigops habroptilus*) (accession number AY309456) with a similarity of 86%.....196

**Figure 4.8:** showing the comparison of PCR products after amplifying DNA extracted from QIAGEN and Promega products using the ND2 primer sets. Lane 1 is the 100 bp marker; lane 2 is a negative PCR control; lane 3 is a positive control; lane 4 is from 10 barbs of the Cockatiel (*Nymphicus hallandicus*) sample using the DNA-IQ; lane 5 is 10 barbs of the Cockatiel sample using the QIAGEN micro-kit; lane 6 is 10 barbs from the Regent Parrot (*Polytelis anthopeplus*) using the DNA-IQ; and lane 7 is 10 barbs from the Regent Parrot using QIAGEN Micro-kit. The example shows a greater amount of PCR product at the expected size when DNA was extracted using the QIAGEN kit compared to the Promega DNA-IQ kit; these same data were observed when using fewer barbs and when using the ND5 primer sets.....197

**Figure 5.1:** showing amplification of sections of the avian mitochondrial ND2 and ND5 genes from the Short-billed Black Cockatoo (*Calyptorhynchus latirostris*). Lanes 1-8 are amplification from the ND2 gene producing a product of 561 bp and are, left to right, 100 bp marker, negative control, positive control from chicken muscle, 1 barb, 2 barbs, 5 barbs, 10 barbs and 20 barbs. Lanes 9 – 16 are amplifications from the ND5 gene producing a product of 921 bp and are left to right, 1 kb marker, negative control, positive control from chicken muscle, 1 barb, 2 barbs, 5 barbs, 10 barbs and 20 barbs. The DNA amplification from a single feather barb of the 561 bp segment of the ND2 gene is indicated in the red circle and the DNA amplification from a single feather barb of the 921 bp segment of the ND5 gene is indicated in the yellow circle.....203

## LIST OF TABLES

- Table 1.1:** List of techniques which have been used in the species testing. (+) shows advantages of the technique and (-) shows disadvantages of the technique.....13-14
- Table 1.2:** The taxonomy of a range of species.....15
- Table 1.3:** Gene arrangement of the 13 loci including the ND family, COI, COII, COIII, ATP6, ATP8 and *cyt b* of chicken (*G. gallus*), junglefowl (*G. varius*), guineafowl (*Numida meleagris*), human (*Homo sapiens*), sheep (*Ovis aries*), Northern native cat (*Dasyurus hallucatus*), grasshopper (*Gomphocerippus rufus*), frog (*Xenopus laevis*) and snake (*Crotalus horridus*) obtained from GenBank. The check mark symbol (☐) (indicates the same gene arrangement based on human mitochondrial loci (the first column from the left). The arrangements of the loci that different from the human mtDNA are indicated in blue.....23
- Table 1.4:** The length of mitochondrial amino acid sequences of the 13 loci including the ND family, COI, COII, COIII, ATP6, ATP8 and *cyt b* of chicken (*G. gallus*), junglefowl (*G. varius*), guineafowl (*Numida meleagris*), human (*Homo sapiens*), sheep (*Ovis aries*), Northern native cat (*Dasyurus hallucatus*), grasshopper (*Gomphocerippus rufus*), frog (*Xenopus laevis*) and snake (*Crotalus horridus*) obtained from the DNA database. The mammalian species are highlighted in green and the avian species are highlighted in pink.  
..... 24
- Table 1.5:** The genetic codes for vertebrate mtDNA . The base in the left is the first base of the codon, the base along the top of the table is the second base and that at the right is the third base. In a table of the genetic code U is used to indicate uracil rather than T (Thymine). The different of the vertebrate mitochondrial genetic codons from the standard codons are indicated in red.....29
- Table 1.6:** The classification of 20 amino acids .....30
- Table 1.7:** A list of examples of the online tools and the software packages for multiple sequence alignment and phylogenetic tree reconstruction. Multiple alignments can be performed directly on the website of the online programs on the left column while the software packages from the column on the right have to be downloaded and installed on the local computer first.....35
- Table 2.1:** A list of the 102 avian species, 40 Families and 19 Orders used in this study, including their scientific name, common name, accession number of complete mtDNA genomes. Colour shading indicates members within the same family.....60

<b>Table 2.2:</b> A list of the ND2 sequences from nine avian species used in this study including their scientific name, common name, accession numbers and taxonomic group. Colour shading indicates members within the same Order.....	69
<b>Table 2.3:</b> showing primer sites, highlighted in yellow, and the melting temperature (T <sub>m</sub> ) of the primers for the ND2 gene amplification of avian species. The redundant positions in the primers are in blue: M = A or C, R = A or G, Y = C or T and K = G or T.....	70
<b>Table 2.4:</b> Primer pairs and product size for ND2 gene amplification.....	71
<b>Table 2.5:</b> A list of finch species used in this study including scientific name and common name.....	71
<b>Table 2.6:</b> showing primer sites, highlighted in yellow, and the melting temperature (T <sub>m</sub> ) of the primers for the ND5 gene amplification of avian species. The redundant positions in the primers are in blue: M = A or C, R = A or G, Y = C or T and K = G or T.....	72
<b>Table 2.7:</b> Primer pairs and product size for ND5 gene amplification.....	73
<b>Table 2.8:</b> A list of the 12 parrot and cockatoo species including their scientific name, common name and accession numbers.....	74
<b>Table 2.9:</b> showing primer site, highlighted in yellow, and the melting temperature (T <sub>m</sub> ) of the primer for the ND5 gene amplification of parrot and cockatoo species.....	74
<b>Table 2.10:</b> Primer pairs and product size for ND5 gene amplification.....	75
<b>Table 2.11:</b> A list of the parrot and cockatoo species used in this study including their scientific name, common name, and physical appearances of each species.....	76-81
<b>Table 2.12:</b> A list of the unknown species including with their expected common name.....	82-83
<b>Table 2.13:</b> PCR components were combined as the following:.....	88
<b>Table 2.14:</b> PCR cycles for each primer pairs.....	89

<b>Table 2.15:</b> A list of the avian species used for the protein alignment analysis. These species had complete mitochondrial sequences submitted on Genbank.....	93
<b>Table 2.16:</b> A list of finch species used for intra-species variation study.....	96
<b>Table 2.17:</b> A list of parrot and cockatoo species used for intra-species variation study.....	97
<b>Table 3.1:</b> showing protein length of each locus, percent homology, total number of variable sites and percent variable sites within the complete protein sequences of the ND family, COI, COII, COIII and <i>cyt b</i> loci based on 33 avian species chosen from Galliformes, Anseriformes, Falconiformes, Tinamiformes, Struthioniformes, Ciconiiformes, Pelecaniformes, Sphenisciformes, Charadriiformes, Passeriformes, Podicipediformes, Gruiformes and Gaviiformes.....	102
<b>Table 3.2:</b> A list of the number of each finch species used in this study to analyse intra-species variation.....	138
<b>Table 3.3:</b> Multiple alignment result of the ND2 partial sequences from the five finch species. The yellow colour indicates the variable bases within species. The amino acids (using the three letter code) that are encoded from the variable sites are shown in red above each variable site. The numbers in the bracket indicate the base positions on the ND2 gene from each species. The pink boxes indicate the positions where the amino acid has been changed.....	138-140
<b>Table 3.4:</b> Percent homology of the partial sequences of the ND2, <i>cyt b</i> and COI loci within <i>F. montifringilla</i> , <i>F. coelebs</i> , <i>C. chloris</i> , <i>C. Carduelis</i> and <i>C. spinus</i> using the MEGA 5 program. The numbers in the bracket indicate the base positions on each gene and the length (bp) of the sequences from each species. The number (n) of sequences used in this study is indicated in red.....	141
<b>Table 3.5:</b> A list of finch species used in this study and the accession number of the sequences obtained from the database.....	142
<b>Table 3.6:</b> The table of conclusion showing percent homogy within the species (intra-species variation) and between species (inter-species variation) of the 12 finches.....	143

**Table 3.7:** A list of parrot and cockatoo species that formed the examination of a study of intra-species variation.....149

**Table 3.8:** Multiple alignment result of the ND2 partial sequences from the two parrots; *P. alexandri* and *A. ochrocephala* (data not shown, there is no genetic variation within *A. ochrocephala* species) and the two cockatoo species; *C. banksii* and *C. funereus*. The pink colour indicates the variable sites within species. The amino acids that are encoded by the variable sites are shown in red letter above each variable site. For the two sub-species of the *C. funereus*, the genetic variation sites of the *C. funereus latirostris* and this sub-species are highlighted in grey. The numbers in the bracket indicate the base positions on the ND2 gene from each species. The yellow boxes indicate the positions where the amino acid has been changed.....150-151

**Table 3.9:** Multiple alignment result of the ND5 partial sequences from the two parrots; *P. alexandri* and *A. ochrocephala* and the two cockatoos; *C. banksii* and *C. funereus*. The pink colour indicates the variable sites within species. The amino acids that are encoded by the variable sites are shown in red letter above each variable site. For the two sub-species of the *C. funereus*, the genetic variation sites of the *C. funereus latirostris* and this sub-species are highlighted in grey. The numbers in the bracket indicate the base positions on the ND5 gene from each species. The yellow boxes indicate the positions where the amino acid has been changed.....152-153

**Table 3.10:** Percentage of homology (%) within species of the two cockatoos: *C. banksii* and *C. funereus* and the two parrots: *P. alexandri* and *A. ochrocephala* using the MEGA 5 program. The partial sequences of the ND2 and ND5 genes are obtained from this study and the partial sequence of the *cyt b* gene and the COI DNA Barcoding region are obtained from the database. The numbers in the bracket indicate the base positions on each gene and the length (bp) of the sequences from each species. The number of sequences used in this study indicates in red.....155

**Table 3.11:** A list of parrot and cockatoo species used in analysing inter-species variation as part of this study.....157

**Table 3.12:** The comparison of the use of all methods and models provided in MEGA 5 and MrBayes programs.....166



## LIST OF ABBREVIATION

TM	melting temperature
DTT	dithiothreitol
rpm	revolutions per minute
PCR	polymerase chain reaction
dNTPs	deoxynucleotide triphosphates
bp	base pair
$\mu$ L	microlitre
$\mu$ M	micromolar
pmol	picomole
ng	nanogram
mM	millimolar
mL	milliliter
MEGA	Molecular Evolutionary Genetics Analysis
EBI	European Bioinformatics Institute
iTOL	Interactive Tree Of Life
ML	Maximum Likelihood
NJ	Neighbor Joining
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
MP	Maximum Parsimony
ME	Minimum Evolution
MCMC	Markov chain Monte Carlo
HTUs	hypothetical taxonomic units
OTUs	operational taxonomic units
tRNA	transfer RNA

rRNA	ribosomal RNA
DNA	deoxyribonucleic acid
mtDNA	mitochondrial DNA
COI	cytochrome oxidase c subunit 1
COII	cytochrome oxidase c subunit 2
COIII	cytochrome oxidase c subunit 3
ND	NADH dehydrogenase
<i>cyt b</i>	cytochrome b
ATP6	ATP synthase subunit 6
ATP8	ATP synthase subunit 8

**IUPAC (International Union of Pure and Applied Chemistry) codes**

**Nucleotide code**

A	Adenine
C	Cytosine
G	Guanine
T (or U)	Thymine (or Uracil)
R	A or G
Y	C or T
S	G or C
W	A or T
K	G or T
M	A or C
B	C or G or T
D	A or G or T
H	A or C or T
V	A or C or G
N	any base

### **Amino acid code**

A/Ala	Alanine
C /Cys	Cysteine
D/Asp	Aspartic Acid
E/Glu	Glutamic Acid
F/Phe	Phenylalanine
G/Gly	Glycine
H/His	Histidine
I/Ile	Isoleucine
K/Lys	Lysine
L/Leu	Leucine
M/Met	Methionine
N/Asn	Asparagine
P/Pro	Proline
Q/Gln	Glutamine
R/Arg	Arginine
S/Ser	Serine
T/Thr	Threonine
V/Val	Valine
W/Trp	Tryptophan
Y/Tyr	Tyrosine