DNA Barcoding the Birds of New Zealand

A thesis presented in fulfillment of the requirements for the degree of
Doctor of Philosophy in Molecular BioSciences
at Massey University, Auckland, New Zealand

William John Waugh

2011

I dedicate this thesis to my parents Molly and Charles Waugh

both of whom died before it was complete.

I know that my achieving this degree would have

given them great pleasure.

No matter how much I imagined I would miss their

love, guidance and friendship,

nothing prepared me for the reality of the loss that I feel.

I also dedicate this thesis to my wife Fiona

who gave me the opportunity to undertake it and

the love and support needed to complete it.

Abstract

A comprehensive inventory of the life forms on earth is at the heart of any scientific study of evolution and biodiversity. The international "Barcode of Life" project is an attempt to identify the earth's biodiversity, at the species level, using short signature DNA sequences. The hypothesis underlying DNA barcoding is being comprehensively tested in different taxa. A database was constructed of DNA sequences from part of the mitochondrial gene cytochrome *c* oxidase subunit 1 for the avian fauna of New Zealand. To date, 833 sequences from 215 species have been added to this database, of which 628 sequences from 126 species are from native or endemic birds. This represents an average of 5 samples per species (minimum 1, maximum 18) for the latter group, which is the central focus of this thesis. Samples of species, from different geographical locations throughout New Zealand, have been collected to highlight any intraspecific nucleotide variation that may occur. Some samples analysed here were from historical specimens housed in museum collections and required specialised DNA extraction and amplification. These techniques were developed as part of the project and provide a means of collecting DNA barcodes where no modern material is available. In general, DNA barcoding proved effective at identifying avian species in New Zealand. However, some species were highlighted that contained distinct DNA barcode clusters, indicative of possible subspecies or cryptic species while in other cases two or more species that appear to be different share very similar DNA barcodes. Remains from aircraft birdstrikes were identified using this technique in order to inform wildlife management at airports around New Zealand. A review of and outlook for the uses of this technique are given.

Acknowledgements

Firstly, I thank my supervisors Professor David Lambert and Dr Craig Millar for providing me with a very interesting PhD project. This research has not only challenged me intellectually but also forms part of what is becoming regarded as an effort to speed the identification of species and, thereby increase global biological literacy. As someone who has taught biosciences for many years, this is a project tailored to augment my life's work. I thank them for providing guidance and resources, discussions and assistance with manuscripts. I also thank them for obtaining funding for this project and relieving me of the burdens relating to financial issues. I thank the funding bodies, the Allan Wilson Centre and Massey University for providing a stipend and tuition fees for this PhD.

I am indebted to a number people and institutions for providing samples for this project. Sylvia Durant provided many fresh samples from both live and dead specimens, as did Rosemary Tulley, Della Bennet, Bill Pohatu, Jon Eyely, Sondra Taylor, Nick Petkov and Phil Gotlieb. In addition, I thank my supervisors for providing many fresh bird samples from their collections as well as Brian Gill of Auckland War Memorial Museum, Ian Jamieson from Otago Museum and Janeatte Norman of Victoria Museum. Particular thanks to Paul Scofield for providing preserved samples and for assisting me to collect them from the Canterbury Museum's collection.

I am most grateful to Andrew Dodd, Jennie Hay, Jennifer Anderson and Leon Huynen for introducing me to the mysteries of molecular genetics and for their patience and forbearance with me during the early part of my research. Their wisdom and help has served me well throughout this project. I am further indebted to Andrew Dodd and Tamara Sirey for their help in all matters pertaining to data analysis and computing technology. My own levels of understanding in these matters have been elevated to a completely different plane thanks to their unstinting help and guidance. Their tolerance of nothing but the very highest standards has provided me with the best possible example.

I thank David Lambert, Craig Millar, Selina Patel and Allan Baker for all their work helping to build the New Zealand Birds DNA barcode database and for sharing their insights and wisdom with me on everything from the minutia of molecular techniques to the broad strategies for developing this project.

I am grateful to Paul Hebert, Mark Stoeckle, Sujeevan Ratnasingham, Gregory Downs, Riadul Mannan Riad, Janet Topan, Megan Milton, Brianne St. Jacques, Claudia Kleint-Steinke, Natalia Ivanova, Andras Papp, Alex Borisenko, Taika von Königslöw and others in Canada at the DNA barcoding facilities at the University of Guelph, Ontario for helping to develop the Birds of New Zealand project and in assisting in building the database that now resides on the Barcode of Life Data Systems (BOLD) website.

When examining specimens from aircraft birdstrikes, I received invaluable help and data from Max Evans of the New Zealand Civil Aviation Authority, advice and the opportunity to attend a Flight Safety Forum, which he chaired, from John Martin of Air Nelson, birdstrike samples from Della Bennet, Wildlife Control Officer at Christchurch International Airport and, Jon Eyely, Airside Operations Manager as well as Bill Pohatu, Sondra Taylor, Nick Petkov and Phil Gotlieb, Wildlife Control Officers at Wellington International Airport. I also received much helpful advice from Carla Dove of the Smithsonian Institute in Washington, USA.

I thank my colleagues, past and present, at the Allan Wilson Centre/Institute of Natural Sciences in Albany: Hayley Lawrence, Andrew Dodd, Jennie Hay, Leon Huynen, Elmira Mohandesan, Martina Dautel, Katie Hartnup, Gabrielle Beans-Pecon, Jyothsna Visweswerіah, Andrew Cridge, Chris Rodley, Tim Heupink, Bill Peacock, Barbara Binney, Subashchandran Sankarasubaramanian, Muharam Khoussainova, Bupanelo Mbuba, Sujung Lee, Jarod Young, Saumya Agrawal, Diane Brunton, Gabriele Schmidt-Adam, Justin O'Sullivan, Evelyn Sattlegger, Wayne Patrick and others too numerous to mention for their companionship, conversation and company as well as for their help, guidance and support throughout this

research.

Finally I thank my wife Fiona and my two daughters Sarah and Rebecca for providing me with much love, tolerance and support throughout the years of this research, without which it would never have been completed

Thesis Structure, Financial Support and Regulatory Compliance

This thesis begins with a general introduction (Chapter 1), which is a review of DNA barcoding and provides the background and intellectual framework that underpins the thesis. It is an update of a review paper published in BioEssays. Chapter 2 is a general methods section outlining the collection and processing protocols associated with the project. Chapters 3 and 6 have been written as stand-alone scientific papers. Therefore some information provided in the introduction will be briefly outlined again in these chapter introductions. Chapter 4 is an analysis of the data. Chapter 5 outlines and discusses interesting anomalies within the data and Chapter 7 discusses the variety of uses that DNA barcoding has and may be put to and the conclusions and applications of the research findings and potential future research.

Financial Support

Funding for this project was provided by: the Allan Wilson Centre for Molecular Ecology and Evolution, Massey University and the University of Auckland. Personal financial support was provided through a Massey University Doctoral scholarship and an Allan Wilson Centre scholarship.

Ethics Approval

This research had ethics approval from the New Zealand Department of Conservation (DOC) Animal Ethics Committee for sampling.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

List of Tables

Chapter 1

General Introduction to DNA Barcoding

1.1 A brief history of modern taxonomy

The identification and characterisation of living things is fundamental to biological science. Modern binomial taxonomy, founded on principals first proposed by Carl Linnaeus (1707-1778), originates with the publication of the first edition of his *Systema Naturae* in 1735. During a lifetime dedicated to the academic study of living things he published many further editions as well as many other publications, which are generally recognised as the first modern systematic classification of plants and animals. By the tenth edition of *Systema*, he had described more than 12000 species of plants and animals.[1] Since these origins in the mid 18th century, taxonomists have described a further 1.7 to 1.8 million taxa to species level.[2]

Linnaeus's three kingdoms, divided into classes, genera and species,[1] have given way to the current five kingdoms, divided into phyla, classes, orders, families, genera and species.[3] Most of these constructs are measures of phylogenetic relatedness rather than actual biological entities and provide a structure by which, the whole of the living world is catalogued. Species alone have a physical manifestation, which can be examined, sampled and tested and species are still named according to his binomial system.

For much of the 233 years since Linnaeus's death, taxonomists have collected specimens in the field and then painstakingly catalogued and described them according to morphological, anatomical and behavioural characteristics. Museums have acquired huge collections of preserved specimens, which serve as "types" for their species, to be used to identify other specimens.

Unsurprisingly, larger animals have generally been the first to be described, while many smaller organisms remain unknown to science.[4] For example, it is thought

that fewer than 10% of the vertebrates are yet to be identified but within the nematodes a bafflingly large number of species may exist, the vast majority of which have not yet been identified.[5] Moreover, even among larger animals, there are doubts about species identification. Recently, the African elephant, long thought to be a single species, has yielded to modern taxonomic techniques and appears to incorporate two or even three separate species.[6]

Publication of a new species is a slow and complicated business, with much of the effort required being an exhaustive search through publications, periodicals, journals, monographs and books, to determine whether this is indeed a new species and that no other taxonomist has already described and named it. Having established that this species has not previously been described, a holotype (a typical adult example of the species) and some paratypes (other examples that display typical variation found within the species) must be deposited with a museum or other major, recognised institutional collection, where they will be preserved for description verification or research using the original specimens. A description of the new species is then required to be published in an appropriate journal, along with a name selected by the taxonomist according to rules embodied in codes, such as the International Code on Zoological Nomenclature (ICZN),[7] the International Code of Botanical Nomenclature (ICBN)[8] and the International Code of Nomenclature of Bacteria (ICNB)[9].

From the beginning of the 19th century until the early part of the 20th, taxonomy enjoyed a heyday amongst biologists as well as amateurs; however, since the mid 20th century it has been in decline. Today it is estimated that there are approximately 6000 taxonomists actively working in this field of science.[10] Their combined efforts and those of their forebears have, on average, lead to the description of 7000-8000 species per annum since Linnaeus's time.

Taxonomy is, of necessity a specialist field, in which individual scientists develop areas of expertise that are not duplicated around the world. With the death of each taxonomist, the world expert on a particular group of species dies, taking with them

a fund of useful information. In many cases, there has been no other taxonomist working in even remotely related fields and thus, much taxonomical knowledge is lost. The advent of the internet has improved communication and led to the construction of websites aimed at collating taxonomical information and making it generally available.

1.2 The size of the taxonomic task

Estimates of the earth's eukaryotic biota range from an improbably low 3.6 million to an equally improbable 100 million species.[10] Any estimate is no more than an order of magnitude guess and perhaps the most commonly suggested is 10 million. The identification of numbers in this range is an insurmountable workload for taxonomists using current methods. At the present rate of species description, this represents over a thousand years of work. Even allowing for improvements in communications and the impact of the internet, the task is overwhelming.

Clearly there is a need to accelerate and simplify the processes of identification involved and, because of the scale of the problem, new methods will have to be employed. In addition, as more species are described, accessing the enlarged pool of taxonomic knowledge will become even more problematical.

Field biologists, faced with the reality of species diversity, recognise the inadequacy of their own ability to access what *is* known about the biota, let alone what is not. These problems also impact upon people working in other areas such as combating the trade in endangered species, monitoring fisheries, identifying and controlling the spread of pest species or pathogens, identifying extinct lineages and regulating the movement of biological material around the world as well as identifying the remains of living things whose morphology is not apparent.[11-14]

Moreover, phenotypic plasticity and genotypic variation in the features used for identification easily lead to identification errors, with cryptic species or differing life stages adding to the confusion.[15] To compound matters, the task of cataloguing

extant species is lent urgency by currently observed mass extinctions that are widely believed to be anthropogenic in origin.[16, 17]

1.3 Potential solutions to the taxonomic impediment

Recent developments in DNA sequencing technology have introduced the possibility of using variations in short sequences of DNA as labels for species in a process that has become known as DNA barcoding. This DNA label has the potential to reduce the time associated with literature searches necessary in a taxonomic description. The concept has already gained considerable acceptance among those working with species refractory to morphological identification such as viruses,[18] bacteria,[19, 20] protists[21] and Rhodophyta.[22] However, it is apparent that, since morphological techniques are difficult to apply without considerable training, some more rapid system of species identification is required for all taxa. This has lead to the formation of the Consortium for the Barcode of Life (CBOL),[23] which aims to provide such a DNA barcode for every eukaryotic species on the planet.

CBOL is an international organisation devoted to developing DNA barcoding as a global standard in taxonomy. It comprises more than 200 member organisations from more than 50 countries and includes museums, herbaria, zoos, research organisations, governmental and intergovernmental agencies as well as other organisations involved in taxonomic research and biodiversity issues.[23]

Members agree to submit their DNA barcode sequences and voucher specimen data to a public database. CBOL was launched in May 2004 and is overseen by an executive committee that reports to the member organisations. It has four working groups to develop particular aspects of DNA barcoding, an Implementation Board and a Secretariat Office to conduct its business. Within the auspices of CBOL a number of initiatives have been established including the All Birds Barcode Initiative (ABBI),[24] the Fish Barcode Initiative,[25] the Sponge Barcoding Project,[26] the All Leps (Lepidoptera) Barcode Initiative[27] and the International Network for Barcoding Invasive and Pest Species.[28] Another group is exploring the barcoding

of endangered vertebrates. Each of these initiatives aims to obtain DNA barcodes for every species within its group.[23]

The Canadian Centre for DNA Barcoding (a member of CBOL) oversees a website, the Barcode of Life Data Systems (BOLD),[29] that permits the uploading of sequences from the 5' region of the COI gene and returns a species-level identification when one is possible. At present, the site has more than 1200000 sequences from over 100000 species and these numbers are increasing steadily. The site also permits a variety of forms of data analysis for submitted sequences. CBOL works in cooperation with a number of other organisations including the Global Biodiversity Information Facility (GBIF),[30] National Centre for Biotechnology Information (NCBI)[31] and many taxonomic communities and web based projects.

1.4 DNA barcoding

A DNA barcode is a short sequence of nucleotides, taken from an appropriate part of an organism's genome that is used to identify the organism to species level. Intraspecific variation in this sequence is an order of magnitude less than that observed interspecifically and this provides the means by which species are differentiated.[32] It is not part of an exclusively DNA based taxonomy nor is it a tool for phylogenetic reconstruction. It simply provides a means of linking sample specimens directly to existing voucher specimens and taxonomical information.

Central to the efficacy of DNA barcoding is the selection of a suitable segment of DNA. Its mutation rate must be slow enough so that intraspecific variation is minimised but sufficiently rapid to highlight interspecific variation.[15] It must be relatively easy to collect and should have as few insertions or deletions as possible to facilitate sequence alignment. Mitochondrial DNA (mtDNA) offers several advantages over nuclear DNA. According to Drake's observation,[33] the rate of DNA mutation is inversely related to the size of the genome. Hence, nuclear DNA undergoes relatively slow mutation compared with mtDNA and, for this reason, would require a much longer nucleotide sequence than is necessary with mtDNA

in order to provide a barcode capable of differentiating species. Inefficient mtDNA repair mechanisms have also been suggested as a reason for high substitution rates.[34]

Generally, mtDNA occurs in animals as a single double-helical circular molecule containing 13 protein-coding genes, 2 ribosomal genes, a non-protein-coding control region, and several tRNAs (Figure 1.1).[35] Each mitochondrion contains several such circular molecules and, therefore, several complete sets of mitochondrial genes. Furthermore, each cell has between 100 and 10000 mitochondria. Thus, when sample tissue is limited, the mitochondrion offers a relatively abundant source of DNA.

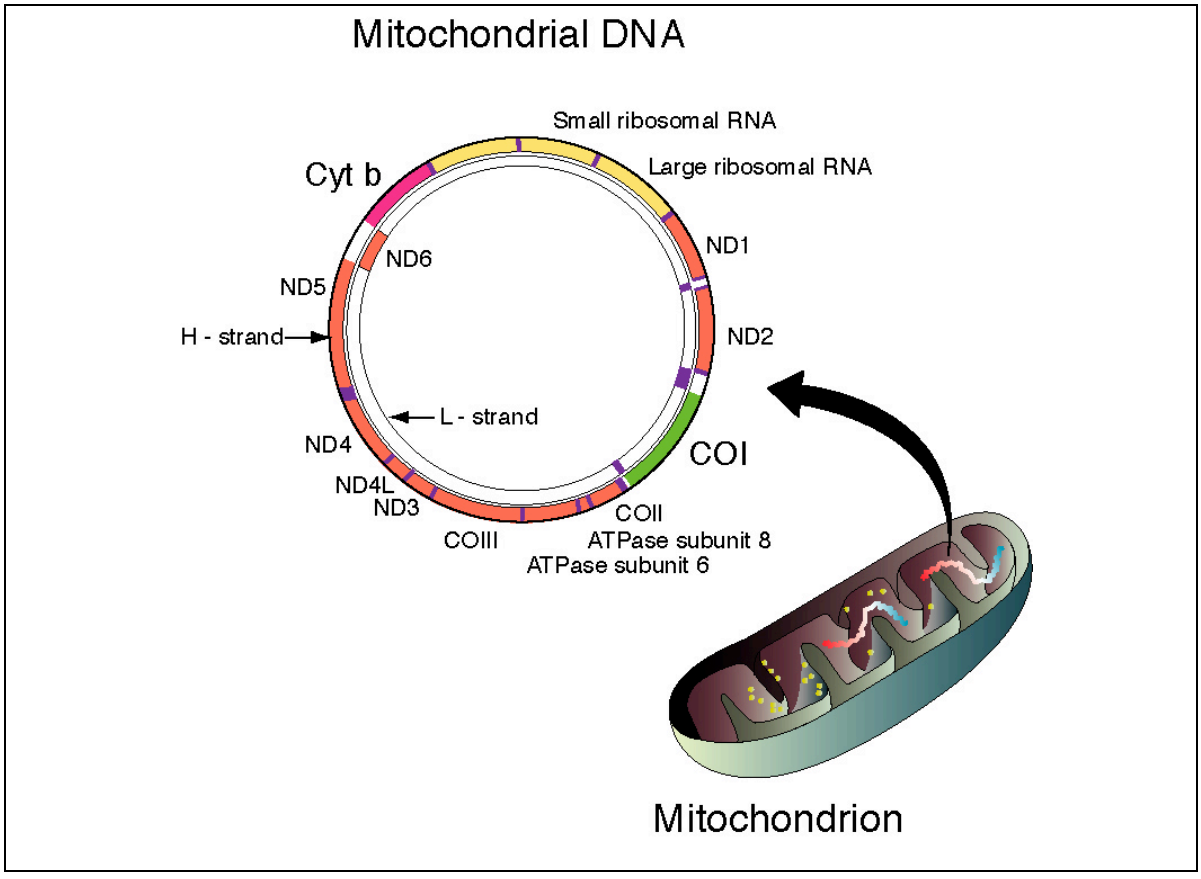


Figure 1.1 The mitochondrial genome of a eukaryote

Cytochrome *c* oxidase is a large transmembrane protein found in the mitochondrion, which is highly conserved across species that employ oxidative

metabolism. It functions as the terminal electron acceptor in the respiratory chain, catalysing the reduction of oxygen to water and pumping protons across the membranes of the cristae.[35, 36] The protein comprises several subunits of nuclear origin and three subunits synthesised in the mitochondrion. The mitochondrial subunits are known as subunits I, II and III.

Cytochrome *c* oxidase subunit I (COI), the catalytic subunit of the enzyme, is predominantly imbedded in the membrane of the mitochondrial crista (Figure 1.2). This structure would indicate a significant level of structural and functional constraint. However, the nucleotides of the gene that codes for it show sufficient variation to differentiate between species. Intraspecific variation in this gene is generally <10% of that observed between species. Moreover, insertions and deletions are rare.[5] Recent studies associated with CBOL have generally selected a 648 bp segment of the COI gene, starting from base 58 at the 5' end, to generate a suitable barcode[23] and this particular segment is referred to throughout this thesis as the "DNA barcode".

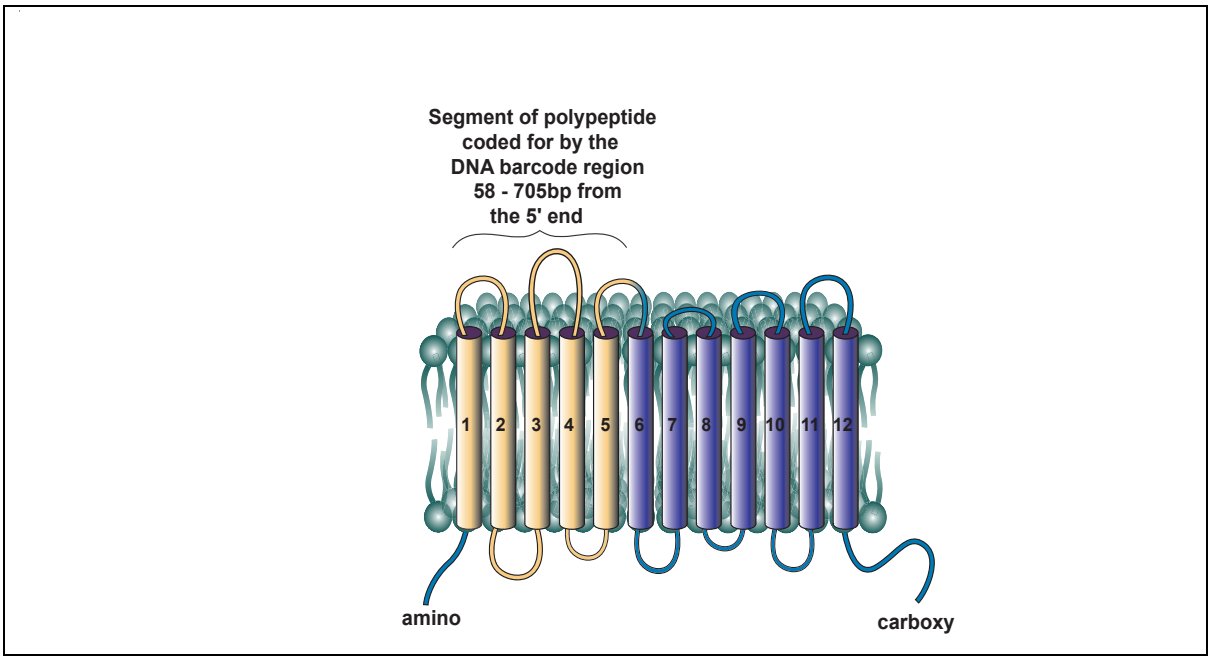


Figure 1.2 A diagrammatic representation of cytochrome *c* oxidase subunit I (COI) imbedded in the membrane of the crista of the mitochondrion, indicating the location of the region used for the DNA barcode.

The appropriate segment of DNA selected for analysis, is then amplified using the polymerase chain reaction (PCR). The amplified segment of the COI gene is sequenced and this sequence, the ‘‘barcode’’, is then matched with existing barcodes from vouchered specimens. The Kimura 2-parameter (K2P) genetic distance correction[37] (see Chapter 4) is used to quantify sequence divergences among individuals because it is the most effective model when distances are low, as is the case with COI barcoding.[15]

1.5 A review of the efficacy of DNA barcoding in published studies

To date, the literature contains a considerable number of fully published studies in which DNA barcoding in animals, using part of the COI gene, has been assessed[15, 38-43] or used to help resolve taxonomic ambiguity[39, 44-92][93-98] or as part of a taxonomic description[84, 99-119] or to link different life stages[44, 49, 120-124] or to examine the contents of foodstuffs[45, 125-127] (Table 1.1). To date, fourteen papers have been published on the DNA barcoding of bird species. Of these, four included databases of bird barcodes for North America,[74, 128] Korea,[119, 129] Argentina[130] and Scandinavia.[131] In addition, DNA sequencing technology has been used for identifying organisms from other Kingdoms including plants,[132-135] bacteria,[19, 20] protists[21, 57, 136-138] and viruses.[18]

Although minor variations occurred in protocols between studies, broadly similar methods were used in each. DNA extracted from tissue samples was prepared using a variety of standard protocols. The sequence blocks of DNA used ranged in length from 350bp to ~1000bp. DNA sequences were then generated using automated sequencers and these were compared both within and between species.

Table 1.1 Design details and major results from selected DNA barcode studies, which utilised part of the cytochrome c oxidase subunit 1 (COI) gene to identify animal taxa. **K2P** = Kimura 2-Parameter

Reference	Sequence length	Taxa (No. and name)	% identified	K2P % intra- *vs* interspecific COI variation	Comment
Ahrens et al. 2007[44]	1600bp	24	92.7		Comparing adult and larval stages of chafer beetles
Armstrong et al. 2005[139]	*ca* 650bp	57 tussock moths 81 fruit fly	93 and 100		Identifying invasive species of tussock moths and fruit flies
Baldwin et al. 2009[49]					Identifying larval fish (Phaeoptyx and Astrapogon, Teleostei: Apogonidae) from Belize
Ball et al. 2005[50]					Identifying of mayflies using DNA barcodes
Ball et al. 2006[51]	617bp				Identifying tussock moths.
Ball et al. 2007[52]	611bp	2		0.3 *vs* 12.8	Used to identify sooty beech scale insects that are otherwise difficult to identify
Banks et al. 2007[53]	647bp	3 oystercatcher spp.	0		These putative species show very little nucleotide difference between species and may be a single species
Barrett et al. 2005[54]	600bp	203 arachnid spp.	100	1.4 *vs* 16.4	Mean intra- and interspecific nucleotide divergences did not overlap except in the case of probable cryptic species.
Borisenko et al. 2008[99]					DNA barcoding used to survey small mammal communities in Surinam.
Brown et al. 2003[100]	648bp	1 lepidopteran sp. *Gnetom gnemon*	100		COI was used as a taxonomic descriptor for a new species.
Campbell et al. 2008[55]					Identifying 'extinct' freshwater mussel species using DNA barcoding.
Cesari et al. 2009[101]					DNA barcoding of *Macrobiotus macrocalix*
Chang et al. 2009[56]	535bp	42 Oligochaet worms		1.3 *vs* 18.7	identifying earthworms

Chantangsi et al. 2007[57]					Barcoding ciliates: a comprehensive study of 75 isolates of the genus Tetrahymena
Chaves et al. 2008[102]					DNA barcoding Brazilian tyrant-flycatchers (Passeriformes: Tyrannidae)
Clare et al. 2007[103]					DNA barcoding of Neotropical bats: species identification and discovery within Guyana.
Costa et al. 2007[58]	650bp				DNA barcoding Crustacea
Dabert et al. 2008[59]					DNA barcoding used to identify feather mites
Day et al. 2008[104]		1 dipteran spp.	100	0.67-0.78 *vs* 7.06	Confirmation of the species status of the blackfly *Simulium galeratum* in Britain using DNA barcoding.
De Astarloa et al. 2008[105]	648bp	1 Alasmabranch spp.	100	*vs* >3	DNA barcoding used as supporting evidence for a new species of longnose skate
Decaens et al. 2008[106]		2 lepidopteran spp.	100	*vs* 5.6-6.6 & 6.7-12.5	Descriptions of two new species of Hemileucinae (Lepidoptera: Saturniidae) based on morphology and DNA barcodes
De Lay et al. 2005[60]	>500bp	60 Nematode worm spp.			Used to identify Nematode worms
Dooh et al. 2006[61]	635bp and 601bp	4 crustacean spp.		1.5 vs 27	Using barcodes to examine the phylogeny of two glacial relict crustacean taxa in North America
Dove et al. 2008[47]	648bp	821 avian samples	96.6		Using DNA barcodes to identify birdstrike material
Efe et al. 2009[62]	684bp	2 avian spp. from the Laridae			Identifying cryptic species within the Sandwich tern complex
Elias-Gutierrez et al. 2008a[63]		61 cladoceran & 21 Copepoda spp.	100	0.82 & 0.79 *vs* 15-20 x as much	DNA barcodes for Cladocera and Copepoda from Mexico and Guatemala
Elias-Gutierrez et al. 2008b[64]		2 cladoceran spp.		? *vs* 14.3%	A new cryptic species (Crustacea, Cladocera, Chydoridae) from the Mexican semi-desert region, highlighted by DNA barcoding.
Elsasser et al. 2009[65]		2 Nematode spp.			Identifying North American guinea worms (Nematoda: Dracunculus) using DNA barcoding
Emery et al. 2009[66]		1 lepidopteran sp			DNA barcoding and morphological analysis to identify specialist floral parasites (Lepidoptera: Coleophoridae: Momphinae: Mompha)
Ferri et al. 2009[67]		Nematode worms			DNA barcoding for the identification of filarioid worms and related parasites.
Fisher et al. 2008[68]		8 ant spp.			A revision of Malagasy species of *Anochetus mayr* and *Odontomachus latreille* based on DNA barcodes.

Foley et al. 2007[69]		17 spp. of dipterans from the genus *Anopheles*			DNA barcoding used to help identify a species-rich anopheline complex
Foottit et al. 2008[70]	658bp	300 aphid spp.	96		DNA barcoding used to identify >300 species of aphid from 130 genera
Foottit et al. 2009[71]	658bp	17 spp. of insects from the Adelgidae			DNA barcoding used to help identify small cryptic insects (*Adelgidae*)
Gomez et al. 2007[72]		1 bryozoan spp. *Celleporella hyalina*			DNA barcoding reveals cryptic speciation of a marine bryozoan taxon confirmed by mating trials
Greenstone et al. 2005[124]	439bp	32 Carabidae spp. and 39 Araneae spp.			COI used to identify species and link different life stages.
Hajibabaei et al. 2006a[39]	311-612bp	521 lepidopteran spp.	97.9%	0.17-0.46 *vs* 4-6	Morphologically distinct sympatric species from three families identified.
Hajibabaei et al. 2006b[40]	100bp	lepidopteran and hymenopteran spp.	100% 93.8%	0.1-0.4 *vs* 5.4-8.9	Using mini-barcodes to identify species
Hajibabaei et al. 2006c[73]					DNA barcodes used to identify primate species
Hebert et al. 2003a[15]	658bp	7 animal phyla 8 insect orders 200 Lepidopteran spp.	96.4 100 100	 Lepidoptera 0.25 *vs* 6.84	The efficacy of COI in identifying species, orders and phyla assessed
Hebert et al. 2003b[140]	>400bp	2238 Annelida, Arthropoda, Chordata, Cnidaria, Echinodermata, Mollusca, Nematoda, Platyhelminthes, and minor phyla	>98	Overall, usually <2 *vs* 11.3	The efficacy of COI in identifying species from eight major and several minor phyla plus a variety of arthropod classes was assessed. Cnidarians showed less COI variation between species than all other taxonomic groups, 94.1% vs 1.9% showing <2% K2P between spp. ($p < 0.0001$).

Hebert et al. 2004a[120]	648bp	10 lepidopteran spp.	100		Ten new taxa identified. Different life stages matched for species
Hebert et al. 2004b[74]	648bp	260 Avian spp.	100	0.43 *vs* 7.93	Four possible cryptic species identified
Hogg et al. 2004[76]	710bp	19 Collembola spp.	100	0.78 *vs* 19.0	Produces high resolution in Collembola species. Possible cryptic species identified.
Hogg et al. 2006[75]	545bp	1? New Zealand Amphipod spp			Divergent lineages of the widespread New Zealand amphipod *Paracalliope fluviatilis* revealed using allozyme and DNA barcode analyses
Holmes et al. 2009[45]		20 shark spp. 7 ray spp.	91.5		Identifying shark species from dried fins for conservation purposes
Hu et al. 2002[77]	≤530bp	7 hookworm spp.	-		Three of the seven species appeared to be possible species complexes based on intraspecific COI variation.
Hu et al. 2005[78]	≈450bp	3 *Progamotaenia* spp. (Platyhelminthes)			Variation at COI suggests that all three species are species complexes
Huang et al. 2007[79]	640bp	28 earthworm spp.	100	<1 *vs* >15.8	Identifying Chinese earthworm species
Hulcr et al. 2007[141]	661bp	*1 Homona mermerodes*	100		Confirming phylogeny of a lepidopteran
Janzen et al. 2005[107]	648bp	~350 spp. lepidopterans	97		Using DNA barcoding to help characterise a large lepidopteran fauna
Johnsen et al. 2010[131]	648bp	296 bird spp.	94	0.24 *vs* 7.95	DNA barcoding Scandinavian birds
Johnson et al. 2008[81]	1000bp	20 taxa within the molluscan genus Lepetodrilus			DNA barcoding of Lepetodrilus limpets reveals cryptic species
Kelly et al. 2007[82]		19 Mollusc spp.			DNA barcoding using chitons (genus Mopalia)
Kerr et al. 2007[128]	X̄ 658bp	546 avian spp.	>95	0.23 *vs* 4.3	Barcoding the birds of North America
Kerr et al. 2009[130]	>550bp X̄ 692bp	500 avian spp.	98.8	0.24 *vs* 6.2	Barcoding birds of Argentina

Lambert et al. 2005[142]	596bp	10 Moa spp.	100	1.24 *vs* 1.3-3.8	COI used to identify extinct species. Possible species synonymy highlighted.
Landry 2007[83]		6 lepidopterans		X̄ 0.16 *vs* 8.1	Barcoding confirms species status of the leek moth genus *Acrolepiopsis* in North America
Langhoff et al. 2009[84]	468bp	12 lepidopteran spp.			DNA barcoding of the endemic New Zealand leafroller moth genera
Linares et al. 2009[109]	648bp	8 lepidopteran spp.	100	0.0043 *vs* 0.0862	DNA barcoding butterflies of Madagascar
Lorenz et al. 2005[110]	~727bp	56 primates		0.011 *vs* -	Problems with taxon specific patterns of 'universal primer' failure; Taxon specific primers developed
Lukhtanov et al. 2009[111]		353 lepidopteran spp.			DNA barcoding Central Asian butterflies
Marra et al. 2009[143]	650bp	*Branta Canadensis*	99-100		DNA barcoding used to identify birds responsible for the crash of US Airways flight 1549
Martinez et al. 2008[85]		*Allorhogus mendocinus*			DNA barcoding used to reassign known species (*Bracon mendocinus*) to a new genus
Mikkelsen et al. 2007[41]		*4 bivalve molusc genera*			DNA barcoding found to be useful for spp. identification within these taxa
Moura et al. 2008[86]		*2 genera of elasmobranchs*			DNA barcoding used to resolve within genera identification problems in deep water sharks for conservation purposes
Nelson et al. 2007[87]	658bp	*9 blowfly spp.*		0.097 *vs* 6.499	Using COI barcodes to identify forensically and medically important blowflies
Ngarmamonpirat et al. 2005[144]	450bp	1 *Gnathastoma spinigerum* (hookworm)	-		COI barcode variation did not match the morphological variation observed in 3rd larval stage of this hookworm
Nolan et al. 2007[88]		Ephemeroptera		0.034 *vs*	DNA barcoding useful for identifying spp. within the New Zealand damselfly genera
Park et al. 2011[129]	648bp	154 bird spp.	98.7	0.2 *vs* 5.0	DNA barcoding Korean birds
32 Paquin et al. 2004[121]	≈1000bp	23 *Cicurina* spp. (Arachnida)	≈100	1.09 *vs* 7.12	Immature specimens identified to species level. Possible species synonymy identified.
Penton et al. 2004[89]	709bp	2 *Daphnia* spp. (Crustacea)	100		Identification of morphologically cryptic species with overlapping distribution

Pfenninger et al. 2007[90]	596bp	7 Chironomus spp.	100		Identifying morphologically cryptic Dipteran species fronm the genus *Chironomus*
Quiroz-Vasquez et al. 2009[112]		Cladoceran spp		*vs* 12.02	DNA barcoding highlights new spp of freshwater cladoceran
Radulovici et al. 2009[91]		Marine crustaceans			DNA barcoding used to identify 80 malacostracan species from the Estuary and Gulf of St Lawrence
Rasmussen et al. 2009[125]	302-652bp	8 spp. of salmon and trout	100	0.04-1.09 *vs* 3.42-12.67	DNA Barcoding of Commercially Important Salmon and Trout Species (Oncorhynchus and Salmo) from North America
Remigio et al. 2003[92]	672bp	70 Gastropod spp.	-		COI used to identify species and higher taxonomic relationships; insertion or deletions more common in COI in this taxonomic group
Rock et al. 2008[113]					DNA barcodes of fish of the Scotia Sea, Antarctica
Scheffer et al. 2006[42]	550bp	7 invasive leafminer spp.	95	<1.9 *vs* > 12.2	DNA barcoding applied to invasive leafminers (Diptera : Agromyzidae) in the Philippines
Seabrook-Davison et al. 2009[94]		3 Quail spp.	100		Ancient DNA resolves identity and phylogeny of New Zealand's extinct and living quail (Coturnix sp.)
Sheffield et al. 2009[114]			100		DNA barcoding of a regional bee (Hymenoptera: Apoidea) fauna
Shirak et al. 2009[115]	619bp	10 cichlid spp.			DNA barcoding of Israeli indigenous and introduced cichlids
Sinclair et al. 2008[95]		13 choronimid spp.	100		Discrimination of *Cricotopus* species (Diptera: Chironomidae) by DNA barcoding
Skevington et al. 2007[93]	658bp		100		DNA barcoding used to identify big-headed flies (Diptera : Pipunculidae)
Smith et al. 2006[145]	658bp	32 dipteran spp.	100	0.17 *vs* 5.78	Fifteen cryptic species found using COI.
Smith et al. 2008[126]	520-621bp	16 NZ fish spp.			DNA barcoding for the identification of smoked fish products
Steinke et al. 2009[116]	X̄ 645bp	391 fish spp.	100	0.42 *vs* 10.81	Producing a DNA database of ornamental fish
Tavares et al. 2008[117]	650bp	60 avian spp.	100		Identification of sister species using DNA barcodes
Vargas et al. 2009[118]		5 sea turtle spp.	100		DNA barcoding of Brazilian sea turtles (Testudines)

Vences et al. 2005[43]	550-650bp	9 Mantellid frog spp. 4 *Aneides* spp. (salamanders)		5.4 (mantellid frogs) and 4.3 (salamanders) vs 20.7 and 13.5	Found high intraspecific variability (7-18%); the use of mitochondrial 16S rRNA gene to supplement COI suggested
Victor et al. 2009[122]					Identification of the larval and juvenile stages of the Cubera Snapper, *Lutjanus cyanopterus*, using DNA barcoding
Ward et al. 2005[146]	655bp	207 fish	100	0.39 vs 9.93	Efficacy of COI at identifying species and higher taxonomic relationships assessed; possible cryptic species identified
Waugh et al 2010[147]	648bp	40	90		Identifying the birds involved in birdstrike
Webb et al. 2006[123]		14 marine invertebrate larvae			Using DNA barcodes to identify marine invertebrate larvae.
Whiteman et al. 2004[97]	379bp	2 Lice spp. (Insecta)	100		Identifying bird parasites with similar morphology
Witt et al. 2006[148]	637bp	3 Amphipod spp.		3.35 *vs* ≤35.2	Cryptic diversity in amphipods studied using DNA barcodes.
Wong et al. 2008[127]	650bp	96 unknown seafood samples	100		Authentication of seafood contents using DNA barcodes.
Yoo et al. 2006[119]	749-751bp	92 avian spp.	100	0.3 *vs* 7.9	DNA barcoding of Korean birds

1.5.1 Efficacy at identifying species

DNA barcoding provided a high degree of taxonomic resolution (>95%) for most species examined in the studies reviewed. Sequence differences were generally much less within species than sequence differences between species. For example, in one study of 13320 congeneric species pairs,[38] intraspecific variation was usually less than 1% and rarely more than 2%, while mean interspecific divergence was 11.3%. This difference between intra- and interspecific divergences at COI was also observed in arachnid,[38, 54, 121, 124] lepidopteran,[15, 38-40, 48, 51, 66, 80, 83, 100, 106, 107, 109, 111, 120] dipteran,[38, 42, 48, 87, 93, 95] fish[45, 86, 105, 115, 126, 127, 139-141] and avian[14, 15, 38, 47, 52, 74, 102, 117, 119, 128-131] species (Table 1.1).

DNA barcoding was generally successful when used for identifying immature specimens,[49, 122] extinct species[14] and individual species at differing stages in their life cycles.[44, 49, 120-124] Furthermore, possible cryptic species were identified in several studies.[41, 53, 54, 62, 64, 70, 72, 74, 77, 78, 81, 89, 90, 96, 112, 120, 141]

Identification difficulties caused by morphological differences between instars in cave dwelling spiders (*Cicurina* spp.) were overcome using this technique, which also aided identification of populations where adult specimens are extremely rare.[121] Potential species identified in this study correspond closely to *a priori* species hypotheses except in the case of *C. caliga* and *C. hoodensis*, which are indistinguishable at this locus and the authors suggest that they may be synonymous. Furthermore, COI barcoding of two other species, *C. madla* and *C. vespera* (a species known only from one female specimen), suggested possible synonymy and indicated a need for further evaluation. This utility was also observed when 10 distinct caterpillars were linked to their morphologically similar adults.[120] While in another study, COI barcoding helped to identify arachnids and carabids of different life stages including eggs, larvae or nymphs and pupae.[124]

In extinct taxa, where full taxonomic descriptions are difficult due to the lack of soft tissue, identification of species can be particularly problematical. However, six species of an extinct ratite bird, the New Zealand moa, were identified using DNA

barcoding when 2.7% sequence divergence at COI was used as the intraspecific threshold. This increased to 10 species when a threshold of 1.24% was used[14]. The species identified were generally confirmed and supported by results from a larger study of moa mtDNA, in which 125 specimens had mitochondrial control region sequences analysed (Figure 1.3).[142]

Potential cryptic species were identified using DNA barcoding among butterflies,[39, 81, 107, 111, 120] flies,[69, 96] birds,[74, 128-131] arachnids,[54] springtails,[76] within the species *Daphnia obtusa*[89] and in three groups of parasitic worms.[77, 78]
Larval caterpillars with distinct colour patterning and food plants were linked with adults that are phenotypically very similar to each other.[120, 143] Divergence at COI was considerable (mean K2P 2.76%; range 0.0-7.95%) and when caterpillar/adult morphology and food plants were mapped onto a neighbour joining tree of the COI divergence, 10 probable new species were revealed that showed covariance between morphological, molecular and ecological characteristics.

Another study found fifteen cryptic species of parasitoid flies that show high host-specificity within a group of what had been thought were three generalist species.[96]

COI-identified cryptic species were not limited to relatively obscure taxa. For example, among 260 North American bird species, K2P distances were 18-fold higher between species than within them; however, in four species (*Tringa solitaia, Sturnella magna, Cisthorus palustris* and *Vireo gilvus*), high intraspecific K2P distances suggested the presence of cryptic species.[74]

Similarly, intraspecific K2P divergence averaged 1.4% among 37 arachnid species but in one species, *Latrodectus hesperus*, it was 3.6%.[54] This divergence was between northern and southern populations and suggested that they are probably separate species; a conclusion supported by the results of another study[144] that examined breeding and pheromones in these populations.

The presence of undescribed species was revealed in a study of springtails from the Canadian Arctic.[76] One species showed up to 13% intraspecific COI divergence compared with the <1% generally observed in species from this group. Likewise, wide sequence variation at COI was observed among members of the species *Daphnia obtusa* collected from 33 North American sites,[89] indicating that this may in fact be two species; one confined to the east and the other more broadly distributed.

The efficacy of DNA barcoding at revealing cryptic species was further demonstrated in three studies of parasitic worms.[77, 78, 143] Analysis of COI fragments in each study found high intraspecific COI variation among tapeworm and hookworm species. Two of the three studies[77, 78] suggested that this might be due to the presence of cryptic species. In the third,[143] the authors conclude that there was little COI and morphological covariance; however, the presence of cryptic species was not precluded.

1.5.2 Problematic Taxa

Some taxonomic groups were not readily resolved to species level. These included benthic Cnidarians,[38] two groups of amphibians,[43] a large group of dipterans,[145] the African cichlid fish[146] and some Gastropod species.[92]

There was little COI divergence between species of benthic Cnidarians, with 94.1% of species pairs showing ≤2% divergence versus 1.9% of species pairs from all other phyla in this study.[38] The same was true for African cichlid fish from Lake Victoria that showed great morphological diversity but little (<1%) COI divergence.[146] Conversely, high intraspecific variation at COI of up to 18% was observed in two amphibian groups (mantellid frogs and salamanders).[43] This overlapped the interspecific variation, making species delineation difficult. Furthermore, variability within the mitochondrial genome of these taxa meant that a mix of primers was required to isolate the required segment. Nonetheless, COI sequences were able to correctly identify species including disparate geographic variants.

Gastropods also proved refractory to COI identification in a study of 70 species aimed at establishing phylogenetic relationships as well as species identification.[92] Insertions and deletions found in the COI gene of two subclasses, Heterobranchia and Patellogastropoda, complicated alignment.

These results with gastropods conformed with the general observation that where DNA barcoding was used to resolve higher taxa, results proved more equivocal. In one study,[54] 87% of genera that contained several species and 67% of families that contained several genera formed cohesive COI groups. Another study correctly assigned 96.4% of 55 taxa to phyla and 100% of 50 taxa to ordinal level (Table 1.1).[15]

1.6 Discussion

In general, DNA barcoding using a segment of the COI gene from the mtDNA of a wide variety of animal phyla was effective at resolving species in the studies reviewed above. However, resolution of higher taxonomic groups was less effective than that observed in species. Indeed, it may be that, to paraphrase Greenstone et al., “COI barcoding is a diagnostic tool for identifying animal species and cannot be expected to serve double duty as a character for deeper phylogenetic reconstruction”.[124] Problems were also encountered with some species that proved refractory to identification.

DNA barcoding showed utility at identifying potential candidates for taxonomic description by highlighting possible cryptic species among those already identified. It drew attention to a number of species, both extant and extinct, that may be synonymous and provided a means of identifying species regardless of life stage or maturity. Furthermore, data from a study using other mtDNA sequences in extinct bird species was confirmed using DNA barcoding.

Nonetheless, some questions relating to this technique remain to be answered. For example, taxa that are undergoing rapid speciation may show little interspecific

COI divergence, thus compromising the resolving power of COI barcoding. New Zealand moas are thought to have undergone rapid speciation prior to extinction,[142] which may account for the relatively low interspecific K2P distances observed in this group.[14] The converse may be true for species that have not undergone recent speciation events. It is also possible that the problems associated with barcoding some cnidarians and amphibians may be replicated in other taxa.

More fundamentally, there must be a clear understanding of species if DNA barcoding is to identify them. There are many definitions including typological, morphological, biological, recognition and phylogenetic species, each relating to differing ideas and having its own relevance. Since species is a real concept, true species probably exist where these definitions come together and a DNA barcode may serve as another stream at this confluence.

However, even accepting a generalised and undefined concept of species; what is a DNA barcode? Although CBOL states that “initially, only cytochrome *c* oxidase 1 is approved as a barcode region, defined relative to the mouse mitochondrial genome as the 648bp region that starts at position 58 and stops at position 705”, very few of those sequences lodged with GenBank are of this precise segment of the gene. A large number are considerably shorter and of those longer than 648bp, many don’t fully overlap the specified segment. This variability complicates and reduces the power of large-scale analyses of these data.

The question also arises whether there should be many different barcodes from differing genes in order to identify the broadest range of species. Indeed, COI does not work for other Kingdoms and Nielsen et al.[147] suggest that the weakest aspect of DNA barcoding is that no single gene will always be invariant within species but different between species. The results of some of the problematic taxa bear this out. Nielsen et al. further suggest that there is a need for statistical protocols to assess whether a sample barcode is sufficiently similar to a known barcode to justify species assignment.[147] Thus, issues of standardisation need to be

addressed if barcoding is to achieve the rigour required of an enduring contribution to science.

Some critics have suggested that COI barcoding is unscientific because it does not set out to test hypotheses, that it generates information not knowledge, that it is just data mining.[148, 149] However, arguably, any experiment generates information that requires interpretation. Moreover, barcoding tests the hypothesis that species can be identified using this technique and in future may be a source of data that will generate other hypotheses. Similar objections, now largely forgotten, were raised at the outset of the Human Genome Project.[150]

A number of people remain skeptical of the utility and efficacy of DNA barcoding. There are those who fear that the promoters of barcoding are seeking to replace conventional taxonomy.[148, 149, 151] These fears have been fuelled by an enthusiasm for DNA taxonomy in some quarters.[152] Some take exception to the use of the term barcode on the basis that it suggests that “each species has a fixed and invariant characteristic like the barcode on a supermarket product”.[153] They also express reservations that sufficient numbers of congeneric species have been sampled or that those samples come from a wide enough distribution to make generalisations about the efficacy of COI barcoding.[153]

In addition, there are concerns that any attempt at producing a universal system for identifying species entailing a centralised database may be seen by third world countries as an attempt by wealthier nations to monopolise taxonomic information. It is also thought that any such system may be more authoritarian and will lack the flexibility inherent in the committee style consensus of existing botanical and zoological codes.

Proponents of barcoding respond that COI barcoding is not a substitute for taxonomy. That it cannot be, since it is only by linking barcodes to fully described voucher specimens that the full power of the technique can be realised and that just as supermarket barcodes would be meaningless without the database of

product details to which they are linked, so DNA barcoding requires a database of taxonomic information to which it links. They acknowledge that the barcode analogy is not an exact one but maintain that, with the current level of accuracy observed, COI barcoding has proved sufficiently discriminatory to demonstrate considerable utility as a tool for differentiating species and, therefore, merits further investigation. With regard to the comments about the numbers of species tested using this method so far, it is only as the database of barcodes builds that the substance of these reservations will either be confirmed or otherwise.

Unquestionably, any progress that is to be made in accelerating species identification will be dependent on the use of new technologies and will employ accessible, easily searchable repositories of taxonomic information. Whether or not this is perceived as a threat by countries around the world will depend largely upon the sensitivity with which the process is managed. CBOL favours an entirely unrestricted open access approach. Moreover, it is unlikely that this database will be completely centralised but will probably follow a distributed model with either geographic centres or moderating centres around the world, playing host to parts of the collected data while other sites will host mirrors of these databases.

Those who express concern that this may be an attempt to divest third world countries of information relating to their biotas, are also probably underestimating the power of modern technology to disseminate information. If the current open access approach is maintained, taxonomic information, much of which resides in relatively inaccessible Northern Hemisphere museums and collections, will become more accessible rather than the reverse. Thus, regions may be able to access and reclaim information relating to their indigenous biotas.

The methods by which species are named and described, however, need not be affected by DNA barcoding. This technique is not “a pretender to the taxonomic throne”,[154] its principal utility is as a searchable label and as just another contributing taxonomic feature. Thus, barcoding may serve to help inform the

debate that generally surrounds species identification but is unlikely to undermine the flexibility of existing codes.

Museums around the world maintain large collections of plant and animal specimens and are, therefore, excellent sources of material for DNA barcoding. Their accumulated experience at curating these collections combined with a need for efficient cataloguing systems suggest the potential for a mutually beneficial relationship with CBOL. Furthermore, these institutions play a central role in the collection and description of new species and may, therefore, be primary beneficiaries of the development of new technologies for identifying known and unknown species. To date, a number of major museums and research institutions around the world have lent support to the consortium; however, some remain sceptical about the utility of a barcode system.

There are a large number of web-based projects, unrelated to the CBOL, that are attempting to collate taxonomic information either as phylogenetic trees or as catalogues of species. An example of these is the Tree of Life (ToL) project[155] that provides a framework in which to electronically publish taxonomic information in a searchable database. It is a collaborative effort and currently consists of approximately 8000 web pages. Each page contains information about a particular taxonomic group and is linked hierarchically to other pages in the form of a phylogeny of life.

Evidence suggests that DNA barcoding can serve as a means of accessing taxonomic information and help in the identification of species. However, even if a complete barcode resource of the world's biota is produced, it cannot achieve its full potential unless the processes involved in obtaining barcodes from specimens and accessing the taxonomic information relating to them are simplified and streamlined so that they can be quickly carried out by relatively unskilled workers in a variety of locations. Those involved with CBOL envisage the development of a hand held device to facilitate this. Although such a device does not represent an

insuperable developmental problem, it requires both capital investment and the determination to ensure that it does reach production.

1.7 Conclusion and outlook

Despite criticisms, DNA barcoding shows potential as a tool for identifying species and may or may not, with refining, show greater efficacy at confirming higher taxa. However, the success of the project will depend on the rationalisation of the many online species identification programmes, so that sufficient resources can be concentrated to develop rapid in situ sequencing. There is also a need for the problems associated with species refractory to COI identification to be resolved. In addition, the support of taxonomists from around the world is a prerequisite and this may prove the most intractable obstacle faced by the CBOL. However, with sufficient impetus and will, none of these problems are insoluble.

The advent of DNA sequencing technologies that allow for rapid analysis of genetic information provides the possibility of a universal system for identifying species. Although not a thorough taxonomic characterisation and despite some difficulties with higher taxa, DNA barcodes may allow users to rapidly identify species and to gain access to databases of taxonomic information relating to known species. In addition, they might highlight those species for which no data are yet available; thus, simplifying and speeding the process of identifying specimens for taxonomic description. To date, evidence from a number of studies largely confirms the feasibility of such a system. It remains to be seen whether this approach will be universally adopted.

1.8 The scope of this thesis

The scope of this thesis was to obtain DNA barcodes for the native and endemic birds of New Zealand, evaluate their performance and test their application in a real world setting. It is the only study to do this in New Zealand to date. Although some introduced, straggler or vagrant species have been included in the database, the central focus of this work has been on native and endemic species.

New Zealand's avifauna encompasses 19 Orders and 67 Families, with 279 species (including stragglers, vagrants and introduced species).[156] There is a high level of endemism, particularly among land birds and there is also a very large group of seabirds, often with wide global or circumpolar distributions. There are an additional 45 that are considered Antarctic species, of which, most are endemic to this region. For a complete list of native and endemic New Zealand birds see Chapter 4.

The distribution of New Zealand and Antarctic birds overlaps. A number of the sea birds with circumpolar distributions that can be found around the New Zealand coast or on and around inshore islands as well as on oceanic outer islands are also found in Antarctic regions. In addition, most Antarctic bird species are occasional visitors to New Zealand shores. For this reason, Barcoding New Zealand Birds includes species from a very large region of approximately 11 million square kilometres, spanning nearly 50° (5500 kilometres) of latitude from the sub-tropical Kermadec Islands to the southern oceanic islands including the Chatham Islands in the West, the Snares, Bounty Islands, Antipodes Islands, Auckland Islands and Campbell Island.

In addition to recovering DNA barcodes from representative samples of as many species as could be obtained, a set of primers has been developed that permit amplification of the barcode region of samples from every bird species in the study area including museum samples with degraded DNA (Chapter 3). Moreover some of the uses of DNA barcoding have also been explored (Chapters 6 and 7). Intra- and interspecific variation at the barcode region has been determined for each species in the database. Neighbour joining trees have been generated for each order of species in the database (Chapter 4). Barcodes have been obtained from taxa where species status is questionable to help inform the debate. A number of species have been found to contain two distinct barcode clusters, indicating the possible presence of cryptic species. Other species that appear to be different share very similar barcodes and may be subspecies or species with high morphological plasticity (Chapter 5). The remains of birds from aircraft birdstrike

incidents have been identified using this technique (Chapter 6). DNA barcoding has been used to determine the species of unidentifiable remains of birds found in the wild and other uses of this technique are discussed (Chapter 7).

1.9 References

1. Anon. *The Life and Times of Carl Linnaeus*. 2009 [cited 2009 15 October]; Available from: http://www.linnean.org/index.php?id=51.

2. Stoeckle, M., *Taxonomy, DNA, and the Bar Code of Life.* BioScience, 2003. **53**(2).

3. Pidwirny, M., *Fundamentals of Physical Geography*, in *Biological Classification of Organisms*. 2006, Michael Pidwirny: Okanagan.

4. Blaxter, M., *Counting angels with DNA.* Nature, 2003. **421**: p. 122-124.

5. Blaxter, M.L., *The promise of a DNA taxonomy.* Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 2004. **359**(1444): p. 669-679.

6. Comstock KE, G.N., Pecon-Slattery J, Roca AL, Ostrander EA, et al., *Patterns of molecular genetic variation among African elephant populations.* Molecular Ecology, 2002. **11**: p. 2489-2498.

7. ICZN. *International Code of Zoological Nomenclature online*. 1999 [cited 2009 15 October]; Available from: http://www.iczn.org/iczn/index.jsp.

8. ICBN. *International Code of Botanical Nomenclature*. Vienna Code 2006 [cited 2009 15 October]; Available from: http://ibot.sav.sk/icbn/main.htm.

9. Buchanan, R.E., *The International Code of Nomenclature of the Bacteria and Viruses.* Systematic Zoology, 1959. **8**(1): p. 27-39.

10. Wilson, E.O., *Taxonomy as a fundamental discipline.* Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 2004. **359**(1444): p. 739-739.

11. Baker, C., Palumbi, SR, *Which Whales Are Hunted? A Molecular Genetic Approach to Monitoring Whaling.* Science, 1994. **265**(5178): p. 1538-1539.

12. Baker, C., Lento, GM, Cipriano, F, Dalebout, ML, Palumbi, SR, *Scientific whaling: source of illegal products for market?* Science, 2000. **290**(5497): p. 1695-6.

13. Besansky, N., Severson, DW, Ferdig, MT, *DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you.* Trends in Parasitology, 2003. **19**(12): p. 545-546.

14. Lambert, D., Baker, A, Huynen, L, Haddrath, O, Hebert, PDN, Millar, CD, *Is a large-scale DNA-based inventory of ancient life possible?* Journal of Heredity, 2005. **96**(3): p. 279-284.

15. Hebert, P., Cywinska, A, Ball, SL, DeWaard, JR, *Biological identifications through DNA barcodes.* Proceedings of the Royal Society of London Series B-Biological Sciences, 2003. **270**(1512): p. 313-321.

16. Anon. *National survey reveals biodiversity crisis*. 1998; Available from: http://www.amnh.org/museum/press/feature/biohall.html.

17. Ananthaswamy, A., *Scene set for next mass extinction.* New Scientist, 2004. **181**(2440): p. 10-11.

18. Mayo, M., Horzinek, MC, *A revised version of the International Code of Virus Classification and Nomenclature.* Arch Virol, 1998. **143**: p. 1645-1654.

19. Unwin, R., Maiden, MCJ, *Multi-locus sequence typing: a tool for global epidemiology.* Trends Microbiol, 2003. **11**: p. 479-487.

20. Woese, C., *Bacterial evolution.* Microbiol Rev, 1987. **51**: p. 221-271.

21. Cavalier-Smith, T., *Only six kingdoms of life.* Proceedings of the Royal Society of London Series B-Biological Sciences, 2004. **271**(1545): p. 1251-1262.

22. Saunders, G.W., *Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1879-88.

23. *Barcode of Life website*. 2009 [cited 2009 15 October]; Available from: http://barcoding.si.edu/.

24. *All Birds Barcoding Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.barcodingbirds.org/#.

25. *Fish Barcode of Life Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.fishbol.org/.

26. Anon. *The Sponge Barcoding Project*. 2010; Available from: http://www.spongebarcoding.org/.

27. *All Leps Barcode of Life*. 2009 [cited 2009 18 October]; Available from: http://www.lepbarcoding.org/.

28. *International Network for Barcoding Invasive and Pest Species (INBIPS)*. 2010; Available from: http://www.barcoding.si.edu/INBIPS.htm.

29. *Barcode of Life Data Systems*. 2009 [cited 2010 22nd July]; Available from: http://www.barcodinglife.org/.

30. *Global Biodiversity Information Facility (GBIF)*. 2009 [cited 2009 15 October]; Available from: http://www.gbif.org/.

31. *National Center for Biotechnology Information (NCBI)*. 2009 [cited 2009 15 October]; Available from: http://www.ncbi.nlm.nih.gov/.

32. Waugh, J., *DNA barcoding in animal species: progress, potential and pitfalls.* Bioessays, 2007. **29**(2): p. 188-197.

33. Drake, J., Charlesworth, B, Charlesworth, D, Crow, JF, *Rates of spontaneous mutation.* Genetics, 1998. **148**(4): p. 1667-86.

34. Shearer, T., Van Oppen, MJ, Romano, SL, Worheide, G, *Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria).* Mol Ecol, 2002. **11**(12): p. 2475-87.

35. *Research Collaboratory for Structural Bioinformatics (RCSB) Protein Databank (PDB)*. 2009 [cited 2009 18 October]; Available from: http://www.rcsb.org/pdb/cgi/explore.cgi?pdbId=2OCC.

36. *Cytochrome Oxidase homepage*. 2009 [cited 2009 18 October]; Available from: http://www.bioc.rice.edu/~graham/CcO.html.

37. Kimura, M., *A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences.* Journal of Molecular Evolution, 1980. **16**: p. 111–120.

38. Hebert, P., Ratnasingham, S, deWaard, JR, *Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species.* Proceedings of the Royal Society of London Series B-Biological Sciences, 2003. **270**: p. S96-S99.

39. Hajibabaei, M., Janzen, DH, Burns, JM, Hallwachs, W, Hebert, PDN, *DNA barcodes distinguish species of tropical Lepidoptera.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(4): p. 968-971.

40. Hajibabaei, M., Smith, MA, Janzen, DH, Rodriguez, JJ, Whitfield, JB, Hebert, PDN, *A minimalist barcode can identify a specimen whose DNA is degraded.* Molecular Ecology Notes, 2006. **6**(4): p. 959-964.

41. Mikkelsen, N.T., C. Schander, and E. Willassen, *Local scale DNA barcoding of bivalves (Mollusca): A case study.* Zoologica Scripta, 2007. **36**(5): p. 455-463.

42. Scheffer, S.J., M.L. Lewis, and R.C. Joshi, *DNA barcoding applied to invasive leafminers (Diptera : Agromyzidae) in the Philippines.* Annals of the Entomological Society of America, 2006. **99**(2): p. 204-210.

43. Vences, M., Thomas, M, Bonett, RM, Vieites, DR, *Deciphering amphibian diversity through DNA barcoding: chances and challenges.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1859-68.

44. Ahrens, D., Monaghan, MT, Vogler, AP, *DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafers (Coleoptera : Scarabaeidae).* Molecular Phylogenetics and Evolution, 2007. **44**(1): p. 436-449.

45. Holmes, B.H., D. Steinke, and R.D. Ward, *Identification of shark and ray fins using DNA barcoding.* Fisheries Research, 2009. **95**(2-3): p. 280-288.

46. Rojo, S., et al., *Testing molecular barcodes: Invariant mitochondrial DNA sequences vs the larval and adult morphology of West Palaearctic Pandasyopthalmus species (Diptera : Syrphidae : Paragini).* European Journal of Entomology, 2006. **103**(2): p. 443-458.

47. Dove, C., Rotzel, NC, Heacker, M, Weigt, LA, *Using DNA barcodes to identify bird species involved in birdstrikes.* Journal of Wildlife Management, 2008. **72**(5): p. 1231-1236.

48. Armstrong, K., Ball, SL, *DNA barcodes for biosecurity: invasive species identification.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**: p. 1813-1823.

49. Baldwin, C., Mounts, JH, Smith, DG, Weigt, LA, *Genetic identification and color descriptions of early life-history stages of Belizean Phaeoptyx and Astrapogon (Teleostei: Apogonidae) with Comments on identification of adult Phaeoptyx.* Zootaxa, 2009(2008): p. 1-22.

50. Ball, S., Hebert, PDN, Burian, SK, Webb, JM, *Biological identifications of mayflies (Ephemeroptera) using DNA barcodes.* Journal of the North American Benthological Society, 2005. **24**(3): p. 508-524.

51. Ball, S., Armstrong, KF, *DNA barcodes for insect pest identification: a test case with tussock moths (Lepidoptera : Lymantriidae).* Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 2006. **36**(2): p. 337-350.

52. Banks, J., Paterson, AM, *A preliminary study of the genetic differences in New Zealand oystercatcher species.* New Zealand Journal of Zoology, 2007. **34**(2): p. 141-144.

53. Ball, S., Armstrong, KF, *Using DNA barcodes to investigate the taxonomy of the New Zealand sooty beech scale insect.* DoC Research and Development 2007. **287**: p. 5-14.

54. Barrett, R., Hebert, PDN, *Identifying spiders through DNA barcodes.* Canadian Journal of Zoology-Revue Canadienne De Zoologie, 2005. **83**(3): p. 481-491.

55. Campbell, D.C., Johnson, PD, Williams, JD, Rindsberg, AK, Serb, JM, Small, KK, Lydeard, C, *Identification of 'extinct' freshwater mussel species using DNA barcoding.* Molecular Ecology Resources, 2008. **8**(4): p. 711-724.

56. Chang, C.H., R. Rougerie, and J.H. Chen, *Identifying earthworms through DNA barcodes: Pitfalls and promise.* Pedobiologia, 2009. **52**(3): p. 171-180.

57. Chantangsi, C., Lynn, DH, Brandl, MT, Cole, JC, Hetrick, N, Ikonomi, P, *Barcoding ciliates: a comprehensive study of 75 isolates of the genus Tetrahymena.* Int J Syst Evol Microbiol, 2007. **57**(Pt 10): p. 2412-25.

58. Costa, F., deWaard, JR, Boutillier, J, Ratnasingham, S, Dooh, RT, Hajibabaei, M, Hebert, PDN, *Biological identifications through DNA barcodes: the case of the Crustacea.* Canadian Journal of Fisheries and Aquatic Sciences, 2007. **64**(2): p. 272-295.

59. Dabert, J., R. Ehrnsberger, and M. Dabert, *Glaucalges tytonis sp n. (Analgoidea, Xolalgidae) from the barn owl Tyto alba (Strigiformes, Tytonidae): compiling morphology with DNA barcode data for taxon descriptions in mites (Acari).* Zootaxa, 2008(1719): p. 41-52.

60. De Ley, P., De Ley, IT, Morris, K, Abebe, E, Mundo-Ocampo, M, Yoder, M, Heras, J, Waumann, D, Rocha-Olivares, A, Burr, AHJ, Baldwin, JG, Thomas, WK, *An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**(1462): p. 1945-1958.

61. Dooh, R.T., S.J. Adamowicz, and P.D.N. Hebert, *Comparative phylogeography of two North American 'glacial relict' crustaceans.* Molecular Ecology, 2006. **15**(14): p. 4459-4475.

62. Efe, M., Tavares, ES, Baker, AJ, Bonatto, SL, *Multigene phylogeny and DNA barcoding indicate that the Sandwich tern complex (Thalasseus sandvicensis, Laridae, Sternini) comprises two species.* Mol Phylogenet Evol, 2009. **52**(1): p. 263-7.

63. Elias-Gutierrez, M., Jeronimo, FM, Ivanova, NV, Valdez-Moreno, M, Hebert, PDN, *DNA barcodes for Cladocera and Copepoda from Mexico and Guatemala, highlights and new discoveries.* Zootaxa, 2008(1839): p. 1-42.

64. Elias-Gutierrez, M. and M. Valdez-Moreno, *A new cryptic species of Leberis Smirnov, 1989 (Crustacea, Cladocera, Chydoridae) from the Mexican semi-desert region, highlighted by DNA barcoding.* Hidrobiologica, 2008. **18**(1): p. 63-74.

65. Elsasser, S., Floyd, R, Hebert, PDN, Schulte-Hostedde, AI, *Species identification of North American guinea worms (Nematoda: Dracunculus) with DNA barcoding.* Molecular Ecology Resources, 2009. **9**(3): p. 707-712.

66. Emery, V.J., J.F. Landry, and C.G. Eckert, *Combining DNA barcoding and morphological analysis to identify specialist floral parasites (Lepidoptera: Coleophoridae: Momphinae: Mompha).* Molecular Ecology Resources, 2009. **9**: p. 217-223.

67. Ferri, E., Barbuto, M, Bain, O, Galimberti, A, Uni, S, Guerrero, R, Ferte, H, Bandi, C, Martin, C, Casiraghi, M, *Integrated taxonomy: traditional approach and DNA barcoding for the identification of filarioid worms and related parasites (Nematoda).* Front Zool, 2009. **6**: p. 1.

68. Fisher, B.L. and M.A. Smith, *A revision of Malagasy species of Anochetus mayr and Odontomachus latreille (Hymenoptera: Formicidae).* PLoS One, 2008. **3**(5): p. e1787.

69. Foley, D., Wilkerson, RC, Cooper, RD, Volovsek, ME, Bryan, JH, *A molecular phylogeny of Anopheles annulipes (Diptera: Culicidae) sensu lato: the most species-rich anopheline complex.* Mol Phylogenet Evol, 2007. **43**(1): p. 283-97.

70. Foottit, R., Maw, HEL, Havill, NP, Ahern, RG, Montgomery, ME, *DNA barcodes to identify species and explore diversity in the Adelgidae (Insecta: Hemiptera: Aphidoidea).* Molecular Ecology Resources, 2009. **9**: p. 188-195.

71. Foottit, R., Maw, HEL, Von Dohlen, CD, Hebert, PDN, *Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes.* Molecular Ecology Resources, 2008. **8**(6): p. 1189-1201.

72. Gomez, A., Wright, PJ, Lunt, DH, Cancino, JM, Carvalho, GR, Hughes, RN, *Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoan taxon.* Proc Biol Sci, 2007. **274**(1607): p. 199-207.

73. Hajibabaei, M., G.A.C. Singer, and D.A. Hickey, *Benchmarking DNA barcodes: an assessment using available primate sequences.* Genome, 2006. **49**(7): p. 851-854.

74. Hebert, P., Stoeckle, MY, Zemlak, TS, Francis, CM *Identification of Birds through DNA Barcodes.* PLoS Biol, 2004. **2**(10).

75. Hogg, I., Stevens, MI, Schnabel, KE, Chapman, MA, *Deeply divergent lineages of the widespread New Zealand amphipod Paracalliope fluviatilis revealed using allozyme and mitochondrial DNA analyses.* Freshwater Biology, 2006. **51**(2): p. 236-248.

76. Hogg, I.D. and P.D.N. Hebert, *Biological identification of springtails (Hexapoda : Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes.* Canadian Journal of Zoology-Revue Canadienne De Zoologie, 2004. **82**(5): p. 749-754.

77. Hu, M., Chilton, NB, Zhu, XQ, Gasser, RB, *Single-strand conformation polymorphism-based analysis of mitochondrial cytochrome c oxidase subunit 1 reveals significant substructuring in hookworm populations.* Electrophoresis, 2002. **23**(1): p. 27-34.

78. Hu, M., Gasser, RB, Chilton, NB, Beveridge, I, *Genetic variation in the mitochondrial cytochrome c oxidase subunit 1 within three species of Progamotaenia (Cestoda : Anoplocephalidae) from macropodid marsupials.* Parasitology, 2005. **130**: p. 117-129.

79. Huang, J., Xu, Q, Sun, ZJ, Tang, GL, Su, ZY, *Identifying earthworms through DNA barcodes.* Pedobiologia, 2007. **51**(4): p. 301-309.

80. Hulcr, J., Miller, SE, Setliff, GP, Darrow, K, Mueller, ND, Hebert, PDN, Weiblen, GD, *DNA barcoding confirms polyphagy in a generalist moth, Homona mermerodes (Lepidoptera : Tortricidae).* Molecular Ecology Notes, 2007. **7**(4): p. 549-557.

81. Johnson, S.B., A. Waren, and R.C. Vrijenhoek, *DNA barcoding of Lepetodrilus limpets reveals cryptic species.* Journal of Shellfish Research, 2008. **27**(1): p. 43-51.

82. Kelly, R., Sarkar, IN, Eernisse, DJ, Desalle, R, *DNA barcoding using chitons (genus Mopalia).* Molecular Ecology Notes, 2007. **7**(2): p. 177-183.

83. Landry, J.F., *Taxonomic review of the leek moth genus Acrolepiopsis (Lepidoptera : Acrolepiidae) in North America.* Canadian Entomologist, 2007. **139**(3): p. 319-353.

84. Langhoff, P., Authier, A, Buckley, TR, Dugdale, JS, Rodrigo, A, Newcomb, RD, *DNA barcoding of the endemic New Zealand leafroller moth genera, Ctenopseustis and Planotortrix.* Molecular Ecology Resources, 2009. **9**(3): p. 691-698.

85. Martinez, J.J., A. Zaldivar-Riveron, and A.G. Saez, *Reclassification of Bracon mendocinus, a gall-associated doryctine wasp, and description of a new closely related species of Allorhogas (Hymenoptera: Braconidae).* Journal of Natural History, 2008. **42**(41-42): p. 2689-2701.

86. Moura, T., Silva, MC, Figueiredo, I, Neves, A, Munoz, PD, Coelho, MM, Gordo, LS, *Molecular barcoding of north-east Atlantic deep-water sharks: species identification and application to fisheries management and conservation.* Marine and Freshwater Research, 2008. **59**(3): p. 214-223.

87. Nelson, L.A., J.F. Wallman, and M. Dowton, *Using COI barcodes to identify forensically and medically important blowflies.* Medical and Veterinary Entomology, 2007. **21**(1): p. 44-52.

88. Nolan, L., Hogg, ID, Sutherland, DL, Stevens, MI, Schnabel, KE, *Allozyme and mitochondrial DNA variability within the New Zealand damselfly genera Xanthocnemis, Austrolestes, and Ischnura (Odonata).* New Zealand Journal of Zoology, 2007. **34**(4): p. 371-380.

89. Penton, E.H., P.D. Hebert, and T.J. Crease, *Mitochondrial DNA variation in North American populations of Daphnia obtusa: continentalism or cryptic endemism?* Mol Ecol, 2004. **13**(1): p. 97-107.

90. Pfenninger, M., Nowak, C, Kley, C, Steinke, D, Streit, B, *Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic Chironomus (Diptera) species.* Mol Ecol, 2007. **16**(9): p. 1957-68.

91. Radulovici, A.E., B. Sainte-Marie, and F. Dufresne, *DNA barcoding of marine crustaceans from the Estuary and Gulf of St Lawrence: a regional-scale approach.* Molecular Ecology Resources, 2009. **9**: p. 181-187.

92. Remigio, E.A. and P.D.N. Hebert, *Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships.* Molecular Phylogenetics and Evolution, 2003. **29**(3): p. 641-647.

93. Skevington, J.H., C. Kehlmaier, and G. Stahls, *DNA Barcoding: Mixed results for big-headed flies (Diptera : Pipunculidae).* Zootaxa, 2007(1423): p. 1-26.

94. Seabrook-Davison, M., et al., *Ancient DNA resolves identity and phylogeny of New Zealand's extinct and living quail (Coturnix sp.).* PLoS One, 2009. **4**(7): p. e6400.

95. Sinclair, C.S. and S.E. Gresens, *Discrimination of Cricotopus species (Diptera: Chironomidae) by DNA barcoding.* Bull Entomol Res, 2008. **98**(6): p. 555-63.

96. Smith, M., Woodley, NE, Janzen, DH, Hallwachs, W, Hebert, PDN, *DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera : Tachinidae).* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(10): p. 3657-3662.

97. Whiteman, N., Santiago-Alarcon, D, Johnson, KP, Parker, PG, *Differences in straggling rates between two genera of dove lice (Insecta: Phthiraptera) reinforce population genetic and cophylogenetic patterns.* Int J Parasitol, 2004. **34**(10): p. 1113-9.

98. Witt, J.D.S., D.L. Threloff, and P.D.N. Hebert, *DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation.* Molecular Ecology, 2006. **15**(10): p. 3073-3082.

99. Borisenko, A., Lim, BK, Ivanova, NV, Hanner, RH, Hebert, PDN, *DNA barcoding in surveys of small mammal communities: a field study in Suriname.* Molecular Ecology Resources, 2008. **8**(3): p. 471-479.

100. Brown, J.W., S.E. Miller, and M. Horak, *Studies on New Guinea moths. 2. Description of a new species of Xenothictis meyrick (Lepidoptera : Tortricidae : Archipini).* Proceedings of the Entomological Society of Washington, 2003. **105**(4): p. 1043-1050.

101. Cesari, M., Bertolani, R, Rebecchi, L, Guidetti, R, *DNA barcoding in Tardigrada: the first case study on Macrobiotus macrocalix Bertolani & Rebecchi 1993 (Eutardigrada, Macrobiotidae).* Molecular Ecology Resources, 2009. **9**(3): p. 699-706.

102. Chaves, A., Clozato, CL, Lacerda, DR, Sari, EHR, Santos, FR, *Molecular taxonomy of brazilian tyrant-flycatchers (Passeriformes: Tyrannidae).* Molecular Ecology Resources, 2008. **8**(6): p. 1169-1177.

103. Clare, E., Lim, BK, Engstrom, MD, Eger, JL, Hebert, PDN, *DNA barcoding of Neotropical bats: species identification and discovery within Guyana.* Molecular Ecology Notes, 2007. **7**(2): p. 184-190.

104. Day, J.C., T.I. Goodall, and R.J. Post, *Confirmation of the species status of the blackfly Simulium galeratum in Britain using molecular taxonomy.* Med Vet Entomol, 2008. **22**(1): p. 55-61.

105. De Astarloa, J., Mabragana, E, Hanner, R, Figueroa, DE, *Morphological and molecular evidence for a new species of longnose skate (Rajiformes: Rajidae: Dipturus) from Argentinean waters based on DNA barcoding.* Zootaxa, 2008(1921): p. 35-46.

106. Decaens, T. and R. Rougerie, *Descriptions of two new species of Hemileucinae (Lepidoptera: Saturniidae) from the region of Muzo in Colombia-evidence from morphology and DNA barcodes.* Zootaxa, 2008(1944): p. 34-52.

107. Janzen, D., Hajibabaei, M, Burns, JM, Hallwachs, W, Remigio, E, Hebert, PDN, *Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1835-45.

108. Lane, C., Lindstrom, SC, Saunders, GW, *A molecular assessment of northeast Pacific Alaria species (Laminariales, Phaeophyceae) with reference to the utility of DNA barcoding.* Mol Phylogenet Evol, 2007. **44**(2): p. 634-48.

109. Linares, M., Soto-Calderon, ID, Lees, DC, Anthony, NM, *High mitochondrial diversity in geographically widespread butterflies of Madagascar: a test of the DNA barcoding approach.* Mol Phylogenet Evol, 2009. **50**(3): p. 485-95.

110. Lorenz, J., Jackson, WE, Beck, JC, Hanner, R, *The problems and promise of DNA barcodes for species diagnosis of primate biomaterials.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1869-77.

111. Lukhtanov, V., Sourakov, A, Zakharov, EV, Hebert, PDN, *DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification.* Molecular Ecology Resources, 2009. **9**(5): p. 1302-1310.

112. Quiroz-Vazquez, P. and M. Elias-Gutierrez, *A New Species of the Freshwater Cladoceran Genus Scapholeberis Schoedler, 1858 (Cladocera: Anomopoda) from the Semidesert Northern Mexico, Highlighted by DNA Barcoding.* Zootaxa, 2009(2236): p. 50-64.

113. Rock, J., Costa, FO, Walker, DI, North, AW, Hutchinson, WF, Carvalho, GR, *DNA barcodes of fish of the Scotia Sea, Antarctica indicate priority groups for taxonomic and systematics focus.* Antarctic Science, 2008. **20**(3): p. 253-262.

114. Sheffield, C., Hebert, PDN, Kevan, PG, Packer, L, *DNA barcoding a regional bee (Hymenoptera: Apoidea) fauna and its potential for ecological studies.* Molecular Ecology Resources, 2009. **9**: p. 196-207.

115. Shirak, A., Cohen-Zinder, M, Barroso, RM, Seroussi, E, Ron, M, Hulata, G, *DNA Barcoding of Israeli Indigenous and Introduced Cichlids.* Israeli Journal of Aquaculture-Bamidgeh, 2009. **61**(2): p. 83-88.

116. Steinke, D., T.S. Zemlak, and P.D. Hebert, *Barcoding nemo: DNA-based identifications for the ornamental fish trade.* PLoS One, 2009. **4**(7): p. e6300.

117. Tavares, E.S. and A.J. Baker, *Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds.* Bmc Evolutionary Biology, 2008. **8**: p. -.

118. Vargas, S.M., F.C.F. Araujo, and F.R. Santos, *DNA barcoding of Brazilian sea turtles (Testudines).* Genetics and Molecular Biology, 2009. **32**(3): p. 608-612.

119. Yoo, H., Eah, JY, Kim, JS, Kim, YJ, Min, MS, Paek, WK, Lee, H, Kim, CB, *DNA barcoding Korean birds.* Molecules and Cells, 2006. **22**(3): p. 323-327.

120. Hebert, P., Penton, EH, Burns, JM, Janzen, DH, Hallwachs, W, *Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator.* Proc Natl Acad Sci U S A, 2004. **101**(41): p. 14812-7.

121. Paquin, P. and M. Hedin, *The power and perils of 'molecular taxonomy': a case study of eyeless and endangered Cicurina (Araneae : Dictynidae) from Texas caves.* Molecular Ecology, 2004. **13**(10): p. 3239-3255.

122. Victor, B., Hanner, R, Shivji, M, Hyde, J, Caldow, C, *Identification of the larval and juvenile stages of the Cubera Snapper, Lutjanus cyanopterus, using DNA barcoding.* Zootaxa, 2009(2215): p. 24-36.

123. Webb, K., Barnes, DKA, Clark, MS, Bowden, DA, *DNA barcoding: A molecular tool to identify Antarctic marine larvae.* Deep-Sea Research Part Ii-Topical Studies in Oceanography, 2006. **53**(8-10): p. 1053-1060.

124. Greenstone, M., Rowley, DL, Heimbach, U, Lundgren, JG, Pfannenstiel, RS, Rehner, SA, *Barcoding generalist predators by polymerase chain reaction: carabids and spiders.* Mol Ecol, 2005. **14**(10): p. 3247-66.

125. Rasmussen, R.S., M.T. Morrissey, and P.D.N. Hebert, *DNA Barcoding of Commercially Important Salmon and Trout Species (Oncorhynchus and Salmo) from North America.* Journal of Agricultural and Food Chemistry, 2009. **57**(18): p. 8379-8385.

126. Smith, P.J., S.M. Mcveagh, and D. Steinke, *DNA barcoding for the identification of smoked fish products.* Journal of Fish Biology, 2008. **72**(2): p. 464-471.

127. Wong, E.H.K. and R.H. Hanner, *DNA barcoding detects market substitution in North American seafood.* Food Research International, 2008. **41**(8): p. 828-837.

128. Kerr, K., Stoeckle, MY, Dove, CJ, Weigt, LA, Francis, CM, Hebert, PDN, *Comprehensive DNA barcode coverage of North American birds.* Molecular Ecology Notes, 2007. **7**(4): p. 535-543.

129. Park, H., Yoo, HS, Jung, G and Kim CB, *New DNA barcodes for identification of Korean birds.* Genes & Genomics, 2011. **33**: p. 91-95.

130. Kerr, K.C., Lijtmaer, D A, Barreira, A S, Hebert, PDN, Tubaro, P L, *Probing evolutionary patterns in neotropical birds through DNA barcodes.* PLoS One, 2009. **4**(2): p. e4379.

131. Johnsen, A., Rindal, E, Ericson, PGP, Zuccon, D, Kerr, KCR, Stoeckle, MY and Lifjeld, JT, *DNA barcoding Scandinavian Birds reveals divergent lineages in trans-Atlantic species.* J. Ornithol, 2010. **151**(3): p. 565-578.

132. Chase, M.W., et al., *A proposal for a standardised protocol to barcode all land plants.* Taxon, 2007. **56**(2): p. 295-299.

133. Chase, M.W. and M.F. Fay, *Ecology. Barcoding of plants and fungi.* Science, 2009. **325**(5941): p. 682-3.

134. Chase, M., Salamin, N, Wilkinson, M, Dunwell, JM, Kesanakurthi, RP, Haidar, N, Savolainen, V, *Land plants and DNA barcodes: short-term and long-term goals.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**(1462): p. 1889-1895.

135. Erickson, D., Spouge, J, Resch, A, Weigt, LA, Kress, WJ, *DNA Barcoding in Land Plants: Developing Standards to Quantify and Maximize Success.* Taxon, 2008. **57**(4): p. 1304-1316.

136. Nassonova, E., Smirnov, A, Fahrni, J, Pawlowski, J, *Barcoding Amoebae: Comparison of SSU, ITS and COI Genes as Tools for Molecular Identification of Naked Lobose Amoebae.* Protist, 2009.

137. Saunders, G.W., *Marine macroalgae as an exemplar of DNA barcoding among protistan lineages challenges widely held perspectives in biodiversity and biogeography.* Journal of Phycology, 2007. **43**: p. 21-21.

138. Scicluna, S.M., B. Tawari, and C.G. Clark, *DNA barcoding of Blastopystis.* Protist, 2006. **157**(1): p. 77-85.

139. Ward, R., Holmes, BH, White, WT, Last, PR, *DNA barcoding Australasian chondrichthyans: results and potential uses in conservation.* Marine and Freshwater Research, 2008. **59**(1): p. 57-71.

140. Ward, R.D., B.H. Holmes, and G.K. Yearsley, *DNA barcoding reveals a likely second species of Asian sea bass (barramundi) (Lates calcarifer).* Journal of Fish Biology, 2008. **72**(2): p. 458-463.

141. Ward, R., Zemlak, TS, Innes, BH, Last, PR, Hebert, PDN, *DNA barcoding Australia's fish species.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**(1462): p. 1847-1857.

142. Baker, A., Huynen, LJ, Haddrath, O, Millar, CD, Lambert, DM, *Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: the giant moas of New Zealand.* Proc Natl Acad Sci U S A, 2005. **102**(23): p. 8257-62.

143. Ngarmamonpirat, C., Waikagul, J, Petmitr, S, Dekumyoy, P, Rojekittikhun, W, Anantapruti, MT, *Analysis of sequence variation in Gnathostoma spinigerum mitochondrial DNA by single-strand conformation polymorphism analysis and DNA sequence.* Parasitology International, 2005. **54**(1): p. 65-68.

144. Kasumovic, M., Andrade, MCB., *Discrimination of airborne pheromonesby mate searching male westernblack widow sspiders (Latrodectushesperus) species and poipulation responses.* Can. J. Zool., 2004. **82**: p. 1027-1034.

145. Meier, R., et al., *DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success.* Syst Biol, 2006. **55**(5): p. 715-28.

146. Rubinoff, D., *Utility of mitochondrial DNA barcodes in species conservation.* Conserv Biol, 2006. **20**(4): p. 1026-33.

147. Nielsen, R., Matz, M., *Statistical approaches to DNA barcoding.* Systematic Biology, 2006. **55**: p. 162-169.

148. Ebach, M.C. and C. Holdrege, *DNA barcoding is no substitute for taxonomy.* Nature, 2005. **434**(7034): p. 697.

149. Lipscomb, D., Platnick, M. and Wheeler, Q., *The intellectual content of taxonomy: a comment on DNA taxonomy.* TRENDS in Ecology and Evolution, 2003. **18**: p. 65-66.

150. Lewontin, R., *It ain't necessarily so: the dream of the human genome and other illusions*. 2000: New York Review of Books 330.

151. Moritz, C. and C. Cicero, *DNA barcoding: promise and pitfalls.* PLoS Biol, 2004. **2**(10): p. e354.

152. Tautz, D., Arclander, P, Thomas, RH, Vogler,AP., *A plea for DNA taxonomy.* Trends in Ecology and Evolution, 2003. **18**: p. 63-65.

153. Seberg, O. and G. Petersen, *How many loci does it take to DNA barcode a crocus?* PLoS One, 2009. **4**(2): p. e4598.

154. Hebert, P., Gregory, TR, *The promise of DNA barcoding for taxonomy.* Syst Biol, 2005. **54**(5): p. 852-9.

155. *Tree of Life*. 2010 [cited 2010 19th July]; Available from: http://tolweb.org.

156. *Checklist of the birds of New Zealand*. 4th ed, ed. B. Gill, Bell, BD, Chambers, GK,Medway, DG, Palma, RL, Scofield, RP, Tennyson AJD and Worthy TH. Vol. 1. 2010: Te Papa Press.

Chapter 2

Methods: DNA barcode generation and BOLD submission

2.1 Overview of sample collection

Ideally, DNA extracted to generate COI sequences for the development of a database such as that produced for this thesis should come from fresh samples of muscle, blood or feathers. The samples should be part of a museum or similar institutional bird collection and be fully vouchered. Details of collection, including date, collector, field ID, location where collected with map coordinates and altitude as well as the institution in which the specimen is deposited should, be given. Any other relevant details such as age, sex and any other distinguishing characteristics of the specimen should also be provided. Identification details should include full taxonomic identification to species level, the name of the identifier with contact details, along with a voucher number. There should be photographs of the particular specimen and the trace file from the sample should be of high quality.[1]

Achieving these standards is dependent on the availability of large collections of properly vouchered fresh tissue in these institutions. The major museums and natural history collections of the Northern Hemisphere have such collections;[2] elsewhere in the world, they are few and far between and tend to be incomplete. For example, in New Zealand, only Auckland and Otago museums keep fresh bird tissue but these are for only a small percentage of the New Zealand avifauna. Comprehensive collections of bird specimens exist in other New Zealand Museums but they are generally made up of preserved study skins of varying age, usually with highly degraded DNA.

Collections of fresh samples are also held by a number of other institutions such as universities, the Department of Conservation (DoC) or the National Institute of Water and Atmospheric Research (NIWA). However, these collections are often

not vouchered. Bird sanctuaries and zoos can also provide fresh samples but again these are usually not vouchered.

2.2 Samples obtained for this study

Samples obtained for this study fell into the following categories:

i. Vouchered fresh tissue samples from the following museums – Auckland War memorial Museum (88 specimens), Otago Museum in Dunedin (45 specimens) and Victoria Museum in Melbourne (101 specimens).

ii. Vouchered samples from preserved skins from Canterbury Museum in Christchurch (236 samples).

iii. Fresh tissue samples (mainly blood) from the collections of David Lambert at the Allan Wilson Centre (197 specimens), Massey University, Albany (13 specimens) and Craig Millar at the Allan Wilson Centre at the University of Auckland (304 specimens).

iv. Fresh tissue samples from the following bird sanctuaries and zoos - Whakatane Bird Rescue, Sylvia Durrant Bird Rescue in Rothesay Bay, Auckland (40 specimens), Auckland Zoo (10 specimens) and Wellington Zoo (1 specimen).

v. Fresh tissue sent by wildlife control officers at Wellington International Airport and Christchurch International Airport (46 specimens).

Specimens, from which samples were obtained, were identified by a number of field biologists and museum taxonomists (see Appendix E). Tissues used for DNA extraction were most frequently blood and less frequently muscle or feathers. In the case of birds sampled from study skins, muscle tissue from a toe-pad was used or skin from a toe on some of the smaller birds sampled.

2.3 Tissue collection and recording

The majority of the fresh tissue from which samples were obtained was blood (see section 2.2). Blood has the advantage that it can be collected from live specimens, although this generally precludes the possibility that the specimen can be vouchered. In general, blood was collected in the field from the tarsal vein of the leg of birds using a needle and syringe. The blood was then stored in lysis buffer.[3]

Muscle tissue can only be collected from dead specimens and so is more difficult to obtain. However, muscle tissue preserved in alcohol or frozen in either a standard -20°C freezer or a laboratory -80°C freezer is the most common fresh tissue available from museums in New Zealand. In addition, tissue can be removed from the toe pads of study skins.

Feathers were obtained from dead or living specimens. Only one or two down feathers from the breast of the bird are sufficient to provide enough DNA for barcoding purposes. Although largely dead tissue, the base of the proximal part of the central shaft of the feather (the calamus) contains an opening (the proximal umbilicus), which has some residual blood within it and DNA is extracted from this after digestion of the surrounding β-keratin, from which the feather is largely made.[4]

Initially, specimen details of samples as well as collection data and, when obtained, a FASTA file of the sequence were recorded in a database generated in FileMaker Pro 8.5 v2© and included the details shown in Table 2.1 and Figure 2.1.

Table 2.1 Sample data by category entered in a database for each sample sequenced for the Barcoding New Zealand Birds project. The database software used was FileMaker Pro 8.5 v2©

Reference data	Taxonomy	Collection data	Other information
Series number	Scientific name	Collector	Maori name
Creation date	Order	Collection location	Status
Modification date	Family	Collection date	Picture
Reference number	Group name	Sample type	GenBank#
Sample obtained	Common name	Source of sample	Sequence

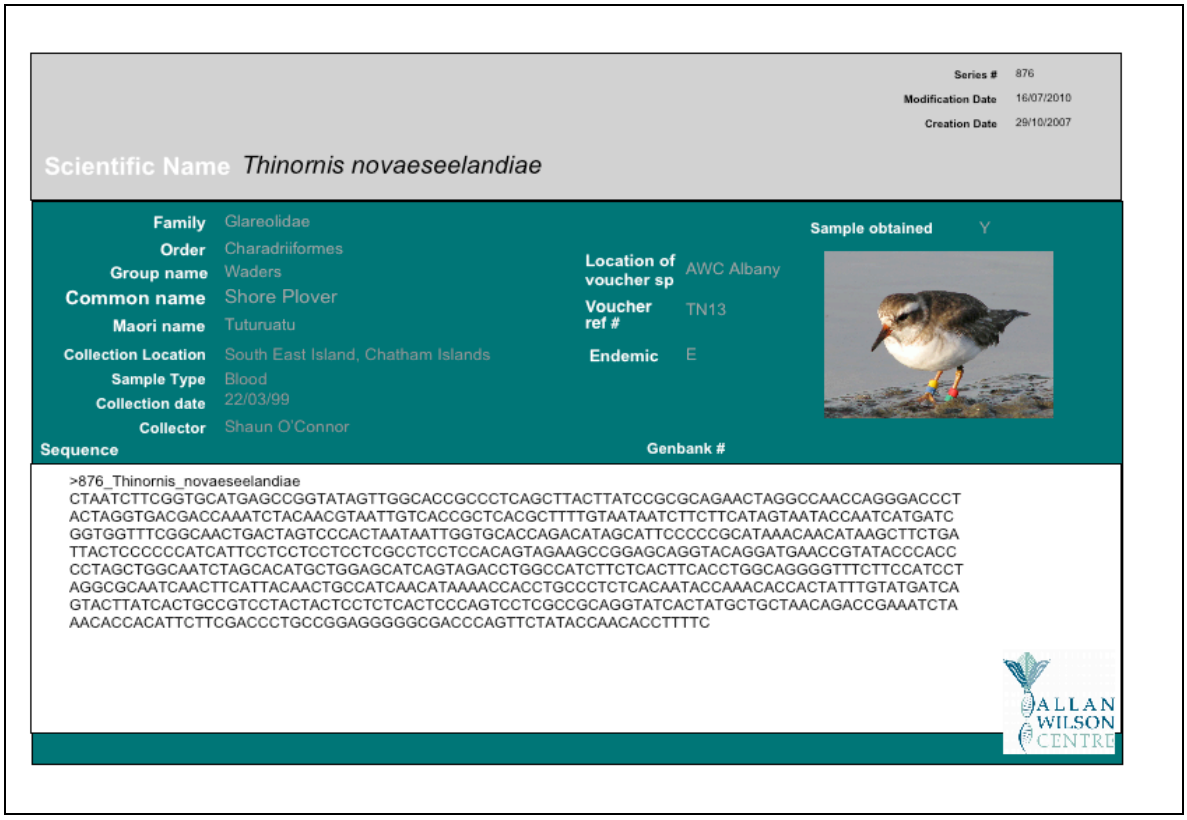


Figure 2.1 An example of a data form from the database of birds sampled. Initially a form such as the one above was completed for all specimens received, with sequence details being added when obtained. The database software used was FileMaker Pro 8.5 v2©

As the project progressed, the Barcode of life Data Systems administration developed and provided a standard format spreadsheet divided into four sheets: voucher information, taxonomy, specimen details and collection details (see

Appendix B). In addition, they provided protocols for data, sequence trace and picture submission to the BOLD database. Following the development of these spreadsheets and protocols, data were directly entered into the submission spreadsheets and these were uploaded to the BOLD website; however a record of each sample was still kept on the FileMaker Pro database. Data for all specimens can be recovered from the BOLD database by opening the Barcoding New Zealand Birds project, selecting view all records and then clicking on the sequence page for sequence details or the specimen page for specimen details (Figure 2.2).

Figure 2.2 Part of the Barcoding New Zealand Birds project on the BOLD website, indicating how to access sequence or specimen data for each sample.[5] Details of the handling history of specimens and sequences can be obtained by clicking the specimen number or sequence number (circled in the figure), which provide a link to this information.

2.4 DNA extraction, amplification and sequencing of fresh samples

The following are protocols followed by the author of this thesis when carrying out extraction, amplification and sequencing procedures. Similar protocols were carried out in Allan Baker's laboratory where some of the samples were processed.

All muscle tissue was macerated with a sterile scalpel blade prior to extraction. DNA was extracted from samples using either standard proteinase K digestion and a phenol/chloroform extraction followed by ethanol precipitation or for some samples, by one step Chelex® 100 resin (Bio-rad Laboratories)/proteinase K digestion. For the latter method, approximately 2µL of blood or 1mm^3 of tissue or 1-2 feather bulbs were digested overnight at 55°C (with slow end over end rotation) in 80µL containing 40µL of Chelex100/water slurry, 38µL of MilliQ water and 2µL of 10mg/mL proteinase K. The next day 1-2µL of the upper layer of the mix was used directly for DNA amplification by polymerase chain reaction (PCR). Negative controls were included for each set of extractions.

PCRs were carried out in 25µL or 10µL volumes consisting of 10x PCR buffer (Invitrogen), 2.5mM $MgCl_2$, 1mg/mL BSA, 0.5µM of each primer, 0.2mM of each dNTP, 0.1U of Platinum *Taq* DNA polymerase (Invitrogen) and 1-2µL of DNA extract (~0.1-5ng).

The primers used for amplification of the barcoding region (648bp starting at base 58 from the 5' terminus of COI) were a forward primer, AWCF1 (5'-CGCYTWAACAYTCYGCCATCTTACC) and a reverse primer AWCR6 (5'-ATTCCTATGTAGCCGAATGGTTCTTT).[6] For some samples an alternative reverse primer, AWCR3 (5'-ATGCTCGGGTGTCTACGTCT) was used.[6] DNA was amplified using an Applied Biosystems GeneAmp 9700 thermal cycler and the cycling conditions for samples was as follows: 2 min at 94^{o}C followed by 30 cycles of 30s at 94^{o}C, 20s at 54^{o}C and 30s at 72^{o}C followed in turn by 15 cycles of 30s at 94^{o}C, 20s at 50^{o}C and 30s at 72^{o}C.

Amplification products were visualised by electrophoresis in 1.2% agarose/0.5 x Tris Borate EDTA buffer, stained with 50ng/mL ethidium bromide and viewed under UV light (Figure 2.3). Positive amplifications were purified using the Agencourt AMPure PCR Purification System (Beckman Coulter) or the DNA Clean & Concentrator-25™ kit (Zymo Research). All purified PCR products were then cycle sequenced in both forward and reverse direction using Big Dye v3.1 chemistry and analysed on an ABI Prism 3130xl genetic analyser.

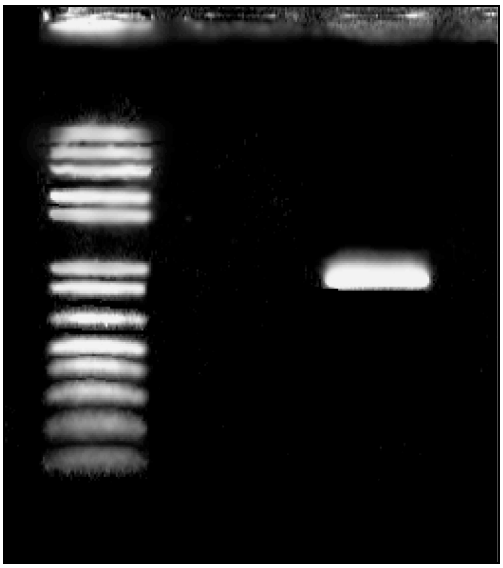


Figure 2.3 Amplification product of approximately 700bp covering the DNA barcode region of the COI gene, visualised by electrophoresis in 1.2% agarose/0.5 x Tris Borate EDTA buffer, stained with 50ng/mL ethidium bromide and viewed under UV light.

Forward and reverse sequences were aligned and edited down to the 648bp barcode segment using Sequencher 4.6 (Gene Codes Corporation). The majority of the sequences were also edited for quality. Only those sequences which achieved PHRED scores of 20 or above for at least 90% of individual bases were acceptable.

2.5 DNA extraction, amplification and sequencing of historical samples

During the design of primers for the birds of New Zealand, two sets of internal primer pairs were developed for dealing with fragmented DNA from historical samples (see Chapter 3). One set amplified three segments that, when spliced together covered the entire 648bp barcode region of the COI gene. The other

amplified five segments. Both proved equally effective in this application. However the protocols for treatment of samples for the testing of these primers were somewhat different from those used for fresh samples.

Dried skins used to test the efficacy of these internal primers were provided by Canterbury Museum. Samples collected before about 1970 were preserved using arsenical soap. After this date, borax was used. However, preparation of the skins did not involve any special preservation on the legs and toe pads. These were allowed to dry out naturally. All samples were taken from the toe pad region of specimens.

Between 5 and 64mg of tissue was used per extraction and all tissue was macerated with a sterile scalpel blade prior to extraction. DNA isolation from dried skins was preformed using proteinase K/DTT/SDS digestion followed by a DNeasy tissue extraction kit for the post lysis steps according to the manufacturer's instructions (Qiagen). Negative controls were included for each batch of extractions. No more than five samples were extracted per batch. All pre-PCR steps were carried out in a separate laboratory dedicated to ancient DNA work.[6]

Protocols for PCR were similar to those used for fresh samples except that 1X PCR buffer (Invitrogen) was used instead of 10X PCR buffer and 4% DMSO was used with the 5' tagged primers. In addition, a slightly larger quantity of DNA template (3-4μL) was used. The thermal cycling conditions differed and were as follows: 2 min at 94°C followed by 10 cycles of 20s at 94°C, 20s at 55°C and 20s at 72°C followed in turn by 30 cycles of 20s at 94°C, 20s at 50°C and 20s at 72°C and a final 4 min at 72°C.

PCR products were visualised and sequenced in the same way as fresh samples. Once the fragments were sequenced, the complete DNA barcode for each historical sample was assembled by concatenating sequences using the software programmes Sequencher 4.6 (Gene Codes Corporation) and Vector NTI (Invitrogen).

2.6 Analyses of sequence data

Following editing of sequences, they were uploaded in batches to the BOLD website. Firstly, the data spreadsheet was completed and emailed to BOLD (Appendix B). On receipt of these, BOLD data managers uploaded the data in them, creating new entries in the Barcoding New Zealand Birds project. Individual specimens could be uploaded directly to the project but this proved inefficient when dealing with large numbers of samples. Once this was completed, the sequence files were uploaded. What follows is the protocol for sequence file uploading.

The data.xls spreadsheet was completed with all the data relating to the submitted sequence files (Figure 2.4). A list of the files was then created in a folder by opening a terminal window, navigating to the folder where the trace (ab1 or scf) and score (phd) files were placed and then running the following commands:

MacOS ls *.ab1>ab1.txt **or** ls *.scf>scf.txt **and** ls *.phd.1 > phd.txt

These commands generate lists of all the files in the current folder. The files were then saved into text files called ab1.txt or scf.txt and phd.txt. Within the text files, the data were moved into appropriate columns. Process IDs were obtained by clicking on “Data Spreadsheets” under the Downloads menu on the left side of the project console. Process IDs that were assigned to each Sample ID were found in the Core Lab Book.

Filename (.ab1)	Score File (.phd.1)	Forward PCR Primer	Reverse PCR Primer	Sequencing Primer	Read Direction	Process ID
JW874.AWCF1.ab1	JW874.AWCF1.phd.1	AWCF1	AWCR6	AWCF1	Forward	NZCOI001-08
JW874.AWCR6.ab1	JW874.AWCR6.phd.1	AWCF1	AWCR6	AWCR6	Reverse	NZCOI001-08
JW779.AWCF1.ab1	JW779.AWCF1.phd.1	AWCF1	AWCR4	AWCF1	Forward	NZCOI002-08
JW779.AWCR4.ab1	JW779.AWCR4.phd.1	AWCF1	AWCR4	AWCR4	Reverse	NZCOI002-08
JW780.AWCF1.ab1	JW780.AWCF1.phd.1	AWCF1	AWCR4	AWCF1	Forward	NZCOI003-08
JW780.AWCR4.ab1	JW780.AWCR4.phd.1	AWCF1	AWCR4	AWCR4	Reverse	NZCOI003-08
JW782.AWCF1.ab1	JW782.AWCF1.phd.1	AWCF1	AWCR4	AWCF1	Forward	NZCOI004-08
JW782.AWCR4.ab1	JW782.AWCR4.phd.1	AWCF1	AWCR4	AWCR4	Reverse	NZCOI004-08
JW783.AWCF1.ab1	JW783.AWCF1.phd.1	AWCF1	AWCR4	AWCF1	Forward	NZCOI005-08
JW783.AWCR4.ab1	JW783.AWCR4.phd.1	AWCF1	AWCR4	AWCR4	Reverse	NZCOI005-08
JW784.AWCF1.ab1	JW784.AWCF1.phd.1	AWCF1	AWCR4	AWCF1	Forward	NZCOI006-08
JW784.AWCR4.ab1	JW784.AWCR4.phd.1	AWCF1	AWCR4	AWCR4	Reverse	NZCOI006-08
JW785.AWCF1.ab1	JW785.AWCF1.phd.1	AWCF1	AWCR4	AWCF1	Forward	NZCOI007-08
JW785.AWCR4.ab1	JW785.AWCR4.phd.1	AWCF1	AWCR4	AWCR4	Reverse	NZCOI007-08
JW786.AWCF1.ab1	JW786.AWCF1.phd.1	AWCF1	AWCR4	AWCF1	Forward	NZCOI008-08
JW786.AWCR4.ab1	JW786.AWCR4.phd.1	AWCF1	AWCR4	AWCR4	Reverse	NZCOI008-08

Figure 2.4 An example of a data.xls sheet to accompany trace (.ab1 or .scf) and quality (.phd) files for uploading to the BOLD website.[5]

These components (Trace files, Score files and Spreadsheet) were then placed in a folder and compressed into a single file of not greater than 190MB before uploading to the BOLD website. Uploading entailed clicking on the link “Trace Files” (Figure 2.5) in the Uploads panel of the project. The zipped folder of files was then selected and the “submit” button pressed.[5]

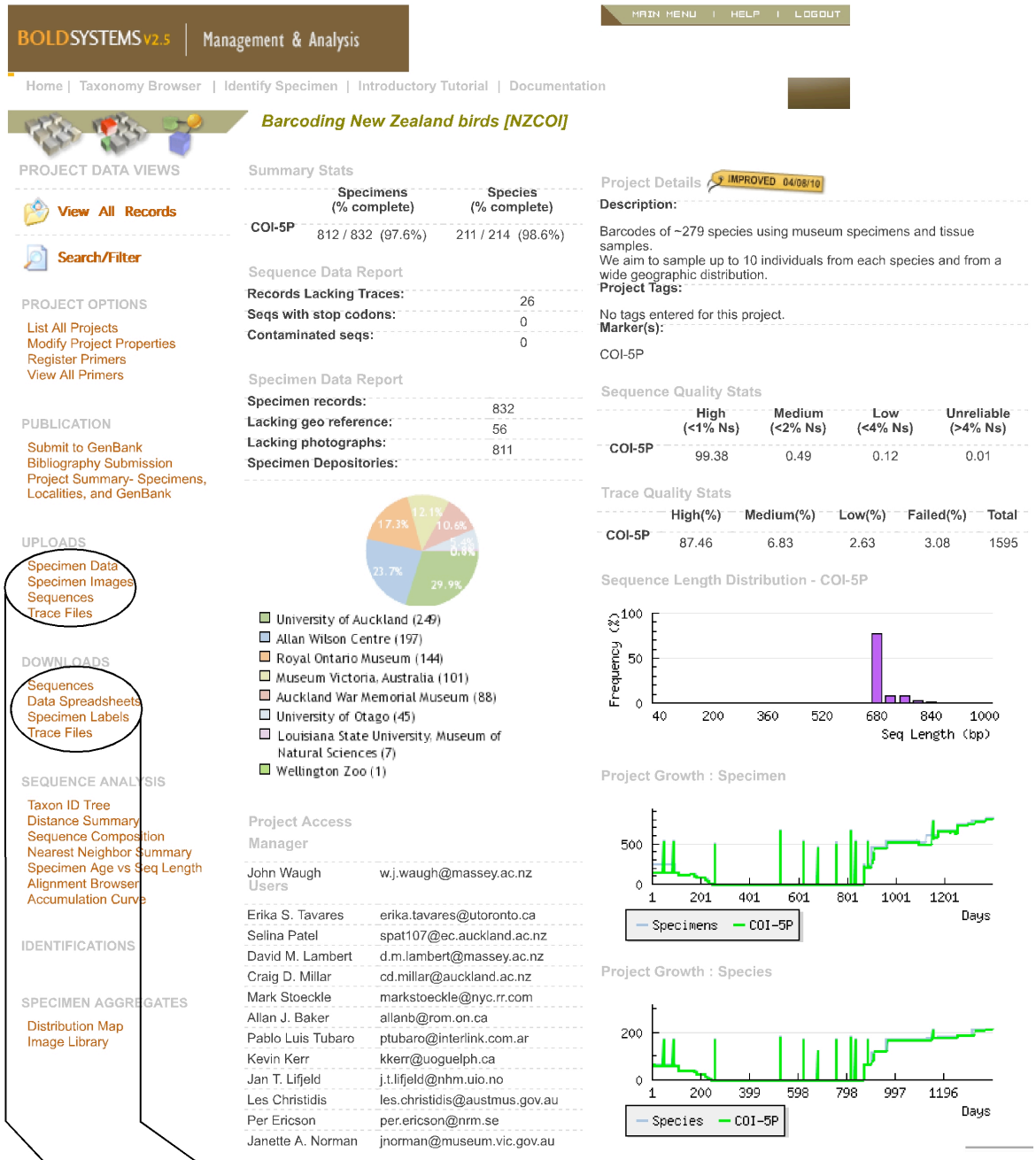


Figure 2.5 Summary page for the Barcoding New Zealand Birds project on the BOLD website[5] highlighting the links to the trace file upload and download protocols.[5]

2.7 References

1. Ratnasingham, S. and P.D.N. Hebert, *BOLD: The Barcode of Life Data System (www.barcodinglife.org)*. Molecular Ecology Notes, 2007. **7**(3): p. 355-364.

2. Stoeckle, M.a.W., K., *A global snapshot of avian tissue collections: state of the enterprise.* Auk, 2009. **126**: p. 684-687.

3. Seutin, G., White, BN, Boag, PT, *Preservation of avian blood and tissue samples for DNA analyses.* Canadian Journal of Zoology, 1991. **69**: p. 82-90.

4. *Cornell Laboratory of Ornithology*. 2009 [cited 2009 26th November]; Available from: http://www.birds.cornell.edu/.

5. *Barcode of Life Data Systems*. 2009 [cited 2010 22nd July]; Available from: http://www.barcodinglife.org/.

6. Patel, S., Waugh, J, Millar, CD, Lambert, DM, *Conserved primers for DNA barcoding historical and modern avian samples.* Molecular Ecology Resources, 2009. **10**(3): p. 431-438.

Chapter 3

Primer design

3.1 An overview of primer design

The design of suitable primers is central to any DNA barcoding project. The effective amplification and sequencing of the 648bp barcode region of the COI gene is dependent upon primers suitable to the study species. These should bind sufficiently close to the barcode region to allow adequate signal strength through the region of interest and far enough from this same region so that the noise of early read sequence does not encroach upon it. The binding sites and in particular the 3' end of the binding site must be highly conserved within the study species. This is relatively easy when there is only one species or perhaps a few similar species being investigated but in a project such as Barcoding the Birds of New Zealand, where the ideal primers must accommodate several hundred species from 16 orders, it poses a considerable challenge. Furthermore, as far as possible, the primers should not incorporate regions of nuclear mitochondrial DNA transpositions (numts). Guidelines for optimising primer design include:

- Primers should be approximately 18-22 bases in length. This is long enough for adequate specificity, and short enough for primers to bind easily to the template at the annealing temperature
- Base composition should be 40-60% (G+C).
- Primers should end (3') in a G or C, or CG or GC.
- Primer Melting Temperature (T_m), the temperature at which one half of the DNA duplex dissociates to become single stranded, should be in the range of 52-58°C.
- Primer secondary structures such as hairpins, self-dimers or cross-dimers should be avoided.
- Three or more Cs or Gs in a row at the 3'-end of primers should also be avoided as this may promote mispriming at G or C-rich sequences[1]

Within these parameters, which represent ideal circumstances that seldom prevail, there is a degree of flexibility and they should be viewed as

guidelines only.

A range of published primers is available through CBOL,[2] ABBI[3] and other sources (Appendix C);[4-9] however, these are customised for particular groups and may not be suitable for even closely related species. Furthermore, at the outset of this project, there was only a small fraction of the number of primers that are available today. Nonetheless, the starting point for most projects is to try those primers already published and see how effective they are for the species to be barcoded.

3.2 Primer design for fresh samples from New Zealand Birds

3.2.1 Introduction

In the case of the birds of New Zealand, although some groups, particularly within the Passerines, had had much of their barcode region amplified using published primers, many had not. Furthermore, the binding sites of these primers were either too close to the barcode region, resulting in the first ten or twenty bases of the segment being unreadable, or too far from the region, causing that part of the barcode region furthest from the primer to have poor signal and low quality. Thus, new primers had to be developed and, because of New Zealand's diverse avifauna, developing a set of universal primers required considerable analysis and testing.

3.2.2 Method

One hundred and forty five mitochondrial sequences from New Zealand and Antarctic birds available on GenBank (see Appendix D for accession numbers) plus one standard edited sequence from a domestic chicken were aligned using ClustalX version 0.1 (copyright 2003 Ramu & Co.). The alignment took four days to run on an Apple Macintosh Powermac G5 Quadcore with a CPU speed of 2.5GHz. It included the barcode region plus 1000bp on either side of it (Figure 3.1) and was used to locate highly conserved areas either side of the barcode region, which would be suitable to test as primer binding sites. Ideally, these sites should be ≤200 and ≥40bp from the region of interest.

Figure 3.1 Part of the mass alignment used to identify suitable locations for primer binding sites. It is apparent from this alignment that the first two bases of each codon are generally highly conserved and that the third base is not, and this provides the variability that DNA barcoding utilises to identify species.

Suitable sites of nucleotides conserved across all or most species were identified and putative primers designed to fit them. The primers were tested using Amplify 3.1.$_4$ for Mac OSX, a Macintosh freeware program for simulating and testing polymerase chain reactions available from

http://engels.genetics.wisc.edu/amplify/.

Those that achieved the highest test scores were ordered from Sigma-Aldrich Biotechnology and laboratory tested on fresh samples, including blood, muscle and feather tissue. Several primers amplified the barcode region but three proved more effective than others and amplified all samples from the species of New Zealand birds tested (Table 3.1).

Table 3.1 Preliminary results of primer testing during the Barcoding New Zealand Birds project. The names of some of these primers were changed to a standardised format (e.g. AWCR6) on acceptance

Primer name	No. of potential dimers	Sequence	Comments	Tm
intF1	1	CAGAACTAGGCCAACCAGG	Slight move to right might help. Needs elongation	53˚C
intR1	-	TTATGCGTGGGAATGCTATG	Appears quite good. Needs elongation	51˚C
intF2	1	TTATAATCGGTGCCCCAGAC	Slight alteration needed. Needs elongation	51˚C
intR2	-	TGCTAGGTGGAGGGAGAAGA	Modification needed. Two potential R binding points	54˚C
intF3	-	TTTATCACAACCGCCATCAA	Seems OK for some spp. but needs adjustment left for broader utility. Needs elongation	50˚C
intR3	1	TGCTCCTAGGATGGAGGAGA	Two binding sites (272-291 and 524-543) problematic	54˚C
intR4	-	TTGATGGCGGTTGTGATAAA	This looks better than results suggest. Could be improved. Needs elongation	50˚C
intR5	-	GGGTCGAAGAATGTGGTGTT	Looks OK. Needs elongation	51˚C
taka1rev	1	CAAAGGCATGTGCGGTGACAATT	Needs left shift	57˚C
taka2fwd	1	CGGCAGCCTATTAGGTGATGACC	Needs to be altered to avoid primer dimer	58˚C
taka2rev	-	CTAGTAGTAGTAGGAAGGATGG	Looks good. Needs elongation	50˚C
taka3fwd	-	CAACATAAGCTTCTGACTCCTT	Looks good. Needs elongation	52˚C
taka3rev	1	ATGTTGATGGCAGTTGTAATGA	See intR4. Needs elongation	51˚C
taka4fwd		GTGTCTCATCCATTCTAGGT	Doesn't seem to work at all	48˚C
taka4rev	1	TCTCCTCTGCCAGCAGGGTC	OK but produces dimer	57˚C
taka5fwd	-	ACCGACCGAAACCTAAACAC	Seems OK. Needs elongation	51˚C
taka5rev	1	GATGGAGAGTATGGCCCATA	Outside COI. Needs elongation	50˚C
intF4	-	GGATGAACTGTCTACCCACCATTAGC	Needs testing	57˚C
intF5	-	CAAACACCCCTGTTCGTATGCTCC	Needs testing	58˚C
intR6		GTTTGGTACTGTGATAGCGGTGG	Needs testing	56˚C
intR7		ATTGTGATGCCGGCAGCAAGAA	Needs testing	58˚C
impF1	-	CCGAGCAGAACTACGTCAACC	Needs testing	56˚C
impR1	-	ATGTTGTTTATGAGTGGGAATGCTATG	Needs testing	56˚C
impF2	-	CCACTTATAATCGGTGCCCCAGAC	Needs testing	57˚C
impR2	-	ACCTGCTAGGTGGAGGGAGAAGA	Needs testing	58˚C
impF3	-	TCAACTTTATCACAACCGCCATCAA	Needs testing	56˚C
impR4	-	TCATGTTGATGGCGGTTGTGATAAA	Needs testing	56˚C
impR5	-	TGCTGGGTCGAAGAATGTGGTGTT	Needs testing	58˚C

Ultimately, three external primers, one forward (AWCF1) and two reverse (AWCR3 and AWCR6), were designed, which successfully amplified the entire 648bp COI region as a single fragment from fresh samples of all the bird species that were tested.

3.3 Primer design for historical* samples from New Zealand birds

3.3.1 Introduction

The majority of DNA barcoding studies on birds have been carried out using fresh tissue samples from Northern Hemisphere species.[5, 6, 9-11] Many major museums in the Northern Hemisphere have developed collections of samples from vouchered specimens for biochemical and DNA analysis.[12, 13] These samples are either preserved in ethanol or at a low temperature and this form of tissue preservation is spreading to museums around the world. However, the majority of New Zealand's vouchered specimens were collected from the late 19th to the mid 20th century. Relatively few specimens have been collected since the 1970s and, thus, most predate DNA preservation techniques. There are limited tissue samples available for standard DNA barcoding techniques in this region. Moreover, the collection and cataloguing of new specimens is in many cases difficult and expensive.[14]

Historical specimens tend to be >20 years old and have not been preserved in a suitable manner for DNA analysis. The DNA from these specimens is generally highly degraded.[15] This fragmented DNA is difficult to amplify and seldom produces amplicons larger than 300-400bp.[16]

There are three common approaches used for obtaining DNA barcodes from historical specimens or specimens whose DNA may otherwise have been degraded. These include the use of enzymes to repair DNA damage caused by chemical treatment or aging processes, amplifying short or "mini" barcodes that are 100-200bp fragments of the 648bp barcode region, or amplifying several overlapping small segments of the barcode region, so that the entire barcode region can be assembled from these fragments.

The use of DNA-repairing enzymes improves amplification success by

* The term "historical samples" is used throughout to differentiate them from ancient samples (thousands to hundreds of thousands of years old or more) and refers to more recent samples (decades to hundreds of years old) not preserved with DNA recovery in mind and which may contain degraded DNA because of age, desiccation or preservation technique.

repairing damage to DNA caused by aging or chemical treatment. Damage to historical samples includes oxidative and hydrolytic damage and deamination, resulting in the replacement of cytosine with uracil and, consequently, C to T and G to A transitions during PCR.[17, 18] DNA repairing enzymes can repair most of these types of damage; [17] but, are unable to splice fragmented DNA and are too costly to use on a large scale.

Mini-barcodes have proved to be successful for identifying almost all species in several insect genera[19] but they are less accurate at identifying species when compared with full-length barcodes and do not work with all taxa.[20] During the barcoding of birdstrike specimens (Chapter 6), some accurate identifications of bird samples, contaminated with aviation lubricants and cleaning agents, were made using fragments of the barcode region.[21] However, no comprehensive attempt to apply mini-barcoding to avian species has been undertaken. Nor, since fragments from historical samples are not necessarily uniform, does this approach remove the need to construct a database of full-length barcode region sequences against which to match the fragments in order to make these identifications. Thus, a relatively short sequence may be sufficient to identify a specimen but it is still necessary to have complete barcodes in the reference database.

The third approach is based on the principle that, in general, when recovering DNA sequences from historical samples, there is an inverse relationship between amplicon size and the rate of amplification success. Thus, by amplifying a number of small overlapping regions, a complete sequence of the 648bp COI barcode region can be assembled from fragmented DNA in historical samples. This approach has been successfully used with ancient DNA[22] and provides a means of generating some of the DNA barcodes for the birds of New Zealand database using historical samples. Central to this approach is the development of suitable internal primers for amplifying the overlapping fragments. During this study, two sets of PCR primers that amplify the COI barcoding region from historical samples in three or five short segments were developed.

3.3.2 Method

Once again, primers for the amplification of historical and modern tissues were designed based on the alignment (using the Clustal X algorithm) of DNA sequences obtained from New Zealand and Antarctic avian species plus a number of COI sequences available on GenBank. In this case, highly conserved regions *within* the barcode region were identified as areas of interest for the synthesis of potential internal primers. Primer integrity was checked using Amplify 3.1.$_4$ for Mac OSX and NetPrimer available from the Premier Biosoft International website.[1]

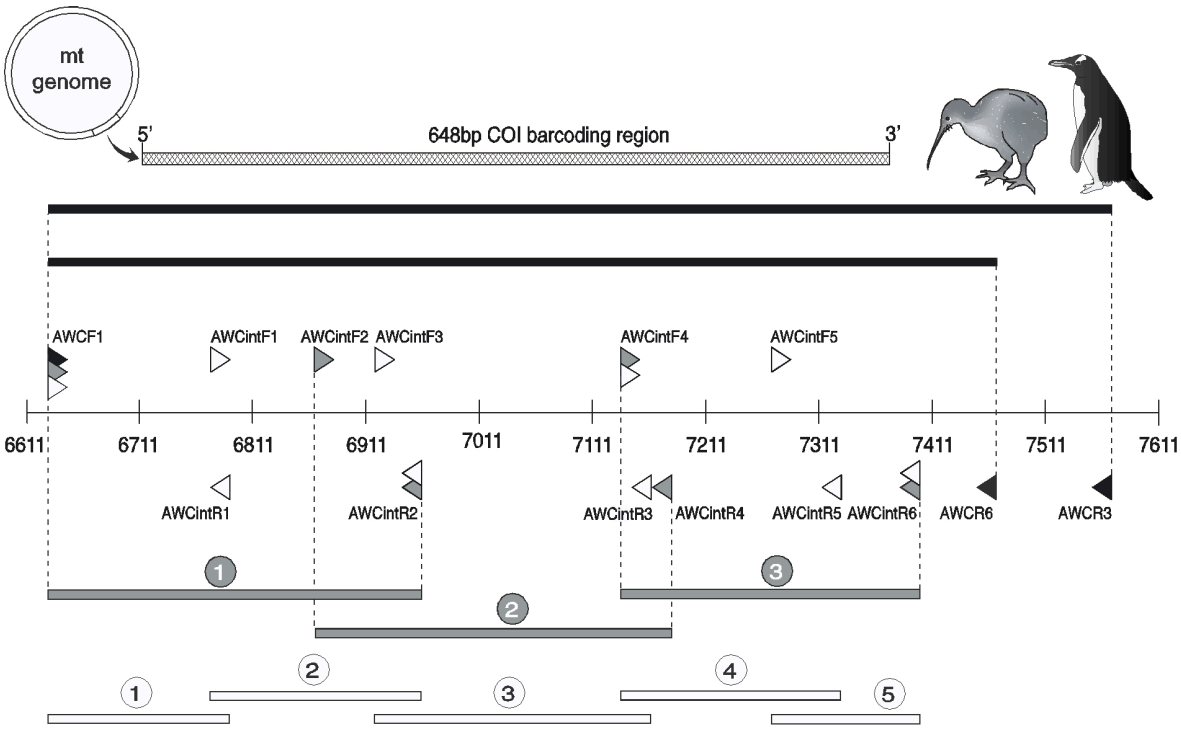


Figure 3.2 A diagrammatic representation of the COI barcoding region used for bird identification with the relative position of the primers (triangles) indicated for amplifying historical and modern samples. The three black triangles and the two black bars indicate the combination of primers used, and resulting DNA fragments generated, for modern samples. Similarly, the grey/white triangles and the grey/white bars represent the combination of primers used and resulting DNA fragments generated, for historical samples. The position of each primer and the amplified DNA fragments are given relative to the chicken mitochondrial sequence.

Two sets of internal primers were designed to span the 648bp COI region. The first set contained three primer pairs resulting in three overlapping fragments (average length ~310bp; Figure 3.2) and the second primer set contained five primer pairs resulting in five shorter overlapping fragments

(average length ~ 190bp; Figure 3.2) to span the same 648bp region. Each internal primer contained a randomly generated 18bp tag attached to its 5' end in order to improve the sequencing of the resulting PCR products. Forward internal primers all contained an identical Ftag sequence while reverse internal primers contained an identical Rtag sequence. Internal primer pairs were tested on a range of modern samples obtained from 15 avian orders prior to being used to amplify DNA from historical samples. The external forward primer, AWCF1 was also used as part of the primer set for historical samples as were the reverse primers, AWCR3 and AWCR6 but were modified with Ftags and Rtags when used in this application. Primer pairs intended for modern specimens did not contain the 18bp tags. All primer sequences and their expected fragment lengths are shown in Table 3.2.

3.4 Results

Two sets of 5' tagged PCR primers were designed and tested, which amplified short fragments of the barcode region from historical samples. One set produced three fragments of approximately 310bp. The second produced five shorter fragments of approximately 190bp. Following testing on modern samples from 15 avian orders, historical samples (age range 16-115 years) from 16 of 27 species (59%) from 9 different orders were sequenced using both primer sets (Table 3.3).[14] Since the success rate of the three fragment and five fragment primer sets on historical samples was similar, the three primer pair set appears to be the best option. AWCF1/AWCR6, that amplify the entire barcode region, were used as a control with no products being amplified.

Historical samples from the Anseriformes, Pelecaniformes and Psittaciformes were not successfully sequenced. However, since all but one specimen from these orders were aged >80 years and since their modern representatives *were* successfully amplified, the problem may be related to age rather than intrinsic problems with these orders. An alternate final primer pair (AWCintF6/AWCintR7) was designed for the Charadriiformes and Anseriformes since AWCintF5/AWCintR6 were ineffective for some

species within these orders.

Table 3.2 PCR primers and 5' tags used for the amplification of the 648 bp COI region as one, three or five fragments. The position of each primer is given relative to the chicken mitochondrial sequence. The 5' tags were only added to the individual primers when PCR amplifying and sequencing ancient samples as either three or five fragments. AWCintF6 and AWCintR7 represent alternate primer pairs used for the orders Charadriiformes and Anseriformes

Forward primer	Sequence 5' to 3'	Position (bp)*	Reverse Primer	Sequence 5' to 3'	Position (bp)*	Fragment size (bp)*
Primers for modern samples						
AWCF1	CGCYTWAACAYTCYGCCATCTTACC	6625–6649	AWCR6	ATTCCTATGTAGCCGAATGGTTCTTT	7446–7471	848
AWCF1			AWCR3	ATGCTCGGGTGTCTACGTCTAT	7542–7563	936
First set of internal primers containing 3 primer pairs						
AWCF1			AWCintR2	ATGTTGTTTATGAGTGGGAATGCTATG		328
AWCintF2	ATAATCGGAGGCTTCGGAAACTGA	6873–6896	AWCintR4	TGGGAKAGGGCTGGTGGTTTTATGTT	7161–7186	314
AWCintF4†	TCCTCAATCCTGGGAGCAATCAACTT	7119–7144	AWCintR6‡	GGATTAGGATGTAGACTTCTGGGTG	7371–7395	278
Second set of internal primers containing 5 primer pairs						
AWCF1			AWCintR1	CCTGGTTGACCTAGTTCTGCTCG	6765–6787	163
AWCintF1	CCGAGCAGAACTACGTCAACC	6764–6784	AWCintR2		6926–6952	189
AWCintF3	ATCGGAGCCCCAGACATAGCATT	6912–6934	AWCintR3	TTGATGGCTGTTGTGATAAAGTTGAT	7137–7162	251
AWCintF4†			AWCintR5	TGCTGGGTCGAAGAATGTGGTGTT	7299–7322	204
AWCintF5	GGCATCACCATACTACTAACAGACCG	7266–7291	AWCintR6‡			130
5' tags for internal primers						
Ftag	AGTCGACGCTTCTAGCTT	—	Rtag	CATGCTACCTGCTACTGT	—	

*Basepairs.
†AWCintF6 – TAGGGGCAATCAACTTCATCACAAC (7129–7153).
‡AWCintR7 – ACGTRTGAGATAATTCCGAATCC (7401–7423).

A number of previously described avian COI primers[5] designed to amplify a single fragment for modern avian samples from the Northern Hemisphere failed to amplify DNA from many New Zealand and Antarctic birds (data not shown). This could be attributed to the fact New Zealand and the Antarctic have the largest number of endemic seabirds compared with other geographic regions such as North America for which these avian primers were originally designed. Moreover, orders like the Apterygiformes (kiwi) are unique to this region.

Table 3.3 Results of PCR success of modern and historical samples for the two primer sets designed to amplify DNA from historical specimens. PCR success is defined as fragments amplified (Y = amplified; N = not amplified; NT = not tested)

	Modern specimens			Historical specimens		
Order	**3 fragments**	**5 fragments**	**No. of species**	**3 fragments**	**5 fragments**	**No. of Species**
Anseriformes	Y*	Y*	2	N	N	0/2
Apterygiformes	Y	Y	2	NT	NT	—
Charadriiformes	Y*	Y*	2	Y*	Y*	2/4
Ciconiiformes	Y	Y	2	Y	Y	2/2
Columbiformes	Y	Y	2	NT	NT	—
Coraciiformes	Y	Y	2	NT	NT	—
Falconiformes	Y	Y	2	Y	Y	1/2
Galliformes	Y	Y	2	NT	NT	—
Gruiformes	Y	Y	2	Y	Y	1/2
Passeriformes	Y	Y	2	NT	NT	—
Pelecaniformes	Y	Y	2	N	N	0/2
Podcipediformes	Y	Y	1	NT	NT	—
Procellariiformes	Y	Y	2	Y	Y	9/10
Psittaciformes	Y	Y	2	N	N	0/1
Sphenisciformes	Y	Y	2	Y	Y	1/2

*Not all primer pairs were successful, alternate primer pairs have been supplied for these orders (see Table 3.2)

The results of primer success using the newly designed primers AWCF1, AWCR6 and AWCR3, are shown in Table 3.3. The only previously published avian primer to have some success on New Zealand and Antarctic birds was the COIbirdR2 primer[23] when paired with the AWCF1 primer. This primer pair worked extremely well for passerines (Table 3.4). The newly designed AWCF1/AWCR6 primer pair worked well for most species, with the forward primer AWCF1 proving to be universally successful at binding appropriately and the main reverse primer AWCR6 proving to be highly successful in the majority of cases. In avian species, for

which the AWCR6 primer did not prove successful, the reverse primer AWCR3 appeared to work instead. AWCR3 worked particularly well for the Apterygiformes, where all five species of this order could only be amplified with the AWCF1/AWCR3 primer pair.

Table 3.4 PCR primer success on 17 different avian orders. The numbers represent the species of each order from which the target sequence was successfully amplified, using different combinations of PCR primers. **— = failed to amplify**

Order	AWCF1/R6	AWCF1/R3	AWCF1/R6 & AWCR3	AWCF1/ COIbirdR2	No. of species per order
Anseriformes	16	—	—	1	16
Apterygiformes	—	5	—	—	5
Charadriiformes	25	4	4	1	25
Ciconiiformes	9	—	—	2	9
Columbiformes	3	—	—	3	3
Coraciiformes	1	—	—	1	1
Cuculiformes	2	—	—	—	2
Falconiformes	2	—	—	1	2
Galliformes	1	—	—	1	1
Gruiformes	7	3	2	4	8
Passeriformes	17	6	1	20	20
Pelecaniformes	7	2	—	—	7
Podcipediformes	1	2	—	—	3
Procellariiformes	33	5	3	5	35
Psittaciformes	4	3	2	1	5
Sphenisciformes	6	1	1	1	6
Strigiformes	1	1	1	—	1
Total	**135**	**32**	**14**	**41**	**149**

3.5 Discussion

The development of primers that will reliably amplify the region or regions of interest is central to the success of the Barcode of Life project. To date, approaching 400, including 9 from the Barcoding New Zealand Birds project, have been registered with BOLD (Appendix C). Because of the variability that occurs at the barcode region of the COI gene, different taxonomic groups may require specific primers and, since these groups are not uniformly distributed about the planet, isolated countries such as New Zealand can and do experience difficulties with primers that have been developed elsewhere. Moreover, in the case of historical samples, the 648bp COI barcoding region can only be recovered in smaller fragments requiring multiple primer sets leading to the idea that primers must be customised to suit.

However, the generation of taxon-specific primers is an impractical one. A realistic solution to this problem is to generate a small set of primers that will work effectively on all avian species, both historical and modern. The primers generated for the Barcoding New Zealand Birds project are an example of such sets (Figure 3.2). These primers, designed to amplify the barcode region of historical and modern samples from New Zealand and Antarctic bird species proved generally successful with only a few exceptional taxa, such as the Apterygiformes requiring specialised primers. It is probable that these primers may be effective for avian orders from other geographic regions. All primer pairs tested generated high quality sequence for the region of interest.

Primers designed for modern samples had binding sites that were neither too close nor too far from the barcode region, allowing for good overlapping bi-directional sequence while avoiding the loss of sequence data which occurs at the beginning of each sequence read.

To maximise chances of obtaining amplifiable DNA from historical specimens samples were taken from the toe pad area where the risk of damage from preservatives is minimal. Specimen age should also be carefully considered and where possible, younger specimens should be selected for DNA analysis.

The 648bp barcode region for historical avian specimens can now be amplified using tagged internal primers as either three or five fragments, depending on the DNA quality of individual samples. The set of primers containing three primer pairs should be used in preference because the likelihood of achieving success with the set of three primers pairs is the same as that with the set of five primer pairs. If one set of internal primers does not prove successful the other set is not likely to be either. In the event that not all overlapping fragments are amplified there is the added possibility of using one or more of these smaller fragments as a mini barcode once this approach has been confirmed to work for all avian taxa.

The addition of 5’ tags to each of the internal primers provides a number of advantages. Not only does it result in an increase in the read length of each sequence, thus allowing for better overlap of individual contigs, but it has been shown to greatly improve sequence quality.[24] This is crucial for historical samples where DNA damage is frequent. An added advantage is the possibility of higher throughput sequencing, as only two primers (forward and reverse tags) are required instead of the many individual forward and reverse primers.

The disadvantage, however, is that the processes of extracting DNA, PCR amplification and sequencing from historical specimens can be taxing and time consuming. Thus, modern specimens should be utilised for barcoding purposes wherever possible. This method should be used as a way of adding to an already existing database where there is a lack of modern samples or for cases of rare and endangered species where sample collection is problematic.

There is an onus upon museums and other institutions to ensure the storage of well preserved, vouchered samples of as many species as possible, from which, high quality DNA can be extracted. The maintenance of such collections is increasingly being adopted around the world. However, until such time as these collections become comprehensive and for rare or even extinct species, study skins and bones provide the only vouchered source of DNA. While that is the case, there is a need for methods, such as those outlined above, to provide a means of generating DNA barcodes.

3.6 References

1. *PCR Primer Design Guidelines*. 2010 [cited 2010 21st July]; Available from: http://www.premierbiosoft.com/.

2. *Consortium for the Barcode of Life*. 2009 [cited 2009 11th November]; Available from: http://barcoding.si.edu/.

3. *All Birds Barcoding Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.barcodingbirds.org/#.

4. Chaves, A., Clozato, CL, Lacerda, DR, Sari, EHR, Santos, FR, *Molecular taxonomy of brazilian tyrant-flycatchers (Passeriformes: Tyrannidae).* Molecular Ecology Resources, 2008. **8**(6): p. 1169-1177.

5. Hebert, P., Stoeckle, MY, Zemlak, TS, Francis, CM *Identification of Birds through DNA Barcodes.* PLoS Biol, 2004. **2**(10).

6. Kerr, K., Stoeckle, MY, Dove, CJ, Weigt, LA, Francis, CM, Hebert, PDN, *Comprehensive DNA barcode coverage of North American birds.* Molecular Ecology Notes, 2007. **7**(4): p. 535-543.

7. Lohman, D.J., D.M. Prawiradilaga, and R. Meier, *Improved COI barcoding primers for Southeast Asian perching birds (Aves: Passeriformes).* Molecular Ecology Resources, 2009. **9**(1): p. 37-40.

8. Tavares, E.S. and A.J. Baker, *Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds.* Bmc Evolutionary Biology, 2008. **8**: p. -.

9. Yoo, H., Eah, JY, Kim, JS, Kim, YJ, Min, MS, Paek, WK, Lee, H, Kim, CB, *DNA barcoding Korean birds.* Molecules and Cells, 2006. **22**(3): p. 323-327.

10. Johnsen, A., Rindal, E, Ericson, PGP, Zuccon, D, Kerr, KCR, Stoeckle, MY and Lifjeld, JT, *DNA barcoding Scandinavian Birds reveals divergent lineages in trans-Atlantic species.* J. Ornithol, 2010. **151**(3): p. 565-578.

11. Park, H., Yoo, HS, Jung, G and Kim CB, *New DNA barcodes for identification of Korean birds.* Genes & Genomics, 2011. **33**: p. 91-95.

12. Arctander, P., Fjeldsa, J, *Avian Tissue collections for DNA analysis.* Ibis, 1994. **136**: p. 359-360.

13. Seutin, G., White, BN, Boag, PT, *Preservation of avian blood and tissue samples for DNA analyses.* Canadian Journal of Zoology, 1991. **69**: p. 82-90.

14. Patel, S., Waugh, J, Millar, CD, Lambert, DM, *Conserved primers for DNA barcoding historical and modern avian samples.* Molecular Ecology Resources, 2009. **10**(3): p. 431-438.

15. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-15.

16. Zimmermann, J., et al., *DNA damage in preserved specimens and tissue samples: a molecular assessment.* Front Zool, 2008. **5**: p. 18.

17. Evans, T., *DNA damage.* NEB Expressions, 2007. **2**(1): p. 1-3.

18. Hoss, M., Jaruga, P, Zastawny, TH, Dizdaroglu, M, Paabo, S, *DNA damage and DNA sequence retrieval from ancient tissues.* Nucleic Acids Res, 1996. **24**(7): p. 1304-7.

19. Hajibabaei, M., Smith, MA, Janzen, DH, Rodriguez, JJ, Whitfield, JB, Hebert, PDN, *A minimalist barcode can identify a specimen whose DNA is degraded.* Molecular Ecology Notes, 2006. **6**(4): p. 959-964.

20. Hajibabaei, M., Singer, GAC, Hebert, PDN, Hickey, DA, *DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics.* Trends in Genetics, 2007. **23**(4): p. 167-172.

21. Waugh, J., Evans, MW, Millar, CD, Lambert, DM, *Birdstrikes and barcoding: can DNA methods help make the airways safer?* Molecular Ecology Resources, 2010. **9999**(9999).

22. Millar, C.D., Huynen, L., Subramanian, S. Mohandesan E. & Lambert D.M., *New developments in ancient genomics.* Trends in Ecology & Evolution, 2008. **23**: p. 386-393.

23. Kerr, K.C., Lijtmaer, D A, Barreira, A S, Hebert, PDN, Tubaro, P L, *Probing evolutionary patterns in neotropical birds through DNA barcodes.* PLoS One, 2009. **4**(2): p. e4379.

24. Binladen, J., Gilbert, MT, Bollback, JP, Panitz, F, Bendixen, C, Nielsen, R, Willerslev, E, *The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing.* PLoS One, 2007. **2**(2): p. e197.

Chapter 4

Data Analysis

4.1 Overview of data

More than one third of all known bird species have had at least one (mean 6.2) DNA barcode(s) recorded on the Barcode of Life Data Systems (BOLD) database.[1] The All Birds Barcoding Initiative (ABBI)[2] is one of several initiatives overseen by the Consortium for the Barcode of Life,[3] each aimed at a particular component of the world's biota such as Lepidoptera,[4] invasive pest species,[5] polar species[6] or fish.[7] As part of its aim to establish a public archive of DNA barcodes for the approximately 10000 species of birds in the world, ABBI has divided the world into eight regions: Palaearctic (including Europe, North Africa, The Middle East, Russia and North-East Asia), Afrotropical (Central and Southern Africa), Indomalayan (India and South-East Asia), Nearctic (North America), Neotropical (South and Central America), Oceania (Pacific Ocean and its islands) Antarctic and Australasia.[2] The first region to have a DNA barcode database of its avifauna published was the Nearctic[8, 9] then Korea,[10, 11] where 154 of 450 species have had DNA barcodes recovered and recorded. In addition, large projects of 500 of the Neotropical birds of Argentina and 296 Palearctic birds of Scandinavia have also been recently published.[12, 13]

One of the regions that is absent from this group of published databases is Australasia, comprising Australia, New Zealand and a number of offshore islands. With its widely scattered landmasses and vast areas of ocean from equatorial regions to the sub-Antarctic, spanning more than 70 degrees of longitude, Australasia has a large and diverse avifauna. New Zealand, the region's most isolated major landmass 2000km to the South-East of Australia, hosts about 3% of the bird species of the world.[14] These species are characterised by a high level of endemism and include a large group of Southern pelagic species as well as the world's greatest concentration of both living and extinct ratite species. These

unique features make this an important fauna to examine and characterise using DNA barcodes. Furthermore, many species found in this region are quite distinct from those in other published projects,[8-13] having undergone a very long period of isolation or because they belong to the southern circumpolar avifauna that has not been examined in this way before. Thus, they sometimes require different approaches to the isolation and amplification of appropriate sequences of DNA (Chapter 3).

Sequences of the barcode region of the mitochondrial COI gene came from several sources. Some were obtained from samples provided by a variety of institutions and people to the author of this thesis (31%), Selina Patel at the University of Auckland (49%) and Allan Baker and his team at the Royal Ontario Museum (20%), others came from existing files in GenBank (see Appendix D for accession numbers; see Appendix E for a complete list of all specimens in the Barcoding New Zealand Birds project with a breakdown of contributing sources). All barcodes along with specimen and sequence details have been lodged with the Barcode of Life Data Systems (BOLD)[1] and, with the exception of those already lodged with GenBank, can be found in the project named “Barcoding New Zealand Birds” on this site.

Fresh tissue samples for DNA barcoding were provided by Auckland War Memorial Museum and Otago Museum in New Zealand as well as Victoria Museum in Australia. In addition, fresh samples were provided by Wellington Zoo, Whakatane Bird Rescue and Sylvia Durrant Bird Rescue in Rothesay Bay, Auckland as well as from the collections of David Lambert and Craig Millar. Historical museum samples (see Chapter 3) were provided by Paul Scofield of Canterbury Museum, Christchurch.

While the overall number of bird species found in the greater New Zealand region is approximately 328, this number includes Antarctic species, species that have been introduced by human agency, stragglers (irregular visitors), rare vagrants (transitory and unexpected), migratory birds that spend some part of their lives in

New Zealand but do not breed here, as well as extinct and probably extinct species. Of the remainder, 71 species are endemic to New Zealand (breed only in New Zealand) and a further 79 are native species (naturally occurring and breeding in New Zealand and elsewhere).[15] These last two groups form the main taxa presented in these data (Table 4.2), but additional species, which have been introduced or which fall into some of the other categories have also had their DNA barcodes sequenced and are included in these data.

According to Fleming (1976), offshore islands such as the Chatham Islands, Auckland and Campbell islands, the Snares and the Kermadecs tend to have their own representatives of many taxa due to allopatric divergence.[16] Molecular data that exists, however, do not always support this conclusion,[17] and the divergence in nucleotides observed in these birds is generally less than that normally indicative of congeneric species. In addition, the collection and sequencing of a large group such as the birds of New Zealand is subject to the law of diminishing returns. Obtaining samples for the last few species is very much more time consuming than for the first samples. For these reasons, some species, particularly those of offshore island variants of mainland species do not appear in this study but will, with time, be added to the project.

Nonetheless, for these data to be useful for the identification of New Zealand birds, they should include all species, both common and uncommon, native, endemic or other found within this region. However, all the introduced species and many of the vagrants, stragglers and migratory species have already had DNA barcodes recovered elsewhere and these are publicly available through the BOLD website and GenBank. Thus, less emphasis has been placed on these groups and many have not been comprehensively sequenced for the Barcoding New Zealand Birds project. To date, 833 DNA barcode sequences from 215 bird species (77% of the total NZ avifauna) have been lodged in the Barcoding New Zealand Birds database with BOLD. Of these, 628 sequences from 126 species are native or endemic to New Zealand representing 84% of this group.

4.2 Data analysis methods

Initial analyses of these data utilised a standard Kimura 2-Parameter (K2P)[18] genetic distance correction, both within species and between all species related at ordinal level, to produce taxon identification trees, nearest neighbour analyses and distance summaries. These analyses were carried out using the online sequence analysis tools available on the BOLD website.[1] Further analysis was carried out using Bayesian inference of phylogeny,[19] employing a variety of software to build phylogenetic trees that could be compared with those generated by the K2P method. In some selected cases, where species were not clearly differentiated by either or both of these methods, character-based identification was employed.[20]

4.2.1 Kimura 2-Parameter genetic distance correction

In general, the longer two species have diverged from each other over evolutionary time, the greater the number of nucleotide differences between them. Thus, sequence distance approximates evolutionary distance. To more precisely infer evolutionary relationships, various models of how sequence changes have occurred are applied. The simplest of these is that all possible nucleotide substitutions (e.g. tansitions and/or transversions) are equally likely and that substitutions occur at the same rate in all lineages. Using these assumptions, the number of nucleotide differences between sequences from different species can be used to calculate evolutionary distance according to the equation:[21]

$$D = k/n$$

$$Var(D) = D(1-D)/n$$

(***D*** = distance, ***k*** = the number of nucleotide substitutions and ***n*** = sequence length)

The rate of change would increase with a slope equal to twice the product of the mutation rate and time since sequences from diverging species would both mutate from the common ancestor. This divergence occurs linearly to start with but, with time, the increase slows, finally reaching an asymptote of 0.75 (Figure 4.1). This is because there are only four types of nucleotide and a random substitution based on these four possibilities would substitute an identical nucleotide for one quarter of them by chance alone. Moreover, as the time increases so the probability of a second substitution at any one site increases and, consequently, the tally of

differences between the diverging species slows.[21]

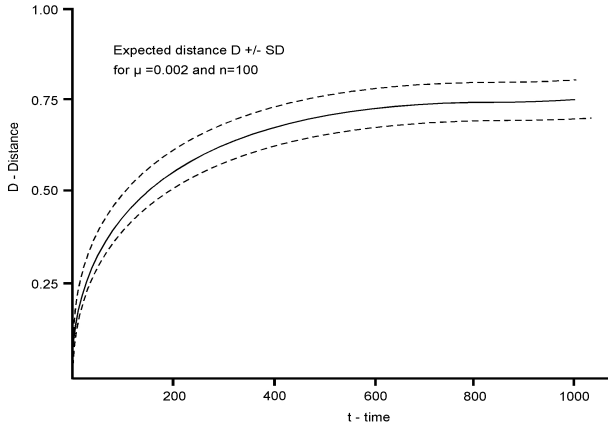


Figure 4.1 An asymptotic divergence with time in the evolution of DNA, assuming that all mutations occur with equal frequency. **μ** = mutation rate, **n** = sequence length[21]

Thus, according to Figure 4.1, a pair of diverging species at time 30 are expected to have a distance of 0.20 and at time 60, a distance of 0.40, distances which can be easily distinguished. However, a pair at time 500 and time 1000 have distances of 0.70 and 0.75, which would be more difficult to tell apart. Yet, in both cases, there is a doubling of the divergence between species pairs.[21]

The Jukes-Cantor correction to distance[22] was developed to correct this and is depicted by:[23]

$$D_{JC} = -(3/4) \log n ((1-(4/3)D)$$
$$Var(D_{JC}) = D(1-D)/[n(1-(4/3)D)^2]$$

(***D_{JC}*** = Jukes-Cantor corrected distance, ***D*** = uncorrected distance, ***n*** = sequence length)

This correction shows that divergence is a logarithmic function of time, and the plot thus formed is linear. However, a drawback of this correction is that as time increases so does the variance about the now straightened plot (Figure 4.2), to the point where as it approaches the asymptote, variance approaches infinity. Thus, as time increases, the Jukes-Cantor correction becomes less reliable.[24]

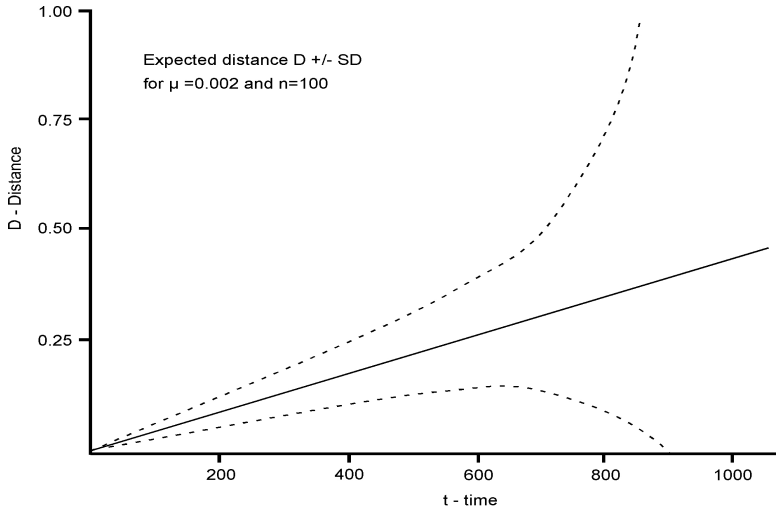


Figure 4.2 The Jukes-Cantor correction to distance model that produces a linear correlation between time and genetic distance.[24] Variance increases exponentially with time making the model less reliable as the time increases. **μ** = mutation rate, **n** = sequence length[23]

The K2P distance correction, developed by Kimura in 1980,[18] corrected this drawback and introduced a model with two parameters, one for the transition rate (when a purine is substituted with a purine, or a pyrimidine with a pyrimidine) and one for the transversion rate (a substitution of a purine with a pyrimidine, or vice versa). This was because it had become clear that all nucleotide substitutions do not occur at equal rates. Usually, transitions occur more frequently than transversions.

The K2P correction indicates that when the rates of transitions and transversions are assumed to be α and β and the observed rates are P and Q, then:

$$D_{K2P} = -(1/2)\log n(1-2P-Q) - (1/4)\log n(1-2Q)$$

$$Var(D_{K2P})=[c^2_1P+c^2_3Q-(C_1p+c_3Q)^2]/n$$

where D_{K2P} = Kimura 2-parameter corrected distance. **n** = sequence length, ***P*** = observed number of transitional differences, ***Q*** = observed number of transversional differences, $c_1 = 1/(1-2P-Q)$, $c_2 = 1/(1-2Q)$ and $c_3 = 1/2(c_1=c_2)$[25]

This model retains the linearity of the Jukes-Cantor correction but reduces the variance with time, making this correction more reliable over longer time periods.[18] It has been adopted by BOLD and forms the basis of their distance-based system for distinguishing species.

There has been some discussion over whether a distance-based system for distinguishing species is a retrograde step.[20, 26-28] Critics say that it is a return to phenetics, a system of classification of organisms based on overall or observable similarities rather than on phylogenetic or evolutionary relationships.[20, 28] They suggest that a character-based system, in which the presence or absence of certain nucleotides is used as a diagnostic tool for distinguishing species, is preferable.[20, 26-29] BOLD uses K2P distance measures for most species identifications but if this measure is unable to distinguish species, for example, hybridising parapatric species, a character-based approach is employed instead.[30]

The K2P trees generated by the BOLD software employ varied distance metrics to produce a neighbour-joining tree based on nucleotide or amino acid sequences.[31] K2P is the default metric and, although pairwise-distance and the Jukes-Cantor correction are also occasionally used, they were not used for the trees produced in this study.

The main virtue of the use of neighbour-joining trees relative to other methods is their computational efficiency. They can be used on very large data sets for which other means of phylogenetic analysis (minimum evolution, maximum parsimony, maximum likelihood) are computationally prohibitive.[32] Thus, they are suitable for this online application where rapid computation is essential. Distance in these trees is given as a percentage along with a scale to measure that distance on the tree. Further analysis using PAUP (available from http://paup.csit.fsu.edu/) was required to produce the non-parametric bootstrap values associated with the branches of these trees.

4.2.2 Bayesian inference of phylogeny

Sequence data were further analysed using Bayesian inference of phylogeny[19, 33] to provide a comparison with the K2P neighbour joining trees. Bayesian inference of phylogeny is based upon a measure termed the posterior probability distribution

of trees, which is the probability of a tree conditional upon observations. The Bayesian approach to phylogenetic inference provides a method for simultaneously estimating trees and obtaining measurements of uncertainty for every branch. It focuses on the posterior probability of hypotheses and this posterior probability is proportional to the product of the prior probability and the likelihood of (or evidence for) that hypothesis. This is accomplished using Bayes' theorem. However, the posterior probability distribution of trees is impossible to calculate analytically; instead, the trees are calculated using programs for Bayesian estimation of phylogeny such as BAMBE (available from http://www.mathcs.duq.edu/larget/bambe.html), MrBayes (available from http://mrbayes.csit.fsu.edu/download.php) or Beast (available from http://beast.bio.ed.ac.uk/Main_Page). These utilise a simulation algorithm called the Markov chain Monte Carlo (MCMC) to approximate the posterior probabilities of trees. For the purposes of this study, MrBayes is the software that has been used (see MrBayes 3.1 Manual[34] for details of operation).

The MCMC algorithm represents a series of steps that form a conceptual chain. At each step, a new location in parameter space is proposed as the next link in the chain. This proposed location is usually similar to the present one because it is generated by random changes to a few of the parameters in the present link of the chain. The relative posterior-probability at the new location is calculated. If the new location has a higher posterior-probability than that of the present location of the chain, the move is accepted and the proposed location becomes the next link in the chain. The cycle is then repeated. Alternatively, if the proposed location has a lower posterior-probability, the move will be accepted only some of the time. Thus, small steps downward are sometimes accepted, whereas large steps downward are prevented. If the proposed location is rejected, the present location is added as the next link in the chain: the last two links in the chain will be identical and the cycle is repeated (Figure 4.3).[19]

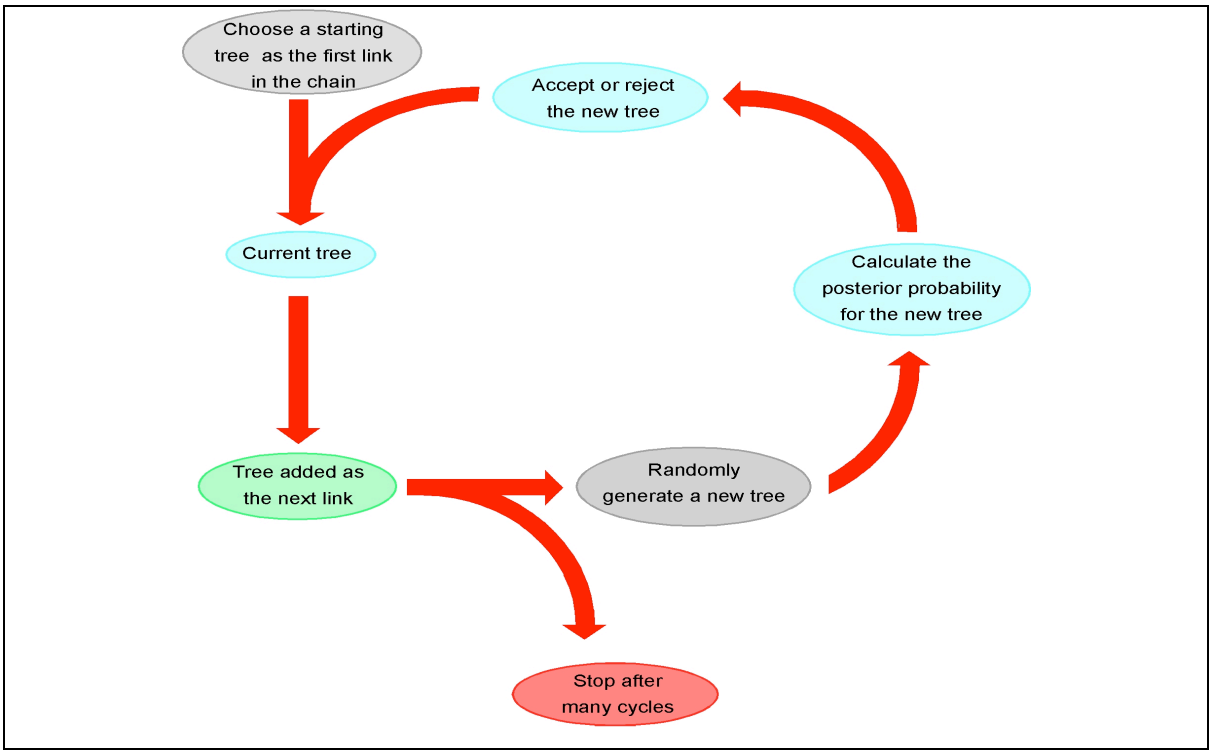


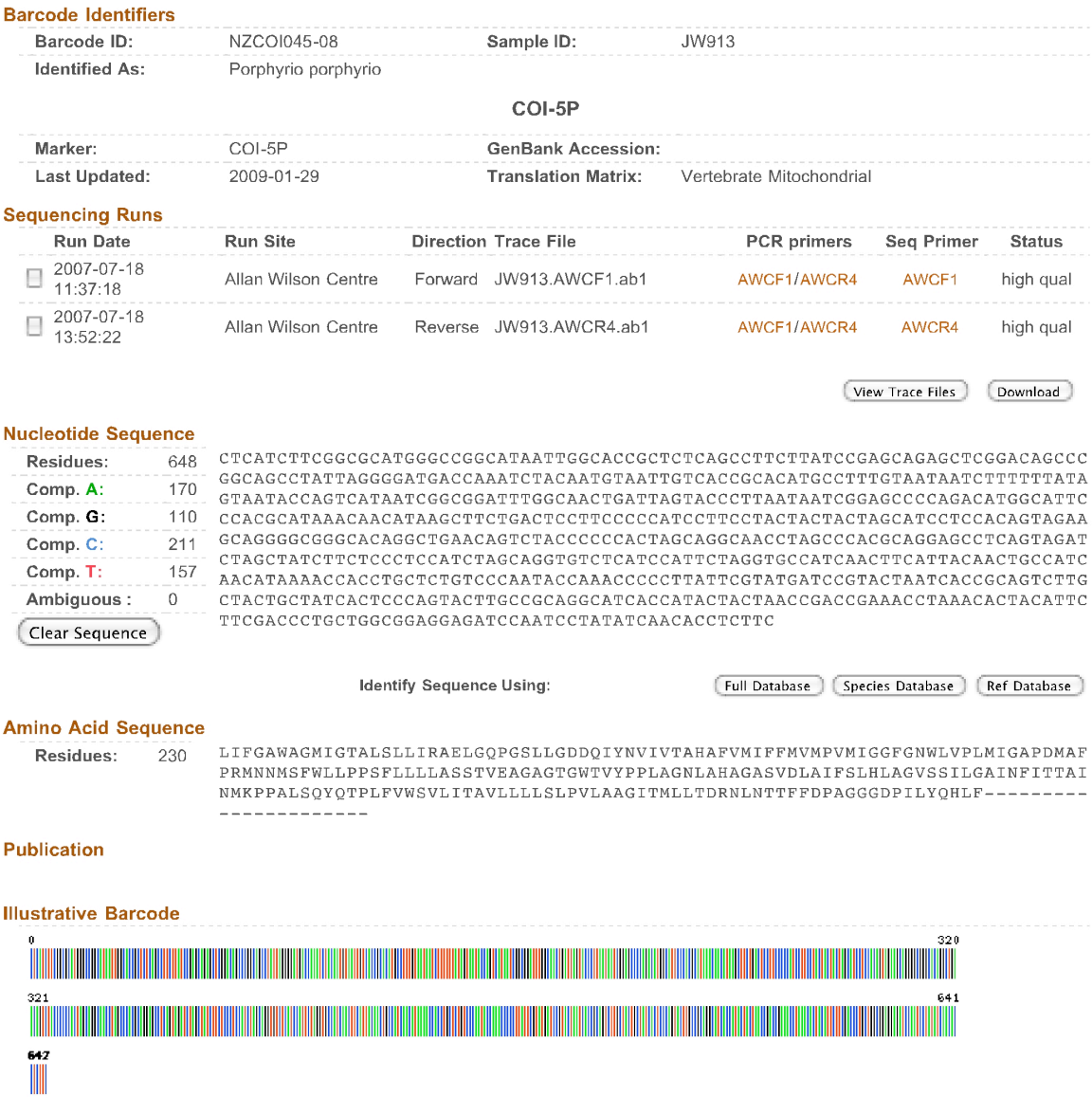
Figure 4.3 Flow chart of the series of steps in the Markov chain Monte Carlo (MCMC) algorithm showing the cycles of random tree generation, calculation of posterior probability, followed by acceptance or rejection. These cycles occur many thousands of times before acceptance of a final tree.[19, 33]

Nucleotide sequences analysed using MrBayes were aligned using ClustalX available from http://www.clustal.org/ and the alignments adjusted by eye. Appropriate models of evolution for each dataset were assessed using MrModeltest available from http://www.abc.se/~nylander/. Phylogenetic trees were constructed using MrBayes. Input to MrBayes was as a character matrix in Nexus file format (Table 4.1). Burnin time was determined by assessment of the number of generations required for the standard deviation of split to reach a value of <0.01. Generations preceding the convergence statistic were discarded. One million cycles were run before a phylogenetic tree with the highest possible posterior probability was produced and, by contrast to K2P calculations of distance, this took considerable computing power and time.[34] Bayesian consensus trees were viewed in Treeview available from http://www.treeview.net/, then edited and converted to jpeg format using Canvas™X scientific imaging software available from ACD Systems of America, Inc..

Table 4.1 An example of Nexus file format, a simple text (ASCII) file beginning with #NEXUS on the first line. The rest of the file is divided into different blocks identifying and characterising different parameters. The example below is for the New Zealand **anseriformes.nex** file.

```
#NEXUS
begin data;
  dimensions ntax=12 nchar=648;
  format datatype=dna interleave=no gap=-;
  matrix
  Anas_castanea               CTTATCTTCGGGGCA ... TACCAACACCTATTC
  A_gracilis                  CTTATCTTCGGGGCA ... TACCAACACCTATTC
  A_nesiotis                  CTTATCTTCGGGGCA ... TACCAACACCTGTTT
  A_platyrhynchos             CTTATCTTCGGGGCA ... TACCAACACCTATTT
  A_platyrhynchos             CTTATCTTCGGGGCA ... TACCAACACCTATTT
  A_platyrhynchos             CTTATCTTCGGGGCA ... TACCAACACCTATTT
  A_rhynchotis                CTTATCTTCGGGGCA ... TACCAACACCTATTT
  A_superciliosa              CTTATCTTCGGGGCA ... CTTATCTTCGGGGCA
  Aythya novaeseelandiae      CTTATCTTCGGGGCA ... TACCAACACCTATTC
  Hymenolaimus malacorhynchos CTTATTTTTGGCGCA ... TACCAACACCTATTC
  Hymenolaimus malacorhynchos CTTATTTTTGGCGCA ... TACCAACACCTATTC
  Hymenolaimus malacorhynchos CTTATTTTTGGCGCA ... TACCAACACCTATTC
 ;
end;
```

In the Bayesian trees produced by this method, the horizontal distance is a measure of genetic distance in nucleotide substitutions read according to the scale provided. In addition, the Bayesian trees have a posterior probability for each branch[34] and a BOLD sequence page number for each specimen.[1] The sequence page number permits access to all the available specimen details from the project website, including sample and identification numbers, taxonomy, collection and identification information, nucleotide and deduced amino acid sequence, trace files, primers used and the quality of sequence (Figures 4.4 and 4.5).



Barcode Identifiers

Barcode ID:	NZCOI045-08	**Sample ID:**	JW913
Identified As:	Porphyrio porphyrio		

COI-5P

Marker:	COI-5P	**GenBank Accession:**	
Last Updated:	2009-01-29	**Translation Matrix:**	Vertebrate Mitochondrial

Sequencing Runs

Run Date	Run Site	Direction	Trace File	PCR primers	Seq Primer	Status
2007-07-18 11:37:18	Allan Wilson Centre	Forward	JW913.AWCF1.ab1	AWCF1/AWCR4	AWCF1	high qual
2007-07-18 13:52:22	Allan Wilson Centre	Reverse	JW913.AWCR4.ab1	AWCF1/AWCR4	AWCR4	high qual

View Trace Files | Download

Nucleotide Sequence

Residues:	648
Comp. A:	170
Comp. G:	110
Comp. C:	211
Comp. T:	157
Ambiguous :	0

Clear Sequence

CTCATCTTCGGCGCATGGGCCGGCATAATTGGCACCGCTCTCAGCCTTCTTATCCGAGCAGAGCTCGGACAGCCC
GGCAGCCTATTAGGGGATGACCAAATCTACAATGTAATTGTCACCGCACATGCCTTTGTAATAATCTTTTTTATA
GTAATACCAGTCATAATCGGCGGATTTGGCAACTGATTAGTACCCTTAATAATCGGAGCCCCAGACATGGCATTC
CCACGCATAAACAACATAAGCTTCTGACTCCTTCCCCCATCCTTCCTACTACTACTAGCATCCTCCACAGTAGAA
GCAGGGGCGGGCACAGGCTGAACAGTCTACCCCCCACTAGCAGGCAACCTAGCCCACGCAGGAGCCTCAGTAGAT
CTAGCTATCTTCTCCCTCCATCTAGCAGGTGTCTCATCCATTCTAGGTGCCATCAACTTCATTACAACTGCCATC
AACATAAAACCACCTGCTCTGTCCCAATACCAAACCCCCTTATTCGTATGATCCGTACTAATCACCGCAGTCTTG
CTACTGCTATCACTCCCAGTACTTGCCGCAGGCATCACCATACTACTAACCGACCGAAACCTAAACACTACATTC
TTCGACCCTGCTGGCGGAGGAGATCCAATCCTATATCAACACCTCTTC

Identify Sequence Using: Full Database | Species Database | Ref Database

Amino Acid Sequence

Residues:	230

LIFGAWAGMIGTALSLLIRAELGQPGSLLGDDQIYNVIVTAHAFVMIFFMVMPVMIGGFGNWLVPLMIGAPDMAF
PRMNNMSFWLLPPSFLLLLASSTVEAGAGTGWTVYPPLAGNLAHAGASVDLAIFSLHLAGVSSILGAINFITTAI
NMKPPALSQYQTPLFVWSVLITAVLLLLSLPVLAAGITMLLTDRNLNTTFFDPAGGGDPILYQHLF---------
----

Publication

Illustrative Barcode

Figure 4.4 A sequence record for a *Puffinus griseus* specimen.

Trees constructed using both K2P and Bayesian methods (Figures 4.6 - 4.27) were produced for all New Zealand avian orders, with five exceptions. Columbiformes, Coraciiformes and Strigiformes, which each contain just one New Zealand species and, therefore, no tree was possible. In addition, the Cuculiformes contained only four specimens, two from each of two species that showed no unusual divergence or convergence so no tree was presented for this order. Likewise, the Podicipediformes had only three specimens and each of these was from a different family so no tree was produced for this group either.

Specimen Identifiers

Sample ID :	JW913	**Museum ID :**	TV-1
Isolate / Field Num:		**Collection Code :**	
Deposited In :	Allan Wilson Centre		

Taxonomy

Identifier :	Ewan J. Grant-Mackie
phylum :	Chordata
class :	Aves
order :	Gruiformes
family :	Rallidae
genus :	Porphyrio
species :	Porphyrio porphyrio
Taxonomy Note :	

Specimen Details

Voucher Type :	
Tissue Type :	
Extra Info :	
Sex :	
Reproduction :	
Life Stage :	A
Note :	

Collection Data

Collectors :	Ewan Grant-Mackie
Date Collected :	19-Jun-2007
Country :	New Zealand
State/Province:	North Island
Region/County :	Coromandel
Sector :	Thames Valley
Exact Site :	
Latitude :	-37.1528
Longitude :	175.595
Coord. Source :	
Coord.Accuracy :	
Elevation :	18 m
Elevation Accuracy:	
Depth:	
Depth Accuracy:	

Photographs

No images available

Figure 4.5 A specimen data record for a *Puffinus griseus* specimen.

Because the Barcoding New Zealand Birds project is ongoing, trees generated using K2P and Bayesian techniques were sometimes from slightly different datasets. This is because, in general, Bayesian analysis was carried out before the K2P analysis and in the interim, the dataset may have grown.

Where species that were difficult to differentiate using COI haplotypes and which both K2P and Bayesian techniques could not clearly assign, a separate K2P tree was produced to help clarify their status, using all the available sequences from this project, plus any others available from the BOLD website and from GenBank. These trees with their associated distance measures and nearest neighbour

analyses are presented in Chapter 5 along with a discussion of the species involved. In addition, nucleotides within the barcode region of the COI gene, that are unique to each of these species, have been sought, in order to ascertain the value of a character-based system in resolving them.

Where species have separate clusters of barcodes, indicative of possible subspecies or cryptic species, similar intensive examination has been carried out (see Chapter 5).

4.3 Results

To date, a total of 628 samples from 126 (84%) of the 150 native and endemic species of New Zealand from 16 avian orders have had their DNA barcode sequence recorded for this project. Details of the number of samples per species are given in Table 4.2. In addition, a further 205 sequences from 89 species that fall outside the categories of native or endemic have also been recorded (see Appendix E for a complete list of species from this project).

Table 4.2 List of native (n=79) and endemic (n=71) New Zealand bird species showing the number of tissue samples (total n=628) per species sequenced for the Barcoding New Zealand Birds project; blanks indicate species yet to be sequenced

Order / Common Name	Genus species	Status	No. of individuals with DNA barcodes
APTERYGIFORMES			
Southern Brown Kiwi	*Apteryx australis*	Endemic	13
Great Spotted Kiwi	*Apteryx haastii*	Endemic	5
North Island Brown Kiwi	*Apteryx mantelli*	Endemic	1
Little Spotted Kiwi	*Apteryx owenii*	Endemic	7
Okarito Brown Kiwi	*Apteryx rowi*	Endemic	1
PODICIPEDIFORMES			
Great Crested Grebe	*Podiceps cristatus*	Native	2
New Zealand Dabchick	*Poliocephalus rufopectus*	Endemic	1
Australasian Grebe	*Tachybaptus novaehollandiae*	Native	2
PROCELLARIIFORMES			
Cape Petrel	*Daption capense*	Native	5
Royal Albatross	*Diomedea epomophora*	Endemic	15
Wandering Albatross	*Diomedea exulans*	Native	2
White-bellied Storm-Petrel	*Fregetta grallaria*	Native	
Black-bellied Storm-Petrel	*Fregetta tropica*	Native	1
Southern Fulmar	*Fulmarus glacialoides*	Native	

Order / Common Name	Genus species	Status	No. of individuals with DNA barcodes
Gray-backed Storm-Petrel	*Garrodia nereis*	Native	5
Hall's Giant Petrel	*Macronectes halli*	Native	
Wilson's Storm-Petrel	*Oceanites oceanicus*	Native	1
Fulmar Prion	*Pachyptila crassirostris*	Native	1
Antarctic Prion	*Pachyptila desolata*	Native	1
Fairy Prion	*Pachyptila turtur*	Native	1
Broad-billed Prion	*Pachyptila vittata*	Native	3
White-faced Storm-Petrel	*Pelagodroma marina*	Native	4
South Georgia Diving-Petrel	*Pelecanoides georgicus*	Native	3
Common Diving-Petrel	*Pelecanoides urinatrix*	Native	6
Light-mantled Albatross	*Phoebetria palpebrata*	Native	2
White-chinned Petrel	*Procellaria aequinoctialis*	Native	4
Gray Petrel	*Procellaria cinerea*	Native	1
Parkinson's Petrel	*Procellaria parkinsoni*	Endemic	1
Westland Petrel	*Procellaria westlandica*	Endemic	8
Chatham Petrel	*Pterodroma axillaris*	Endemic	5
White-naped Petrel	*Pterodroma cervicalis*	Endemic	1
Cook's Petrel	*Pterodroma cookii*	Endemic	10
Mottled Petrel	*Pterodroma inexpectata*	Endemic	7
White-headed Petrel	*Pterodroma lessonii*	Native	6
Great-winged Petrel	*Pterodroma macroptera*	Native	5
Magenta Petrel	*Pterodroma magentae*	Endemic	5
Soft-plumaged Petrel	*Pterodroma mollis*	Native	
Kermadec Petrel	*Pterodroma neglecta*	Native	6
Black-winged Petrel	*Pterodroma nigripennis*	Native	10
Pycroft's Petrel	*Pterodroma pycrofti*	Endemic	
Little Shearwater	*Puffinus assimilis*	Native	5
Buller's Shearwater	*Puffinus bulleri*	Endemic	1
Flesh-footed Shearwater	*Puffinus carneipes*	Native	11
Fluttering Shearwater	*Puffinus gavia*	Endemic	5
Sooty Shearwater	*Puffinus griseus*	Native	16
Hutton's Shearwater	*Puffinus huttoni*	Endemic	1
Wedge-tailed Shearwater	*Puffinus pacificus*	Native	6
Buller's Albatross	*Thalassarche bulleri*	Endemic	
Shy Albatross	*Thalassarche cauta*	Native	1
Gray-headed Albatross	*Thalassarche chrysostoma*	Native	6
Black-browed Albatross	*Thalassarche melanophris*	Native	10
SPHENISCIFORMES			
Rockhopper Penguin	*Eudyptes chrysocome*	Native	9
Fiordland Penguin	*Eudyptes pachyrhynchus*	Endemic	11
Snares Penguin	*Eudyptes robustus*	Endemic	1
Erect-crested Penguin	*Eudyptes sclateri*	Endemic	3
Little Penguin	*Eudyptula minor*	Native	3
Yellow-eyed Penguin	*Megadyptes antipodes*	Endemic	16
PELECANIFORMES			

Order / Common Name	Genus species	Status	No. of individuals with DNA barcodes
Australasian Gannet	*Morus serator*	Native	15
Red-tailed Tropic Bird	*Phaethon rubricauda*	Native	1
Campbell Islands Shag	*Phalacrocorax campbelli*	Endemic	
Great Cormorant	*Phalacrocorax carbo*	Native	1
Rough-faced Shag	*Phalacrocorax carunculatus*	Endemic	
Bronze Shag	*Phalacrocorax chalconotus*	Endemic	
Auckland Islands Shag	*Phalacrocorax colensoi*	Endemic	
Pitt Island Shag	*Phalacrocorax featherstoni*	Endemic	
Little Pied Cormorant	*Phalacrocorax melanoleucos*	Native	2
Chatham Islands Shag	*Phalacrocorax onslowi*	Endemic	
Spotted Shag	*Phalacrocorax punctatus*	Endemic	7
Bounty Islands Shag	*Phalacrocorax ranfurlyi*	Endemic	
Little Black Cormorant	*Phalacrocorax sulcirostris*	Native	1
Pied Cormorant	*Phalacrocorax varius*	Native	1
Masked Booby	*Sula dactylatra*	Native	5
CICONIIFORMES			
Australasian Bittern	*Botaurus poiciloptilus*	Native	3
Great Egret	*Egretta alba*	Native	1
White-faced Heron	*Egretta novaehollandiae*	Native	1
Rufous Night-Heron	*Nycticorax caledonicus*	Native	1
Royal Spoonbill	*Platalea regia*	Native	1
ANSERIFORMES			
Brown Teal	*Anas aucklandica*	Endemic	1
Gray Teal	*Anas gracilis*	Native	1
Australian Shoveler	*Anas rhynchotis*	Native	1
Pacific Black Duck	*Anas superciliosa*	Native	1
New Zealand Scaup	*Aythya novaeseelandiae*	Endemic	2
Blue Duck	*Hymenolaimus malacorhynchos*	Endemic	11
Paradise Shelduck	*Tadorna variegata*	Endemic	1
Chestnut-chested Shelduck	*Tadorna tadornoides*	Native	1
FALCONIFORMES			
Australasian Harrier	*Circus approximans*	Native	3
New Zealand Falcon	*Falco novaeseelandiae*	Endemic	10
GRUIFORMES			
Australian Coot	*Fulica atra*	Native	1
Weka	*Gallirallus australis*	Endemic	6
Takahe	*Porphyrio mantelli*	Endemic	13
Purple Swamphen	*Porphyrio porphyrio*	Native	13
Marsh Crake	*Porzana pusilla*	Native	
Spotless Crake	*Porzana tabuensis*	Native	1
Auckland Islands Rail	*Rallus pectoralis*	Native	1
Banded Rail	*Rallus philippensis*	Native	3
CHARADRIIFORMES			
Wrybill	*Anarhynchus frontalis*	Endemic	2

Order / Common Name	Genus species	Status	No. of individuals with DNA barcodes
Common Noddy	*Anous stolidus*	Native	
White Capped Noddy	*Anous tenuirostris*	Native	
Banded Dotterel	*Charadrius bicinctus*	Endemic	4
Black-fronted Dotterel	*Charadrius melanops*	Native	
Red-breasted Dotterel	*Charadrius obscurus*	Endemic	
Black-fronted Tern	*Chlidonius albostriatus*	Endemic	1
Subantarctic Snipe	*Coenocorypha aucklandica*	Endemic	
Chatham Islands Snipe	*Coenocorypha pusilla*	Endemic	
White Tern	*Gygis alba*	Native	
Chatham Oystercatcher	*Haematopus chathamensis*	Endemic	1
Pied Oystercatcher	*Haematopus ostralegus*	Native	2
Variable Oystercatcher	*Haematopus unicolor*	Endemic	3
Pied Stilt	*Himantopus himantopus*	Native	1
Black Stilt	*Himantopus novaezelandiae*	Endemic	5
Caspian Tern	*Hydroprogne (Sterna) caspia*	Native	1
Black-billed Gull	*Larus bulleri*	Endemic	2
Black-backed Gull	*Larus dominicanus*	Native	6
Red-billed Gull	*Larus novaehollandiae*	Native	10
Grey Ternlet (Noddy)	*Procelsterna cerulea*	Native	
Sooty Tern	*Sterna fuscata*	Native	1
Fairy Tern	*Sterna nereis*	Native	5
White-fronted Tern	*Sterna striata*	Native	3
Antarctic Tern	*Sterna vittata*	Native	
Shore Plover	*Thinornis novaeseelandiae*	Endemic	8
COLUMBIFORMES			
New Zealand Pigeon	*Hemiphaga novaeseelandiae*	Endemic	8
PSITTACIFORMES			
Yellow-crowned Parakeet	*Cyanoramphus auriceps*	Endemic	
Red-crowned Parakeet	*Cyanoramphus novaezelandiae*	Native	7
Antipodes Parakeet	*Cyanoramphus unicolor*	Endemic	3
New Zealand Kaka	*Nestor meridionalis*	Endemic	9
Kea	*Nestor notabilis*	Endemic	7
Kakapo	*Strigops habroptilus*	Endemic	9
CUCULIFORMES			
Shining Cuckoo	*Chrysococcyx lucidus*	Native	5
Long-tailed Cuckoo	*Eudynamys taitensis*	Endemic	2
STRIGIFORMES			
Morepork	*Ninox novaeseelandiae*	Native	5
CORACIIFORMES			
Sacred Kingfisher	*Todiramphus sanctus*	Native	8
PASSERIFORMES			
Rifleman	*Acanthisitta chloris*	Endemic	9
New Zealand Bellbird	*Anthornis melanura*	Endemic	12

Order / Common Name	Genus species	Status	No. of individuals with DNA barcodes
Australasian Pipit	*Anthus novaeseelandiae*	Native	6
Fernbird	*Bowdleria punctata*	Endemic	1
Kokako	*Callaeas cinerea*	Endemic	1
Grey Warbler	*Gerygone igata*	Endemic	11
Welcome Swallow	*Hirundo tahitica*	Native	2
Whitehead	*Mohoua albicilla*	Endemic	5
Brown Creeper	*Mohoua novaeseelandiae*	Endemic	8
Yellowhead	*Mohoua ochrocephala*	Endemic	8
Stitchbird	*Notiomystis cincta*	Endemic	7
New Zealand Robin	*Petroica australis*	Endemic	18
Tomtit	*Petroica macrocephala*	Endemic	8
Chatham Robin	*Petroica traversi*	Endemic	13
Saddleback	*Philesturnus carunculatus*	Endemic	17
Tui	*Prosthemadera novaeseelandiae*	Endemic	12
Grey Fantail	*Rhipidura fuliginosa*	Native	3
(South Island) Rock Wren	*Xenicus gilviventris*	Endemic	
Silver-eye	*Zosterops lateralis*	Native	6

Each of the single New Zealand species from the orders Columbiforme (*Hemiphaga novaeseelandiae*: n=8), Coraciiforme (*Todiramphus sanctus*: n=9) and Strigiforme (*Ninox novaeseelandiae*: n=5) showed no unusual COI sequence clusters and normal intraspecific nucleotide divergences at this region of the mitochondrial genome (Table 4.3). The same was true for the four Cuculiforme and three Podicipediforme specimens. In general, the tree topologies obtained from NJ K2P and Bayesian inference were very similar in all the orders examined except for a few weakly supported nodes. Complete details of individual specimens can be obtained from the BOLD website using the sequence identity number provided with each specimen on the terminal branches of the trees.

Table 4.3 Intraspecific K2P percent distance variation within the single New Zealand representatives of the orders Columbiforme, Coraciiforme and Strigiforme

Species	n=	Mean	Max	SE
Hemiphaga novaeseelandiae	8	0.189	0.625	0.036
Todiramphus sanctus	8	0	0	0
Ninox novaeseelandiae	5	0.313	0.781	0.121

4.3.1 Anseriformes

The New Zealand Anseriformes (ducks, geese and swans) is a medium-sized order (Table 4.2) that shows clear distinction between species using both K2P and Bayesian techniques (Figures 4.6 and 4.7). The trees generated for this order (which include seven of the eight native and endemic species plus four other species) by these two methods show similar topology; all species were reciprocally monophyletic in both trees with node support of 0.97-1.0 for the three species with multiple individuals in the Bayesian analysis. Interspecific distances in NJ trees ranged from 0.311 to 9.261%, while maximum intraspecific distance was 0.156%. All species show distinct COI sequence differences from each other and, according to both techniques, the largest group, *Hymenolaimus malacorhynchos*, shows slight intraspecific variation for five of the eleven specimens. These five specimens are all from the central North Island of New Zealand.

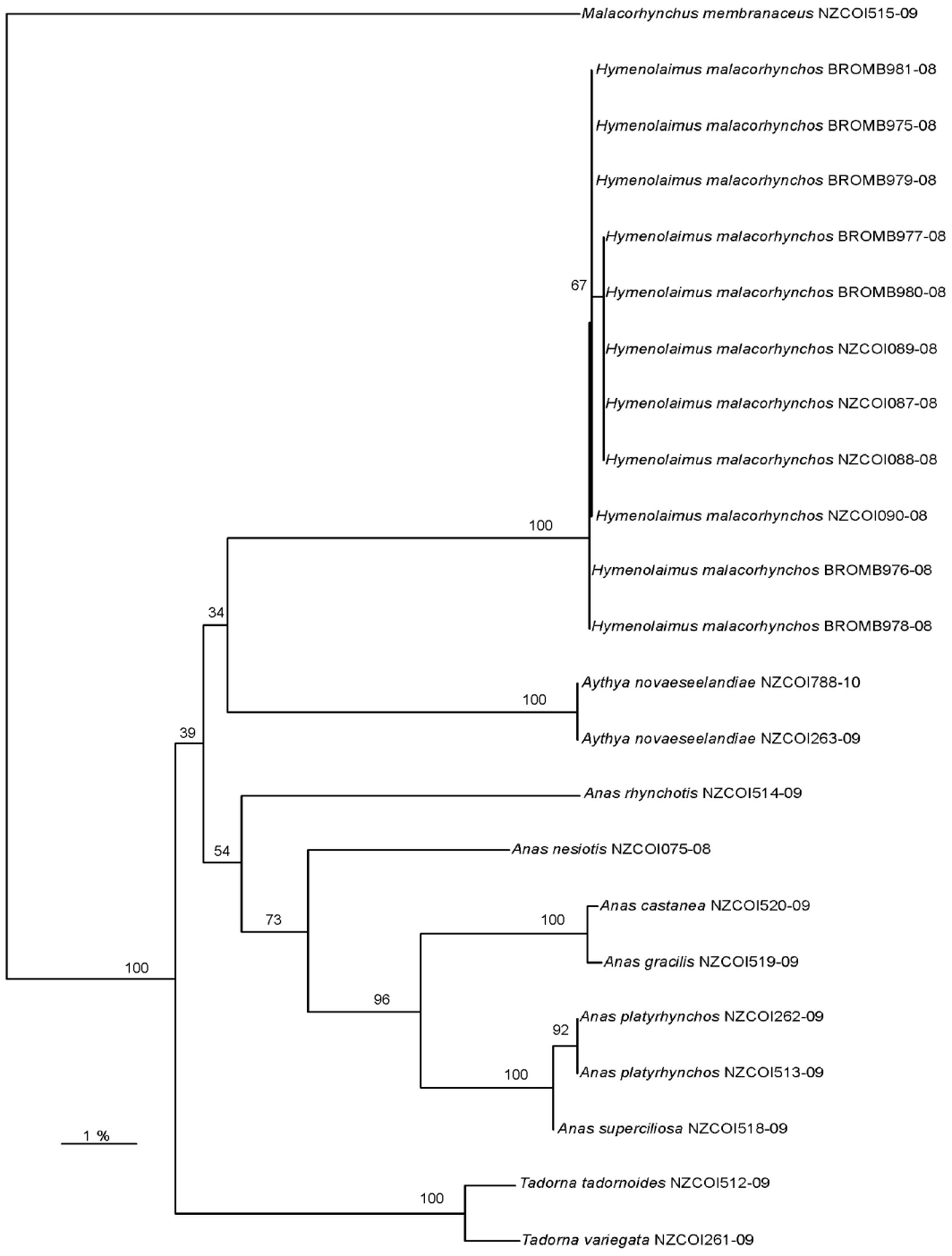


Figure 4.6 New Zealand Anseriformes K2P neighbour joining tree with bootstrap values for all major branches.

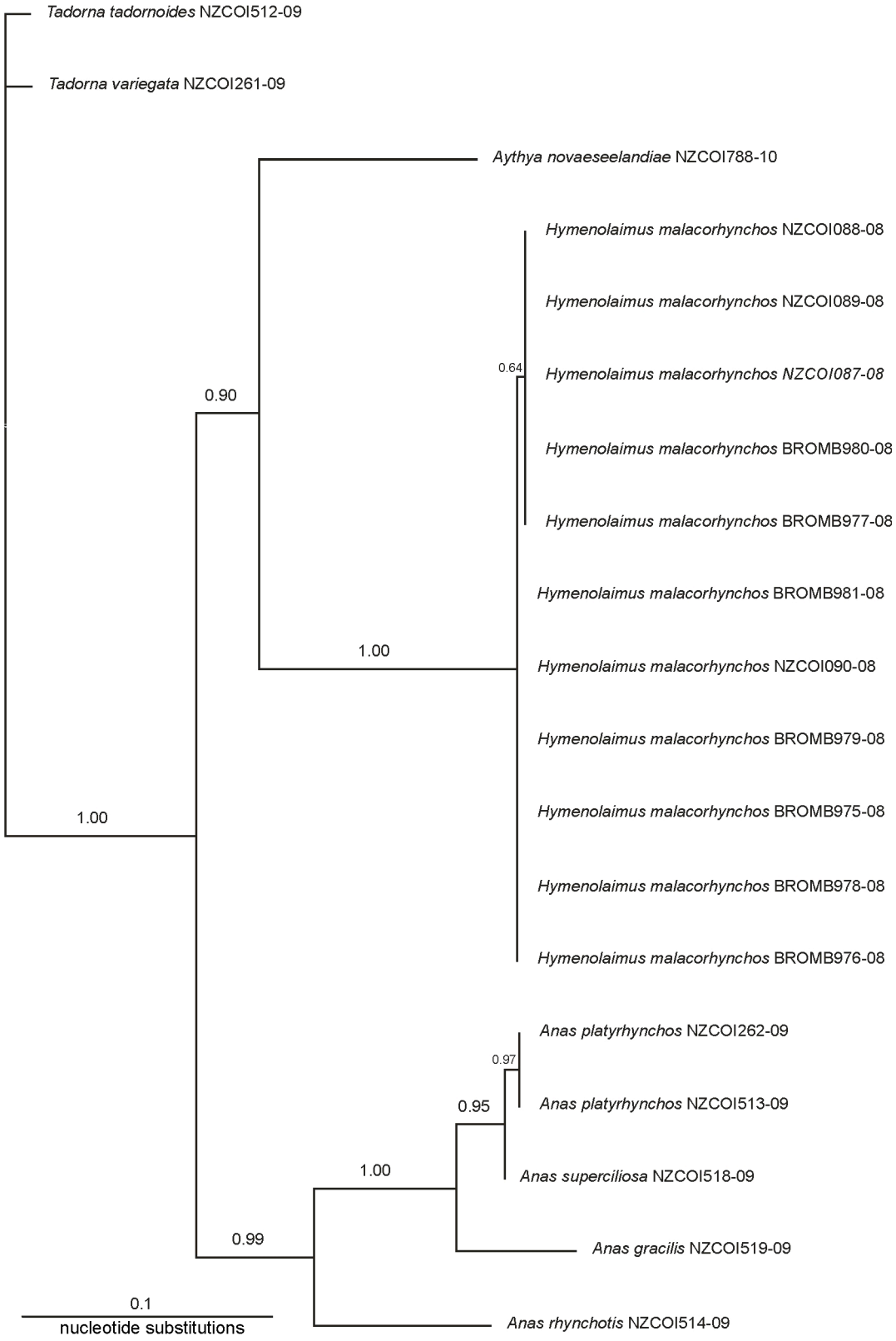


Figure 4.7 New Zealand Anseriformes Bayesian tree with posterior probability for each branch.

4.3.2 Apterygiformes

When subjected to K2P or Bayesian analysis, the Apterygiformes (Kiwis) formed complicated trees in which the acknowledged species generally formed distinct clades. Until recently there were only three recognised species, *Apteryx australis*, *A. owenii* and *A. haastii*.[35] *A. mantelli* was treated as a subspecies of *A. australis* and *A. rowi* has only recently been recognised.[36] DNA barcode data trees identified distinct clades for all these groups; however, *A. australis* is clearly a species with quite high COI variability and has close evolutionary links with both *A.mantelli* and *A. rowi*. Because of the complexity of this group, DNA sequences from all relevant specimens from the BOLD website and GenBank have been used and an enlarged K2P tree was generated (Figure 4.8). The Bayesian tree was, however, produced using only specimens from this project (Figure 4.9).

The maximum intraspecific K2P distance (6.062%) overlapped the mean interspecific distance (4.601%) and nearest neighbour analysis showed 0% distance to nearest neighbour for *A. australis*, *A. mantelli* and *A. owenii* (see Chapter 5). The removal of three apparently anomalous specimens (NZCOI491, NZCOI492 and NZCOI790) of the 47 examined resulted in less intraspecific variation and greater nearest neighbour divergence within these species.

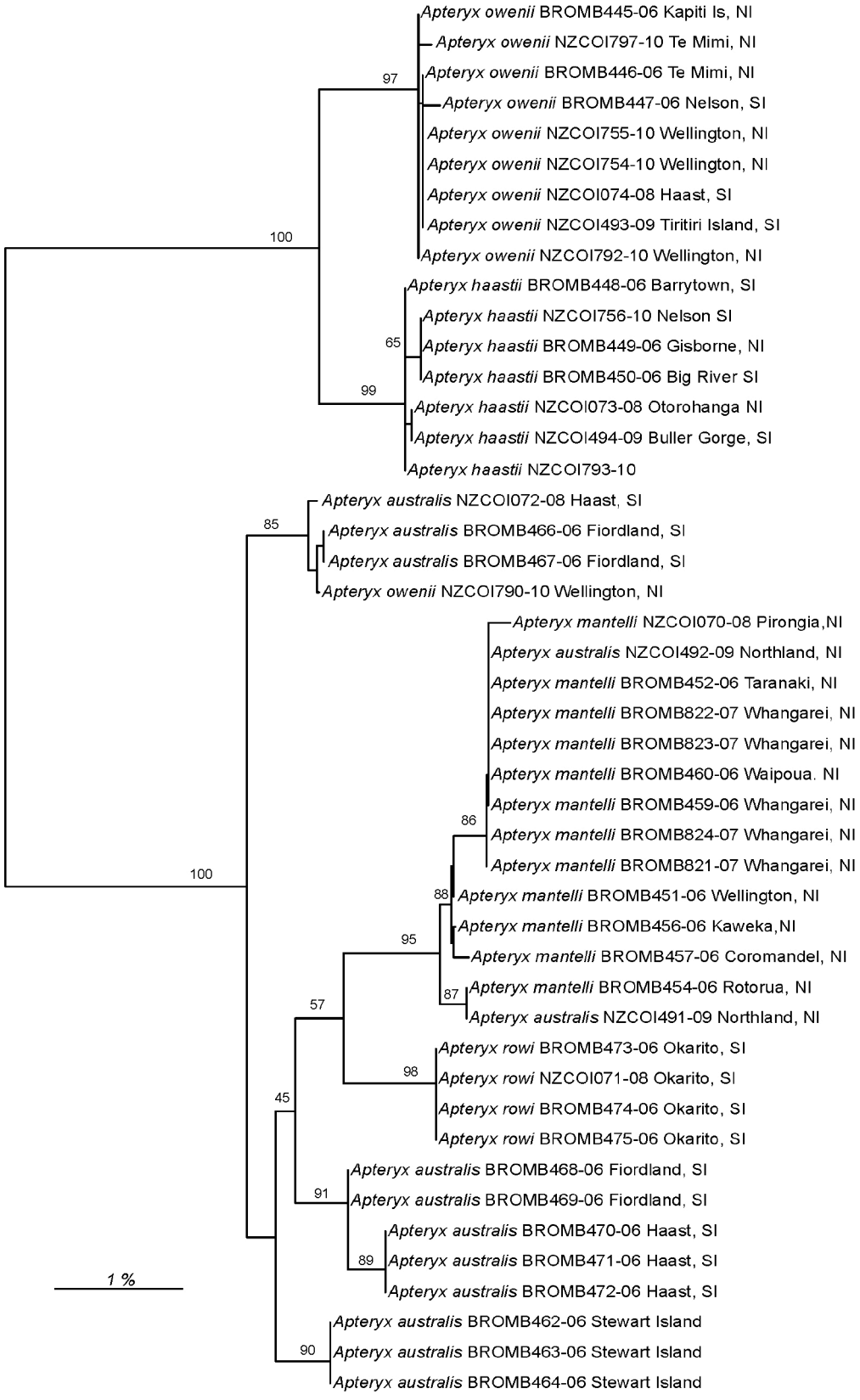


Figure 4.8 New Zealand Apterygiformes K2P neighbour joining tree with bootstrap values for all major branches.

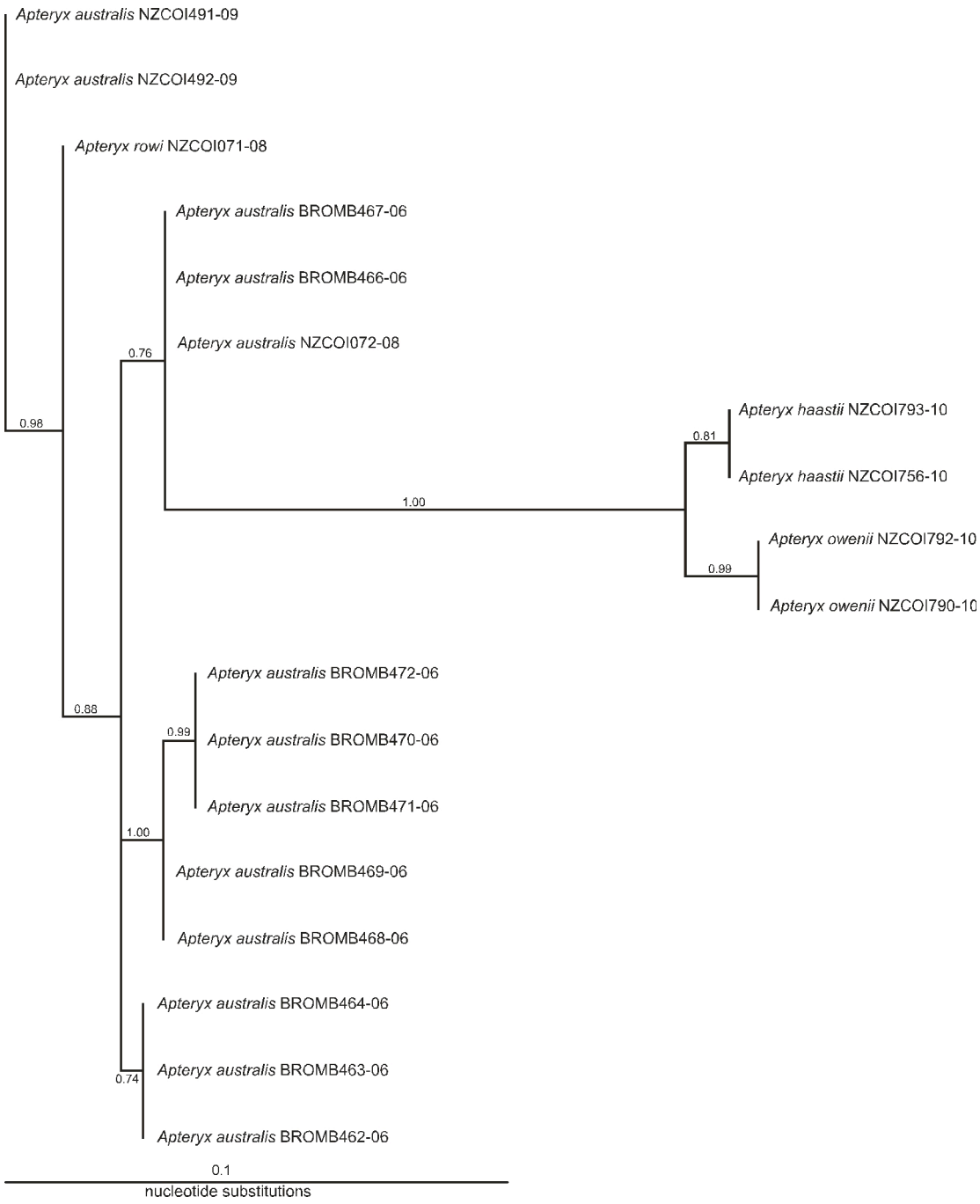


Figure 4.9 New Zealand Apterygiformes Bayesian tree with posterior probabilities given for each branch and BOLD sequence page number for each specimen.

4.3.3 Charadriiformes

The Charadriiformes are the second largest order of native and endemic birds in New Zealand (Table 4.2). Within this order two genera (*Haematopus and Larus*) were not fully distinguished using DNA barcoding, with ambiguities showing in both K2P and Bayesian trees (Figures 4.10 and 4.11).

A larger K2P tree of 112 *Haematopus unicolor*, ten *H. finschi* and three *H. chathamensis* (Figure 5.18) which included entries from the Barcoding New Zealand Birds project as well as other BOLD (records unpublished) and all GenBank entries, also showed little divergence in COI region nucleotides. *H. chathamensis* was a distinct clade within this larger tree, but it only differed from the mainland species by less than 0.1%. Distance summary and nearest neighbour analyses also failed to differentiate these species (Chapter 5.3.2), with distance to nearest neighbour for *H. chathamensis* being less than the mean intraspecific variation for this species (0.14% *vs* 0.18%). The other two species (*H. finschi* and *H. unicolor*) had a nearest neighbour distance of 0%. Nucleotides differed within all 125 specimens at just six sites. Only one of these sites was diagnostic for *H. chathamensis* within this group. The other five appeared to be randomly distributed among the three species.

Members of the genus *Larus* also showed little genetic distance in both K2P and Bayesian trees (Figures 4.10 and 4.11). In general the Black-backed Gull, *L. dominicanus*, showed a relatively distinct clade, with just one anomalous specimen but *L. bulleri* and *L. novaehollandiae* were indistinguishable using this method (Chapter 5.3.4). Mean interspecific distance fell well within the intraspecific maxima, and nearest neighbour analysis found 0% distance between the two species. There were differences at three nucleotide sites for the ten *L. novaehollandia* and seven *L. bulleri* specimens. None of these was diagnostic for either species.

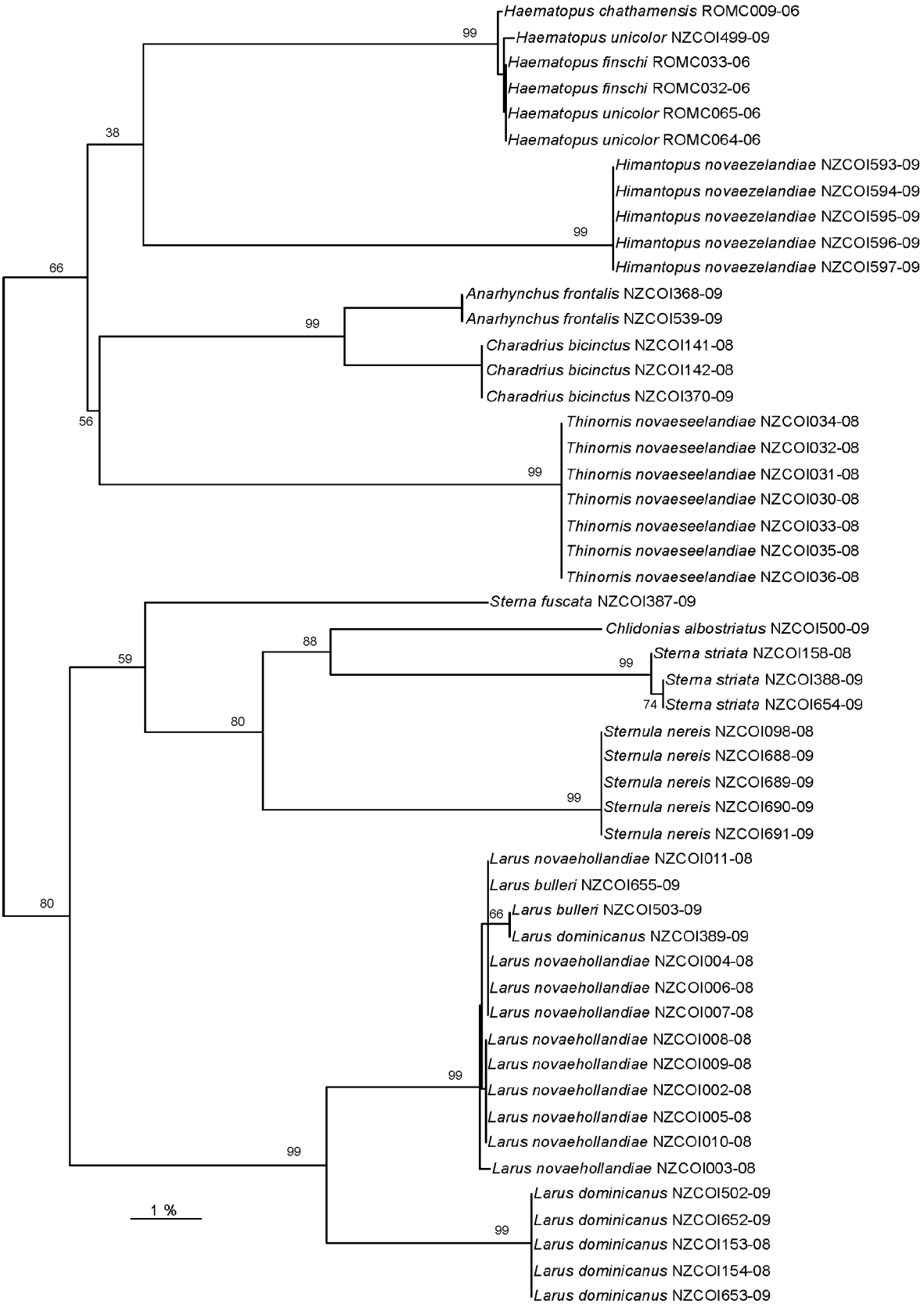


Figure 4.10 New Zealand Charadriiformes K2P neighbour joining tree with bootstrap values for all major branches.

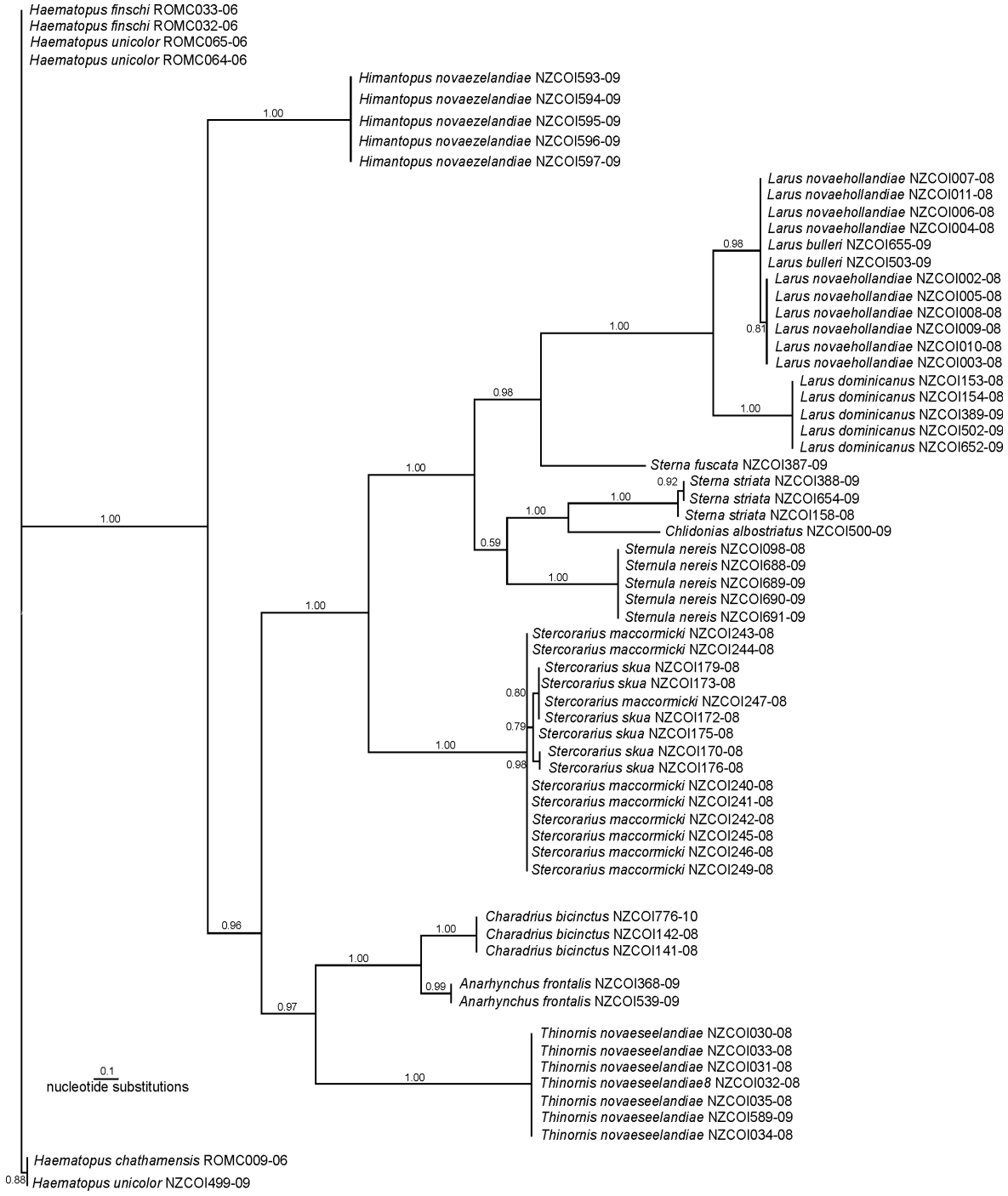


Figure 4.11 New Zealand Charadriiformes Bayesian tree with posterior probabilities given for each branch.

4.3.4 Ciconiformes, Falconiformes and Gruiformes

The orders Ciconiforme, Falconiforme and Gruiforme showed large genetic distance between species when analysed using both Bayesian and NJ K2P (5.181-12.511%) techniques. The one exception was two specimens of the Ciconiiforme *Egretta novaehollandiae,* one of which NZCOI532-09 appears to be a result of contamination and differs from the other at 102 sites (15.74%) within the barcode region. After matching itself in the BOLD search engine, the next closest were various Passerines. Other than this error, DNA barcodes clearly resolved all New Zealand species within these (Figures 4.12 – 4.17). For example, among Gruiformes, five species were analysed, all with multiple individuals; all were monophyletic in the NJ tree and had support at species nodes in the Bayesian tree of 0.98 to 1.00.

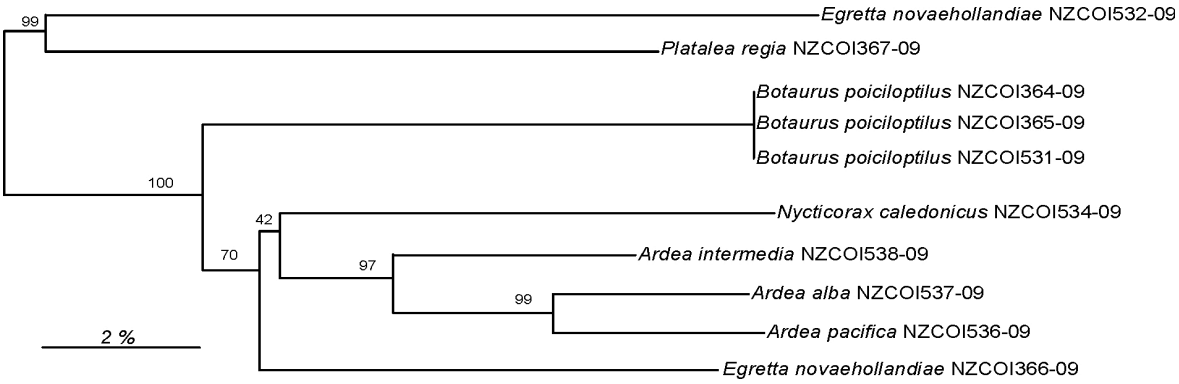


Figure 4.12 New Zealand Ciconiiformes K2P neighbour joining tree with bootstrap values for all major branches.

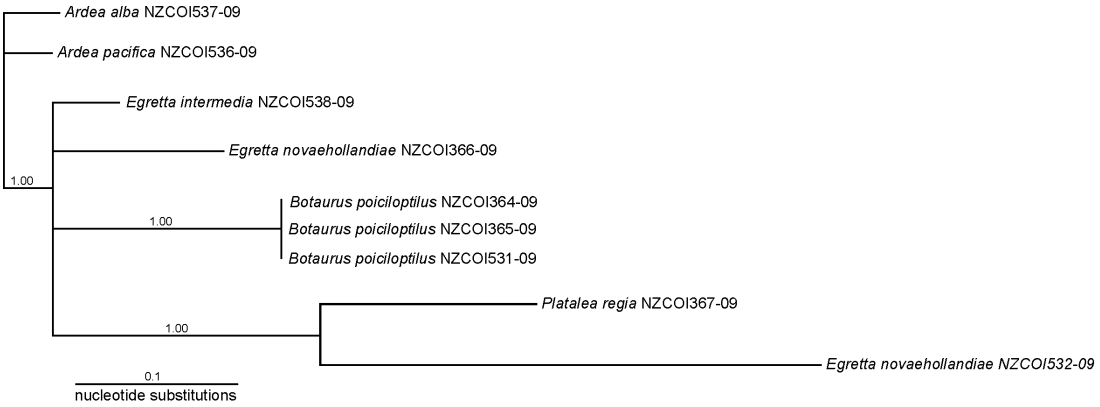


Figure 4.13 New Zealand Ciconiiformes Bayesian tree with posterior probabilities given for each branch.

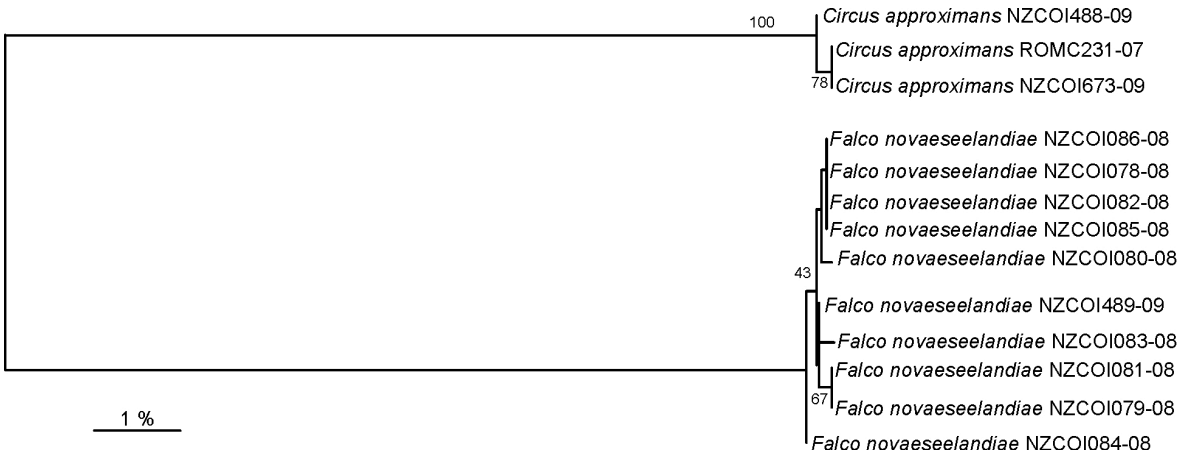


Figure 4.14 New Zealand Falconiformes K2P neighbour joining tree with bootstrap values for all major branches.

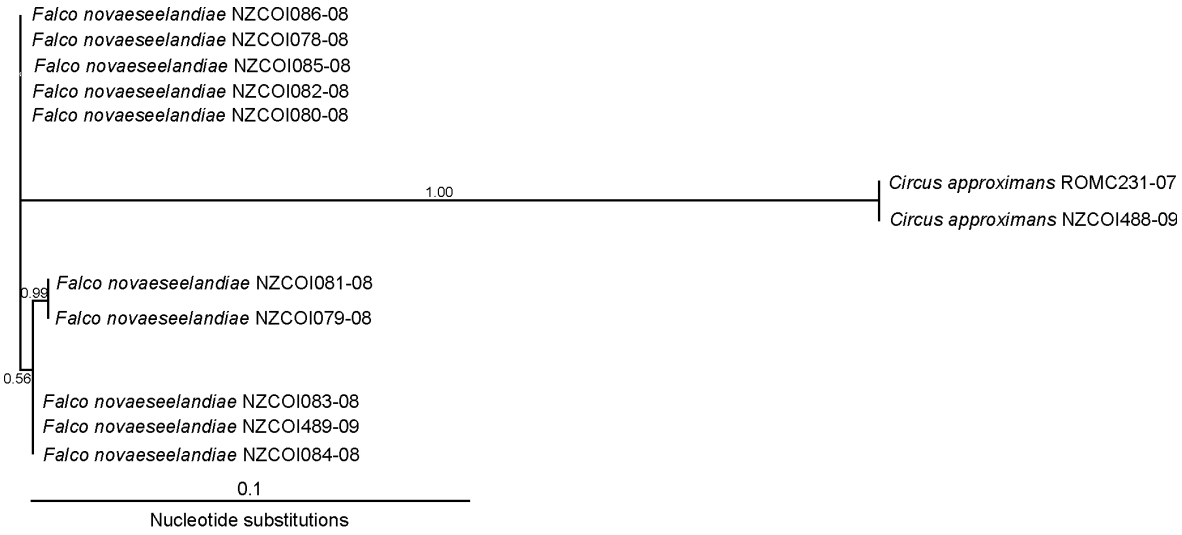


Figure 4.15 New Zealand Falconiformes Bayesian tree with posterior probabilities given for each branch.

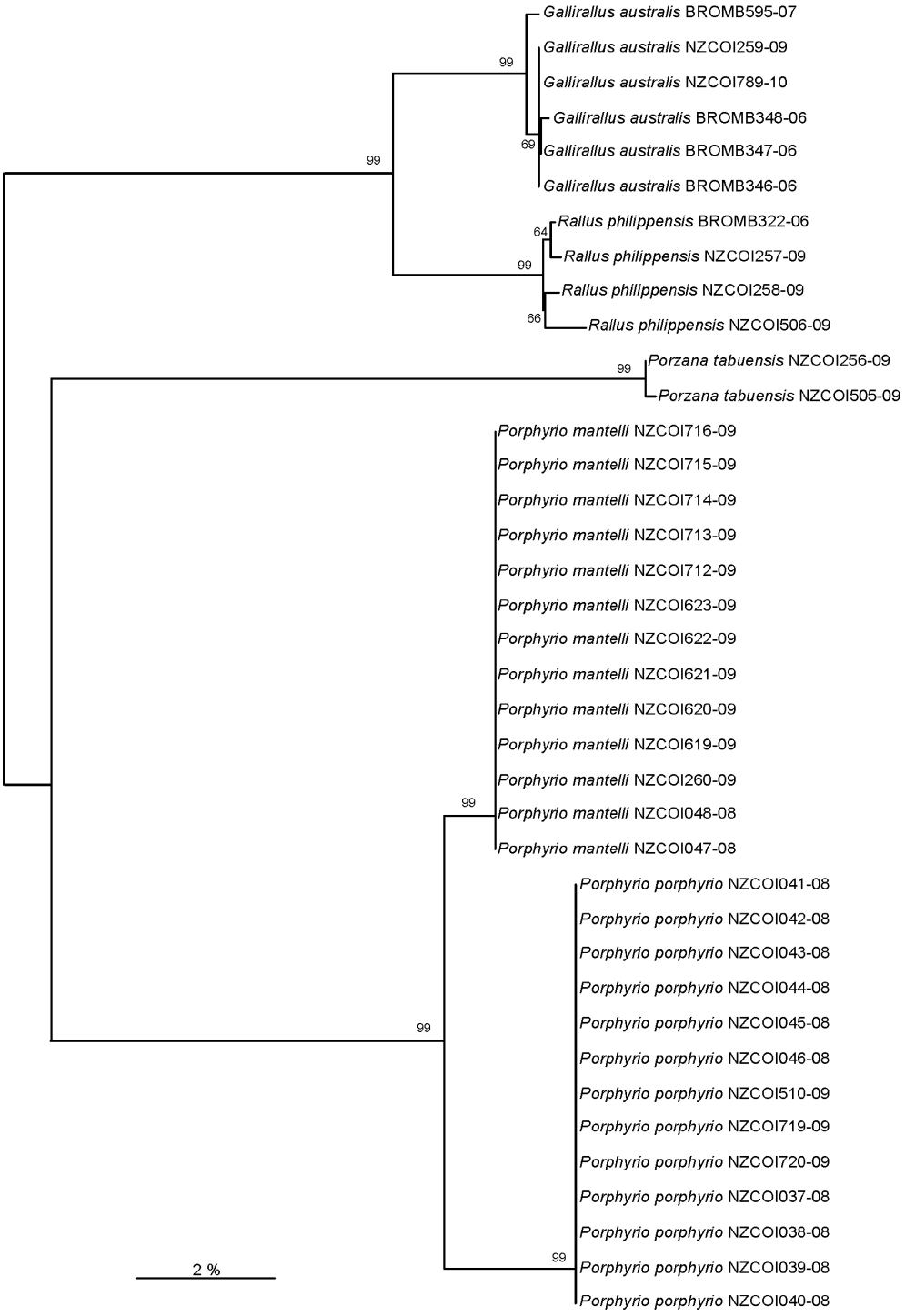


Figure 4.16 New Zealand Gruiformes K2P neighbour joining tree with bootstrap values for all major branches.

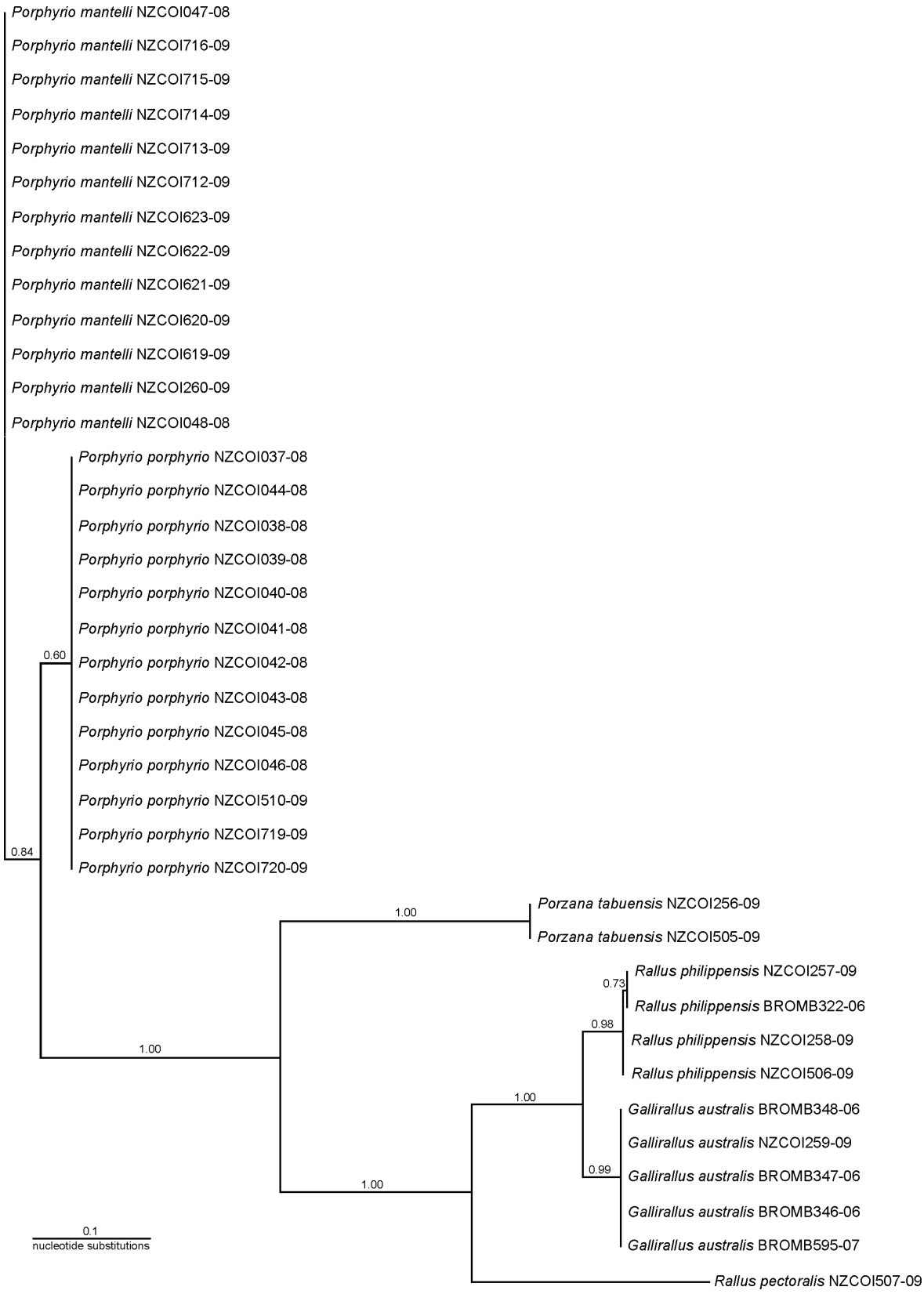


Figure 4.17 New Zealand Gruiformes Bayesian tree with posterior probabilities given for each branch.

4.3.5 Passeriformes

In general, species of the order Passeriforme were accurately distinguished using DNA barcodes; however, two species (*Acanthisitta chloris* and *Petroica australis*) showed distinct barcode clusters (Figures 4.18 and 4.19). Twenty species, including 19 of 20 native or endemic species were analysed of which, 16 were represented by >1 individual. Support at species nodes was 0.93 to 1.00. Support for clusters within species occurred in several species (*Acanthasita chloris* 0.70/0.80; *Mohoua novaeseelandiae* 0.88/0.99 and *Petroica australis* 0.99/1.00).

Acanthisitta chloris showed two distinct clades (Chapter 5.2.1). These clades correlate with their origins in either the Marlborough region of the South Island or Hawkes Bay in the North Island. Overall mean and maximum K2P distances, and the standard error (SE) were 1.142%, 2.216% and 0.147% compared with 0.186%, 0.311% and 0.048% for the South Island specimens and 0.198%, 0.395% and 0.081% for the North Island specimens.

Specimens of *Mohoua novaeseelandiae* showed two well supported clusters in both trees; however, all specimens came from the same area of the South Island (Kaikoura) and within species distance was slight (0-0.781%). Thus, these appear to be variable COI haplotypes within the species.

Petroica australis also showed two distinct clades, one of North Island specimens and one of specimens from the South Island and Stewart Island (Chapter 5.2.2). Overall mean K2P distance for this group was 2.043% with a maximum of 4.527% (SE 0.172%). However, the North Island specimens, which came from three locations (Rotorua, Waikato and Hawkes Bay), had identical COI haplotypes (specimens from Tiritiri Matangi Island in the Hauraki Gulf are originally from the Rotorua region and have been relocated for conservation purposes). Among South Island specimens from five locations (Marlborough Sounds, Nelson, Ulva Island in the Stewart Island region, Breaksea Island in Fiordland and Eglinton Valley in the Lake Wakatipu region), there was more variation in K2P distance than among the North Island specimens (mean 0.207%, max. 0.467%, SE 0.038%) but this was

still tenfold less than that observed when the two clades were examined together. The single Stewart Island specimen showed no sequence divergence from two of the other South Island specimens from the Wakatipu and Marlborough regions. It also showed little divergence (0.2% - 0.37%) from three other South Island specimens from Fiordland, Nelson and Marlborough.

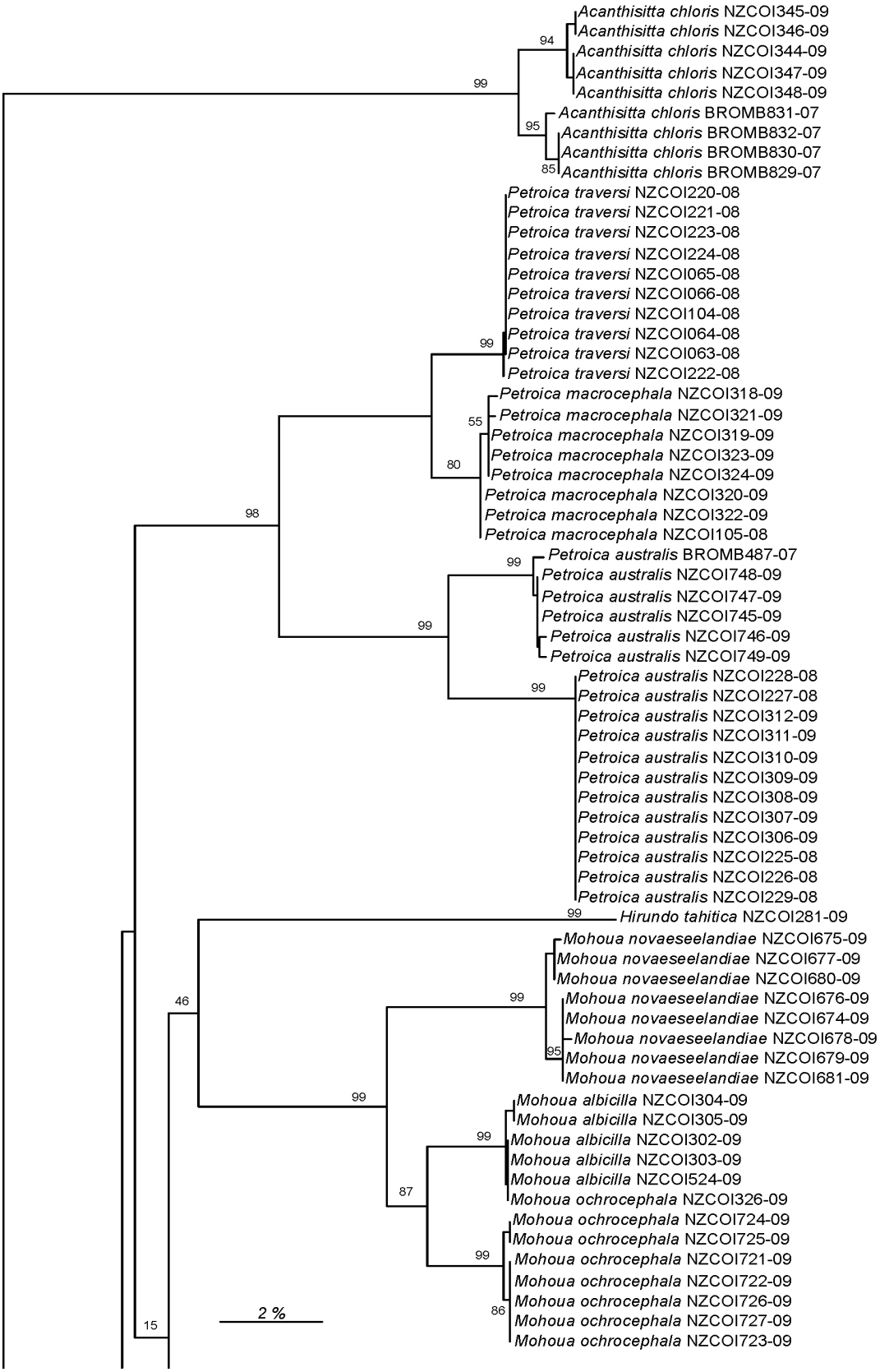


Figure 4.18 New Zealand Passeriformes K2P neighbour joining tree with bootstrap values for all major branches.

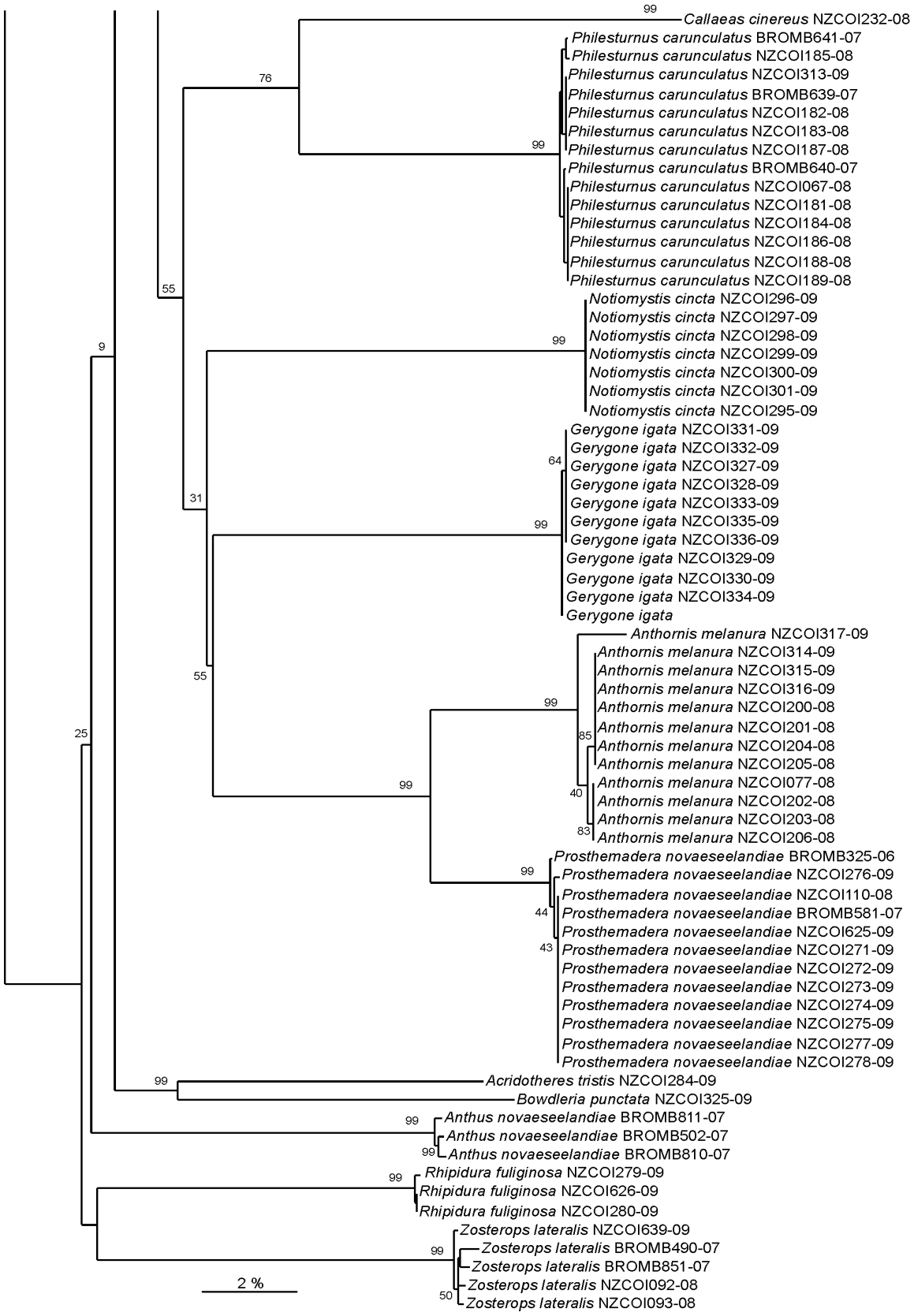


Figure 4.18 New Zealand Passeriformes K2P neighbour joining tree with bootstrap values for all major branches.**cont'd**

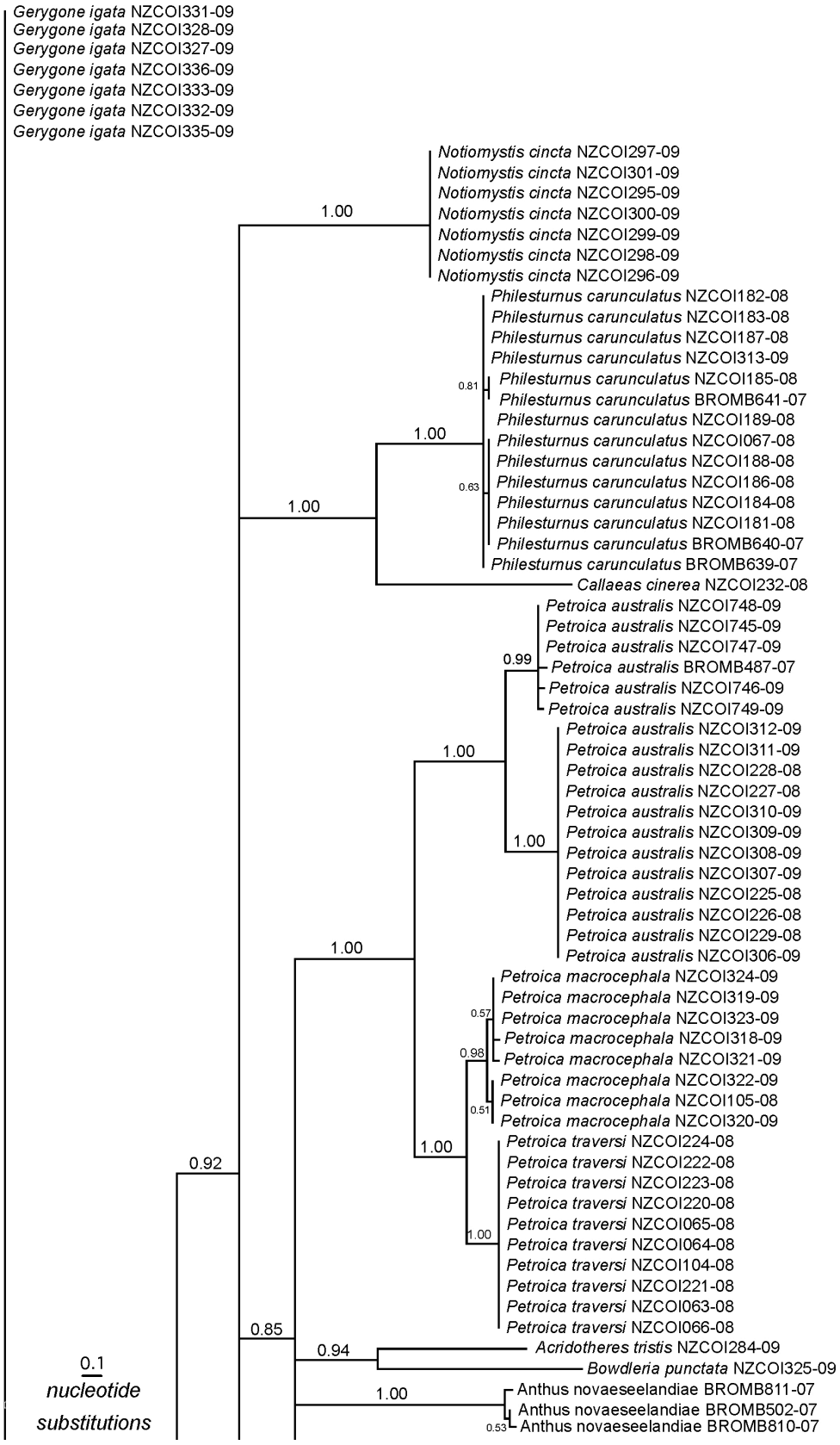


Figure 4.19 New Zealand Passeriformes Bayesian tree with posterior probabilities given for each branch.

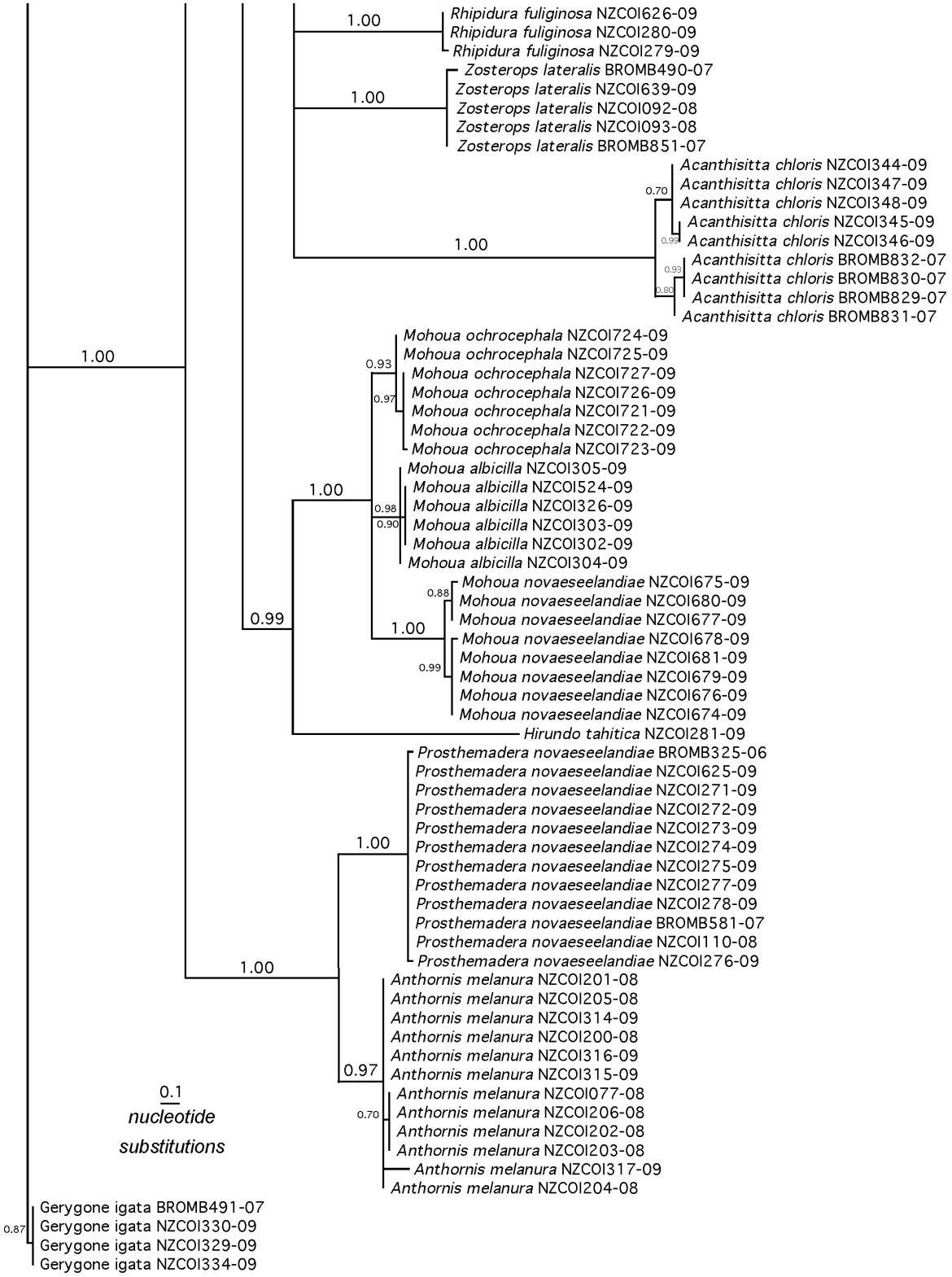


Figure 4.19 New Zealand Passeriformes Bayesian tree with posterior probabilities given for each branch. **cont'd**

4.3.6 Pelecaniformes

All the species of the Pelecaniformes from the Barcoding New Zealand Birds project were accurately differentiated using DNA barcoding (Figures 4.20 and 4.21). Both K2P and Bayesian trees clearly separated all species tested. Neither tree differentiated the genus *Leucocarbo* from the genus *Phalocorcorax,.* This may suggest the need for taxonomic revision; however, DNA barcoding is a system for differentiating species, not higher taxonomic groups.[37]

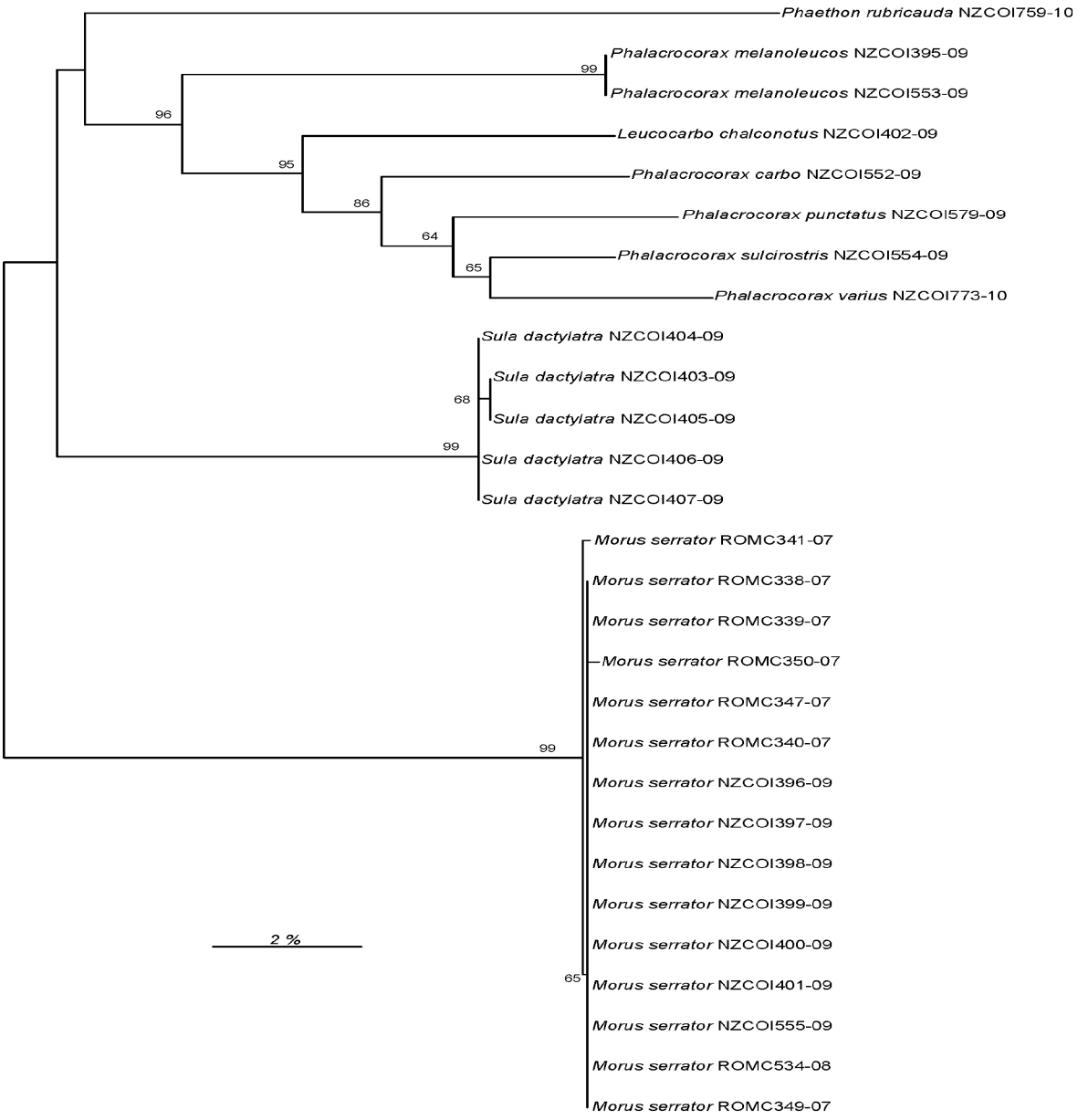


Figure 4.20 New Zealand Pelecaniformes K2P neighbour joining tree with bootstrap values for all major branches.

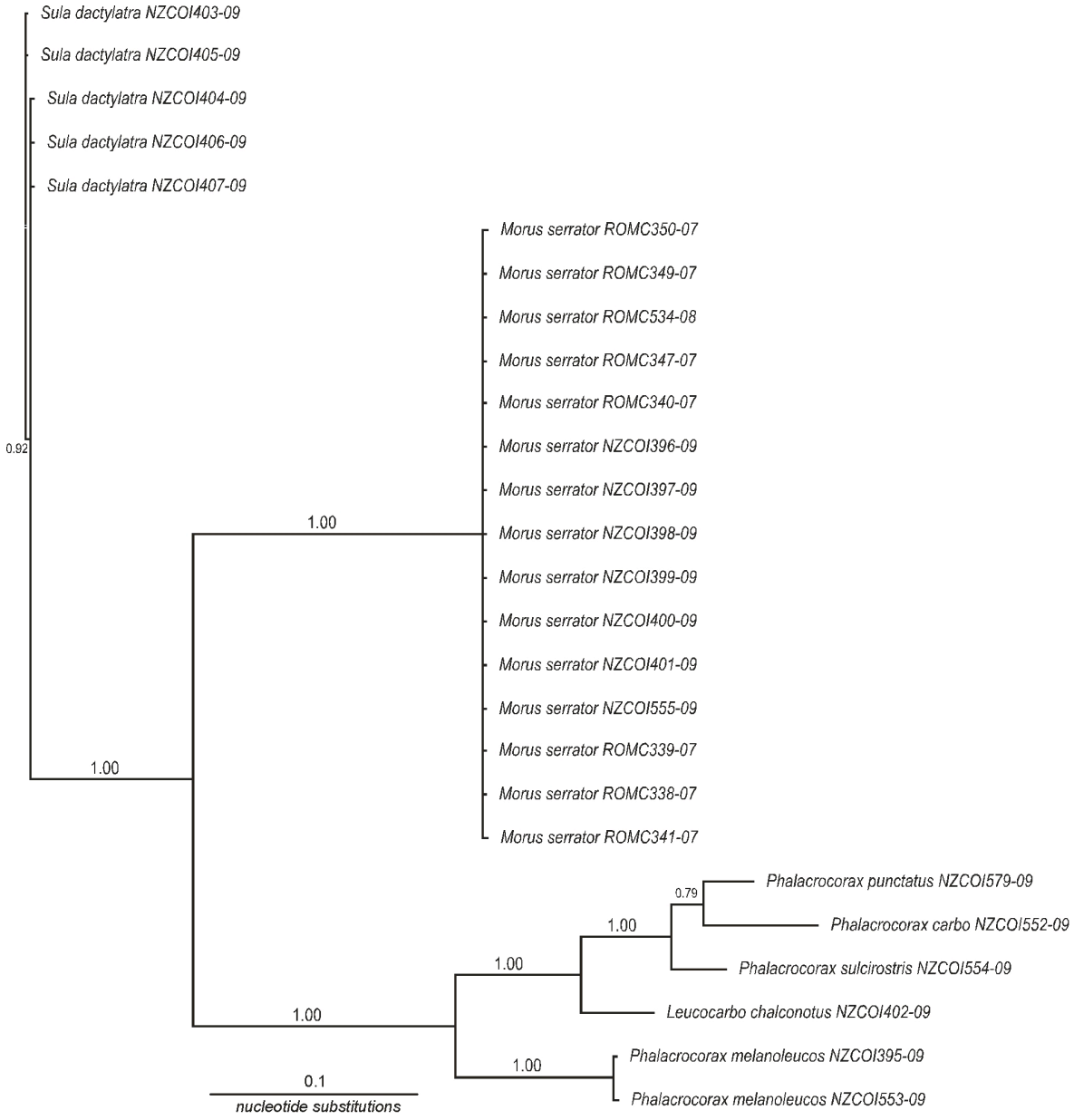


Figure 4.21 New Zealand Pelecaniformes Bayesian tree with posterior probabilities given for each branch.

4.3.7 Procellariiformes

The Procellariiformes are the largest avian order in New Zealand with 30 native and 13 endemic species (Figure 4.22 and 4.23). This order is characterised by many species that travel great distances around the Southern Ocean, sub-Antarctic, Pacific region and beyond. Some of these species were difficult to distinguish using barcodes (*Thalassarche melanophris* and *T. chrysostoma*; *Puffinus assimilis*. *P. gavia* and *P. griseus*) or contained quite distinct barcode clusters within species (*Pelagodroma marina* and *Pterodroma axillaris*).

Thalassarche melanophris and *T. chrysostoma* appear to be more closely related than most congeneric species but, nevertheless, yield distinct clades (Chapter 5.3.3). Analysis suggests a nearest neighbour distance of 0.47%, which compares with the intraspecific distance of 0.17% for *T. melanophris* and 0.22% for *T. chrysostoma*.

Puffinus assimilis. *P. gavia* and *P. griseus* produced a confusing result that showed high intraspecific variation and no nearest neighbour distance (Figures 4.22 – 4.23; Chapter 5.3.5). As with other ambiguous results, the possible reason for this is incorrect identification of specimens and is discussed in Chapter 5.

There were only four samples of *Pelagodroma marina* (Figures 4.22 and 4.23; Chapter 5.2.4). Three of these had similar DNA barcodes but the fourth one, NZCOI414-09, differed from the other three at 31 sites (4.78%). Such a difference would not normally be associated with a single species.

Pterodroma axillaris showed high intraspecific variation at the COI barcoding site in both K2P and Bayesian trees with three distinct clades being identified (Figures 4.22 and 4.23; Chapter 5.2.5).

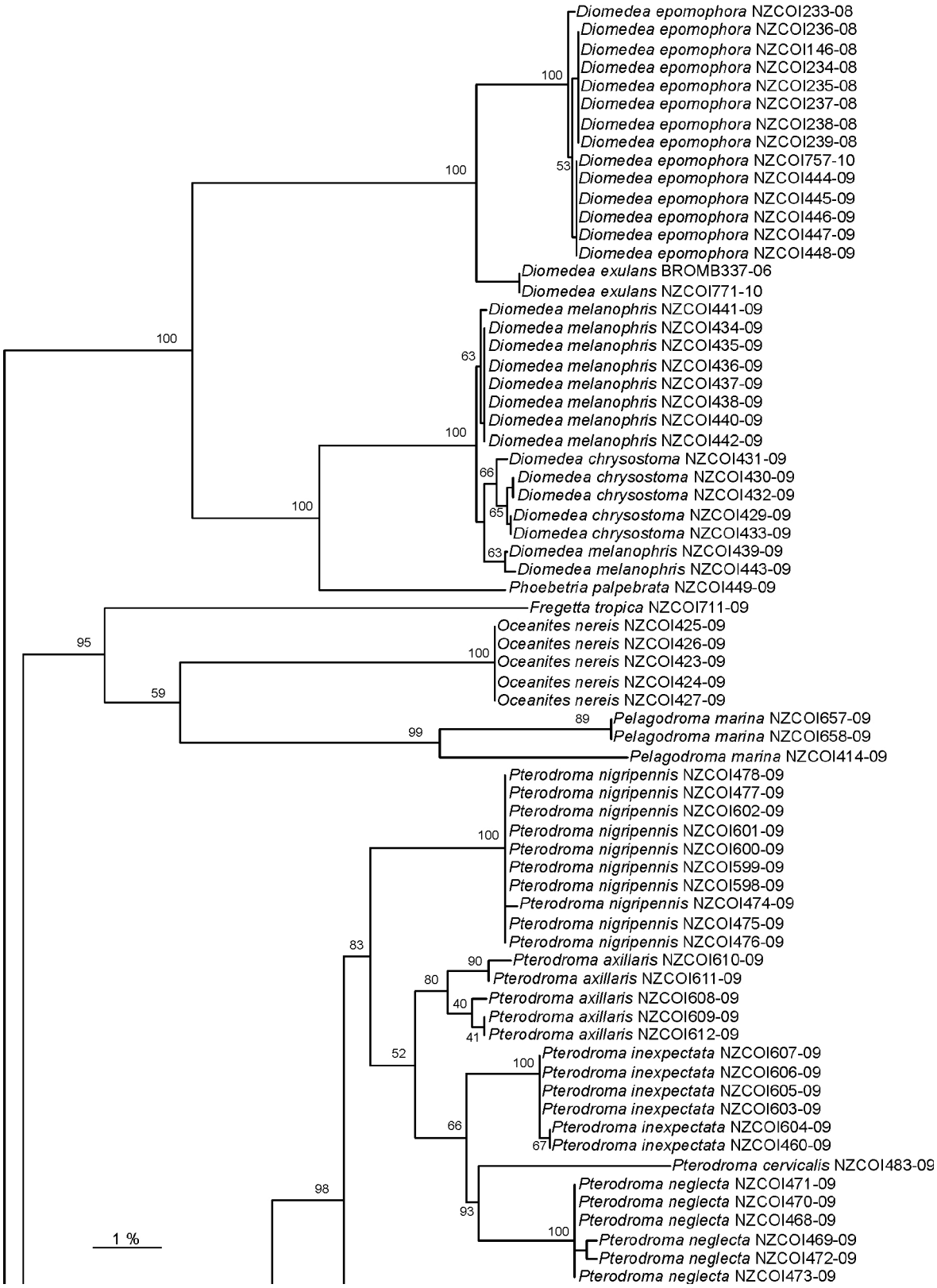


Figure 4.22 New Zealand Procellariiformes K2P neighbour joining tree with bootstrap values for all major branches.

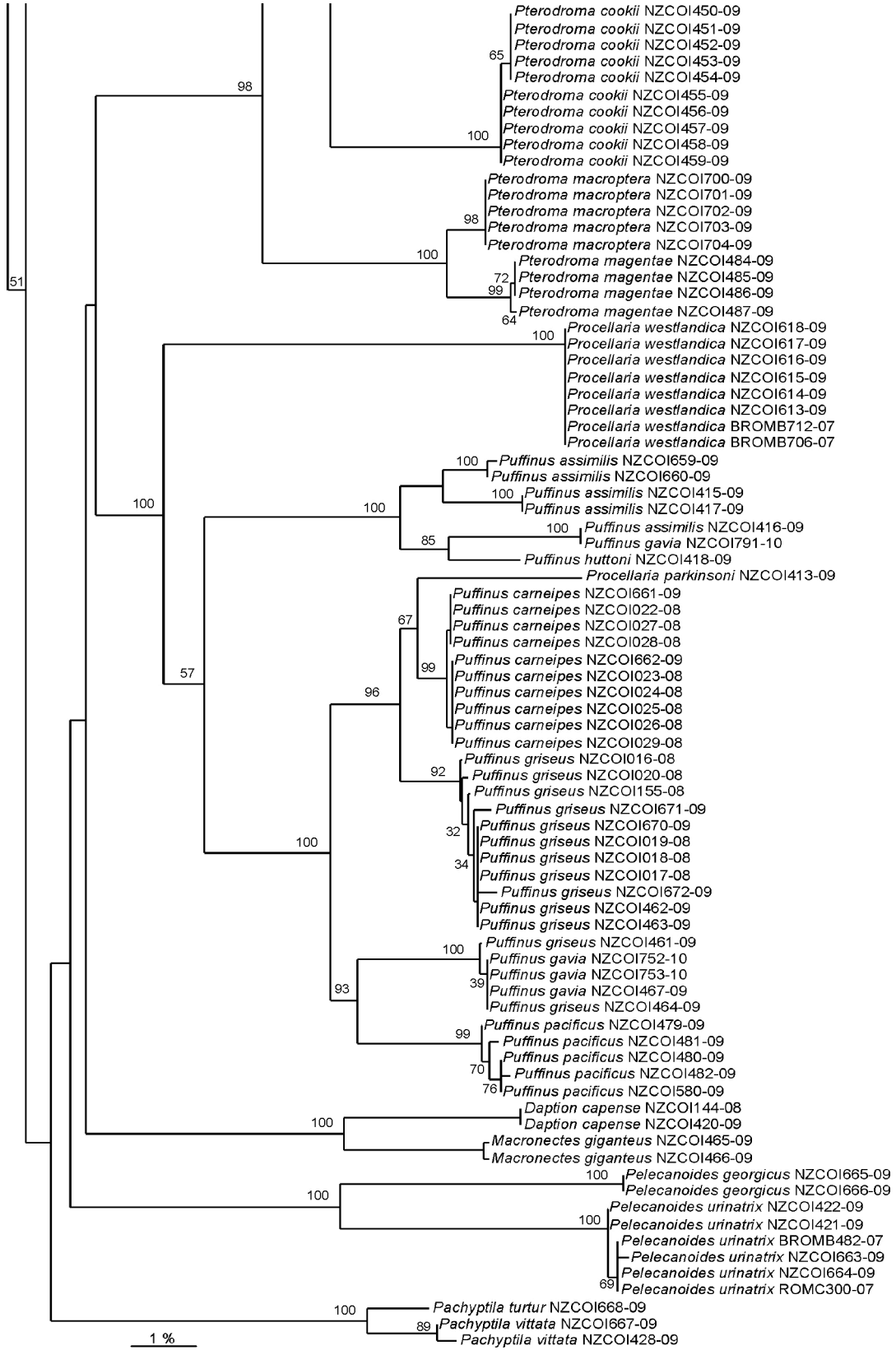


Figure 4.22 New Zealand Procellariiformes K2P neighbour joining tree with bootstrap values for all major branches.**cont'd**

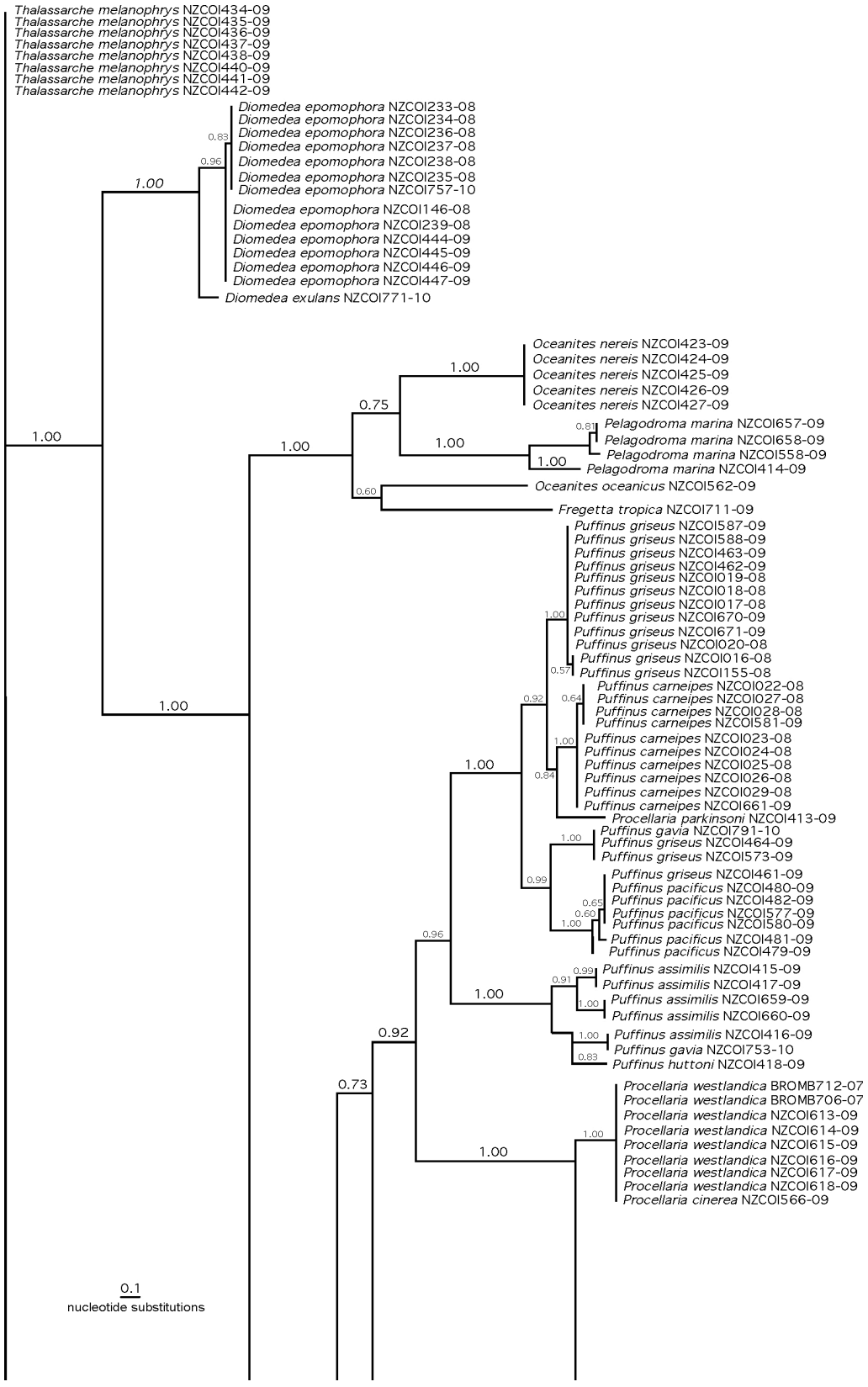


Figure 4.23 New Zealand Procellariiformes Bayesian tree with posterior probabilities given for each branch.

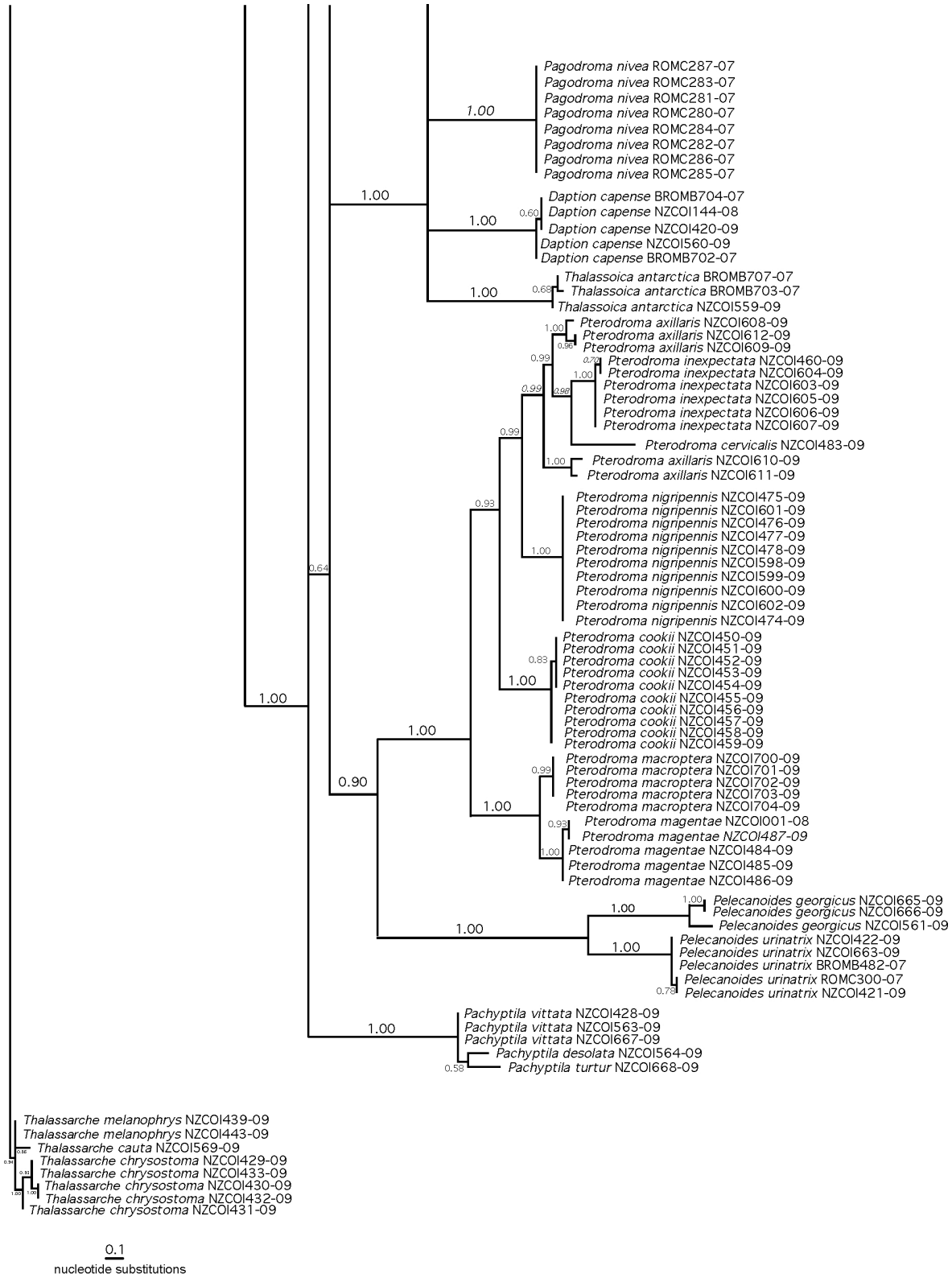


Figure 4.23 New Zealand Procellariiformes Bayesian tree with posterior probabilities given for each branch.**cont'd**

4.3.8 Psittaciformes

With the exception of species from the genus *Cyanorhamphus*, the New Zealand Psittaciformes were readily distinguished using DNA barcodes (Figures 4.24 and 4.25); clear separation was apparent between all other New Zealand species within this order.

Two species from the genus *Cyanorhamphus* (*C. novaezelandiae* and *C. unicolor*) showed little sequence divergence (Chapter 5.3.6). Only two nucleotide sites were diagnostic for the two species out of a total of four sites where there was any divergence. However, of seven specimens identified as *C. novaezelandiae*, four came from Antipodes Island and should, therefore, be *C. hochstetteri* according to current taxonomy. The remaining three *C. novaezelandiae* came from Tiritiri Matangi and Little Barrier Islands in the Hauraki Gulf, where birds of many species have been relocated for conservation purposes. Thus, their species allocation is probably correct.

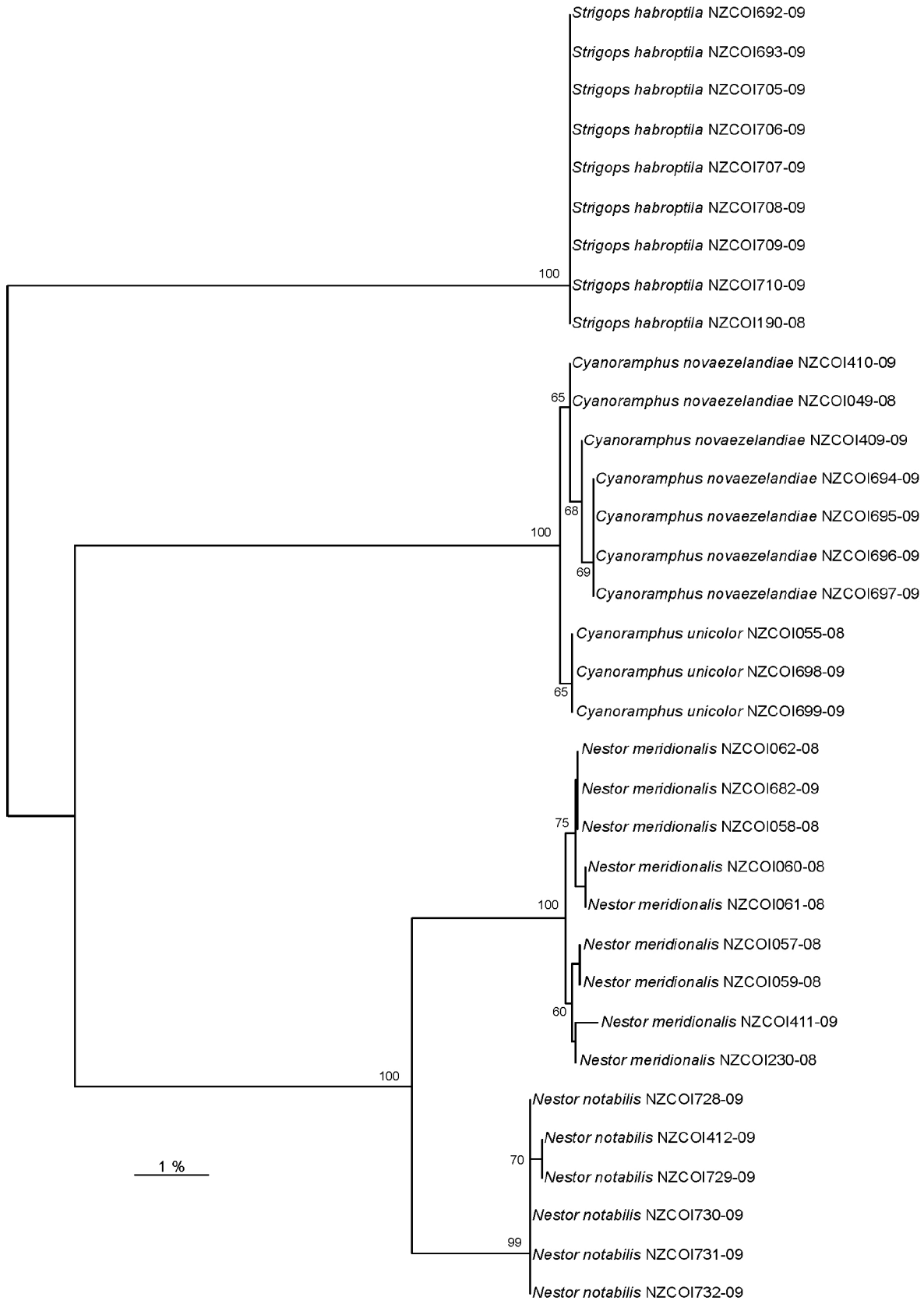


Figure 4.24 New Zealand Psittaciformes K2P neighbour joining tree with bootstrap values for all major branches.

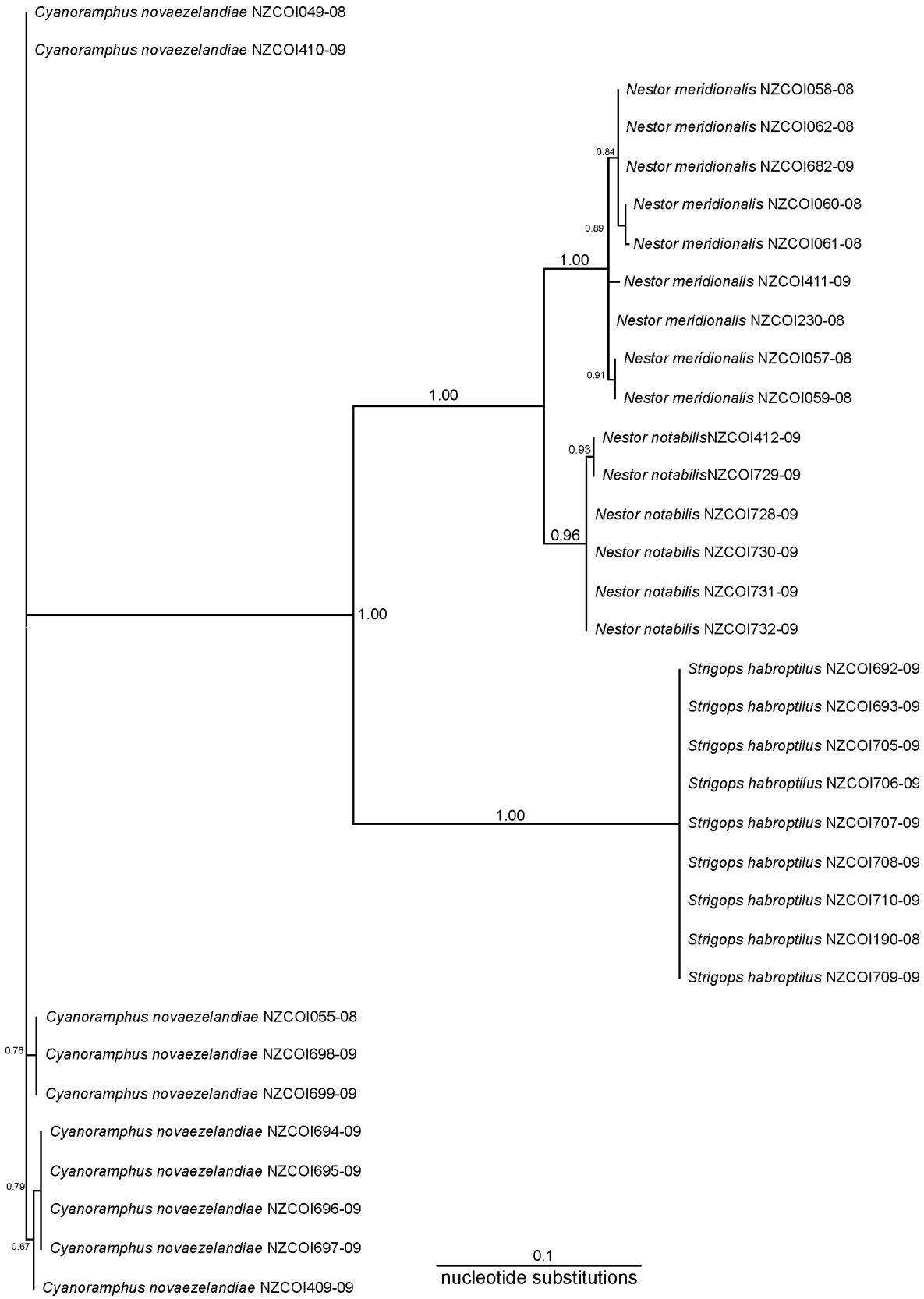


Figure 4.25 New Zealand Psittaciformes Bayesian tree with posterior probabilities given for each branch.

4.3.9 Sphenisiformes

DNA barcoding accurately differentiated species within the order Sphenisiforme (Figures 4.26 and 4.27). However, for two species, *Eudyptes chrysocome* and *Eudyptula minor,* there was considerable intraspecific sequence divergence.

37 *Eudyptes chrysocome* specimens from Campbell Island, Antipodes Island and Phillip Island in Australia and other unidentified locations showed considerable sequence divergence within this species (see Chapter 5.2.6). Four distinct clades and a number of isolated individuals were apparent. The greatest inter-clade divergence was 2.701% with an average divergence of 0.883%. Because the bulk of these specimens were from published data on GenBank (accession numbers DQ137172 and DQ525781-DQ525800)[38] and gave no indication of the geographic location of where they were obtained, it is not possible to determine whether this sequence divergence is allopatric or sympatric in origin.

Unusually high intraspecific variation within *Eudyptula minor* samples that was particularly apparent in the Bayesian tree (Figure 4.27) prompted assembly of a larger tree of 60 specimens from all sources (see Chapter 5.2.3). Samples in this larger tree came from a wide distribution that included Western Australia and Victoria in Australia, and a range of geographical locations within the North and South Islands of New Zealand. The resulting phylogenetic distance tree contained two distinct clades. A small amount of intra-clade variation occurred (<0.01%); however the inter-clade variation was >1.8%. Although this sequence divergence is not great, the difference between intra- and inter-clade variation is nearly twenty-fold. Of the 25 samples in the first clade, 9 were from Victoria and 14 are from Western Australia; two others had no geographic location data. Sixteen of the remaining 35 specimens were from the North Island, 18 from the South Island and the final one had no geographic location indicated.

There was no apparent difference between North and South Island specimens from New Zealand or between specimens from the Canterbury/Banks Peninsular

region and the rest of New Zealand nor were there any significant differences between specimens from the two Australian regions.

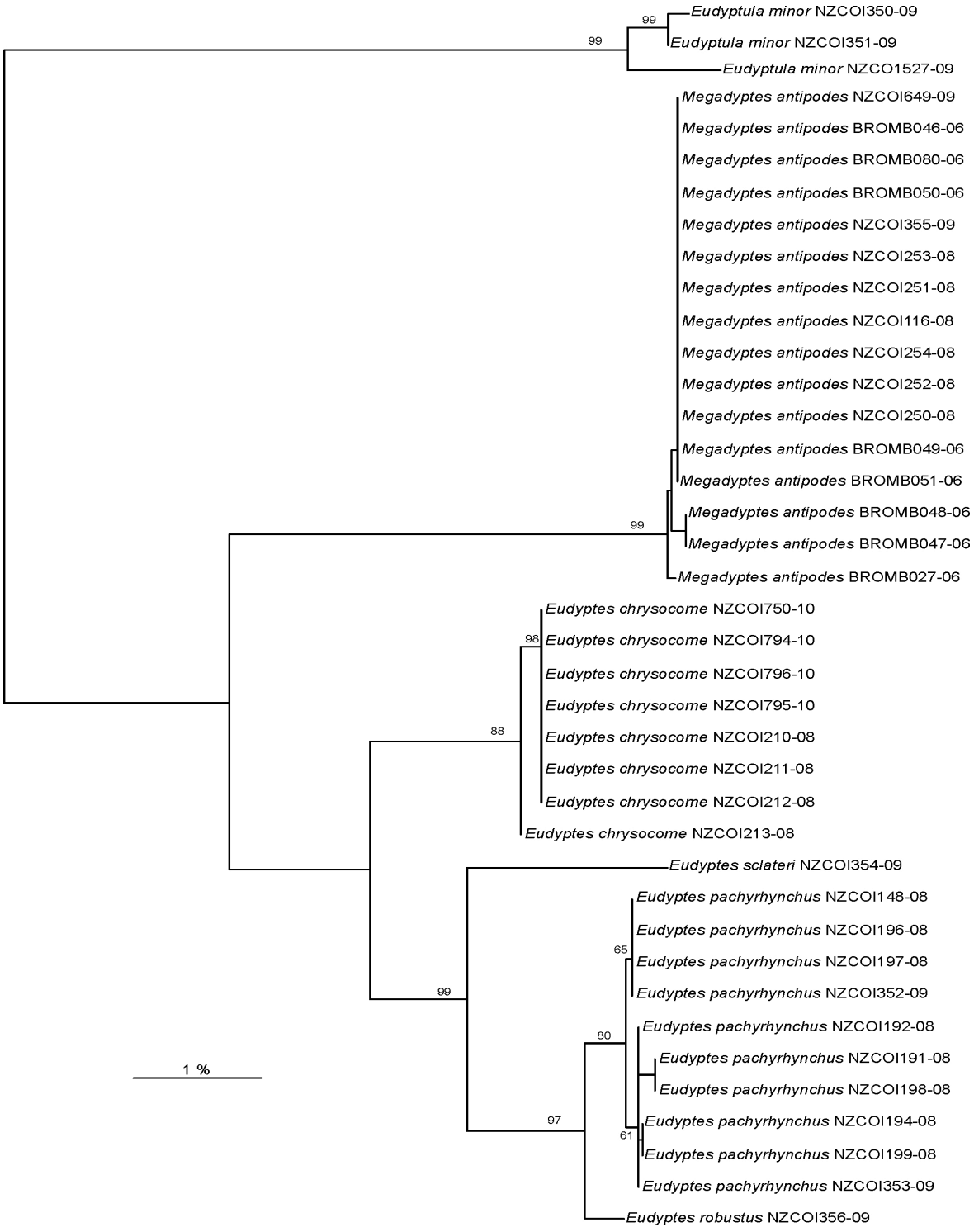


Figure 4.26 New Zealand Sphenisciformes K2P neighbour joining tree with bootstrap values for all major branches.

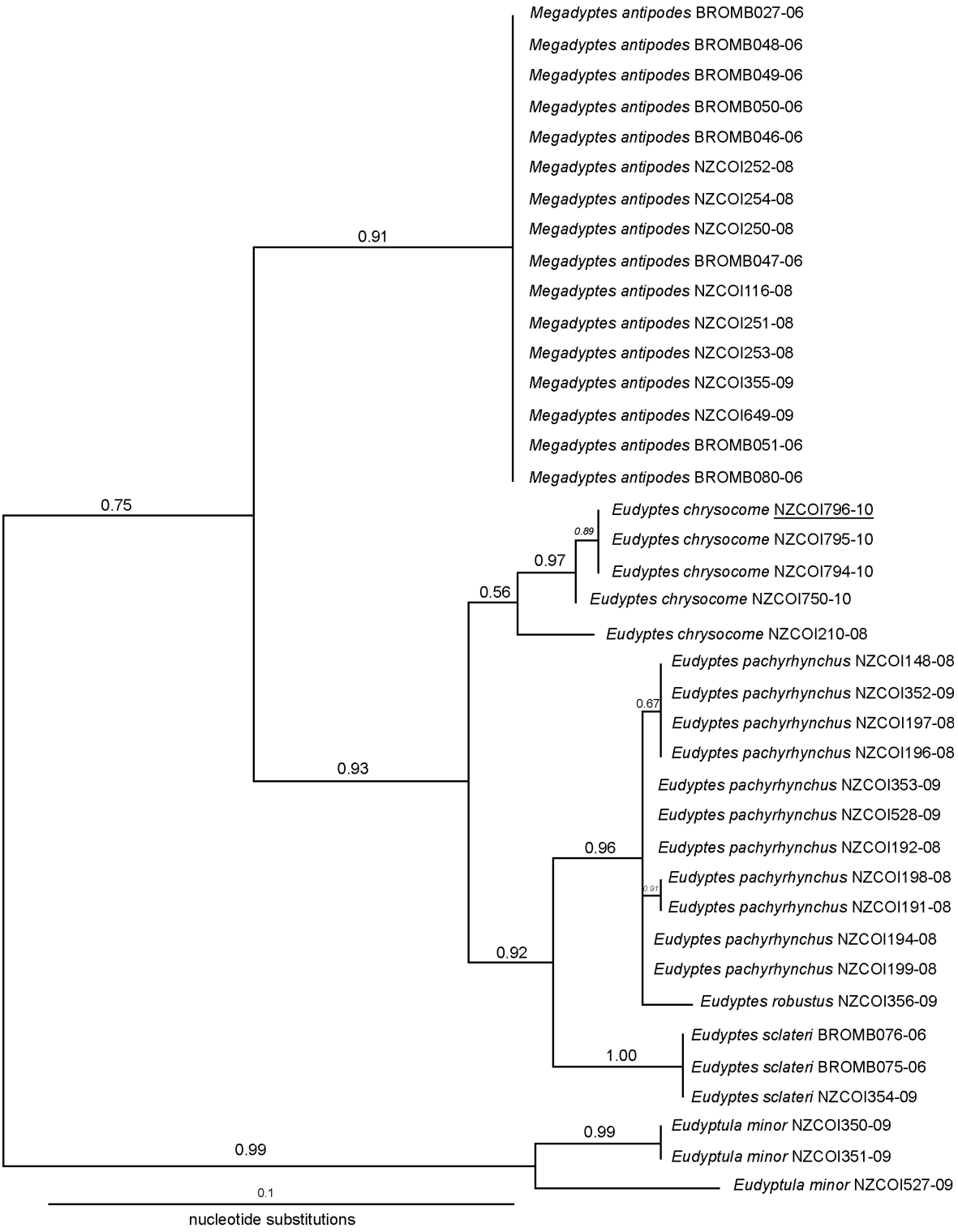


Figure 4.27 New Zealand Sphenisciformes Bayesian tree with posterior probabilities given for each branch.

4.4 References

1. *Barcode of Life Data Systems*. 2009 [cited 2010 22nd July]; Available from: http://www.barcodinglife.org/.

2. *All Birds Barcoding Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.barcodingbirds.org/#.

3. *Consortium for the Barcode of Life*. 2009 [cited 2009 11th November]; Available from: http://barcoding.si.edu/.

4. *All Leps Barcode of Life*. 2009 [cited 2009 18 October]; Available from: http://www.lepbarcoding.org/.

5. *International Network for Barcoding Invasive and Pest Species (INBIPS)*. 2010; Available from: http://www.barcoding.si.edu/INBIPS.htm.

6. *Polar Boarcode of Life*. 2010; Available from: http://www.polarbarcoding.org/.

7. *Fish Barcode of Life Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.fishbol.org/.

8. Hebert, P., Stoeckle, MY, Zemlak, TS, Francis, CM *Identification of Birds through DNA Barcodes.* PLoS Biol, 2004. **2**(10).

9. Kerr, K., Stoeckle, MY, Dove, CJ, Weigt, LA, Francis, CM, Hebert, PDN, *Comprehensive DNA barcode coverage of North American birds.* Molecular Ecology Notes, 2007. **7**(4): p. 535-543.

10. Yoo, H., Eah, JY, Kim, JS, Kim, YJ, Min, MS, Paek, WK, Lee, H, Kim, CB, *DNA barcoding Korean birds.* Molecules and Cells, 2006. **22**(3): p. 323-327.

11. Park, H., Yoo, HS, Jung, G and Kim CB, *New DNA barcodes for identification of Korean birds.* Genes & Genomics, 2011. **33**: p. 91-95.

12. Kerr, K.C., Lijtmaer, D A, Barreira, A S, Hebert, PDN, Tubaro, P L, *Probing evolutionary patterns in neotropical birds through DNA barcodes.* PLoS One, 2009. **4**(2): p. e4379.

13. Johnsen, A., Rindal, E, Ericson, PGP, Zuccon, D, Kerr, KCR, Stoeckle, MY and Lifjeld, JT, *DNA barcoding Scandinavian Birds reveals divergent lineages in trans-Atlantic species.* J. Ornithol, 2010. **151**(3): p. 565-578.

14. *Checklist of the birds of New Zealand*. 4th ed, ed. B. Gill, Bell, BD, Chambers, GK,Medway, DG, Palma, RL, Scofield, RP, Tennyson AJD and Worthy TH. Vol. 1. 2010: Te Papa Press.

15. *New Zealand recognised bird names (NZRBN) database*. 2003 [cited 2010 28 January]; Available from: http://www.bird.org.nz/nzrbn.htm.

16. Fleming, C., *New Zealand as a minor source of terrestrial plants and animals in the Pacific.* Tuatara, 1976. **22**: p. 30-37.

17. Trewick, S.A. and G.C. Gibb, *Vicars, tramps and assembly of the New Zealand avifauna: a review of molecular phylogenetic evidence.* Ibis, 2010. **152**(2): p. 226-253.

18. Kimura, M., *A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences.* Journal of Molecular Evolution, 1980. **16**: p. 111–120.

19. Holder, M. and P.O. Lewis, *Phylogeny estimation: traditional and Bayesian approaches.* Nat Rev Genet, 2003. **4**(4): p. 275-84.

20. Rach, J., et al., *Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata.* Proc Biol Sci, 2008. **275**(1632): p. 237-47.

21. *Simple counts as a distance measure*. 2010 [cited 2010 19th August]; Available from: http://helix.mcmaster.ca/721/distance/node2.html#SECTION00011000000000000000.

22. Jukes, T.H., *Recent advances in studies of evolutionary relationships between proteins and nucleic acids.* Space Life Sci, 1969. **1**(4): p. 469-90.

23. *Jukes - Cantor Correction*. Computational Biology 2010; Available from: http://helix.mcmaster.ca/721/distance/node3.html.

24. Jukes, T., Cantor, CR, *Evolution of protein molecules.*, ed. H. Munro. 1969, New York: Academic Press. 21-123.

25. *Kimura 2-parameter Correction*. Computational Biology 2010 [cited 2010 19th August]; Available from: http://helix.mcmaster.ca/721/distance/node4.html.

26. Aliabadian, M., Kaboli, M, Nijman, V, Vences, M, *Molecular identification of birds: performance of distance-based DNA barcoding in three genes to delimit parapatric species.* PLoS One, 2009. **4**(1): p. e4119.

27. DeSalle, R., Egan, MG, Siddall, M, *The unholy trinity: taxonomy, species delimitation and DNA barcoding.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1905-16.

28. DeSalle, R., *Phenetic and DNA taxonomy; a comment on Waugh.* BioEssays, 2007. **29**(12): p. 1289-1290.

29. Waugh, J., Huynen, L, Millar, C, Lambert, D, *DNA barcoding of animal species - response to DeSalle.* Bioessays, 2008. **30**(1): p. 92-93.

30. *Barcode of Life website*. 2009 [cited 2009 15 October]; Available from: http://barcoding.si.edu/.

31. Ratnasingham, S. and P.D.N. Hebert, *BOLD: The Barcode of Life Data System (www.barcodinglife.org)*. Molecular Ecology Notes, 2007. **7**(3): p. 355-364.

32. Atteson, K., *The performance of neighbor-joining algorithms of phylogeny reconstruction*. Lecture Notes in Computer Science, ed. T. Jiang, and Lee, D. Vol. 1276. 1997, Berlin: Springer-Verlag.

33. Pittman, J., Huang, E, Nevins, J, Wang, Q, West, M, *Bayesian analysis of binary prediction tree models for retrospectively sampled outcomes.* Biostatistics, 2004. **5**(4): p. 587-601.

34. Ronquist, F., Huelsenbeck, JP, and van der Mark, P., *MrBayes 3.1 Manual*. 2005, Florida State University: Tallahassee.

35. Heather, B., Robertson, H., *The Field Guide to the Birds of New Zealand*. 3rd ed. 2005, Auckland: Penguin Books.

36. Tennyson, A., Palma, RL, Robertson, HA, Worthy, TH, Gill, BJ., *A new species of kiwi (Aves, Apterygiformes) from Okarito, New Zealand.* Records of the Auckland Museum, 2003. **40**: p. 55-64.

37. Greenstone, M., Rowley, DL, Heimbach, U, Lundgren, JG, Pfannenstiel, RS, Rehner, SA, *Barcoding generalist predators by polymerase chain reaction: carabids and spiders.* Mol Ecol, 2005. **14**(10): p. 3247-66.

38. Baker, A., Pereira, SL, Haddrath, OP, Edge, KA, *Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling.* Proceedings of the Royal Society B-Biological Sciences, 2006. **273**(1582): p. 11-17.

Chapter 5

Species for which DNA barcoding produced ambiguous results

5.1 An overview of ambiguous results

Considerable and sometimes acrimonious debate has arisen over the role of DNA barcoding in taxonomy.[1-11] Initially, some taxonomists saw it as a threat to their expertise and a drain on limited resources.[7] Others thought that there was a move afoot to replace taxonomy with DNA sequences.[10, 11] However, with the passage of time and in the light of how it *has* been used, DNA barcoding is becoming an increasingly accepted part of a species description.[12]

For taxonomic purposes, haplotypic variation in part of the COI gene, in the absence of other evidence, can never be used as an indicator of species delimitation.[13] Morphological information is also never used by itself without corroborating data. To look only at a portion of the COI gene of an organism and then call this organism a species would be to ignore the intellectual and scientific content of taxonomy.[10, 14] On the other hand this data can be used to augment taxonomy and to provide information to precipitate or assist discussions about species delimitation.[4] Furthermore, a sequence of DNA is biological data, obtained from living material, as apposed to conceptual data derived from taxonomic understanding of phylogeny and is, therefore, intrinsically objective.

Early in the use of DNA barcoding, it became apparent that it might be useful for highlighting potential cryptic species and informing the debate around their status.[15-17] For example, Kerr et al. (2007)[18] found fifteen putative species among the birds of North America that contained two distinct barcode clusters, which might represent cryptic species. Thus, the analysis of barcode sequences for the birds of New Zealand might be expected to highlight some species in which such

clusters are found and, perhaps, others where species that appear to be different share very similar barcodes.

While speciation may result in genetic divergence, speciation is not caused by genetic divergence. Ultimately, it is for the taxonomists to conclude whether these DNA clusters are real expressions of taxonomic differences or merely extreme intraspecific genetic variation. Likewise, for those species that appear to share very similar barcodes, taxonomists must decide whether this is an example of morphological/behavioural variation within a single species or an example of speciation associated with very little divergence at the molecular level[19, 20] Conversely, there may not have been enough time since the speciation event for the genetic divergences typically observed between species to occur.[21]

There are many isolated populations of birds in New Zealand and the regions covered by this thesis. This is especially true of species found on offshore islands. These populations have frequently been accorded species status of their own, although, in many cases there is a lively debate over whether they really constitute separate species.[22] This debate is lent a political slant because, in general, these populations are small and some are under threat. Conservation agencies, with finite resources, are more likely to direct their attention at conserving endangered species than at subspecies or even just local populations. Thus, accurate species identification may influence the deployment or otherwise of these resources.

It is beyond the scope of this thesis to fully characterise species within taxonomic groups that showed either divergent or convergent COI haplotypes (see Chapter 4). What follows is a discussion of the species for which DNA barcoding provided ambiguous results. An attempt is made here to give an overview of any debate that may surround their species assignment and to present the evidence obtained in this study that may contribute to or stimulate this debate. Throughout this chapter, it is fully acknowledged that DNA barcodes alone are insufficient evidence for any such assignment.

Analyses of COI sequences from species collected for this thesis revealed a total of six species that were found to have two distinct DNA barcode clusters within them indicating the possible presence of cryptic species or subspecies. In addition, six groups of congeneric species had very small interspecific genetic distance, indicating that they may be subspecies or just one species.

5.2 Species that contained DNA barcode clusters

5.2.1 Acanthisitta chloris

Both K2P and Bayesian trees (Figures 4.18 and 4.19) suggest that *Acanthasita chloris*, the New Zealand rifleman, has two distinct barcode clusters (Figure 5.1). Samples of this species came from two locations, four specimens were from Marlborough in the northern part of the South Island and five came from Hawkes Bay on the east coast of the North Island. According to Heather & Robertson (2005),[23] these birds are known to have poor dispersal abilities across water or open habitats and, thus, the two populations would be unlikely to interbreed.

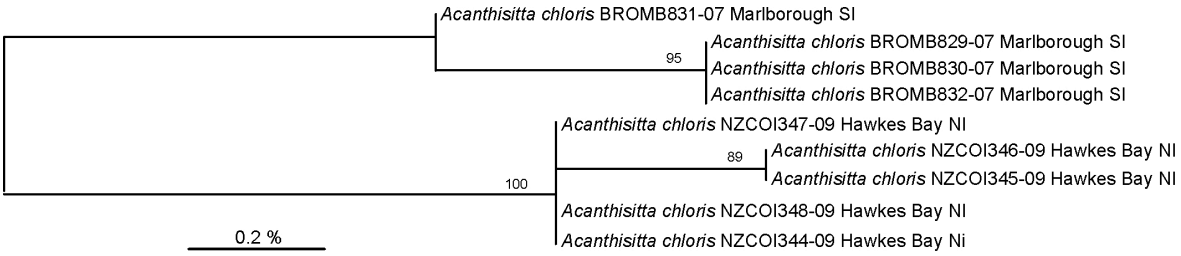


Figure 5.1 A*canthisitta chloris* K2P neighbour joining tree with bootstrap values for all major branches.

There are two recognised subspecies, *A. chloris chloris* (the South Island Rifleman) and *A. chloris granti* (the North Island Rifleman).[24] DNA barcoding distinguished these at eleven nucleotide sites within the 648 bp barcode region. These eleven nucleotides could serve as a character-based means of distinguishing birds from these locations. Mean and maximum K2P distance (Figure 5.2) for within the Marlborough (0.186% and 0.311%) and Hawkes Bay groups (0.198% and 0.395%) were considerably lower than that observed when both groups were combined (1.142% and 2.216%). The divergence observed

between these two groups does not match the tenfold difference between intra- and interspecific variation or the minimum distance of 2.7% suggested by Hebert et al.2003[25] as being indicative of separate species; however, these two recognised subspecies show clear divergence in COI haplotypic diversity.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	9	1	36	0	1.142	2.216	0.147
Within South Island Specimens	5	1	10	0	0.186	0.311	0.048
Within North Island Specimens	4	1	6	0	0.198	0.395	0.081

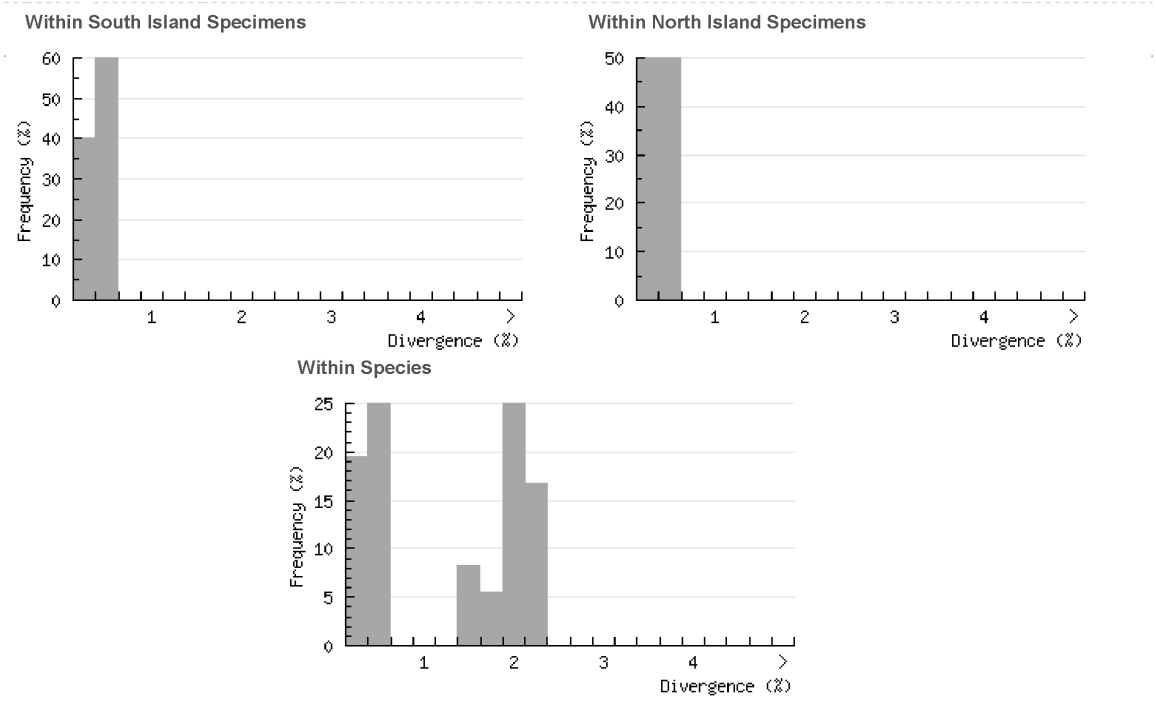


Figure 5.2 K2P distance summary for *Acanthisitta chloris*.

5.2.2 Petroica australis

Until August 2010, *Petroica australis* was considered one species with three subspecies, *P. australis longipes* (the North Island robin), *P. australis australis* (the South Island robin) and *P. australis rakiura* (the Stewart Island robin).[26] However, a new Checklist of the Birds of New Zealand, published in August 2010[24] concurs with Holdaway et al. (2001)[27] and recognises the North Island robin as a separate species, *P. longipes* as well as the South Island robin *P. australis australis* with a Stewart Island subspecies, *P. australis rakiura*.

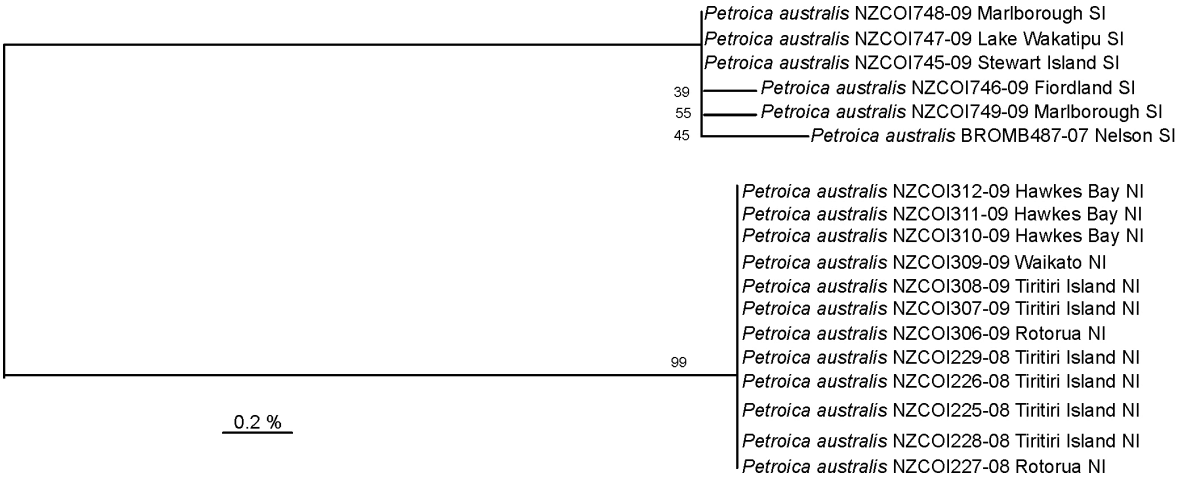


Figure 5.3 ***Petroica australis*** K2P neighbour joining tree with bootstrap values for all major branches.

Interestingly, DNA barcodes provide supporting evidence for this revision, although the status of the Stewart Island birds as a separate subspecies is not supported. Both K2P and Baysian trees of COI sequences from these species, assembled prior to publication of the Checklist's revision, identified two distinct barcode clusters (Figures 4.18, 4.19 and 5.3). These correlate with specimens from the North and South Islands. Very close to a tenfold difference in intra- versus interspecific sequence divergence (mean 0.207 *vs* 2.043%, maximum 0.467 *vs* 4.527%) in the barcode region of the COI gene occurs between specimens of these two newly recognised species (Figure 5.4). Within the 12 North Island specimens there was no divergence at all; all haplotypes were identical. In South Island specimens there was very little divergence (mean 0.207%, maximum 0.467%). The haplotypes from the South Island specimens and the single Stewart Island specimen were identical (minimum 0% maximum 0.467%). Sequences differed diagnostically between North and South Island specimens at 27 sites in the barcode region compared with no sites in the North Island specimens and four sites among the six South/Stewart Island specimens.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	18	1	153	0	2.043	4.527	0.172
Within North Island Specimens	12	1	66	0	0	0	0
Within South Island Specimens	6	1	15	0	0.207	0.467	0.038

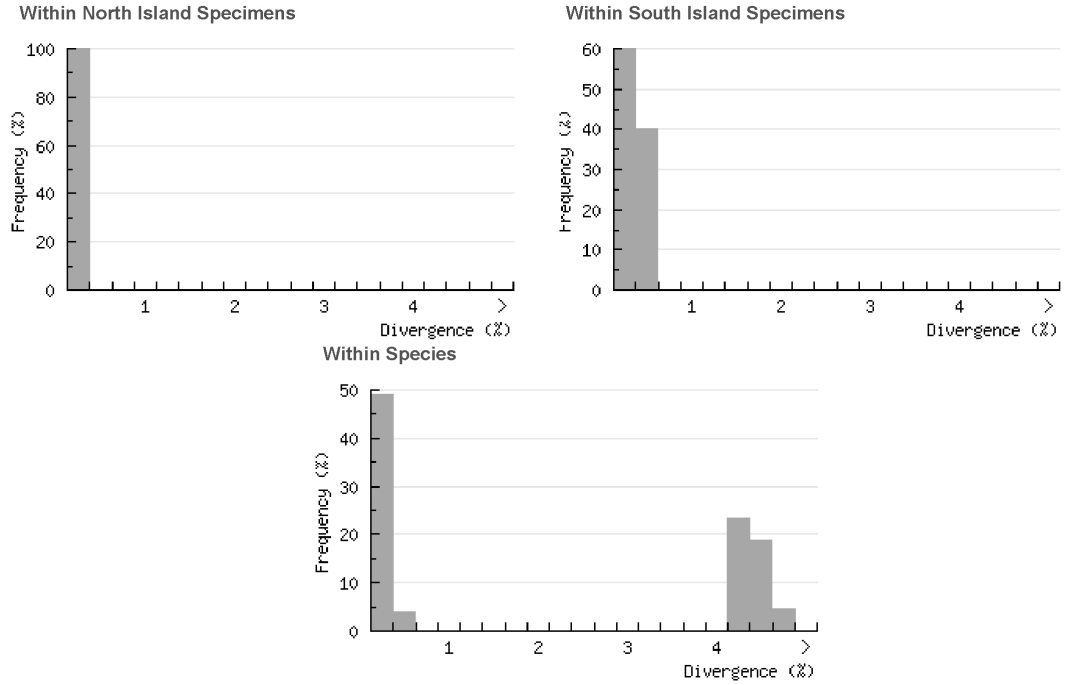


Figure 5.4 Distance summary for *Petroica australis*

5.2.3 Eudyptula minor

Up to six subspecies of *Eudyptula minor* (the little blue penguin) have been recognised in the past, based on morphological characters such as back colour, flipper pattern and measurements as well as distribution (Table 5.1)[28] and the genus was considered to contain two species until 1976 when they were amalgamated.[23, 29]

Baker et al. (2006)[30] reported a genetic split between a white-flippered variant found on the Banks Peninsular and birds from other areas. Two mtDNA clades were identified by Banks et al. (2008)[31] who observed divergence between a group made up of birds from Australia with some from Otago and little blue penguins from the rest of New Zealand. *E. minor* has been the subject of ongoing debate but is currently recognised as a single species without designation of

subspecies in the Checklist of the Birds of New Zealand because “In view of the continuing uncertainty of the taxonomic status of these various populations, including white-flippered birds often classified as *E. albosignata* or *E. minor albosignata*, we have placed all the little blue penguins in one species *E. minor* and not recognised any subspecies.”[24]

Table 5.1 The six subspecies that have been recognised in *Eudyptula minor in the past*.[28] The current Checklist of the Birds of New Zealand[24] does not recognise any subspecies.

Subspecies	Distribution
Eudyptula minor novaehollandiae	Southern Australia and Tasmania
Eudyptula minor iredalei	Northern North Island (NZ)
Eudyptula minor variabilis	Southern North Island and Cook Strait (NZ)
Eudyptula minor albosignata	Eastern South Island (NZ)
Eudyptula minor minor	Western South Island and Stewart Island (NZ)
Eudyptula minor chathamenssis	Chatham Islands

A K2P tree of 60 *E. minor* DNA barcodes showed two very distinct clades based on geographical location (Figures 5.5 and 5.6). Twenty-three out of 25 specimens in one clade came from Australia (9 from Victoria, 14 from Western Australia). The remaining two specimens within the clade had no geographical location data. Of the 35 specimens in the second clade, 16 came from the North Island and 18 came from the South Island with one having no geographical location data.

If one ignores the three specimens for which there are no geographic data (BWA004-06, GBIR0188-06 and GBIR0252-06), the divergence between the Australian and New Zealand specimens is nearly twenty times (19.32) that of the divergence found within the New Zealand specimens and more than thirty times (34.07) that observed within the Australian specimens (Table 5.2).

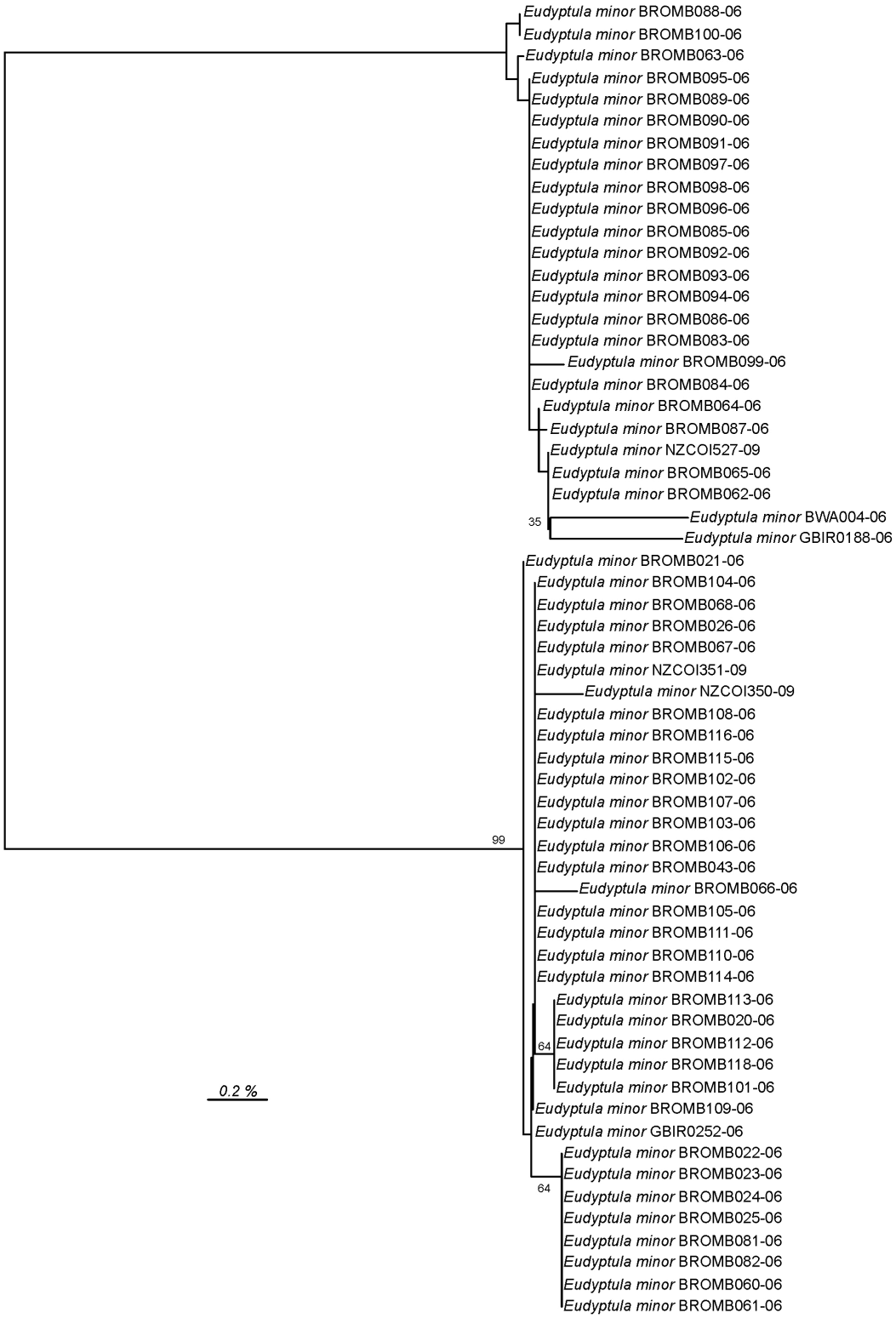


Figure 5.5 ***Eudyptula minor*** K2P neighbour joining tree with bootstrap values for all major branches.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	57	1	1596	0	1.874	4.21	0.046
Within NZ Specimens	34	1	561	0	0.097	0.323	0.004
Within Australian Specimens	23	1	253	0	0.055	0.292	0.005

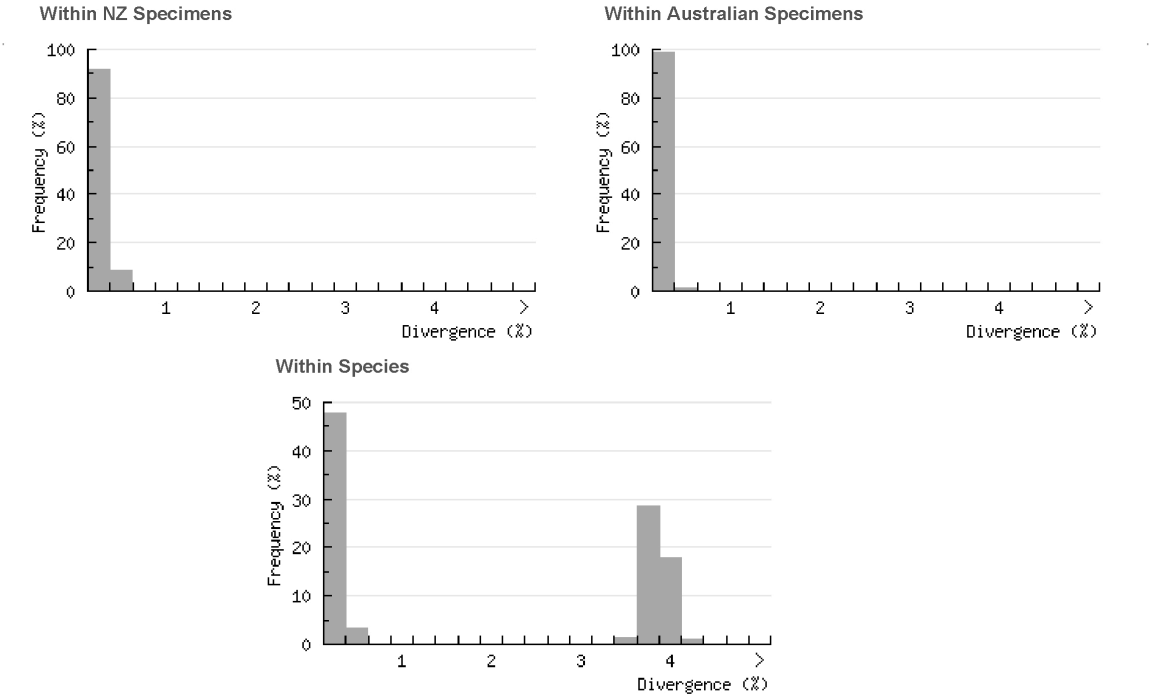


Figure 5.6 Distance summary for *Eudyptula minor*

Analysis of an alignment of New Zealand and Australian birds (Figure 5.7) highlighted 23 diagnostic sites within the barcode region, each of which differed between birds from the two geographic locations but did not differ between birds from the same geographic location. There were only three other nucleotide positions that differed that were not diagnostic and represented within clade variation.

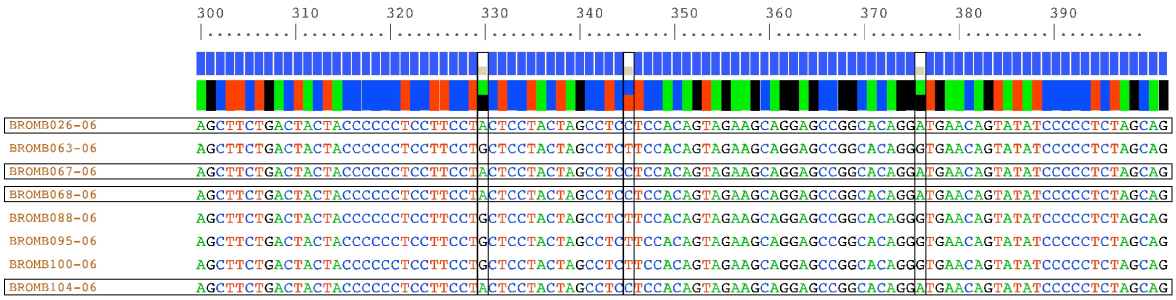


Figure 5.7 Part of an alignment of *Eudyptula minor* specimens from New Zealand (framed with black lines) and Australia (unframed) showing 3 of the 23 sites at which specimens from these two regions differed.

Contrary to the findings of Baker et al. (2006),[30] there was no apparent difference in COI haplotypes between birds (n=10) from the region in which the white-flippered morphotype are found (Banks Peninsular and Canterbury) and those from other regions (Figure 5.5). One sample obtained from Genbank and identified by Baker et al. (2006)[30] as *E. albosignata* (a name sometimes given to the white-flippered morphotype) had no geographic data associated with it and is also not significantly different at the barcode region from the other little blue penguins of New Zealand.

Table 5.2 K2P distance measures for 57 specimens of *Eudyptula minor* from Australasia. Distance measures are also given separately for specimens from New Zealand and from Australia.

	n	Mean distance (%)	Maximum distance (%)	Std error (%)
Within species	57	1.874	4.210	0.046
NZ specimens	34	0.097	0.323	0.004
Aus. specimens	23	0.055	0.292	0.005

5.2.4 Pelagodroma marina

The COI sequence variation within these four specimens (Figures 5.8 and 5.9) is probably due to mislabelling of specimen NZCOI414-09, which differed from the other three specimens of this species at 31 nucleotide sites (4.78%). Such a variation is generally indicative of species level separation.[18]

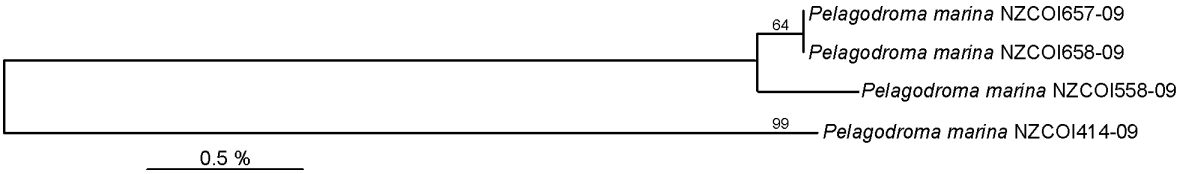


Figure 5.8 ***Pelagodroma marina*** K2P neighbour joining tree with bootstrap values for all major branches.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	4	1	6	0	2.793	5.389	1.015

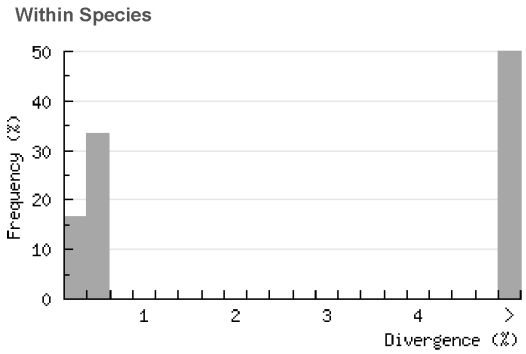


Figure 5.9 Distance summary for *Pelagodroma marina*

5.2.5 Pterodroma axillaris

This species showed high intraspecific K2P variation (Figures 5.10 and 5.11) at the barcode region (mean 0.975%, max. 2.556%). Although this is quite high compared with most other species, it is not extraordinary. What makes this particular case interesting is that *Pterodroma axillaris* is a rare endemic species that is thought to breed only on South East Island and Pitt Island in the Chatham Islands.[23, 24] Numbers are estimated at 500-1000 individuals.[33]

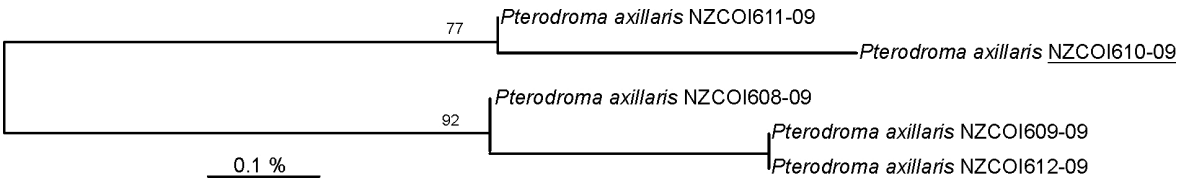


Figure 5.10 ***Pterodroma axillaris*** K2P neighbour joining tree with bootstrap values for all major branches.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	5	1	10	0	0.975	2.556	0.239

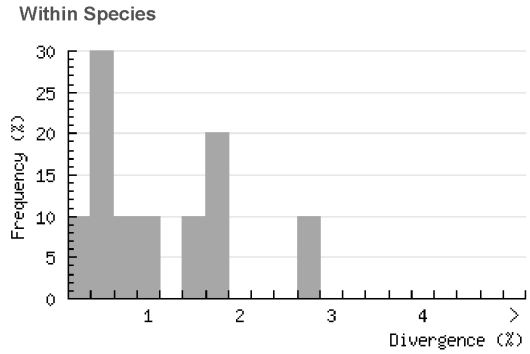


Figure 5.11 Distance summary for *Pterodroma axillaris*

High levels of genetic variation in endangered species from this genus have been previously reported. For example, Lawrence et al (2008)[34] reported similar diversity in the critically endangered magenta petrel, *P. magentae*. They observed twenty-one haplotypes when samples of the mitochondrial cytochrome b gene and both copies of a fragment of the duplicated domain I of the control region from 80 birds were sequenced. This diversity may be related to the relatively long lifespan of these birds (30-40 years[34]), with birds surviving from times past when populations were much larger. Alternatively, there may be some unknown breeding colonies in other parts of the Pacific in which other populations of this species are found. Indeed, *P. axillaris* has been observed as far away as the coast of Peru.[35]

5.2.6 Eudyptes chrysocome

Eudyptes chrysocome is another species that demonstrated high levels of COI nucleotide divergence. This was particularly apparent in the Bayesian tree (Figure 4.27). Mean K2P within species distance was 0.883% with a maximum of 2.701%. When the species was examined more closely (Figure 5.12), four distinct clades were identifiable. However, this species has a wide circumpolar distribution and has breeding colonies on Campbell, Auckland and Antipodes Islands as well as off Cape Horn, and the Falkland Islands.[28] Furthermore, it is known to hybridise with

E. chrosolophus in the Falkland Islands.[36] Thus, relatively wide intraspecific variation might be expected in this species.

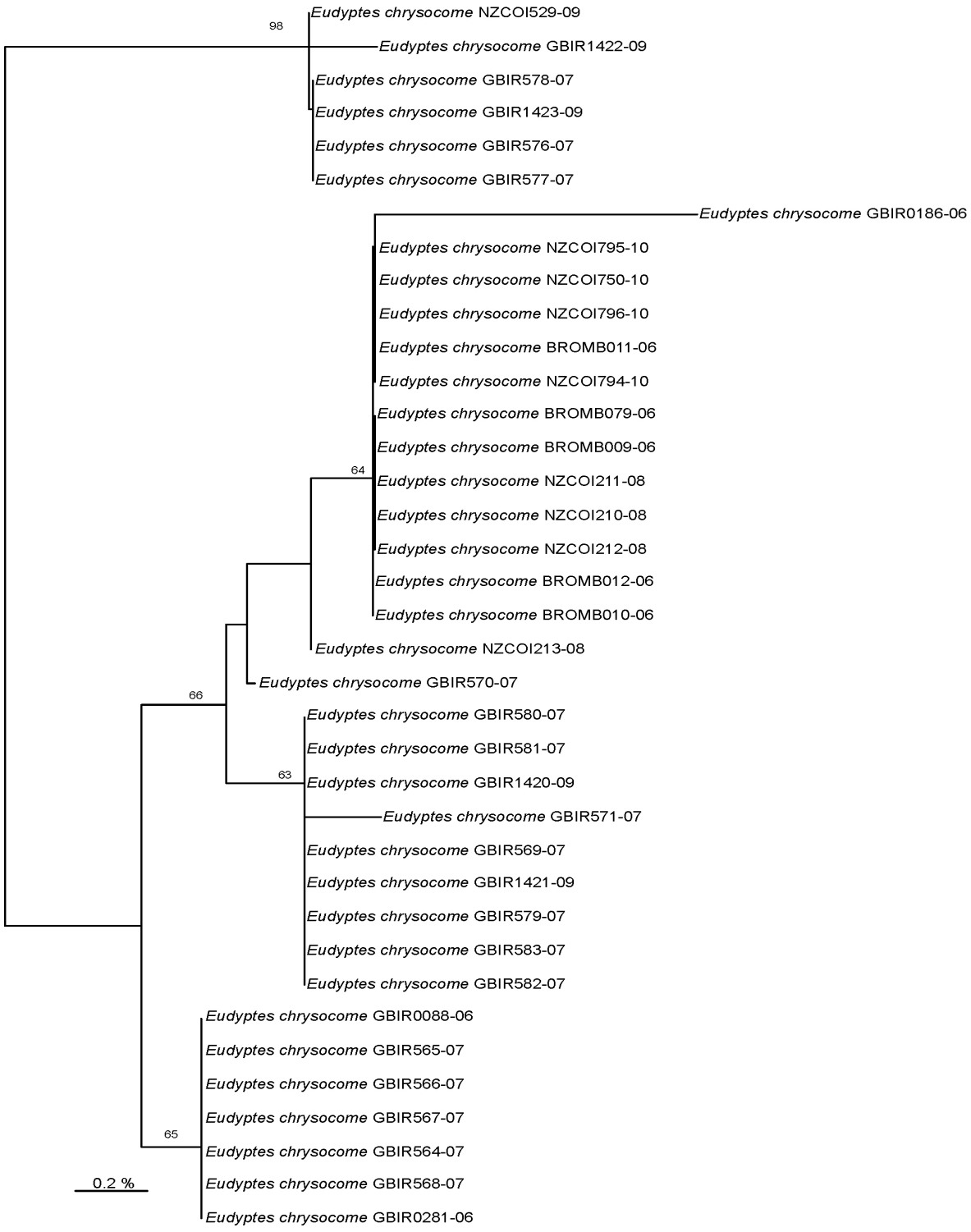


Figure 5.12 *Eudyptes chrysocome* K2P neighbour joining tree with bootstrap values for all major branches.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	37	1	666	0	0.883	2.701	0.025

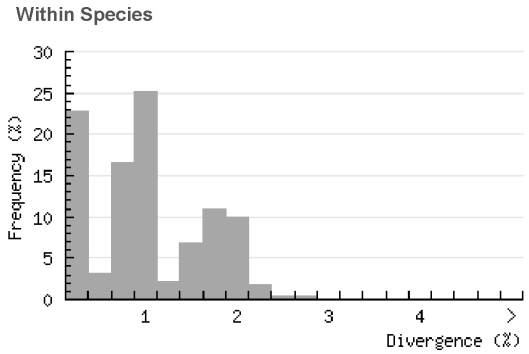


Figure 5.13 Distance summary for *Eudyptes chrysocome*

5.3 Species that were indistinguishable using DNA barcodes

5.3.1 Apteryx species

The recently published Checklist of the Birds of New Zealand[24] "cautiously" recognises a new species of brown kiwi, *Apteryx mantelli* (the North Island brown kiwi) along with the extant *A. rowi* (Okarito brown kiwi) and *A. australis australis* (the South Island brown kiwi) based on Holdaway et al. (2001)[27] and Palma et al. (2003).[37] In addition, a Stewart Island subspecies (*A. australis lawryi*) is recognised.[24] Mitochondrial genetic diversity observed within the brown kiwi[38-40] is not closely associated with morphological diversity making species delimitation difficult.[40]

DNA barcodes of these species generated complicated trees with considerable within species divergence (Figure 5.14), particularly in *A. australis,* and little divergence between *A. mantelli* and *A. australis*. Intraspecific variation in these species (mean 0.762%, max. 6.062%) overlapped that found between species (mean 4.601%, max. 7.954%). There was no distance to nearest neighbour for *A. mantelli, A. ownii* and *A. australis* (Figures 5.15a and 5.16a). However, if three specimens (NZCOI491, NZCOI492 and NZCOI790) with ambiguous sequences and for whom irrefutable provenance was not available were removed from this tree, a clearer picture emerged (Figure 5.15b and 5.16b). Mean intra- versus interspecific divergence (0.356 vs 4.756%) was more than the tenfold difference

suggested as the threshold for species by Hebert et al. 2003.[25] All the species and subspecies formed distinct clades. *A. australis* remained quite variable with three subgroups apparent, one of which is the subspecies *A. australis lawryi.* This subspecies is clearly differentiated by DNA barcoding with the three samples forming their own distinct clade (Figure 5.14).

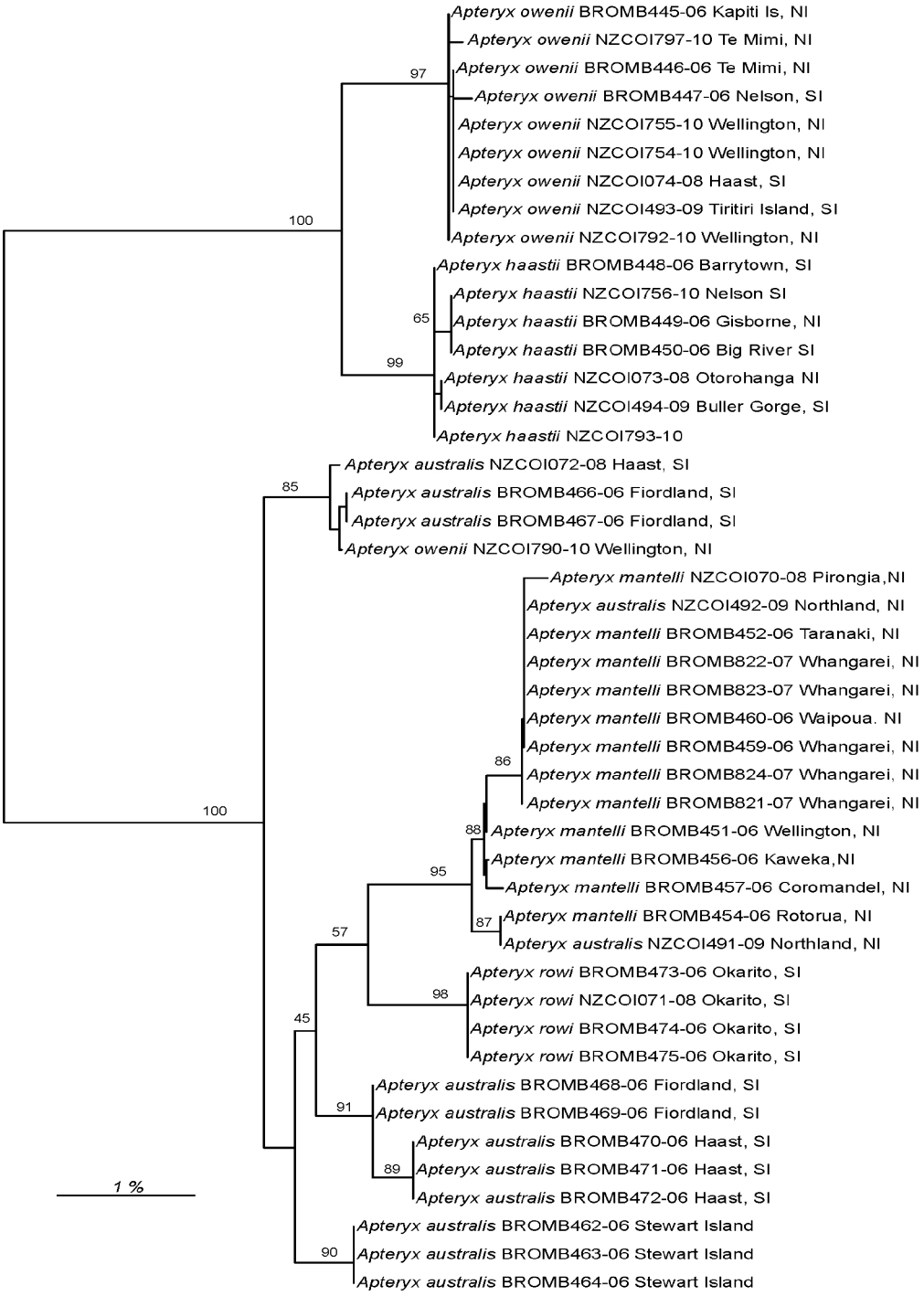
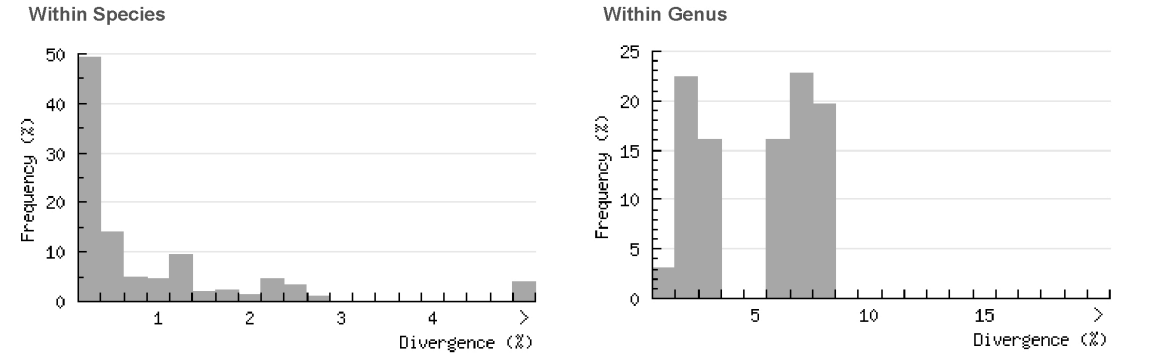


Figure 5.14 New Zealand Apterygiformes K2P neighbour joining tree with bootstrap values for all major branches.

(a)

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	47	5	223	0	0.762	6.062	0.084
Within Genus	47	1	858	0	4.601	7.954	0.084



(b)

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	44	5	191	0	0.356	1.568	0.032
Within Genus	44	1	755	1.214	4.756	7.954	0.086

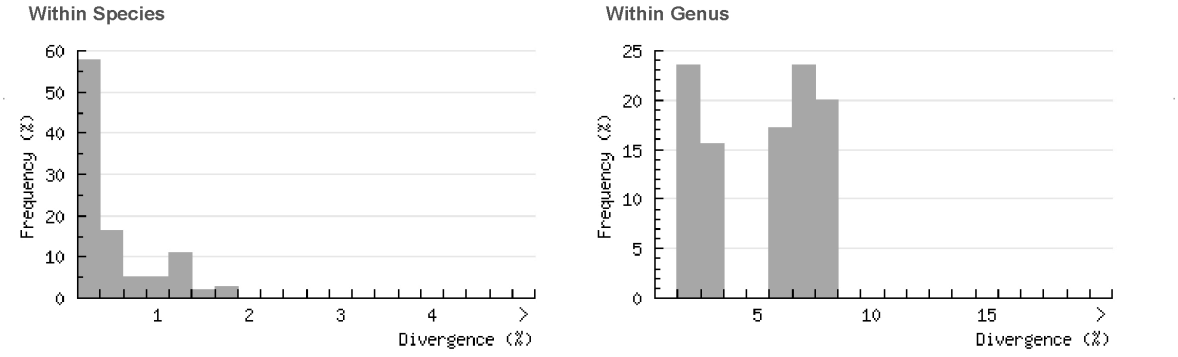
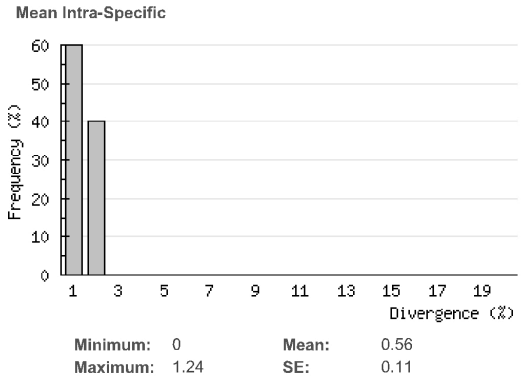
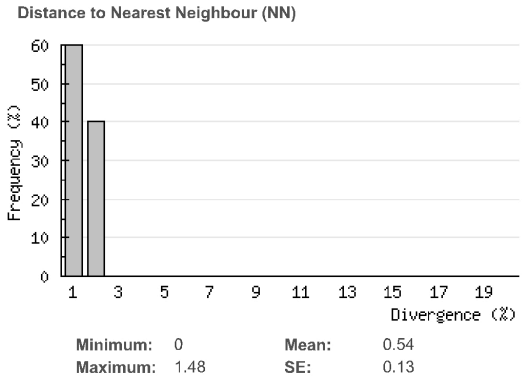


Figure 5.15 Distance summaries for 47 *Apteryx species* before (a) and after (b) removal of three anomalous specimens (NZCOI491, NZCOI492 and NZCOI790).

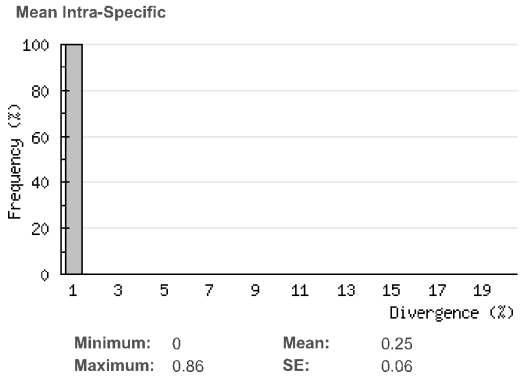
(a)

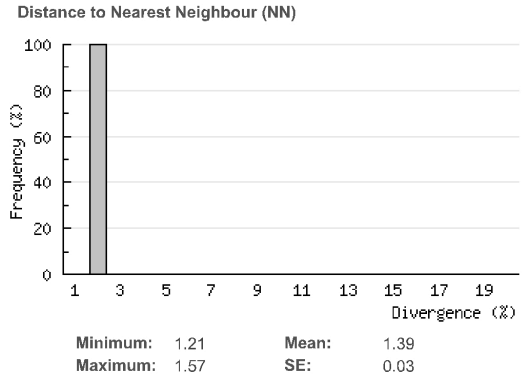




Order	Family	Species	Mean Intra-Sp.	Maximum Intra-Sp.	Nearest Species	Nearest Neighbour	Distance to NN
Struthioniformes	Apterygidae	*Apteryx australis*	1.24	2.54	*Apteryx mantelli*	BROMB454-06	0
Struthioniformes	Apterygidae	*Apteryx haastii*	0.09	0.17	*Apteryx owenii*	BROMB445-06	1.21
Struthioniformes	Apterygidae	*Apteryx mantelli*	0.24	0.87	*Apteryx australis*	NZCOI491-09	0
Struthioniformes	Apterygidae	*Apteryx owenii*	1.22	6.06	*Apteryx australis*	BROMB467-06	0
Struthioniformes	Apterygidae	*Apteryx rowi*	0	0	*Apteryx mantelli*	BROMB451-06	1.48

(b)





Order	Family	Species	Mean Intra-Sp.	Maximum Intra-Sp.	Nearest Species	Nearest Neighbour	Distance to NN
Struthioniformes	Apterygidae	*Apteryx australis*	0.86	1.57	*Apteryx mantelli*	BROMB451-06	1.57
Struthioniformes	Apterygidae	*Apteryx haastii*	0.09	0.17	*Apteryx owenii*	BROMB445-06	1.21
Struthioniformes	Apterygidae	*Apteryx mantelli*	0.24	0.87	*Apteryx rowi*	BROMB473-06	1.48
Struthioniformes	Apterygidae	*Apteryx owenii*	0.07	0.29	*Apteryx haastii*	NZCOI073-08	1.21
Struthioniformes	Apterygidae	*Apteryx rowi*	0	0	*Apteryx mantelli*	BROMB451-06	1.48

Figure 5.16 Nearest neighbour analyses for *Apteryx species* before (a) and after (b) removal of three anomalous specimens (NZCOI491, NZCOI492 and NZCOI790).

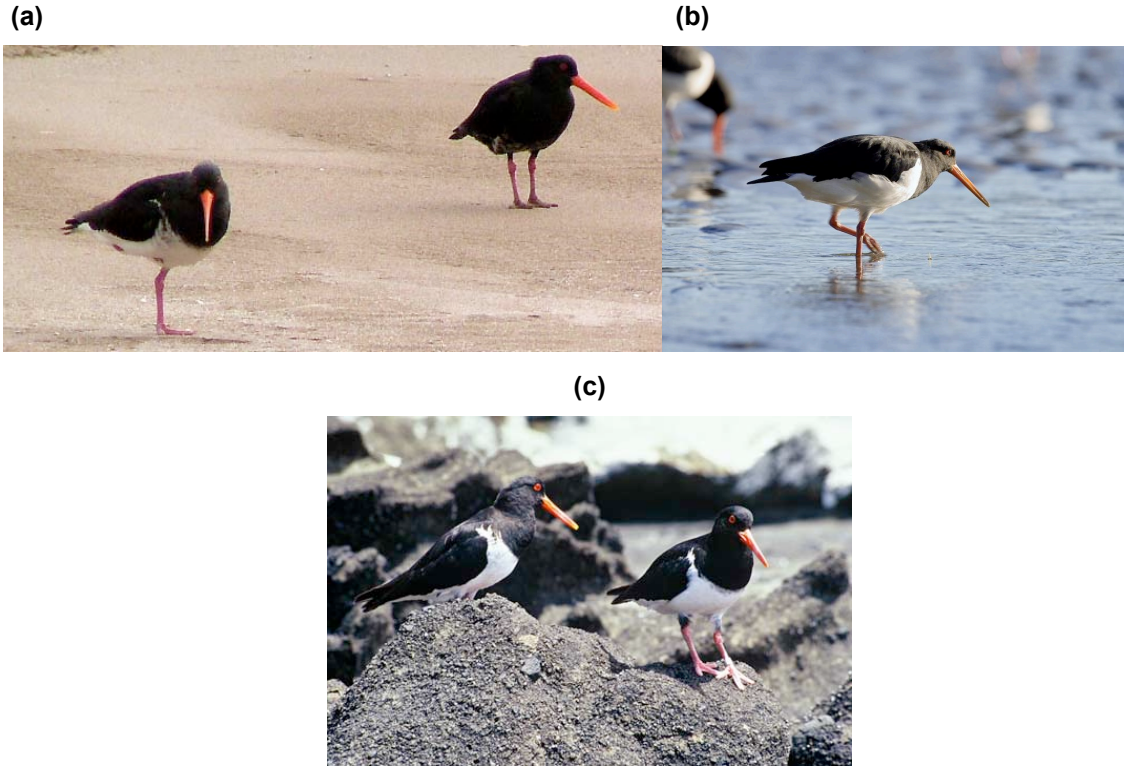
5.3.2 Haematopus species

There are three endemic species of the genus *Haematopus* recognised in the Checklist of New Zealand Birds*, H. unicolor (Figure 5.17a), H. finschi* (Figure 5.17b) and *H. chathamensis (Figure 5.17c)*.[24] *H. unicolor* is found around the coastline of both North and South Islands as well as Stewart Island.[41] *H. finschi* is, likewise, found around the coasts of all three main islands of New Zealand but generally breeds in inland areas to the east of the Southern Alps in the South Island.[28, 41] Both species are known to have annual migration patterns that take them all around New Zealand.[24, 28] *H. chathamensis* is found only in the Chatham Islands.[24, 28, 41] *H. fischi* is pied as is *H. chathamensis* but *H. unicolor* has a plumage that varies from black to pied - more black as latitude increases (northern 43% black, central 85% black and southern 94% black).[24]

According to the Checklist of New Zealand Birds[24] *H. chathamensis* (Figure 5.17c) is supported as a separate species by a recent study of mitochondrial DNA by Banks and Patterson (2007).[22] However, in fact the latter authors found little divergence in mtDNA between the New Zealand species despite testing a variety of mitochondrial markers. These included the hypervariable region 1 of the control region (262 nucleotides), the gene coding for the 12S ribosomal subunit (435 nucleotides) as well as the COI barcoding region (647 nucleotides) and the cytochrome *b* gene (389 nucleotides). Out of this total of 1733 nucleotides only nine differed (0.5%) within the three New Zealand species.

Banks and Paterson (2007)[22] examined samples from four birds, one from each of the New Zealand species and one European oystercatcher (*H. ostralegus ostralegus*). *H. finschi* was previously recognised as a subspecies of the European oystercatcher, *H. ostralegus finschi*. The European specimen diverged from the *H. chathamensis* by 21 nucleotides or 1.2%* within these four mtDNA regions and from the other New Zealand specimens by 12 nucleotides (0.7%).[22]

* This percentage is incorrectly given as 2.1% in Banks and Paterson (2007).

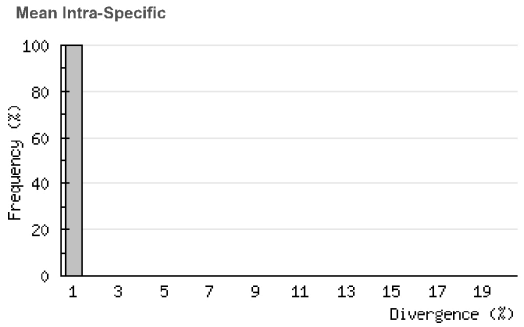


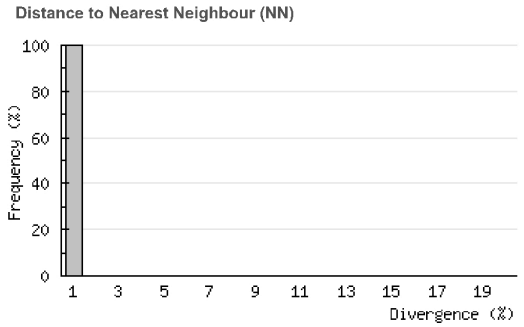
Photos TeAra.govt.nz

Figure 5.17 *Haematopus finschi* (a), *H. unicolor* (b) and *H. chathamensis* (c)

DNA barcoding failed to differentiate 125 New Zealand individuals from the genus *Haematopus* (Figure 5.18). *H. chathamensis* was the only species that formed a distinct clade but mean intraspecific variation for the three specimens of this species was higher than the distance to nearest neighbour (Figure 5.19: 0.18% *vs* 0.14%). There was very little intraspecific COI haplotypic variation for the three taxonomically recognised species (Figure 5.20: mean 0.041%, max. 0.341%) and the same was true for interspecific divergence (mean 0.109%, max. 0.437%). *H. unicolor* and *H. finschi* had identical haplotypes.

Figure 5.18. **New Zealand members of the genus *Haematopus*** K2P neighbour joining tree with bootstrap values for all major branches.

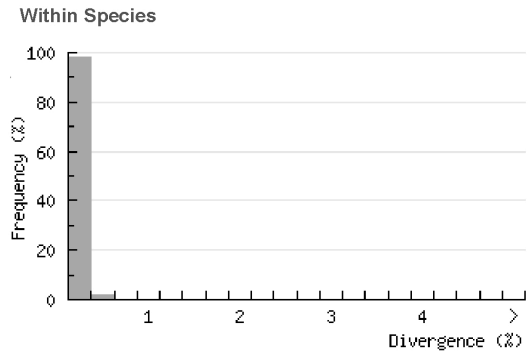




Order	Family	Species	Mean Intra-Sp.	Maximum Intra-Sp.	Nearest Species	Nearest Neighbour	Distance to NN
Charadriiformes	Haematopodidae	*Haematopus chathamensis*	0.18	0.26	*Haematopus finschi*	ROMC407-07	0.14
Charadriiformes	Haematopodidae	*Haematopus finschi*	0.06	0.27	*Haematopus unicolor*	ROMC467-07	0
Charadriiformes	Haematopodidae	*Haematopus unicolor*	0.04	0.34	*Haematopus finschi*	ROMC431-07	0

Figure 5.19 Nearest neighbour analysis for New Zealand *Haematopus* species

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	125	3	6264	0	0.041	0.341	0.001
Within Genus	125	1	1486	0	0.109	0.437	0.003



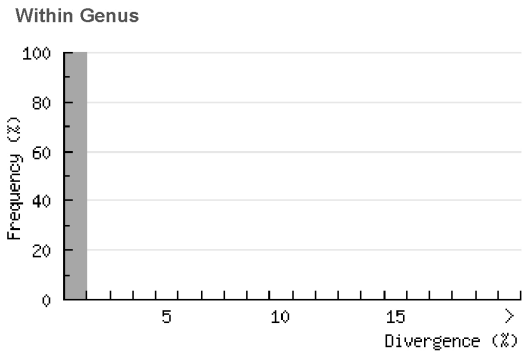


Figure 4.20 Distance summary for New Zealand *Haematopus* species

When COI sequences from each of these species were analysed using a linear search to collect nearest neighbours from a global alignment of all reference sequences,[42] *H. unicolor* matched 100% with *H. finschi*, 99.84% with *H. chathamensis* and 99.84% with *H. longirostrus* (an oystercatcher from Tasmania, New Guinea and elsewhere). A 99.84% match would generally be sufficient to identify species, allowing for some intraspecific variation (Figure 5.21). It must be noted that these three species are similar in appearance, regularly fly long

distances and morphologically similar members of this genus are distributed widely around the world. That there should be three endemic species in New Zealand seems surprising in a bird so capable of dispersal. DNA barcode evidence here combines with morphological similarity, dispersal ability and commonality of range to suggest that these species may be worthy of closer analysis.

BOLD Systems – Identification Result 31/08/10 4:17 PM

BOLDSYSTEMS v2.5 | Management & Analysis

PRINT | HELP

Specimen Identification Request

Search Request:
Type : COI FULL LENGTH (at least 640bp) DATABASE

Search Result:
Identification Summary :

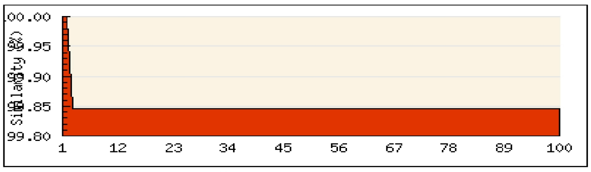
Taxonomic Level	Taxon Assignment	Probability of Placement (%)
phylum	Chordata	100
class	Aves	100
order	Charadriiformes	100
family	Haematopodidae	100
genus	Haematopus	100
species	Haematopus unicolor	100

A species level match could not be made, the queried specimen is likely to be one of the following :

-Haematopus unicolor
-Haematopus finschi
-Haematopus chathamensis
-Haematopus longirostris

Tree Based Identification

Distance Summary :



Similarity scores of the top 100 matches

TOP 20 Matches :

Phylum	Class	Order	Family	Genus	Species	Specimen Similarity (%)
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	100
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*finschi*	100
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*chathamensis*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*chathamensis*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*chathamensis*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*longirostris*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*longirostris*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*longirostris*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*longirostris*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*longirostris*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84
Chordata	Aves	Charadriiformes	Haematopodidae	Haematopus	*unicolor*	99.84

Figure 5.21 The results of a linear search to collect nearest neighbours from a global alignment of all reference sequences.[42] The search was based on a 648bp segment of the COI gene from a specimen of *Haematopus unicolor* (sequence page reference NZCOI499-09). No identification could be made because this sequence was identical to one from H. finschi and very similar to others from *H. chathamensis* and *H. longirostris*. More than 100 *Haematopus* specimens in the database matched this sequence for at least 99.84% of their nucleotides in the barcode region.

5.3.3 Thalassarche melanophris and T. chrysostoma

Both the K2P and Bayesian trees grouped these two species closely together (figure 4.22, 4.23 and 5.22). However, distance analyses indicated two distinct clades that were not highly divergent (Figure 5.23: mean intra- *vs* interspecific K2P distance 0.181 *vs* 0.811%). Distance to nearest neighbour for these species was 0.47% (Figure 5.24). These two smaller albatross species share a common circumpolar range and both are known to breed on Campbell Island between August and October. They nest in large close-knit colonies that intermingle between species.[23, 43] Thus, some hybridisation may occur. Conversly, they may be recently diverged lineages.

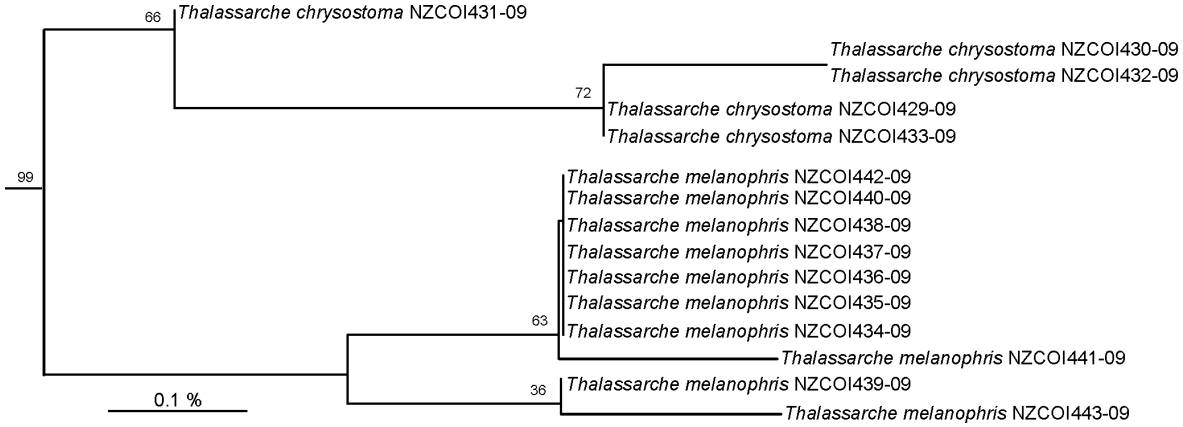


Figure 4.22 *Thalassarche chrysostoma* and *T.melanophrys* K2P neighbour joining tree with bootstrap values for all major branches.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	15	2	55	0	0.181	0.623	0.025
Within Genus	15	1	50	0.467	0.811	1.096	0.026

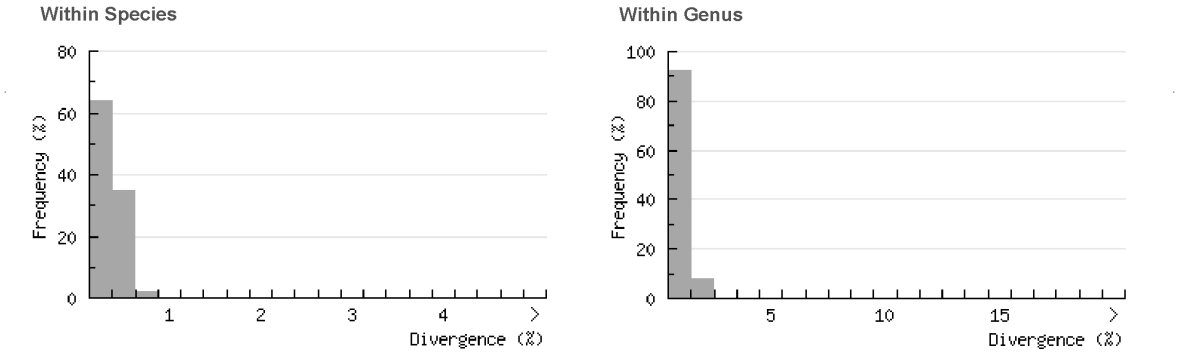
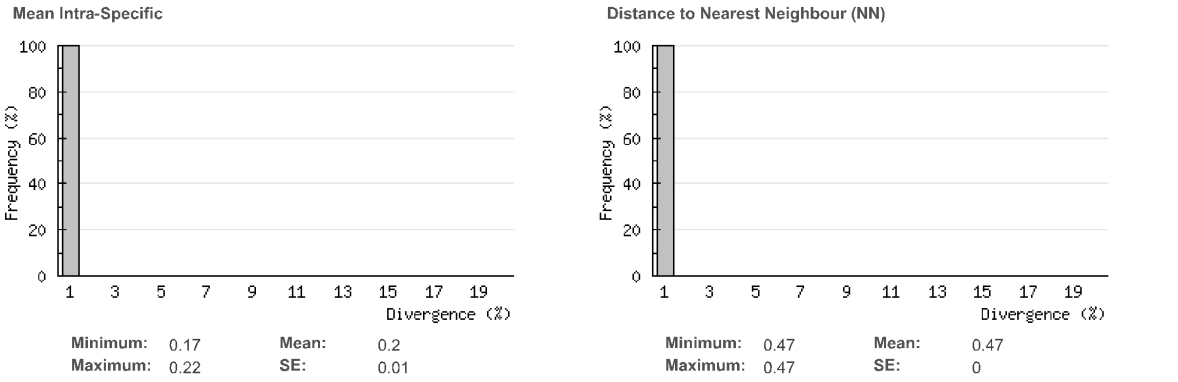


Figure 5.23 Distance summary for *Thalassarche chrysostoma* and *T.melanophrys*



Order	Family	Species	Mean Intra-Sp.	Maximum Intra-Sp.	Nearest Species	Nearest Neighbour	Distance to NN
Procellariiformes	Diomedeidae	*Thalassarche chrysostoma*	0.22	0.47	*Thalassarche melanophris*	NZCOI434-09	0.47
Procellariiformes	Diomedeidae	*Thalassarche melanophris*	0.17	0.62	*Thalassarche chrysostoma*	NZCOI431-09	0.47

Figure 5.24 Nearest neighbour analysis for *Thalassarche chrysostoma* and *T.melanophrys*

5.3.4 Larus species

(a) **(b)**

Photos TeAra.govt.nz

Figure 5.25 *Larus novaehollandiae* (a) and *L. bulleri* (b)

Hybridisation occurs frequently in the genus *Larus* complicating species identification.[44] There are three recognised species in New Zealand, *L. dominicanus* (the Southern Black-backed Gull), *L. novaehollandiae* (the Red-billed Gull) and *L. bulleri* (the Black-billed Gull).[24] Morphologically, *L. dominicanus* is quite distinct from the other two, being approximately three times heavier and with a wingspan of 60cm compared with 37cm for the other two.[23] *L. novaehollandiae* (Figure 5.25a) is found from Africa through to the Chatham Islands, whereas *L. bulleri* (Figure 5.25b) is endemic to New Zealand. Both smaller birds are widely distributed around New Zealand. [23, 24, 28]

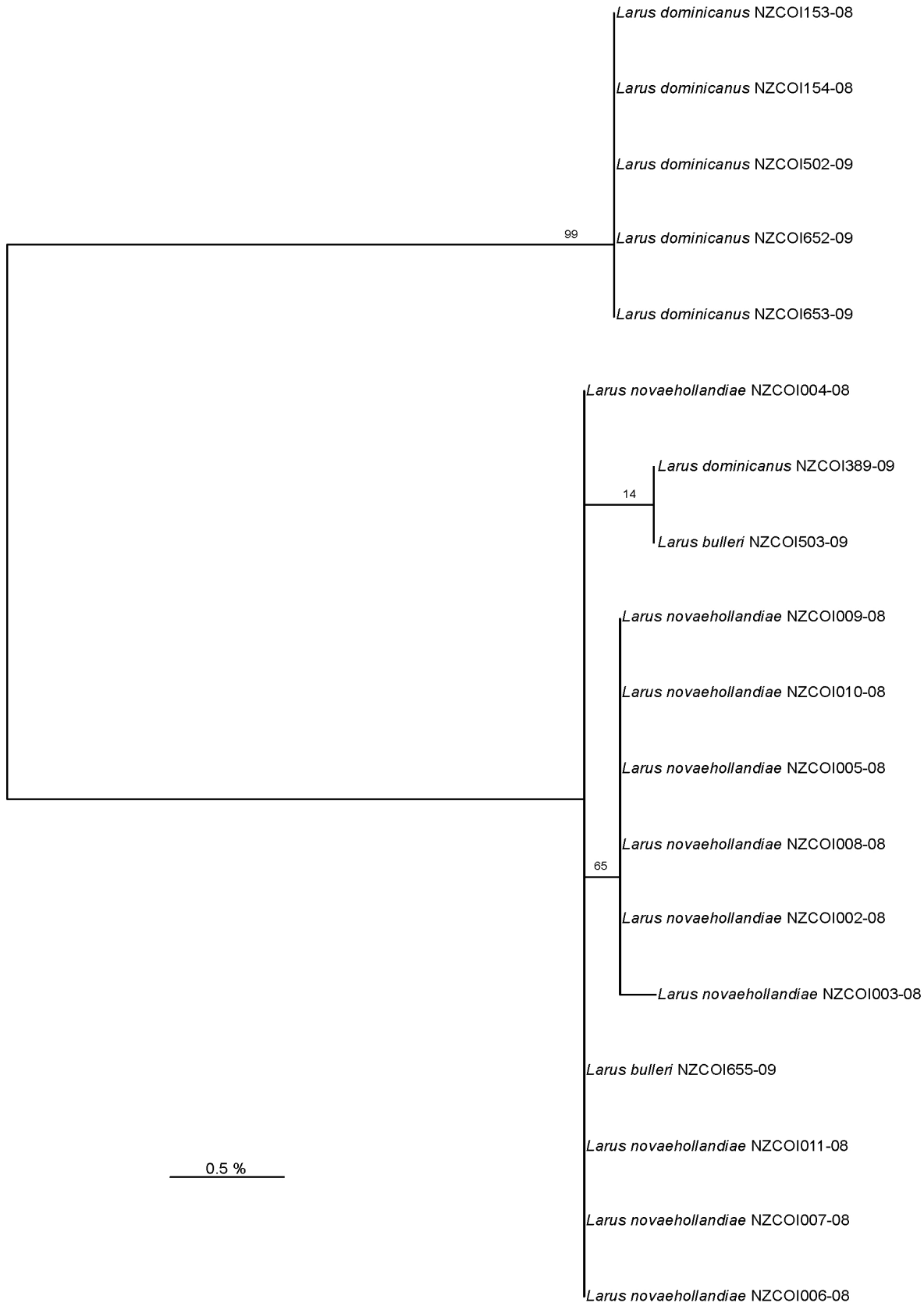


Figure 5.26 New Zealand *Larus* species K2P neighbour joining tree with bootstrap values for all major branches.

Little COI divergence was found between these two smaller species (Figure 5.26). Mean intraspecific variation was 0.143% and the interspecific divergence was 0.33% (Figure 5.27). Nearest neighbour analysis indicated 0 distance between these two species (Figure 4.28). This finding was supported is by another study[45] using maximum parsimony, maximum likelihood and Bayesian analysis on segments of mtDNA including partial cytochrome *b* and control region, which found little distance between *L. bulleri* and *L. novaeholandiae*.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	17	2	66	0	0.143	0.513	0.018
Within Genus	17	1	70	0	0.33	0.668	0.022

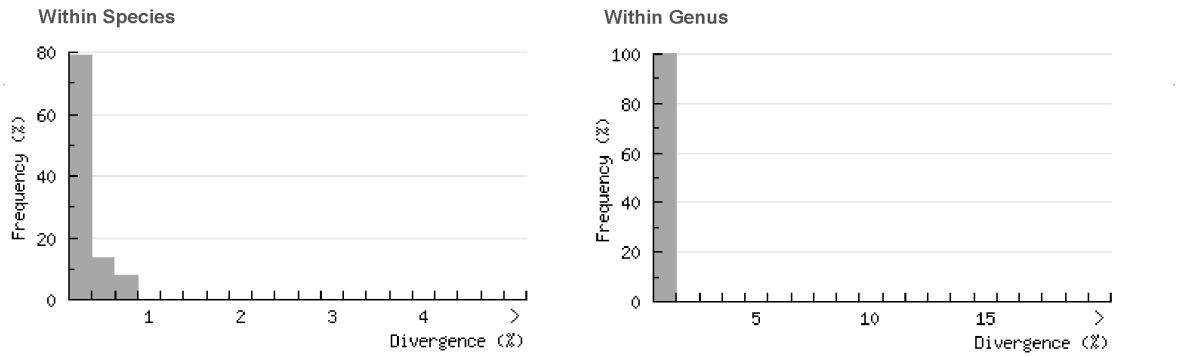
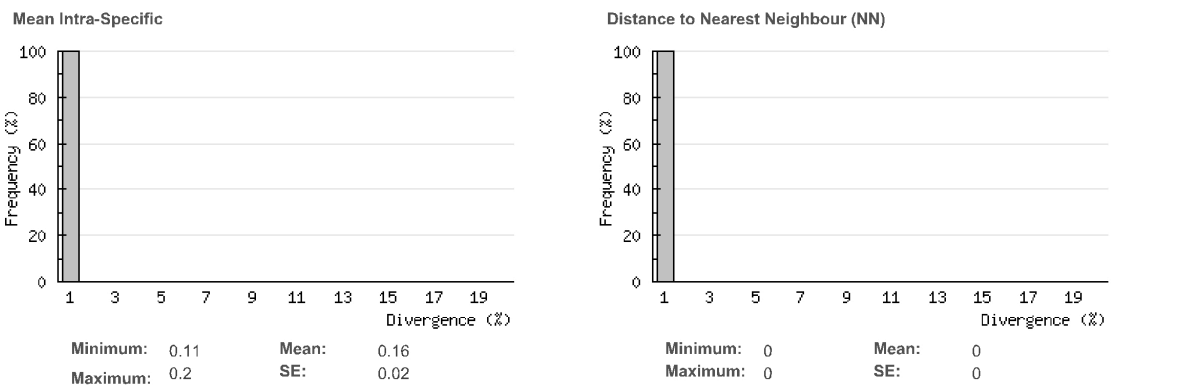


Figure 5.27 Distance summary for *Larus novaehollandiae* and *L. bulleri*



Order	Family	Species	Mean Intra-Sp.	Maximum Intra-Sp.	Nearest Species	Nearest Neighbour	Distance to NN
Charadriiformes	Laridae	*Larus bulleri*	0.2	0.51	*Larus novaehollandiae*	NZCOI005-08	0
Charadriiformes	Laridae	*Larus novaehollandiae*	0.11	0.31	*Larus bulleri*	BROM552-07	0

Figure 5.28 Nearest neighbour analysis for *Larus novaehollandiae* and *L. bulleri*

Although *L. dominicanus* is the only species of its type within New Zealand and, therefore, easy to identify in this region, there are other species (e.g. *L. fuscus*) from which it cannot be differentiated using COI haplotypes. *L. dominicanus* is found throughout the Southern Hemisphere including the west coast of Africa to Namibia. *L. fuscus,* which is morphologically very similar, is found from Iceland through the Atlantic coast of Europe to the west coast of Africa. Thus, opportunities for hybridisation may certainly have occurred in the recent past. Indeed these may be one biological species. This kind of taxonomic confusion is common among the Laridae and may be associated with other phenomena such as the ring species of the Herring Gull (*L. argentatus*), with which *L. fuscus* forms a complex.[45]

5.3.5 Puffinus assimilis, P. gavia and P. griseus

Both Puffinus assimilus and P. gavia are relatively small shearwaters (200-300g). *P. griseus* is an altogether larger bird (800g) and quite easily distinguishable from the other two. However the tree produced for these species appeared confusing as were the distance summary and nearest neighbour analysis (Figures 5.29, 5.30a and 5.31a). Interestingly a number of these samples and notably the ones that appear ambiguous from each species were collected from a single site at Opoutere on the Coromandel Peninsular in New Zealand. Once again, the removal of these samples (NZCOI416-09, NZCOI791-10, NZCOI461-09 and NZCOI 464-09) produced a much more coherent tree. In particular, the distance to nearest neighbour increased from 0% for all three species to 8.49% for *P. assimilis* and 4.19% for the other two species (Figures 5.30b and 5.31b). It seems highly likely that samples from this site have been mislabelled or somehow mixed up.

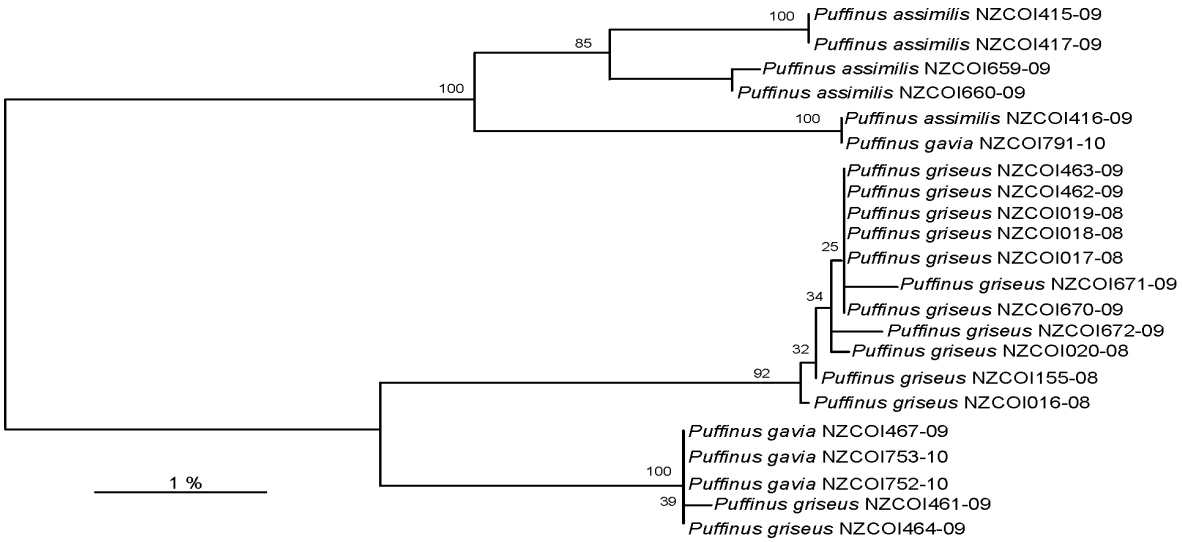
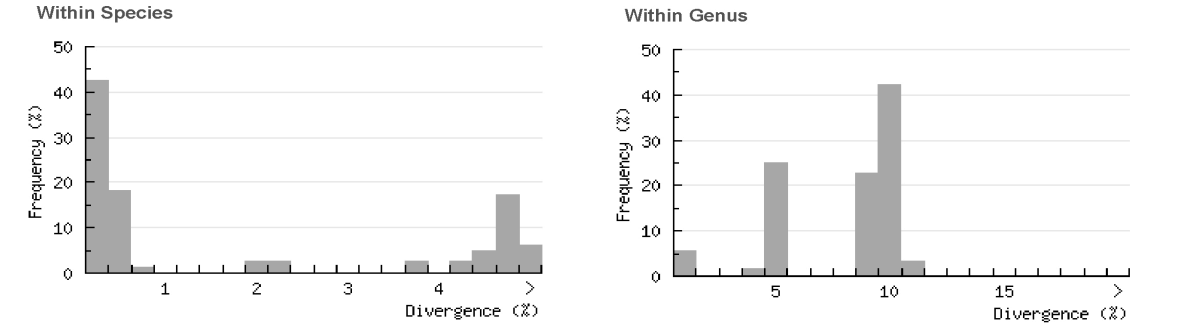


Figure 5.29 ***Puffinus assimilis, P. gavia*** **and** ***P. griseus*** K2P neighbour joining tree with bootstrap values for all major branches.

(a)

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	21	3	82	0	1.83	9.003	0.263
Within Genus	21	1	128	0	7.501	10.179	0.25



(b)

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	18	3	64	0	0.42	4.203	0.102
Within Genus	18	1	89	4.19	7.6	10.179	0.255

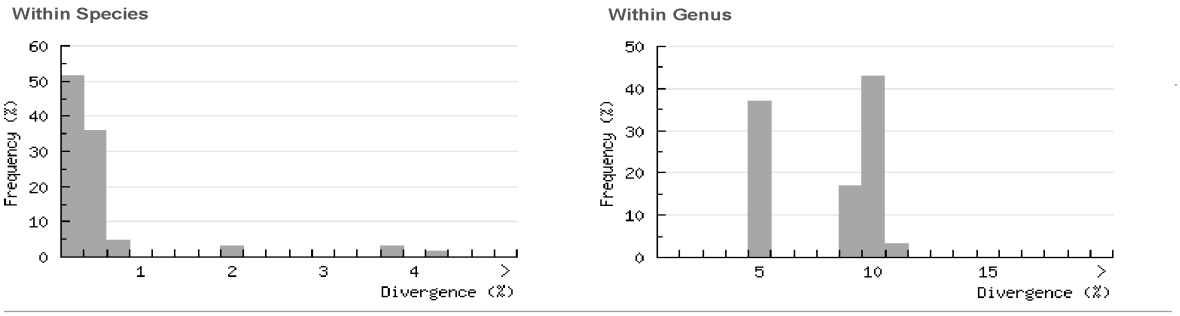
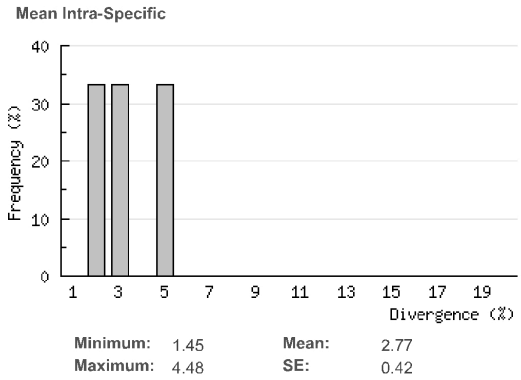
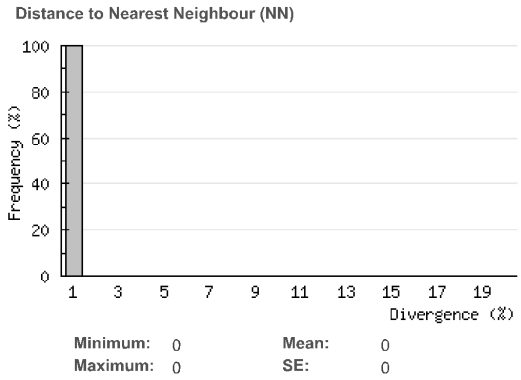


Figure 5.30 Distance summary for *Puffinus assimilus*, *P. gavia* and *P. griseus* before (a) and after (b) removal of four ambiguous specimens (NZCOI416-09, NZCOI791-10, NZCOI461-09 and NZCOI 464-09).

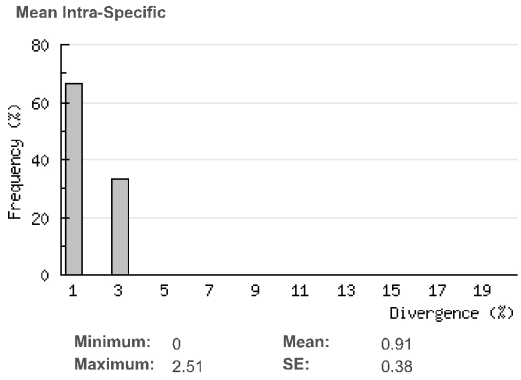
(a)

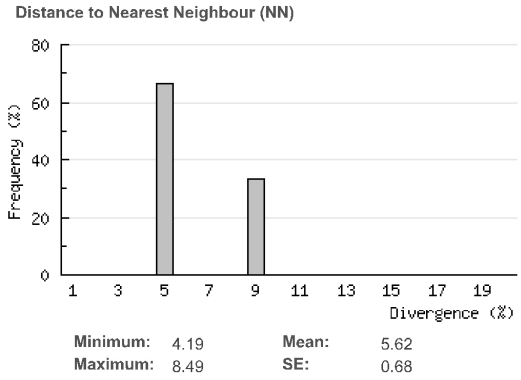




Order	Family	Species	Mean Intra-Sp.	Maximum Intra-Sp.	Nearest Species	Nearest Neighbour	Distance to NN
Procellariiformes	Procellariidae	*Puffinus assimilis*	2.37	4.37	*Puffinus gavia*	NZCOI791-10	0
Procellariiformes	Procellariidae	*Puffinus gavia*	4.48	9	*Puffinus griseus*	NZCOI464-09	0
Procellariiformes	Procellariidae	*Puffinus griseus*	1.45	5.04	*Puffinus gavia*	NZCOI752-10	0

(b)





Order	Family	Species	Mean Intra-Sp.	Maximum Intra-Sp.	Nearest Species	Nearest Neighbour	Distance to NN
Procellariiformes	Procellariidae	*Puffinus assimilis*	2.51	4.2	*Puffinus gavia*	NZCOI752-10	8.49
Procellariiformes	Procellariidae	*Puffinus gavia*	0	0	*Puffinus griseus*	NZCOI016-08	4.19
Procellariiformes	Procellariidae	*Puffinus griseus*	0.22	0.62	*Puffinus gavia*	NZCOI752-10	4.19

Figure 5.31 Nearest neighbour analysis for *Puffinus assimilus*, *P. gavia* and *P. griseus* before (a) and after (b) removal of four ambiguous specimens (NZCOI416-09, NZCOI791-10, NZCOI461-09 and NZCOI 464-09)

5.3.6 Cyanorhamphus unicolor and *C. navaezelandiae*

It appears that misidentification may also be the cause of some of the lack of sequence divergence for these apparently two species. Four specimens described as *Cyanorhamphus novaezelandiae* came from Antipodes Island (Figure 5.32), home to *C. unicolor* and what was formerly regarded as a subspecies of *C. novaezelandiae, C. novaezelandiae hochstetteri* but which is now recognised as a

species *C. hochstetteri.*[23, 24, 46] The three remaining birds came from Little Barrier Island or Tiritiri Island in the Hauraki Gulf. Both these islands are used as relocation sites for endangered bird species. For this reason there is little doubt over their species identification and they are probably *C. novaezelandiae novaezelandiae* as identified.[47] The divergence between them and *C. hochstetteri* is very small (Figure 5.33: mean 0.143%, max. 0.311%) but may possibly be enough to support a subspecies status. The *C. unicolor* clade is quite distinct but the mean distance between it and *C. n. novaezelandiae* is 0.367%, considerably less than is normally associated with a separate species. Interestingly, *C. unicolor* and *C. hochstetteri*, both from Antipodes Island, have the greatest divergence (mean 0.631%, max. 0.645%).

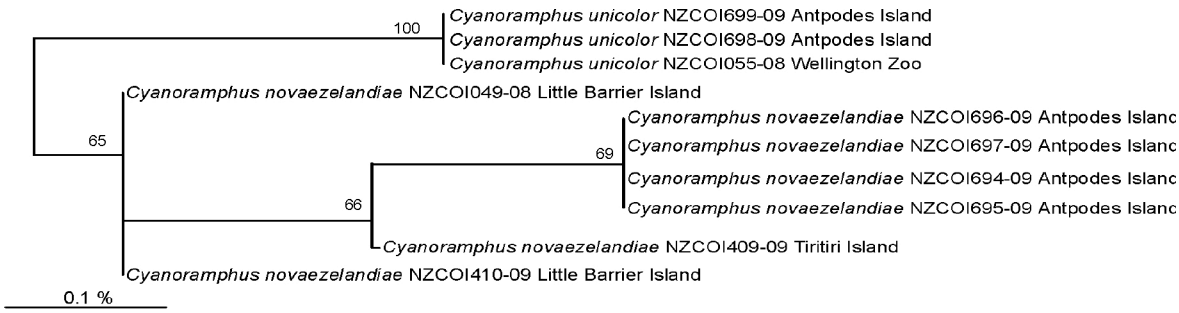


Figure 5.32 *Cyanorhamphus* species K2P neighbour joining tree with bootstrap values for all major branches.

	n	Taxa	Comparisons	Min Dist(%)	Mean Dist(%)	Max Dist(%)	SE Dist(%)
Within Species	10	2	24	0	0.143	0.311	0.027
Within Genus	10	1	21	0.311	0.518	0.645	0.03

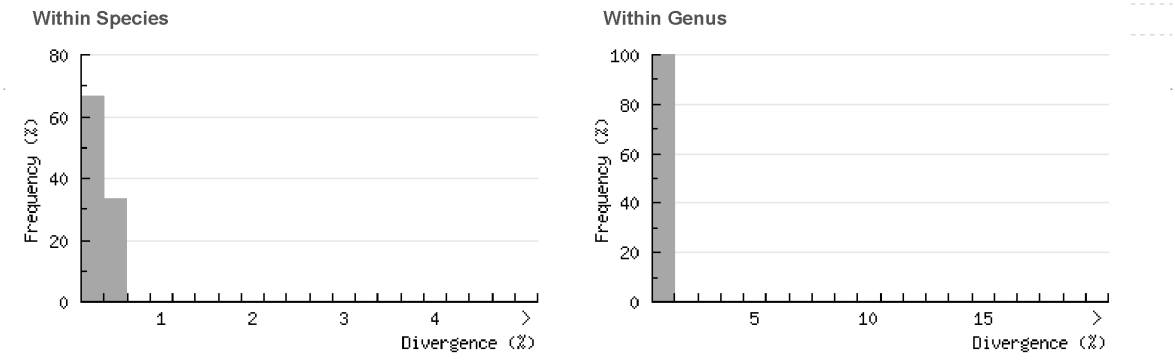
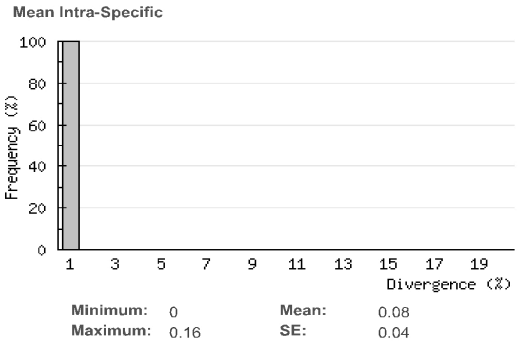
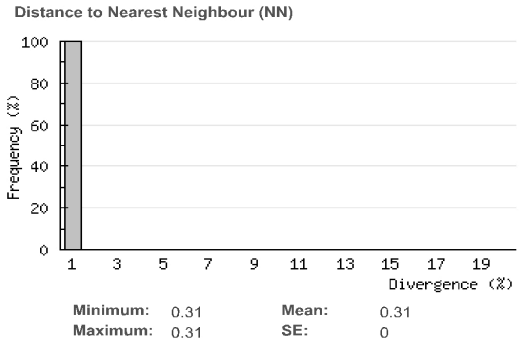


Figure 5.33 Distance summary for *Cyanorhramphus novaezelandiae* and *C. unicolor*





Order	Family	Species	Mean Intra-Sp.	Maximum Intra-Sp.	Nearest Species	Nearest Neighbour	Distance to NN
Psittaciformes	Psittacidae	*Cyanoramphus novaezelandiae*	0.16	0.31	*Cyanoramphus unicolor*	NZCOI698-09	0.31
Psittaciformes	Psittacidae	*Cyanoramphus unicolor*	0	0	*Cyanoramphus novaezelandiae*	NZCOI049-08	0.31

Figure 5.34 Nearest neighbour analysis for *Cyanorhramphus novaezelandiae* and *C. unicolor*

5.4 Discussion

DNA barcoding alone does not fully characterise species; however, in the majority of cases it accurately identifies New Zealand bird species. The delimitation of these species is the role of taxonomists and in many cases species boundaries are the subject of some debate. Molecular techniques have revolutionised this branch of science and analysis of DNA barcodes can provide a useful input to these debates. In a number of cases within these data, DNA barcode clusters that correlate with geographic separation have been observed. In addition, a number of groups of species showed little nucleotide divergence at the DNA barcode region.

DNA barcodes supported sub-species status in North and South Island rifleman populations. They supported the recent recognition of the North Island robin as a species distinct from South Island robins and they have identified Australian and New Zealand members of the genus *Eudyptula* as worthy of further examination to determine whether they should be regarded as separate species. In each of these cases, the evidence of COI divergence was correlated with geographic separation of the populations involved as well as by other evidence presented in the debates relating to their status.[24, 27, 31]

New Zealand representatives of the genus *Haematopus*, particularly those from the mainland were indistinguishable using DNA barcodes. Even *H. chathamenisis* varied at only one nucleotide site from the other two species. Given other factors such as morphological similarity, dispersal ability and overlapping ranges, this group might benefit from a re-evaluation in the light of this and other[22] molecular evidence.

Conversely, unexpectedly high nucleotide divergence was observed within members of the endangered species *Pterodroma axillaris*. This may be a reflection of the long lifespan these kinds of birds are known for;[34] the survivors of the recent decline in population having greater variation at this site than would normally be expected in so small a population. On the other hand, it may be an indication of unknown breeding sites and, perhaps, larger numbers than are currently identified.[34] Either way, it is important information for conservation biologists. This kind of nucleotide divergence is to be expected in widely distributed species, as was the case with *Eudyptes chrysocome*, a penguin with a wide circumpolar distribution[36] and *P. axillaris* is also known to have a large range.[35]

The impact of incorrectly identified specimens on any database is considerable. With time and as the database for each species grows, errors in species identification become more obvious. However, until that time, such errors have the power to cause much confusion.

Several cases in point were identified during the analysis of these data. Within the Apterygiformes of New Zealand, three specimens were probably incorrectly identified, with the result that nearest neighbour trees and distance data were confused and the power of this system substantially undermined. Removal of these three samples from an overall dataset of 47 specimens completely changed the outcome and resolved what appeared to be significant overlaps between species. The same was true for three species from the genus *Puffinus*; Once

again, removal of what were probably incorrectly labelled samples from this group resolved the overlap that had existed between these species.

One of the four specimens of *Pelagodroma marina* also appeared to be incorrectly identified or possibly an undescribed species with a very wide nucleotide divergence from the other three. There are, doubtless, many other examples of this kind of confusion in the BOLD database and some facility for crosschecking sequence data is clearly required. In the mean time, it highlights the need for contributors to this database to be scrupulously accurate in species identification as well as the need for regular editing of submitted sequences when such anomalies become apparent. At present, facilities for rapid online editing of sequences submitted to BOLD are limited.

Incorrectly or rather inadequately labelled specimens from the genus *Cyanorhamphus* also resulted in confusion and the need for considerable research to determine actual species. This was partly due to the fact that their species status has changed in recent times.[24] However, in this case it was compounded by the fact that there is very little COI nucleotide divergence between species of this genus in New Zealand. The three clades identified by K2P distance based phylogenies and phylogenies based on Bayesian inference were correlated with three recognised species but they showed considerably less divergence than is normally associated with separation at the species level. Interestingly, the two species found on Antipodes Island (*C. unicolor* and *C. hochstetteri*) showed greater divergence than was observed between either of them and *C. novaezelandiae*. This genus has been much studied in New Zealand[46, 48] because it is a difficult group taxonomically.[24] Increasingly accessible molecular techniques may help to accurately resolve the debates that surround it.

It seems clear that where there is extensive hybridisation between congeneric species, DNA barcoding becomes relatively ineffective as a means of species identification. This was particularly the case for members of the genus *Larus* and, to a lesser extent, for the two *Thalassarche* species, *T. melanophrus* and *T.*

chrysostoma, which appear to be more closely related than is usual for separate species. Identification of members of the genus *Cyanorhamphus* may also suffer from this problem.

Thus, DNA barcoding provides a useful stream of data beyond simply identifying species. It provides a means of highlighting ambiguous sequences from species, which, if supported by other lines of evidence, may lead to the identification of new subspecies or species or lead to the amalgamation of two or more species. At the very least, it is another character in a species description. The evidence presented here supports the use of DNA barcoding to assist in the delimitation of species as well as its more established role of simple species identification.

5.5 References

1. Godfray, H.C.J., *To boldly sequence.* TRENDS in Ecology and Evolution, 2006. **21**(11): p. 603-604.

2. Song, H., Buhay, JE, Whiting, MF, Crandall, KA, *Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified.* Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13486-91.

3. Meier, R., Zhang, G. and Ali, F., *The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification.* Stystematic Biology, 2008.

4. DeSalle, R., Egan, MG, Siddall, M, *The unholy trinity: taxonomy, species delimitation and DNA barcoding.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1905-16.

5. Rubinoff, D., S. Cameron, and K. Will, *A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification.* J Hered, 2006. **97**(6): p. 581-94.

6. Roe, A.D. and F.A.H. Sperling, *Patterns of evolution of mitochondrial cytochrome c oxidase I and II DNA and implications for DNA barcoding.* Molecular Phylogenetics and Evolution, 2007. **44**(1): p. 325-345.

7. Moritz, C. and C. Cicero, *DNA barcoding: promise and pitfalls.* PLoS Biol, 2004. **2**(10): p. e354.

8. Seberg, O. and G. Petersen, *How many loci does it take to DNA barcode a crocus?* PLoS One, 2009. **4**(2): p. e4598.

9. Hajibabaei, M., Singer, GAC, Hebert, PDN, Hickey, DA, *DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics.* Trends in Genetics, 2007. **23**(4): p. 167-172.

10. Lipscomb, D., Platnick, M. and Wheeler, Q., *The intellectual content of taxonomy: a comment on DNA taxonomy.* TRENDS in Ecology and Evolution, 2003. **18**: p. 65-66.

11. Ebach, M.C. and C. Holdrege, *DNA barcoding is no substitute for taxonomy.* Nature, 2005. **434**(7034): p. 697.

12. Rubinoff, D., *DNA barcoding evolves into the familiar.* Conserv Biol, 2006. **20**(5): p. 1548-9.
13. Desalle, R., *Species discovery versus species identification in DNA barcoding efforts: response to Rubinoff.* Conserv Biol, 2006. **20**(5): p. 1545-7.

14. Seberg, O., Humphries, CJ, Knapp, S, Stevenson DW Perersen et al., *Shortcuts in systematics? A commentary on DNA based taxonomy.* Trends in Ecology and Evolution, 2005. **18**: p. 63-65.

15. Hebert, P., Penton, EH, Burns, JM, Janzen, DH, Hallwachs, W, *Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator.* Proc Natl Acad Sci U S A, 2004. **101**(41): p. 14812-7.

16. Paquin, P. and M. Hedin, *The power and perils of 'molecular taxonomy': a case study of eyeless and endangered Cicurina (Araneae : Dictynidae) from Texas caves.* Molecular Ecology, 2004. **13**(10): p. 3239-3255.

17. Penton, E.H., P.D. Hebert, and T.J. Crease, *Mitochondrial DNA variation in North American populations of Daphnia obtusa: continentalism or cryptic endemism?* Mol Ecol, 2004. **13**(1): p. 97-107.

18. Kerr, K., Stoeckle, MY, Dove, CJ, Weigt, LA, Francis, CM, Hebert, PDN, *Comprehensive DNA barcode coverage of North American birds.* Molecular Ecology Notes, 2007. **7**(4): p. 535-543.

19. Lambert, D., Paterson, HE, *Morphological resemblance and its relationship to genetic distance measures. Reprinted in The Recognition Concept of Species*, ed. S. McEvey. 1993: Johns Hopkins University Press.

20. Ferguson, J.W.H., *On the use of genetic divergence for identifying species.* Biological Journal of the Linnean Society, 2002. **75**: p. 509–516.

21. Waugh, J., Evans, MW, Millar, CD, Lambert, DM, *Birdstrikes and barcoding: can DNA methods help make the airways safer?* Molecular Ecology Resources, 2010. **9999**(9999).

22. Banks, J., Paterson, AM, *A preliminary study of the genetic differences in New Zealand oystercatcher species.* New Zealand Journal of Zoology, 2007. **34**(2): p. 141-144.

23. Heather, B., Robertson, H., *The Field Guide to the Birds of New Zealand*. 3rd ed. 2005, Auckland: Penguin Books.

24. *Checklist of the birds of New Zealand*. 4th ed, ed. B. Gill, Bell, BD, Chambers, GK,Medway, DG, Palma, RL, Scofield, RP, Tennyson AJD and Worthy TH. Vol. 1. 2010: Te Papa Press.

25. Hebert, P., Cywinska, A, Ball, SL, DeWaard, JR, *Biological identifications through DNA barcodes.* Proceedings of the Royal Society of London Series B-Biological Sciences, 2003. **270**(1512): p. 313-321.

26. *New Zealand recognised bird names (NZRBN) database*. 2003 [cited 2010 28 January]; Available from: http://www.bird.org.nz/nzrbn.htm.

27. Holdaway, R., Worthy, TH, Tennyson AJD, *A working list of breeding bird species of the New Zealand region at first human contact.* NZ J. Zoology, 2001. **28**(2): p. 119-187.

28. Clements, J., *The Clements checklist of birds of the World*. 6 ed. 2007: Cornell University.

29. Banks, J., Mitchell, AD, Waas, JR, Paterson, AM, *An unexpected pattern of molecular divergence within the blue penguin (Eudyptula minor) complex. .* Notornis, 2002. **49**(1): p. 29-38.

30. Baker, A., Pereira, SL, Haddrath, OP, Edge, KA, *Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling.* Proceedings of the Royal Society B-Biological Sciences, 2006. **273**(1582): p. 11-17.

31. Banks, J., Cruickshank, RH, Drayton, GM, Paterson, AM, *Few genetic differences between Victorian and Western Australian blue penguins, Eudyptula minor* NZ J. Zoology, 2008. **35**(3).

32. *All Birds Barcoding Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.barcodingbirds.org/#.

33. Sullivan, W., Wilson, K, *Differences in habitat selection between Chatham petrels (Pterodroma axillaris) and broad-billed prions (Pachyptila vittata): implications for*

management of burrow competition. New Zealand Journal of Ecology, 2001. **25**(2): p. 65-69.

34. Lawrence, H., Taylor, GA, Crockett, DE, Millar, CD, Lambert, DM, *New genetic approach to detecting individuals of rare and endangered species.* Conserv Biol, 2008. **22**(5): p. 1267-76.

35. Force, M., Cotton, JM, Rowlett, RA, Ballance LT, *First records of Chatham Island petrel Pterodroma axillaris in the eastern Pacific Ocean.* Marine Ornithology, 2009. **37**: p. 277-279.

36. White, R., Clausen, AP, *Rockhopper Eudyptes chrysocome chrysocome X Macaroni E. chrysolophus penguin hybrids apparently breeding in the Falkland Islands.* Marine Ornithology, 2002. **30**: p. 40-42.

37. Palma, R.W., TH, Tennyson, AJD, *Resolution of the status of the taxon Apteryx maxima.* Tuhinga: Records of the Museum of New Zealand Te Papa Tongarewa, 2003. **14**: p. 1-9.

38. Burbridge, M., Colburn, RM, Robertson, HA, Baker, AJ, *Molecular and other biological evidence supports the recognition of at least three species of brown kiwi.* Conservation Genetics, 2003. **4**: p. 167-177.

39. Baker, A., Daugherty, CH, Colbourne, R, McLennan, JL, *Flightless brown kiwis of New Zealand possess extremely subdivided population structure and cryptic species like small mammals.* Proc Natl Acad Sci U S A, 1995. **92**(18): p. 8254-8.

40. Shepherd, L.D. and D.M. Lambert, *Ancient DNA and conservation: lessons from the endangered kiwi of New Zealand.* Mol Ecol, 2008. **17**(9): p. 2174-84.

41. Robertson, C., Hyvonen, P, Fraser, MJ, Pickard, CR, *Atlas of bird distribution in New Zealand 1999-2004*. 2007, Wellington: Ornithological Society of New Zealand, Inc.

42. Ratnasingham, S. and P.D.N. Hebert, *BOLD: The Barcode of Life Data System (www.barcodinglife.org)*. Molecular Ecology Notes, 2007. **7**(3): p. 355-364.

43. Waugh, S., Weimerskirch, H, Moore PJ, Sugar, PM, *Population dynamics of Black-browed and Grey-headed Albatrosses Diomedea melanophrys and D. chrysostoma at Campbell Island, New Zealand, 1942-96.* Ibis, 1999. **141**: p. 216-225.

44. Bell, D., *Genetic differentiation, geographic variation and hybridization in gulls of the Larus.* The Condor, 1996. **98**: p. 527-546.

45. Pons, J.M., A. Hassanin, and P.A. Crochet, *Phylogenetic relationships within the Laridae (Charadriiformes : Aves) inferred from mitochondrial markers.* Molecular Phylogenetics and Evolution, 2005. **37**(3): p. 686-699.

46. Boon, W., Kearvell, JC, Daugherty, CH, Chambers, GK, *Molecular systematics and conservation of kakariki (Cyanoramphus spp.).* Science for Conservation, 2001. **176**: p. 1-46.

47. *Birds of Tiritiri Matangi*. 2010 [cited 2010 7th September]; Available from: http://www.tiritirimatangi.org.nz/default.htm.

48. Boon, W., Kearvell, JC, Daugherty, CH, Chambers, GK, *Molecular sysstematics of Cyanoramphus parakeets:conservation of orange-fronted and Forbes' parakeets.* Bird Conservation International, 2000. **10**: p. 211-239.

Chapter 6

Aircraft Birdstrike

6.1 Introduction to birdstrike

DNA barcoding using a 648bp segment of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene is an established method of identifying animal species.[1-3] It provides a molecular means of identifying species from a variety of tissue types including blood, muscle, skin and feathers. One of the key advantages of DNA barcoding is that accurate species identifications can be made from very small amounts of sample (micrograms) that are otherwise unidentifiable.

Species identification is, however, dependent on the existence of a reference database of DNA barcodes from vouchered specimens (specimens identified by a taxonomist and catalogued within a museum or similar collection). Producing this database is a global effort, coordinated by the Consortium for the Barcode of Life.[4] The associated Barcode of Life Data Systems (BOLD) website[5] provides a repository for DNA barcodes as well as an "Identification Engine". The identification engine matches a DNA sequence from a sample with known barcodes from vouchered specimens, thus providing a means of identifying species. At the present time the database has over 1 200 000 DNA barcodes from more than 100 000 species,[5] although this number increases daily.

DNA barcoding has been shown to be particularly effective in identifying a wide range of Northern Hemisphere bird species[6-8] and the world's avian fauna has been chosen as one of the first groups on which to test its overall efficacy. Birds have been intensively studied using a range of techniques such as morphometrics, cytogenetics and behaviour analysis and consequently, the number of avian species have remained relatively stable.[7] Several projects to gather the DNA barcodes of birds from different geographic regions have already been published[6-8] and several more are presently being undertaken. To date, over 20 000 birds

from over 3500 species have had their DNA barcodes recorded.[9] This database represents a useful resource for identifying tissue samples.

Applications of DNA barcoding include the identification of species difficult to identify in the field,[10] the identification of species refractory to other techniques[10] as well as matching different life stages of the same species.[11] Important among these applications is the identification of species involved in birdstrike on aircraft (Figure 6.1).

Figure 6.1 This Boeing 737 passenger jet struck a flock of more than 200 starlings while climbing after takeoff from Düsseldorf airport in Germany in October 2009. It returned and made a safe landing back at Düsseldorf on one engine 45 minutes later. The cost of repairs exceeded $US1 000 000.(Photograph Juergen Kienast)

The global cost of damage, delays and cancellations caused by birdstrike is estimated in the billions of US dollars each year; moreover, birdstrike has caused considerable loss of human life.[12] In a 14 year period to 2003 in the United States, the average reported cost of direct damage to aircraft from birdstrikes was

$US103000 per incident.[13] This rose to $US147000 per incident when other costs such as aircraft down time were included.

Birdstrike events are not unusual. In the 19 years to the end of 2008, there were 87384 reported birdstrike incidents on civil aircraft in the United States.[14] Transport Canada reported 6848 birdstrikes between 1991 and 1999.[15] In New Zealand, a country of approximately four million people, there were 3199 reported birdstrike incidents in the seven years to October 2009.[16] However, evidence suggests that only approximately 20% of birdstrikes are reported worldwide by aircrew or ground maintenance staff.[15]

Moreover, reported birdstrikes have increased.[13-16] In the United States, there were fewer than 2000 in 1990. By 2009 that figure had risen to 7516.[14] These figures are reflected elsewhere[15, 16] and are probably a result of higher air traffic volumes, more diligent reporting and increases in the populations of birds frequently associated with birdstrike.[13-16]

Damage caused by birdstrikes is related to impact force, which is affected by several factors (impact speed, bird weight, bird density, bird rigidity, angle of impact, impact surface shape and impact surface rigidity[15]). However, an approximation of the force of impact can be derived from the following equation: $F=2\pi r^2pv^2/3$, where F is the force of impact, r is the radius of a hypothetically spherical bird, p is the density of the bird and v is the aircraft's velocity.[15] Thus, impact force is proportional to the square of the aircraft's velocity. Modern jets have takeoff and approach speeds of 240-360 km/h.[17, 18] At these speeds a 2kg bird will have an impact of 15000-40000kg. Light aircraft tend to have much lower approach speeds of 120km/h or less. Thus, larger aircraft are subject to greater impact forces with the majority of serious incidents involving medium to larger bird species[15] but even small birds can cause major damage.[19]

Most birdstrike incidents (>90%) occur in the proximity of airports at low altitude during landing (~60%) and take off (~37%) procedures.[14, 15] However, the

behaviour patterns of bird species in areas adjacent to airports is reasonably predictable and can often be changed by appropriate wildlife management intervention.[15] Successful wildlife management at airports is pivotal to reducing the risk of birdstrike. Its effectiveness is largely dependent on accurate data about the species that represent the greatest hazard to aircraft. Methods of management start with ensuring the environment in the locality of the airport is not unnecessarily attractive to birds (relocating landfill sites, reducing artificial ponds etc.). However, the location of many airports is unavoidably in areas that may be attractive to birds and, furthermore, airports themselves, with their large open grassy spaces, are highly attractive to some species.[20]

With the impact forces involved, often all that remains of the bird after a birdstrike incident is a smear of blood, a feather or two or a little piece of tissue. Until recently, the only means of identifying these samples was to use a system of feather identification that requires access to a comprehensive library of feathers and considerable expertise on the part of the identifier.[19] Blood or other tissue was unidentifiable until the advent of DNA sequencing technologies. Approximately a quarter of all reported birdstrikes are for species that cannot be identified by morphological means.[13, 14, 16] Thus, the data set used by wildlife management personnel is significantly incomplete. A number of institutions and agencies around the world now employ DNA barcoding to identify the species involved.[21, 22] This approach is effective with all types of tissue recovered and the data collected can be used to inform wildlife management programs at airports.

With ever increasing hazard management requirements being adopted world wide, there is a greater emphasis on species identification from birdstrike incidents. Thus, DNA barcoding may become a normal practice for agencies charged with reducing birdstrike risk.

6.2 Method

Data on birdstrikes in New Zealand were supplied by the Chairman of the National Wildlife Hazard Committee of the New Zealand Civil Aviation Authority (NZCAA).[16] Birdstrike data were collected from 68 airports around the country (New Zealand), ranging from small country airfields to international terminals.

Samples for DNA barcoding were collected at Christchurch International Airport in the South Island[23] and Wellington International Airport in the North Island of New Zealand.[24] Once collected, they were sent to the laboratory by “normal mail” for analysis.

Sample collection employed basic forensic techniques adapted from methods used by police scene of crime officers (SOCO) for collection of DNA samples. Sampling kits consisted of a Ziploc® collection bag containing a pair of examination gloves, a cultiplast® ⊂∈ 0051 sterile swab in a plastic tube, a 5ml plastic ampoule of sterile water, two alcohol swabs, a data sheet and a set of instructions (Figure 6.2). The tip of the plastic tube was cut to allow air to circulate and dry the sample after collection. Two drops of sterile water were dropped onto the sterile swab; the blood smear was then wiped with the swab to collect the blood sample before returning it to the cut tube. Some samples (BS18, 19, 22, 23, 27, 31 and 32) were collected on Whatman FTA® cards. Since surfaces on an aircraft may be contaminated with aviation fuel, cleaning agents and lubricants known to damage DNA,[25] samples were collected from the most uncontaminated areas available.

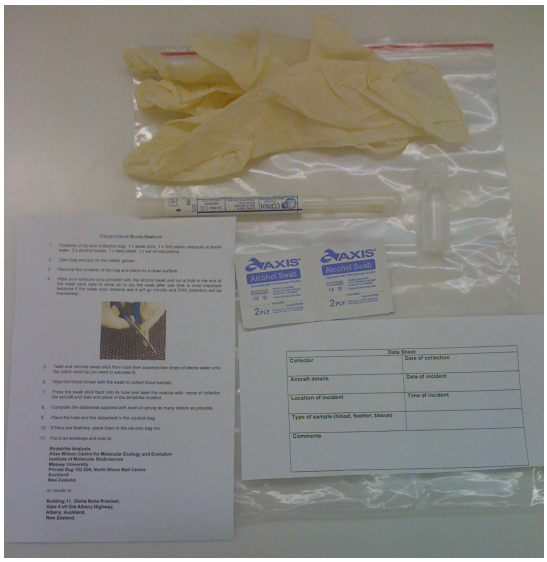


Figure 6.2 Sample collection kit

In the laboratory, DNA was extracted from samples using either standard proteinase K digestion and a phenol/chloroform extraction and ethanol

precipitation or for later samples, by one step Chelex® 100 resin (Bio-rad Laboratories)/proteinase K digestion.

For the latter method, approximately 1mm^2 of blood or tissue or 1-2 feather bulbs were digested overnight at 55°C (with slow end over end rotation) in 80µL containing 40µL of Chelex100/water slurry, 38µL of MilliQ water and 2µL of 10mg/mL proteinase K. The next day 1-2µL of the upper layer of the mix was used directly for DNA amplification by polymerase chain reaction (PCR).

PCRs were carried out in 25µL or 10µL volumes consisting of 10x PCR buffer (Invitrogen), 2.5mM $MgCl_2$, 1mg/mL BSA, 0.5µM of each primer, 0.2mM of each dNTP, 0.4U of Platinum *Taq* DNA polymerase (Invitrogen) and 1-2µL of DNA extract (~0.1-5ng).

The primers used for amplification of the barcoding region (~650bp of the 5' terminus of COI) were a forward primer, AWCF1 (5'-CGCYTWAACAYTCYGCCATCTTACC), and a reverse primer AWCR4 (5'-ATGCTCGGGTGTCTACGTCT). For some samples an alternative reverse primer, AWCR6 (5'-ATTCCTATGTAGCCGAATGGTTCTTT) was used.[26] DNA was amplified using an Applied Biosystems GeneAmp 9700 thermal cycler and the cycling conditions for samples was as follows: 2 min at 94^{o}C followed by 30 cycles of 30s at 94^{o}C, 20s at 54^{o}C and 30s at 72^{o}C followed in turn by 15 cycles of 30s at 94^{o}C, 20s at 50^{o}C and 30s at 72^{o}C.

Amplification products were visualised by electrophoresis in 1.2% agarose/0.5 x Tris Borate EDTA buffer, stained with 50ng/mL ethidium bromide and viewed under UV light. Positive amplifications were purified using the Agencourt AMPure PCR Purification System (Beckman Coulter) or the DNA Clean & Concentrator-25™ kit (Zymo Research). All purified PCR products were then cycle sequenced in both forward and reverse direction using Big Dye v3.1 chemistry and analysed on an ABI Prism 3130xl genetic analyser.

Sequences were edited using Sequencher 4.6 (Gene Codes Corporation) and identifications were carried out by matching a FASTA representation of the sequence to files lodged on the BOLD website using their identification engine.[5]

6.3 Results

6.3.1 Birdstrike statistics

There were 3199 reported birdstrikes in New Zealand in the seven years to October 2009 or approximately 470 per annum (Figure 6.3). The five most commonly identified species associated with birdstrike using morphological techniques were Spur-winged Plover (*Vanellus miles*) 28%, House Sparrow (*Passer domesticus*) 18%, Southern Black-backed Gull (*Larus dominicanus*) 5%, other Gull (*Larus*) species 5% and Oystercatcher (*Haematopus*) species 4%. Of the remainder, 20% were all other identified species and 20% were not identified (Figure 6.3).

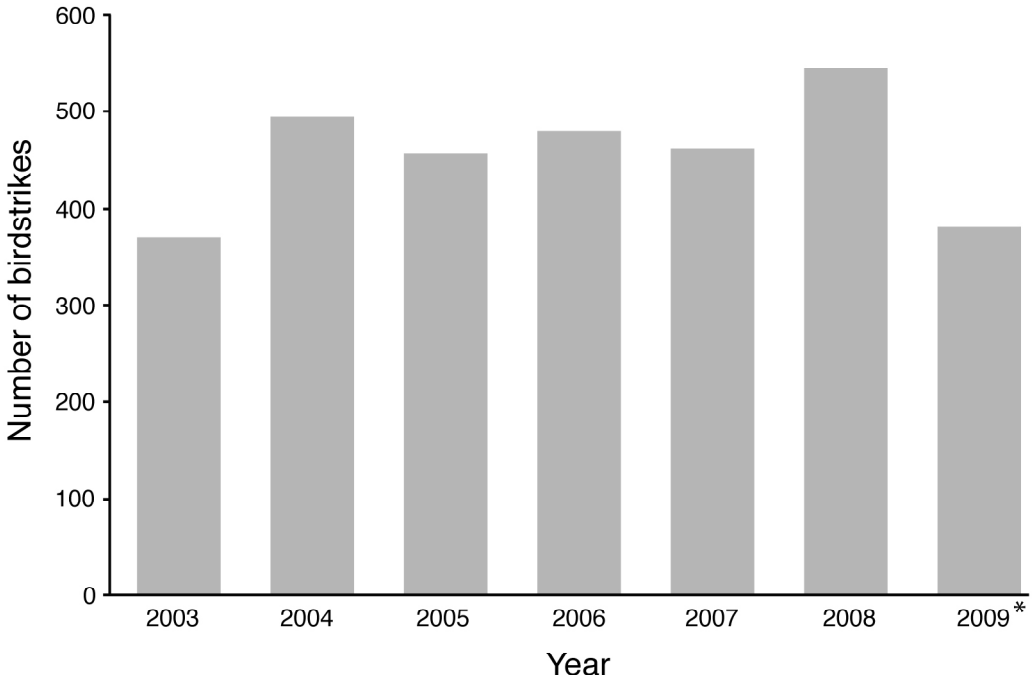


Figure 6.3 Total annually reported birdstrikes from airports in New Zealand 2003-2009 (*Data for 2009 is incomplete and ends at 19 October).[16]

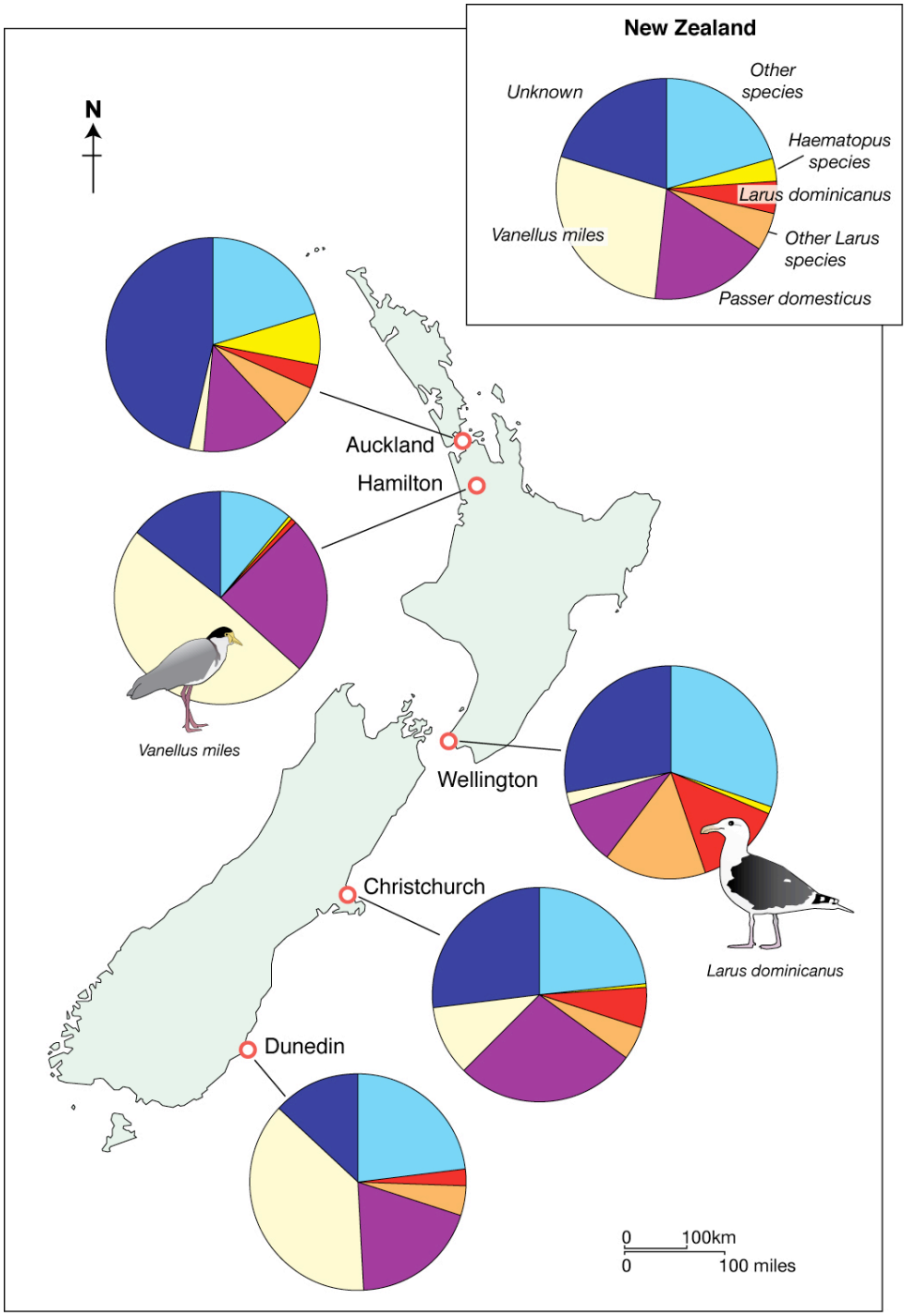


Figure 6.4 The species most highly represented in reported birdstrikes in New Zealand, overall and at the main airports within New Zealand between 1 January 2003 and 19 October 2009.[16]

As would be expected, the major airports, with much higher traffic densities and greater size, experienced the highest number of birdstrike events.[16] However, there was a significant (χ^2 df=4 p <0.001) difference between the main airports

(Auckland, Hamilton, Wellington, Christchurch and Dunedin) in the percentage of birdstrike species that were not identified (Figure 6.4).

Year by year data suggest a trend towards a higher proportion of unidentified species, however, the R value for this trend is 0.73 (Figure 6.5).

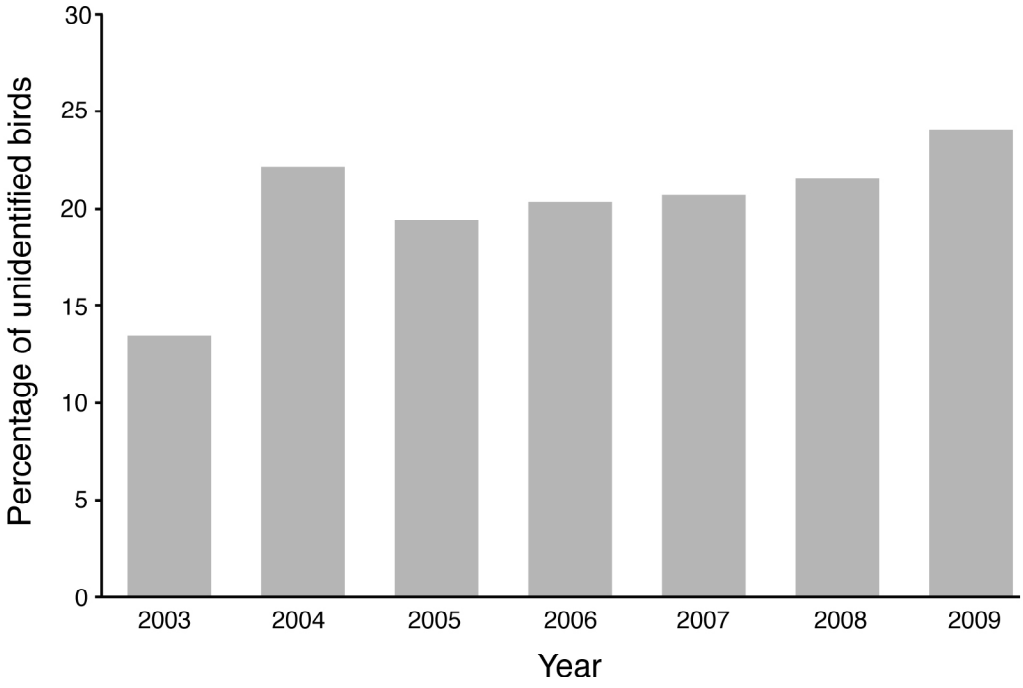


Figure 6.5 The percentage of birds involved in birdstrike incidents between 1 January 2003 and 19 October 2009 in New Zealand whose species or grouping was not identified.[16]

6.3.2 DNA barcoding

In this study, DNA barcoding was generally successful at identifying samples from birdstrike incidents. In 92.5% (N = 37/40) of cases where unidentified birdstrike sample material was tested, sufficient DNA was extracted for a positive identification of the species involved to be made (Table 6.1). There was no difference in DNA recovery between those samples collected using SOCO protocols and those collected using the Whatman FTA® cards. Of the three cases where no DNA based identification was possible (BS21, BS26 and BS37), two were incorrectly collected (the tube was not cut to allow drying of the sample with the result that the sample became contaminated) and the other was heavily

contaminated with engine lubricants.

For the three samples identified as *Haematopus* species (samples BS12, BS31 and BS36), some debate exists over the status of the various New Zealand species and the identification engine could not differentiate between species within this genus despite high quality sequence being obtained from these samples (for a discussion of the New Zealand members of the genus *Haematopus* see Chapter 5). When sequences from these two samples were matched with those on the BOLD database using their identification engine, they could not be distinguished from four putative species (*H. finschi, unicolor, longirostris* and *chathamensis*).

Samples BS19 and BS24, possibly from the genus *Aquila*, could not be accurately identified using the BOLD identification engine because, despite high quality DNA being extracted in both cases (Phred score >20 BS19 97.5% and BS24 97.4%), no record of this species exists on the BOLD website. The same applies to the unknown Passerine BS35 (Phred score >20 98.9%) and to the unknown Charadriiforme BS39 (Phred score >20 81.2%).

6.4 Discussion

Birdstrike statistics from around the world show similar trends that correlate with global changes in agricultural methods, environmental awareness and the need to report and record details of these events as required under civil aviation law.[13-16, 27] A reduction in the use of pesticides such as DDT has resulted in increases in bird populations, as has greater urban green space and set aside wilderness areas.[14] In the United States, 14 species of bird with a body mass ≥4kg have shown significant increases in population size over the past three decades.[14] This trend is also apparent in species of birds associated with birdstrike in New Zealand and Canada.[15, 16]

Table 6.1 Birdstrike samples identified using DNA barcoding. Samples were collected from Christchurch and Wellington International Airports.

ID no. (% match)	Date of incident	Common Name	Species
BS1 (98.3)	03/06/08	Australasian Harrier Hawk	*Circus approximans*†
BS2 (99.8)	05/06/08	Yellowhammer	*Emberiza citrinella**
BS3 (100)	16/06/08	Spur-winged Plover	*Vanellus miles*
BS10 (100)	07/09/08		
BS4 (100)	23/06/08	White-fronted Tern	*Sterna striata*‡
BS5 (100)	25/06/08	Southern Black-backed Gull	*Larus dominicanus*‡
BS7 (100)	27/07/08		
BS15 (99.8)	05/11/08		
BS20 (100)	25/02/09		
BS28 (100)	30/05/09		
BS33 (100)	26/09/09		
BS40 (100)	23/13/10		
BS6 (100)	11/07/08	Australian Magpie	*Gymnorhina tibicen**
BS14 (99.7)	17/10/08		
BS8 (100)	04/08/08	Goldfinch	*Carduelis carduelis**
BS9 (100)	11/08/08	Banded Dotterel	*Charadrius bicinctus*†
BS11 (100)	16/10/08	Starling	*Sturnus vulgaris**
BS12 (100)	16/10/08	Oystercatcher	*Haemotopus spp. (unicolor, finschi)*
BS31§ (100)	21/08/09		
BS36§ (100)	03/01/10		
BS13 (99.8)	17/10/08	Black-fronted Tern	*Chlidonius albostriatus*†
BS16 (100)	25/11/08	Black-billed Gull	*Larus bulleri*†
BS17 (98.8)	25/11/08	Little Owl	*Athene noctua**
BS18§ (94.4)	29/01/09	Paradise Shelduck hybrid	*Tadorna spp.*
BS19§ (93.4)	05/02/09	Unknown	Closest match *Aquila spp.*
BS24 (93.4)	05/02/09		
BS21	23/03/09	No DNA recovered	N/A
BS26	14/04/09		
BS37	12/01/10		
BS22§ (100)	20/03/09	Skylark	*Alauda arvensis**
BS32§ (100)	21/08/09		
BS23§ (99.7)	23/03/09	New Zealand Pipit	*Anthus novaeseelandiae*‡
BS25 (100)	26/04/09	Blackbird	*Turdus merula**
BS27§ (100)	19/05/09	Rock Pigeon	*Columba livia**
BS34§ (100)	07/11/09		
BS38 (99.7)	14/01/10		
BS29 (100)	26/06/09	Greenfinch	*Carduelis chloris**
BS30 (100)	10/07/09	Black Swan	*Cygnus atratus**
BS35§ (91.9)	05/12/09	Unknown	*Passerine*
BS39§	06/03/10	Unknown	Closest match *Hydroprogne sp.*

* Introduced species
† Endemic species
‡ Native species
§ Sample collected on FTA card

Along with increases in bird populations, there has been an increase in aircraft movements too.[14, 15] In the United States, aircraft movements have increased by 56% since 1980[14] and this is reflected around the world.[15] Significant changes to aircraft design may also have unwittingly increased the risks associated with birdstrike. Most commercial air carriers have been replacing older three and four engined aircraft with more modern and more powerful two engined aircraft.[14] These new engines have a greater overall frontal area, which increases the chance of strikes.[14, 15] Thus, these aircraft may be more vulnerable to birdstrike and in particular to the multiple ingestion of birds as occurred with US Airways Airbus 320 flight 1549, which made a forced landing in the Hudson River after ingesting Canada geese into both engines (Figure 6.6).[22] It is expected that the risk, frequency and potential severity of birdstrikes will increase in the coming decade.[14]

Figure 6.6 US Airways Airbus 320 flight 1549 after making a forced landing in the Hudson River on 15 January 2009. The aircraft suffered a multiple ingestion of Canada geese into both engines with a resultant complete loss of power. All passengers and crew survived. The aircraft was a total loss at a cost of $US55-65 million. (Photo Associated Press)[14]

However, analysis of the data for New Zealand shows no statistically significant change in birdstrike numbers for the years 2003-2009. Furthermore, although a

positive trend is apparent in the number of birdstrike species that are unidentified (Figure 5), this was also not statistically significant.

In common with many other countries, the location of New Zealand airports mainly in coastal or rural areas predisposes them to higher risk from birdstrike.[20] Moreover, the proximity of bird sanctuaries and breeding areas to many airports adds to this risk.[20] Nonetheless, a high degree of awareness and active wildlife management programmes within New Zealand have, so far, prevented the worst outcomes from these events.

The fact that many of the people reporting birdstrikes have limited knowledge of birds, including pilots, aircraft engineers and ground maintenance staff, means that within the NZCAA data, and, indeed, data collected all around the world, species identification is frequently inaccurate.[28] Even for an expert, identification of a bird seen approaching at >300km/h from the cockpit of an aircraft would be of questionable accuracy. Thus, the number of incorrectly identified or unidentified species is probably far higher than the reported 20%.

Moreover, many of the NZCAA identifications represent groups such as duck, finch, owl or pigeon and do not give the species. For example, there are two *Larus* species apart from *L. dominicanus* in New Zealand (*L. novaehollandiae and L. bulleri*) that are simply identified as “gulls”. No attempt is made to differentiate *Haematopus* species in the NZCAA data and there is evidence that the large numbers of the House Sparrows (*Passer domesticus*) listed may, in many cases, be several small passerine bird species grouped together under this heading. However, since there are no other species that resemble the Spur-winged Plover (*Vanellus miles*) in New Zealand and because this species is so frequently involved in birdstrike, even the relatively unskilled observer can correctly identify it. Thus, data for this species are probably reasonably accurate.

Within these limitations, there is significant ($p < 0.001$) variation in the percentage of specimens that could not be identified at each of the main airports (Figure 2).

This may be because in some airports, a relatively small number of species are involved in incidents whereas in others, the avifauna is more diverse and identification is correspondingly more difficult. Within New Zealand, the Civil Aviation Rule Part 12 requires that all birdstrikes and near misses must be reported.[27] Furthermore, the International Civil Aviation Organisation (ICAO) has a requirement, that all countries should forward wildlife strike reports to ICAO for inclusion in the ICAO Bird Strike Information System (IBIS) database.[29] However, in some cases, species cannot be identified and it is not mandatory to do so. Nonetheless, where possible, this should be done in order to inform targeted wildlife management. Dolbeer et al. 2009[14] state that "The identification of the exact species struck is particularly important" since "This species information is critical for biologists developing and implementing wildlife risk management programs at airports because a problem that cannot be measured or defined cannot be solved."

A variety of new and ingenious ways are being devised to combat the threat from birdstrike. Many of these are dependent on a good knowledge of the species involved. Species differ from place to place but, despite local variation, the overall percentage that cannot be identified by conventional means is comparable in the United States, Canada and New Zealand (20-25%).[13-16] However, analysis of the New Zealand data indicates that the overall number of unidentified species is much higher. The United Kingdom Birdstrike Committee reports a much higher rate of unidentified species (60%),[30] as does the German Committee on Prevention of Birdstrikes.[31] DNA barcoding can improve the number and accuracy of identifications where other means are difficult or impossible.

Within this study, DNA sequences for six samples could not be identified to species level. For samples BS12, BS31 and BS36 (*Haematopus spp.*), some debate exists over the species status of the three taxa (*H. finschi, unicolor* and *chathamensis*) found in the New Zealand region.[32-36] Banks and Paterson (2007)[37] found little divergence between the New Zealand species despite testing a variety of mitochondrial markers, including the hypervariable region 1 of the

control region, the gene coding for the small ribosomal subunit (12S) as well as the COI barcoding region and the Cytochrome *b* gene. Thus, failure to resolve these samples to species level may be an artifact of incorrect taxonomy rather than any inherent problem with the method. Conversely, speciation within this genus may have been associated with little divergence at the molecular level[38, 39] or perhaps there has been little time for the genetic divergence typically observed between species. In either case DNA barcoding will be incapable of distinguishing them.

Two samples from the Falconiformes, possibly from the genus *Aquila* (samples BS19 and BS24), one Passeriforme (BS35) and one Charadriiforme (BS39) could not be identified to species level using DNA barcoding because the BOLD database does not contain any matching sequence. The first two originated from the same incident but were separately submitted for identification by different airport authorities. They were found in the damaged wing flap of an aircraft in New Zealand that had recently arrived from Australia. High quality sequence from both samples revealed no matches with database species. New Zealand Falconiformes have had their DNA barcodes characterised and these are available on the BOLD identification engine; however, the BOLD website does not have comprehensive DNA barcodes for the avifauna of Australia. It is probable that the birdstrike occurred in Australia. Subsequently, morphological identification techniques used on feather and bone material suggested that this might be a Black Kite (*Milvus migrans*).[23] The Passerine sample was recovered from an aircraft inbound from Nadi in Fiji and, despite the recovery of good DNA sequence, would not match with any sequence on the BOLD database. The Charadriiforme (probably a Tern) was the only bird of New Zealand Origin that could not be identified. These birds are one of the few groups that have not had their DNA barcodes comprehensively characterised within the New Zealand avifauna.

The above examples demonstrate some of the issues that remain to be resolved with the application of DNA barcoding to birdstrikes. DNA barcoding has identified many potential synonymous and cryptic species in various projects, including this one, around the world and further taxonomic work is required to resolve these.

Nonetheless, identification to genus level, in the case of the *Haematopus* species, will probably provide sufficient information for wildlife management purposes in this particular case. However, the unidentified Falconiforme and Passeriforme that originated abroad as well as the New Zealand Charadriiforme are a result of insufficient data within the BOLD database and highlight the need for further development of the BOLD database for species from more regions around the world.

Within the barcoded sample data here presented, there appears to be a number of species that were over represented. The Spur-winged Plover (*Venellus miles*), a medium sized bird (350g) is a relatively recent migrant to New Zealand with the first breeding pair being noted in 1932[36] but it is implicated in 28% of all morphologically described birdstrike events around the country[15] and its numbers are increasing. However, only 5% (BS3 & BS10) of the samples tested (χ^2 df=1 $p<0.05$) were of this species. Christchurch and Wellington have below national average birdstrikes with *V. miles* but these figures may also be an artifact of sample size. In addition, since this species is so frequently implicated in birdstrike, airport staff may have become skilled at identifying remains from this species and, therefore, do not typically send it to be DNA barcoded as often as other less frequently observed species.

Seventeen percent (N=7) of the samples DNA tested were Southern Black-backed Gulls (*Larus dominicanus*) compared with a morphologically described national average of 6%. Of these, five came from Wellington International Airport. This represents >38.5% of samples sent from Wellington (χ^2 df=1 $p<0.001$). This airport is situated on an isthmus that is surrounded by sea. There is a breeding colony of this large gull species on an island about three kilometres away. *Larus dominicanus* is also known to prey on other species of bird and these factors may account for its unusual preponderance in the samples sent from Wellington.

Of the remaining birds represented more than once in the DNA tested specimens, Rock Pigeons (*Columba livia*) occurred in 7.5% of DNA tested specimens versus

1.6% of the morphologically identified specimens, Oystercatchers (*Haematopus spp*. see above) in 7.5% versus 2.5%, Skylarks (*Alauda arvensis*,) occurred in 5% versus 0.8% and Australian Magpies (*Gymnorhina tibicen*,) occurred in 5% versus 0.9% (Table 1). Thus, the frequency of the DNA identified samples did not reflect that of the morphologically identified samples.

Had the frequency of different species identified using DNA barcoding been very similar to the frequency of the morphologically identified specimens, the argument that not much is to be gained by carrying out DNA tests would be compelling. However, the opposite is true and it appears that much of the data supplied to the NZCAA may be inaccurate. Bearing in mind that, to date, DNA barcoding has only been used when morphological means have failed, the figure of 85% identifications of these samples provides a benchmark for the number of species that could be identified and if applied to all samples would provide wildlife officers with an almost complete data set to inform their management strategies. Since DNA barcoding is based on very rapidly developing technologies, costs are likely to be reduced and processing times decreased in the future.

For more than a century, the hazards of air travel have been steadily decreasing. More sophisticated aircraft design and engineering, improved air traffic management and greater emphasis on hazard management within the industry have led to a situation where the safety of air travel rivals or exceeds that of almost every other form of transportation. However, despite this, the high death to injury ratios and the number of fatalities associated with aircraft crashes, continue to cause public concern. One hazard that has increased is that from birdstrikes. The aim of reporting birdstrikes is to reduce the hazard from this source but without accurate identification of the species involved, the data are of limited use to wildlife managers. Morphological identification combined with circumstantial data can be very accurate when carried out by experts. In the absence of this expertise or where samples are otherwise unidentifiable, DNA barcoding provides a cost effective and accurate means of identifying samples collected from these incidents that may be adopted in order to comply with increasingly stringent regulations.

6.5 References

1. Rubinoff, D., *DNA barcoding evolves into the familiar.* Conserv Biol, 2006. **20**(5): p. 1548-9.

2. Waugh, J., *DNA barcoding in animal species: progress, potential and pitfalls.* Bioessays, 2007. **29**(2): p. 188-197.

3. Frezal, L., Leblois, R, *Four years of DNA barcoding: current advances and prospects.* Infect Genet Evol, 2008. **8**(5): p. 727-36.

4. *Consortium for the Barcode of Life*. 2009 [cited 2009 11th November]; Available from: http://barcoding.si.edu/.

5. *Barcode of Life Data Systems*. 2009 [cited 2010 22nd July]; Available from: http://www.barcodinglife.org/.

6. Hebert, P., Stoeckle, MY, Zemlak, TS, Francis, CM *Identification of Birds through DNA Barcodes.* PLoS Biol, 2004. **2**(10).

7. Yoo, H., Eah, JY, Kim, JS, Kim, YJ, Min, MS, Paek, WK, Lee, H, Kim, CB, *DNA barcoding Korean birds.* Molecules and Cells, 2006. **22**(3): p. 323-327.

8. Kerr, K., Stoeckle, MY, Dove, CJ, Weigt, LA, Francis, CM, Hebert, PDN, *Comprehensive DNA barcode coverage of North American birds.* Molecular Ecology Notes, 2007. **7**(4): p. 535-543.

9. *All Birds Barcoding Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.barcodingbirds.org/#.

10. Hebert, P., Ratnasingham, S, deWaard, JR, *Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species.* Proceedings of the Royal Society of London Series B-Biological Sciences, 2003. **270**: p. S96-S99.

11. Hebert, P., Penton, EH, Burns, JM, Janzen, DH, Hallwachs, W, *Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator.* Proc Natl Acad Sci U S A, 2004. **101**(41): p. 14812-7.

12. Allan, J., *A heuristic risk assessment technique for birdstrike management at airports.* Risk Analysis, 2006. **26**(3): p. 723-729.

13. Cleary, E., Dolbeer, RA, *Wildlife hazard management at airports*. 2nd ed. 2005: United States Federal Aviation Authority.

14. Dolbeer, R., Wrigh, SE, Weller, J, Begier, MJ, *Wildlife strikes to civil aircraft in the United Stated 1990-2008*. 2009, Federal Aviation Administration. p. 61.

15. MacKinnon, B., *Sharing the skies*. 2nd ed. 2004: Transport Canada.

16. Evans, M., *New Zealand birdstrike statistics to 0ctober 2003-2009*. 2008: Wellington.

17. Anon. *Airport Reference Code and Approach Speeds for Boeing Airplanes*. 2009 [cited 2009 25th June]; Available from: http://www.boeing.com/commercial/airports/faqs/arcandapproachspeeds.pdf.

18. Anon and *Airliner takeoff speeds*. 2009 [cited 2009 13th August]; Available from: http://www.aerospaceweb.org/question/performance/q0088.shtml.

19. Anon. *Natural History Highlight. CSI for Birds: Scientists Use Forensic Techniques to Improve Airport Safety*. 2009 [cited 2009 17th August]; Available from: http://www.mnh.si.edu/highlight/feathers/.

20. Caithness, T., Williams, MJ, Bull, RM,, *Birds and aircraft: a problem on some New Zealand Airfields*. 1967, Department of Internal Affairs publication: Wellington. p. 58-62.

21. Dove, C., Rotzel, NC, Heacker, M, Weigt, LA, *Using DNA barcodes to identify bird species involved in birdstrikes.* Journal of Wildlife Management, 2008. **72**(5): p. 1231-1236.

22. Marra, P., Dove, CJ, Dolbeer, R, Dahlan, NF, Heacker, M, Whatton, JF, Diggs, NE, France, C, Henkes, GA, *Migratory Canada geese cause crash of US Airways Flight 1549.* Frontiers in Ecology and the Environment, 2009. **7**(6): p. 297-301.

23. Bennet, D., *Birdstrikes at Christchurch iInternational Airport*. 2009.

24. Eyely, J., *Birdstrikes at Wellington International Airport*, J. Waugh, Editor. 2009.

25. Mroz, R., Schins, RP, Li, H, Jimenez, LA, Drost, EM, Holownia, A, MacNee, W, Donaldson, K, *Nanoparticle-driven DNA damage mimics irradiation-related carcinogenesis pathways.* Eur Respir J, 2008. **31**(2): p. 241-51.

26. Patel, S., Waugh, J, Millar, CD, Lambert, DM, *Conserved primers for DNA barcoding historical and modern avian samples.* Molecular Ecology Resources, 2009. **10**(3): p. 431-438.

27. *Civil Aviation Rules Part 12, Amendment 2*. 2009 [cited 2009 12 November]; Available from: http://www.caa.govt.nz/rules/Amendments/Part_012_Amendment_02.pdf.

28. Linnell, M., Conover, MR, Ohashi, TJ, *Biases in Bird Strike Statistics Based on Pilot Reports.* The Journal of Wildlife Management, 1999. **63**(3): p. 997-1003.

29. *ICAO Position and Activities with respect to Bird Strikes*. 2001 [cited 2009 12 November]; Available from: http://www.icao.int/icao/en/ro/nacc/acilac/02_pinos_icao_birdcontrol1.pdf.

30. Riensema, K., *Notes of Meeting held Thursday 23 April 2009* U.B.C.M. (UKBSC), Editor. 2009, CAA Safety Regulation Group.

31. Morgenroth, C., *German birdstrike statistics*. 2009, Deutscher Ausschuss zur Verhütung von Vogelschlägen im Luftverkehr e.V.

32. Oliver, W.R.B., *New Zealand birds*. 2nd ed. 1955, Wellington: A. H. & A. W. Reed.

33. Baker, A., *Ecological and behavioural evidence for the systematic status of New Zealand oystercatchers (Charadriiformes: Haematopodidae).* Life Science Contributions of the Royal Ontario Museum, 1974. **96**: p. 1-34.

34. Falla, R., Sibson, RB, Turbott, EG, *The new guide to the birds of New Zealand*. 1975, Auckland: Collins.

35. Turbott, E.G., *Checklist of the birds of New Zealand*. 3rd ed. 1990, Auckland: Random Century.

36. Heather, B., Robertson, H., *The Field Guide to the Birds of New Zealand*. 3rd ed. 2005, Auckland: Penguin Books.

37. Banks, J., Paterson, AM, *A preliminary study of the genetic differences in New Zealand oystercatcher species.* New Zealand Journal of Zoology, 2007. **34**(2): p. 141-144.

38. Lambert, D., Paterson, HE, *Morphological resemblance and its relationship to genetic distance measures. Reprinted in The Recognition Concept of Species*, ed. S. McEvey. 1993: Johns Hopkins University Press.

39. Ferguson, J.W.H., *On the use of genetic divergence for identifying species.* Biological Journal of the Linnean Society, 2002. **75**: p. 509–516.

Chapter 7

General Conclusion

7.1 Introduction

Following the development of suitable primers for the birds of New Zealand and its region (Chapter 3), DNA barcodes were successfully obtained from most of the native and endemic as well as a number of introduced, vagrant or migratory species. Samples from those that remain to be barcoded (usually rare species from offshore and sub-Antarctic islands) are gradually being sequenced and added to the Barcoding New Zealand Birds database as they become available. No technical problems remain that will impede the completion of this project.

With some exceptions, generally explicable by taxonomic uncertainty, DNA barcodes differentiated New Zealand bird species. They supported recent taxonomic changes to some species (section 5.2.2), separately identified subspecies (section 5.2.1) and possible subspecies (section 5.3.3) and highlighted potential new species (section 5.2.3). They also identified some species whose status may need revision in the light of molecular, behavioural, morphological and geographic data (sections 5.2.3, 5.3.2 and 5.3.6).

DNA barcodes provided information of potential interest to people charged with the conservation of endangered species (section 5.2.5) as well as those whose job it is to reduce risk from aircraft birdstrike (Chapter 6). They highlighted the genetic variability found in widely distributed species (section 5.2.6) as well as the lack of variability in some isolated species or species that may have undergone genetic bottlenecks in the recent past (Chapter 4). However, in cases where hybridization may occur between congeneric species (sections 5.3.2, 5.3.3 and 5.3.4), they proved less effective.

In all of the above, DNA barcodes in isolation were insufficient evidence to determine species boundaries and, of course, formed only a rudimentary part of species descriptions. However, when supported by morphological, geographic and behavioural data, this molecular data adds another objective dimension to a species description. Furthermore, when used purely for species identification purposes, DNA barcoding proved generally successful. In over 95% of New Zealand bird species examined, an accurate species allocation could be made using DNA barcodes alone.

The promise of a system that offers rapid and accurate identification of species from even minute samples is undeniable. Possible applications, after only the briefest consideration, begin to become apparent and, these are many and varied. A number of analytical barriers have been overcome and the costs associated with DNA sequencing have been reduced through the use of high throughput devices.[1] The approach has proved effective in a variety of taxonomic groups from diverse geographical settings (see Chapter 1 for details of studies). Moreover, a wide group of people and organisations have become involved and some of the intellectual and technical barriers that have appeared before it have been resolved.[2-4]

Early objections to DNA barcoding[5-10] (see Chapter 1) have largely died down as perceived threats posed by what were seen as extravagant aims for the project have proved unfounded. Many, though by no means all taxonomists (the group most directly affected) now see DNA barcoding as a useful tool in their armoury rather than as a threat to their livelihoods.[11, 12] Others, who objected to DNA barcoding not as a concept but because of the distance methods used for differentiating species,[13] have been mollified by the inclusion of a character based system where K2P distance fails to identify species.

Some problems that have occurred with various taxa are gradually being resolved. On the other hand, solutions to some quite significant problems remain elusive. Technologically, DNA barcoding has advanced with the rapid developments in

general DNA sequencing technology.[1] This has tended towards large centralised high throughput sequencing devices that are appropriate for the accumulation of the DNA barcode database, requiring as it does, extraction and sequencing of DNA from large numbers of samples. At present, laboratory facilities using standardardised protocols are required to extract all DNA sequences. This limits the applicability of this system of species identification. No major technological developments specific to DNA barcoding have been made.

The future of barcoding lies not in an endless large-scale expansion of the BOLD database. It will have to be updated and maintained but the real usage will come from individuals and organisations using the database for identification purposes.[14] Their requirements will be for small-scale, portable, rapid sequencing devices to examine biological material of interest. To date, no equipment for this kind of sequencing has been developed but as time passes and the database increases in size, the need for such a device is building. If and when such equipment becomes available, a major step towards rapid, accurate species identification will have been taken.

The above is not to say that real world applications have yet to be found for DNA barcoding. Indeed, apart from forming an increasingly normal part of any species description,[15-17] barcoding is used in many applications where other means of identification are difficult or impossible. Some of these have already been enumerated elsewhere in this thesis (Chapters 1, 5 and 6) and can be seen in summary in Table 7.1.

Table 7.1 Published examples of current usage of DNA barcoding.

Use	References
Identification of species at different life stages	[18-22]
Identification of processed food contents	[23, 24]
Identification of remains from aircraft birdstrike	[25-27]
Identification of morphologically similar species	[28-49]
Assist in conservation projects	[50, 51]
Identification of possible cryptic species	[52-64]
Identification of species important to forensics	[65]
Identification of extinct species	[66, 67]
Identification of host-specific parasites	[68-70]
Identification of invasive species	[71-73]
Used to confirm species status	[74-77]
Used to reassign known species to a new genus	[78, 79]
Used to help manage trade in ornamental fish	[80]

7.2 Present and future applications of DNA barcoding

Analysis of the published examples of the use of DNA barcoding suggests three broad areas where this technology has proved effective and indicates avenues of potential usage. What follows are some examples from these broad categories of use:

7.2.1 Accurate species identification

Where species cannot readily be identified by other means, DNA barcoding provides a relatively simple, second line, approach that offers a high degree of accuracy. Some examples of this kind of use can be found in ecological studies, the study of disease vectors, or for identification of species that are hard to identify by other means:

- By analysing the DNA barcodes of gut contents, researchers have accurately established food webs and examine predator/prey relationships in the wild[81]

- Samples from ancient permafrost cores[82] and soil, [83] generally difficult to identify using traditional methods, have been analysed using DNA barcodes to determine the array of species present at different locations and in different times

- Identification of the mosquitoes of India, some of which spread a variety of diseases[84]
- Identification of the black flies of Central and South America, some of which spread river blindness disease[85]
- Identification of the freshwater snails of Cameroon's crater lakes, some known to transmit parasites to humans and livestock[86]
- Identification of the poplar tree fungus, to monitor the distribution and spread of this fungal pathogen[87]
- Identification of parasitic species in Mexico, including micro- and macro-organisms[88]
- Identification of species that are hard to identify by other means ranging from nematode[68] and oligochaete worms[42] to aphids,[41] rhodophytes[89] and, occasionally even large animals[90]
- Identification of otherwise unidentifiable remains from birds involved in aircraft birdstrike[25-27]

7.2.2 Verification of claims made about biological material

The contents of food, herbal medicines and other products, hard to identify accurately by morphological means, can often be identified using DNA barcodes:

- Seafood, which poses particular problems due to the number of species that are marketed and the fact that it is often processed; several incidents of incorrect labelling have been identified[23, 24, 91]
- Wild bush meat and other animal products obtained from endangered species but labelled for trading purposes as coming from legal sources have been accurately identified and some efforts made to curb this multi-billion dollar trade[92]
- Whale meat labelled as coming from species that are not endangered and sold for food has been identified as coming from rare and endangered species[93, 94]

- Identification of the pods from star anise (*Illicium verum*), used in teas, herbal remedies and cooking; they cannot be morphologically distinguished from otherwise identical seedpods of a sister species, *Illicium anisatin*, which contains neurotoxins[95]

- Countering the trade in rhinoceros horn[96] and tiger parts[97] used in traditional Chinese medicine

- Identification of the smuggled eggs of endangered bird species[98]

7.2.3 Biological data collection

The third major use of DNA barcoding is data collection for potential use in a variety of fields. Examples include:

- DNA barcoding the bees of the world, essential to agriculture: approximately 2000 described species have now had DNA barcodes recorded[99]

- Producing a DNA barcode database for the grasses of the world, including the most important grain crops such as wheat, rice, and maize, as well as many of the world's most problematic invasive species[100]

- Identification of Heliothinae moths of Australia, a group of 365 species, including some serious crop pests and biosecurity threats[101]

- DNA barcoding the fishes of India,[102] Mexico,[103] Russia,[104] Australia,[105] Canada and the United States.[106]

- Identifying of the fish parasites of Canada.[107]

- Creating a DNA barcode database of the marine macroalgae of Canada, organisms that are impossible to identify using other approaches.[89]

- Creating a DNA barcode database of the birds of North America,[108] Korea,[109, 110] Argentina,[111] Scandinavia[112] and New Zealand

Thus, though the future of DNA barcoding will undoubtedly lie partly in field biology, it may also lie in any area where identification of biological material is

necessary. These may be related to medicine, conservation, agriculture, fisheries and other commercial activities or for controlling the movement of biological materials around the world. Furthermore, the accumulation of a very large database of DNA barcodes from many species of diverse life forms, is likely to prove a valuable resource for biologists interested in the study of evolution, systematics, biodiversity, bioinformatics or molecular biology.

Where other means of identification are difficult or impossible is clearly an area of potential growth for the use of DNA barcoding. This ranges from identification of morphologically similar organisms to verifying the contents of processed biological material. Both food and medicinal product labelling requires, amongst other things, a list of ingredients.[113] However, testing the claims made within these labels is often costly and time consuming, with the result that it is frequently overlooked. DNA barcoding is a relatively simple way of carrying out regular analysis of processed food and medicinal products and represents an area of future growth for this technology. Similar analysis is also a requirement for other regulatory bodies whose role is to police fisheries or the trade in biological materials.

7.3 Progress, potential and pitfalls

To the scientist and to the entrepreneur, the DNA barcoding project is one of great interest. However, before committing to such an undertaking, prudence would dictate that a careful analysis of the progress, potential and pitfalls of this project be made.

The stated aim of CBOL is “to develop DNA barcoding as a global standard for the identification of biological species”.[114] This is a tall order by any measure. It requires an enormous co-operative effort from biologists around the world to amass the databases of DNA barcodes of as many species as possible. It also requires that the arguments against amassing such a database (cost, questionable accuracy, problems with some taxa, oversimplification of complex biological questions;[5-10] see Chapter 1 for details of these debates) be effectively countered by solutions to the problems highlighted by these criticisms. It requires a major

effort of co-operation between the worlds of science, commerce and technology to develop fast effective and portable sequencing devices. Above all, it requires sufficient momentum to be gathered so that resources are available to solve the problems that have arisen and will arise along the road to achieving CBOL's aim.

The most public project in molecular biology of recent times has been the Human Genome Project (HGP). It required the development of molecular techniques, software and analytical tools to manage the data generated and a cooperative effort from many researchers to amass the data required.[115] It also needed investment to back the development of tools and research. Much of the work carried out was of a pioneering nature. This is not the case for most of the developments of the Barcode of Life project. Nonetheless, the HGP is a model of project development that may act as a guide to the Barcode of Life project. Analysis of the stages through which the HGP progressed from inception to completion suggests that in the lifetime of a large scientific project in the biological domain that is not entirely publicly funded, there are a number stages through which it must progress before final achievement of its goals (Figure 8.1).

One of the crucial times for such a project is the period after its announcement and during the building of momentum for the project.[116] Unless a critical mass of scientific interest is developed, the project will eventually wither and die. In the case of DNA barcoding that moment probably came and went after about 200000 sequences had been lodged on the BOLD database.[117] At this point it represented a sufficiently large resource of genetic information that its utility was apparent to most, if not all. With a database that now exceeds 1200000 sequences from over 100000 species or almost 8% of the total number of nucleotide sequences lodged with GenBank, the value of this resource is undeniable. This, despite unresolved problems that include difficulties with species that have undergone slow[118] or fast[119] evolution or with species that show high intraspecific nucleotide variation[120, 121] and with attempts at ameliorating some of the shortcomings of DNA barcoding by integrating its use with other data sources

(nuclear DNA, cytosystematic, morphological or behavioural) having been largely ignored.[122]

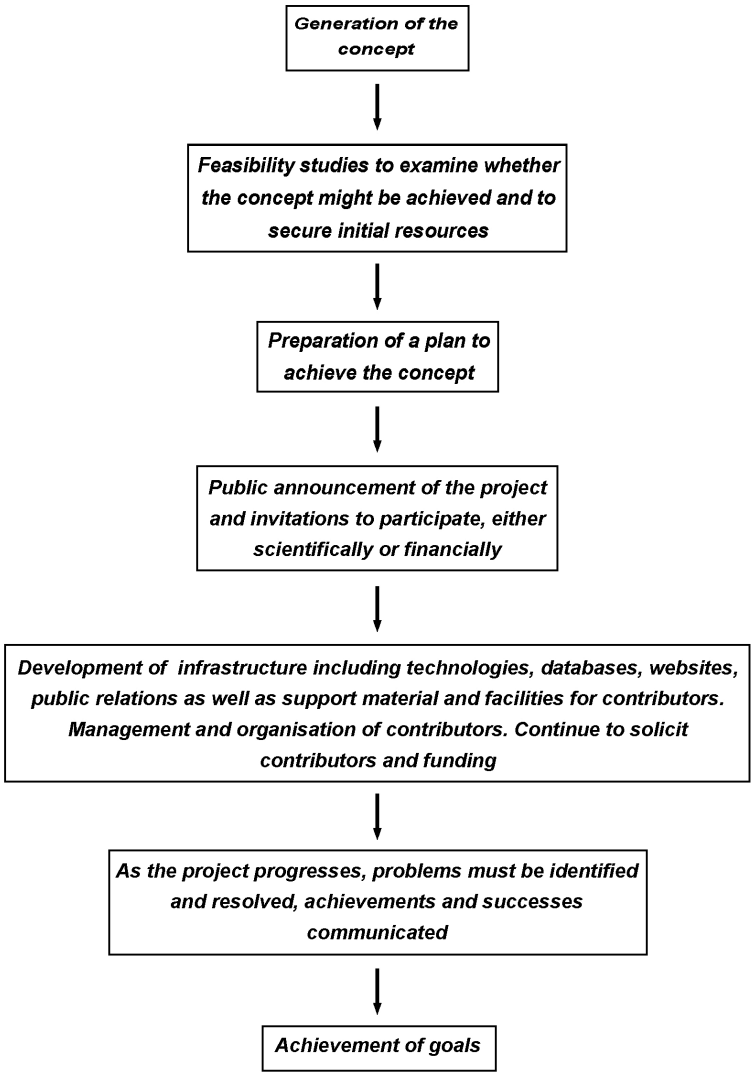


Figure 7.1 Stages of development of a major biological project such as the Human Genome Project.[115]

A brief examination of the literature via PubMed finds more than 90 publications on the subject in 2010. The initial impetus provided by a few individuals appears to have been overtaken by the project itself. It is reasonable, therefore, to assume that in time and with the continued efforts of those involved, the primary goal of CBOL may be achieved. It is useful then to examine what has been achieved and to identify what remains to be done.

There have been some problems that have been continuously addressed and appear to be near resolution. A DNA barcode for plants for example, has been a long sought goal.[123-127] The barcoding of plants has been limited by intrinsically lower rates of mitochondrial COI sequence evolution in plant genomes than that observed in animals;[128] however, recent research suggests other DNA regions may be suitable as barcodes and identifies two chloroplast DNA genes (rbcL and matK)[129] and a nuclear ribosomal internal transcribed spacer (ITS2)[130] that may prove useful for barcoding land plants.

The infrastructure to manage the DNA barcode project is well advanced[114, 131-137] and, while it may need updating and augmentation from time to time, is more or less complete. The BOLD website databases are now growing at an almost exponential rate (150000 samples were added between May and August 2010[132]) as they are updated with new sequences of specimens from a wide array of animal, fungal, protistan and plant phyla. The website incorporates a rapid search engine for matching submitted sequences with those lodged in its database and has tools for visualising and aligning sequences, generating phylogenetic trees, summarising genetic distance between selected specimens and nearest neighbour analysis as well as tools for monitoring the development of projects. Moreover, it serves as an information portal for people interested in any aspect of DNA barcoding.[132] CBOL also oversees regular international conferences on DNA barcoding.[114]

The overall project has been subdivided into a number of smaller campaigns (Table 7.2) and these are set up to coordinate the development of different parts of the BOLD database. Many thousands of scientists are contributing to these campaigns. As of August 2010, there were approximately 5700 registered users of the BOLD website but registration is not a prerequisite for using the databases for many processes including species identification. Thus, the total number of actual users is probably higher than this. (Megan Milton, BOLD Lead Data Manager, personal communication) and, although there is a very large task ahead, the project is under way and the databases are growing.

Table 7.2 A summary of campaigns overseen by the Consortium for the Barcode of Life (CBOL).[131]

Campaign	Aims
Formicidae Barcode of Life	Launched in 2008 with the aim of barcoding all of the world's more than 12,000 ant species.
Bee Barcode of Life Initiative	Coordinates the assembly of a standardized reference sequence library for all 20,000 bee species.
All Birds Barcoding Initiative	A campaign to collect DNA barcodes from five or more individuals of all of the approximately 10,000 bird species.
Trichoptera Barcode of Life	A long-term project to barcode the world's approximately 13,000 species of caddisflies.
Coral Reef Barcode of Life	A detailed barcode study of fishes at the Great Barrier Reef.
Fish Barcode of Life Initiative	A global effort to coordinate assembly of a standardised reference sequence library for all fish species.
All Fungi Barcoding	Provides up-to-date information on fungal barcoding and facilitates communication and collaboration among researchers interested in fungi.
HealthBOL	Coordinates initiatives to barcode vectors, pathogens, and parasites for the betterment of human health around the world.
Lepidoptera Barcode of Life	Aims to build a barcode library for all butterflies and moths with sub-campaigns for selected families.
Mammal Barcode of Life	Aims to build a comprehensive reference library of DNA barcodes for the global mammal fauna.
Mosquito Barcoding Initiative	Project aimed at producing a global operational system for identifying mosquitoes in two years.
Marine Barcode of Life	A joint effort of the Consortium for the Barcode of Life (CBOL) and the Census of Marine Life (CoML) to barcode life in the world's oceans.
Polar Barcode of Life	Coordinates barcoding projects in Arctic and Antarctic marine, freshwater and terrestrial ecosystems.
Shark Barcode of Life	Aims to barcode the 1000 marine and 100 freshwater shark species.
Sponge Barcoding Project	The barcoding project for the complete taxonomic range of Porifera.
Tipulidae Barcode of Life	Coordinates barcoding of the world's crane flies.
Odonata Barcode of Life	Coordinates barcoding of the world's dragonflies and damselflies.
Ephemeroptera Barcode of Life	Coordinates barcoding of the world's mayflies.
Lumbricidae Barcode of Life	Coordinates barcoding of the world's earthworms.

DNA barcoding projects contributing to these campaigns are currently under way in Australia, Argentina, Brazil, Canada, China, Colombia, Costa Rica, France, Germany, India, Kenya, Madagascar, Mexico, the Netherlands, New Zealand, Norway, Panama, Papua New Guinea, Peru, Portugal, Russia, Saudi Arabia, South Africa, South Korea, the United Kingdom and the United States.[131] These include projects on organisms from four of the five Kingdoms and many Phyla

(Monera are also identified by molecular means but there is no BOLD database or campaign for them).

The campaign to which research from this thesis contributes is the All Birds Barcode Initiative (ABBI). This project has amassed over 20000 sequences from more than 3500 species, over a third of the world's bird species. Of these, 833 high quality sequences from 215 species belong to the Barcoding New Zealand Birds project. This very nearly completes a barcode database for the birds of this region, only the second region in the world, after North America,[108] to do so. The associated Barcoding Antarctic Birds is, when other projects are combined with it, almost complete too, with a total of 18 penguin species, 8 skuas, 1 gull as well as several Procellariiformes of the 45 known Antarctic species sequenced. DNA barcodes for two other regions (Korea[109] and Argentina[111]) have been published but they are far from complete.

With the completion of the Barcoding New Zealand Birds project, potential future efforts should be based in the Oceania region, since these species are most likely to stray into the New Zealand region. Australia, with an avifauna in excess of 700 species, still has a very large number of birds that require barcoding, many of which are occasional visitors to New Zealand shores. The birds of the Fiji, Cook, Tonga and French Polynesian Island groups are not only species rich[138] but also mostly have not had DNA barcode sequences recorded for them. For example, of the 77 native and endemic species found in Fiji (27 endemic and 50 native), only two have a DNA barcode recorded with BOLD.[132]

The Barcode of Life Project seems to be in admirable health and well on track to achieve its goals. With rapidly growing databases, the currently available technology is suitable for the mass sequencing that is required. Much of this is carried out in dedicated facilities at the University of Guelph, Ontario but also in other laboratories around the world. However, several issues remain unresolved. A crisis is approaching when, unless new portable sequencing technology that is able to match sequences in the field with online databases is developed, the broad

utility of this system will remain beyond reach. The development of such equipment has always been part of the aims of the DNA barcoding project and according to CBOL, “there is no major technological barrier. The components are available but require integration and, in some cases, miniaturisation.”[131] Nonetheless, to date, little has been achieved.

In addition to those previously itemised in this chapter, early problems with the barcoding system (plants,[124] cnidarians,[118] some amphibian[121] and molusc[139] species that could not be identified using DNA barcodes) have rarely been publicly revisited or resolved. Issues around the need for multiple site barcoding for accurate identification of some species have also been neglected.[140] While so many species remain to have their DNA barcodes sequenced, it seems that the tendency is to simply move on from awkward species and try something else. Nonetheless, the generally positive results from the work undertaken for this thesis contribute to the growing number of successes achieved under the umbrella of the Barcode of Life organisation and add to the global applicability of this system of species identification within the class Aves.

The potential for DNA barcoding is great. However, as with all large undertakings, the devil lies in the detail. The early development of such a project is generally about broad concepts and creating momentum and this, to a large extent has been done. Nonetheless, a number of problems remain unresolved and apart from the most obvious ones, a DNA barcode for plants for example, little appears to be being done to address them. Furthermore, despite rapid developments in sequencing technology during this period, none has been specific to the DNA barcoding project and although there is still time to address this, it indicates an indifference by the major players in the sequencing industry that does not bode well for DNA barcoding. It must be hoped that, with time, solutions will be found for these problems, because without them the use of DNA barcoding will remain the domain of the specialist scientist alone and many of the claims made by its protagonists will not be achieved.

There are many current and potential applications for DNA barcoding and as the more obvious are adopted, so more creative uses are being found for this technology. They range from the purely practical identification of species to the more academic analyses of DNA sequences. A number of problems have arisen that are still to be resolved. These are either technological or relate to taxonomic groups that are hard to identify using DNA barcodes. They limit the utility of this method of species identification. Pressure to overcome these problems is growing as the database expands, gradually increasing its potential as it does so.

As work from this thesis and many other projects demonstrates, DNA barcoding is a useful tool. Its universal applicability is yet to be established; however, with continued development its uses will multiply, thus, securing its place not only in the science of biology but also in the world at large.

7.4 References

1. Magklaras, G. (2009) *High Throughput Sequencing and the IT architecture Part 1 : Volume dimensioning and filesystems*. EBMnet.news **15**.

2. Hajibabaei, M., deWaard, JR, Ivanova, NV, Ratnasingham, S, Dooh, RT, Kirk, SL, Mackie, PM, Hebert, PD, *Critical factors for assembling a high volume of DNA barcodes.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1959-67.

3. Hajibabaei, M., Singer, GA, Clare, EL, Hebert, PD, *Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring.* BMC Biol, 2007. **5**: p. 24.

4. Rubinoff, D., *DNA barcoding evolves into the familiar.* Conserv Biol, 2006. **20**(5): p. 1548-9.
5. Ebach, M.C. and C. Holdrege, *DNA barcoding is no substitute for taxonomy.* Nature, 2005. **434**(7034): p. 697.

6. Ebach, M.C. and C. Holdrege, *More taxonomy, not DNA barcoding.* Bioscience, 2005. **55**(10): p. 822-823.

7. Lipscomb, D., Platnick, M. and Wheeler, Q., *The intellectual content of taxonomy: a comment on DNA taxonomy.* TRENDS in Ecology and Evolution, 2003. **18**: p. 65-66.

8. Moritz, C. and C. Cicero, *DNA barcoding: promise and pitfalls.* PLoS Biol, 2004. **2**(10): p. e354.

9. Seberg, O. and G. Petersen, *How many loci does it take to DNA barcode a crocus?* PLoS One, 2009. **4**(2): p. e4598.

10. Seberg, O., Humphries, CJ, Knapp, S, Stevenson DW Perersen et al., *Shortcuts in systematics? A commentary on DNA based taxonomy.* Trends in Ecology and Evolution, 2005. **18**: p. 63-65.

11. Schindel, D.E. and S.E. Miller, *DNA barcoding a useful tool for taxonomists.* Nature, 2005. **435**(7038): p. 17.

12. Francis, C., Borisenko, AV, Ivanova, NV, Eger, JL, Lim, BK, Guillen-Servent, A, Kruskop, SV, Mackie, I, Hebert, PD, *The role of DNA barcodes in understanding and conservation of mammal diversity in southeast Asia.* PLoS One, 2010. **5**(9).

13. DeSalle, R., M.G. Egan, and M. Siddall, *The unholy trinity: taxonomy, species delimitation and DNA barcoding.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1905-16.

14. Dasmahapatra, K.K. and J. Mallet, *Taxonomy: DNA barcodes: recent successes and future prospects.* Heredity, 2006. **97**(4): p. 254-5.

15. Nguyen, H.D. and K.A. Seifert, *Description and DNA barcoding of three new species of Leohumicola from South Africa and the United States.* Persoonia, 2008. **21**: p. 57-69.

16. Yoshitake, H., Kato, T, Jinbo, U, Ito, M, *A new Wagnerinus (Coleoptera : Curculionidae) from northern Japan: Description including a DNA barcode.* Zootaxa, 2008(1740): p. 15-27.

17. Hirose, M. and E. Hirose, *DNA barcoding in photosymbiotic species of Diplosoma (Ascidiacea: Didemnidae), with the description of a new species from the southern Ryukyus, Japan.* Zoolog Sci, 2009. **26**(8): p. 564-8.

18. Ahrens, D., Monaghan, MT, Vogler, AP, *DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafers (Coleoptera : Scarabaeidae).* Molecular Phylogenetics and Evolution, 2007. **44**(1): p. 436-449.

19. Baldwin, C., Mounts, JH, Smith, DG, Weigt, LA, *Genetic identification and color descriptions of early life-history stages of Belizean Phaeoptyx and Astrapogon (Teleostei: Apogonidae) with Comments on identification of adult Phaeoptyx.* Zootaxa, 2009(2008): p. 1-22.

20. Greenstone, M., Rowley, DL, Heimbach, U, Lundgren, JG, Pfannenstiel, RS, Rehner, SA, *Barcoding generalist predators by polymerase chain reaction: carabids and spiders.* Mol Ecol, 2005. **14**(10): p. 3247-66.

21. Paquin, P. and M. Hedin, *The power and perils of 'molecular taxonomy': a case study of eyeless and endangered Cicurina (Araneae : Dictynidae) from Texas caves.* Molecular Ecology, 2004. **13**(10): p. 3239-3255.

22. Webb, K., Barnes, DKA, Clark, MS, Bowden, DA, *DNA barcoding: A molecular tool to identify Antarctic marine larvae.* Deep-Sea Research Part Ii-Topical Studies in Oceanography, 2006. **53**(8-10): p. 1053-1060.

23. Smith, P.J., S.M. Mcveagh, and D. Steinke, *DNA barcoding for the identification of smoked fish products.* Journal of Fish Biology, 2008. **72**(2): p. 464-471.

24. Wong, E.H.K. and R.H. Hanner, *DNA barcoding detects market substitution in North American seafood.* Food Research International, 2008. **41**(8): p. 828-837.

25. Dove, C., Rotzel, NC, Heacker, M, Weigt, LA, *Using DNA barcodes to identify bird species involved in birdstrikes.* Journal of Wildlife Management, 2008. **72**(5): p. 1231-1236.

26. Marra, P., Dove, CJ, Dolbeer, R, Dahlan, NF, Heacker, M, Whatton, JF, Diggs, NE, France, C, Henkes, GA, *Migratory Canada geese cause crash of US Airways Flight 1549.* Frontiers in Ecology and the Environment, 2009. **7**(6): p. 297-301.

27. Waugh, J., Evans, MW, Millar, CD, Lambert, DM, *Birdstrikes and barcoding: can DNA methods help make the airways safer?* Molecular Ecology Resources, 2010. **9999**(9999).

28. Skevington, J.H., C. Kehlmaier, and G. Stahls, *DNA Barcoding: Mixed results for big-headed flies (Diptera : Pipunculidae).* Zootaxa, 2007(1423): p. 1-26.

29. Ball, S., Hebert, PDN, Burian, SK, Webb, JM, *Biological identifications of mayflies (Ephemeroptera) using DNA barcodes.* Journal of the North American Benthological Society, 2005. **24**(3): p. 508-524.

30. Ball, S., Armstrong, KF, *Using DNA barcodes to investigate the taxonomy of the New Zealand sooty beech scale insect.* DoC Research and Development 2007. **287**: p. 5-14.

31. Chang, C.H., R. Rougerie, and J.H. Chen, *Identifying earthworms through DNA barcodes: Pitfalls and promise.* Pedobiologia, 2009. **52**(3): p. 171-180.

32. Chantangsi, C., Lynn, DH, Brandl, MT, Cole, JC, Hetrick, N, Ikonomi, P, *Barcoding ciliates: a comprehensive study of 75 isolates of the genus Tetrahymena.* Int J Syst Evol Microbiol, 2007. **57**(Pt 10): p. 2412-25.

33. Chaves, A., Clozato, CL, Lacerda, DR, Sari, EHR, Santos, FR, *Molecular taxonomy of brazilian tyrant-flycatchers (Passeriformes: Tyrannidae).* Molecular Ecology Resources, 2008. **8**(6): p. 1169-1177.

34. Costa, F., deWaard, JR, Boutillier, J, Ratnasingham, S, Dooh, RT, Hajibabaei, M, Hebert, PDN, *Biological identifications through DNA barcodes: the case of the Crustacea.* Canadian Journal of Fisheries and Aquatic Sciences, 2007. **64**(2): p. 272-295.

35. Dabert, J., R. Ehrnsberger, and M. Dabert, *Glaucalges tytonis sp n. (Analgoidea, Xolalgidae) from the barn owl Tyto alba (Strigiformes, Tytonidae): compiling morphology with DNA barcode data for taxon descriptions in mites (Acari).* Zootaxa, 2008(1719): p. 41-52.

36. De Ley, P., De Ley, IT, Morris, K, Abebe, E, Mundo-Ocampo, M, Yoder, M, Heras, J, Waumann, D, Rocha-Olivares, A, Burr, AHJ, Baldwin, JG, Thomas, WK, *An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**(1462): p. 1945-1958.

37. Elsasser, S., Floyd, R, Hebert, PDN, Schulte-Hostedde, AI, *Species identification of North American guinea worms (Nematoda: Dracunculus) with DNA barcoding.* Molecular Ecology Resources, 2009. **9**(3): p. 707-712.

38. Ferri, E., Barbuto, M, Bain, O, Galimberti, A, Uni, S, Guerrero, R, Ferte, H, Bandi, C, Martin, C, Casiraghi, M, *Integrated taxonomy: traditional approach and DNA barcoding for the identification of filarioid worms and related parasites (Nematoda).* Front Zool, 2009. **6**: p. 1.

39. Foley, D., Wilkerson, RC, Cooper, RD, Volovsek, ME, Bryan, JH, *A molecular phylogeny of Anopheles annulipes (Diptera: Culicidae) sensu lato: the most species-rich anopheline complex.* Mol Phylogenet Evol, 2007. **43**(1): p. 283-97.

40. Foottit, R., Maw, HEL, Von Dohlen, CD, Hebert, PDN, *Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes.* Molecular Ecology Resources, 2008. **8**(6): p. 1189-1201.

41. Foottit, R., Maw, HEL, Havill, NP, Ahern, RG, Montgomery, ME, *DNA barcodes to identify species and explore diversity in the Adelgidae (Insecta: Hemiptera: Aphidoidea).* Molecular Ecology Resources, 2009. **9**: p. 188-195.

42. Huang, J., Xu, Q, Sun, ZJ, Tang, GL, Su, ZY, *Identifying earthworms through DNA barcodes.* Pedobiologia, 2007. **51**(4): p. 301-309.

43. Kelly, R., Sarkar, IN, Eernisse, DJ, Desalle, R, *DNA barcoding using chitons (genus Mopalia).* Molecular Ecology Notes, 2007. **7**(2): p. 177-183.

44. Mikkelsen, N.T., C. Schander, and E. Willassen, *Local scale DNA barcoding of bivalves (Mollusca): A case study.* Zoologica Scripta, 2007. **36**(5): p. 455-463.

45. Nolan, L., Hogg, ID, Sutherland, DL, Stevens, MI, Schnabel, KE, *Allozyme and mitochondrial DNA variability within the New Zealand damselfly genera Xanthocnemis, Austrolestes, and Ischnura (Odonata).* New Zealand Journal of Zoology, 2007. **34**(4): p. 371-380.

46. Pfenninger, M., Nowak, C, Kley, C, Steinke, D, Streit, B, *Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic Chironomus (Diptera) species.* Mol Ecol, 2007. **16**(9): p. 1957-68.

47. Quiroz-Vazquez, P. and M. Elias-Gutierrez, *A New Species of the Freshwater Cladoceran Genus Scapholeberis Schoedler, 1858 (Cladocera: Anomopoda) from the Semidesert Northern Mexico, Highlighted by DNA Barcoding.* Zootaxa, 2009(2236): p. 50-64.

48. Radulovici, A.E., B. Sainte-Marie, and F. Dufresne, *DNA barcoding of marine crustaceans from the Estuary and Gulf of St Lawrence: a regional-scale approach.* Molecular Ecology Resources, 2009. **9**: p. 181-187.

49. Sinclair, C.S. and S.E. Gresens, *Discrimination of Cricotopus species (Diptera: Chironomidae) by DNA barcoding.* Bull Entomol Res, 2008. **98**(6): p. 555-63.

50. Holmes, B.H., D. Steinke, and R.D. Ward, *Identification of shark and ray fins using DNA barcoding.* Fisheries Research, 2009. **95**(2-3): p. 280-288.

51. Moura, T., Silva, MC, Figueiredo, I, Neves, A, Munoz, PD, Coelho, MM, Gordo, LS, *Molecular barcoding of north-east Atlantic deep-water sharks: species identification and application to fisheries management and conservation.* Marine and Freshwater Research, 2008. **59**(3): p. 214-223.

52. Smith, M., Woodley, NE, Janzen, DH, Hallwachs, W, Hebert, PDN, *DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera : Tachinidae).* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(10): p. 3657-3662.

53. Barrett, R., Hebert, PDN, *Identifying spiders through DNA barcodes.* Canadian Journal of Zoology-Revue Canadienne De Zoologie, 2005. **83**(3): p. 481-491.

54. Efe, M., Tavares, ES, Baker, AJ, Bonatto, SL, *Multigene phylogeny and DNA barcoding indicate that the Sandwich tern complex (Thalasseus sandvicensis, Laridae, Sternini) comprises two species.* Mol Phylogenet Evol, 2009. **52**(1): p. 263-7.

55. Elias-Gutierrez, M. and M. Valdez-Moreno, *A new cryptic species of Leberis Smirnov, 1989 (Crustacea, Cladocera, Chydoridae) from the Mexican semi-desert region, highlighted by DNA barcoding.* Hidrobiologica, 2008. **18**(1): p. 63-74.

56. Gomez, A., Wright, PJ, Lunt, DH, Cancino, JM, Carvalho, GR, Hughes, RN, *Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoan taxon.* Proc Biol Sci, 2007. **274**(1607): p. 199-207.

57. Hebert, P., Penton, EH, Burns, JM, Janzen, DH, Hallwachs, W, *Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator.* Proc Natl Acad Sci U S A, 2004. **101**(41): p. 14812-7.

58. Hebert, P., Stoeckle, MY, Zemlak, TS, Francis, CM *Identification of Birds through DNA Barcodes.* PLoS Biol, 2004. **2**(10).

59. Hogg, I.D. and P.D.N. Hebert, *Biological identification of springtails (Hexapoda : Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes.* Canadian Journal of Zoology-Revue Canadienne De Zoologie, 2004. **82**(5): p. 749-754.

60. Hogg, I., Stevens, MI, Schnabel, KE, Chapman, MA, *Deeply divergent lineages of the widespread New Zealand amphipod Paracalliope fluviatilis revealed using allozyme and mitochondrial DNA analyses.* Freshwater Biology, 2006. **51**(2): p. 236-248.

61. Hu, M., Gasser, RB, Chilton, NB, Beveridge, I, *Genetic variation in the mitochondrial cytochrome c oxidase subunit 1 within three species of Progamotaenia (Cestoda : Anoplocephalidae) from macropodid marsupials.* Parasitology, 2005. **130**: p. 117-129.

62. Johnson, S.B., A. Waren, and R.C. Vrijenhoek, *DNA barcoding of Lepetodrilus limpets reveals cryptic species.* Journal of Shellfish Research, 2008. **27**(1): p. 43-51.

63. Penton, E.H., P.D. Hebert, and T.J. Crease, *Mitochondrial DNA variation in North American populations of Daphnia obtusa: continentalism or cryptic endemism?* Mol Ecol, 2004. **13**(1): p. 97-107.

64. Witt, J.D.S., D.L. Threloff, and P.D.N. Hebert, *DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation.* Molecular Ecology, 2006. **15**(10): p. 3073-3082.

65. Nelson, L.A., J.F. Wallman, and M. Dowton, *Using COI barcodes to identify forensically and medically important blowflies.* Medical and Veterinary Entomology, 2007. **21**(1): p. 44-52.

66. Lambert, D., Baker, A, Huynen, L, Haddrath, O, Hebert, PDN, Millar, CD, *Is a large-scale DNA-based inventory of ancient life possible?* Journal of Heredity, 2005. **96**(3): p. 279-284.

67. Seabrook-Davison, M., et al., *Ancient DNA resolves identity and phylogeny of New Zealand's extinct and living quail (Coturnix sp.).* PLoS One, 2009. **4**(7): p. e6400.

68. Hu, M., Chilton, NB, Zhu, XQ, Gasser, RB, *Single-strand conformation polymorphism-based analysis of mitochondrial cytochrome c oxidase subunit 1 reveals significant substructuring in hookworm populations.* Electrophoresis, 2002. **23**(1): p. 27-34.

69. Ngarmamonpirat, C., Waikagul, J, Petmitr, S, Dekumyoy, P, Rojekittikhun, W, Anantapruti, MT, *Analysis of sequence variation in Gnathostoma spinigerum mitochondrial DNA by single-strand conformation polymorphism analysis and DNA sequence.* Parasitology International, 2005. **54**(1): p. 65-68.

70. Whiteman, N., Santiago-Alarcon, D, Johnson, KP, Parker, PG, *Differences in straggling rates between two genera of dove lice (Insecta: Phthiraptera) reinforce population genetic and cophylogenetic patterns.* Int J Parasitol, 2004. **34**(10): p. 1113-9.

71. Armstrong, K., Ball, SL, *DNA barcodes for biosecurity: invasive species identification.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**: p. 1813-1823.

72. Ball, S., Armstrong, KF, *DNA barcodes for insect pest identification: a test case with tussock moths (Lepidoptera : Lymantriidae).* Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 2006. **36**(2): p. 337-350.

73. Scheffer, S.J., M.L. Lewis, and R.C. Joshi, *DNA barcoding applied to invasive leafminers (Diptera : Agromyzidae) in the Philippines.* Annals of the Entomological Society of America, 2006. **99**(2): p. 204-210.

74. Day, J.C., T.I. Goodall, and R.J. Post, *Confirmation of the species status of the blackfly Simulium galeratum in Britain using molecular taxonomy.* Med Vet Entomol, 2008. **22**(1): p. 55-61.

75. De Astarloa, J., Mabragana, E, Hanner, R, Figueroa, DE, *Morphological and molecular evidence for a new species of longnose skate (Rajiformes: Rajidae: Dipturus) from Argentinean waters based on DNA barcoding.* Zootaxa, 2008(1921): p. 35-46.

76. Hulcr, J., Miller, SE, Setliff, GP, Darrow, K, Mueller, ND, Hebert, PDN, Weiblen, GD, *DNA barcoding confirms polyphagy in a generalist moth, Homona mermerodes (Lepidoptera : Tortricidae).* Molecular Ecology Notes, 2007. **7**(4): p. 549-557.

77. Landry, J.F., *Taxonomic review of the leek moth genus Acrolepiopsis (Lepidoptera : Acrolepiidae) in North America.* Canadian Entomologist, 2007. **139**(3): p. 319-353.

78. Fisher, B.L. and M.A. Smith, *A revision of Malagasy species of Anochetus mayr and Odontomachus latreille (Hymenoptera: Formicidae).* PLoS One, 2008. **3**(5): p. e1787.

79. Martinez, J.J., A. Zaldivar-Riveron, and A.G. Saez, *Reclassification of Bracon mendocinus, a gall-associated doryctine wasp, and description of a new closely related species of Allorhogas (Hymenoptera: Braconidae).* Journal of Natural History, 2008. **42**(41-42): p. 2689-2701.

80. Steinke, D., T.S. Zemlak, and P.D. Hebert, *Barcoding nemo: DNA-based identifications for the ornamental fish trade.* PLoS One, 2009. **4**(7): p. e6300.

81. Clare, E., Barber, BR, Fraser, EE, Adams, AM., *Genetic Identification of Prey Species in Guano: Spatial and Temporal Variation in Food Webs*, in *Third International Barcode of Life Conference*. 2009: Mexico.

82. Munch, K., Boomsma, W, Willerslev, E, Nielsen, R, *Fast phylogenetic DNA barcoding.* Philos Trans R Soc Lond B Biol Sci, 2008. **363**(1512): p. 3997-4002.

83. Rougerie, R., Decaens, T, Deharveng, L, Porco, D, James, SW, Chang, CH, Richard, B, Potapov, M, Suhardjono, Y, Hebert, PDN, *DNA barcodes for soil animal taxonomy.* Pesquisa Agropecuaria Brasileira, 2009. **44**(8): p. 789-802.

84. Kumar NP, R.A., Natarajan R, Jambulingam P., *DNA barcodes can distinguish species of Indian mosquitoes (Diptera: Culicidae).* J Med Entomol., 2007. **44**(1): p. 1-7.

85. Rivera, J. and D.C. Currie, *Identification of Nearctic black flies using DNA barcodes (Diptera: Simuliidae).* Molecular Ecology Resources, 2009. **9**: p. 224-236.

86. Stothard, J., Webster, BL, Weber, T, Nyakaana, S, Webster, JP, Kazibwe, F, Kabatereine, NB, Rollinson, D, *Molecular epidemiology of Schistosoma mansoni in Uganda: DNA barcoding reveals substantial genetic diversity within Lake Albert and Lake Victoria populations.* Parasitology, 2009: p. 1-12.

87. Feau, N., Vialle, A, Allaire, M, Tanguay, P, Joly, DL, Frey, P, Calland, BE, Hamelin, RC, *Fungal pathogen (mis-) identifications: A case study with DNA barcodes on Melampsora rusts of aspen and white poplar.* Mycological Research, 2009. **113**: p. 713-724.

88. de Leon, G.P. and A. Choudhury, *Parasite inventories and DNA-based taxonomy: lessons from helminths of freshwater fishes in a megadiverse country.* J Parasitol. **96**(1): p. 236-44.

89. Saunders, G.W., *Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1879-88.

90. Ward, R., Holmes, BH, White, WT, Last, PR, *DNA barcoding Australasian chondrichthyans: results and potential uses in conservation.* Marine and Freshwater Research, 2008. **59**(1): p. 57-71.

91. Lowenstein, J., Burger, J, Jeitner, CW, Amato, G, Kolokotronis, S, Gochfeld, M, *DNA barcodes reveal species-specific mercury levels in tuna sushi that pose a health risk to consumers.* Biology Letters, 2010.

92. Burgess-Herbert, S., Mwenja, i, Opyene, O, *From Science to Action: Taking DNA Barcoding to Battle Against the Bushmeat*, in *24th Annual Meeting of the Society for Conservation Biology*. 2010: Edmonton, Alberta, Canada.

93. Baker, C., Lento, GM, Cipriano, F, Dalebout, ML, Palumbi, SR, *Scientific whaling: source of illegal products for market?* Science, 2000. **290**(5497): p. 1695-6.

94. Baker, C., Palumbi, SR, *Which Whales Are Hunted? A Molecular Genetic Approach to Monitoring Whaling.* Science, 1994. **265**(5178): p. 1538-1539.

95. Ottesen, A., Ziobro, GC., *Use of "DNA Barcoding" to Detect Adulteration of Star Anise (Illicium verum) by Other Illicium Species*, in *12th Annual FDA Science Forum*. 2006: Washington, DC, Conference Center.

96. WWF. *Rhinoceros - Threats*. 2010; Available from: http://wwf.panda.org/what_we_do/endangered_species/rhinoceros/rhino_threats/.

97. Shepherd, C., Magnus, N., *Nowhere to hide: the trade in sumatran tiger*. 2004, TRAFFIC SOUTHEAST ASIA.

98. Reuters, *Genetic tests help track food web, climate change*. 2009.

99. Magnacca, K.N. and M.J. Brown, *Mitochondrial heteroplasmy and DNA barcoding in Hawaiian Hylaeus (Nesoprosopis) bees (Hymenoptera: Colletidae).* BMC Evol Biol, 2010. **10**: p. 174.

100. Ragupathy, S., Newmaster, SG, Murugesan, M, Balasubramaniam, V, *DNA barcoding discriminates a new cryptic grass species revealed in an ethnobotany study by the hill tribes of the Western Ghats in southern India.* Molecular Ecology Resources, 2009. **9**: p. 164-171.

101. Mitchell, A. *DNA barcoding (an emerging global standard for species identification) could revolutionize biosecurity diagnostics*. in *Global Biosecurity Conference*. 2010. Brisbaine.

102. Lakra, W. *DNA Barcoding Indian Fishes*. in *Third International DNA Barcoding Conference*. 2009. Mexico.

103. Valdez-Moreno, M. *Fish BOL Strategies Iin mesoAmerica*. 2009. Mexico.

104. Bogutskaya, N., Naseka, A, Mogilevsky, N. *Barcoding of the freshwater fish fauna of Russia: a pilot project*. in *Third InternationalBarcode of Life Conference*. 2009. Mexico.

105. Ward, R., Zemlak, TS, Innes, BH, Last, PR, Hebert, PDN, *DNA barcoding Australia's fish species.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**(1462): p. 1847-1857.

106. Ward, R., Anner, RH, Hebert, PDN., *The campaign to DNA barcode all fishes, FISH-BOL.* Journal of Fish Biology, 2009. **74**: p. 329-356.

107. Locke, S., McLaughlin, JD, Marcogliese, DJ., *DNA barcodes show cryptic diversity and a potential physiological basis for host specificity among Diplostomoidea (Platyhelminthes: Digenea) parasitizing freshwater fishes in the St. Lawrence River, Canada.* Molecular Ecology, 2010. **19**(13): p. 2813-2827.

108. Kerr, K., Stoeckle, MY, Dove, CJ, Weigt, LA, Francis, CM, Hebert, PDN, *Comprehensive DNA barcode coverage of North American birds.* Molecular Ecology Notes, 2007. **7**(4): p. 535-543.

109. Yoo, H., Eah, JY, Kim, JS, Kim, YJ, Min, MS, Paek, WK, Lee, H, Kim, CB, *DNA barcoding Korean birds.* Molecules and Cells, 2006. **22**(3): p. 323-327.

110. Park, H., Yoo, HS, Jung, G and Kim CB, *New DNA barcodes for identification of Korean birds.* Genes & Genomics, 2011. **33**: p. 91-95.

111. Kerr, K.C., Lijtmaer, D A, Barreira, A S, Hebert, PDN, Tubaro, P L, *Probing evolutionary patterns in neotropical birds through DNA barcodes.* PLoS One, 2009. **4**(2): p. e4379.

112. Johnsen, A., Rindal, E, Ericson, PGP, Zuccon, D, Kerr, KCR, Stoeckle, MY and Lifjeld, JT, *DNA barcoding Scandinavian Birds reveals divergent lineages in trans-Atlantic species.* J. Ornithol, 2010. **151**(3): p. 565-578.

113. *Codex Alimentarius*, C.A.C.a.J.F.W.F.S. Programme, Editor. 2010, Codex Alimentarius Commission and Joint FAO/WHO Food Standards Programme: Rome.

114. *Consortium for the Barcode of Life*. 2009 [cited 2009 11th November]; Available from: http://barcoding.si.edu/.

115. *Human Genome Project Time Line*. 2010; Available from: http://www.ornl.gov/sci/techresources/Human_Genome/project/timeline.shtml.

116. Jordan, E., *The Human Genome Project: Where Did It Come From, Where Is It Going?* Am. J. Hum. Genet., 1992. **51**: p. 1-6.

117. Zhaurova, K. *Genomes of Other Organisms: DNA Barcoding and Metagenomics*. 2008 [cited 2010 14th August]; Available from: http://www.nature.com/scitable/topicpage/genomes-of-other-organisms-dna-barcoding-and-662.

118. Shearer, T., Van Oppen, MJ, Romano, SL, Worheide, G, *Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria).* Mol Ecol, 2002. **11**(12): p. 2475-87.

119. Baker, A., Huynen, LJ, Haddrath, O, Millar, CD, Lambert, DM, *Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: the giant moas of New Zealand.* Proc Natl Acad Sci U S A, 2005. **102**(23): p. 8257-62.

120. Won, Y., Sivasundar, A, Wang, Y, Hey, J, *On the origin of Lake Malawi cichlid species: a population genetic analysis of divergence.* Proc Natl Acad Sci U S A, 2005. **102 Suppl 1**: p. 6581-6.

121. Vences, M., Thomas, M, Bonett, RM, Vieites, DR, *Deciphering amphibian diversity through DNA barcoding: chances and challenges.* Philos Trans R Soc Lond B Biol Sci, 2005. **360**(1462): p. 1859-68.

122. Rubinoff, D., Holland BS., *Between two extremes: Mitochondrial DNA is neither the panacea nor the nemesisof phylogenetic and taxonomic inference.* Systematic bBiology, 2005. **54**(6): p. 952-961.

123. Chase, M.W., et al., *A proposal for a standardised protocol to barcode all land plants.* Taxon, 2007. **56**(2): p. 295-299.

124. Kress, W., Wurdack, KJ, Zimmer, EA, Weigt, LA, Janzen, DH, *Use of DNA barcodes to identify flowering plants.* Proc Natl Acad Sci U S A, 2005. **102**(23): p. 8369-74.

125. Chase, M.W. and M.F. Fay, *Barcoding of Plants and Fungi.* Science, 2009. **325**(5941): p. 682-683.

126. Chase, M., Salamin, N, Wilkinson, M, Dunwell, JM, Kesanakurthi, RP, Haidar, N, Savolainen, V, *Land plants and DNA barcodes: short-term and long-term goals.*

Philosophical Transactions of the Royal Society B-Biological Sciences, 2005. **360**(1462): p. 1889-1895.

127. Hollingsworth, P.M., *DNA barcoding plants in biodiversity hot spots: progress and outstanding questions.* Heredity, 2008. **101**(1): p. 1-2.

128. Cowan, R., Chase, MW, Kress, WJ, Savolainen, V, *300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants.* Taxon, 2006. **55**(3): p. 611-616.

129. *A DNA barcode for land plants.* Proc Natl Acad Sci U S A, 2009. **106**(31): p. 12794-7.

130. Chen, S., Yao, H., et al., *Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species.* PLoS One, 2010. **5**(1): p. -.

131. *Barcode of Life website*. 2009 [cited 2009 15 October]; Available from: http://barcoding.si.edu/.

132. *Barcode of Life Data Systems*. 2009 [cited 2010 22nd July]; Available from: http://www.barcodinglife.org/.

133. *All Birds Barcoding Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.barcodingbirds.org/#.

134. *Fish Barcode of Life Initiative*. 2009 [cited 2009 18 October]; Available from: http://www.fishbol.org/.

135. *All Leps Barcode of Life*. 2009 [cited 2009 18 October]; Available from: http://www.lepbarcoding.org/.

136. *International Network for Barcoding Invasive and Pest Species (INBIPS)*. 2010; Available from: http://www.barcoding.si.edu/INBIPS.htm.

137. *Polar Boarcode of Life*. 2010; Available from: http://www.polarbarcoding.org/.

138. Watling, D., *'A Guide to the Birds of Fiji and Western Polynesia including American Samoa, Niue, Samoa, Tokelau, Tonga, Tuvalu and Wallis & Futuna'*. 2001, Naadi: Environmental Consultants (Fiji) 272.

139. Remigio, E.A. and P.D.N. Hebert, *Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships.* Molecular Phylogenetics and Evolution, 2003. **29**(3): p. 641-647.

140. Fazekas, A., Burgess, KS, Kesanakurti, PR, Graham, SW, Newmaster, SG, Husband, BC, Percy, DM, Hajibabaei, M, Barrett, SC, *Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well.* PLoS One, 2008. **3**(7): p. e2802.

Appendix A

Papers associated with this study

Papers 1 and 2

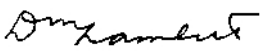
A review paper on DNA barcoding suggested by David Lambert and published in BioEssays in 2007, of which John Waugh was the sole author, forms the core of Chapter 1 of this thesis.[1] The information in this paper has been substantially updated for this chapter to include data up to and including the first half of 2010. John Waugh carried out all the research and prepared the manuscript. A commentary on the paper by Rob DeSalle, subsequently published in the same journal,[2] was followed by a response co-authored by John Waugh, Leon Huynen, Craig Millar and David Lambert.[3] David Lambert and Craig Millar suggested that this paper be written and substantially contributed to it. Leon Huynen contributed the moa data to the paper.

Auckland 16th November 2010

John Waugh

Leon Huynen

Craig Millar



David Lambert

References

1. Waugh, J., *DNA barcoding in animal species: progress, potential and pitfalls.* Bioessays, 2007. **29**(2): p. 188-197.

2. DeSalle, R., *Phenetic and DNA taxonomy; a comment on Waugh.* BioEssays, 2007. **29**(12): p. 1289-1290.

3. Waugh, J., Huynen, L, Millar, C, Lambert, D, *DNA barcoding of animal species - response to DeSalle.* Bioessays, 2008. **30**(1): p. 92-93.

DNA barcoding in animal species: progress, potential and pitfalls

John Waugh

Summary

Despite 250 years of work in systematics, the majority of species remains to be identified. Rising extinction rates and the need for increased biological monitoring lend urgency to this task. DNA sequencing, with key sequences serving as a "barcode", has therefore been proposed as a technology that might expedite species identification. In particular, the mitochondrial cytochrome *c* oxidase subunit 1 gene has been employed as a possible DNA marker for species and a number of studies in a variety of taxa have accordingly been carried out to examine its efficacy. In general, these studies demonstrate that DNA barcoding resolves most species, although some taxa have proved intractable. In some studies, barcoding provided a means of highlighting potential cryptic, synonymous or extinct species as well as matching adults with immature specimens. Higher taxa, however, have not been resolved as accurately as species. Nonetheless, DNA barcoding appears to offer a means of identifying species and may become a standard tool. *BioEssays* 29:188–197, 2007.

Introduction

The identification and characterisation of living things is fundamental to biological science. Modern taxonomy, with its origins in the mid 18th century, has described about 1.7 million species.[1] In addition, using morphological and behavioural techniques, much has been learnt about the relationships of living things with each other.

Allan Wilson Centre for Molecular Ecology and Evolution, Institute of Molecular BioSciences, Massey University, Private Bag 102 904, North Shore Mail Centre, Auckland, New Zealand.
E-mail: w.j.waugh@massey.ac.nz
DOI 10.1002/bies.20529
Published online in Wiley InterScience (www.interscience.wiley.com).

Abbreviations: BOLD, Barcode of Life Data Systems; CBOL, Consortium for the Barcode of Life; GBIF, Global Biodiversity Information Facility; K2P, Kimura 2 parameter; mtDNA, Mitochondrial DNA; NCBI, National Center for Biotechnology Information; PCR, Polymerase Chain Reaction; ToL, Tree of Life.

Unsurprisingly, larger animals have generally been the first to be described, while many smaller organisms remain unknown to science.[2] For example, it is thought that fewer than 10% of the vertebrates are yet to be identified but within the nematodes a bafflingly large number of species may exist, the vast majority of which has not been identified.[3] However, even among larger animals, there are doubts about species identification. The African elephant, long thought to be a single species, is now the subject of a debate based on nuclear and mitochondrial genomes over whether it incorporates two separate species.[4–6]

The earth's biota may contain between 10 and 100 million eukaryotic species.[7,8] The identification of numbers in this range represents an insurmountable workload for taxonomists using current methods. Even allowing for improvements in communications and the impact of the internet, the task is overwhelming. Moreover, phenotypic plasticity and genotypic variation in the features used for identification easily lead to identification errors and cryptic species or differing life stages can add to the confusion.[9] To compound matters, the task of cataloguing extant species is lent urgency by currently observed mass extinctions that are widely believed to be anthropogenic in origin.[10,11]

Field biologists, faced with the reality of species diversity, recognise the inadequacy of their own ability to access what *is* known about the biota, let alone what is not. These problems also impact upon people working in other areas such as combating the trade in endangered species, monitoring fisheries, identifying and controlling the spread of pest species or disease, identifying extinct lineages and regulating the movement of biological material around the world.[12–15]

Clearly there is a need to accelerate and simplify the processes of identification involved and, because of the scale of the problem, new methods will have to be employed. In addition, as more species are described, accessing the enlarged pool of taxonomic knowledge will become even more problematic.

Recent developments in DNA-sequencing technology have introduced the possibility of using variations in short sequences of DNA as labels for species in a process that has become known as DNA barcoding. The concept has already gained considerable acceptance among those working with species refractory to morphological identification such as viruses,[16] bacteria,[17,18] protists[19] and Rhodophyta.[20]

BioEssays 29:188–197, © 2007 Wiley Periodicals, Inc.

However, it is apparent that, since morphological techniques are difficult to access and apply without considerable training, some more rapid system of species identification is required for all taxa. This has led to the formation of the Consortium for the Barcode of Life (CBOL),[21] which aims to provide such a DNA barcode for every species on the planet (Box 1).

This review examines the progress, potential and pitfalls of DNA barcoding in animal species.

The DNA-barcoding process

A DNA barcode is a short sequence of nucleotides taken from an appropriate part of an organism's genome that is used to identify it at species level. Intraspecific variation in this sequence is an order of magnitude less than that observed interspecifically and this provides the means by which species are differentiated. It is not part of a DNA taxonomy nor is it a tool for phylogenetic reconstruction. It simply provides a means of linking sample specimens directly to existing voucher specimens and taxonomical information.

Central to the efficacy of DNA barcoding is the selection of a suitable segment of DNA. Its mutation rate must be slow enough so that intraspecific variation is minimised but sufficiently rapid to highlight interspecific variation.[9] It must be relatively easy to collect and should have as few insertions or deletions as possible to facilitate sequence alignment.

Mitochondrial DNA (mtDNA) offers several advantages over nuclear DNA. According to Drake's observation,[22] the rate of DNA mutation is inversely related to the size of the genome. Hence, nuclear DNA undergoes relatively slow mutation compared with mtDNA and, for this reason, would require a much longer nucleotide sequence than is necessary with mtDNA in order to provide a barcode capable of differentiating species.

In animals, mtDNA occurs as a single double-helical circular molecule containing 13 protein-coding genes, 2 ribosomal genes, a non-protein-coding control region, and several tRNAs (Fig. 1).[23] Each mitochondrion contains several such circular molecules and, therefore, several complete sets of mitochondrial genes. Furthermore, each cell has several mitochondria. Thus, when sample tissue is limited, the mitochondrion offers a relatively abundant source of DNA.

Cytochrome *c* oxidase is a large transmembrane protein found in the mitochondrion, which is highly conserved across

Box 1

The Consortium for the Barcode of Life (CBOL) is an international organisation devoted to developing DNA barcoding as a global standard in taxonomy. It comprises more than 120 member organisations from 45 countries and includes museums, herbaria, zoos, research organisations, governmental and intergovernmental agencies as well as other organisations involved in taxonomic research and biodiversity issues. Members agree to submit their DNA barcode sequences and voucher specimen data to a public database. CBOL was launched in May 2004 and is overseen by an executive committee that reports to the member organisations. It has five working groups to develop particular aspects of DNA barcoding, a Scientific Advisory Board and a small Secretariat Office to conduct its business. Within the auspices of CBOL a number of initiatives have been established including the All Birds Barcode Initiative, the Fish Barcode Initiative, the All Leps (Lepidoptera) Barcode Initiative and the International Network for Barcoding Invasive and Pest Species. Another group is exploring the barcoding of endangered vertebrates. Each of these initiatives aims to obtain DNA barcodes for every species within its group. The Canadian Centre for DNA Barcoding (a member of CBOL) oversees a website for Barcode of Life Data Systems (BOLD) that permits the uploading of sequences from the 5′ region of the COI gene and returns a species-level identification when one is possible. At present the site has more than 165,000 sequences from almost 20,000 species and these numbers are increasing steadily. The site also permits a variety of forms of data analysis for submitted sequences. CBOL works in cooperation with a number of other organisations including the Global Biodiversity Information Facility (GBIF), National Centre for Biotechnology Information (NCBI) and many taxanomic communities and web based projects.

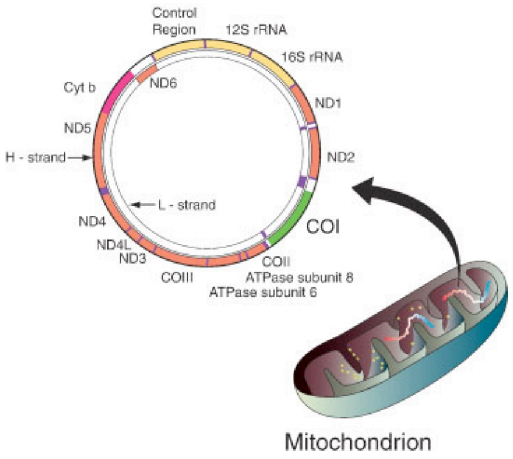


Figure 1. A diagrammatic representation of a mitochondrial genome based on complete mitochondrial DNA sequences from a variety of bird species.

species that employ oxidative phosphorylation for metabolism. It functions as the terminal electron acceptor in the respiratory chain, catalysing the reduction of oxygen to water and pumping protons across the membranes of the cristae.[23,24] The protein comprises several subunits of nuclear origin and three subunits synthesised in the mitochondrion. The mitochondrial subunits are known as subunits I, II and III. Cytochrome *c* oxidase subunit I (COI), the catalytic subunit of the enzyme, is predominantly imbedded in the membrane of the mitochondrial crista (Fig. 2). This structure would indicate a significant level of structural and functional constraint. However, the nucleotides of the gene that codes for it show sufficient variation to differentiate between species. Conversely, intraspecific variation in this gene is generally <10% of that observed between species. Moreover, insertions and deletions are rare.[3]

Recent studies associated with CBOL have generally selected a 648 bp segment of the COI gene, starting from the 5′ end, to generate a suitable barcode (Fig. 2).[21]

Having selected an appropriate segment of DNA for analysis, it must first be extracted from the sample specimen and amplified using the polymerase chain reaction (PCR). The amplified segment of the COI gene is sequenced and this sequence, the "barcode", is then matched with existing barcodes or material from voucher specimens. The Kimura 2-parameter (K2P) genetic distance correction is used to quantify sequence divergences among individuals because it is the most effective model when distances are low, as is the case with COI barcoding.[9]

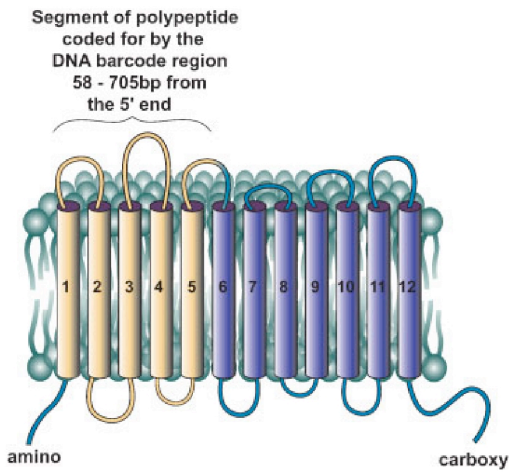


Figure 2. The predicted transmembrane structure for cytochrome *c* oxidase subunit 1 (COI). The area highlighted in yellow that includes five of the twelve transmembrane regions is coded for by the sequence designated by CBOL as the DNA barcode region.

The efficacy of DNA barcoding

To date, the literature contains a considerable number of fully published studies of animals in which DNA barcoding, using part of the COI gene, has been assessed[9,25–36] or used to help resolve taxonomic ambiguity[15,32–42] or used as part of a taxonomic description[43] (Table 1). DNA barcoding was employed to resolve species within narrow taxonomic groupings in some of the studies[15,25–27,31–36,40–43] or to identify higher taxa from wider assemblages of animals in others.[9,25,30] In addition, DNA-sequencing technology has been used for identifying organisms from other Kingdoms including plants,[44,45] bacteria,[17,18] protists[19] and viruses;[16] however, these are not reviewed here.

Although minor variations in protocols occurred between studies, broadly similar methods were used in each. DNA extracted from tissue samples was prepared using a variety of standard protocols. The sequence blocks of DNA used ranged in length from 350 bp to ~1000 bp. COI profiles were then generated using automated sequencers and these profiles were compared both within and between species as well as between higher taxa in some studies.[9,25,30]

Species

DNA barcoding provided a high degree of taxonomic resolution (>95%) for most species examined in the studies reviewed.[9,25–28,30,32–36,40–42] Each species had a unique nucleotide sequence at COI with only slight intraspecific K2P divergence. For example, in one study of 13,320 congeneric species pairs,[28] intraspecific variation was usually less than 1% and rarely more than 2%, while mean interspecific divergence was 11.3%. This difference between intra- and interspecific divergences at COI was also observed in arachnid,[25,32,40] lepidopteran,[9,33,36] (Fig. 3), dipteran,[34] avian,[26] fish[31] and Collembola[27] species (Table 1).

DNA barcoding was generally successful when used for identifying immature specimens,[40] extinct species[15] and individual species at differing stages in their life cycles.[32,36,40] Furthermore, possible cryptic species were identified in several studies.[25–27,31,34,36–39,41]

Identification difficulties caused by morphological differences between instars in cave-dwelling spiders (*Cicurina* spp.) were resolved using this technique, which also aided identification of populations where adult specimens are extremely rare.[40] Potential species identified in this study correspond closely to *a priori* species hypotheses except in the case of *C. caliga* and *C. hoodensis*, which contain COI sequences so similar that it is hypothesised they are synonymous. Furthermore, COI barcoding of two other species, *C. madla* and *C. vespera* (a species known only from one female specimen), suggested possible synonymy and indicated a need for further evaluation. This utility was also observed when 10 distinct caterpillars were linked to their morphologically similar adults.[36] While in another study, COI

Table 1. Design details and major results from selected studies that utilised part of the cytochrome *c* oxidase subunit 1 gene to provide a DNA barcode for identifying animal taxa

Reference	Sequence length	No. of taxa	% identified	K2P % intra vs interspecific COI variation	Comment
Barrett et al. 2005[23]	600bp	203 arachnid spp.	100	1.4 vs 16.4	Mean intra- and interspecific nucleotide divergences did not overlap except in the case of probable cryptic species.
Brown et al. 2003[41]	648bp	1 Lepidopteran sp. *Gnetom gnemon*	100		COI was used as a taxonomic descriptor for a new species.
Greenstone et al. 2005[30]	439bp	32 Carabidae spp. and 39 Araneae spp.			COI used to identify species and link different life stages.
Hajibabaei et al. 2006[31]	311–612bp	521 Lepidopteran spp.	97.9%	0.17–0.46 vs 4–6	Morphologically distinct sympatric species from three families identified.
Hebert et al. 2003a[7]	658bp	7 animal phyla	96.4		The efficacy of COI at identifying species, orders and phyla assessed.
		8 insect orders	100		
		200 Lepidopteran spp.	100	Lepidoptera 0.25 vs 6.84	
Hebert et al. 2003b[26]	>400bp	2238 Annelida, Arthropoda, Chordata, Cnidaria, Echinodermata, Mollusca, Nematoda, Platyhelminthes, and minor phyla	>98	Overall, usually <2 vs 11.3	The efficacy of COI at identifying species from eight major and several minor phyla plus a variety of arthropod classes was assessed. Cnidarians showed less COI variation between species than all other taxonomic groups, 94.1% vs 1.9% showing <2% K2P between spp. ($p < 0.0001$).
Hebert et al. 2004a[34]	648bp	10 Lepidopteran spp.	100		Ten new taxa identified. Different life stages matched for species.
Hebert et al. 2004b[24]	648bp	260 Avian spp.	100	0.43 vs 7.93	Four possible cryptic species identified.
Hogg et al. 2004[25]	710bp	19 Collembola spp.	100	0.78 vs 19.0	Produces high resolution in Collembola species. Possible cryptic species identified.
Hu et al. 2002[35]	≤530bp	7 Hookworm spp.	—		Three of the seven species appeared to be possible species complexes based on intraspecific COI variation.
Hu et al. 2005[36]	≈450bp	3 *Progamotaenia* spp. (Platyhelminthes)			Variation at COI suggests that all three species are species complexes.
Lambert et al. 2005[13]	596bp	10 Moa spp.	100		COI used to identify extinct species. Possible species synonymy highlighted.
Lorenz et al. 2005[27]	727bp	56 primates		0.011 vs -	Problems with taxon specific patterns of 'universal primer' failure. Taxon specific primers developed.
Ngarmamonpirat et al. 2005[37]	450bp	1 *Gnathastoma spinigerum* (hookworm)	—		COI barcode variation did not match the morphological variation observed in 3rd larval stage of this hookworm.
32 Paquin et al. 2004[38]	≈1000bp	23 *Cicurina* spp.(Arachnida)	≈100	1.09 vs 7.12	Immature specimens identified to species level. Possible species synonymy identified.
Penton et al. 2004[39]	709bp	2 *Daphnia* spp.(Crustacea)	100		Identification of morphologically cryptic species with overlapping distribution.
Remigio et al. 2003[28]	672bp	70 Gastropod spp.	—		COI used to identify species and higher taxonomic relationships. Insertion or deletions more common in COI in this taxonomic group.
Smith et al. 2006[32]	658bp	32 Dipteran spp.	100	0.17 vs 5.78	Fifteen cryptic species found using COI.
Vences et al. 2005[33]	550–650bp	9 Mantellid frog spp. 4 *Aneides* spp. (salamanders)		5.4 (mantellid frogs) and 4.3 (salamanders) vs 20.7 and 13.5	Found high intraspecific variability (7–18%). The use of mitochondrial 16S rRNA gene to supplement COI suggested.
Ward et al. 2005[29]	655bp	207 fish	100	0.39 vs 9.93	Efficacy of COI at identifying species and higher taxonomic relationships assessed. Possible cryptic species identified.
Whiteman et al. 2004[40]	379bp	2 Lice spp. (Insecta)	100		Study of bird parasites with similar morphology that can be identified using COI.

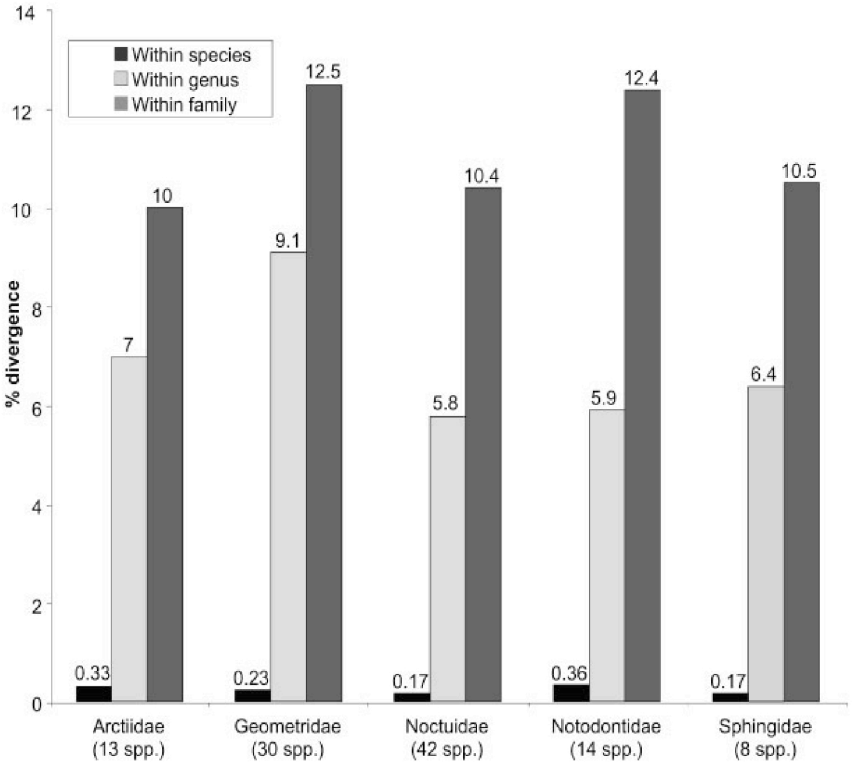


Figure 3. Nucleotide divergence in a 617 bp segment of the COI gene in five lepidopteron families at species, genus and family level. Data from Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003a Proc R Soc Lond B 270:313–322

barcoding helped to identify arachnids and carabids of different life stages including eggs, larvae or nymphs and pupae.[32]

In extinct taxa, where full taxonomic descriptions are difficult due to the lack of soft tissue, identification of species can be particularly problematical. However, 6 species of an extinct ratite bird, the New Zealand moa, were identified using DNA barcoding when 2.7% sequence divergence at COI was used as the intraspecific threshold. This increased to 10 species when a threshold of 1.24% was used.[15] The species identified were generally confirmed and supported by results from a larger study of moa mtDNA, in which 125 specimens had mitochondrial control region sequences analysed (Fig. 4).[46]

Potential cryptic species were identified using DNA barcoding among butterflies,[36] flies,[34] birds,[26] arachnids,[25] springtails,[27] within the species *Daphnia obtusa*[41] and in three groups of parasitic worms.[37–39]

Larval caterpillars with distinct colour patterning and food plants were linked with adults that are phenotypically very similar to each other.[36] Divergence at COI was considerable (mean K2P 2.76%; range 0.0–7.95%) and when caterpillar/adult morphology and food plants were mapped onto a neighbour joining tree of the COI divergence, 10 probable new species were revealed that showed covariance between morphological, molecular and ecological characteristics.

Another study found fifteen cryptic species of parasitoid flies that show high host-specificity within a group of what had been thought were three generalist species.[34]

COI-identified cryptic species were not limited to relatively obscure taxa. For example, among 260 North American bird species, K2P distances were 18-fold higher between species than within them; however, in four species (*Tringa solitaia, Sturnella magna, Cisthorus palustris* and *Vireo gilvus*), high intraspecific K2P distances suggested the presence of cryptic species.[26]

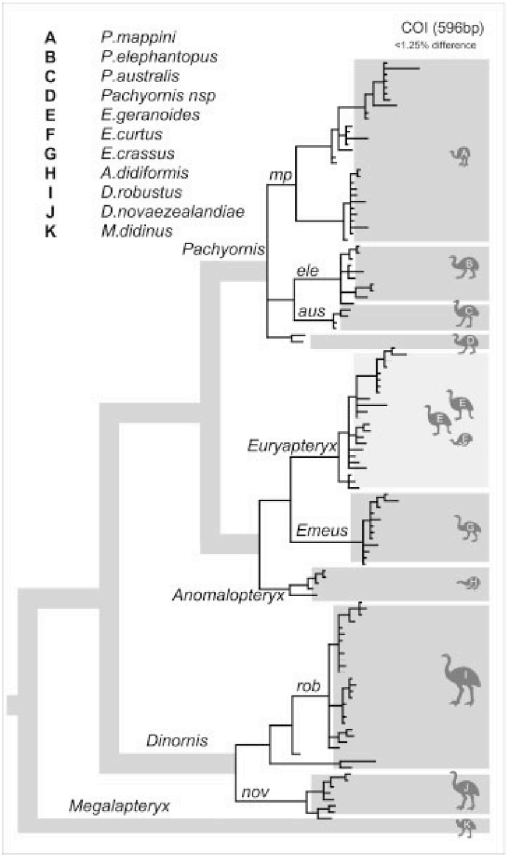


Figure 4. A comparison of the efficacy of mitochondrial control region (656 bp) with COI sequences (596 bp) at determining moa lineages.[13,44] The tree was constructed using control region mtDNA sequences. Dark grey boxes represent COI sequences that differ by less than 1.25%. Euryapteryx species (light grey box) appear to represent a complex of overlapping subgroups. The three lineages suggested by control region sequences for *D. robustus* and the two for *P. mappini* are reduced to two lineages using COI data.

Similarly, intraspecific K2P divergence averaged 1.4% among 37 arachnid species but in one species, *Latrodectus hesperus*, it was 3.6%.[25] This divergence was between northern and southern populations and suggested that they are probably separate species; a conclusion supported by the results of another study[47] that examined breeding and pheromones in these populations.

The presence of undescribed species was revealed in a study of springtails from the Canadian Arctic.[27] One species showed up to 13% intraspecific COI divergence compared with the <1% generally observed in species from this group. Likewise, wide sequence variation at COI was observed among members of the species *Daphnia obtusa* collected from 33 North American sites,[41] indicating that this may in fact be two species; one confined to the east and the other more broadly distributed.

The efficacy of DNA barcoding at revealing cryptic species was further demonstrated in three studies of parasitic worms.[37–39] Analysis of COI fragments in each study found high intraspecific COI variation among tapeworm and hookworm species. Two of the three studies[37,38] suggested that this might be due to the presence of cryptic species. In the third,[39] the authors conclude that there was little COI and morphological covariance; however, the presence of cryptic species was not precluded.

Problematic taxa

Some taxonomic groups were not readily resolved to species level. These included benthic Cnidarians,[28] two groups of amphibians[35] and some Gastropod species.[30]

There was little COI divergence between species of benthic Cnidarians, with 94.1% of species pairs showing <2% divergence versus 1.9% of species pairs from all other phyla in this large study.[28] According to Shearer et al,[48] Cnidarians show particularly slow mutation rates in their mitochondrial DNA when compared with other taxa and this may impair the resolving power of COI in this group.

Conversely, high intraspecific variation at COI of up to 18% was observed in two amphibian groups (mantellid frogs and salamanders).[35] This overlapped the interspecific variation, making species delineation difficult. Furthermore, variability within the mitochondrial genome of these taxa meant that a mix of primers was required to isolate the required segment. Nonetheless, COI sequences were able to correctly identify species including disparate geographic variants.

Gastropods also proved refractory to COI identification in a study of 70 species aimed at establishing phylogenetic relationships as well as species identification.[30] Insertions and deletions found in the COI gene of two subclasses, Heterobranchia and Patellogastropoda, complicated alignment.

These results with gastropods conformed to the general observation that where DNA barcoding was used to resolve higher taxa, results proved more equivocal. In one study,[25] 87% of genera that contained several species and 67% of families that contained several genera formed cohesive COI groups. Another study correctly assigned 96.4% of 55 taxa to phyla and 100% of 50 taxa to ordinal level (Table 1).[9]

Discussion

There is currently an acknowledged biodiversity crisis of anthropogenic origin.[10,11] It is the result of the destruction of habitats and unsustainable harvesting of natural products. Many species are becoming extinct without our ever having registered their existence. The fact that the majority of

Eukaryotes remain unknown to science has focussed attention on the overwhelming task that taxonomists face in trying to identify them. However, moments of crisis can precipitate novel and creative solutions. To quote Quentin Wheeler[49] "within the field of taxonomy, there is presently a conjunction of new theories, technological advances and urgent needs." The first two elements of this conjunction represent opportunities, the third, according to Plato, "is the mother of invention." Into this environment comes DNA barcoding; not, perhaps, a solution to the crisis but a tool that may help in its resolution.

DNA barcoding claims neither to replace taxonomy nor to reconstruct phylogenies. It does not absolve governments and funding bodies of the need to inject new life into the science of taxonomy. However, it may prove a useful tool for taxonomists and the many other agencies and individuals interested in species identification.

In general, DNA barcoding was effective at resolving species in the studies reviewed. However, resolution of higher taxonomic groups was less effective. Indeed, to paraphrase Greenstone et al., COI barcoding is a diagnostic tool for identifying animal species and cannot be expected to serve double duty as a character for deeper phylogenetic reconstruction.[32]

DNA barcoding showed utility at identifying potential candidates for taxonomic description by highlighting possible cryptic species among those already identified. It drew attention to a number of species, both extant and extinct, that may be synonymous and provided a means of identifying species regardless of life stage or maturity. Furthermore, data from a study using other mtDNA sequences in extinct bird species was confirmed using DNA barcoding.

Any system that is designed to identify species needs to tackle the issue of what a species is. There are many concepts including typological, phylogenetic, morphological, biological, isolation and mate recognition, to name a few. Each has its own perspective, none is universally accepted. However, it is beyond the scope of this paper to debate the relative merits of these species concepts. Moreover, DNA barcoding does not answer philosophical or ontological questions about the nature of species or higher taxa, it simply identifies highly correlated sequences, which, if it works, are derived from highly correlated (i.e. conspecific) individuals. Furthermore, once species limits have been defined, DNA barcoding may raise unexpected questions about what those limits are. The results presented in this review suggest that, empirically, DNA barcoding accurately identified species in >95% of cases.

Some questions relating to this technique remain to be answered. For example, what is a DNA barcode? Although CBOL states that "only cytochrome *c* oxidase 1 is approved as a barcode region, defined relative to the mouse mitochondrial genome as the 648 bp region that starts at position 58 and stops at position 705",[21] very few of the sequences lodged with GenBank are of this precise segment of the gene. A large number are considerably shorter and of those longer than 648 bp, many do not fully overlap the specified segment. This variability complicates and reduces the power of large-scale analyses of these data.

Another question arises: is it necessary to use many barcodes from differing genes in order to identify the broadest range of species? COI does not work for other Kingdoms and Nielsen et al.[50] suggest that the weakest aspect of DNA barcoding is that no single gene will always be invariant within species but different between species. The results of some of the problematic taxa bear this out. They further suggest that there is a need for statistical protocols to assess whether a sample barcode is sufficiently similar to a known barcode to justify species assignment.[50] Thus, issues of standardisation need to be addressed if barcoding is to achieve the rigour required of an enduring contribution to science.

In addition, taxa that are undergoing rapid speciation show little interspecific COI divergence, thus compromising the resolving power of COI barcoding. For example, New Zealand moas are thought to have undergone rapid speciation prior to extinction,[46] which may account for the relatively low interspecific K2P distances observed in this group.[15] The same may be true of the cichlid fish in Africa, a group known for its rapid speciation.[51] The converse may be true for species that have not undergone recent speciation events. Moreover, some groups, such as the Cnidarians, show slow mitochondrial DNA-sequence evolution resulting in negligible interspecific variation.

Critics of DNA barcoding suggest that it is unscientific because it does not set out to test hypotheses, that it generates information not knowledge.[52,53] However, arguably, any experiment generates information that requires interpretation. Moreover, barcoding tests the hypothesis that species can be identified using this technique and in future may be a source of data that will generate other hypotheses. Furthermore, it is possible to make similar comments about the invention of the microscope, perhaps the most-important scientific invention of the past millennium.

A number of people remain sceptical of the utility and efficacy of DNA barcoding. There are those who fear that the promoters of barcoding are seeking to replace conventional taxonomy.[52–55] These fears have been fuelled by an enthusiasm for DNA taxonomy in some quarters outside CBOL.[56] Some take exception to the use of the term barcode on the basis that it suggests that "each species has a fixed and invariant characteristic like the barcode on a supermarket product".[54] They also express reservations that sufficient numbers of congeneric species have been sampled or that those samples come from a wide-enough distribution to make generalisations about the efficacy of COI barcoding.[54] In addition, there are concerns that any attempt at producing a universal system for identifying species entailing a centralised database may be seen by third world countries as an attempt

by wealthier nations to monopolise taxonomic information.[54] It is also thought that any such system may be more authoritarian and will lack the flexibility inherent in the committee style consensus of existing botanical and zoological codes.

Proponents of barcoding respond that COI barcoding is not a substitute for taxonomy.[57,58] That it cannot be, since it is only by linking barcodes to fully described voucher specimens that the full power of the technique can be realised and that just as supermarket barcodes would be meaningless without the database of product details to which they are linked, so DNA barcoding requires a database of taxonomic information to which it links. They acknowledge that the barcode analogy is not an exact one but maintain that, with the current level of accuracy observed, COI barcoding has proved sufficiently discriminatory in trials to demonstrate considerable utility as a tool for differentiating species and, therefore, merits further investigation. With regard to the comments about the numbers of species tested using this method so far, it is only as the database of barcodes builds that the substance of these reservations will either be confirmed or otherwise.

Unquestionably, any progress that is to be made in accelerating species identification will be dependent on the use of new technologies and will employ accessible, easily searchable repositories of taxonomic information. Whether or not this is perceived as a threat by countries around the world will depend largely upon the sensitivity with which the process is managed. CBOL favours the open access approach of the Global Biodiversity Information Facility (GBIF)[59] and works in coordination with them.

Those who express concern that this may be an attempt to divest third world countries of information relating to their biotas, are also probably underestimating the power of modern technology to disseminate information. If the current open access approach is maintained, taxonomic information, much of which resides in relatively inaccessible Northern Hemisphere museums and collections, will become more accessible rather than the reverse. Thus, regions may be able to access and reclaim information relating to their indigenous biotas.

The methods by which species are named and described, however, will not be affected by DNA barcoding. This technique is not "a pretender to the taxonomic throne",[60] its principal utility is as a searchable label rather than as a contributing taxonomic feature. Thus, barcoding may serve to help inform the debate that generally surrounds species identification but is unlikely to undermine the flexibility of existing codes.

Museums around the world maintain large collections of plant and animal specimens and are, therefore, excellent sources of material for DNA barcoding. Their accumulated experience at curating these collections combined with a need for efficient cataloguing systems suggest the potential for a mutually beneficial relationship with CBOL. Furthermore, these institutions play a central role in the collection and description of new species and may, therefore, be primary beneficiaries of the development of new technologies for identifying known and unknown species. To date, several major museums and research institutions around the world have lent support to the consortium; however, some remain sceptical about the utility of a barcode system.

Evidence suggests that DNA barcoding can serve as a means of accessing taxonomic information and help in the identification of species. However, even if a complete barcode resource of the world's biota is produced, it cannot achieve its full potential unless the processes involved in obtaining barcodes from specimens and accessing the taxonomic information relating to them are simplified and streamlined so that they can be quickly carried out by relatively unskilled workers in a variety of locations. Those involved with CBOL envisage the development of a hand-held device to facilitate this. Although such a device does not represent an insuperable developmental problem, it requires both capital investment and the determination to ensure that it does reach production. However, taxonomy is not a science that tends to attract this level of funding.

The proliferation of unrelated web-based initiatives represents a possible impediment to the accretion of this investment. There are a large number of projects, unrelated to CBOL, that are attempting to collate taxonomic information either as phylogenetic trees or as catalogues of species. Some of these initiatives show clear areas of overlap, others are markedly dissimilar and have much to offer each other. An example of these is the Tree of Life (ToL) project[61] that provides a framework in which to electronically publish taxonomic information in a searchable database. It is a collaborative effort and currently consists of approximately 4000 web pages. Each page contains information about a particular taxonomic group and is linked hierarchically to other pages in the form of a phylogeny of life. The CBOL and ToL initiatives share links with the GBIF and are complimentary, with CBOL focused on species and ToL on higher classification. DNA barcodes may supply a key to rapidly linking specimens to ToL information and might expedite navigation around this large and growing database. ToL may provide taxonomic information to which DNA barcodes link. Symbiosis between these projects can facilitate the process of species identification and may also lead to as yet unenvisaged use of ToL information, thus, increasing the value of both initiatives.

Conclusion and outlook

DNA barcoding shows considerable potential as a system for identifying species that may allow users to link specimens to databases of taxonomic information as well as highlighting those species for which no data are yet available. It is not a thorough taxonomic description nor is it a tool for phylogenetic reconstruction. However, it may help speed the work of

taxonomists and others interested in species identification. To date, evidence from a number of studies largely confirms the feasibility of such a system.

The success of DNA barcoding will depend on the rationalisation and coordination of the many online species identification programmes, so that sufficient resources can be concentrated to develop rapid in situ sequencing. There is also a need for the problems associated with species refractory to COI identification to be resolved. Furthermore, the support of taxonomists, from around the world, is a vital prerequisite that may prove the most intractable obstacle faced by CBOL. However, with sufficient impetus and dedication, none of these problems are insoluble.

In the 270 years since Carl Linnaeus first published his *Systema Natura*, approximately 1.7 million species have been described. There may be as many as 100 million Eukaryotes that remain unknown to science. If that is true, at the present rate, it will take about 600 generations of scientists or 16,000 years before the job is complete. Clearly, systems for accelerating these processes are required. Use of electronic media can speed progress and help to reduce duplication or the loss of information that occurs with the death of each taxonomist but new methods are necessary if a breakthrough is to be made. Regardless of whether DNA barcoding is this breakthrough, it may prove to be a very useful taxonomic tool. Given the enormity of the task of identifying the world's biota and the many other potential tasks for which it might be employed, it would seem imprudent to ignore the promise of DNA barcoding.

Acknowledgments

I am grateful to David Lambert, Craig Millar, Jennifer Hay, Philip Smith, Gimme Walter and Andrew Dodd for their useful comments and for proof-reading this manuscript.

References

1. Stoeckle M. 2003. Taxonomy, DNA, and the Bar Code of Life. 2003. BioScience 53:2.
2. Blaxter ML. 2003. Counting angels with DNA. Nature 421:122–124.
3. Blaxter ML. 2004. The promise of a DNA taxonomy. Phil Trans R Soc Lond B 359:669–679.
4. Comstock KE, Georgiadis N, Pecon-Slattery J, Roca AL, Ostrander EA, et al. 2002. Patterns of molecular genetic variation among African elephant populations. Molecular Ecology 11:2489–2498.
5. Roca AL, Geordiadis N, O'Brien SJ. 2005. Cytonuclear genomic dissociation in African elephant species. Nature genetics 37:96–100.
6. Debruyne R. 2004. Contribution of molecular phylogeny and morphometrics to the systematics of African elephants. J Soc Biol 198:335–342.
7. May RM. 1988. How many species are there on Earth?. Science 241: 1441–1449.
8. Whitfield J. 2003. DNA barcodes catalogue animals. Nature. News Service. Available on URL http://www.nature.com/news/2003/030512/full/0305127.html. Accessed 11 December 2005.
9. Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003a. Biological identifications through DNA barcodes. Proc R Soc Lond B 270:313–322.
10. American Museum of Natural History (AMNH). 1998. National survey reveals biodiversity crisis. Press release, available on URL http://www.amnh.org/museum/press/feature/biohall.html. Accessed 24 May 2006.
11. Ananthaswamy A. 2004. Scene set for next mass extinction. New Scientist Mar 27.
12. Baker CS, Palumbi SR. 1994. Which whales are hunted? A molecular approach to monitoring whaling. Science 265:1538–1539.
13. Baker CS, Dalebout ML, Lavery S, Ross HA. 2003. www.DNA-surveillance: applied molecular taxonomy for species conservation and discovery. Trends in Ecol and Evol 18:271–272.
14. Besansky NJ, Severson DW, Ferdig MT. 2003. DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you. Trends Parasitol 19:545–546.
15. Lambert DM, Baker A, Huynen L, Haddrath O, Hebert PDN,, et al. 2005. Is a large-scale DNA-based inventory of ancient life possible? J Hered 96:1–6.
16. Mayo MA, Horzinek MC. 1998. A revised version of the International Code of Virus Classification and Nomenclature. Arch Virol 143:1645–1654.
17. Woese CR. 1987. Bacterial evolution. Microbiol Rev 51:221–271.
18. Unwin R, Maiden MCJ. 2003. Multi-locus sequence typing: a tool for global epidemiology. Trends Microbiol 11:479–487.
19. Cavalier-Smith T. 2004. Only six kingdoms of life. Proc R Soc Lond *B* 271:1251–1262.
20. Saunders GW. 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. Phil Trans R Soc Lond 360:1879–1888.
21. Barcode of Life website. http://www.barcodinglife.org/ . Accessed 3 May 2006.
22. Drake JW, Charlesworth B, Charlesworth D, Crow JF. 1998. Rates of spontaneous mutation. Genetics 148:1667–1686.
23. Research Collaboratory for Structural Bioinformatics (RCSB) Protein Databank (PDB). available on URL http://www.rcsb.org/pdb/cgi/explore.cgi?pdbId=2OCC . Accessed 19 May 2006.
24. Cytochrome c oxidase homepage. available on URL http://www.bioc.rice.edu/~graham/CcO.html . Accessed 19 May 2006.
25. Barrett RDH, Hebert PDN. 2005. Identifying spiders through DNA barcoding. Can J Zool 83:481–489.
26. Hebert PDN, Stoeckle MY, Zemiak TS, Francis CM. 2004b. Identification of birds through DNA barcodes. PLoS Biol 2004b 2:1657–1663.
27. Hogg ID, Hebert PDN. 2004. Biological identification of springtails (Collembola: Hexapoda) from the Canadian Arctic, using mitochondrial DNA barcodes. Can J Zool 82:749–754.
28. Hebert PDN, Ratnasingham S, deWaard JR. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc R Soc Lond B 270:1–4.
29. Lorenz JG, Jackson WE, Beck JC, Hanner R. 2005. The problems and promise of DNA barcodes for special diagnosis of primate biomaterials. Phil Trans R Soc Lond 360:1869–1877.
30. Remigio EA, Hebert PDN. 2003. Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships. Mol Phylogenet Evol 29:641–647.
31. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. Phil Trans R Soc Lond 360:1847–1857.
32. Greenstone MH, Rowley DL, Heimbach U, Lundgren JG, Pfannenstiel RS, et al. 2005. Barcoding generalist predators by polymerase chain reaction: carabids and spiders. Molecular Ecology 14:3247–3266.
33. Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN. 2006, DNA barcodes distinguish species of tropical Lepidoptera. PNAS 103: 968–971.
34. Smith MA, Woodley NE, Janzen DH, Hallwachs W, Hebert PDN. 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). PNAS 103:3657–3662.
35. Vences M, Thomas M, Bonnett RM, Vieites DR. 2005. Deciphering amphibian diversity through DNA barcoding: chances and challenges. Philosophical Transactions of the Royal Society B 360:1859–1868.
36. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004a. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. PNAS 101:14812–14817.

37. Hu M, Chilton NB, Zhu X, Gasser RB. 2002. Singe-strand conformation polymorphism-based analysis of mitochondrial cytochrome *c* oxidase subunit 1 reveals significant substructuring in hookworm populations. Electrophoresis 23:27–34.
38. Hu M, Gasser RB, Chilton NB, Beveridge I. 2005. Genetic variation in the mitochondrial cytochrome *c* oxidase subunit 1 within three species of *Progamotaenial* (Cestoda: Anoplocephalidae) from macropodid marsupials. Parasitology 130:117–129.
39. Ngarmamonpirat C, Waikagul J, Petmitr S, Dekumyoy P, Rojekittikhun W, et al. 2005. Analysis of sequence variations in *Gnathostoma spinigerum* mitochondrial DNA by single-strand conformation polymorphism analysis and DNA sequence. Parasitology International 4:65–68.
40. Paquin P, Hedin M. 2004. The power and perils of 'molecular taxonomy': a case study of eyeless and endangered *Circurina* (Araneae: Dictynidae) from Texas caves. Molecular Ecol 13:3239–3255.
41. Penton EH, Hebert PDN, Crease TJ. 2004. Mitochodrial DNA variation in North American populations of *Daphnia obtuse*: continentalism or cryptic endemism? Molecular Ecology 13:97–107.
42. Whiteman NK, Santiago-Alarcon D, Johnson KP, Parker PG. 2004. Differences in straggling rates between two genera of dove lice (Insect: Phthiraptera) reinforce population genetic and cophylogenetic patterns. Int J Parasitol 34:1113–1119.
43. Brown JW, Miller SE, Horak M. 2003. Studies on New Guinea moths. 2. Description of a new species of Xenothictis Meyrick (Lepidoptera: Tortricidae: Archipini). Proc Entomol Soc Wash 105:1043–1050.
44. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. PNAS 102:8369–8374.
45. Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, et al. 2005 The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92:142–165.
46. Baker AJ, Huynen LJ, Haddrath O, Millar CD, Lambert DM. 2005. Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: the giant moas of New Zealand. PNAS 102: 8257–8262.
47. Kasumovic MM, Andrade MCB. 2004. Discrimination of airborne pheromones by mate searching male western black widow spiders (*Latrodectus Hesperus*): species- and population-specific responses. Can J Zool 82:1027–1034.
48. Shearer TL, Van Oppen MJH, Romano SL, Wörheide G. 2002. Slow mitochondrial DNA sequence evolution in the Anthazoa (Cnidaria). Molecular Ecology 11:2475–2487.
49. Wheeler QD. 2004. Taxonomic triage and the poverty of phylogeny. Philosophical Transactions of the Royal Society B 360:571–583.
50. Nielsen R. Matz M. 2006 Statistical approaches for DNA barcoding. Systematic Biology 55:162–169.
51. Won YJ, Sivasundar A, Wang Y, Hey J. 2005. On the origin of Lake Malawi cichlid species: A population genetic analysis of divergence. PNAS 102:6581–6586.
52. Lipscomb D, Platnick N, Wheeler Q. 2003. The intellectual content of taxonomy: a comment on DNA taxonomy. Trends in Ecology and Evolution 18:65–66.
53. Ebach MC, Holdrege C. 2005. DNA barcoding is no substitute for taxonomy. Nature 434:697.
54. Moritz C, Cicero C. 2004. DNA Barcoding: promise and pitfalls. PLoS Biology 2:1529–1531.
55. Seberg O, Humphries CJ, Knapp S, Stevenson DW, Petersen G, et al. 2003. Shortcuts in systematics? A commentary on DNA-based taxonomy. Trends in Ecology and Evolution 18:63–65.
56. Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP. 2003. A plea for DNA taxonomy. Trends in Ecology and Evolution 18:70–74.
57. Hebert PDN, Gregory TR. 2005. The promise of DNA barcoding for taxonomy. Systematic Biology 54:852–859.
58. Schindel DE, Miller SE. 2005. DNA barcoding a useful tool for taxonomists. Nature 435:17.
59. The Global Biodiversity Information Facility website. http://www.gbif.org/ Accessed 1 November 2006.
60. Smith V. 2005. DNA barcoding: perspectives from a "partnerships for Enhancing Expertise in Taxonomy" (PEET) debate. Systematic Biology 54:841–844.
61. Tree of Life website. http://tolweb.org/tree/phylogeny.html Accessed 6 May 2006.

DNA barcoding of animal species—response to DeSalle

In issue number 29.12 of this journal, DeSalle[1] recently commented on the review by Waugh,[2] which outlined the conceptual basis and progress on the worldwide effort to DNA barcode animal species. DeSalle's paper is a useful clarification of a number of the issues relating to the analytical approach to DNA barcoding data and to the role that DNA barcoding might play in modern biology. He criticised the use of distance-based approaches to DNA barcoding and pointed out that these methods have a number of drawbacks. Moreover, he also pointed out that such distance-based approaches are not typically used by taxonomists in the process of describing new species. Instead, DeSalle[1] argued for the use of character-based or diagnostic approaches to both species discovery and species identification. He suggested that this approach was preferable because first, diagnostic DNA sequence characters can be easily extracted from COI sequences without assuming any particular model of nucleotide substitution (*sic* evolution). Second, character-based methods do not require arbitrary cutoffs, as do distance-based methods. Third a DNA diagnostic can be determined without reference to other organisms. Fourth, it avoids the confounding effects of distance-based methods that occur as a result of anagenetic changes. Finally, DeSalle[1] suggested that diagnostic-based techniques are more adaptable to the possible development of handheld DNA-based identification devices at some time in the future.

Conceptually, there is much to recommend DeSalle's approach. Species represent a major component of the diversity that is inherent in nature. Of course, there is also diversity at the individual and population levels, for example. Although there has been over the centuries much debate about species definitions, there has been little disagreement about the reality of species and their importance to the diversity of life. DNA barcoding is not directed at species definition but to the problem of their identification. DeSalle points out that his approach is consistent with the methods of taxonomy and we would point out that it is also consistent with the biological reality of species and with their existence, independent of other species.

Notwithstanding DeSalle's comments, a distance-based approach to DNA barcoding appears generally successful. Specifically, the Kimura 2 parameter[3] model of nucleotide substitution, when applied to the measurement of intra- and interspecific variation in mitochondrial cytochrome c oxidase I (COI) gene sequences, has been widely used. The COI gene (58–705 from the 5′ end) works well as an identifier of a large number of animal species. It is hard to gainsay the evidence of more than 300,000 successful identifications from over 30,000 formally described species, using the distance-based approach to the analysis of DNA barcodes.[4]

Nonetheless, if, as Desalle remarks, a character-based approach can easily be applied to DNA barcoding, it could have potential advantages and should certainly be explored. In an earlier paper, DeSalle et al.[5] successfully applied this approach using various DNA sequences from a number of taxa. Similarly, we have compared COI barcode sequences from 11 species of the extinct New Zealand moa, which we had previously investigated using distance-based analyses of DNA barcodes. In this case, we used a character-based or diagnostic approach to identify any unique nucleotides at specific sites within the COI region. These would then enable species identification. Results showed that unique nucleotides within the barcode region of moa species correctly identified 10 of the 11 putative species (Fig. 1). However, all species can be identified using a combination of the COI characters from the dataset available. By comparison, a K2P distance analysis of the same data also correctly identified all 11 species.[6]

Hence, in this case, as in many others, both character and distance-based approaches to the analysis of DNA-barcoding data are successful. However, it is not difficult to imagine a situation in which some species can be identified based on one or a small number of nucleotide characters. In this situation, it is unlikely that a distance-based approach would be successful. Thus, diagnostic or character-based approaches may generally be more successful than distance based measures.

At the more general level, DeSalle[1] also suggested that taxonomy should be regarded as species discovery, while DNA barcoding should be about species identification. This distinction between these two concepts is useful and DeSalle is correct that we should be mindful of blurring the distinction between them. However, just as taxonomy must be about both species discovery and identification, so too DNA barcoding can contribute to both. If it becomes possible to rapidly identify known species using DNA barcoding, unknown species can then be more easily identified. We also agree with DeSalle[1] on the importance of the distinction between of DNA taxonomy, and species discovery and identification. The former represents the concept that information in DNA sequences is sufficient to identify and discover species. However, few authors[7] have proposed DNA taxonomy as an alternative to existing approaches and Waugh[2] did not either.

No authors that we know of have suggested that DNA barcoding is without merit. Some authors have criticized the costs which, they argue, drain resources from other areas of biodiversity science while many others such as DeSalle[1] have argued that DNA barcoding will not and should not replace taxonomy per se. DeSalle's[1] central point is about the analysis of DNA-barcoding data and about the importance of

BioEssays 30:92–93, © 2007 Wiley Periodicals, Inc.

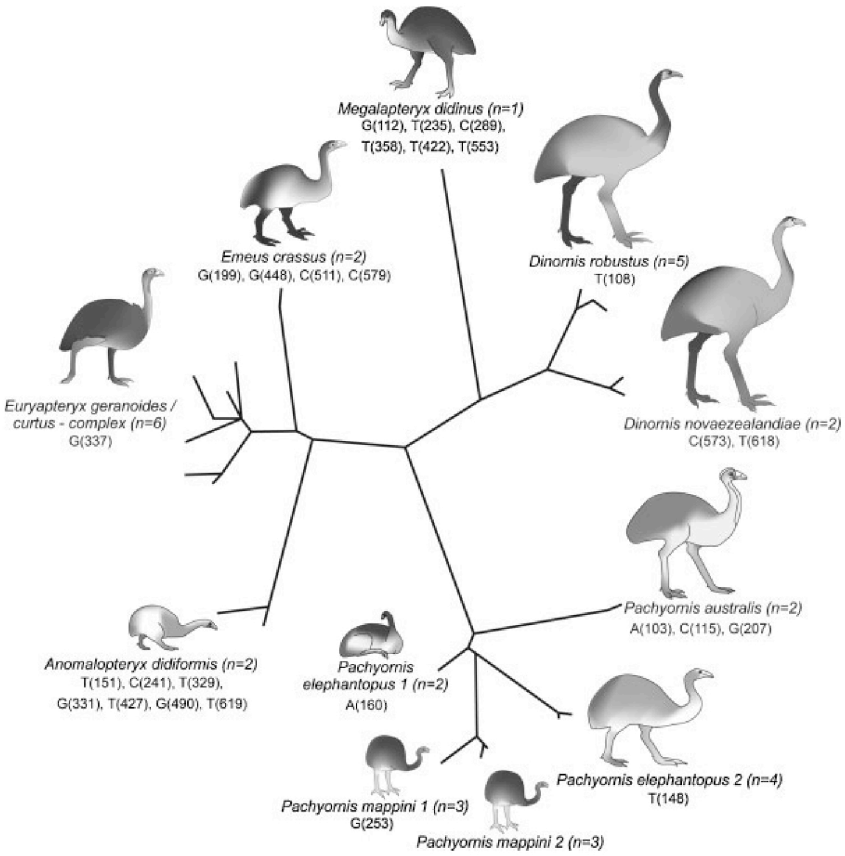


Figure 1. An unrooted phylogram of 596bp COI sequences belonging to 11 species of extinct New Zealand moa. The number of samples examined is shown in brackets, together with the number of unique diagnostic nucleotide sites. The nucleotide sites are numbered from the beginning of the barcode region, defined in moa as 58–705 from the 5′ end.[6]

the distinction between some of the central concepts involved. On both these important issues we agree.

References

1. DeSalle R. 2007. Phenetic and DNA taxonomy; a comment on Waugh. BioEssays 29:1289–1290.
2. Waugh J. 2007. DNA barcoding in animal species: progress, potential and pitfalls. BioEssays 29:188–197.
3. Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120.
4. Barcode of Life data systems website http://www.barcodinglife.org/views/login.php accessed 11th September, 2007.
5. Desalle R, Egan MG, Siddal M. 2005. The unholy trinity: taxonomy, species delimitation and DNA barcoding. Phil Trans R Soc B 360:1905–1916.
6. Lambert DM, Baker A, Huynen L, Hadrath O, Hebert PDN, et al. 2005. Is a large-scale DNA-based inventory of ancient life possible? J Hered 96: 279–284.
7. Tautz D, Archtander P, Minelli A, Thomas RH, Vogler AP. 2003. A plea for DNA taxonomy. TREE 18:70–74.

John Waugh, Leon Huynen, Craig Millar and David Lambert

Allan Wilson Centre for Molecular Ecology and Evolution
Institute of Molecular BioSciences, Massey University
Private Bag 102 904, North Shore Mail Centre
Auckland, New Zealand
E-mail: w.j.waugh@massey.ac.nz

DOI 10.1002/bies.20698
Published online in Wiley InterScience (www.interscience.wiley.com).

Paper 3

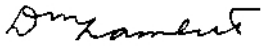
Chapter 3 is based on a paper that reports the design and the efficacy of a suite of PCR and sequencing primers. These primers were designed to amplify the avian COI barcoding region from modern and historical samples. This paper was published in Molecular Ecology Resources in 2009.[1] The first author was Selina Patel and the paper was co-authored by John Waugh, Craig Millar and David Lambert. Some additional information on primer design is included in this chapter. John Waugh produced the multiple bird sequence alignment of the COI barcode region including approximately 1000 bases either side of this region. The sequence alignment was used in the initial phases of primer design for all the primers described here. He designed several of the "external" primers that were subsequently used and both computer and laboratory tested them. He also designed and computer tested the initial versions of the "internal" primers for historical samples. These primers were tested, substantially modified and had 5' tags added to them by Selina Patel. All historical samples used in the study were collected from study skins in Canterbury Museum by John Waugh with the help of Paul Scofield.

Auckland 16th November 2010

Selina Patel

John Waugh

Craig Millar



David Lambert

References

1. Patel, S., Waugh, J, Millar, CD, Lambert, DM, *Conserved primers for DNA barcoding historical and modern avian samples.* Molecular Ecology Resources, 2009.

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2009) doi: 10.1111/j.1755-0998.2009.02793.x

DNA BARCODING

Conserved primers for DNA barcoding historical and modern samples from New Zealand and Antarctic birds

SELINA PATEL,* JOHN WAUGH,† CRAIG D. MILLAR* and DAVID M. LAMBERT†‡
*Allan Wilson Centre for Molecular Ecology and Evolution, University of Auckland, Private Bag 92019, Auckland, New Zealand, †Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, Private Bag 102 904 NSMC, Auckland, New Zealand, ‡Griffith School of Environment and School of Biomolecular and Physical Sciences, Griffith University, 170 Kessels Road, Nathan, Qld 4111, Australia

Abstract

Our ability to DNA barcode the birds of the world is based on the effective amplification and sequencing of a 648 base pair (bp) region of the mitochondrial cytochrome *c* oxidase (COI or *cox*1) gene. For many geographic regions the large numbers of vouchered specimens necessary for the construction of a DNA barcoding database have already been collected and are available in museums and other institutions. However, many of these specimens are old (>20 years) and are stored as either fixed study skins or dried skeletons. DNA extracted from such historical samples is typically degraded and, generally, only short DNA fragments can be recovered from such specimens making the recovery of the barcoding region as a single fragment difficult. We report two sets of conserved primers that allow the amplification of the entire DNA barcoding region in either three or five overlapping fragments. These primer sets allow the recovery of DNA barcodes from valuable historical specimens that in many cases are unique in that they are unable or unlikely to be collected again. We also report three new primers that in combination allow the effective amplification from modern samples of the entire DNA barcoding region as a single DNA fragment for 17 orders of Southern Hemisphere birds.

Keywords: 5′-tags, avian primers, COI, *cox*1, DNA barcoding, species identification

Received 3 June 2009; revision received 24 July 2009; accepted 25 July 2009

Introduction

At the heart of DNA barcoding is the proposition that individuals can be rapidly and easily assigned to the species level using standardized DNA sequences from the mitochondrial genome (Hebert *et al.* 2003a, b). A considerable body of research has been carried out that provides support for this proposition and there is currently intense international interest in this area of research. Furthermore, it has been suggested that the DNA barcoding approach, if successful, will complement modern taxonomy while at the same time revolutionising the way we routinely identify species in the future (Hajibabaei *et al.* 2007).

Correspondence: David M. Lambert, Fax: +61 7 373 57459; E-mail: d.lambert@griffith.edu.au

The World's avian fauna has been identified as a good candidate to test the general applicability of DNA barcoding because birds have been extensively studied using a range of approaches including morphology, genetics and behaviour. This has resulted in a stable and mature taxonomy (Gill 2007; Baker *et al.* 2009). However, to date, the majority of DNA barcoding studies on birds have been conducted using modern samples from Northern Hemisphere species (Hebert *et al.* 2004; Yoo *et al.* 2006; Kerr *et al.* 2007). It remains to be seen whether DNA barcoding will accurately identify species from non-continental regions where biogeographic patterns, mutation and speciation rates may have been quite different (Moritz & Cicero 2004).

New Zealand and the Antarctic have a distinctive avian fauna of moderate size (comprising of approximately 290 native or indigenous species) with a high level of endemism (Heather & Robertson 2005). In order to

establish a DNA barcode database of these birds it is necessary to have access to a large number of vouchered museum specimens and to effectively amplify and sequence the DNA barcoding region from samples obtained from these specimens.

A 648 bp region at the 5′ end of the mitochondrial gene cytochrome *c* oxidase subunit 1 (COI or *cox*1) has been identified as the DNA barcoding region for most animal species, including birds (Hebert *et al.* 2003a, 2004). Two types of samples are available from vouchered museum specimens, fresh or modern tissue (blood, feather or muscle that has been preserved at low temperatures or in ethanol) and historical specimens (usually study skins or skeletons). Although many museums now collect and preserve samples from vouchered specimens for biochemical and DNA analysis (Seutin *et al.* 1991; Arctander & Fjeldsa 1994), this is a relatively recent practice. As a result, there are only a relatively small number of modern tissue samples available from vouchered specimens for New Zealand and Antarctic birds and this is likely to be the case for many regions worldwide. Furthermore, the collection of new or additional vouchered specimens is in many cases difficult, expensive and time consuming. The vast majority of historical specimens are older than 20 years in age and have not been preserved with DNA analyses in mind. Samples from these specimens typically produce highly degraded DNA (Lindahl 1993). The size of the DNA fragments recovered from historical samples depends on the preservation method, sample age and a number of environmental factors (Hofreiter *et al.* 2001). The degraded nature of the recovered DNA from historical samples makes it difficult to amplify a fragment larger than 300–400 bp in length as well as resulting in low rates of amplification (Zimmermann *et al.* 2008).

There are a number of approaches used to tackle the problem of generating barcodes from historical museum specimens. Short or minimalist barcodes have been recovered by amplifying only a selected 100–200 bp fragment of the originally proposed ~650 bp COI region. In a number of insect genera mini-barcodes have proved to be successful in identifying almost all species (Hajibabaei *et al.* 2006). However, mini-barcodes show limited success compared with the full-length barcode and only work within confined taxonomic groups (Hajibabaei *et al.* 2006). Although potentially useful for the identification of individual samples, this approach does not resolve the issue of having to construct an initial full-length database from which to make these identifications. Furthermore, whether it can be applied to avian species is yet to be determined.

Another approach is to use a range of DNA repair enzymes. These enzymes can increase the amplification success rate by repairing some of the damage that is induced in DNA as a result of chemical treatment or aging. The most common types of DNA damage to historical and ancient DNA are oxidative and hydrolytic damage (Hoss *et al.* 1996). Deamination, a type of hydrolytic damage common to cytosine residues, results in the replacement of cytosine with uracil, ending in C to T and G to A transitions during polymerase chain reaction (PCR) (Evans 2007). Treatment with DNA repair enzymes can result in the reversal of most types of DNA damage leading to higher amplification success rates (Evans 2007). However, this approach has a number of shortcomings specifically its inability to repair fragmented DNA and its high cost when processing large numbers of specimens.

The ability to recover DNA sequences from historical samples is in part dependent on the size of the PCR product targeted, generally the smaller the amplicon size the higher the rate of amplification success. Therefore, for any section of DNA, complete sequences can be recovered from historical samples by targeting a number of smaller overlapping regions. Once obtained, these overlapping fragments can be assembled and this approach has been used very successfully in the study of ancient DNA (Millar *et al.* 2008). Hence, the ability to DNA barcode the birds of the World will be facilitated by the availability of a set of conserved primers that will allow the amplification of the COI barcoding region from historical samples.

We report the development of two sets of PCR primers that reliably amplify from historical samples the COI barcoding region in three or five short segments. In addition, one of these primers, in combination with two others, also allows the effective amplification of the entire COI barcoding region as a single fragment from modern samples from 17 orders of Southern Hemisphere birds tested.

Materials and methods

Primer design

Primers for the amplification of historical and modern tissue were designed based on an alignment (using the Clustal_X algorithm) of sequences obtained from New Zealand and Antarctic avian species in addition to a number of COI sequences available on GenBank (accession numbers available in supplementary information Table S1). Highly conserved regions were targeted as areas of interest for the synthesis of potential primers. Primer integrity was checked using Amplify 3X, available from http://engels.genetics.wisc.edu/amplify/ and NetPrimer available on the PREMIER biosoft International website http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html.

Two sets of internal primers (Table 1) were designed to span the 648 bp COI region for historical samples considered to have low quality DNA. The first primer set contained three primer pairs resulting in three overlapping fragments (average length ~310 bp; Fig. 1). The second primer set contained five primer pairs resulting in five shorter overlapping fragments (average length ~190 bp; Fig. 1) to span the same 648 bp region. Each internal primer contained a randomly generated 18 bp tag attached to its 5′ end in order to improve the sequencing of the resulting PCR products. Forward internal primers all contained an identical Ftag sequence and similarly all reverse internal primers contained an identical Rtag sequence. Internal primer pairs were tested on a range of modern samples represented from 15 avian orders prior to being used to amplify DNA from historical samples.

Three primers, one forward and two reverse were designed to amplify the entire 648 bp COI region as a single fragment to be used for modern specimens. The forward primer, AWCF1 was also used as part of the primer set for historical samples. Primer pairs intended for modern specimens did not contain the 18 bp tags. All primer sequences and their expected fragment lengths are shown in Table 1.

DNA extraction from historical and modern samples

Historical specimens in the form of dried skins used for this study were obtained from Canterbury Museum (Christchurch, New Zealand). Until 1970 all specimens were preserved with arsenical soap. After this date the preservative used was changed to borax. Both methods are commonly used in museums worldwide. Preparation did not involve any special preservation on the legs and toe pads however. These were allowed to dry out naturally (P Scofield 2009, personal cummunication). All samples were collected from the foot and toe region of the specimens.

Between 5 and 64 mg of tissue was used per extraction and all tissue was macerated with a sterile scalpel blade prior to extraction. DNA isolation from dried skins was performed using proteinase K/DTT/SDS digestion followed by a DNeasy tissue extraction kit for the post-lysis steps according to the manufacturer's instructions (Qiagen). Negative controls were included for each set of extractions. No more than five samples were extracted per set. All pre-PCR steps were carried out in a separate laboratory solely dedicated to ancient DNA work. DNA from modern samples (fresh tissue) was extracted in a different laboratory using a standard proteinase K digestion and a modified version of the phenol/chloroform method (Sambrook *et al.* 1989; Lawrence *et al.* 2008).

PCR amplification and sequencing

PCRs were run in 25 μL volume reactions and included 1x PCR buffer (Invitrogen), 2.5 mM $MgCl_2$, 1 mg/mL BSA, 0.5μM forward and reverse primers, 4% DMSO (only for 5′-tagged primers), 0.2 mM dNTPs, 1U of Platinum *Taq* DNA polymerase (Invitrogen) and 3-4 μL of DNA template from the final elution for historical samples and 25–50 ng of DNA for modern samples. The thermal cycling conditions for historical samples was as follows: 2 min at 94 °C followed by 10 cycles of 20 s at 94 °C, 20 s at 55 °C and 20 s at 72 °C followed in turn by 30 cycles of 20 s at 94 °C, 20 s at 50 °C and 20 s at 72 °C and a final 4 min at 72 °C. In the case of modern samples the thermal cycling conditions were altered and consisted of 2 min at 94 °C followed by 35 cycles of 30 s at 94 °C, 30 s at 57.5 °C and 30 s at 72 °C and a final 4 min at 72 °C. PCR products were visualized on a 2% agarose gel. The remaining volume of PCR products were purified using the Agencourt AMPure PCR Purification system (Beckman Coulter). All purified PCR products were then cycle sequenced using Big Dye 3.1 chemistry and subsequently analysed on an ABI Prism 3130xl genetic analyzer. Sequencing was carried out bidirectionally to produce an overlap of at least 648 bp of the COI region for all modern samples. In the case of the historical samples, three or five overlapping fragments were amplified which covered the same COI region. In each case, these fragments were bidirectionally sequenced from independent PCR reactions. Once the fragments were sequenced, the complete DNA barcode for each historical sample was assembled by concatenating sequences using the software programmes Sequencher 4.6 (Gene Codes Corporation) and Vector NTI (Invitrogen). Only those sequences which reached PHRED scores of 20 or above for at least 90% of individual bases were deemed acceptable.

Results

We designed and tested two sets of 5′-tagged PCR primers that reliably amplified short DNA fragments from historical samples for the barcoding region. The first set of primers containing three primer pairs generated three fragments approximately 310 bp in size whereas the second set containing five primer pairs generated shorter fragments of approximately 190 bp in size. Both primer sets were first tested on a range of modern samples with representative species from 15 different orders (Table 2) prior to be being used on historical specimens. Of a total of 29 species, both primer sets worked on all species tested.

Twenty seven historical samples were utilized in this study belonging to nine different orders (See supporting

Table 1 PCR primers and 5′-tags used for the amplification of the 648 bp COI region as either one, three or five fragments. The position of each primer is given relative to the chicken mitochondrial sequence. The 5′-tags were only added to the individual primers when PCR amplifying and sequencing ancient samples as either three or five fragments. AWCintF6 and AWCintR7 represent alternate primer pairs used for the orders Charadriiformes and Anseriformes [Correction added after online publication Dec 8 2009: Sequences and positions of forward primers AWCintF2 and AWCintF3 were swapped. Sequences and positions of reverse primers AWCintR4 and AWCintR3 were swapped.]

Forward primer	Sequence 5′ to 3′	Position (bp)*	Reverse Primer	Sequence 5′ to 3′	Position (bp)*	Fragment size (bp)*
Primers for modern samples						
AWCF1	CGCYTWAACAYTCYGCCATCTTACC	6625–6649	AWCR6	ATTCCTATGTAGCCGAATGGTTCTTT	7446–7471	848
AWCF1			AWCR3	ATGCTCGGGTGTCTACGTCTAT	7542–7563	936
First set of internal primers containing 3 primer pairs						
AWCF1			AWCintR2	ATGTTGTTTATGAGTGGGGAATGCTATG		328
AWCintF2	ATAATCGGAGGCTTCGGAAACTGA	6873–6896	AWCintR4	TGGGAKAGGGCTGGTGGTTTTTATGTT	7161–7186	314
AWCintF4†	TCCTCAATCCTGGGAGCAATCAACTT	7119–7144	AWCintR6‡	GGATTAGGATGTAGACTTCTGGGTG	7371–7395	278
Second set of internal primers containing 5 primer pairs						
AWCF1			AWCintR1	CCTGGTTGACCTAGTTCTGCTCG	6765–6787	163
AWCintF1	CCGAGCAGAACTACGTCAACC	6764–6784	AWCintR2		6926–6952	189
AWCintF3	ATCGGAGCCCCCAGACATAGCATT	6912–6934	AWCintR3	TTGATGGCTGTTGTGATAAAGTTGAT	7137–7162	251
AWCintF4†			AWCintR5	TGCTGGGTCGAAGAATGTGGTGTT	7299–7322	204
AWCintF5	GGCATCACCATACTACTAACAGACCG	7266–7291	AWCintR6‡			130
5′tags for internal primers						
Rtag	AGTCGACGCTTCTAGCTT	—	Rtag	CATGCTACCTGCTACTGT	—	

*Basepairs.
†AWCintF6 – TAGGGGCAATCAACTTCATCACAAC (7129–7153).
‡AWCintR7 – ACGTRTGAGATAATTCCGAATCC (7401–7423).

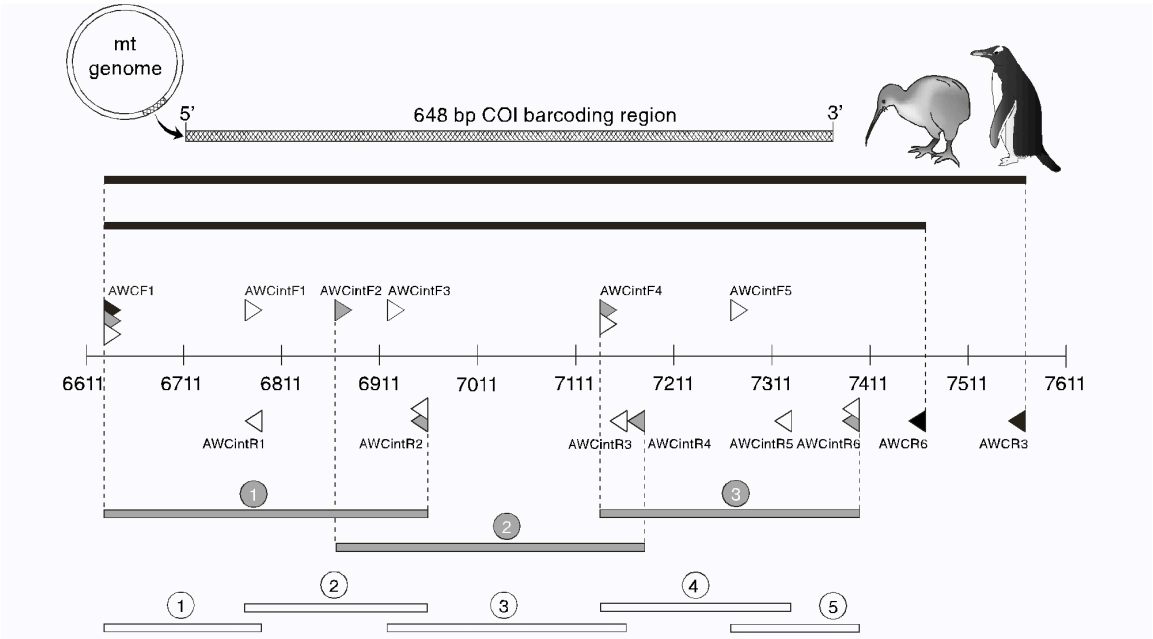


Fig. 1 A diagrammatic representation of the COI barcoding region used for birds with the relative position of the primers (triangles) indicated for amplifying historical and modern samples. The three black triangles and the two black bars indicate the combination of primers used and DNA fragments generated for modern samples. Similarly, the grey or white triangles and the grey or white bars represent the primer combinations and resulting DNA fragments generated for historical samples. The position of each primer and the amplified DNA fragments are given relative to the chicken mitochondrial genome.

information Table S2 for species list). The age of the historical specimens tested ranged from 16 to 115 years. PCR products were amplified at similar success rates from both the three and five primer pair sets for the same given samples tested suggesting that the three primer pair set is the best option (Table 2). The AWCF1/AWCR6

Table 2 Results of PCR success of modern and historical samples for the two primer sets designed to amplify DNA from historical specimens. PCR success is defined as fragments amplified (Y = amplified; N = not amplified; NT = not tested)

	Modern specimens			Historical specimens		
Order	3 fragments	5 fragments	No. of species	3 fragments	5 fragments	No. of Species
Anseriformes	Y*	Y*	2	N	N	0/2
Ciconiiformes	Y	Y	2	Y	Y	2/2
Charadriiformes	Y*	Y*	2	Y*	Y*	2/4
Falconiformes	Y	Y	2	Y	Y	1/2
Gruiformes	Y	Y	2	Y	Y	1/2
Pelecaniformes	Y	Y	2	N	N	0/2
Procellariiformes	Y	Y	2	Y	Y	9/10
Sphenisciformes	Y	Y	2	Y	Y	1/2
Psittaciformes	Y	Y	2	N	N	0/1
Passeriformes	Y	Y	2	NT	NT	—
Apterygiformes	Y	Y	2	NT	NT	—
Coraciiformes	Y	Y	2	NT	NT	—
Galliformes	Y	Y	2	NT	NT	—
Columbiformes	Y	Y	2	NT	NT	—
Podicipediformes	Y	Y	1	NT	NT	—

*Not all primer pairs were successful, alternate primer pairs have been supplied for these orders (see Table 1).

primer pair which amplifies the entire 648 bp COI fragment was used as a control for historical specimens with no products being amplified.

Of the 27 historical samples tested, 16 samples successfully amplified the regions of interest resulting in an amplification success rate of 59%. Furthermore, analysis found PCR success to be strongly correlated with specimen age. When separated into age groups we found samples which were 80 years or less in age had a higher average amplification success rate of 77% (n = 13) than their older counterparts of greater than 80 years of age which had an average amplification success rate of 50% (n = 12). However there was no significant difference found between the two age groups (Fisher's exact test, P-value = 0.2262). Two of the 27 samples were of unknown ages and could not be classified into age groups; neither produced positive PCR results.

Historical specimens belonging to the orders Anseriformes, Pelecaniformes and Psittaciformes did not successfully amplify the three fragments or the five fragments. However, their modern representatives were successfully amplified suggesting that the problem was due to DNA damage rather than issues with primer binding. Bar one, all samples were over 80 years of age, stressing again the importance of age on PCR success. For the orders Charadriiformes and Anseriformes the last primer pair (AWCintF5/AWCintR6) posed a problem for some species. We were able to design an alternate primer pair (AWCintF6/AWCintR7) for these orders (refer to Table 1). This primer pair could have future use for other untested orders if a similar problem is encountered.

A number of previously described avian COI primers (Hebert *et al.* 2004; Kerr *et al.* 2009) designed to amplify a single fragment for modern samples from the Northern Hemisphere species had limited success when applied to samples from the New Zealand and Antarctic avifauna (data not shown). In particular, the amplification of DNA from tube-nosed seabirds (Procellariiformes) and kiwi (Apterygiformes) proved to be difficult. A similar result was found for the internal primers AvMiF1 and AvMiR1 where preliminary tests on Procellariiformes resulted in the amplification of multiple fragments. This difficulty could be attributed to the fact that New Zealand and Antarctica have a more diverse and larger number of endemic seabirds than other geographic regions like North America for which these avian primers were originally designed. Moreover, orders like the Apterygiformes (kiwi) are unique to New Zealand.

The successful amplification of the entire barcoding region using the newly designed primers AWCF1, AWCR6 and AWCR3, are shown in Table 3. The only previously published avian primer to have applicability to New Zealand and Antarctic birds was the COIbirdR2 primer (Kerr *et al.* 2009) when paired with our AWCF1 primer. This primer pair worked extremely well for Passerines (Table 3). The newly designed AWCF1/AWCR6 primer pair worked well for most orders of birds, with the forward primer AWCF1 proving to be universally successful at binding appropriately and the main reverse primer AWCR6 proving to be highly successful in the majority of cases. AWCR3 worked well in species for which AWCR6 did not prove successful.

Table 3 PCR primer success on 17 different avian orders. The numbers represent the species of each order from which the target sequence was successfully amplified, using different combinations of PCR primers

Order	AWCF1/R6	AWCF1/R3	AWCF1/R6 & AWCR3	AWCF1/ COIbirdR2	No. of species per order
Anseriformes	16	—	—	1	16
Apterygiformes	—	5	—	—	5
Charadriiformes	25	4	4	1	25
Ciconiiformes	9	—	—	2	9
Columbiformes	3	—	—	3	3
Coraciiformes	1	—	—	1	1
Cuculiformes	2	—	—	—	2
Falconiformes	2	—	—	1	2
Galliformes	1	—	—	1	1
Gruiformes	7	3	2	4	8
Passeriformes	17	6	1	20	20
Pelecaniformes	7	2	—	—	7
Podcipediformes	1	2	—	—	3
Procellariiformes	33	5	3	5	35
Psittaciformes	4	3	2	1	5
Sphenisciformes	6	1	1	1	6
Strigiformes	1	1	1	—	1
Total	135	32	14	41	149

AWCR3 worked particularly well in Apterygiformes, where all five species of this order could only be amplified with the AWCF1/AWCR3 primer pair.

The majority of the sequences generated using the AWCF1/AWCR3/AWCR6 primers fell into the category of high quality as stipulated by the Barcode of Life Data System (BOLD) where mean PHRED scores were >40 (Ratnasingham & Hebert 2007).

Discussion

COI sequence variation between taxonomic groups has led to difficulties in the design of a single primer pair with universal applicability. Furthermore, in the case of historical samples, the 648 bp COI barcoding region can only be recovered in smaller fragments requiring multiple primer sets leading to the proposition that primers must be customized to suit. However, the generation of taxa-specific primers is impractical. A realistic solution to this problem is to generate a small set of primers that will work effectively on all avian species, both historical and modern. The primers generated by us are an example of such a set (Fig. 1). These primers, designed to amplify the barcode region of historical and modern samples from New Zealand and Antarctic bird species proved generally successful with only a few exceptional taxa, such as the Apterygiformes requiring specialized primers. As these primers were designed using an alignment of sequences from a range of both Northern and Southern Hemisphere species, it is likely that they will be effective for avian orders from many geographic locations. It is important to note that there are a number of other primers published that are not tested in this study which could prove to be highly useful (Tavares & Baker 2008; Lohman *et al.* 2009). However, our primers have a very broad applicability, having been tested on a large number of species from 17 avian orders. All primers generated high quality sequences for the region of interest. Primers designed for modern samples had binding sites at least 80 bp from the start and end of the barcode region to allow for good overlapping bi-directional sequence and to make allowance for the loss of sequence data which occurs at the beginning of each sequence read.

To maximize the chance of obtaining amplifiable DNA from historical specimens we suggest sampling from the toe pad area where the risk of damage from preservatives is minimal. Specimen age should also be carefully considered and where possible, younger specimens should be selected for DNA analysis.

The 648 bp barcode region for historical avian specimens can now be amplified using tagged internal primers as either three or five fragments, depending on the DNA quality of individual samples. We suggest using the set of primers containing three primer pairs as the likelihood of achieving success with the set of three primers pairs is the same as that from the set of five primer pairs. As our results demonstrate, if one set of internal primers does not prove successful the other set is unlikely to work either. In the event that not all overlapping fragments are amplified there is the added possibility of using one or more of these smaller fragments as a mini-barcode once this approach has be confirmed to work for avian taxa.

The addition of 5′-tags to each of the internal primers provides a number of advantages. Not only does it result in an increase in the read length of each sequence, thus allowing for better overlap of individual contigs, but it has been shown to greatly improve sequence quality (Binladen *et al.* 2007). This is crucial for historical samples where DNA damage is frequent. The 5′-tags also allow for large-scale sequencing as well as being more cost effective, as only two primers (forward and reverse tags) are required instead of the many individual forward and reverse primers.

The disadvantage, however, is that the processes of DNA extraction, PCR amplification and sequencing from historical specimens can be taxing and time consuming. Thus, modern specimens should be utilized for barcoding purposes wherever possible. We suggest using DNA barcodes generated from historical samples as a way of adding to an already existing database where there is a lack of their modern counterparts or for cases of rare and endangered species where sample collection is problematic.

There is now an onus upon museums and other institutions to ensure the storage of well preserved vouchered samples of as many species as possible, from which, high quality DNA can be extracted. The maintenance of such collections is increasingly being adopted around the world. However, until such time as these collections become comprehensive and for rare or now extinct species, study skins and bones provide the only vouchered source of DNA. While that is the case, there is a need for methods, such as those outlined above, to provide a means of generating DNA barcodes.

Acknowledgements

Funding for this study was provided by the University of Auckland's Vice Chancellor's Development Fund and the Allan Wilson Centre for Molecular Ecology and Evolution. We would like to thank Leon Huynen for help with primer design and invaluable advice on ancient DNA techniques and also Judith Robins for allowing us to make use of the Department of Anthropology ancient DNA Laboratory at the University of Auckland. We would like to thank all the people who kindly donated samples to make this study possible and in particular Paul Scofield for providing us with all the historical specimens from Canterbury museum.

References

Arctander P, Fjeldsa J (1994) Avian tissue collections for DNA analysis. *Ibis*, **136**, 359–360.

Baker AJ, Tavares ES, Elbourne RF (2009) Countering criticisms of single mitochondrial DNA gene barcoding in birds. *Molecular Ecology Resources*, **9**, 257–268.

Binladen J, Thomas M, Gilbert P, Campos PF, Willerslev E (2007) 5'-Tailed sequencing primers improve sequencing quality of PCR products. *BioTechniques*, **42**, 174–176.

Evans T (2007) DNA damage. *NEB expressions*, **2.1**, 1–3.

Gill FB (2007) *Ornithology*, 3rd edn. W.H. Freeman, New York.

Hajibabaei M, Smith AM, Janzen DH *et al.* (2006) A minimalist barcode can identify a specimen whose DNA is degraded. *Molecular Ecology Notes*, **6**, 959–964.

Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, **23**, 167–172.

Heather BD, Robertson HA (2005) *The Field Guide to the Birds of New Zealand*, Rev. edn. Penguin Books, Auckland.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, 313–321.

Hebert PDN, Ratnasingham S, de Waard JR (2003b) Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **270**, S96–S99.

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of Birds through DNA Barcodes. *PloS Biology*, **2**, 1657–1663.

Hofreiter M, Serre D, Poinar HN, Kuch M, Paabo S (2001) Ancient DNA. *Nature Reviews Genetics*, **2**, 353–359.

Hoss M, Jaruga P, Zastawny TH, Dizdaroglu M, Paabo S (1996) DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Research*, **24**, 1304–1307.

Kerr KCR, Stoeckle MY, Dove CJ *et al.* (2007) Comprehensive DNA barcode coverage of North American birds. *Molecular Ecology Notes*, **7**, 535–543.

Kerr KCR, Lijtmaer DA, Barreira AS, Hebert PDN, Tubaro PL (2009) Probing Evolutionary Patterns in Neotropical Birds through DNA Barcodes. *PloS One*, **4**, e4379.

Lawrence HA, Taylor GA, Millar CD, Lambert DM (2008) High mitochondrial and nuclear genetic diversity in the world's most endangered seabird, the Chatham Island Taiko (*Pterodroma magentae*). *Conservation Genetics*, **9**, 1293–1301.

Lindahl T (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.

Lohman DJ, Prawiradilaga DM, Meier R (2009) Improved COI barcoding primers for Southeast Asian perching birds (Aves: Passeriformes). *Molecular Ecology Resources*, **9**, 37–40.

Millar CD, Huynen L, Subramanian S, Mohandesan E, Lambert DM (2008) New developments in ancient genomics. *Trends in Ecology & Evolution*, **23**, 386–393.

Moritz C, Cicero C (2004) DNA Barcoding: promise and Pitfalls. *PloS Biology*, **2**, 1529–1531.

Ratnasingham S, Hebert PDN (2007) bold: the Barcode of Life Data System. *Molecular Ecology Notes*, **7**, 355–364.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.

Seutin GW, Bradley N, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, **69**, 82–90.

Tavares E, Baker A (2008) Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. *BMC Evolutionary Biology*, **8**, 81.

Yoo H, Eah J, Kim J *et al.* (2006) DNA Barcoding Korean Birds. *Molecules and Cells*, **22**, 323–327.

Zimmermann J, Hajibabaei M, Blackburn D *et al.* (2008) DNA damage in preserved specimens and tissue samples: a molecular assessment. *Frontiers in Zoology*, **5**, 18.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Accession numbers of species used as part of the alignment for primer design.

Table S2 List of historical specimens from Canterbury museum.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Paper 4

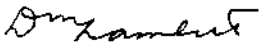
A paper published in Molecular Ecology Resources in 2010 forms the core of Chapter 6 of this thesis.[1] John Waugh was the principal author and carried out the laboratory work with the exception of three samples that were sequenced by Selina Patel at the University of Auckland. He also carried out the analysis and interpretation of the sequence data and birdstrike statistics. The data on birdstrikes from New Zealand airports were provided by Max Evans of the New Zealand Civil Aviation Authority. David Lambert and Craig Millar contributed to the writing of the manuscript and assisted with the figures as well as providing funding for the laboratory work carried out for this paper.

Auckland 16th November 2010

John Waugh

Max Evans

Craig Millar



David Lambert

References

1. Waugh, J., Evans, MW, Millar, CD, Lambert, DM, *Birdstrikes and barcoding: can DNA methods help make the airways safer?* Molecular Ecology Resources, 2010. **9999**(9999).

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2010) doi: 10.1111/j.1755-0998.2010.02884.x

DNA BARCODING

Birdstrikes and barcoding: can DNA methods help make the airways safer?

JOHN WAUGH,* MAX W. EVANS,† CRAIG D. MILLAR‡ and DAVID M. LAMBERT§
*Allan Wilson Centre for Molecular Ecology and Evolution, Institute of Natural Sciences, Massey University, Private Bag 102 904, NSMC, Auckland, †New Zealand Civil Aviation Authority, PO Box 31 441 Lower Hutt, ‡Allan Wilson Centre for Molecular Ecology and Evolution, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand, §Griffith School of Environment and School of Biomolecular and Physical Sciences, Griffith University, 170 Kessels Road, Qld 4111, Australia

Abstract

While flying remains one of the safest means of travel, reported birdstrikes on aircraft have risen. This is a result of increased aircraft flight movements, changes in agricultural methods and greater environmental awareness contributing to growing populations of hazardous bird species, as well as more diligent reporting of incidents. Measures to mitigate this hazard require accurate data about the species involved; however, the remains of birds from these incidents are often not easy to identify. Reported birdstrikes include a substantial number where the species cannot be determined from morphology alone. DNA barcoding offers a reliable method of identifying species from very small amounts of organic material such as blood, muscle and feathers. We compare species identification based on morphological criteria and identifications based on mitochondrial cytochrome *c* oxidase subunit I DNA barcoding methods for New Zealand species. Our data suggest that DNA-based identification can substantially add to the accuracy of species identifications, and these methods represent an important addition to existing procedures to improve air safety. In addition, we outline simple and effective protocols for the recovery and processing of samples for DNA barcoding.

Keywords: avian, birdstrike, COI, *cox*1, DNA barcode, species identification

Received 10 January 2010; revision accepted 26 April 2010

Introduction

DNA barcoding using a 648-bp segment of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene is now an established method of identifying animal species (Rubinoff 2006; Waugh 2007; Frézal & Leblois 2008). The within-species variation in nucleotide sequences for this region of the gene is typically about 10 fold less than that observed between congeneric species (Hebert *et al.* 2003). Thus, DNA barcoding provides a molecular means of identifying species from a variety of tissue types including blood, muscle, skin and feathers. One of the key advantages of DNA barcoding is that in most cases, accurate species identifications can be made from very small amounts of sample (micrograms) that are otherwise unidentifiable.

Species identification is however dependent on the existence of a reference database of DNA barcodes from vouchered specimens (specimens identified by a taxonomist and catalogued within a museum or similar collection). Producing this database is a global effort, co-ordinated by the Consortium for the Barcode of Life available at http://barcoding.si.edu/. The associated Barcode of Life Data Systems (BoLD) Website (http://www.barcodeoflife.org/) provides a repository for DNA barcodes as well as an 'Identification Engine'. The identification engine matches a DNA sequence from a sample with known barcodes from vouchered specimens, thus providing a means of identifying species. At the present

Correspondence: John Waugh, Fax: +64-9-441-8142;
E-mail: W.J.Waugh@massey.ac.nz

time, the database has >800 000 DNA barcodes from >70 000 different species (BoLD). This database represents a useful resource for identifying tissue samples.

DNA barcoding has been shown to be particularly effective in identifying a wide range of northern hemisphere bird species (Hebert *et al.* 2004a; Yoo *et al.* 2006; Kerr *et al.* 2007), and the world's avian fauna has been chosen as one of the first groups on which to test its overall efficacy. Birds have been intensively studied using a range of techniques such as morphometrics, cytogenetics and behaviour analysis, and consequently, the number of avian species has remained relatively stable (Hebert *et al.* 2004a). Several projects to gather DNA barcodes of birds from different geographical regions have already been published (Hebert *et al.* 2004a; Yoo *et al.* 2006; Kerr *et al.* 2007) and several more are presently being undertaken. To date, >17 000 birds from nearly 3000 species have had their DNA barcodes recorded (BoLD 2010). Of these, nearly 1000, comprising around 200 bird species* (including >90% of the native and endemic species), are from the New Zealand region and are available for matching with DNA recovered from sample tissue (BoLD 2010). Many introduced or circumpolar species in New Zealand have also been DNA barcoded both in New Zealand and elsewhere. Samples for barcoding have been recovered from birds that are as widely distributed as possible.

Applications of DNA barcoding include the identification of species difficult to identify in the field (Hebert *et al.* 2003), the monitoring of fisheries (Schander & Willassen 2005), the identification of species refractory to other techniques (Hebert *et al.* 2003), as well as matching different life stages of the same species (Hebert *et al.* 2004b). Important among these applications is the identification of species involved in birdstrikes on aircraft (Dove *et al.* 2007; Marra *et al.* 2009; Yang *et al.* 2009).

The global cost of damage, delays and cancellations caused by birdstrikes is estimated in the billions of US dollars each year; moreover, birdstrikes have caused considerable loss of human life (Allan 2006). In a 19-year period to 2008 in the United States, the average reported cost of direct damage to civil aircraft birdstrikes was $US117 787 per incident (Dolbeer *et al.* 2009). This rose to $US153 790 per incident when other costs such as aircraft down time were included.

Birdstrike events are not unusual. In the same 19-year period (1990–2008), there were 89 727 reported birdstrike incidents in the United States (Dolbeer *et al.* 2009). Transport Canada reported 6848 birdstrikes between 1991 and 1999 (Transport Canada 2004). In New Zealand, a country of approximately four million people, there were 3199 reported birdstrike incidents in the 7 years to October 2009 (M. W. Evans, personal communication). However, evidence suggests that only approximately 20% of birdstrikes are reported worldwide by aircrew or ground maintenance staff (Transport Canada 2004). Moreover, reported birdstrikes have increased (Dolbeer *et al.* 2009; Transport Canada 2004; M. W. Evans, personal communication). In the United States, 1759 birdstrikes were reported to the Federal Aviation Authority in 1990. By 2008 that figure had risen to 7516 (Dolbeer *et al.* 2009). These figures are reflected elsewhere (Transport Canada 2004; M. W. Evans, personal communication).

Most birdstrike incidents (>90%) occur in the proximity of airports at low altitude during landing and take off procedures (Dolbeer *et al.* 2009; Transport Canada 2004). Damage caused by birdstrikes is related to impact force, and impact force is proportional to the square of the aircraft's velocity. Thus, at normal take off and landing speeds, a 2-kg bird may have an impact force of 15 000–40 000 kg. The majority of serious incidents involve medium to larger bird species (Transport Canada 2004), but even small birds can cause major damage (Smithsonian National Museum 2009). The behaviour patterns of bird species in areas adjacent to airports is sometimes predictable and can often be changed by appropriate wildlife management intervention (Transport Canada 2004). Methods of management start with ensuring the environment in the locality of the airport is not unnecessarily attractive to birds (relocating landfill sites, reducing artificial ponds etc.). However, the location of many airports is unavoidably in areas that may be attractive to birds, and furthermore, airports themselves, with their large open grassy spaces, are highly attractive to some species (Caithness *et al.* 1967).

Often, all that remains of the bird after a birdstrike incident is a smear of blood, feathers or tissues. Until recently, these samples were identified by using feather identification that requires access to a comprehensive library of feathers and considerable expertise on the part of the identifier (Dove *et al.* 2007) or by a system of keratin protein electrophoresis (Ouellet 1994). Blood or other tissue was unidentifiable until the advent of DNA sequencing technologies. Approximately, a quarter of all reported birdstrikes are for species that cannot be identified by morphological means (Dolbeer *et al.* 2009; Transport Canada 2004; M. W. Evans, personal communication). Several institutions around the world now employ DNA barcoding to identify the species involved including the authors of this paper, the Smithsonian Institute and Yang *et al.* in China (Dove *et al.* 2007; Marra *et al.* 2009; Yang *et al.* 2009). This approach is effective with all types of tissue recovered, and the data collected can be used to inform wildlife management programs at airports.

*This database is being continually updated.

With ever increasing hazard management requirements being adopted worldwide, there is a greater emphasis on species identification from birdstrike incidents. The aim of this study is to present data for birdstrike incidents in New Zealand, compare the use of morphological and DNA barcoding techniques in identifying birdstrike species and to outline simple and effective protocols for the collection and laboratory processing of samples from aircraft.

Method

Data on birdstrikes in New Zealand were supplied by the Chairman of the National Wildlife Hazard Committee of the New Zealand Civil Aviation Authority (NZCAA) (M. W. Evans, personal communication). Birdstrike data were collected from 68 airports around the country (New Zealand), ranging from small country airfields to international terminals.

Samples for DNA barcoding were collected at Christchurch International Airport in the South Island (D. Bennet, personal communication) and Wellington International Airport in the North Island of New Zealand (J. Eyely, personal communication). Once collected, they were sent to the laboratory by 'normal mail' for analysis.

Sample collection employed forensic techniques adapted from methods used by police scene of crime officers (SOCO) for collection of DNA samples. Sampling kits consisted of a Ziploc® collection bag containing a pair of examination gloves, a cultiplast® ∈ 0051 sterile swab in a plastic tube, a 5-mL plastic ampoule of sterile water, two alcohol swabs, a data sheet and a set of instructions. The tip of the plastic tube was cut to allow air to circulate and dry the sample after collection. Two drops of sterile water were dropped onto the sterile swab; the blood smear was then wiped with the swab to collect the blood sample before returning it to the cut tube. Some samples (BS18, 19, 22, 23, 27, 31, 32, 34, 35, 36 and 39) were collected on Whatman FTA® cards (for details of use, see http://www.whatman.com/FTANucleicAcidCollectionStorageandPurification.aspx). Because surfaces on an aircraft may be contaminated with aviation fuel, cleaning agents and lubricants known to damage DNA, samples were collected from the most uncontaminated areas available.

In the laboratory, DNA was extracted from samples using either standard proteinase K digestion and a phenol/chloroform extraction and ethanol precipitation or for later samples, by one step Chelex® 100 resin (Bio-rad Laboratories)/proteinase K digestion.

PCRs were carried out in 25-μL or 10-μL volumes consisting of 10x PCR buffer (Invitrogen), 2.5 mM $MgCl_2$, 10 ng/mL BSA, 0.5 μM of each primer, 0.2 mM of each dNTP, 0.4 U of Platinum *Taq* DNA polymerase (Invitrogen) and 1–2 μL of DNA extract (approximately 0.1–5 ng).

The primers used for amplification of the barcoding region (approximately 650 bp of the 5′ terminus of COI) were a forward primer, AWCF1, and a reverse primer AWCR4 (Patel *et al.* 2009). For some samples, an alternative reverse primer, AWCR6 was used (Patel *et al.* 2009). DNA was amplified using an Applied Biosystems GeneAmp 9700 thermal cycler, and the cycling conditions for samples was as follows: 2 min at 94°C followed by 30 cycles of 30 s at 94°C, 20 s at 54°C and 30 s at 72°C followed in turn by 15 cycles of 30 s at 94°C, 20 s at 50°C and 30 s at 72°C.

Amplification products were visualized by electrophoresis in 1.2% agarose/0.5 × Tris Borate EDTA buffer, stained with 50 ng/mL ethidium bromide and viewed under UV light. Positive amplifications were purified using the Agencourt AMPure PCR Purification System (Beckman Coulter) or the DNA Clean & Concentrator-25™ kit (Zymo Research). All purified PCR products were then cycle sequenced in both forward and reverse direction using Big Dye v3.1 chemistry and analysed on an ABI Prism 3130xl genetic analyser (see Dove *et al.* 2007 or Hebert *et al.* 2004a for other suitable methods).

Sequences were edited using Sequencher 4.6 (Gene Codes Corporation), and identifications were carried out by matching a FASTA representation of the sequence to files lodged on the BoLD Website using their identification engine (BoLD 2010). An unambiguous match of >98% was considered a successful identification.

Results

Birdstrike statistics

There were 3199 reported birdstrikes in New Zealand in the 7 years to October 2009 or approximately 470 per annum (Fig. 1). Identification of the birds involved was by morphological methods, and the five most commonly identified taxa associated with birdstrikes were Spur-winged Plover (*Vanellus miles*) 28%, House Sparrow (*Passer domesticus*) 18%, Southern Black-backed Gull (*Larus dominicanus*) 5%, other Gull (*Larus*) species 5% and Oystercatcher (*Haematopus*) species 4%. Twenty per cent were all other identified species, and 20% were not identified (Fig. 2). As would be expected, the major airports, with much higher traffic densities and greater size, experienced the highest number of birdstrike events (M. W. Evans, personal communication). However, there was a significant (χ^2 df = 4 $P < 0.001$) difference between main airports in the percentage of birdstrike species that were not identified (Fig. 2). Year-by-year data suggest a trend towards a higher proportion of unidentified species; however, the R value for this trend is 0.73 (Fig. 3).

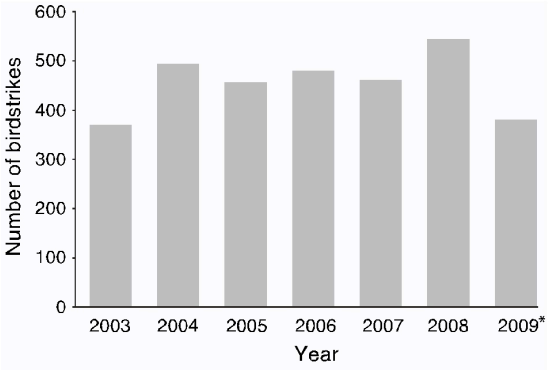


Fig. 1 Total annually reported birdstrikes from airports in New Zealand 2003–2009 (*Data for 2009 is incomplete and ends at 19th October).

DNA barcoding

In this study, DNA barcoding was generally successful at identifying samples from birdstrike incidents. In 92.5% ($N = 40$) of cases where unidentified birdstrike sample material was tested, sufficient DNA was extracted for identification purposes. There was no difference in DNA recovery between those samples collected using SOCO protocols and those collected using the Whatman FTA® cards. Of the three cases where no DNA-based identification was possible (BS21, BS26 and BS37), two (BS21 and BS37) were incorrectly collected (the tube was not cut to allow drying of the sample with the result that the sample became contaminated with mould) and the other was heavily contaminated with engine lubricants.

For the three samples identified as *Haemotopus* species (BS12, BS31 and BS36), some debate exists over the status of the various New Zealand species, and the identification engine could not differentiate between species within this genus, despite high quality sequence being obtained from all these samples. When sequences from these samples were matched with those on the BoLD database using their identification engine, they could not be distinguished from four putative species (*finschi, unicolor, longirostris* and *chathamensis*). Samples BS19 and BS24, possibly from the genus *Aquila*, could not be accurately identified using the BoLD identification engine because, despite high-quality DNA being extracted in both cases (Phred score >20 BS19 97.5% and BS24 97.4%), no record of this species exists on the BoLD Website. The same applies to the unknown Passerine BS35 (98.9%) and to the unknown Charadriiforme BS39 (81.2%).

Discussion

Birdstrike statistics from around the world show similar trends that reflect global changes in agricultural methods, environmental awareness and the need to report and record details of these events as a resource for those charged with reducing their incidence (Dolbeer *et al.* 2009; Transport Canada 2004; M. W. Evans, personal communication). A reduction in the use of pesticides such as DDT has resulted in increases in bird populations, as has greater urban green space and set aside wilderness areas (Dolbeer *et al.* 2009).

Analysis of the data for New Zealand shows no statistically significant change in birdstrike statistics for the years 2003–2009. Furthermore, although a positive trend is apparent in the number of birdstrike species that are unidentified (Fig. 3), this was also not statistically significant. In common with many other countries, the location of New Zealand airports in mainly coastal or rural areas predisposes them to higher risk from birdstrike. Moreover, the proximity of bird sanctuaries and breeding areas to many airports adds to this risk. Nonetheless, a high degree of awareness and active wildlife management programmes have, so far, prevented the worst outcomes from these events.

The NZCAA data presented here are frequently nonspecific. For example, there are two Gull species apart from the Black-backed Gull that are only identified as 'Gulls'. No attempt is made to differentiate Duck, Dotterel, Finch, Heron, Owl, Oystercatcher, Pigeon, Stilt or Tern species in the these data. Furthermore, the fact that many of the people reporting birdstrikes to the NZCAA have limited knowledge of birds means that identifications may be questionable. Many small species are incorrectly identified as sparrows. Thus, the number of incorrectly identified or unidentified species is probably far higher than the reported 20%. However, because there is no other species that resembles the Spur-winged Plover (*Vanellus miles*) in New Zealand and because this species is so frequently involved in birdstrikes, even the relatively unskilled observer can correctly identify it. Thus, data for this species are probably reasonably accurate.

There is significant ($P < 0.001$) variation in the percentage of specimens that could not be identified at each of the main airports (Fig. 2). This may be because in some airports, only a limited number of species are involved in incidents whereas in others, the avifauna is more diverse and identification is correspondingly more difficult. Within New Zealand the Civil Aviation Rule Part 12 it is required that all bird strikes and near misses must be reported (Civil Aviation Authority of New Zealand 2006). Furthermore, the International Civil Aviation Organisation (ICAO) has a requirement that all countries should forward wildlife strike reports to ICAO for inclusion in the ICAO Bird Strike Information System database (International Civil Aviation Authority 2001). However, in some cases, species cannot be identified and it is not mandatory to do so.

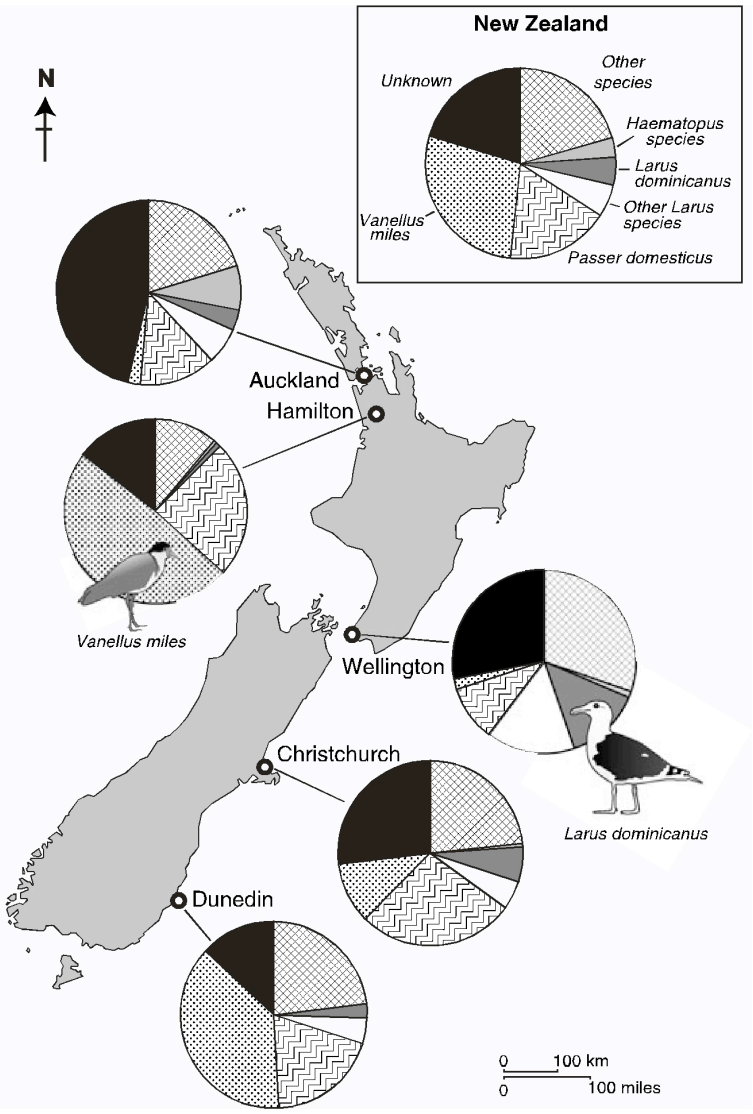


Fig. 2 The species most highly represented in reported birdstrikes in New Zealand overall and at the main airports within New Zealand between 1st January 2003 and 19th October 2009.

A variety of new and ingenious ways are being devised to combat the threat from birdstrike. Many of these require a good understanding of the species involved. Species differ from place to place but, despite local variation, the overall percentage that cannot be identified by conventional means is comparable in the United States, Canada and New Zealand (20–25%) (Dolbeer *et al.* 2009; Transport Canada 2004; M. W. Evans, personal communication). However, analysis of the New Zealand data indicates that the overall number of unidentified species is much higher. The United Kingdom Birdstrike Committee reports 60% of species involved in birdstrikes are unidentified (UK Birdstrike Committee Meeting (UKBSC) 2009), as does the German Committee on Prevention of Birdstrikes (C. Morgenroth, personal communication).

DNA sequences for six samples could not be identified to species level. For samples BS12, BS31 and BS36 (*Haemotopus spp.*), some debate exists over the species status of the three taxa (*ostralegus, unicolor* and *chathamensis*) found in the New Zealand region (Oliver 1955; Baker 1974; Falla *et al.* 1975; Turbott 1990; Heather & Robertson 2005). Banks & Paterson (2007) found little difference between the New Zealand species despite testing a variety of mitochondrial markers, including the hypervariable region (HVR I) of the control region, the gene coding for the small ribosomal subunit (12S) as well as the COI barcoding region and the Cytochrome *b* gene. Thus, fail-

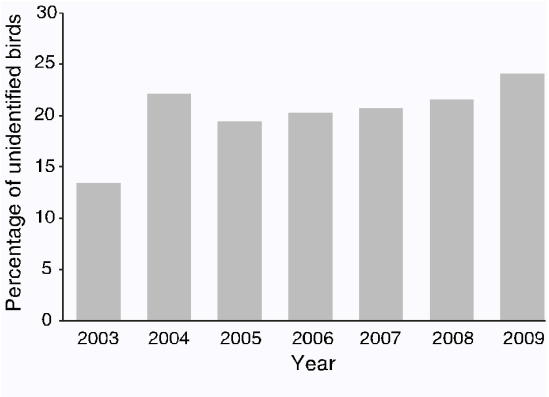


Fig. 3 The percentage of unidentified birdstrikes in New Zealand by year between 1st January 2003 and 19th October 2009.

ure to resolve these samples to species level may be an artefact of incorrect taxonomy rather than any inherent problem with the method. Conversely, speciation within this genus may have been associated with little divergence at the molecular level (Lambert & Paterson 1993; Ferguson 2002) or perhaps there has been little time for the genetic divergence typically observed between species. In either case, DNA barcoding will be incapable of distinguishing them.

Two samples from the Falconiformes, possibly from the genus *Aquila* (samples BS19 and BS24), one Passeriforme (BS35) and one Charadriiforme (BS39) could not be identified to species level using DNA barcoding because the BoLD database does not contain any matching sequence. The first two originated from the same incident but were separately submitted for identification. They were found in the damaged wing flap of an aircraft in New Zealand that had recently arrived from Australia. High-quality sequence from both samples revealed no matches with database species. New Zealand Falconiformes have had their DNA barcodes characterized and these are available in the BoLD database; however, the BoLD Website does not have comprehensive DNA barcodes for the avifauna of Australia. It is probable that the birdstrike occurred in Australia. The Passerine sample was recovered from an aircraft inbound from Nadi in Fiji and, despite the recovery of good DNA sequence, would not match with any sequence on the BoLD database. The Charadriiforme (probably a Tern) was the only bird of New Zealand Origin that could not be identified. These birds are one of the few groups that have not had their DNA barcodes comprehensively characterized within the New Zealand avifauna.

The above examples demonstrate some of the issues that remain to be resolved with the application of DNA barcoding to birdstrikes. DNA barcoding has identified many potential synonymous and cryptic species and further taxonomic work is required to resolve these. Nonetheless, identification to genus level, in the case of the *Haematopus* species, will probably provide sufficient information for wildlife management purposes in this particular case. However, the unidentified Falconiforme, Passeriforme and Charadriiforme highlight a need for further development of the BoLD database to include species from more regions around the world.

Within the barcoded sample data, there appear to be a number of species that seem to be unusually represented. The Spur-winged Plover (*Venellus miles*), a medium sized bird (350 g), is a relatively recent migrant to New Zealand with the first breeding pair being noted in 1932 (Heather & Robertson 2005), but it is implicated in 28% of all morphologically described birdstrike events around the country (M. W. Evans, personal communication) and its numbers are increasing. However, only 5% (BS3 & BS10) of the samples tested (χ^2 df = 1 $P < 0.05$) were of this species. Christchurch and Wellington have below national average birdstrikes with *V. miles*, but these figures may also be an artefact of sample size. In addition, because it is so frequently implicated in birdstrikes, airport staff may have become skilled at identifying remains from this species and, therefore, may not send it to be DNA barcoded as often as other less frequently observed species.

Seventeen and a half per cent ($N = 7$) of the samples DNA tested were Southern Black-backed Gulls (*Larus dominicanus*) compared with a morphologically described national average of 6%. Of these, five came from Wellington International Airport. This represents >38.5% of samples sent from Wellington (χ^2 df = 1 $P < 0.001$). This airport is situated on an isthmus that is surrounded by sea. There is a breeding colony of this large gull species on an island about three kilometres away. *Larus dominicanus* is also known to prey on other species of bird, and these factors may account for its unusual preponderance in the samples sent from this source.

Of the remaining birds represented more than once in the DNA tested specimens, Rock Pigeons (*Columba livia*) occurred in 7.5% of DNA tested specimens vs. 1.6% of the morphologically identified specimens, Oystercatchers (*Haematopus spp.* see above) in 7.5% vs. 2.5%, Skylarks (*Alauda arvensis*,) occurred in 5% vs. 0.8% and Australian Magpies (*Gymnorhina tibicen*,) occurred in 5% vs. 0.9% (Table 1). Thus, the frequency of the DNA identified samples did not reflect that of the morphologically identified samples.

Had the frequency of different species identified using DNA barcoding been very similar to the frequency of the morphologically identified specimens, the argument that not much is to be gained by carrying out DNA tests would be compelling. However, the opposite is true, and it appears that much of the data supplied to the NZCAA

Table 1 Results of DNA barcode analysis of specimens recovered from aircraft birdstrike incidents during the study. Samples are identified of occurrence of the birdstrike incidents between 3rd June 2008 and 14th January 2010

ID no. (% match)	Date of incident	Common name	Species
BS1 (98.3)	03/06/08	Australasian Harrier Hawk	*Circus approximans*†
BS2 (99.8)	05/06/08	Yellowhammer	*Emberiza citrinella**
BS3 (100)	16/06/08	Spur-winged Plover	*Vanellus miles*
BS10 (100)	07/09/08		
BS4 (100)	23/06/08	White-fronted Tern	*Sterna striata*‡
BS5 (100)	25/06/08	Southern Black-backed Gull	*Larus dominicanus*‡
BS7 (100)	27/07/08		
BS15 (99.8)	05/11/08		
BS20 (100)	25/02/09		
BS28 (100)	30/05/09		
BS33 (100)	26/09/09		
BS40 (100)	23/13/10		
BS6 (100)	11/07/08	Australian Magpie	*Gymnorhina tibicen**
BS14 (99.7)	17/10/08		
BS8 (100)	04/08/08	Goldfinch	*Carduelis carduelis**
BS9 (100)	11/08/08	Banded Dotterel	*Charadrius bicinctus*†
BS11 (100)	16/10/08	Starling	*Sturnus vulgaris**
BS12 (100)	16/10/08	Oystercatcher	*Haematopus spp. (unicolor ostralegus)*
BS31§ (100)	21/08/09		
BS36§ (100)	03/01/10		
BS13 (99.8)	17/10/08	Black-fronted Tern	*Chlidonius albostriatus*†
BS16 (100)	25/11/08	Black-billed Gull	*Larus bulleri*†
BS17 (98.8)	25/11/08	Little Owl	*Athene noctua**
BS18§ (94.4)	29/01/09	Paradise Shelduck hybrid	*Tadorna spp.*
BS19§ (93.4)	05/02/09	Unknown	Closest match *Aquila spp.*
BS24 (93.4)	05/02/09		
BS21	23/03/09	No DNA recovered	N/A
BS26	14/04/09		
BS37	12/01/10		
BS22§ (100)	20/03/09	Skylark	*Alauda arvensis**
BS32§ (100)	21/08/09		
BS23§ (99.7)	23/03/09	New Zealand Pipit	*Anthus novaeseelandiae*‡
BS25 (100)	26/04/09	Blackbird	*Turdus merula**
BS27§ (100)	19/05/09	Rock Pigeon	*Columba livia**
BS34§ (100)	07/11/09		
BS38 (99.7)	14/01/10		
BS29 (100)	26/06/09	Greenfinch	*Carduelis chloris**
BS30 (100)	10/07/09	Black Swan	*Cygnus atratus**
BS35§ (91.9)	05/12/09	Unknown	*Passerine*
BS39§	06/03/10	Unknown	Closest match *Hydroprogne sp.*

*Introduced species.
†Endemic species.
‡Native species.
§Sample collected on FTA card.

may be inaccurate. Bearing in mind that, to date, DNA barcoding has only been used when morphological means have failed, the figure of 82.5% identifications of these otherwise unidentifiable samples provides a benchmark for the number of species that could be identified and if applied to all samples would provide wildlife officers with an almost complete data set to inform their management strategies. Because DNA barcoding is based on very rapidly developing technologies, costs are likely to be reduced and processing times decreased in the future.

For more than a century, the hazards of air travel have been steadily decreasing. More sophisticated aircraft design and engineering, improved air traffic management and greater emphasis on hazard management within the industry have led to a situation where the safety of air travel rivals or exceeds that of almost every other form of transportation. However, despite this, the

high death to injury ratios and the number of fatalities associated with aircraft crashes continue to cause public concern. One hazard that has increased is that from birdstrikes. The aim of reporting birdstrikes is to reduce the hazard from this source but without accurate identification of the species involved, the data are of limited use to wildlife managers. Morphological identification techniques combined with circumstantial data can be very accurate when carried out by experts. In the absence of this expertise or where samples are otherwise unidentifiable, DNA barcoding provides a cost effective and accurate means of identifying samples collected from these incidents that is likely to be adopted to comply with increasingly stringent regulations.

Acknowledgements

We are indebted to Della Bennet, Jon Eyely and Bill Pohatu for sending samples for identification, Andrew Cridge and Leon Huynen for commenting on this manuscript, the NZCAA for allowing us to use data relating to birdstrike records in New Zealand and to Leon Huynen and Selina Patel for helping with the laboratory processing of samples.

References

Allan J (2006) A heuristic risk assessment technique for birdstrike management at airports. *Risk Analysis*, **26**, 723–729.

Baker AJ (1974) Ecological and behavioural evidence for the systematic status of New Zealand oystercatchers (Charadriiformes: Haematopodidae). *Life Science Contributions of the Royal Ontario Museum*, **96**, 1–34.

Banks JC, Paterson AM (2007) A preliminary study of the genetic differences in New Zealand oystercatcher species. *New Zealand Journal of Zoology*, **34**, 141–144.

Barcode of Life Data Systems (BoLD) website http://www.barcodeoflife.org/

Caithness TA, Williams MJ, Bull RM (1967) Birds and aircraft: a problem on some New Zealand Airfields. Department of Internal Affairs publication, Wellington. 58–62.

Civil Aviation Authority of New Zealand (2006) Civil aviation rules part 12, amendment 2. Accessed on 12 November 2009. http://www.caa.govt.nz/rules/Amendments/Part_012_Amendment_02.pdf

Consortium for the Barcode of Life (CBOL) website http://barcoding.si.edu/

Dolbeer RA, Wright SE, Weller J, Begier MJ (2009) Wildlife strikes to civil aircraft in the United States 1990-2008. Accessed 26/02/10. http://wildlife-mitigation.tc.faa.gov/wildlife/downloads/BASH90-08.pdf

Dove CJ, Rotzel NC, Heacker M, Weigt LA (2007) Using DNA barcodes to identify bird species involved in birdstrikes. *The Journal of Wildlife Management*, **72**, 1231–1236.

Falla RA, Sibson RB, Turbott EG (1975) *The New Guide to the Birds of New Zealand*. Auckland, Collins.

Ferguson JWH (2002) On the use of genetic divergence for identifying species. *Biological Journal of the Linnean Society*, **75**, 509–516.

Frézal L, Leblois R (2008) Four years of DNA barcoding: current advances and prospects. *Infection Genetics and Evolution*, **8**, 727–736.

Heather BD, Robertson HA (2005) *The Field Guide to the Birds of New Zealand*. Auckland, Viking.

Hebert PDN, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B*, **270**, S96–S99.

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004a) Identification of birds through DNA barcodes. *PLoS Biology*, **2**, 1657–1663.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004b) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 14812–14817.

International Civil Aviation Authority, (2001) ICAO position and activities with respect to bird strikes. Accessed 12 November 2009. http://www.icao.int/icao/en/ro/nacc/acilac/02_pinos_icao_birdcontrol1.pdf

Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, Hebert PDN (2007) Comprehensive DNA barcode coverage of North American birds. *Molecular Ecology Notes*, **7**, 535–543.

Lambert DM, Paterson HE (1993) Morphological resemblance and its relationship to genetic distance measures. In: *Reprinted in The Recognition Concept of Species* (ed. McEvey S), pp. 63–75. Johns Hopkins University Press, Baltimore, Maryland.

Marra PP, Dove CJ, Dolbeer R *et al.* (2009) Migratory Canada geese cause crash of US Airways Flight 1549. *Frontiers in Ecology and the Environment*, **7**, 297–301.

Oliver WRB (1955) *New Zealand Birds*, 2nd edn. A. H. & A. W. Reed, Wellington.

Ouellet H (1994) Keratin protein electrophoresis and the identification of feather remains: new developments and update. *Proceedings of the Bird Strike Committee Europe*, **22**, 499–512.

Patel S, Waugh J, Millar CD, Lambert DM (2009) Conserved primers for DNA barcoding historical and modern avian samples. *Molecular Ecology Resources*, **10**, 431–438.

Rubinoff D (2006) DNA barcoding evolves into the familiar. *Conservation Biology*, **20**, 1548–1549.

Schander C, Willassen E (2005) What can biological barcoding do for marine biology? *Marine Biology Research*, **1**, 79–83.

Smithsonian National Museum (2009) Smithsonian National Museum of Natural History. Accessed 17/08/09 http://www.mnh.si.edu/highlight/feathers/

Transport Canada (2004) *Sharing the Skies: An Aviation Industry Guide to the Management of Wildlife*, 2nd edn. Civil Aviation, Transport Canada.

Turbott EG (1990) *Checklist of the Birds of New Zealand*, 3rd edn. Random Century, Auckland.

UK Birdstrike Committee Meeting (UKBSC). Notes of Meeting held Thursday 23 April (2009) CAA Safety Regulation Group

Waugh J (2007) DNA barcoding in animal species: progress, potential and pitfalls. *Bioessays*, **29**, 188–197.

Yang R, Wu X, Yan P, Li X (2009) Using DNA barcodes to identify a bird involved in a birdstrike at a Chinese airport. *Molecular Biology Reports*. Accessed 26/02/10 http://www.springerlink.com/content/f6762m2818704m32/

Yoo HS, Eah JY, Kim JS *et al.* (2006) DNA barcoding Korean birds. *Molecules and cells*, **22**, 323–327.

Appendix B

BOLD data upload spreadsheets

Data Submission Protocol

This protocol assists in the submission of bulk data to BOLD. It describes the necessary format of the data that is required for a correct submission.

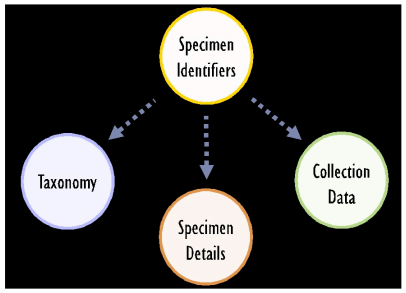
Whenever a bulk submission is sent to the data manager, the following pieces of information need to be sent along with the standard submission spreadsheet:

I. Project title
II. Project code
III. Project manager
IV. Priority Level (High, Intermediate or Low)
V. Submission type (New Records or Update)*

** If type is Update : Please specify which worksheets (Voucher Info, Taxonomy, Specimen Details, or Collection Data) need to be updated.*

See next page for more information.

The data spreadsheet consists of 4 worksheets, a specimen identifier worksheet (voucher info) that is referenced by a taxonomy, specimen details, and collection data sheet (Field definitions on Table 1 to 4).



Sample ID	ID associated with sample begin sequenced (often an extension of field or Museum Id)
Field ID	Specimen identifier from a private collection or Field number from a collection event.
Museum ID	Catalog number in curated collection for a vouchered specimen
Collection Code	Code associated with a given collection
Institution Storing	Institution where specimen is vouchered (controlled list)
Sample Donor	Individual or institution responsible for providing specimen or tissue sample.
Donor Email	Email of the sample donor

Table 1 : Field definitions for Voucher Info page on accompanying spreadsheet

Full Taxonomy	Full taxonomy consisting of plylum, class, order, family, subfamily(optional), genus, species binomial
Identifier	Primary individual responsible for providing taxanomic identification of the specimen
Identifier Email	Email address of the primary identifier
Identifier Institution	Institution of the identifier

Table 2 : Field definitions for Taxonomy page on accompanying spreadsheet

Sex	Male/female/hermaphrodite (controlled list)
Reproduction	Sexual/asexual/cyclic parthenogenesis (controlled list)
Life Stage	Adult/immature (controlled list)
Extra Info	User Specified Characteristic (free text) - Can be displayed on tree or used to sort records, max 50 characters
Notes	Free text or XML tagged text. All XML text should be surrounded by the XML start(<xml>) and stop (</xml>) tags.

Table 3 : Field definitions for Specimen Details page on accom-

Collectors	Comma delimited list of collectors
Collection Date	Date of collection, must be in MMM-DD-YYYY format.
Country/FAO region	ISO Country name, Ocean, or FAO region
State/Province	States and provinces (Controlled list according to Getty Geographical Thesaurus)
Region	Park, county, district, lake or river
Sector	Sector of park or county
Exact Site	Description of collection location
GPS Coordinates	Latitude & Longitude in "degrees.decimal degrees" format (e.g. 45.837)
Elevation/Depth	Elevation or depth in meters

Table 4 : Field definitions for Collection Data page on accompanying spreadsheet

Data Submission Protocol Continued

There are two types of submissions: "New Submission" and "Update".

1. *New Submission*

New submissions are project specific, so that their data can be associated with a project on BOLD. If records are submitted that need to be entered into different projects on BOLD, a separate file for each project needs to be sent.

The minimal requirements for a new submission on BOLD are:
•Voucher Info Page - Sample ID
•Voucher Info Page - Field ID and/or Museum voucher ID
•Voucher Info Page - Institution Storing
•Taxonomy Page - Phylum

Other useful information:
Try to be unique when making up the format of your sample IDs. If the sample IDs you provide are not original to BOLD, you will have to change them before the data can go online.
The more data and details you provide with your new submissions, the less time it will take later to try and update the blanks.

2. *Updated Submission*

The quickest way to update data is to download the Data Spreadsheet from BOLD containing the records that need to be modified. Only download the worksheets and records that will be affected by the update (e.g. if the taxonomy needs to be updated only download the Taxonomy worksheet, if specimen details and collection date need to be update only download the Specimen Details and Collection Data worksheets, etc.). Once the worksheets are downloaded, modify the data and copy it into the standard submission spreadsheet. The submitted update should reflect what the data should be on BOLD.

NOTE: All empty fields will be blanked! The computer cannot distinguish between "blank: do not update this field or "blank: delete the content of this field".

Updates to Voucher Info are slightly different from updates to Taxonomy, Specimen Details, and Collection Data.

a.) Updates to Voucher Info

Identical to new submissions, updates to the voucher info are project specific. The records need to be split into their corresponding project.

b.) Updates to Taxonomy, Specimen Details, and Collection Data

Updates to taxonomy, specimen details, and collection data are project independent. Records from any number of projects can be submitted in one submission spreadsheet, and the number of records are (in theory) infinite for this type of update.

Voucher Information						
Sample ID	Field ID	Museum voucher ID	Collection Code	Institution Storing	Sample Donor	Donor Email
JW791		P491		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW792		P492		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW794		P494		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW795		P495		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW796		P496		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW807		P497		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW808		P423		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW809		P424		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW810		P535		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW812		P537		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW813		P539		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW814		P545		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW816		P551		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW817		P553		Allan Wilson Centre	Craig Millar	cd.millar@auckland.ac.nz
JW876	57130		SP13	Allan Wilson Centre	Shaun O'Connor	smoconnor@doc.govt.nz
JW879	28298		SP18	Allan Wilson Centre	Shaun O'Connor	smoconnor@doc.govt.nz
JW880	57106		SP28	Allan Wilson Centre	Shaun O'Connor	smoconnor@doc.govt.nz
JW881	57105		SP31	Allan Wilson Centre	Shaun O'Connor	smoconnor@doc.govt.nz
JW882	57101		SP33	Allan Wilson Centre	Shaun O'Connor	smoconnor@doc.govt.nz
JW883	28279		SP38	Allan Wilson Centre	Shaun O'Connor	smoconnor@doc.govt.nz
JW884	57135		SP41	Allan Wilson Centre	Shaun O'Connor	smoconnor@doc.govt.nz

Taxonomy									
Sample ID	Phylum	Class	Order	Family	Genus	Species	Identifier	Identifier Email	Identifier Institution
JW791	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	griseus	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW792	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	griseus	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW794	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	griseus	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW795	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	griseus	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW796	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	griseus	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW807	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW808	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW809	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW810	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW812	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW813	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW814	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW816	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW817	Chordata	Aves	Procellariiformes	Procellariidae	Puffinus	carneipes	Craig Millar	cd.millar@auckland.ac.nz	Allan Wilson Centre
JW876	Chordata	Aves	Charadriiformes	Glareolidae	Thinornis	novaeseelandiae	Shaun O'Connor	smoconnor@doc.govt.nz	Dept of Conservation
JW879	Chordata	Aves	Charadriiformes	Glareolidae	Thinornis	novaeseelandiae	Shaun O'Connor	smoconnor@doc.govt.nz	Dept of Conservation
JW880	Chordata	Aves	Charadriiformes	Glareolidae	Thinornis	novaeseelandiae	Shaun O'Connor	smoconnor@doc.govt.nz	Dept of Conservation
JW881	Chordata	Aves	Charadriiformes	Glareolidae	Thinornis	novaeseelandiae	Shaun O'Connor	smoconnor@doc.govt.nz	Dept of Conservation
JW882	Chordata	Aves	Charadriiformes	Glareolidae	Thinornis	novaeseelandiae	Shaun O'Connor	smoconnor@doc.govt.nz	Dept of Conservation
JW883	Chordata	Aves	Charadriiformes	Glareolidae	Thinornis	novaeseelandiae	Shaun O'Connor	smoconnor@doc.govt.nz	Dept of Conservation
JW884	Chordata	Aves	Charadriiformes	Glareolidae	Thinornis	novaeseelandiae	Shaun O'Connor	smoconnor@doc.govt.nz	Dept of Conservation

Specimen Details		
Sample ID	Sex	Life Stage
JW791	female	adult
JW792	female	adult
JW794	male	adult
JW795	male	adult
JW796	female	adult
JW807	female	adult
JW808	male	adult
JW809	female	adult
JW810	male	adult
JW812	male	adult
JW813	male	adult
JW814	male	male
JW816	male	immature
JW817	male	adult
JW876	male	adult
JW879	male	adult
JW880		immature
JW881		immature
JW882		immature
JW883	female	adult
JW884		adult

Collection Information										
Sample ID	Collectors	Collection Date	Region	Country	State/Province	Region	Exact Site	Latitude	Longitude	Elevation
JW791	Craig Millar	26/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW792	Craig Millar	26/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW794	Craig Millar	26/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW795	Craig Millar	26/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW796	Craig Millar	26/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW807	Craig Millar	26/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW808	Craig Millar	1/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW809	Craig Millar	1/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW810	Craig Millar	1/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW812	Craig Millar	1/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW813	Craig Millar	1/12/96	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW814	Craig Millar	4/10/97	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW816	Craig Millar	4/10/97	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW817	Craig Millar	4/10/97	Oceania	New Zealand	Auckland	North-West Auckland	Bethells Beach	S 36' 53' 43	E 174' 26' 46	8 m
JW876	Shaun O'Connor	28/2/99	Oceania	New Zealand	Chatham Islands	South East Island	Thinornis Bay South	S 44' 20' 52	E 176' 10' 26	5 m
JW879	Shaun O'Connor	3/02/99	Oceania	New Zealand	Chatham Islands	South East Island	Nog's Folly	S 44' 20' 52	E 176' 10' 26	35 m
JW880	Shaun O'Connor	3/03/99	Oceania	New Zealand	Chatham Islands	South East Island	Island Point	S 44' 20' 52	E 176' 10' 26	50 m
JW881	Shaun O'Connor	3/03/99	Oceania	New Zealand	Chatham Islands	South East Island	Thinornis Bay	S 44' 20' 52	E 176' 10' 26	5 m
JW882	Shaun O'Connor	3/03/99	Oceania	New Zealand	Chatham Islands	South East Island	Front Landing	S 44' 20' 52	E 176' 10' 26	5 m
JW883	Shaun O'Connor	3/04/99	Oceania	New Zealand	Chatham Islands	South East Island	West Clears	S 44' 20' 52	E 176' 10' 26	15 m
JW884	Shaun O'Connor	3/08/99	Oceania	New Zealand	Chatham Islands	South East Island	Clears	S 44' 20' 52	E 176' 10' 26	15 m

Appendix C

BOLD primer list

BOLDSYSTEMS V2.5 | Management & Analysis

PRINT | HELP

Primer List

Register New Primer

User Primers

Primers : 9

Code	Marker	Submitter	Reference
AWCR3	COI-5P	John Waugh	John Waugh
AWCR2	COI-5P	John Waugh	Selena Patel
AWCR6	COI-5P	John Waugh	John Waugh
AWCR1	COI-5P	John Waugh	Jennifer Anderson
AWCR5	COI-5P	John Waugh	John Waugh
AWCR4	COI-5P	John Waugh	John Waugh
AWCF2	COI-5P	John Waugh	Selina Patel
AWCR7	COI-5P	John Waugh	Selina Patel
AWCF1	COI-5P	John Waugh	Jennifer Anderson

Public Primers

Primers : 368

Code	Marker	Submitter	Reference
1709Fg	COI-5P	Xin Zhou	Zhou, X., K. M. Kjer, and J. C. Morse. 2...
1709Fs	COI-5P	Xin Zhou	Zhou, X., K. M. Kjer, and J. C. Morse. 2...
A2590	COI-5P	Gontran Sonet	Normark et al. 1999
AHyFu-F	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
AHyFu-R	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
AHyLe-Fa	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
AHyMe-F	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
AminF	COI-5P	Ralph Imondi	Direct submission from Ralph Imondi, Coa...
AminR	COI-5P	Ralph Imondi	Direct submission from Ralph Imondi, Coa...
ANTMR1D	COI-5P	Alex Smith	
AspR1	COI-5P	-	Jean-Marc Moncalvo (U of T)
AvMiF1	COI-5P	Kevin Kerr	Kerr et al. (2009) Probing evolutionary ...
AvMiR1	COI-5P	Kevin Kerr	Kerr et al. (2009) Probing evolutionary ...
BatL5310	COI-5P	Nataly Ivanova	Robins JH, Hingston M, Matisoo-Smith E, ...
BirdF1	COI-5P	Kevin Kerr	Hebert et al 2004
BirdR1	COI-5P	Paul Hebert	Hebert et al 2004
BirdR2	COI-5P	Kevin Kerr	Hebert et al 2004
BirdR3	COI-5P	Kevin Kerr	Hebert et al 2004
C113R	COI-5P	Alex Smith	Modified by M. Alex Smith from Quek, S. ...
C1-J-1580	COI-5P	Karl Magnacca	K.N. Magnacca & M.J.F. Brown
C1-N-2191	COI-5P	Taika von Konigslow	Simon C., Frati F., Beckenbach A., Cresp...
C1-N-2263	COI-5P	Karl Magnacca	K.N. Magnacca & M.J.F. Brown
C_ANTMR1D	COI-5P	Alex Smith	Alex Smith
CAS18sF1	COI-5P	Megan A. Milton	Ji, Y.-J., D.-X. Zhang, and L.-J. He. 20...
C_FishF1t1	COI-5P	Nataly Ivanova	Ivanova NV, Zemlak TS, Hanner RH, Hebert...
C_FishF1t1/C_FishR1t1	COI-5P	Daniela Denti	Natalia V. Ivanova
C_FishR1t1	COI-5P	Nataly Ivanova	Ivanova NV, Zemlak TS, Hanner RH, Hebert...
Chelicerate-F1	COI-5P	Sarah Adamowicz	Barrett and Hebert. 2005. Identifying sp...
Chelicerate-R1	COI-5P	Sarah Adamowicz	Barrett and Hebert. 2005. Identifying sp...
Chelicerate-R2	COI-5P	Sarah Adamowicz	Barrett and Hebert. 2005. Identifying sp...
CO1UnivF	COI-5P	Shadi K. Shokralla	Mehrdad Hajibabaei
COH6	COI-5P	Dirk Steinke	Schubart et al 2007
COI50L-cal	COI-5P	Megan A. Milton	Rebecca Elbourne
COI746H-cal	COI-5P	Megan A. Milton	Rebecca Elbourne
COI748H	COI-5P	Megan A. Milton	Rebecca Elbourne, unpublished
COI908aH2	COI-5P	Rebecca F. Elbourne	Rebecca Elbourne, unpublished
COIaR	COI-5P	Megan A. Milton	Oliver Haddrath
COIART	COI-5P	Megan A. Milton	Erika S Tavares
COIbirdF1	COI-5P	Lee Weigt	Hebert et al 2004
COIbirdR1	COI-5P	Lee Weigt	Hebert et al 2004
COIbirdR2	COI-5P	Lee Weigt	Hebert et al 2004
COI_E-	COI-5P	David Porco	Title: Molecular phylogeny of naidid wor...
COI(F)	COI-5P	Ralph Imondi	Direct submission from R. Imondi, Coasta...
COILWW26F	COI-5P	Natasha R. Serrao	Wilson-Wilde Forensic Sci Med Pathol, 2...
COI(R)	COI-5P	Ralph Imondi	Direct submission from R. Imondi, Coasta...
COL6	COI-5P	Dirk Steinke	Schubart et al 2007
cox1. dino.2.r	COI-5P	Rowena F. Stern	Designed by Behzad Imanian, not publishe...
cox1.dino.f	COI-5P	Rowena F. Stern	DESIGNED BY BEHZAD IMANIAN, NOT PUBLISHE...
Cox1-RF578-F	COI-5P	Anna E. Elz	Hyde JR (Hyde, John R.), Vetter RD (Vett...
Cox1-RF-R	COI-5P	Anna E. Elz	Hyde JR (Hyde, John R.), Vetter RD (Vett...
Cox1Tyr-RF-F	COI-5P	Anna E. Elz	Hyde JR (Hyde, John R.), Vetter RD (Vett...
cox42F	COI-5P	Aron J. Fazekas	Cho, Y., Y.-L. Qiu, P. Kuhlman and J.D. ...
COXF	COI-5P	-	Iwatani et al - 2005
COXR	COI-5P	-	Iwatani et al - 2005
CrustDF1	COI-5P	Dirk Steinke	Steinke 2007 unpublished

CrustDR1	COI-5P	Dirk Steinke	Steinke 2007 unpublished
CrustF1	COI-5P	Filipe Costa	Costa F.O., deWaard, J. R., Boutillier, ...
CrustF2	COI-5P	Filipe Costa	Costa F.O., deWaard, J. R., Boutillier, ...
CRYPCOIR	COI-5P	Margaret J. Beaton	M. Beaton
C_VF1di	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
C_VF1di-t1	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
C_VF1LFt1	COI-5P	Nataly Ivanova	Ivanova NV, Zemlak TS, Hanner RH, Hebert...
C_VR1di	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
C_VR1di-t1	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
C_VR1LRt1	COI-5P	Nataly Ivanova	Ivanova NV, Zemlak TS, Hanner RH, Hebert...
dgHCO-2198	COI-5P	Christopher Meyer	Meyer, C. 2003. Molecular systematics of...
dgLCO-1490	COI-5P	Christopher Meyer	Meyer, C. 2003. Molecular systematics of...
DiamF1	COI-5P	Gary Saunders	Clayden, S.L. & Saunders, G.W. 2010. Rec...
DiamF2	COI-5P	Gary Saunders	Gary Saunders unpublished
DiamF3	COI-5P	Gary Saunders	Gary W Saunders unpublished.
DiamR1	COI-5P	Gary Saunders	Gary Saunders unpublished
DiamR2	COI-5P	Sarah E. Hamsher	Gary Saunders unpublished
DumR1	COI-5P	Gary Saunders	Saunders, G.W. 2005. Applying DNA barcod...
Ech2modF1	COI-5P	Erin Corstorphine	D. Eernisse pers. comm. for use with Hen...
enhANTr1	COI-5P	Alex Smith	Designed by M. Alex Smith
EnhLepR1	COI-5P	Mehrdad Hajibabaei	Mehrdad Hajibabaei
enhTACr1	COI-5P	Alex Smith	Designed by M. Alex Smith
enhWASPr1	COI-5P	Alex Smith	Designed by M. Alex Smith
FalcoFA	COI-5P	Paul Hebert	P.D.N. Hebert
FF2d	COI-5P	Nataly Ivanova	Ivanova NV, Zemlak TS, Hanner RH, Hebert...
Fl3-F	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
Fl3-R	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
Fl4b-F	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
FlFC3	COI-5P	Brianne St. Jacques	Unpublished data
Fish1	COI-5P	Mark Polinski	Ward et al. DNA barcoding Australia's fi...
Fish-BCH	COI-5P	Megan A. Milton	Baldwin, CC, JH Mounts, DG Smith, LA Wei...
Fish-BCL	COI-5P	Megan A. Milton	Baldwin, CC, JH Mounts, DG Smith, LA Wei...
FlSHCOX1F	COI-5P	Megan A. Milton	John Hyde designed these primers as a va...
FishF2	COI-5P	Brianne St. Jacques	Ward RD, TS Zemlak, BH Innes, PR Last an...
FishR2	COI-5P	Brianne St. Jacques	Ward RD, TS Zemlak, BH Innes, PR Last an...
FR1d	COI-5P	Nataly Ivanova	Ivanova NV, Zemlak TS, Hanner RH, Hebert...
FR1d_t1	COI-5P	Nataly Ivanova	Ivanova NV, Zemlak TS, Hanner RH, Hebert...
FucR1	COI-5P	Gary Saunders	Gary W. Saunders unpublished
GazF1	COI-5P	Gary Saunders	Saunders, G.W. 2005. Applying DNA barcod...
GazF2	COI-5P	Gary Saunders	Saunders, G.W. 2005. Applying DNA barcod...
GazR1	COI-5P	Gary Saunders	Saunders, G.W. 2005. Applying DNA barcod...
GazR2	COI-5P	Gary Saunders	Lane, C.E., Lindstrom, S. & Saunders, G....
GazR3	COI-5P	Gary Saunders	Clayden, S.L. & Saunders, G.W. 2010. Rec...
GazR4	COI-5P	Gary Saunders	Saunders, G.W. 2008. A DNA barcode exami...
GazR5	COI-5P	Gary Saunders	Saunders, G.W. 2009. Routine DNA barcodi...
GHalF	COI-5P	Gary Saunders	Saunders, G.W. 2008. A DNA barcode exami...
GHalR	COI-5P	Gary Saunders	Clarkston, B.E. & Saunders, G.W. 20.10. ...
GminF	COI-5P	Gary Saunders	Saunders, GW. unpublished
GminR1	COI-5P	Gary Saunders	Gary W. Saunders unpublished.
GrF1	COI-5P	Gary Saunders	Clarkston, B.E. & Saunders, G.W. 20.10. ...
GrF2	COI-5P	Gary Saunders	Gary Saunders unpublished
GrR1	COI-5P	Gary Saunders	Gary Saunders
GrR2	COI-5P	Gary Saunders	Gary Saunders unpublished
GWSF	COI-5P	Gary Saunders	Saunders, G.W. 2009. Routine DNA barcodi...
GWSF1	COI-5P	Gary Saunders	Gary W. Saunders
GWSF2	COI-5P	Gary Saunders	Gary W. Saunders unpublished
GWSF5	COI-5P	Gary Saunders	Saunders, G.W. 2009. Routine DNA barcodi...
GWSFa	COI-5P	Gary Saunders	Gary W Saunders unpublished
GWSFi	COI-5P	Gary Saunders	Gary W. Saunders unpublished
GWSFn	COI-5P	Gary Saunders	Le Gall, L. & Saunders, G.W. DNA barcod...
GWSFt	COI-5P	Gary Saunders	Gary W Saunders unpublished
GWSR	COI-5P	Gary Saunders	Saunders, GW. unpublished
GWSR2	COI-5P	Gary Saunders	Gary W. Saunders unpublished
GWSR3	COI-5P	Gary Saunders	Saunders, G.W. 2009. Routine DNA barcodi...
GWSR4	COI-5P	Gary Saunders	Gary W Saunders unpublished
GWSR5	COI-5P	Gary Saunders	Saunders, G.W. 2009. Routine DNA barcodi...
GWSRi	COI-5P	Gary Saunders	Gary W Saunders unpublished
GWSRx	COI-5P	Gary Saunders	Gary W. Saunders unpublished.
H600	COI-5P	Julien Lorion	Lorion Julien
H7005	COI-5P	Dirk Steinke	Hafner et al., 1994. Disparate rates of ...
HCO2198	COI-5P	Sujeevan Ratnasingham	Folmer et al. 1994
HCO2198_t1	COI-5P	Robin M. Floyd	Robin Floyd
HCOoutout	COI-5P	Roberto Guidetti	Prendini L, Weygoldt P, Wheeler WC (2005...
ITS1	COI-5P	Alex Smith	White, T. J., T. Bruns, S. Lee, and J. T...
ITS2	COI-5P	Alex Smith	White, T. J., T. Bruns, S. Lee, and J. T...
JB3	COI-5P	Sean A. Locke	Bowles J, Blair D, McManus D. 1995. A mo...
JB45	COI-5P	Sean A. Locke	Bowles J, Blair D, McManus D. 1995. A mo...
K698	COI-5P	Vazrick Nazari	Caterino & Sperling 1999
KEdtmR	COI-5P	Sarah E. Hamsher	Evans KM, Wortley AH, Mann DG. 2007. An ...
KEint2F	COI-5P	Sarah E. Hamsher	Evans KM, Wortley AH, Mann DG. 2007. An ...
KEintR	COI-5P	Sarah E. Hamsher	Evans KM, Wortley AH, Mann DG. 2007. An ...
L1_DCHIM	COI-5P	Michael G. Trizna	Virgilio M.; De Meyer M.; White I.M., an...
LCO1490	COI-5P	Sujeevan Ratnasingham	Folmer et al. 1994
LCO1490_t1	COI-5P	Robin M. Floyd	Robin Floyd
LCO_Hym	COI-5P	Karl Magnacca	Folmer et al., 1994
LCOI1490	COI-5P	Sujeevan Ratnasingham	Folmer O, Black M, Hoeh W, Lutz R, Vrije...
LepF1	COI-5P	Paul Hebert	

LepF1-short	COI-5P	Rodolphe Rougerie	Hebert et al. (2004) - Ten species in on...
LepF1_t1	COI-5P	Paul Hebert	Hebert et al. 2004
LepR1	COI-5P	Paul Hebert	
LepR1-short	COI-5P	Rodolphe Rougerie	Hebert et al. (2004) - Ten species in on...
LepR1_t1	COI-5P	Paul Hebert	Hebert et al. 2004
LHM1R	COI-5P	Hai D. T. Nguyen	Hai D.T. Nguyen
LHM2R	COI-5P	Hai D. T. Nguyen	Hai D.T. Nguyen
LHM3F	COI-5P	Hai D. T. Nguyen	Hai D.T. Nguyen
LHM4F	COI-5P	Hai D. T. Nguyen	Hai D.T. Nguyen
LHM5F	COI-5P	Hai D. T. Nguyen	Hai D.T. Nguyen
M13-29FishR2_t1	COI-5P	Fabio Colombo	FABIO COLOMBO
M1329HCO2198	COI-5P	Cristian Bernardi	FABIO COLOMBO
M13COXF	COI-5P	Monica B. J. Moniz	Mmodification of Iwatani et al 2005
M13dinoCO1F2	COI-5P	Kavish P. Chandra	Ferrell, J. F. 2008. The evaluation of D...
M13F	COI-5P	Nataly Ivanova	Messing J (1983) New M13 vectors for clo...
M13Forward	COI-5P	Monica B. J. Moniz	Boutin-Ganache I., Raposo, M., Raymon, M...
M13GazF1mod	COI-5P	Monica B. J. Moniz	Mmodification of B. Gemeinholzer, BGBM,...
M13MD1R	COI-5P	Monica B. J. Moniz	Modification of Kaczmarska et al. 2008
M13R	COI-5P	Nataly Ivanova	Messing J (1983) New M13 vectors for clo...
M13rev-bg	COI-5P	Monica B. J. Moniz	Mmodification of B. Gemeinholzer, BGBM,...
MEPTR1_t1	COI-5P	Xin Zhou	Xin Zhou, Sarah J. Adamowicz, Luke M. Ja...
MLepF1	COI-5P	Paul Hebert	
MLepR1	COI-5P	Mehrdad Hajibabaei	Mehrdad Hajibabaei
MplatCOX1dF	COI-5P	Sean A. Locke	Moszczynska, A. Locke, S.A., McLaughlin,...
MplatCOX1dR	COI-5P	Sean A. Locke	Moszczynska, A. Locke, S.A., McLaughlin,...
NancyM	COI-5P	Agnes Schrett-Major	Pfunder, M., Holzgang, O., Frey, J. E. (...
Nancy_short	COI-5P	Karl Magnacca	Simon et al., 1994
Nematode-COI-F1	COI-5P	Constantine Christopoulos	Kanzaki & Futai, 2002
Nematode-COI-R2	COI-5P	Constantine Christopoulos	Kanzaki & Futai, 2002
OribF	COI-5P	David Porco	Heethoff, M., K. Domes, M. Laumann, M. M...
Pat	COI-5P	Karl Magnacca	Simon et al. 1994
PenF1	COI-5P	-	Jean-Marc Moncalvo (U of T)
Peziz-F	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
Peziz-R	COI-5P	Scott R. Gilmore	Gilmore et al Mol. Ecol. Res. (in press)...
plat-diploCOX1F	COI-5P	Brianne St. Jacques	Moszczynska A, Locke S, McLaughlin JD, M...
plat-diploCOX1R	COI-5P	Brianne St. Jacques	Moszczynska A, Locke S, McLaughlin JD, M...
polyHCO	COI-5P	Dirk Steinke	not published yet - designed by D. Stein...
polyLCO	COI-5P	Christy M. Carr	Christina M. Carr
PolyshortCOIR	COI-5P	Christy M. Carr	Christina M. Carr
R179	COI-5P	Samuel Iglesias	Luchetti EA, Iglésias SP, Sellos DY. Chi...
R1_DCHIM	COI-5P	Michael G. Trizna	Virgilio M.; De Meyer M.; White I.M., an...
R6036R	COI-5P	Nataly Ivanova	Robins JH, Hingston M, Matisoo-Smith E, ...
RonM	COI-5P	Nataly Ivanova	Pfunder M, Holzgang O, Frey JE (2004) De...
RonM_t1	COI-5P	Nataly Ivanova	Pfunder M, Holzgang O, Frey JE (2004) De...
RonMWASPdeg_t1	COI-5P	Alex Smith	M. Alex Smith
ScyF1	COI-5P	Gary Saunders	Clarkston, B.E. & Saunders, G.W. 20.10. ...
siM13F	COI-5P	Amy C. Driskell	standard
siM13R	COI-5P	Amy C. Driskell	standard
SP6	COI-5P	Agnes Schrett-Major	www.promega.com: pGEM-T and pGEM-T Easy ...
SueF1	COI-5P	Gary Saunders	Clayden, S.L. & Saunders, G.W. 2010. Rec...
SueR1	COI-5P	Gary Saunders	Clayden, S.L. & Saunders, G.W. 2010. Rec...
sym-C1-J-1718	COI-5P	Gontran Sonet	Simon et al. 1994
T3	COI-5P	Erik J. van Nieukerken	commercially available
T7promotor	COI-5P	Erik J. van Nieukerken	commercially available
TelF1	COI-5P	Agnes Dettai	Dettai, A., Lautredou, A.-C., Bonillo, C...
TelR1	COI-5P	Agnes Dettai	Dettai, A., Lautredou, A.-C., Bonillo, C...
T-LepF1-short	COI-5P	Erik J. van Nieukerken	Hebert et al. (2004). Ten species in one...
T-LepR1-short	COI-5P	Erik J. van Nieukerken	Hebert et al. (2004). Ten species in one...
TY-J-1460	COI-5P	Taika von Konigslow	Simon, C., F. Frati, A. Beckenbach, B. C...
UBC6 F	COI-5P	Yvonne Linton	Simon et al., 1994
UBC9 R	COI-5P	Yvonne Linton	Simon et al., 1994
Uni-MinibarR1_t1	COI-5P	Nataly Ivanova	Meusnier I, Singer GAC, Landry JF, et al...
VF1	COI-5P	Bob Ward	Ward et al. 2005
VF1d	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
VF1d_t1	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
VF1i	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
VF1i_t1	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
VF2	COI-5P	Nataly Ivanova	Ward RD, Zemlak TS, Innes BH, Last PR, H...
VF2_t1	COI-5P	Nataly Ivanova	Ward RD, Zemlak TS, Innes BH, Last PR, H...
VR1	COI-5P	Nataly Ivanova	Ward RD, Zemlak TS, Innes BH, Last PR, H...
VR1d	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
VR1d_t1	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
VR1i	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
VR1i_t1	COI-5P	Nataly Ivanova	Ivanova NV, deWaard JR, Hebert PDN (2006...
VR1_t1	COI-5P	Nataly Ivanova	Ward RD, Zemlak TS, Innes BH, Last PR, H...
CAS28sB1deg	ITS2	Alex Smith	Jie et al. Molecular Ecology Notes (2003...
CAS5p8sFtdeg	ITS2	Alex Smith	modified by MAS from Jie et al 2003 MEN ...
ITSP5	ITS2	Gary Saunders	Tai, V., Lindstrom, S.C. & Saunders, G.W...
MAS-ITS2Fdeg	ITS2	Alex Smith	modified by Alex Smith from Campbell et ...
MAS-ITS2Rdeg	ITS2	Alex Smith	Modified by Alex Smith from BioControl 4...
16Sar-L	16S	Andrew J. Crawford	Kessing B, Croom H, Martin A, et al. (19...
16Sbr-H	16S	Andrew J. Crawford	Kessing B, Croom H, Martin A, et al. (19...
16SgaF	16S	Jeremy deWaard	Yamamoto, S. and T. Sota. 2007. Phylogen...
16SgaR	16S	Jeremy deWaard	Yamamoto, S. and T. Sota. 2007. Phylogen...
16SrRNA F	16S	Raniero Lorenzetti	Amplificazione 16SrRNA per identificazio...
new16SrRNAR	16S	Raniero Lorenzetti	Primer imbastito da me ed in grado di ri...
nr LR-J-12887	16S	Vazrick Nazari	new

18E-5	18S	Alex Smith	Hillis and Dixon 1991 Ribosomal DNA. Qua...
18GW-5R	18S	Alex Smith	White, T. J., T. Bruns, S. Lee, and J. W...
18H-3	18S	Alex Smith	http://www.entomology.cornell.edu/BeePhy...
18H-3R	18S	Alex Smith	http://www.entomology.cornell.edu/BeePhy...
18L	18S	Jeremy deWaard	Weigmann et al. 1999
CAM	18S	Davoodbasha Mubarak Ali	Mubarak Ali, D and N. Thajuddin (Unpubli...
FLR3	18S	Pedro M. Antunes	Gollotte, A., D. van Tuinen, and D. Atki...
FLR4	18S	Pedro M. Antunes	Gollotte, A., D. van Tuinen, and D. Atki...
rc18H	18S	Jeremy deWaard	Weigmann et al. 1999
WormA	18S	Sean A. Locke	Littlewood D, Olson P 2001. Small subuni...
WormB	18S	Sean A. Locke	Littlewood D, Olson P 2001. Small subuni...
1500R	28S	Sean A. Locke	Olson P, Cribb T, Tkach V, Bray R, Littl...
28SD2D3Fwd	28S	Alejandro Zaldivar-Riveron	Belshaw and Quicke 1997
28SD2D3Rev	28S	Alejandro Zaldivar-Riveron	Mardulyn and Whitfield 1999
28S D9-10 F1	28S	Christy M. Carr	Brown et al. 1999
28S D9-10 R1	28S	Christy M. Carr	Brown et al. 1999
D2B	28S	Alex Smith	Saux C, Fisher BL, Spicer GS (2004) Drac...
D3Ar	28S	Alex Smith	Saux C, Fisher BL, Spicer GS (2004) Drac...
F63	28S	Ian Hogg	Bissett, A., J. A. E. Gibson, S. N. Jarm...
LR0R	28S	Isabelle P. Meusnier	Vilgalys, R. and M. Hester (1990). "Rapi...
LR5	28S	Isabelle P. Meusnier	Vilgalys, R. and M. Hester (1990). \"Rap...
LSU5	28S	Sean A. Locke	Olson P, Cribb T, Tkach V, Bray R, Littl...
R635	28S	Ian Hogg	Bissett, A., J. A. E. Gibson, S. N. Jarm...
1385R	rbcL	Gary Saunders	Manhart. Phylogenetic analysis of green ...
CfD	rbcL	Gary Saunders	Hamsher & Saunders unpubl.
DPrbcL7	rbcL	Gary Saunders	Amato, A. 2006. Plastid phylogeny and ch...
DPrbcL7Jones	rbcL	Sarah E. Hamsher	Jones HM, Simpson GE, Stickle AJ, & Mann...
F57	rbcL	Hana Kucera	Hommersand, M.H., Fredericq, S., and Fre...
GrbcLFi	rbcL	Gary Saunders	Kucera and Saunders unpubl.
NDrbcL5	rbcL	Sarah E. Hamsher	Evans KE, Wortley AH, Simpson GE, Chepur...
NolrbcL-f	rbcL	Sarah E. Deel	Nolan, K. A., T. Gurney, Jr., L. Roberts...
NolrbcL-r	rbcL	Sarah E. Deel	Nolan, K. A., T. Gurney, Jr., L. Roberts...
rbcLajf634R	rbcL	David S. Gernandt	Fazekas, A.J. et al. 2008. Multiple Mult...
rbcLrev	rbcL	Hana Kucera	Vis, M. L., Harper, J. T. & Saunders, G....
rbcLrevNew	rbcL	Hana Kucera	unpublished
RR4	rbcL	Hana Kucera	Kucera and Saunders unpublished
RrIf	rbcL	Hana Kucera	Kucera and Saunders, unpublished
RrIfPey	rbcL	Hana Kucera	Kucera and Saunders, Unpublished
RrIr	rbcL	Hana Kucera	Kucera and Saunders, Unpublished
F428fabales	matK	Natasha De Vere	J Allainguillaume
matK_1R_kim	matK	Cassio Van Den Berg	Ki-Joong Kim, unpublished
matK2.1F	matK	Brianna Chouinard	Phase 2 protocols for DNA Barcoding as s...
matK2F-r	matK	Masha L. Kuzmina	Wang et al. 1999, AJB
matK-Equi-F	matK	Masha L. Kuzmina	http://www.kew.org/barcoding/update.html
matK_Equisetum	matK	Diana M. Percy	Phase 2 protocols for DNA Barcoding as s...
25F1R	ITS	Monica B. J. Moniz	Kogame, K., Horiguchi, T. & Masuda, M. ...
5.8S	ITS	Bryn T. M. Dentinger	Vilgalys lab
5.8S-Pr	ITS	Hermann Voglmayr	Voglmayr H, Constantinescu O. Revision a...
5.8SR	ITS	Bryn T. M. Dentinger	Vilgalys lab
CAS5p8sB1d	ITS	Alex Smith	Modified by M. Alex Smith from Ji, Y.-J....
Clino_ITS_F	ITS	Sean A. Locke	Gustinelli A, Caffara M, Florio D, Otach...
Cole A	ITS	Kavish P. Chandra	Coleman et al. 1994
Cole B	ITS	Kavish P. Chandra	Coleman et al. 1994
ITS1-F	ITS	Bryn T. M. Dentinger	Gardes and Bruns, 1993
ITS2.2	ITS	Sean A. Locke	Cribb T, Anderson G, Adlard R, Bray R. 1...
ITS2F	ITS	Xin Zhou	YANWEI LI, XIN ZHOU, GUI FENG, HAOYUAN H...
ITS3	ITS	Hai D. T. Nguyen	White, T.J., Bruns, T., Lee, S., and Tay...
ITS4	ITS	Hai D. T. Nguyen	White, T.J., Bruns, T., Lee, S., and Tay...
ITS4-B	ITS	Bryn T. M. Dentinger	Gardes and Bruns, 1993
ITS4BR	ITS	Agathe M. Vialle	Vialle et al., 2009 - Evaluation of mito...
ITS5	ITS	Hai D. T. Nguyen	White, T.J., Bruns, T., Lee, S., and Tay...
ITS5-P2	ITS	Hermann Voglmayr	Voglmayr H, Constantinescu O. Revision a...
ITSBD1	ITS	Taika von Konigslow	Luton, K., Walker, D., and Blair, D. 199...
ITSBD2	ITS	Taika von Konigslow	Luton, K., Walker, D., and Blair, D. 199...
ITS_BR	ITS	Sean A. Locke	Overstreet RM, Curran SS, Pote LM, King ...
ITSD1	ITS	Taika von Konigslow	Galazzo, D.E., Dayanandan, S., Marcoglie...
ITSD2	ITS	Taika von Konigslow	Galazzo, D.E., Dayanandan, S., Marcoglie...
ITS_DigH	ITS	Sean A. Locke	Tkach, V.V., Pawlowski, J. & Sharpilo, V...
ITSG4	ITS	Hana Kucera	Tai, V., Lindstrom, S.C., and Saunders, ...
ITSP1	ITS	Hana Kucera	Tai, V., Lindstrom, S.C., and Saunders, ...
ITSR1	ITS	Bridgette E. Clarkston	Tai, V., Lindstrom, S.C. & Saunders, G.W...
ITSR1s	ITS	Gary Saunders	Milstein, D. & Saunders, G.W. unpubl.
K1R1	ITS	Hana Kucera	Lane, C.E., Mayes, C., Druehl, L.D., and...
KG4	ITS	Daniel C. McDevit	Lane, C.E., Mayes, C., Druehl, L.D., and...
KP5	ITS	Daniel C. McDevit	Lane, C.E., Mayes, C., Druehl, L.D., and...
LR1850	ITS	Monica B. J. Moniz	Bhattacharya et al., 1996
NS7m	ITS	Monica B. J. Moniz	Bhattacharya 1996
SR12cF	ITS	Imke D. Lang	Yoshihito Takano and Takeo Horigushi. 20...
TW13	ITS	Bryn T. M. Dentinger	White, T.J., Bruns, T., Lee, S., and Tay...
UNLO28S22	ITS	Megan A. Milton	White TJ, Bruns T, Lee S, Taylor J, 1990...
UNUP18S42	ITS	Megan A. Milton	White TJ, Bruns T, Lee S, Taylor J, 1990...
DPrbcL1	rbcLa	Sarah E. Hamsher	Jones HM, Simpson GE, Stickle AJ, & Mann...
GrbcLF	rbcLa	Hana Kucera	Kucera and Saunders, Unpublished
GrbcLnF	rbcLa	Hana Kucera	Kucera and Saunders, unpublished
GrbcLR	rbcLa	Hana Kucera	Kucera and Saunders, unpublished
NDrbcL11	rbcLa	Sarah E. Hamsher	Evans KE, Wortley AH, Simpson GE, Chepur...
rbcLa-F	rbcLa	Masha L. Kuzmina	Soltis P. 1992 Proc Nat Acad Sci USA, vo...

CB3A	CYTB	Nicolas M. Hubert	Kocher, T. D., Thomas, W. K., Meyer, A.,...
CytbH15915	CYTB	Nicolas M. Hubert	Irwing, D. M., Kocher, T. D., Wilson, A....
CytbL14724	CYTB	Nicolas M. Hubert	Palumbi, S., Martin, A., Romano, S., McM...
CytTHAL	CYTB	Nicolas M. Hubert	Bernardi, G. , Bucciarelli, G., Costagli...
Glufish	CYTB	Nicolas M. Hubert	Sevilla, R., Diez, A., Noren, M., Mouche...
MCB398	CYTB	Nataly Ivanova	Verma SK, Singh L (2003) Novel universal...
MCB869	CYTB	Nataly Ivanova	Verma SK, Singh L (2003) Novel universal...
TruccytB	CYTB	Nicolas M. Hubert	Sevilla, R., Diez, A., Noren, M., Mouche...
D2dnB	28S-D2	Xin Zhou	Zhou, X., K. M. Kjer, and J. C. Morse. 2...
D2F	28S-D2	Jeanette Nordqvist	Campbell et al. 1993
D2R	28S-D2	Jeanette Nordqvist	Campbell et al. 1993
D2up4	28S-D2	Xin Zhou	Zhou, X., K. M. Kjer, and J. C. Morse. 2...
T16N	28S-D2	Gary Saunders	Saunders unpublished
T24	28S-D2	Hana Kucera	Harper, JT and Saunders, GW 2001. The ap...
T24U	28S-D2	Gary Saunders	Hamsher & Saunders unpublished
Bear	EF1-alpha	Vazrick Nazari	new
BJ	EF1-alpha	Vazrick Nazari	new
Bo	EF1-alpha	Vazrick Nazari	new
Buck	EF1-alpha	Vazrick Nazari	new
Cho (E234F)	EF1-alpha	John J. Wilson	Reed RD, Sperling FA (1999) Interaction ...
EF1aLepF1	EF1-alpha	Jeremy deWaard	Yamamoto, S. and T. Sota. 2007. Phylogen...
EF1aLepF2	EF1-alpha	Jeremy deWaard	Yamamoto, S. and T. Sota. 2007. Phylogen...
EF1aLepF3	EF1-alpha	Jeremy deWaard	Yamamoto, S. and T. Sota. 2007. Phylogen...
EF1aR	EF1-alpha	Jeremy deWaard	Kawakita et al., 2004 A. Kawakita, A. Ta...
Hutch	EF1-alpha	Vazrick Nazari	new
Juke (E600rc)	EF1-alpha	John J. Wilson	Reed RD, Sperling FA (1999) Interaction ...
Laverne	EF1-alpha	Vazrick Nazari	new
Luke	EF1-alpha	Vazrick Nazari	new
Petra	EF1-alpha	Vazrick Nazari	new
Tweeky	EF1-alpha	Vazrick Nazari	new
Verdi3	EF1-alpha	Vazrick Nazari	new
12SJ-L	12S	Virginia Leon	Goebel, A.M., Donnelly, J.M., Atz, M.E.,...
12SK-H	12S	Virginia Leon	Goebel, A.M., Donnelly, J.M., Atz, M.E.,...
Rh1039r	Rho	Agnes Dettai	Chen, W.J., Bonillo, C., Lecointre, G., ...
Rh1073r	Rho	Agnes Dettai	Chen, W.J., Bonillo, C., Lecointre, G., ...
Rh193	Rho	Agnes Dettai	Chen, W.J., Bonillo, C., Lecointre, G., ...
Rod-F2w	Rho	Andrea Bernard	Sevilla et al. 2007. Primers and polymer...
Rod-R4n	Rho	Andrea Bernard	Sevilla et al. 2007. Primers and polymer...
Diam23Sr1	UPA	Sarah E. Hamsher	Hamsher & Saunders unpublished
p23SnewR	UPA	Hana Kucera	Clarkston and Saunders, in press
p23SrV_f1	UPA	Sarah E. Hamsher	Sherwood, AR and Presting, GG. 2007. Uni...
p23SrV_r1	UPA	Hana Kucera	Sherwood, AR and Presting, GG. 2007. Uni...
coa3772	COI-3P	Jeremy deWaard	Monteiro and Pierce, 2001 A. Monteiro an...
COI-IIR	COI-3P	Jeremy deWaard	Yamamoto, S. and T. Sota. 2007. Phylogen...
cos2183	COI-3P	Jeremy deWaard	Simon et al., 1994 C. Simon, F. Frati, A...
LCOLong	COI-3P	Karl Magnacca	K.N. Magnacca & M.J.F. Brown
Lepwg-1	Wnt1	Vazrick Nazari	new
modLepwg-2	Wnt1	Vazrick Nazari	new
CAD1028R	CAD	John J. Wilson	Wahlberg N, Wheat CW (2008) Genomic outp...
CAD743nF	CAD	John J. Wilson	Wahlberg N, Wheat CW (2008) Genomic outp...
TufAR	tufA	Daniel C. McDevit	Gary W Saunders unpublished
TufGF4	tufA	Daniel C. McDevit	Gary W Saunders unpublished

Appendix D

GenBank accession numbers

AY666445	AY666343
AF168063	AY666177
AY666220	AY666260
AY666438	AY666265
AF168057	U76059
AY666293	AY666590
AY666271	AY666249
AY666197	AY666175
AY666180	AY666435
AY666298	AY666317
AF197831	AY666202
AF197830	AY666356
AF197826	AY666233
AF197827	AY666379
AF197842	AY666255
AY666174	AY666522
AY666366	AY666306
AY666444	AY666284
AY666525	AY666386
AY666391	AF168070
AY666282	AF168069
AY666236	AY666510
AY666321	AY666489
AY666335	AY666498
AF197858	AY666285
AF197860	AY666402
AF197854	AY666502
AY666494	AY666276
DQ385165	AY666390
DQ385174	AY666176
AY940785	AY666337
AY666418	AY666434
AY940779	AY666219
AY942876	AY666212
AY944307	AY666339
AY927186	AY666223
AY927188	AY666369
AF394279	AY666407
AY942877	AY666277
AY940783	AY666374
AF272596	DQ385176
AY942878	DQ385168
AY666289	DQ385170
AF168061	AY666287
AF168068	AY666471
AY666198	AF168058
AY666596	
AY666263	
AY666309	
AY666308	
AY666318	
AY666307	
AY666291	
AY666412	
AY666305	
AY666371	
AY666359	
AY666222	
AY666272	
AY666286	

Appendix E

Barcoding New Zealand Birds contributing sources

A complete list of the Barcoding New Zealand Birds project. All specimen page numbers prefixed with **JW** were sequenced by John Waugh and all those prefixed with **SP** were sequenced by Selina Patel. All others were sequenced by Allan Baker's group with the exception of five samples prefixed with **LH,** which were contributed by Leon Huynen. Note *Diomedea melanophrys* is now known as *Thalassarche melanophrys*.

Identification	Specimen Page	Sequence Page	Length [Ambig] COI-5P	Record Flags	Extra Info	Set
Acanthisitta chloris	RIFL 004	BROMB832-07	776 [0n]			
Acanthisitta chloris	RIFL 003	BROMB831-07	776 [0n]			
Acanthisitta chloris	RIFL 002	BROMB830-07	779 [0n]			
Acanthisitta chloris	RIFL 001	BROMB829-07	779 [0n]			
Acanthisitta chloris	SP113	NZCOI344-09	648 [0n]			
Acanthisitta chloris	SP114	NZCOI345-09	648 [0n]			
Acanthisitta chloris	SP115	NZCOI346-09	648 [0n]			
Acanthisitta chloris	SP116	NZCOI347-09	648 [0n]			
Acanthisitta chloris	SP117	NZCOI348-09	648 [0n]			
Acanthiza lineata	1B-961	BROMB584-07	839 [0n]			
Acanthiza lineata	1B-960	BROMB583-07	739 [1n]			
Acridotheres tristis	SP050	NZCOI284-09	648 [0n]			
Alauda arvensis	AJB 5562	BROMB498-07	790 [0n]		ssp:arvensis	
Alauda arvensis	AJB 5561	BROMB496-07	719 [0n]		ssp:arvensis	
Alauda arvensis	MKP 1687	BROMB805-07	773 [0n]		ssp: arvensis	
Alauda arvensis	MKP 1686	BROMB804-07	808 [0n]	4	ssp: arvensis	
Alauda arvensis	JW217	NZCOI069-08	648 [0n]	1		
Alauda arvensis	SP109	NZCOI340-09	648 [0n]			
Anarhynchus frontalis	SP153	NZCOI368-09	648 [0n]			
Anarhynchus frontalis	SP152	NZCOI539-09	648 [0n]			
Anas castanea	SP033	NZCOI520-09	648 [0n]			
Anas gracilis	SP032	NZCOI519-09	648 [0n]			
Anas nesiotis	JW707	NZCOI075-08	648 [0n]		Cambell Island Teal	
Anas platyrhynchos	SP016	NZCOI262-09	648 [0n]			
Anas platyrhynchos	SP017	NZCOI513-09	648 [0n]			
Anas platyrhynchos	SP331	NZCOI624-09	0			
Anas rhynchotis	SP027	NZCOI514-09	648 [0n]			
Anas superciliosa	SP031	NZCOI518-09	648 [0n]			
Anthornis melanura	JW693	NZCOI077-08	648 [0n]			
Anthornis melanura	JW999	NZCOI200-08	694 [0n]			
Anthornis melanura	JW1000	NZCOI201-08	694 [0n]			
Anthornis melanura	JW1001	NZCOI202-08	694 [0n]			
Anthornis melanura	JW1002	NZCOI203-08	694 [0n]			

Anthornis melanura	JW1003	NZCOI204-08	694 [0n]		2	
Anthornis melanura	JW1004	NZCOI205-08	694 [0n]		2	
Anthornis melanura	JW1005	NZCOI206-08	694 [0n]		2	
Anthornis melanura	SP082	NZCOI314-09	648 [0n]		2	
Anthornis melanura	SP083	NZCOI315-09	648 [0n]		2	
Anthornis melanura	SP084	NZCOI316-09	648 [0n]		2	
Anthornis melanura	SP085	NZCOI317-09	648 [0n]		2	
Anthus novaeseelandiae	AJB 5570	BROMB502-07	818 [0n]		3	ssp:novaeseelandiae
Anthus novaeseelandiae	MKP 2234	BROMB855-07	776 [0n]		2	
Anthus novaeseelandiae	MKP 1697	BROMB811-07	753 [0n]		2	ssp: novaeseelandiae
Anthus novaeseelandiae	MKP 1696	BROMB810-07	753 [0n]		2	ssp: novaeseelandiae
Aphrodroma brevirostris	B-13545	BROMB701-07	749 [0n]		2	
Aphrodroma brevirostris	1B-1033	BROMB619-07	734 [0n]		2	
Apteryx australis	R34157	BROMB472-06	715 [0n]		3	
Apteryx australis	R35021	BROMB471-06	802 [0n]		2	
Apteryx australis	TFCK	BROMB470-06	746 [0n]		4	
Apteryx australis	JAM 1	BROMB469-06	701 [0n]		3	
Apteryx australis	CD1141	BROMB468-06	725 [0n]		2	
Apteryx australis	CD1150	BROMB467-06	725 [0n]		2	
Apteryx australis	DC4	BROMB466-06	744 [0n]		2	
Apteryx australis	RA 0887	BROMB464-06	732 [0n]		3	
Apteryx australis	RA 0886	BROMB463-06	755 [0n]		2	
Apteryx australis	RA 0882	BROMB462-06	760 [0n]		2	
Apteryx australis	LH006	NZCOI072-08	584 [0n]			
Apteryx australis	SP317	NZCOI491-09	648 [0n]		2	Chick
Apteryx australis	SP318	NZCOI492-09	648 [0n]		2	Chick
Apteryx haastii	GS06	NZCOI793-10	712 [0n]		2	
Apteryx haastii	M1B	NZCOI756-10	722 [0n]		2	
Apteryx haastii	LH004	NZCOI073-08	584 [0n]			
Apteryx haastii	SP320	NZCOI494-09	648 [0n]		2	
Apteryx mantelli	LH002	NZCOI070-08	584 [1n]			
Apteryx owenii	O-20583	NZCOI797-10	708 [0n]		2	
Apteryx owenii	O-20579	NZCOI792-10	712 [0n]		2	
Apteryx owenii	Trig001	NZCOI790-10	799 [0n]		2	
Apteryx owenii	O-20556	NZCOI755-10	722 [0n]		2	
Apteryx owenii	O-20580	NZCOI754-10	722 [0n]		2	
Apteryx owenii	LH005	NZCOI074-08	584 [1n]			
Apteryx owenii	SP319	NZCOI493-09	648 [0n]		2	
Apteryx rowi	LH003	NZCOI071-08	584 [0n]			
Ardea alba	SP150	NZCOI537-09	648 [0n]		2	
Ardea intermedia	SP151	NZCOI538-09	648 [0n]		2	
Ardea novaehollandiae	SP143	NZCOI366-09	648 [0n]		2	

Ardea pacifica	SP149	NZCOI536-09	648 [0n]		2		
Arenaria interpres	SP489	NZCOI769-10	648 [0n]				
Aythya australis	SP029	NZCOI516-09	648 [0n]		2		
Aythya novaeseelandiae	SP508	NZCOI788-10	648 [0n]				
Aythya novaeseelandiae	SP018	NZCOI263-09	648 [0n]		2		
Botaurus poiciloptilus	SP140	NZCOI364-09	648 [0n]		2		
Botaurus poiciloptilus	SP141	NZCOI365-09	648 [0n]		2		
Botaurus poiciloptilus	SP142	NZCOI531-09	648 [0n]		2		
Bowdleria punctata	SP093	NZCOI325-09	648 [0n]		2		
Branta canadensis	SP019	NZCOI264-09	648 [0n]		2		
Branta canadensis	SP020	NZCOI265-09	648 [0n]		2		
Branta canadensis	SP021	NZCOI266-09	648 [0n]		2		
Branta canadensis	SP022	NZCOI267-09	648 [0n]		2		
Branta canadensis	SP023	NZCOI268-09	648 [0n]		2		
Branta canadensis	SP024	NZCOI269-09	648 [0n]		2		
Branta canadensis	SP025	NZCOI270-09	648 [0n]		2		
Calidris acuminata	SP502	NZCOI782-10	648 [0n]				
Calidris canutus	SP168	NZCOI381-09	648 [0n]		2		
Calidris canutus	SP169	NZCOI382-09	648 [0n]		2		
Calidris canutus	SP170	NZCOI383-09	648 [0n]		2		
Calidris canutus	SP171	NZCOI384-09	648 [0n]		2		
Calidris canutus	SP172	NZCOI385-09	648 [0n]		2		
Calidris canutus	SP173	NZCOI542-09	648 [0n]		2		
Calidris ferruginea	SP182	NZCOI547-09	648 [0n]		2		
Calidris melanotos	SP488	NZCOI768-10	648 [0n]				
Calidris minuta	SP490	NZCOI770-10	648 [0n]				
Calidris ruficollis	SP494	NZCOI774-10	648 [0n]				
Callaeas cinereus	JW1031	NZCOI232-08	660 [2n]		2		
Carduelis carduelis	AJB 5566	BROMB500-07	817 [0n]		2		ssp:britannica
Carduelis carduelis	AJB 5564	BROMB499-07	727 [0n]		1		ssp:britannica
Carduelis carduelis	AJB 5555	BROMB493-07	730 [0n]		1		ssp:britannica
Carduelis carduelis	AJB 5574	BROMB483-07	730 [0n]		2		ssp:britannica
Carduelis carduelis	JW1061	NZCOI139-08	648 [0n]		2		
Carduelis carduelis	SP052	NZCOI286-09	648 [0n]		2		
Carduelis carduelis	SP407	NZCOI630-09	648 [0n]		2		
Carduelis chloris	1B-1101	BROMB748-07	730 [0n]		2		
Carduelis chloris	SP056	NZCOI290-09	648 [1n]		2		
Carduelis chloris	SP057	NZCOI291-09	648 [0n]		2		
Carduelis chloris	SP058	NZCOI292-09	648 [0n]		2		
Carduelis chloris	SP059	NZCOI293-09	648 [0n]		2		
Carduelis chloris	SP341	NZCOI635-09	648 [0n]		2		
Carduelis chloris	SP342	NZCOI636-09	648 [0n]		2		
Carduelis chloris	SP343	NZCOI637-	648		2		

	Carduelis chloris	SP343	09	[0n]		2	
☑	Carduelis flammea	AJB 5578	BROMB508-07	732 [0n]		2	ssp:cabaret
☑	Carduelis flammea	AJB 5575	BROMB505-07	738 [0n]		1	ssp:cabaret
☑	Carduelis flammea	AJB 5572	BROMB504-07	830 [0n]		3	ssp:cabaret
☑	Carduelis flammea	AJB 5563	BROMB497-07	735 [0n]		1	ssp:cabaret
☑	Carduelis flammea	MKP 2026	BROMB850-07	791 [0n]		2	ssp: cabaret
☑	Catharacta skua	SP157	NZCOI371-09	648 [0n]		2	Band no.M-62037
☑	Catharacta skua	SP158	NZCOI372-09	648 [0n]		2	Band no.M-62045
☑	Catharacta skua	SP159	NZCOI373-09	648 [0n]		2	Band no.M-62046
☑	Catharacta skua	SP160	NZCOI374-09	648 [0n]		2	Band no.M-45818
☑	Catharacta skua	SP161	NZCOI375-09	648 [0n]		2	Band no.M-34054
☑	Cereopsis novaehollandiae	SP035	NZCOI522-09	648 [0n]		2	
☑	Charadrius bicinctus	SP496	NZCOI776-10	648 [0n]			
☑	Charadrius bicinctus	JW145	NZCOI141-08	726 [0n]		2	
☑	Charadrius bicinctus	JW1062	NZCOI142-08	648 [0n]		2	
☑	Charadrius bicinctus	SP155	NZCOI370-09	648 [0n]		2	
☑	Charadrius leschenaultii	SP184	NZCOI549-09	648 [0n]		2	
☑	Charadrius mongolus	SP505	NZCOI785-10	648 [0n]			
☑	Charadrius ruficapillus	SP185	NZCOI550-09	648 [0n]		2	
☑	Charadrius veredus	SP501	NZCOI781-10	648 [0n]			
☑	Chenonetta jubata	SP030	NZCOI517-09	648 [0n]		2	
☑	Chlidonias albostriatus	JW1126	NZCOI500-09	648 [0n]		2	
☑	Chlidonias hybrida	SP498	NZCOI778-10	0			
☑	Chrysococcyx lucidus	1B-945	BROMB600-07	798 [0n]		2	ssp:lucidus
☑	Chrysococcyx lucidus	1B-966	BROMB587-07	806 [0n]		2	ssp:lucidus
☑	Chrysococcyx lucidus	SP136	NZCOI362-09	648 [0n]		2	
☑	Chrysococcyx lucidus	SP139	NZCOI363-09	648 [0n]		2	
☑	Chrysococcyx lucidus	SP137	NZCOI530-09	0			
☑	Circus approximans	AJB 5573	ROMC231-07	746 [0n]		2	gouldi
☑	Circus approximans	SP314	NZCOI488-09	648 [0n]		2	
☑	Circus approximans	SP406	NZCOI673-09	648 [0n]		2	
☑	Columba livia	SP351	NZCOI645-09	648 [0n]		2	
☑	Columba livia	SP352	NZCOI646-09	648 [0n]		2	
☑	Cyanoramphus novaezelandiae	JW206	NZCOI049-08	648 [0n]		2	
☑	Cyanoramphus novaezelandiae	SP213	NZCOI409-09	648 [0n]		2	
☑	Cyanoramphus novaezelandiae	SP214	NZCOI410-09	648 [0n]		2	
☑	Cyanoramphus novaezelandiae	SP421	NZCOI694-09	648 [0n]		2	
☑	Cyanoramphus novaezelandiae	SP422	NZCOI695-09	648 [0n]		2	
☑	Cyanoramphus novaezelandiae	SP423	NZCOI696-09	648 [0n]		2	
☑	Cyanoramphus novaezelandiae	SP424	NZCOI697-09	648 [0n]		2	
☑	Cyanoramphus unicolor	JW371	NZCOI055-08	627 [0n]		1	
☑	Cyanoramphus unicolor	SP425	NZCOI698-09	648 [0n]		2	
☑	Cyanoramphus unicolor	SP426	NZCOI699-09	648 [0n]		2	

Cygnus atratus	SP036	NZCOI523-09	648 [0n]		2	
Cygnus olor	SP034	NZCOI521-09	648 [0n]		2	
Daption capense	B-13539	BROMB704-07	739 [0n]		2	
Daption capense	B-13540	BROMB702-07	737 [1n]		3	
Daption capense	JW53	NZCOI144-08	689 [2n]		2	
Daption capense	SP226	NZCOI420-09	648 [0n]		2	
Daption capense	SP227	NZCOI560-09	648 [0n]		2	
Diomedea cauta	SP246	NZCOI569-09	648 [0n]		2	
Diomedea chrysostoma	SP247	NZCOI429-09	648 [0n]		2	Band no.M-49488
Diomedea chrysostoma	SP248	NZCOI430-09	648 [0n]		2	Band no.M-41704
Diomedea chrysostoma	SP249	NZCOI431-09	648 [0n]		2	Band no.M-45152
Diomedea chrysostoma	SP250	NZCOI432-09	648 [0n]		2	Band no.O-15160
Diomedea chrysostoma	SP251	NZCOI433-09	648 [0n]		2	Band no.M-46395
Diomedea epomophora	SP477	NZCOI757-10	648 [0n]			
Diomedea epomophora	JW705	NZCOI146-08	648 [0n]		2	
Diomedea epomophora	JW1032	NZCOI233-08	694 [0n]		2	chick
Diomedea epomophora	JW1033	NZCOI234-08	694 [0n]		2	
Diomedea epomophora	JW1034	NZCOI235-08	694 [0n]		2	chick
Diomedea epomophora	JW1035	NZCOI236-08	655 [0n]		4	chick
Diomedea epomophora	JW1036	NZCOI237-08	694 [0n]		2	
Diomedea epomophora	JW1037	NZCOI238-08	694 [0n]		2	chick
Diomedea epomophora	JW1038	NZCOI239-08	694 [0n]		2	
Diomedea epomophora	SP264	NZCOI444-09	648 [0n]		2	Band no.R-48706
Diomedea epomophora	SP265	NZCOI445-09	648 [0n]		2	Band no.R-48697
Diomedea epomophora	SP266	NZCOI446-09	648 [0n]		2	Band no.R-48716
Diomedea epomophora	SP267	NZCOI447-09	648 [0n]		2	Band no.R-48715
Diomedea epomophora	SP268	NZCOI448-09	648 [0n]		2	Band no.R-48690
Diomedea epomophora	JW700	NZCOI584-09	0		2	
Diomedea exulans	SP491	NZCOI771-10	648 [0n]			
Diomedea exulans	1B-111	BROMB337-06	767 [0n]		2	ssp: exulans
Diomedea melanophris	SP253	NZCOI434-09	648 [0n]		2	Band no.M-62087
Diomedea melanophris	SP254	NZCOI435-09	648 [0n]		2	Band no.M-48637
Diomedea melanophris	SP255	NZCOI436-09	648 [0n]		2	Band no.O-27730
Diomedea melanophris	SP256	NZCOI437-09	648 [0n]		2	Band no.O-16083
Diomedea melanophris	SP257	NZCOI438-09	648 [0n]		2	Band no.M-62092
Diomedea melanophris	SP258	NZCOI439-09	648 [0n]		2	Band no.M-45767
Diomedea melanophris	SP259	NZCOI440-09	648 [0n]		2	Band no.O-24042
Diomedea melanophris	SP260	NZCOI441-09	648 [0n]		2	Band no.M-64122
Diomedea melanophris	SP261	NZCOI442-09	648 [0n]		2	Band no.M-42046
Diomedea melanophris	SP262	NZCOI443-09	648 [0n]		2	Band no.M-62089
Egretta novaehollandiae	SP144	NZCOI532-09	648 [0n]		2	
Elseyornis melanops	SP156	NZCOI540-09	648 [0n]		2	
Emberiza		NZCOI282-	648			

	citrinella	SP048		09	[0n]		
☑	Emberiza citrinella	SP334		NZCOI627-09	648 [0n]		2
☑	Erythrogonys cinctus	SP500		NZCOI780-10	648 [0n]		
☑	Eudynamys taitensis	SP135		NZCOI361-09	648 [0n]		2
☑	Eudynamys taitensis	SP356		NZCOI650-09	648 [0n]		2
☑	Eudyptes chrysocome	JD 13A		NZCOI796-10	709 [0n]		2
☑	Eudyptes chrysocome	JD 09A		NZCOI795-10	749 [0n]		2
☑	Eudyptes chrysocome	JD 10A		NZCOI794-10	737 [2n]		2
☑	Eudyptes chrysocome	JD 03A		NZCOI750-10	722 [0n]		2
☑	Eudyptes chrysocome	JW1009		NZCOI210-08	648 [0n]		4
☑	Eudyptes chrysocome	JW1010		NZCOI211-08	648 [0n]		4
☑	Eudyptes chrysocome	JW1011		NZCOI212-08	648 [0n]		6
☑	Eudyptes chrysocome	JW1012		NZCOI213-08	658 [0n]		2
☑	Eudyptes chrysocome	SP130		NZCOI529-09	648 [0n]		2
☑	Eudyptes pachyrhynchus	JW21		NZCOI148-08	648 [0n]		2
☑	Eudyptes pachyrhynchus	JW990		NZCOI191-08	686 [0n]		2
☑	Eudyptes pachyrhynchus	JW991		NZCOI192-08	662 [0n]		2
☑	Eudyptes pachyrhynchus	JW993		NZCOI194-08	681 [1n]		2
☑	Eudyptes pachyrhynchus	JW995		NZCOI196-08	686 [0n]		2
☑	Eudyptes pachyrhynchus	JW996		NZCOI197-08	694 [0n]		2
☑	Eudyptes pachyrhynchus	JW997		NZCOI198-08	671 [0n]		2
☑	Eudyptes pachyrhynchus	JW998		NZCOI199-08	694 [0n]		2
☑	Eudyptes pachyrhynchus	SP124		NZCOI352-09	648 [0n]		2
☑	Eudyptes pachyrhynchus	SP125		NZCOI353-09	648 [0n]		2
☑	Eudyptes pachyrhynchus	SP126		NZCOI528-09	648 [0n]		
☑	Eudyptes robustus	SP129		NZCOI356-09	648 [0n]		2
☑	Eudyptes sclateri	EC02		BROMB076-06	689 [0n]		2
☑	Eudyptes sclateri	EC01		BROMB075-06	658 [0n]		2
☑	Eudyptes sclateri	SP127		NZCOI354-09	648 [0n]		2
☑	Eudyptula minor	SP121		NZCOI350-09	648 [0n]		2
☑	Eudyptula minor	SP122		NZCOI351-09	648 [0n]		2
☑	Eudyptula minor	SP123		NZCOI527-09	648 [0n]		2
☑	Falco novaeseelandiae	JW852		NZCOI080-08	648 [0n]		2
☑	Falco novaeseelandiae	JW853		NZCOI081-08	648 [0n]		2
☑	Falco novaeseelandiae	JW854		NZCOI082-08	648 [0n]		2
☑	Falco novaeseelandiae	JW855		NZCOI083-08	648 [0n]		2
☑	Falco novaeseelandiae	JW856		NZCOI084-08	648 [0n]		2
☑	Falco novaeseelandiae	JW857		NZCOI085-08	648 [0n]		2
☑	Falco novaeseelandiae	JW858		NZCOI086-08	648 [0n]		2
☑	Falco novaeseelandiae	JW849		NZCOI078-08	648 [0n]		2
☑	Falco novaeseelandiae	JW851		NZCOI079-08	648 [0n]		2
☑	Falco novaeseelandiae	SP315		NZCOI489-09	648 [0n]		2
☑	Fregata minor	SP205		NZCOI556-09	648 [0n]		2	subadult

Fregetta tropica	SP438	NZCOI711-09	648 [0n]			2	
Fringilla coelebs	AJB 5576	BROMB506-07	731 [0n]			2	ssp:gengleri
Fringilla coelebs	SP051	NZCOI285-09	648 [1n]			2	
Fringilla coelebs	SP335	NZCOI628-09	648 [0n]			2	
Fringilla coelebs	SP336	NZCOI629-09	648 [0n]			2	
Fulica atra	SP001	NZCOI255-09	648 [0n]			2	
Gallinula tenebrosa	SP010	NZCOI508-09	648 [0n]			2	
Gallinula ventralis	SP011	NZCOI509-09	648 [0n]			2	
Gallirallus australis	SP509	NZCOI789-10	648 [0n]				
Gallirallus australis	1B-987	BROMB595-07	809 [0n]			3	
Gallirallus australis	1B-157	BROMB348-06	743 [0n]			1	ssp: australis
Gallirallus australis	1B-156	BROMB347-06	740 [0n]			2	ssp: australis
Gallirallus australis	1B-155	BROMB346-06	818 [0n]			2	ssp: australis
Gallirallus australis	SP008	NZCOI259-09	648 [0n]			2	
Gallirallus philippensis	1B-68	BROMB322-06	742 [0n]			2	ssp: australis
Gallus gallus	JW690	NZCOI150-08	648 [0n]			2	
Gerygone igata	AJB 5553	BROMB491-07	743 [0n]			3	
Gerygone igata	SP095	NZCOI327-09	648 [0n]			2	
Gerygone igata	SP096	NZCOI328-09	648 [0n]			2	Band no.AA5603
Gerygone igata	SP097	NZCOI329-09	648 [0n]			2	Band no.AA5601
Gerygone igata	SP098	NZCOI330-09	648 [0n]			2	Band no.AA5611
Gerygone igata	SP099	NZCOI331-09	648 [0n]			2	Nestling, Band no.AA5627
Gerygone igata	SP100	NZCOI332-09	648 [0n]			2	Nestling, Band no.AA5630
Gerygone igata	SP101	NZCOI333-09	648 [0n]			2	Band no.AA5660
Gerygone igata	SP102	NZCOI334-09	648 [0n]			2	Band no.AA5661
Gerygone igata	SP103	NZCOI335-09	648 [0n]			2	Band no.AA5663
Gerygone igata	SP105	NZCOI336-09	648 [0n]			2	Band no.AA5672
Glareola maldivarum	SP181	NZCOI546-09	648 [0n]			2	
Grus rubicunda	SP013	NZCOI511-09	648 [0n]			2	
Gymnorhina tibicen	JW1059	NZCOI152-08	648 [0n]			2	
Gymnorhina tibicen	SP049	NZCOI283-09	648 [0n]			2	
Gymnorhina tibicen	JW1127	NZCOI501-09	648 [0n]			2	
Gymnorhina tibicen	JW222	NZCOI582-09	0			2	
Haematopus chathamensis	CIOC 7473	ROMC009-06	809 [0n]			2	
Haematopus finschi	AJB 4784	ROMC033-06	878 [0n]			2	syn:Haematopus ostralegus finschi
Haematopus finschi	BR 28767	ROMC032-06	783 [0n]			2	syn:Haematopus ostralegus finschi
Haematopus unicolor	BR 28766	ROMC065-06	830 [0n]			2	voucher:skin; skeleton
Haematopus unicolor	BR 28765	ROMC064-06	861 [0n]			2	voucher:skin; skeleton
Haematopus unicolor	JW1125	NZCOI499-09	648 [0n]			2	
Halcyon sancta	SP187	NZCOI390-09	648 [0n]			2	
Halcyon sancta	SP189	NZCOI391-09	648 [0n]			2	
Halcyon sancta	SP190	NZCOI392-09	648 [0n]			2	
		NZCOI393-	648				

Halcyon sancta	SP191	NZCOI393-09	648 [0n]		2	
Halobaena caerulea	SP482	NZCOI762-10	0			
Halobaena caerulea	SP224	NZCOI419-09	648 [0n]		2	
Hemiphaga novaeseelandiae	SP119	NZCOI349-09	648 [1n]		2	
Hemiphaga novaeseelandiae	SP353	NZCOI647-09	648 [0n]		2	
Hemiphaga novaeseelandiae	SP354	NZCOI648-09	648 [0n]		2	
Hemiphaga novaeseelandiae	SP410	NZCOI683-09	648 [4n]		2	
Hemiphaga novaeseelandiae	SP411	NZCOI684-09	648 [0n]		2	
Hemiphaga novaeseelandiae	SP412	NZCOI685-09	648 [0n]		2	
Hemiphaga novaeseelandiae	SP413	NZCOI686-09	648 [1n]		2	
Hemiphaga novaeseelandiae	SP414	NZCOI687-09	648 [0n]		2	
Heteroscelus brevipes	SP180	NZCOI545-09	648 [0n]		2	
Himantopus himantopus	SP497	NZCOI777-10	648 [0n]			
Himantopus leucocephalus	AJB 4798	ROMC016-06	842 [0n]		2	
Himantopus leucocephalus	AJB 4793	ROMC015-06	821 [0n]		2	
Himantopus novaezelandiae	SP362	NZCOI593-09	648 [0n]		2	Band no. YR-BKW
Himantopus novaezelandiae	SP363	NZCOI594-09	648 [0n]		2	Band no. WR-RR
Himantopus novaezelandiae	SP364	NZCOI595-09	648 [0n]		2	Band no. WR-RY
Himantopus novaezelandiae	SP365	NZCOI596-09	648 [0n]		2	Band no. RY-WY
Himantopus novaezelandiae	SP366	NZCOI597-09	648 [0n]		2	Band no. RY-WG
Hirundo tahitica	SP047	NZCOI281-09	648 [0n]		2	
Hydroprogne caspia	SP507	NZCOI787-10	648 [0n]			
Hymenolaimus malacorhynchos	Hmalac38	BROMB981-08	734 [0n]		2	
Hymenolaimus malacorhynchos	Hmalac2F	BROMB980-08	746 [0n]		2	
Hymenolaimus malacorhynchos	Hmalac4M	BROMB979-08	776 [0n]		2	
Hymenolaimus malacorhynchos	Hmalac3	BROMB978-08	758 [1n]			
Hymenolaimus malacorhynchos	Hmalac1F	BROMB977-08	758 [0n]			
Hymenolaimus malacorhynchos	HmalacHM26	BROMB976-08	758 [1n]			
Hymenolaimus malacorhynchos	Hmalac2_M	BROMB975-08	758 [0n]			
Hymenolaimus malacorhynchos	JW101	NZCOI087-08	648 [0n]		2	
Hymenolaimus malacorhynchos	JW316	NZCOI088-08	648 [0n]		2	
Hymenolaimus malacorhynchos	JW317	NZCOI089-08	648 [0n]		1	
Hymenolaimus malacorhynchos	JW938	NZCOI090-08	648 [0n]		1	
Ixobrychus minutus	SP146	NZCOI533-09	648 [0n]		2	
Larus bulleri	JW1129	NZCOI503-09	648 [0n]		2	
Larus bulleri	SP361	NZCOI655-09	648 [0n]		2	
Larus dominicanus	JW1058	NZCOI153-08	648 [0n]		2	
Larus dominicanus	JW1060	NZCOI154-08	648 [0n]		2	
Larus dominicanus	SP178	NZCOI389-09	648 [0n]		2	
Larus dominicanus	JW1128	NZCOI502-09	648 [0n]		2	
Larus dominicanus	SP358	NZCOI652-09	648 [0n]		2	
Larus dominicanus	SP359	NZCOI653-09	648 [0n]		2	
Larus	JW770	NZCOI002-	648			

	novaehollandiae	JW779	08	[0n]			
	Larus novaehollandiae	JW780	NZCOI003-08	648 [0n]		2	
	Larus novaehollandiae	JW782	NZCOI004-08	648 [0n]		2	
	Larus novaehollandiae	JW783	NZCOI005-08	648 [0n]		2	
	Larus novaehollandiae	JW784	NZCOI006-08	648 [0n]		2	
	Larus novaehollandiae	JW785	NZCOI007-08	648 [0n]		2	
	Larus novaehollandiae	JW786	NZCOI008-08	648 [0n]		2	
	Larus novaehollandiae	JW787	NZCOI009-08	648 [0n]		2	
	Larus novaehollandiae	JW789	NZCOI010-08	648 [0n]		2	
	Larus novaehollandiae	JW790	NZCOI011-08	648 [0n]		2	
	Leucocarbo chalconotus	SP204	NZCOI402-09	648 [0n]		2	
	Limosa lapponica	SP162	NZCOI376-09	648 [0n]		2	
	Limosa lapponica	SP163	NZCOI377-09	648 [0n]		2	
	Limosa lapponica	SP164	NZCOI378-09	648 [0n]		2	
	Limosa lapponica	SP165	NZCOI379-09	648 [0n]		2	
	Limosa lapponica	SP166	NZCOI380-09	648 [0n]		2	
	Limosa lapponica	SP167	NZCOI541-09	648 [0n]		2	
	Limosa limosa	SP503	NZCOI783-10	648 [0n]			
	Macronectes giganteus	1B-929	BROMB613-07	799 [0n]		1	
	Macronectes giganteus	SP287	NZCOI465-09	648 [0n]		2	
	Macronectes giganteus	SP288	NZCOI466-09	648 [0n]		2	
	Macronectes giganteus	SP289	NZCOI574-09	648 [0n]		2	
	Malacorhynchus membranaceus	SP028	NZCOI515-09	648 [0n]		2	
	Megadyptes antipodes	JD 41A	BROMB046-06	725 [0n]		2	
	Megadyptes antipodes	JD 64A	BROMB027-06	842 [0n]		3	
	Megadyptes antipodes	JD 40A	BROMB080-06	830 [0n]		2	
	Megadyptes antipodes	1B-926	BROMB051-06	687 [0n]		2	
	Megadyptes antipodes	JD 150	BROMB050-06	731 [0n]		2	
	Megadyptes antipodes	JD 115	BROMB049-06	783 [0n]		2	
	Megadyptes antipodes	JD 88A	BROMB048-06	756 [0n]		2	
	Megadyptes antipodes	JD 62A	BROMB047-06	722 [0n]		2	
	Megadyptes antipodes	JW014	NZCOI116-08	648 [0n]		2	
	Megadyptes antipodes	JW1049	NZCOI250-08	694 [0n]		2	
	Megadyptes antipodes	JW1050	NZCOI251-08	648 [0n]		4	
	Megadyptes antipodes	JW1051	NZCOI252-08	694 [0n]		2	
	Megadyptes antipodes	JW1052	NZCOI253-08	648 [0n]		4	
	Megadyptes antipodes	JW1053	NZCOI254-08	694 [0n]		2	
	Megadyptes antipodes	SP128	NZCOI355-09	648 [0n]		2	
	Megadyptes antipodes	SP355	NZCOI649-09	648 [0n]		2	
	Megalurus punctatus	AJB 5579	BROMB509-07	766 [0n]		2	ssp:punctatus
	Mohoua albicilla	SP068	NZCOI302-09	648 [0n]		2	
	Mohoua albicilla	SP069	NZCOI303-09	648 [0n]		2	
	Mohoua albicilla	SP071	NZCOI304-09	648 [0n]		2	

Mohoua albicilla	SP072	NZCOI305-09	648 [0n]		2	
Mohoua albicilla	SP070	NZCOI524-09	648 [0n]		2	
Mohoua novaeseelandiae	NC1	NZCOI674-09	648 [0n]		2	
Mohoua novaeseelandiae	NC2	NZCOI675-09	648 [0n]		2	Band no. A177163
Mohoua novaeseelandiae	NC3	NZCOI676-09	648 [0n]		2	
Mohoua novaeseelandiae	NC4	NZCOI677-09	648 [0n]		2	Band no. A177165
Mohoua novaeseelandiae	NC5	NZCOI678-09	648 [1n]		2	
Mohoua novaeseelandiae	NC6	NZCOI679-09	648 [0n]		2	
Mohoua novaeseelandiae	NC7	NZCOI680-09	648 [0n]		2	Band no. A177160
Mohoua novaeseelandiae	NC8	NZCOI681-09	648 [0n]		2	Band no. A177168
Mohoua ochrocephala	SP094	NZCOI326-09	648 [0n]		2	
Mohoua ochrocephala	SP451	NZCOI724-09	648 [0n]		2	
Mohoua ochrocephala	SP452	NZCOI725-09	648 [0n]		2	
Mohoua ochrocephala	SP453	NZCOI726-09	648 [0n]		2	
Mohoua ochrocephala	SP454	NZCOI727-09	648 [0n]		2	
Mohoua ochrocephala	SP448	NZCOI721-09	648 [0n]		2	
Mohoua ochrocephala	SP449	NZCOI722-09	648 [0n]		2	
Mohoua ochrocephala	SP450	NZCOI723-09	648 [0n]		2	
Morus serrator	SSK27	ROMC534-08	730 [0n]		1	Sula serrator
Morus serrator	SSK33	ROMC350-07	730 [0n]		1	
Morus serrator	SSK26	ROMC349-07	739 [0n]		2	
Morus serrator	SSK24	ROMC347-07	733 [0n]		1	
Morus serrator	SSK8	ROMC341-07	744 [0n]		2	
Morus serrator	SSK7	ROMC340-07	749 [0n]		2	
Morus serrator	SSK18	ROMC339-07	733 [0n]		2	
Morus serrator	SSK17	ROMC338-07	739 [0n]		2	
Morus serrator	SP197	NZCOI396-09	648 [0n]		2	Band no. M-74773
Morus serrator	SP198	NZCOI397-09	648 [0n]		2	Band no. M-74937
Morus serrator	SP199	NZCOI398-09	648 [0n]		2	Band no. M-74944
Morus serrator	SP200	NZCOI399-09	648 [0n]		2	Band no. M-74945
Morus serrator	SP201	NZCOI400-09	648 [0n]		2	Band no. M-77205
Morus serrator	SP202	NZCOI401-09	648 [0n]		2	Band no. M-74948
Morus serrator	SP203	NZCOI555-09	648 [0n]		2	
Nestor meridionalis	JW828	NZCOI057-08	648 [0n]		2	
Nestor meridionalis	JW830	NZCOI058-08	648 [0n]		2	
Nestor meridionalis	JW832	NZCOI059-08	648 [0n]		2	
Nestor meridionalis	JW833	NZCOI060-08	648 [0n]		2	
Nestor meridionalis	JW836	NZCOI061-08	648 [0n]		2	
Nestor meridionalis	JW837	NZCOI062-08	648 [0n]		1	
Nestor meridionalis	JW1029	NZCOI230-08	694 [0n]		2	
Nestor meridionalis	SP215	NZCOI411-09	648 [0n]		2	
Nestor meridionalis	SP409	NZCOI682-09	648 [0n]		2	

Nestor notabilis	SP216	NZCOI412-09	648 [0n]		2	In captivity
Nestor notabilis	SP455	NZCOI728-09	648 [0n]		2	
Nestor notabilis	SP456	NZCOI729-09	648 [0n]		2	
Nestor notabilis	SP457	NZCOI730-09	648 [0n]		2	
Nestor notabilis	SP458	NZCOI731-09	648 [0n]		2	
Nestor notabilis	SP459	NZCOI732-09	648 [0n]		2	
Nestor notabilis	SP460	NZCOI733-09	0			
Ninox novaeseelandiae	1B-1114	NZCOI751-10	722 [0n]		2	
Ninox novaeseelandiae	SP131	NZCOI357-09	648 [0n]		2	
Ninox novaeseelandiae	SP132	NZCOI358-09	648 [0n]		2	
Ninox novaeseelandiae	SP133	NZCOI359-09	648 [0n]		2	
Ninox novaeseelandiae	SP134	NZCOI360-09	648 [0n]		2	
Notiomystis cincta	SP061	NZCOI295-09	648 [0n]		2	
Notiomystis cincta	SP062	NZCOI296-09	648 [0n]		2	
Notiomystis cincta	SP063	NZCOI297-09	648 [0n]		2	
Notiomystis cincta	SP064	NZCOI298-09	648 [0n]		2	
Notiomystis cincta	SP065	NZCOI299-09	648 [0n]		2	
Notiomystis cincta	SP066	NZCOI300-09	648 [0n]		2	
Notiomystis cincta	SP067	NZCOI301-09	648 [0n]		2	
Numenius madagascariensis	SP183	NZCOI548-09	648 [0n]		2	
Nycticorax caledonicus	SP147	NZCOI534-09	648 [0n]		2	
Oceanites nereis	SP232	NZCOI423-09	648 [0n]		2	
Oceanites nereis	SP233	NZCOI424-09	648 [0n]		2	Band no. B-56208
Oceanites nereis	SP234	NZCOI425-09	648 [0n]		2	Band no. B-56209
Oceanites nereis	SP235	NZCOI426-09	648 [0n]		2	Band no. B-56212
Oceanites nereis	SP236	NZCOI427-09	648 [0n]		2	Band no. B-56213
Oceanites oceanicus	SP237	NZCOI562-09	648 [0n]		2	
Pachyptila belcheri	SP244	NZCOI567-09	648 [0n]		2	
Pachyptila belcheri	SP381	NZCOI669-09	648 [0n]		2	
Pachyptila crassirostris	SP245	NZCOI568-09	0		2	
Pachyptila desolata	SP240	NZCOI564-09	648 [0n]		2	
Pachyptila salvini	SP241	NZCOI565-09	648 [0n]		2	
Pachyptila turtur	SP380	NZCOI668-09	648 [0n]		1	
Pachyptila vittata	SP238	NZCOI428-09	648 [0n]		2	Band no. D-106539
Pachyptila vittata	SP239	NZCOI563-09	648 [0n]		2	
Pachyptila vittata	SP379	NZCOI667-09	648 [0n]		2	
Pagodroma nivea	WS 591	ROMC287-07	727 [3n]		2	
Pagodroma nivea	WS 590	ROMC286-07	729 [1n]		2	
Pagodroma nivea	WS 589	ROMC285-07	729 [0n]		2	
Pagodroma nivea	WS 588	ROMC284-07	742 [0n]		2	
Pagodroma nivea	WS 581	ROMC283-07	727 [0n]		2	
Pagodroma nivea	WS 580	ROMC282-07	727 [0n]		2	
Pagodroma nivea	WS 553	ROMC281-	727		2	

Pagodroma nivea	WS 553	07	[0n]				
Pagodroma nivea	WS 506	ROMC280-07	727 [0n]				
Passer domesticus	SP106	NZCOI337-09	648 [0n]				
Passer domesticus	SP107	NZCOI338-09	648 [0n]				
Passer domesticus	SP108	NZCOI339-09	648 [0n]				
Passer domesticus	SP466	NZCOI739-09	648 [0n]				
Passer domesticus	SP467	NZCOI740-09	648 [0n]				
Passer domesticus	SP468	NZCOI741-09	648 [0n]				
Passer domesticus	SP469	NZCOI742-09	648 [0n]				
Passer domesticus	SP470	NZCOI743-09	648 [0n]				
Passer domesticus	SP471	NZCOI744-09	648 [0n]				
Passer domesticus	SP346	NZCOI640-09	648 [2n]				
Passer domesticus	SP347	NZCOI641-09	648 [0n]				
Passer domesticus	SP348	NZCOI642-09	648 [2n]				
Passer domesticus	SP349	NZCOI643-09	648 [2n]				
Pelagodroma marina	SP218	NZCOI414-09	648 [0n]				
Pelagodroma marina	SP219	NZCOI558-09	648 [0n]				
Pelagodroma marina	SP368	NZCOI657-09	648 [0n]				
Pelagodroma marina	SP369	NZCOI658-09	648 [0n]				
Pelecanoides georgicus	SP231	NZCOI561-09	648 [0n]				
Pelecanoides georgicus	SP377	NZCOI665-09	648 [0n]				
Pelecanoides georgicus	SP378	NZCOI666-09	648 [0n]				
Pelecanoides urinatrix	89-29	BROMB482-07	705 [1n]				ssp: urinator, Pelecanoides urinator
Pelecanoides urinatrix	MKP 1638	ROMC300-07	774 [0n]				urinatrix
Pelecanoides urinatrix	SP228	NZCOI421-09	648 [0n]				Band no. D-106532
Pelecanoides urinatrix	SP229	NZCOI422-09	648 [0n]				Band no. D-106540
Pelecanoides urinatrix	SP375	NZCOI663-09	648 [0n]				Band no. D178694
Pelecanoides urinatrix	SP376	NZCOI664-09	648 [0n]				Band no. D181592
Pelecanus conspicillatus	SP487	NZCOI767-10	648 [0n]				
Petroica australis	AJB 5648	BROMB487-07	731 [0n]				ssp:australis
Petroica australis	JW1024	NZCOI225-08	694 [0n]				
Petroica australis	JW1025	NZCOI226-08	694 [0n]				
Petroica australis	JW1026	NZCOI227-08	655 [0n]				
Petroica australis	JW1027	NZCOI228-08	655 [0n]				
Petroica australis	JW1028	NZCOI229-08	694 [0n]				
Petroica australis	SP073	NZCOI306-09	648 [0n]				
Petroica australis	SP074	NZCOI307-09	648 [0n]				
Petroica australis	SP075	NZCOI308-09	648 [0n]				
Petroica australis	SP076	NZCOI309-09	648 [0n]				
Petroica australis	SP077	NZCOI310-09	648 [0n]				
Petroica australis	SP078	NZCOI311-09	648 [0n]				
Petroica australis	SP079	NZCOI312-09	648 [0n]				

Petroica australis	SP472		NZCOI745-09	648 [0n]			
Petroica australis	SP473		NZCOI746-09	648 [0n]			
Petroica australis	SP474		NZCOI747-09	648 [0n]			
Petroica australis	SP475		NZCOI748-09	648 [0n]			
Petroica australis	SP476		NZCOI749-09	648 [0n]			
Petroica macrocephala	JW264		NZCOI105-08	648 [0n]			Dead from hatchery
Petroica macrocephala	SP086		NZCOI318-09	648 [0n]			
Petroica macrocephala	SP087		NZCOI319-09	648 [0n]			
Petroica macrocephala	SP088		NZCOI320-09	648 [0n]			
Petroica macrocephala	SP089		NZCOI321-09	648 [0n]			
Petroica macrocephala	SP090		NZCOI322-09	648 [0n]			
Petroica macrocephala	SP091		NZCOI323-09	648 [0n]			
Petroica macrocephala	SP092		NZCOI324-09	648 [0n]			
Petroica traversi	JW896		NZCOI063-08	648 [0n]			endemic Chatham Islands
Petroica traversi	JW899		NZCOI064-08	648 [0n]			endemic Chatham Islands
Petroica traversi	JW901		NZCOI065-08	648 [0n]			endemic Chatham Islands
Petroica traversi	JW903		NZCOI066-08	648 [0n]			endemic Chatham Islands
Petroica traversi	JW686		NZCOI104-08	648 [0n]			
Petroica traversi	JW1019		NZCOI220-08	694 [0n]			
Petroica traversi	JW1020		NZCOI221-08	694 [0n]			
Petroica traversi	JW1021		NZCOI222-08	687 [0n]			
Petroica traversi	JW1022		NZCOI223-08	694 [0n]			
Petroica traversi	JW1023		NZCOI224-08	694 [0n]			
Petroica traversi	JW895		NZCOI590-09	0			
Petroica traversi	JW897		NZCOI591-09	0			
Petroica traversi	JW898		NZCOI592-09	0			
Phaethon lepturus	SP480		NZCOI760-10	648 [0n]			
Phaethon rubricauda	SP479		NZCOI759-10	648 [0n]			
Phalacrocorax carbo	SP192		NZCOI552-09	648 [0n]			
Phalacrocorax melanoleucos	SP194		NZCOI395-09	648 [0n]			
Phalacrocorax melanoleucos	SP195		NZCOI553-09	648 [0n]			
Phalacrocorax punctatus	SP324		NZCOI579-09	648 [0n]			
Phalacrocorax sulcirostris	SP196		NZCOI554-09	648 [0n]			
Phalacrocorax varius	SP493		NZCOI773-10	648 [0n]			
Philesturnus carunculatus	36 C.carunculatus		BROMB641-07	775 [1n]			
Philesturnus carunculatus	28 C.carunculatus		BROMB640-07	775 [1n]			
Philesturnus carunculatus	13 C.carunculatus		BROMB639-07	791 [1n]			
Philesturnus carunculatus	JW688		NZCOI067-08	648 [0n]			endemic
Philesturnus carunculatus	JW980		NZCOI181-08	694 [0n]			
Philesturnus carunculatus	JW981		NZCOI182-08	694 [0n]			
Philesturnus carunculatus	JW982		NZCOI183-08	694 [0n]			
Philesturnus carunculatus	JW983		NZCOI184-08	694 [0n]			
Philesturnus			NZCOI185-	694			

	carunculatus	JW984	08	[0n]			
☑	Philesturnus carunculatus	JW985	NZCOI186-08	694 [0n]			
☑	Philesturnus carunculatus	JW986	NZCOI187-08	694 [0n]			
☑	Philesturnus carunculatus	JW987	NZCOI188-08	694 [0n]			
☑	Philesturnus carunculatus	JW988	NZCOI189-08	694 [0n]			
☑	Philesturnus carunculatus	SP081	NZCOI313-09	648 [0n]			
☑	Philesturnus carunculatus	JW689	NZCOI583-09	0			
☑	Philesturnus carunculatus	JW763	NZCOI585-09	0			
☑	Philesturnus carunculatus	JW765	NZCOI586-09	0			
☑	Phoebetria palpebrata	SP269	NZCOI449-09	648 [0n]			Band no.M-62049
☑	Phoebetria palpebrata	SP270	NZCOI572-09	648 [0n]			
☑	Platalea regia	SP145	NZCOI367-09	648 [0n]			
☑	Platycercus eximius	SP212	NZCOI408-09	648 [0n]			
☑	Platycercus eximius	SP367	NZCOI656-09	648 [0n]			
☑	Pluvialis fulva	SP174	NZCOI386-09	648 [0n]			
☑	Pluvialis fulva	SP175	NZCOI543-09	648 [0n]			
☑	Pluvialis squatarola	SP484	NZCOI764-10	648 [0n]			
☑	Podiceps cristatus	SP444	NZCOI717-09	0			Chick
☑	Podiceps cristatus	SP445	NZCOI718-09	648 [0n]			Chick
☑	Poliocephalus poliocephalus	SP323	NZCOI578-09	648 [0n]			
☑	Poliocephalus rufopectus	SP321	NZCOI495-09	648 [0n]			
☑	Porphyrio mantelli	JW708	NZCOI047-08	649 [1n]			Endemic to NZ
☑	Porphyrio mantelli	JW709	NZCOI048-08	648 [0n]			Endemic to NZ
☑	Porphyrio mantelli	SP009	NZCOI260-09	648 [0n]			
☑	Porphyrio mantelli	SP326	NZCOI619-09	648 [0n]			
☑	Porphyrio mantelli	SP327	NZCOI620-09	648 [0n]			
☑	Porphyrio mantelli	SP328	NZCOI621-09	648 [0n]			
☑	Porphyrio mantelli	SP329	NZCOI622-09	648 [0n]			
☑	Porphyrio mantelli	SP330	NZCOI623-09	648 [0n]			
☑	Porphyrio mantelli	SP439	NZCOI712-09	648 [0n]			
☑	Porphyrio mantelli	SP440	NZCOI713-09	648 [0n]			
☑	Porphyrio mantelli	SP441	NZCOI714-09	648 [0n]			
☑	Porphyrio mantelli	SP442	NZCOI715-09	648 [0n]			
☑	Porphyrio mantelli	SP443	NZCOI716-09	648 [0n]			
☑	Porphyrio porphyrio	JW905	NZCOI037-08	648 [0n]			
☑	Porphyrio porphyrio	JW906	NZCOI038-08	648 [0n]			
☑	Porphyrio porphyrio	JW907	NZCOI039-08	648 [0n]			
☑	Porphyrio porphyrio	JW908	NZCOI040-08	648 [0n]			
☑	Porphyrio porphyrio	JW909	NZCOI041-08	648 [0n]			
☑	Porphyrio porphyrio	JW910	NZCOI042-08	648 [0n]			
☑	Porphyrio porphyrio	JW911	NZCOI043-08	648 [0n]			
☑	Porphyrio porphyrio	JW912	NZCOI044-08	648 [0n]			
☑	Porphyrio porphyrio	JW913	NZCOI045-08	648 [0n]			

Porphyrio porphyrio	JW914	NZCOI046-08	648 [0n]				
Porphyrio porphyrio	SP012	NZCOI510-09	648 [0n]				
Porphyrio porphyrio	SP446	NZCOI719-09	648 [0n]				
Porphyrio porphyrio	SP447	NZCOI720-09	648 [0n]				
Porzana tabuensis	SP002	NZCOI256-09	648 [0n]				
Porzana tabuensis	SP003	NZCOI505-09	648 [0n]				
Procellaria cinerea	SP243	NZCOI566-09	648 [0n]				
Procellaria parkinsoni	SP217	NZCOI413-09	648 [0n]				
Procellaria westlandica	B-23382	BROMB712-07	713 [0n]				
Procellaria westlandica	B-23383	BROMB706-07	749 [0n]				
Procellaria westlandica	SP400	NZCOI613-09	648 [0n]				Band no. 37686
Procellaria westlandica	SP401	NZCOI614-09	648 [0n]				Band no. 39121
Procellaria westlandica	SP402	NZCOI615-09	648 [0n]				Band no. 25864
Procellaria westlandica	SP403	NZCOI616-09	648 [0n]				Band no. 38103
Procellaria westlandica	SP404	NZCOI617-09	648 [0n]				Band no. 39284
Procellaria westlandica	SP405	NZCOI618-09	648 [0n]				Band no. 22562
Prosthemadera novaeseelandiae	1B-958	BROMB581-07	794 [0n]				
Prosthemadera novaeseelandiae	1B-74	BROMB325-06	733 [0n]				
Prosthemadera novaeseelandiae	JW237	NZCOI110-08	527 [0n]				
Prosthemadera novaeseelandiae	SP037	NZCOI271-09	648 [0n]				
Prosthemadera novaeseelandiae	SP038	NZCOI272-09	648 [0n]				
Prosthemadera novaeseelandiae	SP039	NZCOI273-09	648 [0n]				
Prosthemadera novaeseelandiae	SP040	NZCOI274-09	648 [0n]				
Prosthemadera novaeseelandiae	SP041	NZCOI275-09	648 [0n]				
Prosthemadera novaeseelandiae	SP042	NZCOI276-09	648 [0n]				
Prosthemadera novaeseelandiae	SP043	NZCOI277-09	648 [0n]				
Prosthemadera novaeseelandiae	SP044	NZCOI278-09	648 [0n]				
Prosthemadera novaeseelandiae	SP332	NZCOI625-09	648 [0n]				
Prunella modularis	SP461	NZCOI734-09	648 [0n]				
Prunella modularis	SP462	NZCOI735-09	648 [0n]				
Prunella modularis	SP463	NZCOI736-09	648 [0n]				
Prunella modularis	SP464	NZCOI737-09	648 [0n]				
Prunella modularis	SP465	NZCOI738-09	648 [0n]				
Prunella modularis	SP337	NZCOI631-09	648 [0n]				
Pterodroma axillaris	SP390	NZCOI608-09	648 [6n]				Band no. D167638
Pterodroma axillaris	SP391	NZCOI609-09	648 [3n]				Band no. D163157
Pterodroma axillaris	SP392	NZCOI610-09	648 [0n]				Band no. D171732
Pterodroma axillaris	SP393	NZCOI611-09	648 [7n]				Band no. D171812
Pterodroma axillaris	SP394	NZCOI612-09	648 [9n]				Band no. D171774
Pterodroma cervicalis	SP309	NZCOI483-09	648 [0n]				
Pterodroma cookii	SP271	NZCOI450-09	648 [0n]				
Pterodroma cookii	SP272	NZCOI451-09	648 [0n]				

Pterodroma cookii	SP273	NZCOI452-09	648 [0n]			
Pterodroma cookii	SP274	NZCOI453-09	648 [0n]			
Pterodroma cookii	SP275	NZCOI454-09	648 [0n]			
Pterodroma cookii	SP276	NZCOI455-09	648 [0n]			
Pterodroma cookii	SP277	NZCOI456-09	648 [0n]			
Pterodroma cookii	SP278	NZCOI457-09	648 [0n]			
Pterodroma cookii	SP279	NZCOI458-09	648 [0n]			
Pterodroma cookii	SP280	NZCOI459-09	648 [0n]			
Pterodroma inexpectata	SP281	NZCOI460-09	648 [0n]			
Pterodroma inexpectata	SP382	NZCOI603-09	648 [0n]			Band no. 14812
Pterodroma inexpectata	SP383	NZCOI604-09	648 [0n]			Band no.
Pterodroma inexpectata	SP384	NZCOI605-09	648 [0n]			Band no. 14824
Pterodroma inexpectata	SP385	NZCOI606-09	648 [0n]			Band no. 14821
Pterodroma inexpectata	SP386	NZCOI607-09	648 [0n]			Band no. 14822
Pterodroma macroptera	SP427	NZCOI700-09	648 [0n]			
Pterodroma macroptera	SP428	NZCOI701-09	648 [0n]			
Pterodroma macroptera	SP429	NZCOI702-09	648 [0n]			
Pterodroma macroptera	SP430	NZCOI703-09	648 [0n]			
Pterodroma macroptera	SP431	NZCOI704-09	648 [0n]			
Pterodroma magentae	JW874	NZCOI001-08	648 [0n]			
Pterodroma magentae	SP310	NZCOI484-09	648 [0n]			Chick, Band No. E108951
Pterodroma magentae	SP311	NZCOI485-09	648 [0n]			Band No. E107404
Pterodroma magentae	SP312	NZCOI486-09	648 [0n]			Band No. E107414
Pterodroma magentae	SP313	NZCOI487-09	648 [0n]			Band No. E107432
Pterodroma neglecta	SP293	NZCOI468-09	648 [0n]			
Pterodroma neglecta	SP294	NZCOI469-09	648 [0n]			
Pterodroma neglecta	SP295	NZCOI470-09	648 [0n]			
Pterodroma neglecta	SP296	NZCOI471-09	648 [0n]			
Pterodroma neglecta	SP297	NZCOI472-09	648 [0n]			Chick
Pterodroma neglecta	SP298	NZCOI473-09	648 [0n]			
Pterodroma nigripennis	SP299	NZCOI474-09	648 [0n]			Chick
Pterodroma nigripennis	SP300	NZCOI475-09	648 [0n]			Chick
Pterodroma nigripennis	SP301	NZCOI476-09	648 [0n]			Chick
Pterodroma nigripennis	SP302	NZCOI477-09	648 [0n]			
Pterodroma nigripennis	SP303	NZCOI478-09	648 [0n]			Chick
Pterodroma nigripennis	SP395	NZCOI598-09	648 [0n]			Band no. D157705
Pterodroma nigripennis	SP396	NZCOI599-09	648 [0n]			Band no. D157706
Pterodroma nigripennis	SP397	NZCOI600-09	648 [0n]			Band no. D157708
Pterodroma nigripennis	SP398	NZCOI601-09	648 [0n]			Band no. D157709
Pterodroma nigripennis	SP399	NZCOI602-09	648 [0n]			Band no. D157710
Puffinus assimilis	SP220	NZCOI415-09	648 [0n]			
Puffinus assimilis	SP221	NZCOI416-09	648 [0n]			

Puffinus assimilis	SP222	NZCOI417-09	648 [0n]		2	
Puffinus assimilis	SP370	NZCOI659-09	648 [0n]		2	
Puffinus assimilis	SP371	NZCOI660-09	648 [0n]		2	
Puffinus carneipes	JW808	NZCOI022-08	648 [0n]		2	
Puffinus carneipes	JW809	NZCOI023-08	648 [0n]		2	
Puffinus carneipes	JW810	NZCOI024-08	648 [0n]		2	
Puffinus carneipes	JW812	NZCOI025-08	648 [0n]		2	
Puffinus carneipes	JW813	NZCOI026-08	648 [0n]		2	
Puffinus carneipes	JW814	NZCOI027-08	648 [0n]		2	
Puffinus carneipes	JW816	NZCOI028-08	648 [0n]		2	
Puffinus carneipes	JW817	NZCOI029-08	648 [0n]		2	
Puffinus carneipes	JW039	NZCOI581-09	0		2	
Puffinus carneipes	SP372	NZCOI661-09	648 [0n]		2	Band no. Z-26806
Puffinus carneipes	SP373	NZCOI662-09	648 [0n]		2	Band no. Z-50491
Puffinus gavia	1B-1256	NZCOI791-10	736 [0n]		2	
Puffinus gavia	MKP 1629	NZCOI753-10	722 [0n]		2	
Puffinus gavia	MKP 1640	NZCOI752-10	722 [0n]		1	
Puffinus gavia	SP290	NZCOI467-09	648 [0n]		2	
Puffinus gavia	SP291	NZCOI575-09	648 [0n]		2	
Puffinus griseus	JW791	NZCOI016-08	648 [0n]		2	
Puffinus griseus	JW792	NZCOI017-08	648 [0n]		2	
Puffinus griseus	JW794	NZCOI018-08	648 [0n]		2	
Puffinus griseus	JW795	NZCOI019-08	648 [0n]		2	
Puffinus griseus	JW796	NZCOI020-08	648 [0n]		2	
Puffinus griseus	JW807	NZCOI155-08	648 [0n]		2	
Puffinus griseus	SP282	NZCOI461-09	648 [0n]		2	
Puffinus griseus	SP283	NZCOI462-09	648 [0n]		2	
Puffinus griseus	SP284	NZCOI463-09	648 [0n]		2	
Puffinus griseus	SP285	NZCOI464-09	648 [0n]		2	
Puffinus griseus	SP286	NZCOI573-09	648 [0n]		2	
Puffinus griseus	JW793	NZCOI587-09	0		2	
Puffinus griseus	JW797	NZCOI588-09	0		2	
Puffinus griseus	SP387	NZCOI670-09	648 [0n]		2	Band no. Z-23739
Puffinus griseus	SP388	NZCOI671-09	648 [0n]		2	Band no. Z-23862
Puffinus griseus	SP389	NZCOI672-09	648 [0n]		2	
Puffinus huttoni	SP223	NZCOI418-09	648 [0n]		2	
Puffinus pacificus	SP304	NZCOI479-09	648 [0n]		2	Chick
Puffinus pacificus	SP305	NZCOI480-09	648 [0n]		2	Chick
Puffinus pacificus	SP306	NZCOI481-09	648 [0n]		2	Chick
Puffinus pacificus	SP307	NZCOI482-09	648 [0n]		2	
Puffinus pacificus	SP308	NZCOI577-09	648 [0n]		2	
Puffinus pacificus	SP325	NZCOI580-09	648 [0n]		2	
Puffinus		NZCOI576-	648			

	tenuirostris	SP292	09	[0n]		
	Pygoscelis adeliae	JW959	NZCOI160-08	655 [0n]		4
	Pygoscelis adeliae	JW960	NZCOI161-08	686 [0n]		2
	Pygoscelis adeliae	JW961	NZCOI162-08	694 [0n]		2
	Pygoscelis adeliae	JW962	NZCOI163-08	647 [0n]		2
	Pygoscelis adeliae	JW963	NZCOI164-08	694 [0n]		2
	Pygoscelis adeliae	JW964	NZCOI165-08	692 [0n]		2
	Pygoscelis adeliae	JW965	NZCOI166-08	655 [0n]		4
	Pygoscelis adeliae	JW966	NZCOI167-08	694 [0n]		2
	Pygoscelis adeliae	JW967	NZCOI168-08	619 [3n]		2
	Pygoscelis adeliae	JW968	NZCOI169-08	647 [0n]		2
	Pygoscelis adeliae	D4	BROMB042-06	712 [0n]		2
	Pygoscelis adeliae	D 20	BROMB041-06	673 [0n]		2
	Pygoscelis adeliae	S 16	BROMB040-06	695 [0n]		2
	Pygoscelis adeliae	C 98	BROMB039-06	710 [0n]		2
	Pygoscelis adeliae	D9	BROMB038-06	766 [0n]		2
	Pygoscelis adeliae	B 59	BROMB028-06	896 [0n]		2
	Pygoscelis adeliae	C 79	BROMB077-06	794 [0n]		2
	Rallus pectoralis	SP007	NZCOI507-09	648 [0n]		2
	Rallus philippensis	SP004	NZCOI257-09	648 [0n]		2
	Rallus philippensis	SP005	NZCOI258-09	648 [0n]		2
	Rallus philippensis	SP006	NZCOI506-09	648 [0n]		2
	Recurvirostra novaehollandiae	SP506	NZCOI786-10	648 [0n]		
	Rhipidura fuliginosa	SP045	NZCOI279-09	648 [0n]		2
	Rhipidura fuliginosa	SP046	NZCOI280-09	648 [0n]		2
	Rhipidura fuliginosa	SP333	NZCOI626-09	648 [0n]		2
	Stercorarius maccormicki	JW1039	NZCOI240-08	694 [0n]		2
	Stercorarius maccormicki	JW1040	NZCOI241-08	694 [0n]		2
	Stercorarius maccormicki	JW1041	NZCOI242-08	694 [0n]		2
	Stercorarius maccormicki	JW1042	NZCOI243-08	694 [0n]		2
	Stercorarius maccormicki	JW1043	NZCOI244-08	642 [0n]		2
	Stercorarius maccormicki	JW1044	NZCOI245-08	646 [0n]		2
	Stercorarius maccormicki	JW1045	NZCOI246-08	694 [0n]		2
	Stercorarius maccormicki	JW1046	NZCOI247-08	694 [0n]		2
	Stercorarius maccormicki	JW1048	NZCOI249-08	694 [0n]		2
	Stercorarius parasiticus	SP179	NZCOI544-09	648 [0n]		2
	Stercorarius skua	JW969	NZCOI170-08	665 [0n]		2
	Stercorarius skua	JW971	NZCOI172-08	694 [0n]		2
	Stercorarius skua	JW972	NZCOI173-08	694 [0n]		2
	Stercorarius skua	JW974	NZCOI175-08	675 [0n]		2
	Stercorarius skua	JW975	NZCOI176-08	694 [0n]		2
	Stercorarius skua	JW978	NZCOI179-08	694 [0n]		2
	Sterna fuscata	SP176	NZCOI387-09	648 [0n]		2

Sterna paradisaea	SP357	NZCOI651-09	648 [0n]		2	Band no. C-34618
Sterna striata	JW1057	NZCOI158-08	648 [0n]		2	
Sterna striata	SP177	NZCOI388-09	648 [0n]		2	
Sterna striata	SP360	NZCOI654-09	648 [0n]		2	
Sternula albifrons	SP483	NZCOI763-10	648 [0n]			
Sternula nereis	JW695	NZCOI098-08	648 [0n]		4	
Sternula nereis	SP415	NZCOI688-09	648 [0n]		2	
Sternula nereis	SP416	NZCOI689-09	648 [0n]		2	
Sternula nereis	SP417	NZCOI690-09	648 [0n]		2	
Sternula nereis	SP418	NZCOI691-09	648 [0n]		2	
Streptopelia chinensis	SP120	NZCOI526-09	648 [0n]		2	
Streptopelia roseogrisea	SP118	NZCOI525-09	648 [0n]		2	
Strigops habroptila	JW989	NZCOI190-08	648 [0n]		6	chick
Strigops habroptila	SP419	NZCOI692-09	648 [0n]		2	
Strigops habroptila	SP420	NZCOI693-09	648 [0n]		2	
Strigops habroptila	SP432	NZCOI705-09	648 [0n]		2	
Strigops habroptila	SP433	NZCOI706-09	648 [0n]		2	
Strigops habroptila	SP434	NZCOI707-09	648 [0n]		2	
Strigops habroptila	SP435	NZCOI708-09	648 [0n]		2	
Strigops habroptila	SP436	NZCOI709-09	648 [0n]		2	
Strigops habroptila	SP437	NZCOI710-09	648 [0n]		2	
Sturnus vulgaris	SP110	NZCOI341-09	648 [0n]		2	
Sturnus vulgaris	SP111	NZCOI342-09	648 [0n]		2	
Sturnus vulgaris	SP112	NZCOI343-09	648 [0n]		2	
Sturnus vulgaris	JW1124	NZCOI498-09	648 [0n]		2	
Sturnus vulgaris	SP350	NZCOI644-09	648 [0n]		2	
Sula dactylatra	SP206	NZCOI403-09	648 [0n]		2	Band no. 0-28447
Sula dactylatra	SP207	NZCOI404-09	648 [0n]		2	Band no. 0-28450
Sula dactylatra	SP208	NZCOI405-09	648 [0n]		2	Band no. 0-28449
Sula dactylatra	SP209	NZCOI406-09	648 [0n]		2	Chick, Band no. 0-28448
Sula dactylatra	SP210	NZCOI407-09	648 [0n]		2	Band no. 0-28419
Sula leucogaster	SP478	NZCOI758-10	648 [0n]			
Sula leucogaster	SP211	NZCOI557-09	648 [0n]		2	
Synoicus ypsilophorus	SP316	NZCOI490-09	648 [0n]		2	
Tachybaptus novaehollandiae	SP485	NZCOI765-10	648 [0n]			
Tachybaptus novaehollandiae	SP322	NZCOI496-09	648 [0n]		2	
Tadorna tadornoides	SP015	NZCOI512-09	648 [0n]		2	
Tadorna variegata	SP014	NZCOI261-09	648 [0n]		2	
Thalassarche chlororhynchos	SP492	NZCOI772-10	648 [0n]			
Thalassarche chrysostoma	SP252	NZCOI570-09	648 [0n]		2	
Thalassarche melanophris	SP263	NZCOI571-09	648 [0n]		2	
Thalasseus bergii	SP486	NZCOI766-10	648 [0n]			
Thalassoica		BROMB707-	749		2	

antarctica	B-13536	07	[0n]			
Thalassoica antarctica	B-13544	BROMB703-07	740 [0n]			
Thalassoica antarctica	SP225	NZCOI559-09	648 [0n]			
Thinornis novaeseelandiae	JW876	NZCOI030-08	648 [0n]			
Thinornis novaeseelandiae	JW879	NZCOI031-08	648 [0n]			
Thinornis novaeseelandiae	JW880	NZCOI032-08	648 [0n]			
Thinornis novaeseelandiae	JW881	NZCOI033-08	648 [1n]			
Thinornis novaeseelandiae	JW882	NZCOI034-08	648 [1n]			
Thinornis novaeseelandiae	JW883	NZCOI035-08	648 [0n]			
Thinornis novaeseelandiae	JW884	NZCOI036-08	648 [0n]			
Thinornis novaeseelandiae	JW878	NZCOI589-09	0			
Threskiornis molucca	SP148	NZCOI535-09	648 [0n]			
Todiramphus sanctus	JW214	NZCOI125-08	648 [0n]			
Todiramphus sanctus	JW285	NZCOI126-08	648 [0n]			
Todiramphus sanctus	JW286	NZCOI127-08	648 [0n]			
Todiramphus sanctus	JW287	NZCOI128-08	648 [0n]			
Tringa nebularia	SP504	NZCOI784-10	0			
Tringa stagnatilis	SP499	NZCOI779-10	648 [0n]			
Turdus merula	JW240	NZCOI094-08	648 [0n]			
Turdus merula	JW284	NZCOI095-08	729 [0n]			
Turdus merula	SP053	NZCOI287-09	648 [0n]			
Turdus merula	SP054	NZCOI288-09	648 [0n]			
Turdus merula	SP055	NZCOI289-09	648 [0n]			
Turdus merula	SP338	NZCOI632-09	648 [0n]			
Turdus merula	SP339	NZCOI633-09	648 [0n]			
Turdus merula	SP340	NZCOI634-09	648 [0n]			
Turdus philomelos	AJB 5558	BROMB495-07	746 [2n]			ssp:clarkei
Turdus philomelos	AJB 5546	BROMB489-07	738 [0n]			ssp:clarkei
Turdus philomelos	AJB 5545	BROMB488-07	725 [0n]			ssp:clarkei
Turdus philomelos	1B-1102	BROMB749-07	704 [0n]			
Turdus philomelos	SP060	NZCOI294-09	648 [0n]			
Turdus philomelos	SP344	NZCOI638-09	648 [0n]			
Vanellus miles	SP481	NZCOI761-10	648 [0n]			
Vanellus miles	JW1056	NZCOI159-08	648 **[10n]**			
Vanellus miles	SP154	NZCOI369-09	648 [0n]			
Xenus cinereus	SP495	NZCOI775-10	648 [0n]			
Zosterops lateralis	AJB 5632	BROMB532-07	719 [0n]			ssp:lateralis
Zosterops lateralis	AJB 5550	BROMB490-07	730 [0n]			ssp:lateralis
Zosterops lateralis	MKP 2000	BROMB851-07	802 [0n]			ssp: lateralis
Zosterops lateralis	JW891	NZCOI092-08	648 [0n]			
Zosterops lateralis	JW894	NZCOI093-08	648 [0n]			
Zosterops lateralis	SP345	NZCOI639-09	648 [0n]			