# CHAPTER 5 CONCLUSION

## 5.1 Summary

The major purpose of this thesis is to find the smallest DNA segment which is capable of distinguishing closely related avian species. The genes of interest used in this thesis were selected from mitochondrial DNA due to its high copy number and rapid rate of evolution. These benefits are useful for identification in forensic analyses.

Firstly, amino acid sequences were aligned to determine domains of coding regions which are highly conserved or highly variable. This study showed that the ND2 and the ND5 genes exhibited lesser conservation compared to the cyt *b* and COI. The COI locus was found to be the most highly conserved gene compared to the ND2, ND5 and cyt *b* loci. Greater variation of the COI locus was within the 3' end of the amino acid sequences; this is not used as part of the Barcoding locus. It would be expected that parts of the cyt *b* protein showing greater conservation would be associated with the active areas of the protein lying within the mitochondrial membrane, although this was not evident from the amino acid alignment. The genetic variation predicted based on the amino acid sequences of the COI, ND2, ND5 and cyt *b* loci aided in choosing the segments of these loci for further study.

Secondly, while the amino acid alignments indicated areas of conservation (which should have been based on function), greater variation for the purposes of species identification may be identified at the nucleotide level. Phylogenetic tree reconstruction of the complete mitochondrial genome sequences based on 102 avian species showed only a few anomalies, however it is unlikely that in a forensic context the whole of the mitochondrial genome will be available for testing. Smaller parts of mitochondrial DNA that can identify species and reconstruct the phylogenetic tree will be more appropriate for forensic applications therefore, the 37 loci of the avian mitochondrial genome were analysed individually for domains of sequence conservation and variation.

Initially, the percentage of homology of each gene was calculated to determine the level of genetic variation of each locus. The COI, cyt *b*, ND2, ND4 and ND5 genes were selected for species identification and phylogenetic tree reconstruction according to their low percentage of homology between different species, as shown in the previous result of protein sequence analysis section. The complete gene sequences of the selected loci (COI, cyt *b*, ND2, ND4 and ND5) are greater than 1,000 bp; 1041 bp for ND2, 1377 bp for ND4, 1818 bp for ND5, 1143 bp for cyt *b*, 1551 bp for COI. As expected, the complete gene sequences of both ND2 and ND5 were able to split the avian species at this higher taxonomic level, correctly grouping the 102 avian species into their appropriate taxonomic groups at a higher taxonomic level including the subclass Neognathae and subclass Paleognathae as well as at lower taxonomic levels, with only a few anomalies.

These initial studies showed that the ND2 and ND5 genes are preferable for avian species identification and phylogenetic reconstruction based on fewer anomalies. However, the use of the complete gene sequences of these two genes (ND2 and ND5) to identify species is still impractical due to the large size of both complete gene sequences. The large fragment length limits successful DNA amplification from highly degraded DNA, therefore the smaller sections of 100 and 450 bp of these loci were analysed. The 100 and 450 bp segments from the ND2 and ND5 genes were compared to the segments from COI, cyt *b* and the ND4 loci, with the result that it demonstrated further that segments of the ND2 and ND5 loci are superior in identifying avian species and reconstructing phylogenetic trees compared to segments of the COI, cyt *b* and ND4 loci. As expected, the segments that can identify species and reconstruct a phylogenetic tree with fewest anomalies were parts of the ND2 gene (base positions 1 - 450), and the ND5 gene (base positions 101 - 550) which are from 5' end of both genes.

**5.2 Applications**

This study has shown that the segments from the 5' end of the ND2 and ND5 loci are able to be used for identify closely related avian species of the Fringillidae (finches), Psittacidae (parrots) and Cacatuidae (cockatoos) families. Additionally, DNA amplifications of the 561 bp segment of the ND2 and the 921 bp of the ND5 loci from a single feather barb demonstrated for the first time in this study, as shown in Figure 5.1.

This included the successful of PCR amplification of part of the ND5 gene (921 bp) using a barb taken from a museum specimen that was taxidermically mounted in 1979. As this sample was collected 32 years prior to the time of analysis, the test confirms that this technique can be applied in forensic analysis for the DNA amplification from small amounts and highly degraded DNA template.

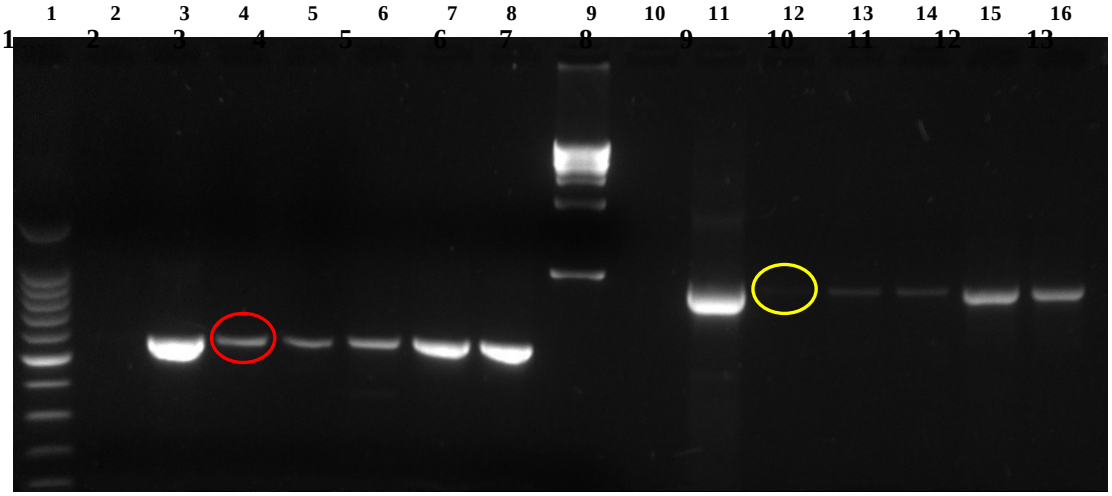


**Figure 5.1: showing amplification of sections of the avian mitochondrial ND2 and ND5 genes from the Short-billed Black Cockatoo (*Calyptorhynchus latirostris*). Lanes 1-8 are amplification from the ND2 gene producing a product of 561 bp and are, left to right, 100 bp marker, negative control, positive 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.**

### 5.3 Validation

Any new marker or tool to be used in the criminal justice system or forensic casework then needs to be validated. Validation includes sensitivity, specificity, stability, reproducibility and accuracy.

It is necessary to performed sensitivity studies to determine the lowest amount of DNA that will generate a result. The sensitivity of both primer sets (FC and RV primers, and upF1 and R1 primers) for the ND2 and ND5 segments need to be tested by performing DNA amplification of the DNA extracted from various concentration of DNA template, such as the DNA extracted from 1, 2, 5, 10, 20, 40 and 80 barbs.

It is essential to performed specificity tests to ensure the specificity for a particular species. Equally it is necessary to demonstrate that there is no cross-reactivity with related species. The possibilities to detect DNA profiles from non-targeted species from the two primer sets (FC and RV primers, and upF1 and R1 primers) for the ND2 and ND5 segments must be determined by using them to amplify the DNA template from various species of birds, such as chicken, turkey, quail, parrots, and cockatoos and not from other organisms such as humans and reptiles.

The Environmental and other factors are tested to determine which, if any, affect the stability of the sample. An initial example is using museum and achiral samples, although exposing feathers to external factors such as rain and temperature fluctuations need to be investigated.

It is important that the test will generate the same result. Linked to this are tests of the accuracy and precision of the test. Reproducibility can be tested by performing many replications of the test and checked the results from each replication are the same.

Blind trial testing and testing by another laboratory will ensure that the test works on unknown samples and by other operators. Finally publication in a peer reviewed journal or dissemination at an international conference will allow outside scrutiny of the test.

**5.4 Dissemination**

The COI and cyt *b* genes are commonly used in species identification and taxonomic studies [1, 2]. In contrast, this study illustrated that the ND2 and ND5 loci are more suitable for species identification and phylogenetic tree reconstruction of the avian species, based on the species used in this study, than the use of COI and cyt *b* loci.

The use of the ND2 gene in avian species identification and phylogenetic tree reconstruction was presented at the International Society for Forensic Genetics (ISFG), Buenos Aries, 2009 and published as a conference proceeding in the journal Forensic Science International (FSI): Genetics Supplement Series. This is presented in Appendix E.

Another paper describing a technique for avian species identification from a single barb based on the segments of the 5' end of the ND2 and the ND5 loci has been accepted by the Journal of Forensic Sciences as a Technical Note and is supplied in Appendix E.

A publication in preparation will detail the phylogenetic tree reconstructions using the 561 and 921 bp segments of the 5' end of the ND2 and the ND5 loci, respectively. This paper will be submitted in the first instance to PLoS ONE.

These data may also be disseminated at the forthcoming forensic conferences:

- the 64th American Academy of Forensic Sciences (AAFS) annual scientific meeting will be held Atlanta, Georgia, USA from 20-25 February 2012.
- the Society for Wildlife Forensic Science (SWFS) meeting will be held in Wyoming, USA from 22-25 May 2012.
- the 6th European Academy of Forensic Science (EAFS) conference will be held in The Hague, the Netherlands from 20 to 24 August 2012.

- the 21st International Symposium on the Forensic Sciences of the Australian and New Zealand Forensic Science Society (ANZFSS), to be held in Hobart, Tasmania from the 23–27 September 2012.
- the 25th International Society for Forensic Genetics (ISFG) congress will be held in Melbourne, Australia from the 2-7 September 2013.

There was a wildlife forensic meeting of the 1st Australia and New Zealand Wildlife Forensics Workshop and Symposium held at the New Zealand Centre for Conservation Medicine (NZCCM) in Auckland from the 1-4 February 2011. The next meeting of the Wildlife Forensics Workshop and Symposium, a New Zealand and Australian Perspective, will be ideal to disseminate the technique for avian species identification. Many bird species are protected by the National Parks and Wildlife Act 1972 (South Australia legislation) such as some parrot and cockatoo species as listed below:

**Psittacidae (parrots)**

- *Glossopsitta pusilla* (Little Lorikeet)
- *Lathamus discolour* (Swift Parrot)
- *Neophema chrysogaster* (Orange-bellied Parrot)
- *Pezoporus occidentalis* (Night Parrot)
- *Pezoporus wallicus* (Ground Parrot)

**Cacatuidae (cockatoos)**

- *Calyptorhynchus banksii graptogyne* (Red-tailed Black-Cockatoo, south east subspecies)
- *Calyptorhynchus lathami halmaturinus* (Glossy Black-Cockatoo, Kangaroo Island subspecies)

### 5.5 The technologies for species identification

Fluorescent fragment length Barcoding (FFLB) is a DNA based method which has been used to identify the African tsetse-transmitted trypanosomes species based on length variation in regions of the 18S and 28Sα ribosomal DNA using fluorescently tagged primers. These primer sets will amplify fragments of various sizes to generate a unique Barcode of each species [3]. A further method uses a variety of sizes of the amplicons for species identification is the use of size variation of the PCR products from the amplification of the mitochondrial D-loop region. This method uses non-fluorescent primer sets and gel electrophoresis or fluorescent primer sets detected by capillary electrophoresis. The sizes of the PCR products range from 350-900 bp which can distinguish different species [4]. The use of melt curve analysis (MCA) is a gel-free technique where different species can be identified by using species-specific primers to amplify a particular amplicon and from which the melt temperature (Tm) of the PCR fragment can be determined [5]. Terminal restriction fragment length polymorphism (T-RFLP) may include the amplification of the Barcode region with a fluorescent forward primer.

The length of the terminal fragments vary depending on the restriction sites within the Barcode region of each species [6, 7]. All these methods above are unable to identify closely related species and unable to identify species from DNA mixtures. Species identification from mixture samples can be performed using species-specific multiplex primers [8]. These primers were designed to amplify fragments of various sizes which are specific for each species. This technique will be able to identify specific species for which the test has been designed, and these multiplex primers will not work on any other species.

This study has used a version of the forensically informative nucleotide sequences (FINS) technique [9]. The method used of amplification of a section of the mitochondrial genome, in this case parts of the ND2 and ND5 loci, are simple and convenient to perform in any forensic science laboratory. A limited study on intra-species variation was performed, although a more comprehensive study is needed, as was performed for mammalian species [1]. As avian species are less studied to date than mammalian species, a main problem that often occurs that there is no reference sequence data submitted on the database for a particular avian species. If the aim is to identify to species level an unknown sample then this can be a real problem in a forensic context. A significant problem is that of a lack of regulation of GenBank as there are sequences on this database that are erroneous.

### 5.6 Future developments

This study has demonstrated that the gene loci that have been successful for species identification of mammalian species, such as the cyt *b* and COI genes are not the most suitable locus for the avian species identification. While the COI Barcoding locus is useful in invertebrate species and cyt *b* in mammalian species, ND2 and ND5 loci are more suitable for avian species. Any gene locus used in species identification and phylogenetic tree reconstruction in different taxonomic groups have to be validated for that particular taxonomic group first. In addition to the lack of regulation of GenBank, an issue is that there are few avian species listed compared to mammalian species. Further, there are few DNA sequences covering the ND2 and ND5 gene loci. If, as proposed in this thesis, these two loci are superior to cyt *b* or COI for avian species identification, then a major hurdle is that already there are many more cyt *b* or COI sequences for the avian species of forensic interest, than for ND2 or ND5. The techniques that have been used since the start of this thesis have improved. Examples include the capability and speed of the PCR machines and novel versions of Taq DNA polymerase allowing for longer fragments to be amplified. DNA sequencing technology allows for ever increasing length of reads. Software packages have been recently released allowing for faster alignment of sequences from which a range of phylogenetic trees can be generated. With the advent of mass parallel DNA sequencing it is feasible to sequence complete mitochondrial genomes very rapidly and hence the argument over which loci is the more informative will no longer be an issue.

# Chapter 5 References


1. Tobe, S.S., A.C. Kitchener, and A.M.T. Linacre, *Reconstructing Mammalian Phylogenies: A Detailed Comparison of the Cytochrome and Cytochrome Oxidase Subunit I Mitochondrial Genes.* PLoS ONE, 2010. 5(11): p. e14156 EP -.

2. Linacre, A. and S. Tobe, *An overview to the investigative approach to species testing in wildlife forensic science.* Investigative genetics, 2011. 2(1): p. 2.

3. Hamilton, P.B., a, and a, *A novel, high-throughput technique for species identification reveals a new species of tsetse-transmitted trypanosome related to the Trypanosoma brucei subgenus, Trypanozoon.* Infection Genetics and Evolution, 2008. 8(1): p. 26-33.

4. Nakamura, H., a, and a, *Forensic species identification based on size variation of mitochondrial DNA hypervariable regions.* International Journal of Legal Medicine, 2009. 123(2): p. 177-184.

5. Berry, O. and S.D. Sarre, *Gel-free species identification using melt-curve analysis.* Molecular Ecology Notes, 2007. 7(1): p. 1-4.

6. Darling, J.A. and M.J. Blum, *DNA-based methods for monitoring invasive species: a review and prospectus.* Biological Invasions, 2007. 9(7): p. 751-765.

7. Dorigo, U., L. Volatier, and J.-F. Humbert, *Molecular approaches to the assessment of biodiversity in aquatic microbial communities.* Water Research, 2005. 39(11): p. 2207-2218.

8. Tobe, S. and A. Linacre, *A multiplex assay to identify 18 European mammal species from mixtures using the mitochondrial cytochrome b gene.* Electrophoresis, 2008. 29(2): p. 340 - 347.

9. Bartlett, S. and W. Davidson, *FINS (forensically informative nucleotide sequencing): a procedure for identifying the animal origin of biological specimens.* BioTechniques, 1992. 12(3): p. 408-411.