CHAPTER 2 MATERIALS AND METHODS

2.1 Sequence Data

The avian mitochondrial DNA genome was studied in detail by looking initially at the complete mitochondrial genome, each complete mitochondrial gene sequence and subsequently partial sections of particular mitochondrial gene. Avian species were obtained from the GenBank DNA database (http://www.ncbi.nlm.nih.gov). All DNA sequences used in the initial study were taken from *c*omplete mitochondrial genome sequences from a wide range of avian species which cover most of avian taxonomic groups and included 102 species, 75 genera, 40 Families and 19 Orders (out of total 32 order for avian species). These are shown in Table 2.1. It should be noted that there are over 10,000 avian species, however the species chosen were based on being representative of the main families and then their availability.

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.

1. Order Galliformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Tetraophasis szechenyii	Buff-throated Partridge	FJ799728	Phasianidae
Pavo muticus	Green peafowl	EU417811	
Francolinus pintadeanus	Chinese Francolin	NC_011817	
Lophura nycthemera	Silver pheasant	EU417810	
Lophura ignita	Crested fireback pheasant	NC_010781	
Syrmaticus soemmerringi ijimae	Ijima copper pheasant	AB164622	
Syrmaticus humiae	Hume's pheasant	NC_010774	
Syrmaticus ellioti	Elliot's pheasant	NC_010771	
Syrmaticus reevesii	Reeves's pheasant	NC_010770	
Phasianus versicolor	Green pheasant	NC_010778	
Tragopan caboti	Cabot's tragopan	NC_013619	
Arborophila rufipectus	Sichuan partridge	FJ194942	
Gallus gallus	Chicken	AY235571	
Gallus sonneratii	Gray junglefowl	AP003320	
Gallus lafayettei	Ceylon junglefowl	AP003325	
Bambusicola thoracica	Chinese bamboo-partridge	EU165706	
Meleagris gallopavo	Turkey	NC_010195	
Coturnix japonica	Japanese quail	AP003195	
Coturnix chinensis	Excalfactoria chinensis	NC_004575	
Polyplectron bicalcaratum	Gray peacock-pheasant	EU417812	
Numida meleagris	Helmeted guineafowl	Helmeted guineafowl	Numididae

2.Order Anseriformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Cygnus columbianus	Tundra swan	NC_007691	Anatidae
Cygnus atratus	Black swan	NC_012843	
Branta canadensis	Canada goose	NC_007011	
Anser albifrons	White-fronted goose	NC_004539	
Anser anser	Domestic goose	NC_011196	
Cairina moschata	Muscovy duck	EU755254	
Anas platyrhynchos	Mallard	EU009397	
Aythya americana	Redhead	AF090337	
Dendrocygna javanica	Lesser whistling duck	NC_012844	
Anseranas semipalmata	Magpie goose	NC_005933	Anseranatidae

3.Order Gruiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Grus canadensis	Sandhill crane	FJ769855	Gruidae
Grus paradisea	Blue crane	FJ769844	
Grus carunculatus	Wattled crane	FJ769843	
Grus antigone	Sarus crane	FJ769854	
Grus rubicunda	Brolga	FJ769853	
Grus vipio	White-naped crane	FJ769852	
Grus virgo	Demoiselle crane	FJ769845	
Grus nigricollis	Black-necked crane	FJ769851	
Grus monacha	Hooded crane	FJ769850	
Grus americana	Whooping crane	FJ769848	
Grus grus	Eurasian crane	FJ769849	
Grus japonensis	Japanese crane	FJ769847	
Grus leucogeranus	Siberian crane	FJ769846	
Balearica pavonine	Crowned crane	FJ769842	
Gallirallus okinawae	Okinawa rail	NC_012140	Rallidae
Rallina eurizonoides sepiaria	Slake	NC_012142	
Porphyrio hochstetteri	South Island takahe	NC_010092	
Rhynochetos jubatus	Kagu	NC_010091	Rhynochetidae

4.Order Charadriiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Arenaria interpres	Ruddy turnstone	NC_003712	Scolopacidae
Haematopus ater	Blackish oystercatcher	NC_003713	Haematopodidae
Synthliboramphus antiquus	Ancient murrelet	NC_007978	Alcidae

5.Order Falconiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Falco peregrinus *Falco tinnunculus* *Falco sparverius* *Micrastur gilvicollis* *Buteo buteo*	Peregrine Falcon Common kestrel American kestrel Lined Forest-falcon Common buzzard	AF090338 NC_011307 NC_008547 DQ780881 NC_003128	Accipitridae
Spizaetus alboniger *Accipiter gentilis* *Spizaetus nipalensis* *Pandion haliaetus*	Blyth's hawk-eagle Goshawk Mountain hawk-eagle Osprey	AP008239 NC_011818 AP008238 NC_008550	Falconidae

6.Order Procellariiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Pterodroma brevirostris	Kerguelen petrel	AY158678	Procellariidae
Thalassarche melanophris	Black-browed albatross	AY158677	Diomedeidae

7.Order Struthioniformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Apteryx owenii *Apteryx haastii*	Little spotted kiwi Great spotted kiwi	NC_013806 NC_002782	Apterygidae
Casuarius casuarius	Southern cassowary	NC_002778	Casuariidae
Struthio camelus	Ostrich	NC_002785	Struthionidae
Dromaius novaehollandiae	Emu	NC_002784	Dromaiidae
Pterocnemia pennata *Rhea americana*	Lesser Rhea Greater rhea	NC_002783 AF090339	Rheidae
Emeus crassus *Anomalopteryx didiformis* *Dinornis giganteus*	Eastern moa Little bush moa Giant moa	NC_002673 AF338714 AY016013	Dinornithidae

8.Order Tinamiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Tinamus major *Eudromia elegans*	Great tinamou Elegant crested-tinamou	NC_002781 AF338710	Tinamidae

9.Order Piciformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Dryocopus pileatus	Pileated woodpecker	NC_008546	Picidae
Pteroglossus azara flavirostris	Ivory billed aracari	NC_008549	Ramphastidae

10.Order Passeriformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Smithornis sharpei	Grey-headed broadbill	NC_000879	Eurylaimidae
Vidua chalybeate	Steelblue widowfinch	AF090341	Viduidae
Taeniopygia guttata	Zebra finch	NC_007897	Estrildidae
Corvus frugilegus	Rook	NC_002069	Corvidae
Pycnonotus taivanus	Taiwan bulbul	NC_013483	Pycnonotidae
Sylvia crassirostris *Sylvia atricapilla* *Acrocephalus scirpaceus*	Eastern orphean warbler Blackcap Eurasian reed warbler	NC_010229 NC_010228 NC_010227	Sylviidae

11.Order Ciconiiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Ardea novaehollandiae *Egretta eulophotes*	White-faced Heron Chinese egret	DQ780878 NC_009736	Ardeidae
Nipponia nippon *Threskiornis aethiopicus* *Platalea leucorodia* *Platalea minor*	Crested ibis Sacred Ibis Eurasian spoonbill Black-faced spoonbill	NC_008132 GQ358927 NC_012772 EF455490	Threskiornithidae
Ciconia boyciana *Ciconia ciconia*	Oriental stork White stork	NC_002196 NC_002197	Ciconiidae

12.Order Gaviiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Gavia stellata	Red-throated loon	NC_007007	Gaviidae

13.Order Phoenicopteriformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Phoenicopterus ruber roseus	Greater flamingo	NC_010089	Phoenicopteridae

14.Order Pelecaniformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Phaethon rubricauda	Red-tailed tropicbird	AP009043	Phaethontidae

15.Order Sphenisciformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Eudyptula minor	Little blue penguin	NC_004538	Spheniscidae

16.Order Psittaciformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Strigops habroptilus	Kakapo	NC_005931	Psittacidae
Melopsittacus undulates	Budgerigar	NC_009134	

17.Order Apodiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Archilochus colubris	Ruby-throated hummingbird	NC_010094	Trochilidae

18.Order Strigiformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Ninox novaeseelandiae	Morepork	NC_005932	Strigidae

19.Order Podicipediformes

Scientific name	Common name	Full mtDNA genome's accession number	Family
Tachybaptus novaehollandiae	Australasian grebe	NC_010095	Podicipedidae

2.2 Mitochondrial protein sequence analysis

This aspect of the thesis starts with the protein alignment of the ND family genes and the COII and COIII genes compared to the cyt *b* and the COI gene loci which have been widely used in species identification and taxonomic studies. The protein sequences of these genes were obtained from GenBank and included 33 avian species. The species were chosen from Galliformes, Anseriformes, Falconiformes, Tinamiformes, Struthioniformes, Ciconiiformes, Pelecaniformes, Sphenisciformes, Charadriiformes, Passeriformes, Podicipediformes, Gruiformes and Gaviiformes. The 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 cyt *b* loci were determined. The sequences corresponding to the genes with the highest variation were used for further analysis.

Regions of sequence where lesser or greater variation was noted were examined at each amino acid position aim to identify conserve and variable sequences within each gene. This information was used to select position for designing primer.

2.3 Primer design

Based on the variation of the protein sequence result from section above, the ND2 and ND5 loci were selected for further analysis. Primers were designed for amplifying several sections of these genes with the aim to find the optimal section which shows the highest power of distinguishing closely related avian species.

2.3.1 ND2 primers

Universal primers to amplify the complete avian ND2 gene were designed from DNA sequences obtained from the GenBank DNA database. All retrieved DNA sequences cover the region before the start codon of the ND2 gene, being within the tRNA-Met gene, to a region after the stop codon, being within the tRNA-Trp gene (around base position 5176 to 6358, based upon the mitochondrial genome of *Gallus gallus*). As the tRNA genes were known to be extremely highly conserved, therefore the tRNA-Met gene and the tRNA-Trp gene were expected to be very similar in all avian species. For this reason, only nine avian species from the Orders Falconiformes, Galliformes and Anseriformes were most likely to be representative of avian species for designing the ND2 primers. The common name, accession numbers and taxonomic group of the species from which the sequences were obtained are found in Table 2.2.

Table 2.2: 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.

Scientific name	Common name	Accession number	Order
Sturnella neglecta	Western meadowlark	FJ154705	Falconiformes
Buteo jamaicensis	Red-tailed hawk	AY987156	
Buteo buteo	Common buzzard	NC_003128	
Falco tinnunculus	Common kestrel	NC_011307	
Pandion haliaetus	Osprey	NC_008550	
Gallus gallus	Chicken	GU261719	Galliformes
Phasianus colchicus	Ring-necked pheasant	JF739859	
Anas americana	American wigeon	AF059163	Anseriformes
Anas platyrhynchos	Mallard Duck	HM010684	

The complete ND2 gene in most avian species is 1041 bp, although it is 1038 bp in some species. Primers were designed to amplify the complete gene sequence of the ND2 gene from all avian species. Multiple sequence alignments were performed using the MEGA program. The primer binding sites and the melting temperature (TM) of each primer are shown in Table 2.3.

Table 2.3: showing primer sites, highlighted in yellow, and the melting temperature (Tm) 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.

Primers	Tm (°C)
upF1	63.2
F2	53.6
R1 **(R1 is a reverse complementary sequence of F2)**	50.9
dnR2	58.6

The upF1 and R1 primers will amplify a partial sequence of the tRNA-Met gene to the middle of the ND2 gene; being a total product size of approximately 561 bp. The primer pair F2 and dnR2 will amplify from the middle of the ND2 gene to the tRNA-Trp gene; being a total product size of approximately 598 bp. When using the primer pair upF1 and the dnR2 to amplify the complete ND2 gene the total product size is approximately 1139 bp. All expected products from those primers are shown in Table 2.4. The overlapping sites of the expected PCR products are shown in Figure 2.1.

Table 2.4: Primer pairs and product size for ND2 gene amplification.

Primer pairs	Product size (bp)
upF1 and R1	561
F2 and dnR2	598
upF1 and dnR2	1130

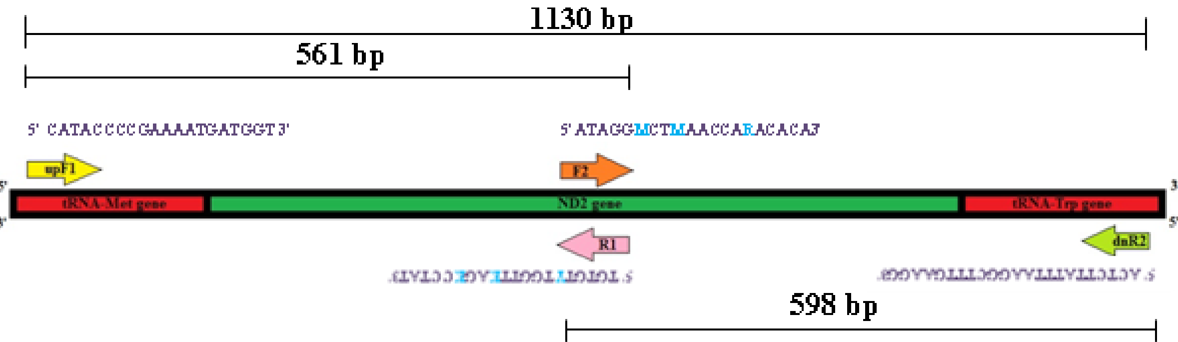


Figure 2.1: A map of the overlapping sites of the expected PCR products generated from the ND2 primers.

To test whether ND2 can be used to separate members of the same genus, DNA samples were extracted from dried blood spots taken from nine members of the family Fringillidae as shown in Table 2.5.

Table 2.5: A list of finch species used in this study including scientific name and common name.

Scientific name	Common name
Fringilla montifringilla	Bramblefinch
Pyrrhula pyrrhula	Bullfinch
Coccothraustes coccothraustes	Hawfinch
Carduelis spinus	Sisken
Carduelis cannabina	Linnet
Carduelis carduelis	Goldfinch
Carduelis cabaret	Lesser Red poll
Carduelis chloris	Greenfinch
Fringilla coelebs	Chaffinch

DNA extraction from the dried blood spot samples can be found in section 2.3.1. The entire ND2 locus was amplified by using two primer pairs, primer upF1 with R1, F2 with dnR2. The PCR products were purified and sequenced in both directions as described in section 2.5 and 2.6, respectively. The PCR product was sequenced in both directions. All sequences were aligned and generated a phylogenetic tree using MEGA program as described in section 2.10.2.2.

2.3.2 ND5 primers

The primer pairs to amplify the complete ND5 gene in all avian species were designed from mitochondrial DNA sequences of 102 avian species (Table 2.1). All sequences cover the region from partial tRNA-Leu gene to the partial cyt *b* gene (around base position 13007 to 16042, based upon the mitochondrial genome of *Gallus gallus*). The alignment result and the primer biding site of each primer are shown in Table 2.6.

Table 2.6: showing primer sites, highlighted in yellow, and the melting temperature (Tm) 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.

Primers	Tm(^{o}C)
Fc	54
RV and FW (RV is a reverse complementary sequence of FW)	54
Rc	54

The primers for complete gene sequence amplification of the ND5 gene were Fc and Rc and the total product size is approximately 1960 bp. This product size is too large for sequencing in one reaction. Alternative primer pairs are the Fc and RV primers and the Fw and Rc primers which were used to amplify two fragments; upstream to the middle of the gene and the middle to downstream of the gene. The Fc and RV primers will amplify the 5’ end part of the ND5 gene with a total product size of approximately 921 bp and the FW and Rc primers will amplify the 3’ end part of ND5 gene with a total product size of approximately 1062 bp, as shown in Table 2.7. The overlapping sites of the expected PCR products are shown in Figure 2.2.

Table 2.7: Primer pairs and product size for ND5 gene amplification.

Primer pairs	Product size (bp)
Fc and Rc	1960
Fc and RV	921
FW and Rc	1062

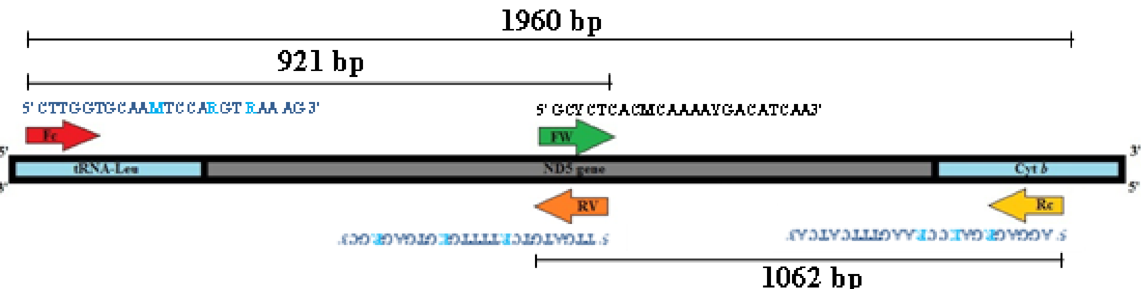


Figure 2.2: A map of the overlapping sites of the expected PCR products generated from the ND5 primers.

2.3.3 The primer for ND5 gene sequencing of the parrot and cockatoo species

The primers for parrot and cockatoo species were designed from the ND5 sequences obtained from GenBank. There were only 12 sequences of parrot and cockatoo species submitted on the DNA database at the time of analysis. Their common name and accession number can be found in Table 2.8. The primer sequences, primer binding region and Tm are shown in Table 2.9.

Table 2.8: A list of the 12 parrot and cockatoo species including their scientific name, common name and accession numbers.

Scientific name	Common name	Accession number
Brotogeris cyanoptera	Cobalt-winged parakeet	NC_015530
Nymphicus hollandicus	Cockatiel	NC_015192
Calyptorhynchus latirostris	Short-billed white-tailed black-cockatoo	JF414243
Calyptorhynchus baudinii	Long-billed white-tailed black-cockatoo	JF414242
Calyptorhynchus lathami	Glossy black-cockatoo	JF414241
Cacatua pastinator butleri	Western corella	JF414240
Cacatua moluccensis	Salmon-crested cockatoo	JF414239
Aratinga pertinax chrysogenys	Brazilian brown-throated parakeet	NC_015197
Strigops habroptilus	Kakapo	NC_005931
Forpus modestus	Dusky-billed parrotlet	HM755882
Agapornis roseicollis	Peach-faced lovebird	NC_011708
Melopsittacus undulatus	Budgerigar	NC_009134

Table 2.9: showing primer site, highlighted in yellow, and the melting temperature (Tm) of the primer for the ND5 gene amplification of parrot and cockatoo species.

Primers	Tm (°C)
RV-PC	52

The PCR products for complete gene sequencing of the ND5 gene from parrot and cockatoo species were prepared using the primer pairs shown in Table 2.10. The overlapping sites of the expected PCR products are shown in Figure 2.3.

Table 2.10: Primer pairs and product size for ND5 gene amplification.

Primer pairs	Product size (bp)
Fc and RV-PC (5'end part)	1212
FW and Rc (3'end part)	1067

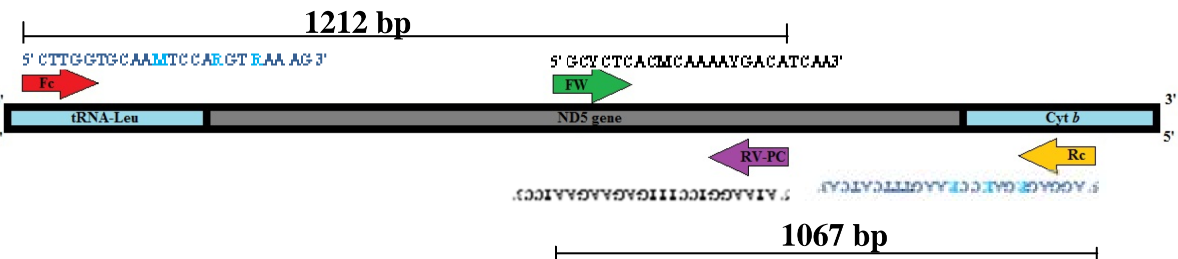


Figure 2.3: A map of the overlapping sites of the expected PCR products from the ND5 primers for the complete ND5 gene sequencing of the parrot and cockatoo species.

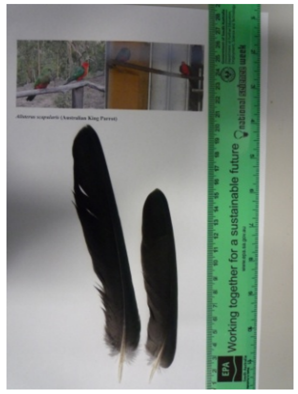
2.4 Sample collection

2.4.1 Voucher specimen

The parrot and cockatoo species were sampled from a private collector: Mr. Thomas Massey, 283 Lower Athelstone Road, Athelstone SA 5076. The feather samples from 19 individuals were plucked from each species. All species were verified independently by Dr. Greg Johnston of the University of South Australia and the Museum of South Australia. Their common name and pictures are shown in Table 2.11.

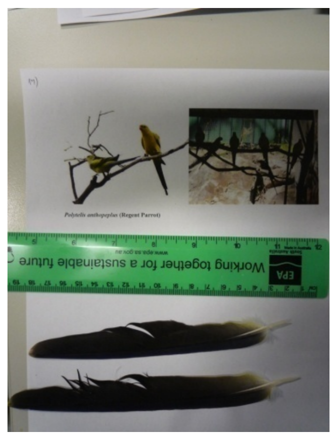
Table 2.11: A list of the parrot and cockatoo species used in this study including their scientific name, common name, and physical appearances of each species.

Scientific name	Common name	Physical appearances
Lophochroa leadbeateri	Major Mitchell's Cockatoo	
Aprosmictus erythropterus	Red-winged Parrot	
Callocephalon fimbriatum	Gang-gang Cockatoo	

Scientific name	Common name	Physical appearances
Alisterus scapularis	Australian King Parrot	
Eolophus roseicapilla	Galah Cockatoo	
Calyptorhynchus banksii	Red-tailed Black Cockatoo	
Glossopsitta pusilla	Little Lorikeet	

Scientific name	Common name	Physical appearances
Psephotus dissimilis	Hooded Parrot	
Calyptorhynchus lathami	Glossy Black Cockatoo	
Neophema splendida	Scarlet-chested Parrot	

Scientific name	Common name	Physical appearances
Trichoglossus haematodus	Rainbow Lorikeet	
Psephotus chrysopterygius	Golden-shouldered Parrot	
Calyptorhynchus funereus	Yellow-tailed Black Cockatoo	
Lathamus discolour	Swift Parrot	

Scientific name	Common name	Physical appearances
Calyptorhynchus latirostris	Carnaby's Black-Cockatoo	
Polytelis anthopeplus	Regent Parrot	
Platycercus elegans adelaide	Adelaide Rosella	

Scientific name	Common name	Physical appearances
Polytelis swainsonii	Superb Parrot	
Polytelis alexandrae	Princess Parrot	

2.4.2 Unknown samples

Shed feathers of unknown species were collected from various places in Adelaide including the bank of the River Torrens, Victoria Square and the car park of Westfield Marion Shopping Centre as shown in the Google map (Figure 2.4). For unknown specimens, their species names were estimated from their feather morphology and the common species that have found in those areas (Table 2.12). This part of the study aims to test whether the sequence from ND2 and ND5 loci can identify unknown species correctly. Derived sequences from the ND2 and ND5 fragments of unknown species were compared to the sequences on GenBank.

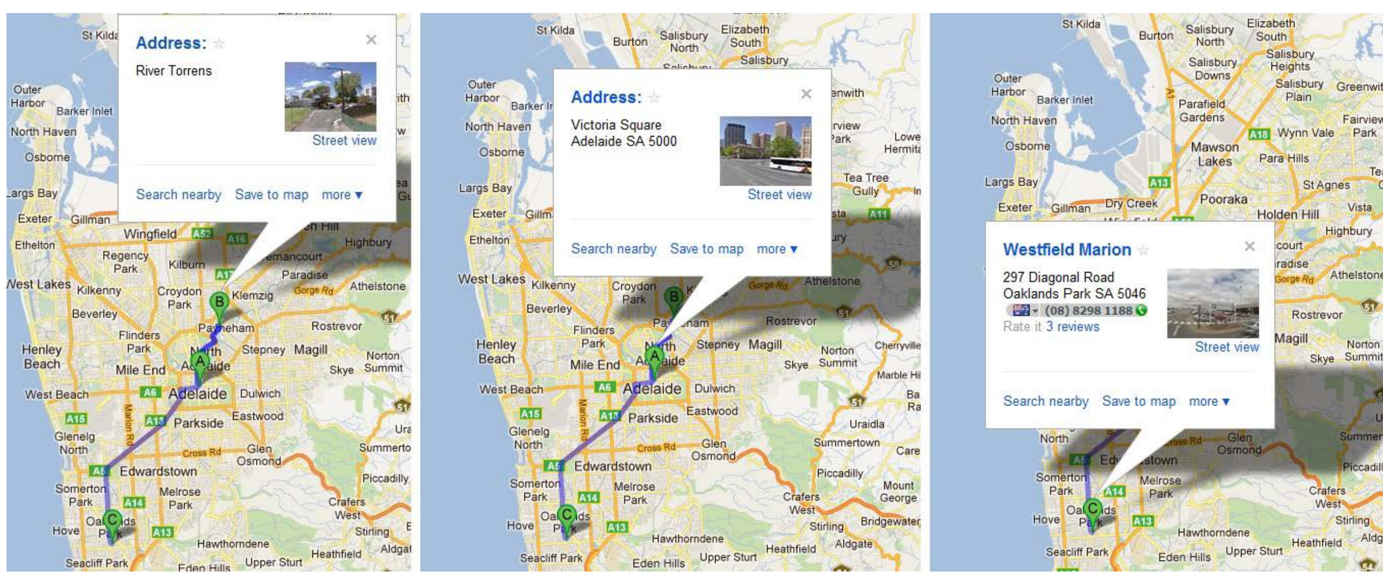
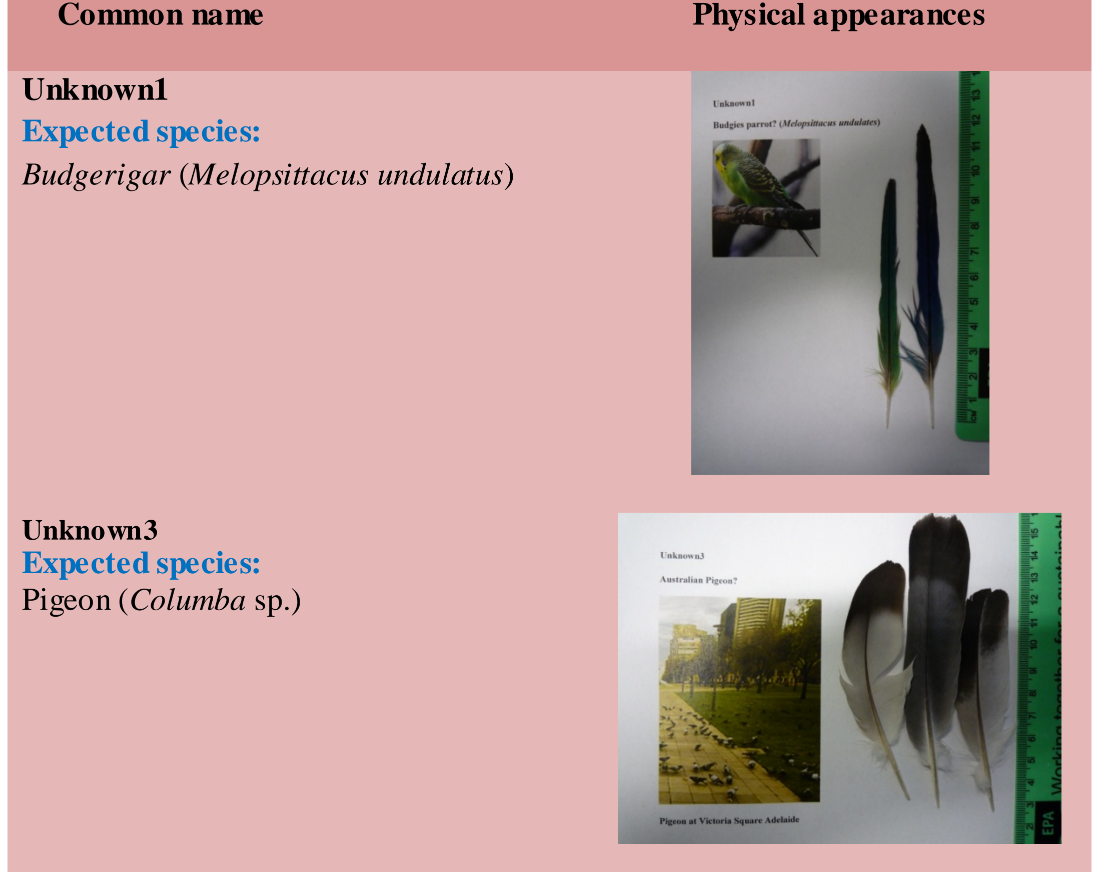


Figure 2.4: showing Google map positions of the River Torrens, Victoria Square and the Westfield Marion Shopping Centre, Adelaide, South Australia.

Table 2.12: A list of the unknown species including with their expected common name.



Common name	Physical appearances
Unknown1 **Expected species:** *Budgerigar* (*Melopsittacus undulatus*)	
Unknown3 **Expected species:** Pigeon (*Columba* sp.)	

Common name	Physical appearances
Unknown4 **Expected species:** Rainbow Lorikeet (*Trichoglossus haematodus*)	

2.5 DNA extraction

2.5.1 DNA Extraction from dried blood spots

The DNA extraction from dried blood spots was performed using the QIAamp® DNA micro Kit (QIAGEN, Doncaster, Victoria). All steps for the DNA extraction and the explanation of each step, including the procedure of the kit, are found in the QIAamp® DNA Micro handbook on pages 19-21 and 8-9, respectively [1].

2.5.2 DNA Extraction from Feathers

DNA extraction from feathers was performed using the QIAamp® DNA micro Kit from QIAGEN. The DNA extraction procedure was modified from the User-Developed Protocol*:* Purification of DNA from nails, hair, or feathers [2]. Requirements prior to extraction are found on page 2 of this protocol:

Requirements prior to extraction

- Prepare an aqueous 1 M Dithiothreitol (DTT) solution; 0.154 g in 1 mL sterile water. Store aliquots at -20 ^{o}C. Thaw immediately before use.
- Buffer ATL and Buffer AL may form precipitates upon storage. If necessary, warm in the microwave for about 10 seconds until the precipitates have fully dissolved.
- Buffer AW1 and Buffer AW2 were supplied as concentrates. Before using for the first time, add the appropriate amount of ethanol (96-100%) as indicated on the bottle to obtain a working solution.
- Preheat a thermomixer, shaking water bath, or rocking platform to 56°C for use in step 3.

The procedure for DNA extraction from feather using the QIAamp® DNA micro Kit

1. Cut a 1 cm piece of calamus and use 2 pieces of calamus per tube or use barbs (2-10 barbs), and transfer them to a 1.5 mL microcentrifuge tube.
2. Add 300 µL of ATL buffer, 20 µL of proteinase K, and 10 µL of 1 M DTT.
3. Incubate at 56 ^{o}C in a thermomixer compact (Eppendorf) until the sample is completely lysed.
4. Vortex for 2 seconds and spin down. Add 300 µL of AL buffer and 300 µL of absolute ethanol and mix.

5. Transfer the mixture into the QIAamp MinElute column placed in a 2 ml collection tube. Centrifuge at 8000 rpm for 1 minute. Discard flow-through and collection tube.
6. Place the column in a new collection tube, add 500 µL of AW1, and centrifuge at 8000 rpm for 1 minute. Discard flow-through and collection tube.
7. Place the column in a new collection tube, add 500 µL of AW2, and centrifuge at 8000 rpm for 1 minute. Discard flow-through and empty the collection tube and reuse it in another centrifugation at 13000 rpm for 3 minutes to dry the membrane of the column.
8. Place the column in a clean 1.5 mL microcentrifuge tube and pipet 50 µL of AE buffer directly onto the QIAamp MinElute membrane. Incubate at room temperature for 5 minutes, and then centrifuge at 8000 rpm for 1 minute to elute.

In this study, the DNA was extracted from two parts of the feather: the calamus and barbs. Two calamus per species were used for DNA extraction (Figure 2.5). For DNA extraction from individual barbs, the numbers of barb varied from 1, 2, 5, 10, 20, 40 and 80 barbs (Figure 2.6).

(1) DNA Extraction from calamus

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

(2) DNA Extraction from barbs

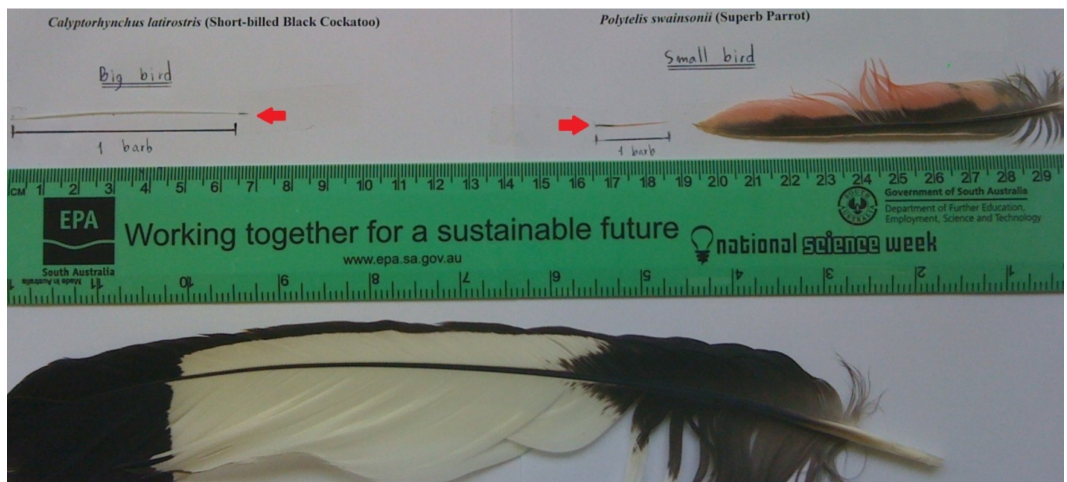


Figure 2.6: showing feather and barb samples from the smaller size bird: superb parrot (*Polytelis swainsonii*) and the larger size bird: short-billed black cockatoo (*Calyptorhynchus latirostris*). The red arrows indicate a barb from each species.

2.6 DNA amplification

Target DNA was amplified using the polymerase chain reaction technique (PCR). All PCRs were set up under UV cabinet which is equipped with UV lamps for sterilization. Any DNA present in the cabinet was destroyed by UV light exposure for about 15 minutes before and after use. A negative control and positive control reaction were set up for every PCR. A PCR product must not be produced in a negative control reaction to ensure that there is no contamination in the reaction mixture. A positive control reaction was set for checking that all PCR reagents were present in the mixture and at appropriate concentration. Muscle DNA from chicken (*Gallus* sp.) was used as a positive control. All PCRs were performed using of 5 units/µL GoTaq® Flexi DNA polymerase (Promega), 5 x colourless GoTaq Flexi Buffer (Promega), 25 mM $MgCl_2$ (Promega), 10 mM dNTP mix

(Promega), 10 µM forward and reverse primers, sterile water and DNA template. PCR reagents were combined in total volume 25 µL and the final concentrations of each component are found in Table 2.13.

Table 2.13: PCR components were combined as the following:

PCR reagents	Reaction mix (µL)	Final concentration
5x GoTaq Buffer	5	1x
25 mM $MgCl_2$	2	2
2 mM dNTPs	2.5	0.2
10 µM Forward Primer	1.5	0.6
10 µM Reverse primer	1.5	0.6
Sterile water	7	-
5 units/µL GoTaq	0.5	2.5 units
DNA template	5	< 250 ng*
Total volume	25	

***Usage Information recommend using template DNA <0.5 µg/50 µL**

The optimum annealing temperatures for each primer pair were found using gradient PCR in the "advance function" on the thermal cycler from Labnet, model: MULTIGENE Gradient. The PCR cycles of each primer pair are showed in Table 2.14.

Table 2.14: PCR cycles for each primer pairs

Primer pairs	Product size(bp)	PCR cycles
upF1 and R1	561	35* cycles at 95 °C for 30 seconds, 55 °C for 30 seconds, and 72 °C for 1.5 minutes
F2 and dnR2	598	35 cycles at 95 °C for 30 seconds, 60 °C for 30 seconds, and 72 °C for 1.5 minutes
upF1 and dnR2	1139	20 cycles at 95 °C for 30 seconds, 60 °C for 30 seconds, and 72 °C for 1.5 minutes and continue 15 cycles at 95 °C for 30 seconds, 60 °C for 30 seconds, and 72 °C for 2 minutes
Fc and Rc	1900	20 cycles at 95 °C for 30 seconds, 60 °C for 30 seconds, and 72 °C for 1.5 minutes and continue 20 cycles at 95 °C for 30 seconds, 60 °C for 30 seconds, and 72 °C for 2 minutes
Fc and RV-PC	1212	35 cycles at 95 °C for 30 seconds, 55 °C for 30 seconds, and 72 °C for 1.5 minutes
Fc and Rv	921	35* cycles at 95 °C for 30 seconds, 56 °C for 30 seconds, and 72 °C for 1.5 minutes
Fw and Rc	1067	35 cycles at 95 °C for 30 seconds, 57 °C for 30 seconds, and 72 °C for 1.5 minutes

***40 cycles for the DNA amplification from 2 barbs and 5 barbs.**

2.7 Gel Electrophoresis

PCR products were separated on a 1% agarose gel (BIO-RAD) with 5 µL of 10 mg/mL ethidium bromide (BIO-RAD) in 1 x Tris/Boric Acid/EDTA buffer (BIO-RAD) at 110 Volts for about 30 minutes. The size of PCR products were estimated by comparison with a 100 bp DNA ladder (Biolabs) or 1 kb EZ load™ (BIO-RAD) depending upon the size of the PCR product being separated. The marker was loaded as 0.4 µg per lane and 5 µL of the PCR products were mixed with 2 µL of Blue/ Orange 6 x loading dye (Promega) and loaded into each lane. The gels were photographed using a Gel Doc™ EZ Imager (BIO-RAD).

2.8 PCR product purification by using Gel extraction kit

The gel were visualised under an UV fluorescence analysis cabinet (SPECTROLINE, Model: CX-20). The gel slab at the position of the expected PCR product was cut and isolated from the gel using the QIAquick Gel Extraction Kit. Gel extraction steps are modified from the QIAquick Spin Handbook on page 25-26 [3].

The procedure for PCR product purification from the gel using the QIAquick Gel Extraction Kit

1. Excise the DNA fragment from the agarose gel with a clean, sharp scalpel.
2. Weigh the gel slice in a 1.5 microcentrifuge tube. Add 3 volumes of QG buffer to 1 volume of gel.
3. Incubate at 50 ^{o}C for 10 minutes. To help dissolve gel, mix by vortexing every 2-3 minutes during incubation.
4. Place a QIAquick spin column in a collection tube.
5. Apply the sample to the column and centrifuge for 1 minute.
6. Discard flow-through and collection tube.
7. Place the column in a new collection tube, add 500 µL of QG buffer and centrifuge for 1 minute to remove all traces of agarose.
8. Add 750 µL of PE buffer to the column, let the column stand for 5 minutes before centrifuging at 8000 rpm for 1 minute.

9. Discard flow-through and collection tube and place the column in a new collection tube and centrifuge at 13000 rpm for 3 minutes to dry the membrane of the column.
10. Place the QIAquick column into a clean 1.5 microcentrifuge tube.
11. Add 50 µL of EB buffer to the centre of the membrane, let the column stand for 5 minutes and then centrifuge for 1 minute.
12. The success of the gel extraction was tested by separating the purified DNA on the gel. Five µL of the purified DNA was mixed with 2 µL of loading dye, and then separated on 1 % agarose gel at 110 Volts for about 30 minutes.

2.9 DNA Sequencing

The purified DNA was sent to the Australian Genome Research Facility Ltd (AGRF) for sequencing. All samples were quantified using the NanoDrop 1000 Spectrophotometer (Thermo Scientific) and then were prepared for sequencing following the guide to AGRF sequencing service for Purified DNA (PD):

10 pmol of a primer* + 30-80 ng of purified DNA + sterile water (in total volume of 12 µL)

***Forward primer or Reverse primer**

All sequencing results were analysed using the Sequence Scanner Software from Applied Biosystems. The sequences were identified species using the Blast program (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.10 Avian mitochondrial DNA analysis

2.10.1 Translated mitochondrial DNA analysis

The complete mtDNA genome of selected bird species were translated to analyse each protein sequences of the COI, COII, COIII, cyt *b*, ND1, ND2, ND3, ND4L, ND4, ND5 and ND6 gene. The avian species that were chosen for the protein alignment are shown in Table 2.15.

The translated sequences were aligned using ClustalW from the online program http://www.genome.jp/tools/clustalw/. The phylogenetic tree of each gene was created using the program from http://itol.embl.de/

Table 2.15: A list of the avian species used for the protein alignment analysis. These species had complete mitochondrial sequences submitted on Genbank.

Scientific name	Common name
Lophura ignita	Crested Fireback
Phasianus versicolor	Green Pheasant
Falco peregrines	Peregrine Falcon
Accipiter gentilis	Northern Goshawk
Egretta eulophotes	Chinese Egret
Nipponia nippon	Asian Crested Ibis
Corvus frugilegus	Eurasian Rook
Vidua chalybeata	Steelblue Widowfinch
Haematopus ater	Blackish Oystercatcher
Arenaria interpres	Ruddy Turnstone
Podiceps cristatus	Great Crested Grebe
Gavia pacifica	Pacific Loon
Eudyptes chrysocome	Macaroni Penguin
Apteryx haastii	Great Spotted Kiwi
Phaethon rubricauda	Red-tailed Tropicbird
Fregata minor	Great Frigate bird
Procellaria cinerea	Grey Petrel
Emeus crassus	Eastern Moa
Anomalopteryx didiformis	Little Bush Moa
Pterocnemia pennata	Lesser Rhea
Rhea americana	Greater Rhea
Gallirallus okinawae	Okinawa Rail
Rallina eurizonoides	Slaty-legged Crake
Dromaius novaehollandiae	Emu
Casuarius casuarius	Southern Cassowary
Tinamus major	Great Tinamou
Eudromia elegans	Elegant Crested Tinamou
Aythya Americana	Redhead
Anas platyrhynchos	Common Mallard
Struthio camelus	Ostrich
Meleagris gallopavo	Wild Turkey
Coturnix japonica	Japanese Quail
Gallus gallus	Chicken

2.10.2 Phylogenetic tree reconstruction

2.10.2.1 MrBayes

MrBayes program estimates phylogenetic tree based on probability distribution of trees, using Bayes theorem and a simulation technique called Markov chain Monte Carlo (or MCMC) to approximate the posterior probability of trees [4-16].

Each sequence has to be a unique name/number and needs to be the exact same length. Any gaps must be filled in with a '-'. The file must be converted to NEX file format. After the calculation is complete, the results from the CON file were used for making a phylogenetic tree using the online program on the iTOL website. Bayesian phylogeny (BP) inferences with MrBayes v. 3.1.2 [5] implemented substitution models and calculated with jModelTest based on 24 models [17, 18]. The program was run for 10 million generations using four Markov chains sampled for every 100 generations by starting from a randomly selected tree. The calculations of a 50 % majority rule consensus tree and posterior probabilities for each split were performed after excluding the first 25000 sampled trees.

2.10.2.2 Molecular Evolutionary Genetics Analysis (MEGA)

In this study, MEGA [19] is a tool for multiple sequence alignment, phylogenetic tree reconstruction and computing the genetics distance between species (GDA). The methods for constructing phylogeny are Neighbour-Joining (NJ), Minimum Evolution (ME), Maximum Parsimony (MP) and UPGMA. Those methods were used for finding the true tree using 11 different models; Hasegawa-Kishino-Yano, Tamura-Nei, General Time Reversible, Number of differences, p-distance, Jukes-Cantor, Kimura 2-parameter, Tajima-Nei, Tamura 3-parameter, Maximum Composite Likelihood and LogDet (Tamura-Kumar). Each method to generate a phylogenetic tree was compared to determine which model can make a tree with fewer anomalies.

2.10.3 Mitochondrial DNA segments analysis

The complete mitochondrial genome, complete gene and partial part of the genes in 102 avian species were aligned and the resulting reconstructed the phylogenetic trees used to identify which part of the gene of interest is the most effective region for distinguishing closely related avian. The sequences from 102 individual species were aligned in three levels: 1) complete genome alignment, 2) complete gene alignment and 3) small fragment alignment. For small fragment analysis of each gene, all sequences were taken at 100 bp and 450 bp (350 bp overlap to the previous fragment).

2.10.4. Inter- and intra-species of the Fringillidae family (finches) at genus and species taxonomic level using partial sequences of the cyt *b*, COI and ND2 loci

The partial sequences of the ND2 gene obtained from this study are from *Fringilla coelebs*, *F. montifringilla*, *Carduelis chloris*, *C. carduelis* and *C. spinus*. The pairwise distance between finch species was calculated by MEGA 5 program using Kimura 2-parameter model. A list of finch species for intra-species variation study can be found in Table 2.16 and all the other sequences for both inter- and intra-species variation study of the cyt *b*, and COI loci are obtained from the DNA database.

Table 2.16: A list of finch species used for intra-species variation study.

Scientific name	Common name	Number of sequences		
		From this study	From the Database	
		ND2	cyt *b*	COI
F. montifringilla	Bramblefinch	6	4	6
F. coelebs	Chaffinch	5	6	6
C. chloris	European greenfinch	5	6	6
C. carduelis	Goldfinch	6	4	6
C. spinus	Eurasian sisken	6	5	6

2.10.5 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*, COI, ND2 and ND5 loci

The partial sequences of the ND2 and ND5 loci obtained from this study are from *Calyptorhynchus banksii*, *Calyptorhynchus funereus*, *Psittacula alexandri* and *Amazona ochrocephala*. The pairwise distance between parrot and cockatoo species was calculated by MEGA 5 program using Kimura 2-parameter model. A list of parrot and cockatoo species analysed for intra-species variation study can be found in Table 2.17 and all the other sequences for both inter- and intra-species variation study of the cyt *b*, and COI loci were obtained from the DNA database.

Table 2.17: A list of parrot and cockatoo species used for intra-species variation study.

Species	Sub-species	Common name	Number of sequences			
			From this study		From the database	
			ND2	ND5	cyt *b*	COI
Calyptorhynchus banksii	*samueli-1* *samueli-2*	Red-tailed Black Cockatoo	6	6	2	2
	naso-1 *naso-2*					
	macrorhynchus-1 *macrorhynchus-2*					
Calyptorhynchus funereus	*latirostris-1* *latirostris-2* *latirostris-3*	White-tailed Black Cockatoo	7	7	2	2
	Funereus-1 *Funereus-2* *Funereus-3* *Funereus-4* *Funereus-5*	Yellow-tailed Black Cockatoo				
Psittacula alexandri	-	Moustached Parakeet	5	6	5	Not on the database
Amazona ochrocephala	-	Yellow-crowned Amazon Parrot	6	4	6	6

2.10.6 Genetic variation avian mitochondrial DNA genes

The mitochondrial genome of 102 avian species consisted of 37 genes comprised 13 polypeptides, two rRNA genes, and 22 tRNA genes. All gene sequences were aligned using the MEGA program to calculate the genetic distance between these species. The p-distance was calculated using this model to estimate the genetic distance between different DNA sequences calculating the number of differences between two sequences divided by the sequence length.

The steps to enter the data into MEGA included: click on the distance tab, calculate overall mean of a standard comparison (p value) and add 1000 bootstrap repetitions to evaluate the variance as shown in Figure 2.7.

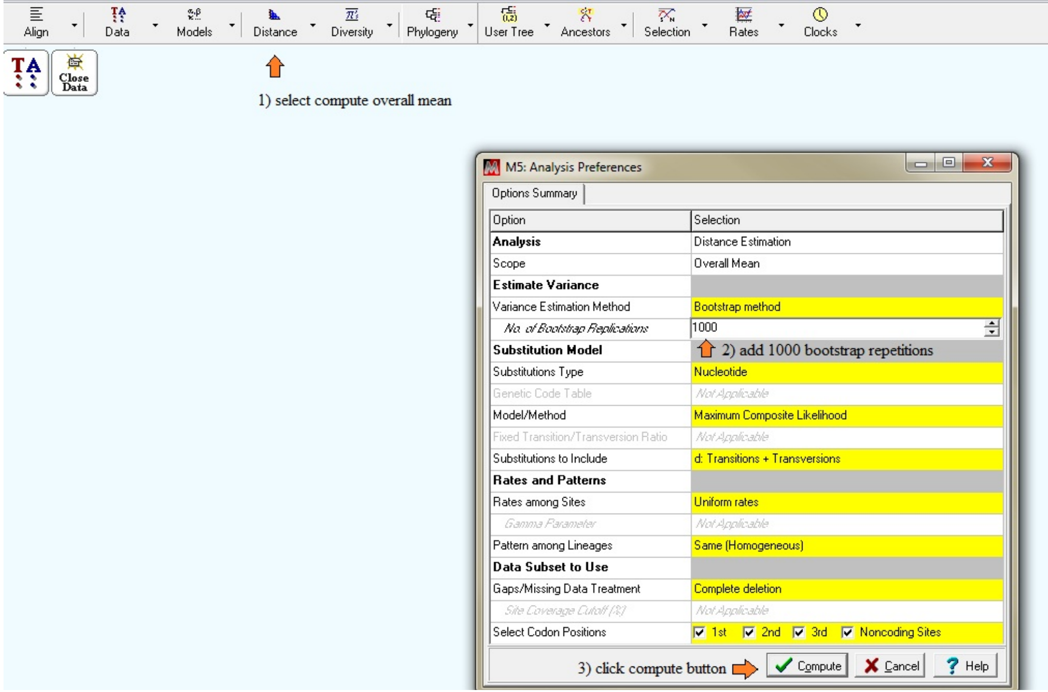


Figure 2.7: showing p-distance calculation using MEGA program. The orange arrows indicate the three steps to compute the overall mean value.

Chapter 2 References

1. QIAGEN, *QIAamp DNA Micro handbook.* QIAGEN, 2010.

2. QIAGEN, *User-Developed Protocol: Purification of DNA from nails, hair, or feathers.* QIAGEN, 2006.

3. QIAGEN, *QIAquick Spin Handbook.* QIAGEN, 2006.

4. Huelsenbeck, J.P. and F. Ronquist, *MRBAYES: Bayesian inference of phylogeny.* Bioinformatics, 2001. **17**: p. 754-755.

5. Ronquist, F. and J.P. Huelsenbeck, *MRBAYES 3: Bayesian phylogenetic inference under mixed models.* Bioinformatics, 2003. **19**: p. 1572-1574.

6. Metropolis, N., et al., *Equations of state calculations by fast computing machines.* J. Chem. Phys, 1953. **21**: p. 1087-1091.

7. Hastings, W.K., *Monte Carlo sampling methods using Markov chains and their applications.* Biometrika, 1970. **57**: p. 97-109.

8. Larget, B. and D. Simon, *Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees.* Mol. Biol. Evol, 1999(16): p. 750-759.

9. Li, S., *Phylogenetic tree construction using Markov chain Monte carlo.* Ph. D. dissertation, Ohio State University, Columbus, 1996.

10. Mau, B., *Bayesian phylogenetic inference via Markov chain Monte carlo methods.* Ph. D. dissertation, University of Wisconsin, Madison, 1996.

11. Mau, B. and M. Newton, *Phylogenetic inference for binary data on dendrograms using Markov chain Monte Carlo.* Journal of Computational and Graphical Statistics, 1997. **6**: p. 122-131.

12. Mau, B., M. Newton, and B. Larget, *Bayesian phylogenetic inference via Markov chain Monte carlo methods.* Biometrics, 1999. **55**: p. 1-12.

13. Newton, M., B. Mau, and B. Larget, *Markov chain Monte Carlo for the Bayesian analysis of evolutionary trees from aligned molecular sequences.* In Statistics in molecular biology (F. Seillier-Moseiwitch, T. P. Speed, and M. Waterman, eds.). Monograph Series of the Institute of Mathematical Statistics, 1999.

14. Rannala, B. and Z. Yang, *Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference.* J. Mol. Evol, 1996. **43**: p. 304-31.

15. Yang, Z. and B. Rannala, *Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte carlo method.* Molecular Biology and Evolution., 1997. **14**: p. 717-724.

16. Geyer, C.J. and P. in, *Markov chain Monte Carlo maximum likelihood.* Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface. (E. M. Keramidas, ed.). Fairfax Station: Interface Foundation, 1991: p. 156-163.

17. Posada D, *jModelTest: Phylogenetic Model Averaging.* Molecular Biology and Evolution, 2008. **25**: p. 1253 –1256.

18. Guindon S and G. O, *A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.* Systematic Biology, 2003: p. 52:696.

19. Tamura K, et al., *MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods.* Molecular Biology and Evolution, 2011. **28**: p. 2731-2739.