Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Birds in a tree: A journey through avian phylogeny, with particular emphasis on the birds of New Zealand

Gillian Claire Gibb

2010

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Genetics

Allan Wilson Centre for Molecular Ecology and Evolution
Institute of Molecular BioSciences
Te Kura Putaiao Koiora-a-Ngota
Massey University
Palmerston North
New Zealand

ABSTRACT

Two main themes to the avian research presented in this thesis are,
1. Deep resolution of birds generally, and
2. Investigation of specific aspects of the New Zealand avifauna.
More specifically, this thesis covers phylogeny, and predictions about palaeognaths, pigeons, pelecaniforms and passerines.

Significant progress is made in resolving the basal branches of Neoaves. This thesis examines whether the six-way basal Neoavian split of Cracraft (2001) is, in principle, resolvable. New mitochondrial genomes are added to improve taxon sampling, break up long branches, and allow testing of the prior assumptions of six Neoavian groups. This research shows the six-way split is resolvable, although more work is required for specific details. From a life-history perspective, it is interesting that the two bird-of-prey groups (falcons and buzzards) are very divergent, and may not be sister groups. Molecular dating supports major diversification of at least 12 Neoavian lineages in the Late Cretaceous. Additionally, novel avian mitochondrial gene orders are investigated and a hypothesis put forward suggesting gene conversion and stable intermediate forms allows an apparently rare event (gene rearrangement) to occur multiple times within Neoaves.

One of Cracraft's six groups, informally called the 'Conglomerati', is particularly difficult to resolve. The pigeons (Columbiformes) lie within the 'Conglomerati', and this chapter examines two aspects along the continuum of pigeon evolution. Firstly the large South Pacific fruit pigeon radiation is examined with mid-length mitochondrial sequences. This clade contains a third of all pigeon species, and has been very successful in island colonisation throughout South East Asia and the Pacific. Secondly, candidates for the closest relative of pigeons are tested using analysis of whole mitochondrial genomes. Highest support was found for the grouping of sandgrouse and pigeon, although they are clearly very divergent.

Also within the 'Conglomerati' is the traditional order Pelecaniformes, and their close allies the Ciconiiformes. These orders (the P&C) are part of an adaptive radiation of seabird water-carnivores, including loons, penguins, petrels and albatrosses. This group is separate from the large shorebird water-carnivore group; although both appear to have begun radiating abut 70 million years ago. The

tropicbird represents a separate, convergent life history and is not part of the Pelecaniformes, nor within the larger seabird water-carnivore group.

Resolution of the basal phylogeny of oscine passerines is important for interpreting the radiation of this group out of the Australasian region. Many endemic New Zealand oscine passerines belong to 'basal corvid' lineages, but have not previously been investigated with mitochondrial DNA. This chapter shows that many 'basal corvid' lineages are actually 'basal passerine' lineages, and there is a discrepancy between nuclear Rag-1 phylogenies (the most commonly used gene in passerine phylogenetics) and other phylogenies, including mitochondrial, that requires further investigation.

Taken as a whole, this thesis adds significantly to our understanding of the evolution of birds, and provides a foundation for future research, not only of phylogenetic relationships, but also of avian life history, long-term niche stability and macroevolution.

Acknowledgements

Firstly to David, thank you for being an amazing supervisor. Thank you for your incredible ability to read and return comments on just about anything overnight, for your wide-ranging knowledge, which I can only aspire to, and also for breakfast meetings!

To my co-supervisors, Matt, Trevor, and Lindell. Although widely dispersed when I started this project, you all ended up in Australia. I really appreciated visiting you all there, and meeting up at various conferences. In between times, thank you for always being at the end of an email.

To Trish, Abby, Matt and Kerryn, thank you for showing me how to do this funny thing called ‘sequencing mitochondrial genomes’ all those years ago. You guys taught me so much and I really appreciate it. Special thanks to Trish – you are the glue that holds this group together.

Thank you to Alan Cooper and Jeremy Austin for hosting me at the Australian Centre for Ancient DNA, and teaching me the principles of ancient DNA work. Thank you to Trevor, Jenny and Nic, for looking after me and taking me travelling around South Australia. With caves, fossils, kangaroos, wombats, and echidnas, you made my trip there something special.

To Steve, I’ve really enjoyed our ‘constructive rants’ on phylogeny, birds, New Zealand, and anything in between. Long may they continue!

Thank you to Joy and Susan for all your hard work behind the scenes and on the front line! You have always been able to ‘get stuff done’ at a moments notice, and I really appreciate that. And to Ann and Cynthia, thank you for being understanding when I put my head round the door for more milk!

On that note, cheers to all who came for a cup of coffee over the years. To Farside and Pleb labers, Boffin lounge and Goldfish bowl inhabitants. While too many to name here, you make the AWC a wonderful community to work in.

Thanks for Saturday games evenings, Friday SciFi DVDs, and lunchtime New York Times crosswords. Geeks and loving it! To Mort, Sue, Barbara, Klaus, Rogerio, Tim, Sylvia, Robin, Bennet, Atheer and Nick. For good food, good wine, good beer (!) and good company.

Thank you to my friends and family. For sanity, music and understanding, even when you're not really sure what I do everyday.

Lastly, thank you to Simon. Without you I might not have started along this route. Thank you for persuading me to do a PhD when I told myself I wasn't going to. Thank you for sharing this path with me.

TABLE OF CONTENTS

ABSTRACT i

ACKNOWLEDGEMENTS iii

LIST OF FIGURES x

LIST OF TABLES xi

1. Chapter One

Introduction

1.1. STRUCTURE OF THE THESIS 1

1.2. BACKGROUND 3

1.2.1. Mitochondrial DNA – why is it useful for phylogeny? 3

1.2.2. Alignments 4

1.2.3. Model choice, rates across sites 4

1.2.4. Gene trees vs. species trees 5

1.2.5. Consilience of induction, generating hypotheses and testing predictions 5

1.3. WHAT DO WE DO WHEN WE KNOW THE TREE OF LIFE? 6

1.3.1. Are the processes of microevolution sufficient to explain macroevolution? 6

1.3.2. Evolutionary-Stable Niche-Discontinuity (ESND) 7

1.4. CONTRIBUTIONS OF OTHERS TO THIS THESE PUBLICATIONS 8

1.5. REFERENCES 9

2. Chapter Two

Vicars, tramps and the assembly of New Zealand avifauna – A review of molecular phylogenetic evidence.

2.1. ABSTRACT 11

2.2. INTRODUCTION 12

2.2.1. A Brief History Of New Zealand 14

2.2.2. Overview Of The New Zealand Avifauna 17

2.3. MOLECULAR VIEW OF THE NZ AVIFAUNA 20

2.3.1. Phylogeny of birds 20

2.3.2. Birds in New Zealand 22

2.4. DISCUSSION 40

2.4.1. Accumulation and Radiation 41
2.4.2. A Zealandian element? 44
2.4.3. Disharmony & Extinction 46
2.4.4. Dating diversification 48
2.5. CONCLUSION 51
2.6. ACKNOWLEDGEMENTS 52
2.7. REFERENCES 52
2.8. SUPPLEMENTARY MATERIAL 71
2.8.1. Latin names of birds referred to in review 71
2.8.2. Peer reviewed molecular publications 72

3. Chapter Three

Mitochondrial Genomes and Avian Phylogeny: Complex Characters and Resolvability without Explosive Radiations

3.1. ABSTRACT 85
3.2. INTRODUCTION 86
3.3. MATERIALS AND METHODS 92
3.3.1. Phylogenetic Analysis 94
3.4. RESULTS 96
3.4.1. Phylogenetic Analysis 99
3.5. DISCUSSION 104
3.6. ACKNOWLEDGMENTS 108
3.7. REFERENCES 109

4. Chapter Four

Two aspects along the continuum of pigeon evolution: a South-Pacific radiation and the relationship of pigeons within Neoaves

4.1. ABSTRACT 115
4.2. INTRODUCTION 116
4.2.1. A South East Asian/Pacific radiation of pigeons. 116
4.2.2. The closest relative of Pigeons. 117
4.2.3. Nuclear versus Mitochondrial DNA and the problem of introns in deep phylogeny 117
4.3. METHODS 119
4.3.1. Taxon sampling 119

4.3.2. DNA extraction and PCR 119
4.3.3. Phylogenetic analyses 120
4.4. RESULTS 123
4.4.1. A South East Asian/Pacific radiation within Pigeons 123
4.4.2. Phylogenetic relationships of Pigeons within Neoaves 126
4.5. DISCUSSION 129
4.5.1. The Ducula–Ptilinopus radiation in a biogeographic context 129
4.5.2. Closest relative of Pigeons 132
4.6. ACKNOWLEDGEMENTS 133
4.7. REFERENCES 133

5. Chapter Five

Beyond Phylogeny: Pelecaniform and Ciconiiform Birds, and Long-Term Niche Stability

5.1. ABSTRACT 139
5.2. **INTRODUCTION** 140
5.2.1. Consilience of Induction 140
5.2.2. History of the Pelecaniformes 141
5.3. METHODS 144
5.3.1. Taxon Sampling, DNA Extraction, PCR and Sequencing 144
5.3.2. Species for Analysis 145
5.3.3. Alignment 146
5.3.4. Phylogenetic Analyses and Tree Building 146
5.3.5. Molecular Dating 147
5.4. RESULTS AND DISCUSSION 147
5.4.1. Details of New Mitochondrial Genomes and Gene Orders 147
5.4.2. Implications of Core Pelecaniform Gene Duplications for Phylogenetic Studies 148
5.4.3. Phylogenetic Analyses 150
5.4.4. Resolution of the Core Pelecaniforms 152
5.4.5. The Pelican is not a Pelecaniform 153
5.4.6. The predicted P&C Group 155
5.4.7. The Nature of the Tropicbird 156
5.4.8. Molecular Dating 157

5.4.9. Water-Carnivory in Neoaves, Adaptive Radiations and Niche Stability 158
5.5. CONCLUSION 160
5.6. **ACKNOWLEDGEMENTS** 161
5.7. REFERENCES 161
5.8. APPENDIX 1 167
5.8.1. Predicting the probabilities of observing splits in a new dataset. 167

6. Chapter Six

New Zealand Passerines Help Resolve Basal Oscine Phylogeny.

6.1. ABSTRACT 169
6.2. INTRODUCTION 170
6.2.1. Oscine biogeography 170
6.2.2. Limitations of current deep oscine phylogenetic resolution 171
6.3. METHODS 172
6.3.1. DNA extraction, PCR and sequencing of modern samples 172
6.3.2. Ancient DNA protocols 173
6.3.3. Alignment 174
6.3.4. Phylogenetic Analyses 177
6.4. RESULTS & DISCUSSION 177
6.4.1. Mitogenomic features 177
6.4.2. Phylogenetic analyses 178
6.4.3. Ancient DNA sequence verification 179
6.4.4. Phylogenetic analysis of the piopio with broader oscine sampling 180
6.5. FUTURE WORK 183
6.5.1. Dating Passeriformes 183
6.6. CONCLUSION 185
6.7. ACKNOWLEDGEMENTS 185
6.8. REFERENCES 185
6.9. SUPPLEMENTARY MATERIAL 190

7. Chapter Seven

Thesis Summation 191
7.1. FUTURE DIRECTIONS: TO PHYLOGENY AND BEYOND 197
7.2. REFERENCES 198

8. Appendix One 201

Pratt, R. C., Gibb, G. C., Morgan-Richards, M., Phillips, M. J., Hendy, M. D., and D. Penny. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol 26: 313-326

9. Appendix Two 215

Phillips, M. J., G. C. Gibb, E. A. Crimp and D. Penny. 2010. Tinamous And Moa Flock Together: Mitochondrial Genome Sequence Analysis Reveals Independent Losses Of Flight Among Ratites. Syst Biol 59:90-107

LIST OF FIGURES

FIGURE 1.1	Five models for the divergence of birds and placental mammals.	6
FIGURE 1.2	Evolutionarily-stable niche-discontinuity between two taxa.	7
FIGURE 2.1	New Zealand and neighbouring islands	15
FIGURE 2.2	Geological timescale	16
FIGURE 2.3	Phylogeny of birds based on multiple genes	21
FIGURE 2.4	Phylogeny of passerine diversity	35
FIGURE 2.5	The effect of taxon sampling on tree interpretation	45
FIGURE 3.1	Gene orders found in avian mitochondrial genome control regions	88
FIGURE 3.2	Comparison between CR(1) and CR(2) in Aracari (*A*) and Osprey (*B*)	97
FIGURE 3.3	Primers used to sequence duplicate tThr-CR gene order	98
FIGURE 3.4	Maximum Likelihood tree of 40 birds	100
FIGURE 3.5	Consensus network for 27 Neoavian birds	102
FIGURE 4.1	Phylogeny of pigeons with emphasis on the SE Asia/Pacific radiation	124
FIGURE 4.2	Bayesian rooted consensus tree of pigeons	123
FIGURE 4.3	Maximum likelihood tree of birds	127
FIGURE 5.1	Comparison of duplicated regions in the three core Pelecaniformes	148
FIGURE 5.2	Maximum likelihood tree from analysis using RAxML, with 31 taxa	150
FIGURE 5.3	Consensus network of the four analytical methods	151
FIGURE 5.4	Maximum likelihood tree from analyses using RAxML of 35 taxa	152
FIGURE 5.5	Maximum clade credibility chronogram of the 31 taxa dataset	158
FIGURE 6.1	Unrooted maximum likelihood analysis of mt genome dataset	178
FIGURE 6.2	Network of Bayesian analysis	180
FIGURE 6.3	Network of Bayesian analysis without *Turnagra* U51734	182
FIGURE 6.4	Consensus cladograms of passerine phylogeny	184
FIGURE S6.1	Piopio specimen *Turnagra capensis*	190
FIGURE 7.1	Network showing splits over 20% of RAxML bootstrap analysis	194

LIST OF TABLES

TABLE 2.1	New Zealand native passerine diversity and endemicity	36
TABLE S2.1	Latin Names of Birds	71
TABLE S2.2	Peer reviewed publications reporting application of molecular data to New Zealand birds	72
TABLE 4.1	Accession numbers and mt genes included in analysis of pigeon phylogeny	120
TABLE 4.2	Accession numbers and species names for birds used in complete mt genome analyses	122
TABLE 4.3	Testing alternative hypotheses about pigeon relatives using the SH test	128
TABLE 5.1	Characters previously used to group pelecaniform families	143
TABLE 6.1	Genbank accessions of additional samples used in piopio analyses	175

CHAPTER ONE

1 INTRODUCTION

1.1 STRUCTURE OF THE THESIS

This thesis has been written as a series of seven manuscripts that are published, accepted, or submitted for publication in a range of journals. Because each manuscript chapter is designed to stand alone, there is some overlap and repetition between chapters. In this thesis mitochondrial gene sequences play an important role in understanding avian evolution in the New Zealand context. There are two main themes to the papers presented here.

1. Deep phylogenetic resolution of birds generally

2. Phylogenetic investigation of specific aspects of the New Zealand avifauna

Early work on avian phylogeny is well summarised in Sibley and Ahlquist (1990) and Sibley and Monroe (1991). Since that time, most of the work on the phylogeny of birds has used molecular data, and this thesis continues that theme. We contend that genetic (molecular) data is the primary data for phylogenetics, and then in combination with morphological, life history and ecological data, we can understand important biological events that have occurred. We are only beginning to scratch the surface of this goal, and so it is also our motivation for future research.

Chapter Two stands somewhat outside the two themes just mentioned in that it is a review of the current molecular phylogenetic literature involving New Zealand birds. The work includes background on New Zealand's biogeography and puts its avifauna in the context of what is known and what is expected of a cluster of large islands in the South Pacific. This chapter has been accepted for publication in the journal Ibis, and will be published in the April 2010 issue.

Chapter Three and **Appendix One** are part of a series focused on resolving the basal nodes of Neoaves using mitochondrial gene-coding sequences (previous research includes Harrison et al. 2004; Slack et al. 2006; Slack et al. 2007). **Chapter Three** investigates the utility of complex characters for phylogenetic reconstruction of Neoaves, analyses seven new mitochondrial genomes to increase resolution in Cracraft's 6-way split of Neoaves (Cracraft 2001), and breaks up long branches to help stabilize some 'problem taxa'. **Appendix One** continues this investigation by using conditional down-weighting to reduce noise relative to signal, increasing taxon sampling of Cracraft's six groups with an additional nine mitochondrial genomes, and introducing formulae for calculating probabilities of finding predefined groupings in the optimal tree. Importantly, these chapters show that progress is being made in resolving Neoaves, and therefore it does not represent an 'explosive radiation' (Poe and Chubb 2004). These chapters were published in Molecular Biology and Evolution in 2007 and 2009, respectively.

The subsequent chapters investigate questions relating to specific avian groups in the context of New Zealand and its surrounds.

Chapter Four focuses on two aspects in the continuum of pigeon phylogeny. Using mid-length mitochondrial gene sequences, the large South-Pacific radiation of frugivorous pigeons (the *Ducula-Ptilinopus* clade) is examined. The New Zealand pigeon, *Hemiphaga novaeseelandiae* is part of this clade. The complete mitochondrial genome of *H. novaeseelandiae* is reported, and used to investigate potential closest relatives of pigeons within Neoaves. This chapter has been submitted to Molecular Phylogenetics and Evolution.

Chapter Five goes beyond phylogeny and investigates a case of long-term niche stability within Neoaves. We present four new mitochondrial genomes, analysis of which helps conclude that a consensus on Pelecaniformes phylogeny has been

reached. We then begin to examine the wider context of Pelecaniformes and Ciconiiformes within the seabird water-carnivore radiation, a group that also includes penguins, loons and albatross. This group shows evidence of an adaptive radiation around 70 Ma. While this chapter does not specifically discuss the New Zealand fauna (although New Zealand species are represented), seabird water-carnivores comprise a significant proportion of the extant New Zealand avifauna. This chapter has been submitted to Systematic Biology.

Chapter Six investigates the resolution of deep oscine phylogeny within the Passeriformes (perching birds). Oscines appear to have originated in Australasia, and many New Zealand endemic species are basal oscines. With this analysis we triple the available oscine mitochondrial genomes and show the basal oscine nodes are resolvable. In addition, new sequences and analysis of the extinct NZ piopio show it is part of the core Corvoidea radiation, rather than being basal to bowerbirds as previously thought. This chapter forms the basis of a paper we intend to submit (with the addition of NZ robin, fantail and grey warbler) to Molecular Phylogenetics and Evolution.

Chapter Seven is a short chapter that discusses, summarises and concludes this thesis. It includes ideas for future work.

1.2 BACKGROUND

1.2.1 Mitochondrial DNA – why is it useful for phylogeny?

Long sequences with sufficient information are required for resolving deep avian divergences. An advantage of using mitochondrial DNA for deep phylogeny is that coding regions comprise most of its length, and are straightforward to align. Therefore a conservative alignment of deeply divergent birds containing 12 protein coding, two rRNA and 22 tRNA genes, with no gaps can be 13 Kb in length. In contrast, the 19 gene, 32 Kb alignment of Hackett *et al.* (2008) reduced to just 9.5 Kb when gaps were removed, and the 2 Kb β-fibrinogen intron 7 alignment of Fain and Houde (2004) reduces to zero when our stricter alignment criteria for indels are used (Morgan-Richards et al. 2008).

1.2.2 Alignments

Nuclear introns can evolve quickly, as they don't have coding constraints. This makes them useful targets for medium depth phylogeny. Unfortunately this trait also means they accumulate indels. In some cases indels can be seen as rare genomic changes, and hence useful for parsimonious tree reconstruction (Steel and Penny 2005). However, as divergences become deeper and indels become more numerous and also overlap, alignments and models that successfully account for gaps become increasingly difficult to achieve. Morgan-Richards *et al.* (2008) showed that at deep divergences, introns with many gaps can provide overriding signal for an incorrect tree, such as the Metaves/Coronaves split found by Fain and Houde (2004).

The dataset of Hackett et al. (2008) is predominantly from nuclear introns (74% of the 32Kb), which are known to be difficult to align at deep divergences (Morgan-Richards et al. 2008). Even if the alignment can be shown to be correct, the introduction of gaps into a dataset, which is highly likely with introns, can be difficult to model correctly (Wong et al. 2008). A contrasting view was provided by Chojnowski et al. (2008), who claimed in the specific case of the clathrin heavy chain genes, introns outperformed exons in basal avian phylogeny. However this study could just as well claim 3 Kb of sequence (the introns) outperforms 535 bp (the exons) when trying to resolve deep divergences.

Long exon datasets (nuclear and mitochondrial) are more likely to resolve deep phylogenetic divergences. However, sequencing divergent nuclear exons can be difficult as the regions adjacent to exons evolve faster, making universal PCR primers less useful and amplification more time consuming. Additionally, concatenating multiple short exons (or any unrelated sequences) from different genes can have its own pitfalls, as outlined below.

1.2.3 Model choice, rates across sites

Having created an alignment, model choice, accounting for rate variation across sites, and data partitioning are important factors in finding the optimal tree. **Chapter Three** provides a cautionary tale of why it is important to account for rates across sites in phylogenetics. Specifically, increased taxon-sampling and RY-coding of the rapidly evolving 3rd codon positions reduces tendencies for long-branch attraction between taxa, and provides stronger statistical support for the

optimal grouping. Additionally, separately modeled protein codon and RNA structural data partitions reduce the influence of branch-length biases found in some earlier studies (e.g. Haddrath and Baker 2001).

1.2.4 Gene trees vs. species trees

In some situations, the best phylogeny recovered from one or more genes may not in fact represent the species tree. While this is more often a problem with closely related taxa (for example within a genus), it cannot be ignored in avian phylogeny. Examples where gene trees do not always represent the species tree include β-fibrinogen intron 7 (in this case likely an alignment problem, Morgan-Richards et al. 2008) and the nuclear exon Rag-1 (Irestedt and Ohlson 2008, discussed in Chapter Eight). In both cases, concatenation with other genes did not remove the signal of the incongruent phylogeny (Ericson et al. 2006; Irestedt and Ohlson 2008). As such, agreement on phylogeny between multiple different datasets will always help in resolving the species tree.

1.2.5 Consilience of induction, generating hypotheses and testing predictions

In the longer term, we need to compare the results of studies using diverse datasets such as nuclear, mitochondrial and morphological characters so that we can evaluate where they agree and use sources of disagreement for further hypotheses and study. This allows us to test quantitatively the agreement between trees (Pratt et al. 2009). Far more confidence can be placed on results when there is agreement between datasets—in our case nuclear and mitochondrial. In general, as taxon sampling improves there is good agreement between nuclear (exon) and mitochondrial trees (Slack et al. 2007).

1.3 WHAT DO WE DO WHEN WE KNOW THE TREE OF LIFE?

Luckily, there is life after phylogeny! Simply resolving radiations within the avian tree, while an admirable and challenging job, is not our end goal. The following sections outline some of the questions that have influenced our thinking, and our goals for going beyond phylogeny.

1.3.1 Are the processes of microevolution sufficient to explain macroevolution?

The resolution of the basal nodes within Neoaves has been known to be a difficult problem for some time, and these basal divergences have been called an "explosive radiation" (see e.g. Poe and Chubb 2004). This expression raises concerns regarding the sufficiency of microevolutionary processes to explain macroevolution (Penny and Phillips 2004). As a test of the sufficiency of microevolution, Penny and Phillips (2004) outlined five models (reprinted here in Fig. 1.1), which can be tested through a combination of phylogeny, dating of lineages and determination of behavioral/life history traits.

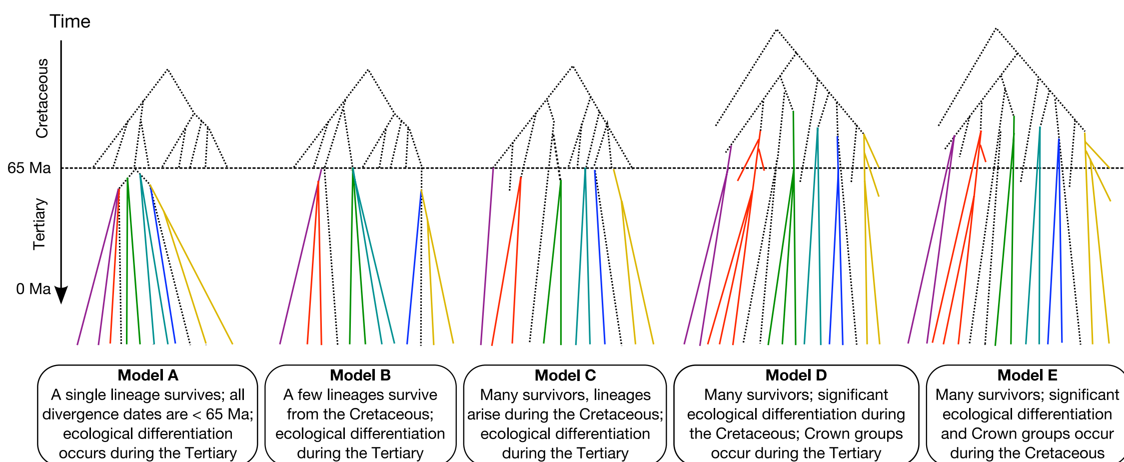

FIGURE 1.1—Five models for the divergence of birds and placental mammals. Each coloured solid line represents an order of modern mammals or birds. Dotted lines represent lineages in the ancestral niche (small insectivores for mammals, and probably ground-foragers for birds). In each case, families and genera evolve during the Tertiary but the models differ in the time both for the origin of orders and infra-orders, and for the ecological transitions from the ancestral niche. The models differ markedly in their implications for the mechanisms of macroevolution. Figure from Penny and Phillips (2004)

The concept of an 'explosive radiation' is expanded further in **Chapter Three**. We prefer to use the term 'adaptive radiation' where it can be seen that lineage divergences have been fast in geological time (such as parts of the basal Neoavian

radiation like the water carnivores, discussed further in **Chapter Five**), but where rates of ecological and morphological change are not known.

1.3.2 Evolutionary-Stable Niche-Discontinuity (ESND)

An adaptive radiation of lineages does not necessarily also mean fast morphological and ecological change. Morphological and ecological change may have potentially occurred over a completely different timescale, and this is often much harder to determine. It is possible for the morphology and ecology of a group to remain stable for long periods of time. The model of M. Phillips in Poole et al. (2003) outlines a method for understanding how groups can stay in an ecologically stable niche for long periods of time (tens of millions of years) (reprinted in Fig 1.2).

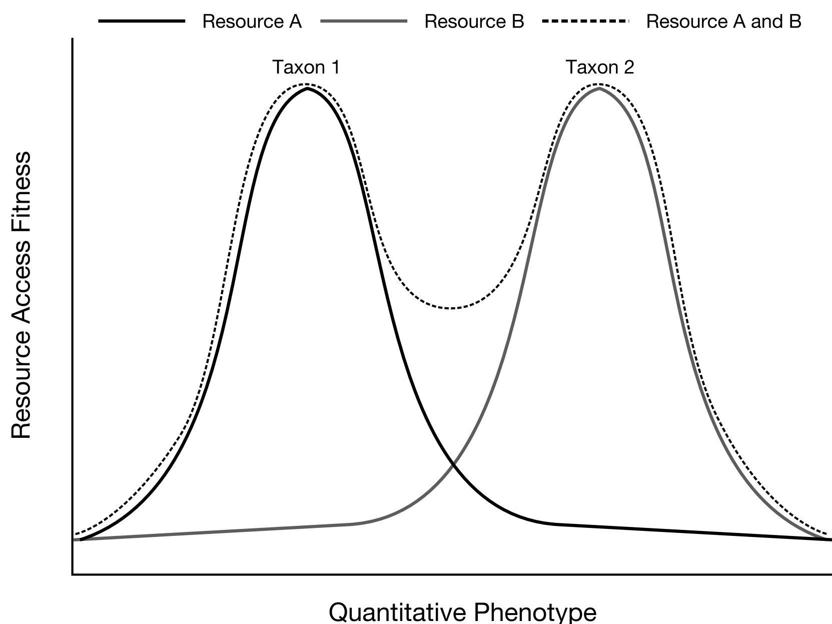

FIGURE 1.2—Evolutionarily-stable niche-discontinuity (ESND) between two taxa. The curves represent relative fitness contribution derived by an organism from access to resources A (black line) and B (grey line), as dependent on a quantitative phenotype. The sum of the curves for resources A + B (dashed line) represents the overall relative fitness of an organism with respect to the quantitative phenotype. The signature of an ESND is a selection pattern creating a valley of low fitness. This is expected to occur where a phenotype shift results in a deleterious trade-off between interspecific and intraspecific competition (i.e. towards the valley). From Poole et al. (2003).

Mapping of morphology, ecology and life history traits (i.e. biology) onto phylogenetic trees is a way to help answer questions in avian evolution, and a start is made here. But first, there is the minor obstacle of resolving and understanding avian phylogeny. The main research of this thesis is based on questions of avian phylogenetic resolution.

1.4 CONTRIBUTIONS OF OTHERS TO THIS THESE PUBLICATIONS

The bulk of the molecular and phylogenetic analyses, and writing of this thesis is my own work. I am the primary author on four of the papers. For these papers I developed the ideas, undertook the molecular work, created the alignments, performed all the phylogenetic analyses, and wrote the manuscripts. My co-authors variously sourced DNA samples, provided some mitochondrial genome sequences, and gave advice and discussion on the manuscript drafts.

Chapter Two is a collaboration with S. Trewick. This manuscript began as an invitation to S. Trewick to speak at the Australasian Ornithological Conference in Perth, 2007, and evolved into a review article for the journal Ibis. I collected and collated all the reference data and avian statistics that form the basis of the review material. We jointly developed the ideas and structure, and wrote the manuscript.

Two further published papers where I am a co-author are included as appendices. For a logical progression of arguments and ideas, it is suggested these appendices are read between **Chapters Three** and **Four**.

1.5 REFERENCES

Chojnowski, J. L., R. T. Kimball, and E. L. Braun. 2008. Introns outperform exons in analyses of basal avian phylogeny using clathrin heavy chain genes. Gene 410:89-96.

Cracraft, J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proc R Soc Lond B Biol Sci 268:459-469.

Ericson, P., C. Anderson, T. Britton, A. Elzanowski, and U. Johansson. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett 2:543-547.

Fain, M. G., and P. Houde. 2004. Parallel radiations in the primary clades of birds. Evolution 58:2558-2573.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Haddrath, O., and A. J. Baker. 2001. Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis. Proc R Soc Lond B Biol Sci 268:939-945.

Harrison, G. L., P. A. McLenachan, M. J. Phillips, K. E. Slack, A. Cooper, and D. Penny. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late Cretaceous. Mol Biol Evol 21:974-983.

Irestedt, M., and J. Ohlson. 2008. The division of the major songbird radiation into Passerida and 'core Corvoidea' (Aves: Passeriformes) — the species tree vs. gene trees. Zoologica Scripta 37:305-313.

Morgan-Richards, M., S. Trewick, A. Bartosch-Härlid, O. Kardailsky, M. J. Phillips, P. A. McLenachan, and D. Penny. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol Biol 8:20.

Penny, D., and M. J. Phillips. 2004. The rise of birds and mammals: are microevolutionary processes sufficient for macroevolution? Trends Ecol Evol 19:516-522.

Poe, S., and A. L. Chubb. 2004. Birds in a bush: Five genes indicate explosive evolution of avian orders. Evolution 58:404-415.

Poole, A. M., M. J. Phillips, and D. Penny. 2003. Prokaryote and eukaryote evolvability. Biosystems 69:163-185.

Pratt, R. C., G. C. Gibb, M. Morgan-Richards, M. J. Phillips, M. D. Hendy, and D. Penny. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol 26:313-326.

Sibley, C. G., and J. E. Ahlquist. 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale University Press, New Haven, Connecticut.

Sibley, C. G., and B. L. Monroe. 1991. Distribution and Taxonomy of Birds of the World. Yale University Press, New Haven, CT.

Slack, K. E., F. Delsuc, P. A. Mclenachan, U. Arnason, and D. Penny. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol Phylogenet Evol 42:1-13.

Slack, K. E., C. M. Jones, T. Ando, G. L. A. Harrison, R. E. Fordyce, U. Arnason, and D. Penny. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol 23:1144-1155.

Steel, M. A., and D. Penny. 2005. Maximum parsimony and the phylogenetic information in multistate characters. Pp. 163-178 *in* V. Albert, ed. Parsimony, phylogeny and genomics. Oxford University Press, Oxford (UK).

Wong, K. M., M. A. Suchard, and J. P. Huelsenbeck. 2008. Alignment uncertainty and genomic analysis. Science 319:473-476.

CHAPTER TWO

2 VICARS, TRAMPS AND THE ASSEMBLY OF NEW ZEALAND BIRDS – A REVIEW OF MOLECULAR PHYLOGENETIC EVIDENCE[1]

2.1 ABSTRACT

The avifauna of New Zealand is taxonomically and ecologically distinctive, as is typical of island biotas. The potential for an old age of New Zealand has however, tended to encourage a popular notion of a 'Moa's ark' based on the attractive proposition that the fauna was isolated when Zealandia broke from Gondwana c. 83 million years ago. Molecular phylogenetics has proved useful for exploring the relative importance of different biogeographic processes. Tramp species have arrived in New Zealand even in the last few hundred years. Others have such close phylogenetic relatives overseas (predominantly Australian) that it is clear their recent ancestors were tramps too. Distinctive taxa with deep phylogenetic ancestry might be vicars that owe their presence to vicariance, but lack of close morphological, taxonomic and phylogenetic affinity provides only tenuous evidence for this. Disproving the alternative possibility that apparent vicars are descended from tramps that dispersed in earlier times remains challenging, but molecular analyses have made startling revalations. Among iconic taxa, the world's largest eagle shared a Pleistocene ancestor with a small Australian eagle, and giant, flightless moa are phylogenetic sisters of much smaller, flying South American

[1] This chapter published as: Trewick, S., and G. C. Gibb. 2010. Vicars, Tramps and the Assembly of New Zealand Birds – A review of molecular phylogenetic evidence. Ibis 152:226-253

tinamous. The New Zealand avifauna is neither isolated nor stable, but demonstrative of prolonged and ongoing colonisation, speciation and extinction.

Keywords—Zealandia, dispersal, vicariance, Neoaves, molecular clock, colonisation, migration.

2.2 INTRODUCTION

The biotas of islands (whether oceanic or other habitat patches) are modelled both as equilibrium communities assembled by migration balanced by extinction (MacArthur and Wilson 1967) and as foci of speciation and adaptation (Darwin 1859). Birds have provided exemplars of both paradigms (e.g. Diamond 1974; Fleischer et al. 1998), making it clear that assembly of biotas even on strictly oceanic islands is complex. Where the islands concerned are deemed continental (Cowie and Holland 2006), and are seemingly of considerable age we might expect the processes of assembly to have stabilised and their contributions to be readily inferred.

New Zealand's biota, however, has been variously described as ill-balanced, unique, small, Gondwanic and a hotspot (e.g. Keast 1971; Daugherty et al. 1993; Gibbs 2006; McDowall 2008). It is well recognised that although endemicity is high, representation among taxonomic groups is highly variable and overall diversity in many groups is low. This is true of the birds, which provide an amenable group for the study of biogeography as they are well observed and taxonomically well resolved (in comparison to most other animal groups in New Zealand). Interestingly, although authors in the 1950s and 1960s were pragmatic about the origins of the New Zealand (NZ) aviafauna and recognised that the assemblage was largely what would be expected on a large oceanic island (e.g. Falla 1953; Fleming 1962; Caughley 1964), by the 1970s and 1980s a shift in biogeographic reasoning had taken place. Acceptance of continental drift, explained by plate tectonics (Tarling and Tarling 1971; Sutherland 1999; Graham 2008), stimulated a resurgence of interest in vicariance as an important process in the formation of the NZ biota and emphasis on its ancient origins. As continental drift would seem to have limited influence on flying birds, evidence for ancient isolation of the biota has generally been drawn from other sections of the biota. However, the distinctiveness

of NZ's avifauna is none the less routinely attributed to 'its long period of isolation' and linked with the observation of low taxonomic diversity (Baker 1991). In fact there is no readily apparent reason to connect low diversity with protracted isolation, as many other much smaller, more isolated and younger islands have high diversity and endemicity in birds and other fauna (Goldberg et al. 2008).

In considering the NZ avifauna, Robert Falla (1953) in his capacity as president of the Royal Australasian Ornithologists Union, discussed 'The Australian element in the avifauna of New Zealand'. Part of Falla's motivation at that time was to encourage greater research and awareness of bird taxonomy and ecology in neighbouring geographic areas. The benefits of this approach are evident in the gradual reduction over time in the number of endemic NZ taxa, by improved systematics (e.g. compare with Mathews and Iredale 1921 treatment). That outward looking perspective is still relevant today as we endeavour to disentangle the evolutionary history of regional biotas. Two important developments since Falla's (1953) paper that have profoundly influenced this endeavour are recognition of the potential importance of continental drift on composition of our biota, and the development and application of DNA based phylogenetic methods. The first had a rapid influence on prediction and interpretation of biogeographic hypotheses, whilst the second continues to develop as the most rewarding source of data with which to better understand the evolution of the species concerned and test biogeographic hypotheses. As advocated by Falla, however, appropriate geographic and taxonomic sampling remains an essential prerequisite in this research.

Caughley (1964) attempted to retrospectively predict the vertebrate composition of the NZ biota given its size, position relative to Australia and inferences about dispersal ability. Caughley (1964) and Falla (1953) made inferences not requiring continental drift to explain distribution patterns, except perhaps for the existence of ratites (kiwi and moa) in New Zealand. 'It must be admitted that the fauna of the NZ archipelago conforms, in the main, to what should be expected on large oceanic islands' (Falla 1953). Even the presence of the iconic New Zealand kiwi and moa have been ascribed to invasions overland (Fleming 1962) or by flying ancestors (De Beer 1956; Mayr 1963). Each of the endemic passerine families might also represent separate colonisations (McDowall 1968). These writers were interested in the possibility, and perhaps presumed, that dispersal was the primary or only source of NZ taxa. The necessity for this explanation diminished with acceptance of

continental drift (Skipworth 1974), which provided a simple mechanism to explain the distribution of lineages that appear to be poor dispersers. However, recent molecular studies cast doubt on assumed dispersal ability and colonisation opportunity (e.g. the Gondwanan vicariance icon Nothofagus beech, arrived after New Zealand formed Cook and Crisp 2005b; Knapp et al. 2005).

Three types of information can illuminate our understanding of the way in which biotas develop: observation of the local and neighbouring modern faunas and floras; the fossil record; and phylogenetic inference from DNA sequence data. The extant biota is our primary point of reference, but it represents a single time-slice through evolution, whereas fossils provide the only means of observing, albeit partial, biotas from other times. Since the 1980s, molecular phylogenetics has rapidly transformed our perceptions of the rates and modes of speciation, and allowed the effective inclusion of time in biogeography and systematics. The emerging field of genomics promises data that will allow detailed analysis of the genetic framework underlying speciation (Butlin 2008). Although molecular phylogenetic and genomic approaches can only inform directly on living organisms and recently extinct organisms, the resulting evolutionary lessons can be applied to interpretation of fossil evidence. Here we review the contribution so far from molecular phylogenetics on the assembly of the NZ avifauna.

2.2.1 A Brief History of New Zealand

New Zealand is an archipelago in the southwest Pacific Ocean consisting of two major islands and numerous smaller ones with a total area a little less than 270,000 km^2 (Fig. 2.1). In the Pacific, only the Japanese archipelago has a greater area. The nearest substantial land area to New Zealand is the continent of Australia, ≥1520 km to the west across the Tasman Sea (Neall and Trewick 2008). Unlike most islands in the Pacific and elsewhere, which are predominantly of volcanic origin, New Zealand is formed primarily of continental crust (Trewick et al. 2007) and is described as a continental island (Cowie and Holland 2006). The formation of modern New Zealand began in the early Miocene when activity on the boundary between the Australian and Pacific continental plates became vigorous (26-22 Ma Campbell and Hutching 2007). This plate boundary developed through the largely submerged continent of Zealandia, such that New Zealand sits astride a major fault zone referred to as the alpine fault (Fig. 2.1). At that stage, most and possibly all of

Zealandia (Landis et al. 2008) was submerged beneath the sea, and even today some 90% of Zealandia is underwater (Mortimer 2004).

FIGURE 2.1—New Zealand and neighbouring islands (modified from Wallis and Trewick 2009). The approximate extent of the submerged continent of Zealandia is indicated with grey fill and dashed line. The approximate position of the Alpine Fault is indicated with a white line.

A spreading ridge producing new oceanic crust had developed by 83 Ma between Zealandia and the Australian/Antarctic part of Gondwana forming the Tasman Sea. This seafloor spreading continued until about 53 Ma but by 75 Ma Zealandia was clearly isolated from Australia and Antarctica, and by the start of the Tertiary (65 Ma) Zealandia was low lying and surrounded by deep seas (Graham 2008) (Fig. 2.2). A process of crustal spreading and thinning is invoked to explain the gradual submergence of Zealandia culminating in the late Oligocene (Campbell and Hutching 2007). Australia and South America finally separated from Antarctica in the late Eocene (~35 Ma) initiating the circumpolar current (Sanmartin and Ronquist 2004). Although biologists frequently refer to New Zealand separating from Australia, New Zealand has never had any direct physical contact with other parts of Gondwana. The question of whether there was any continually habitable land in the Zealandia region through the time of maximum inundation in the

Oligocene is unresolved (Campbell and Hutching 2007; Landis et al. 2008). Importantly, at about the size of India, Zealandia was large and there is little doubt that it began with a substantial terrestrial biota. This Zealandian biota is one possible source of lineages present in the modern NZ biota.

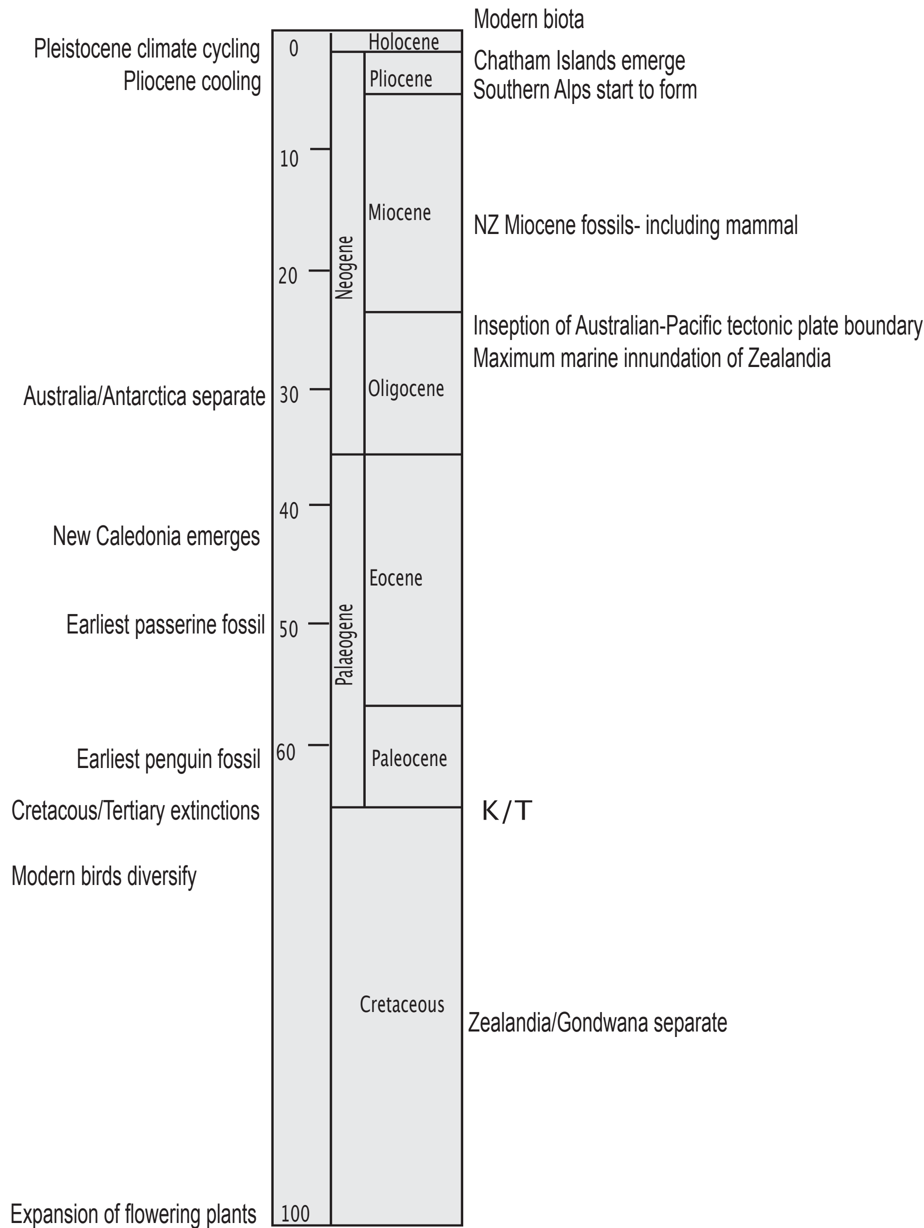

FIGURE 2.2—Geological timescale with some relevant ecological, evolutionary and geophysical events indicated. Values indicate millions of years before present (Ma).

However, almost nothing is known about the Zealandian biota beyond that inferred from the extant faunas of New Zealand and neighbouring islands. Fossils provide some insights into the terrestrial fauna, with evidence of Cretaceous dinosaurs (Long 1998), Miocene mammals, crocodiles and birds (Molnar and Pole 1997; Worthy et al. 2006; Worthy et al. 2007; Worthy et al. 2009), and Plio-Pleistocene

vertebrates (Worthy and Holdaway 2002). NZ fossils of terrestrial animals are relatively few and give only glimpses of the past, although they confirm that assemblages change through time, and can also contribute to dating of phylogeny if appropriate methods are used (see Discussion). However, spatially and temporarily isolated fossils fail to demonstrate continuity of occupation by particular bird lineages through time and provide little information on biogeographic process given the absence of comparable well-sampled fossil biotas for similar times elsewhere.

2.2.2 Overview of the New Zealand avifauna

Approximately 245 birds were breeding in New Zealand on human contact with the islands (estimated from late Holocene fossils and the extant fauna, Holdaway et al. 2001; Worthy and Holdaway 2002). This number is dependent to some degree on opinion about subspecies/species status and on the inclusion/exclusion of recent colonists and vagrants. For instance, Worthy & Holdaway (2002) exclude the self-introduced and highly successful silvereye *Zosterops lateralis* (Robertson et al. 2007b), but include as species many populations that are elsewhere treated as subspecies. Perceptions about composition are also influenced by the inclusion/exclusion of species found in the wider NZ archipelago (from subantarctic islands in the south to Norfolk Island in the north). Offshore islands tend to have their own representatives of many taxa due to allopatry (Fleming 1976); the Chatham Islands, for example, have in some taxonomic treatments more than 10 endemic land birds, each of which are almost certainly sister to a species on mainland New Zealand (Worthy & Holdaway 2002). The avifaunas of offshore islands are perhaps better treated as independent units for the purposes of inventory. A large proportion of the birds occupy freshwater, marine and coastal habitats, including seabirds (57 Procellariiformes, 14 Pelecaniformes, and 10 Sphenisciformes), shorebirds (32 Charadriiformes) and ducks (18 Anseriformes). If we exclude offshore island endemics and seabirds, but include North Island and South Island forms of many taxa as distinct species (following Worthy and Holdaway 2002) there were 95 mainland breeding bird species in New Zealand at human contact, including 29 passerines, 14 shorebirds, 13 ducks and 9 rails.

Analysis of New Zealand's prehuman (Holocene fossil bones in caves and swamps) and modern avifauna indicates two phases of extinction associated with human

colonisation. The first, Polynesian phase, starting around 1280AD (Wilmshurst *et al.* 2008), resulted in extinction of the larger and possibly least fecund taxa including many moa, swan, geese, giant coots and giant eagle (hereafter, see Table S2.1 for latin names). Habitat modification, and introduced predators such as dogs (kuri) and Pacific rats (kiore) must have had some impact, but direct hunting may have been sufficient to take populations below replacement (Turvey & Holdaway 2005). The contents of middens of indicate that hunting of large birds was an important component of Maori (or pre-Maori) culture at that time (Anderson 2003; King 2003). Settlement by Europeans in the 19th century corresponds with extinction of additional endemic birds including Huia, Quail, Laughing Owl, Bittern and Bush Wren. Most of these were forest inhabitants whose populations were depleted by large-scale habitat modification (e.g. pasturalisation), a plethora of mammalian predators (rats, mice, mustelids, cats, pigs, possums) and hunting with guns (see Worthy and Holdaway 2002; Tennyson and Martinson 2006). In total, about 31% of the estimated 245 species of birds breeding in the wider New Zealand region have become extinct since human contact. Of the non-marine taxa, the largest losses were among ducks, ratites, rails and passerines.

The modern element in the NZ avifauna arrived in two ways: intentional introductions mostly mediated by various Acclimatization Societies (Thomson 1922; McDowall 1994) and self introduction. Intentional introductions were primarily of north European species brought to New Zealand as part of a colonisation ethic that prevailed in the 1860s, or biological control (McDowall 1994; Walrond 2008). For birds, these include some 137 species of which about 20% are recorded as successfully 'acclimatizing' including Blackbird, Song Thrush, Rock Pigeon, Skylark, House Sparrow, finches and allies, Little Owl, Rook, Starling and Common Mynah (Veltman et al. 1996). Success of introduced birds was primarily dependent on the number of individuals introduced, and their migratory behaviour (Veltman et al. 1996). Introduction of species that were migratory in their natural range tended to be less successful, although many species introduced to mainland New Zealand quickly reached and colonised other parts of the archipelago over distances up to 900 km (Williams 1953). In general, introduced birds in New Zealand are mostly those of field and hedgerow, reflecting a bias in introduction effort.

A separate group of species have arrived by their own efforts traversing at least 1,500 km of ocean and founding viable NZ populations. Together with introduced species these provide valuable insight into dispersal, colonisation and speciation. The first arrival times of several species are known, including Silvereye 1856, Pacific Swallow 1958, Cattle Egret 1963, Masked Plover 1932, Australian Coot 1954, White Faced Heron 1940, and Black Fronted Dotterel 1954 (Falla 1953; Baker 1991; see also Robertson et al. 2007b). Additional species appear to be aviary escapees including parrots, doves and a kingfisher.

Among the most mobile species visiting NZ are seabirds, including terns, gannets, petrels, penguins and albatross which today nest mostly on offshore islands. Some nested on mainland NZ prior to human contact (Worthy and Holdaway 2002) and a few continue to do so today (e.g. Royal Albatross, Yellow-eyed Penguin, Hutton's Shearwater, Australasian Gannet). Migratory shore birds include the Godwit, Knot, and Ruddy Turnstone. Two species of cuckoo, *Chrysococcyx lucidus* and *Eudynamys taitensis* migrate to NZ to breed. The White Egret *Egretta alba* currently breeds in one part of NZ, but has sometimes reached other parts including Auckland, Campbell, Chatham and Maquarie islands (Falla 1953).

Inferring the timing and process of assembly prior to human contact is much more difficult, but given observation of the modern element it is clear that something can be gleaned from taxonomic distinctiveness. Indeed, evidence of hierachical taxonomic distinctiveness led Falla (1953) and others (e.g. Fleming 1962; Caughley 1964) to propose that colonisation of New Zealand occurred many times. Thus, NZ subspecies are likely to be the more recent colonists including Fantail, Banded Rail, Grey Duck, Spotless Crake, Little Black Shag, Pukeko, White Egret and Morepork. Species endemics such as New Zealand Robin and Tomtit are likely to be the products of earlier arrivals. In some cases it appears that particular lineages have reached New Zealand on more than one occasion and thus contributed several times to bird diversity. Fleming (1974) cited examples including Takahe/Pukeko, Shore Plover/Black Fronted Dotterel, NZ Robin/Tomtit, Black Stilt/Pied Stilt, Weka/Banded Rail, Laughing Owl/Morepork. Molecular and traditional taxonomy have confirmed many of Flemming's examples (see order-specfic sections below, and Table S2.2).

2.3 MOLECULAR VIEW OF THE NEW ZEALAND AVIFAUNA

2.3.1 Phylogeny of birds

Molecular phylogenetic methods applied to neutral DNA markers provide an estimate of the evolutionary relationships among living taxa, the rate of lineage formation, and the relative timing of phylogenetic branching events. What a molecular phylogeny cannot directly reveal is the pattern of ecological and morphological change along lineages. Although in some instances it is likely that phylogenetic lineage formation approximately coincides with 'species' formation, and might even result from speciation (as the mechanism of isolation), in many cases where species formation occurs in allopatry distinctive morphological and behavioural characteristics accumulate over time, after lineage formation. So, species labels and their attendant information such as morphology, behaviour, ecological preference and interactions with other species apply with confidence only to the tips of phylogenetic trees unless some additional information is available (e.g. detailed fossils). Nonetheless, phylogenetic trees provide invaluable hypotheses about branching, speciation and population structure.

The phylogeny of living birds is becoming clearer as more species are sequenced for both mitochondrial and nuclear DNA markers. The most basal split among living birds is between palaeognaths (ratites and tinamous) and neognaths (all other birds). The split of neognaths into Galloanserae (ducks and chickens) and Neoaves is also well established (e.g. Slack et al. 2007). Although many of the basal relationships within Neoaves are difficult to resolve due to short internal branches, there may be a division between orders comprising 'water birds' and 'land birds' (see Hackett et al. 2008 and Fig. 3). Accumulating evidence indicates many lineages of birds arose before and around the Cretaceous/Tertiary boundary (K/T, or Cretaceous/Palaeogene boundary) (e.g. Slack et al. 2006; Brown et al. 2008). This has contributed to a growing opinion that modern birds started diversifying before the global extinctions that mark the end of the Cretaceous (Cooper & Penny 1997). Some authors have proposed that significant modern radiations date to the early Cretaceous (Baker et al. 2007), but this is largely dependent on the choice of molecular calibrations.

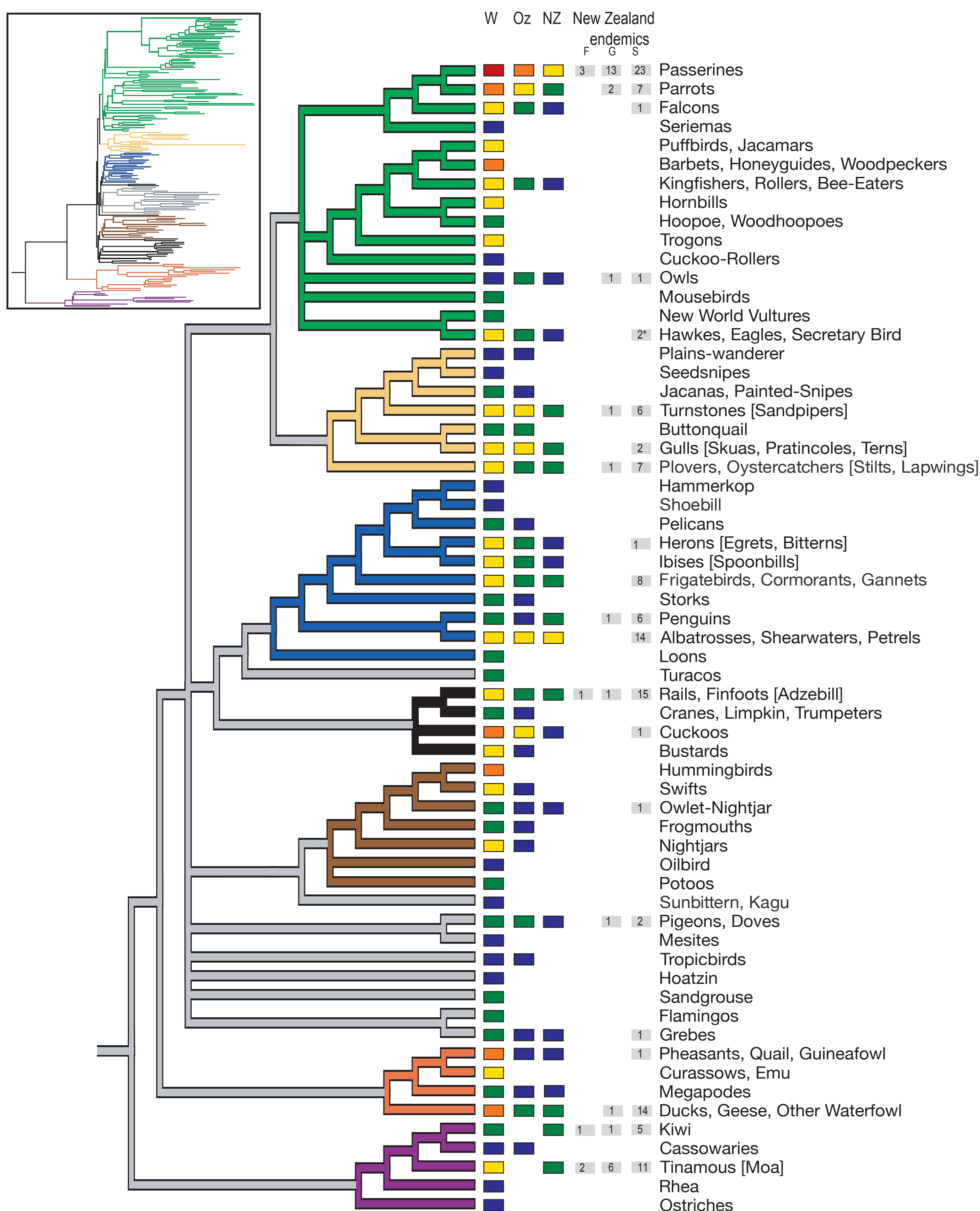

FIGURE 2.3—Phylogeny of birds based on multiple genes (redrawn from Hackett et al. 2008 Fig. 4). Branch colours are those given by Hackett et al. (2008) and indicate major clades supported: land birds (green), charadriiforms (yellow), water birds (blue), core gruiforms and cuckoos (black), apodiforms and caprimulgiforms (brown), galloanserae (orange), and paleognaths (purple). Bird groups not cited by Hackett et al. (2008) are placed on the tree [in square brackets], sometimes tentatively, according to other phylogenetic or taxonomic evidence (e.g. the New Zealand endemic Adzebills (*Aptornis*) are currently assigned to their own family, which is

usually placed tentatively within Rallidae). Coloured boxes indicate species diversity on a base 5 log-scale (1-4 blue, 5-24 green, 25-124 yellow, 125-624 orange, >3125 red), worldwide (W), in Australia (Oz) and in New Zealand (NZ) respectively. Statistics for species numbers came from Clements (2007) (World), Walter Boles pers. comm., (Australia), NZRBN (Robertson and Medway 2003) and Holdaway et al. (2001) (New Zealand). New Zealand taxa do not include Norfolk Island. Also indicated are estimates from current taxonomy of New Zealand endemics per bird group, at Family, Genus and Species level (note endemic species and genera are not always nested within endemics of higher taxonomic levels). New Zealand passerine data is as in Table 2.1. The inset tree provides an indication of molecular rate variation among lineages and short internodes at the base of the Neoaves and comes from Hackett et al. (2008 Fig. 2.3).

Despite substantial advances in avian phylogeny, considerable problems remain. Particularly, long unbroken branches leading to enigmatic monotypic species such as the hoatzin (Sorenson et al. 2003) and major radiations such as pigeons (Hackett et al. 2008), variation in rates of molecular evolution (Phillips et al. 2010), and short internal branches in the neoavian polytomy (Pratt et al. 2009), all cause difficulty in resolution. These uncertainties in avian systematics mean that some parts of deeper bird systematics are unresolved, and that some biogeographic inferences relating to inferences about older events must be treated with caution. There also remain a considerable number of unresolved issues with genus level taxonomy that are likely tractable with additional molecular data (for a pertinent discussion see Christidis and Boles 2008).

2.3.2 Birds in New Zealand

Here we review the available molecular phylogenetic analyses of bird groups that include data from NZ taxa. In the following section, and in Table S2.2, orders are ordered according to the recent work of Christidis and Boles (2008). Direct searches of the NCBI GenBank database (http://www.ncbi.nlm.nih.gov/) were designed to identify all published DNA sequence data for NZ birds, and from there we identified publications reporting analyses of those data (Table S2.2). We also used web searches and citation searches to track information on particular taxa. Important background to understanding the assembly of the NZ avifauna comes from deep phylogenetic analyses of bird orders and families undertaken with large data sets of mitochondrial and nuclear DNA sequences (see multiple orders in Table S2.2). Some of these do not include representatives of NZ species, whilst others do (Slack et al. 2006; Gibb et al. 2007; Morgan-Richards et al. 2008; Pratt et al. 2009), but all contribute to a framework for interpreting evolutionary history of

local lineages (Fig. 2.3). Many consist of local sampling of NZ taxa with overseas relatives as outgroups or actually use NZ species among an outgroup, such as *Acanthasitta* wren in analyses of Passerine evolution (e.g. Barker et al. 2004). All levels provide important insights into the timing and mode of assembly of the modern avifauna.

There are many gaps in available phylogenetic and taxonomic information even among the extant avifauna. Many extant taxa are so similar to overseas species that their phylogenetic placement with molecular data would probably yield few surprises. Nonetheless, additional phylogeographic analyses, especially in the wider geographic context of the Oceania region, would be of interest for many taxa (e.g. warbler, robin, owl and kingfisher). The situation is worse for the recently extinct species with the notable exception of moa, which have been intensively studied. Some extinct lineages have received little or no attention as yet, but would undoubtedly yield useful insights into the manner in which the fauna has assembled. Among these are *Aptornis*, a distinctive flightless gruiform currently placed in a monotypic family, and several less enigmatic taxa including goose, crow or raven, swan, Laughing Owl and Piopio.

Struthioniformes (ratites & tinamous)

The ratites and tinamous are sometimes placed in separate orders, partly on the basis of morphology associated with flight in the tinamous, but there has long been debate about this (e.g. Cracraft 1974; Parkes and Clark 1996). As a result, several molecular studies have constrained ratite monophyly treating tinamous as an outgroup (Cooper et al. 2001; Paton et al. 2002; Pereira and Baker 2006). Living tinamous are restricted to Central and South America whereas ratites live on all southern hemisphere continents. Ratites have usually been inferred to be a monophyletic group derived from a single flightless ancestor (Cracraft 1974). In New Zealand, ratites are represented by two families: moa (extinct Dinornithidae) and kiwi (Apterygidae). Despite their extinction, a large body of molecular data has been gathered for moa from late Pleistocene and Holocene fossil bones (see Table S2.2). This has contributed to revision of their taxonomic diversity including identification of extreme sexual dimorphism that had previously been one source of taxonomic oversplitting (Bunce et al. 2003; Huynen et al. 2003).

Early molecular phylogenetics revealed that the two NZ ratite families do not form a monophyletic group (Cooper et al. 1992; Cooper et al. 2001). This observation implied either that the moa, kiwi and other ratite lineages coexisted prior to separation of Zealandia, which would imply living bird groups evolved much earlier than is supported by other evidence, or that the kiwi ancestor dispersed to Zealandia/New Zealand over water (Cooper et al. 2001). Haddrath and Baker (2001) used whole mitochondrial DNA sequences to date ratite divergence times with an Australian fossil representing the ancestor of emu and cassowary. Their estimates of 62 Ma and 79 Ma respectively for the origin of kiwi and moa supported the notion that kiwi arrived after separation of Zealandia, but moa ancestors might have been present on Zealandia.

However, recent analyses using larger, multigene data sets and more stringent phylogenetic tools have revealed that tinamous and ratites are probably not monophyletic sister groups, but rather the clade of flying tinamous is nested within the ratites (Hackett et al. 2008; Harshman et al. 2008; Phillips et al. 2010). Furthermore, moa and tinamous appear to be sister lineages within ratites (Phillips et al. 2010). This has important evolutionary and biogeographic implications because it indicates that flight has been lost several times in the group rather than being the ancestral and defining condition for ratites (Harshman et al. 2008), and it weakens the logic favouring a vicariant explanation for their presence in New Zealand. Dispersal of ratite lineages becomes more plausible, and analogous to the numerous independent instances of flight loss evident in other groups incuding Rallidae (see Gruiformes below). Recent dating using a suite of calibration points indicates that *both* the kiwi and moa lineages postdate separation of Zealandia and the K/T boundary (ca. 60 Ma, Phillips et al. 2010).

Moa represent the most extensive and best-documented avian radiation in New Zealand. At one time some 27 species were recognised (Archey 1941; Oliver 1949), but this has recently been reduced to 11 taxa associated with a range of lowland and alpine environments (Worthy and Holdaway 2002). This radiation has been dated to late Miocene and Pliocene (10-4 Ma) using an 82 Ma date for the split between moa and an outgroup comprising other ratites and tinamous (Baker et al. 2005). However, the justification for this molecular clock calibration is questionable (see above and Discussion), so moa radiation is likely to be younger and primarily

associated with habitat heterogeneity that developed with the formation of the Southern Alps starting in the Pliocene (Bunce et al. 2009).

Analyses of mtDNA sequence data that included extinct populations reveals kiwi comprise five main lineages that are for the most part spatially partitioned (Shepherd and Lambert 2006). Burbidge et al. (2003) estimated that Brown and Spotted Kiwi diversity might date to the Miocene and splitting in the Brown Kiwi group to the Pliocene, based on the calibration of ratite phylogeny by Haddrath & Baker (2001).

Anseriformes (ducks & geese)

Diversity among ducks and their allies in New Zealand is almost entirely a product of numerous independent colonisations (Worthy and Olson 2002). These have resulted in species level endemism, and it is evident that at least some apparently endemic genera are more likely to be island forms of widespread genera (e.g. *Euryanas finschi* is on the basis of morphology better placed in *Chenonetta* with the Australian Wood Duck, (Worthy and Olson 2002). However, molecular data testing this and other hypotheses do not yet exist. The New Zealand Brown Teal *Anas aucklandica* includes closely related taxa on several offshore islands in a shallow, allopatric radiation allied to the Australian Chestnut Teal *A. castanea* and Grey Teal *A. gracilis* (Kennedy and Spencer 2000). The molecular phylogeny of *Anas,* though only partially resolved, indicates substantial dispersal of different dabbling duck lineages around the globe (Johnson and Sorenson 1999). The Black Duck *Anas superciliosa* appears to have colonised New Zealand twice in recent geological time and support for the subspecies status of the New Zealand Grey Duck population is very weak (Rhymer et al. 2004). The shallow phylogenetic history of *Anas* (Johnson and Sorenson 1998; Johnson and Sorenson 1999) is evident in their capacity to successfully hybridise, as documented in New Zealand between *A. superciliosa* and *A. platyrhynchos* (Gillespie 1985; Rhymer et al. 1994; Rhymer et al. 2004). The endemic Blue Duck *Hymenolaimus malacorhynchos* shows strong phylogeographic structure between North Island and South Island (Robertson et al. 2007a), but its phylogenetic placement is unresolved. The two extinct large geese *Cnemiornis* are thought to be closely allied to the Australian Cape Barren Goose *Cereopsis novaehollandiae*, an inference from morphology supported by a small amount of DNA data (Worthy et al. 1997). The merganser *Mergus australis* is

known from Holocene fossils in mainland New Zealand but the last population to go extinct was that on the Subantarctic Auckland Islands. Other mergansers occur in Europe, Asia, North America and Brazil but are not otherwise known in Australasia.

Galliformes (game birds)

Morphology indicates that the single extinct native galliform, the quail *Coturnix novaezelandiae* was closely related to the Australian Stubble Quail *C. pectoralis*), a relationship confirmed by molecular analysis of the group (Seabrook-Davison et al. 2009). There is a closer relationship between *C. novaezelandiae* and *C. pectoralis*, which are estimated to have diverged in the Pliocene, than there is between the two Australian quails *C. pectoralis* and *C. ypsilophora* (Seabrook-Davison et al. 2009). The latter species has since been introduced to New Zealand (Thomson 1922; Robertson et al. 2007b).

Podicipediformes (grebes)

New Zealand has one endemic grebe, the New Zealand Dabchick *Poliocephalus rufopectus*. The genus includes the Hoary-headed Grebe *P. poliocpehalus,* which is a vagrant from Australia, but there has been no molecular analysis of these taxa. Two other widespread grebes, *Podiceps cristatus* and *Tachybaptus novaehollandiae*, are also recorded in New Zealand. Analysis of complete mitochondrial DNA sequences of these latter species has confirmed the sister relationship of grebes and flamingos (Morgan-Richards et al. 2008) proposed by van Tuinen et al. (2001).

Columbiformes (pigeons)

The New Zealand native pigeon, Kereru *Hemiphaga novaeseelandiae* forms a clade with the monotypic Australian Topknot Pigeon *Lophalaimus antarcticus* and the small genus *Gymnophaps* distributed across Indonesia, New Guinea and Solomon Islands (Pereira et al. 2007; Gibb and Penny 2010). A molecular analysis with multiple fossil calibrations estimated that these lineages diverged in the early to mid Miocene (Pereira et al. 2007). If correct this provides a maximum age for the arrival of the *Hemiphaga* lineage in New Zealand. There is little taxonomic diversity in *Hemiphaga* with an extinct subspecies *H. n. spadicaea* known from Norfolk Island and a sister species *H. chathamensis* from the Chatham Islands.

These islands are no more than 4 Ma, but the distribution appears to be of even more recent origin. Control region sequences reveal low genetic distances and a sister relationship between *H. chathamensis* and the two *H. novaeseelandiae* subspecies consistent with structuring during the Pleistocene (Goldberg et al. 2010).

Apodiformes (owlet nightjar)

The New Zealand Owlet Nightjar *Aegotheles novaezealandiae* is extinct, but is considered to have been most similar to the New Caledonia species *A. savesi*, (Tennyson and Martinson 2006). A molecular analysis of this small genus which has its centre of diversity in Papua New Guinea, indicates that the New Zealand and New Caledonian species are sister to species in New Guinea/Australia (Dumbacher et al. 2003). The comparatively large size, long legs and short wings of *A. novaezealandiae* and *A. savesi* imply a more terrestrial rather than arboreal habit (Dumbacher et al. 2003), which could have been derived independently after island colonisation.

Procellariiformes (tubenosed birds)

Tube-nosed seabirds are among the most mobile birds in the world, although they are also frequently philopatric (Shaffer et al. 2006). A large proportion of the World's Procellariiformes (~45%) take advantage of coastal and island nesting sites in the NZ region where they contribute a large proportion of all avian species (~22%). It is not surprising that most molecular data relating to NZ seabirds deals with population structuring across the oceans (Abbott and Double 2003b) and in many cases focuses on identifying the provenance of birds killed as fisheries by-catch (Abbott et al. 2006). Species that breed in NZ waters tend to have genetically close sister taxa on nearby islands outside the region. For example White-capped Albatross *Diomedea* that nest on the sub-Antarctic Auckland islands are closely related to Shy Albatross on small islands near Tasmania, Australia (Abbott and Double 2003a). Shallow divergences exist between *Pelecanoides georgicus* populations in southern New Zealand and the Indian ocean (Paterson et al. 2000), and among Black-browed *D. melanophrys* and Wandering *D. exulans* albatrosses from southern New Zealand, South America and the Kerguelen islands (Alderman et al. 2005). The Chatham Island endemic Taiko *Pterodroma magentae* forages as far as Chile, and is closely related to other Southern Ocean petrels including *P.*

macroptera and *P. lessonii* (Lawrence et al. 2008). Shearwaters in the NZ region are members of an Australasian clade (Austin 1996; Austin et al. 2004).

Sphenisciformes (penguins)

Penguins have a long history in the region, and early penguin fossils are recorded from South Island, New Zealand (Jadwiszczak 2009). Fossils demonstrate that the penguin lineage extends at least to the early Tertiary and this has provided a valuable molecular clock calibration (Slack et al. 2006; Ho and Phillips 2009). The group may have originated in the Australasian/Antarctic region, and its current distribution continues to be the southern hemisphere. However, extant diversity is much younger (Baker et al. 2006; Ksepka et al. 2006). Penguins are of course highly mobile, seeking land only for reproduction where they seek nesting sites with low predator risk. Speciation appears to have resulted from numerous dispersal events presumably enhanced by subsequent philopatric behaviour, and the modern NZ assemblage reflects this.

The lineage leading to Yellow-eyed Penguin *Megadyptes antipodes* may have evolved in the Miocene (Baker et al. 2006) but has arrived on the NZ mainland, from Subantarctic islands, since the arrival of humans. This appears to have happened after the extinction of a previously undescribed closely related species, *M. waitaha* (Boessenkool et al. 2008). This may indicate repeated waves of colonisation, which is a pattern also evident in Little Blue Penguin *Eudyptula minor*. The diversity of *E. minor* indicates more than one exchange between New Zealand and Australia since the start of the Pleistocene (Banks et al. 2002; Overeem et al. 2008; Peucker et al. 2009) and lineage sharing with Australia might also be explained by extinction and recolonisation. Sister to *Megadyptye*s is the *Eudyptes* group (including Fiordland Crested and Snares penguins), which diversified in the Pliocene (Baker et al. 2006) and are distributed around the Southern Ocean (Bertelli and Giannini 2005).

Pelecaniformes (gannets & cormorants)

Pelecaniform birds are speciose on New Zealand's extensive coastline. About 25% of the world's pelecaniform species breed in the area, and most are cormorants. Cormorants have colonised New Zealand several times, in some cases resulting in island endemics (Kennedy et al. 2000). For example, a group including the Stewart,

Chatham and Campbell Island Shags are closely related to each other, and among their closest relatives are South American cormorants (Kennedy et al. 2000). The Australasian Gannet *Morus serrator* nests in New Zealand and Australia. *Morus serrator* is closely related to the South African Cape Gannet *Morus capensis* and diversification of these gannets is thought to have occurred in the Plio-Pleistocene (Friesen and Anderson 1997).

Ciconiiformes (herons & storks)

The NZ avifauna includes a few widespread Australasian heron and egret species (*Ardea spp., Egretta spp.*). The New Zealand Little Bittern *Ixobrychus novaezelandiae* became extinct at human contact, however the Australasian Bittern *Botaurus poiciloptilus* and vagrant Little Bittern *I. minutus* are now present. Molecular phylogenetics of Ardeidae using cytochrome *b* sequences (Sheldon et al. 2000) yielded comparable results to an earlier DNA-DNA hybridization study (Sheldon 1987). *Ixobrychus* is sister to *Botaurus* (Sheldon et al. 2000) and both genera have worldwide distributions.

Falconiformes (falcons) & Accipitriformes (hawks & eagles)

Molecular phylogenetics has been instrumental in revealing that the falcons, hawks and eagles are not a monophyletic group (Fain and Houde 2004; Ericson et al. 2006; Gibb et al. 2007; Hackett et al. 2008; Morgan-Richards et al. 2008; but see Pratt et al. 2009). The commonest raptor in New Zealand is the Australasian Swamp Harrier *Circus approximans,* which probably arrived sometime after the first humans (Worthy and Holdaway 2002). A much larger and now extinct endemic NZ species (*C. eylesi*) is known only from Holocene fossils. Also represented only by Holocene fossil is Haast's eagle *Harpagornis moorei* an unusually large species with a wingspan up to 3m, whose talon marks have been found in fossil moa hipbones, which survived into Polynesian time. However, Haast's Eagle has been shown to be closely related to the Australian Little Eagle *Hieraaetus morphnoides* and appears to have evolved after colonisation in the Pleistocene (Bunce et al. 2005). Haast's Eagle provides an excellent example of rapid and substantial morphological change that can occur following colonisation of an island. If as has been proposed, *Hierraaetus* is reclassified with *Aquila*, then Haast's Eagle should be *Aquila moorei* (Lerner and Mindell 2005). The New Zealand Falcon *Falco novaeseelandiae*, which also reached the Chatham Islands

(Worthy and Holdaway 2002), is a member of a genus with worldwide distribution and relatively shallow phylogenetic history (Griffiths 1999). Six species of *Falco* occur in Australia, and the Nankeen Falcon *F. cenchroides* has also reached New Zealand at various times (Falla 1953), although the globally distributed and habitat tolerant Peregrine Falcon (*F. peregrinus*) has not been recorded (Caughley 1964).

Gruiformes (rails & cranes)

As for many islands in the Pacific (Steadman 2006), rails are, in terms of diversity, an important component of the NZ avifauna. In all cases, close allies of NZ taxa are known in Australia and the western Pacific. Crakes (*Porzana*) are represented on the mainland and offshore islands by two species that are also found in Australia. Coots (*Fulica*), swamphens (*Porphyrio*), brown rails (*Gallirallus*) and waterhen (*Gallinula*) are represented by recently dispersed and vagrant species (e.g. *Fulica atra*, *Porphyrio porphyio*, *Gallirallus philippensis*) in addition to endemic and often flightless derivatives of earlier but not ancient colonisations (Trewick 1996; Trewick 1997a; Trewick 1997b). Strikingly, although they are often skulking and secretive ground birds, many volant rails are very effective dispersers, such that most Pacific islands have or had prior to arrival of humans one or more often endemic rail species (Steadman 2006). Though effective at moving between islands (e.g. Kirchman 2009), natural selection and/or isolation has been sufficient to yield morphologically distinct endemics. In rails, flight is primarily a predator escape response, but on many oceanic islands that lack mammalian predators, flight appears to be less relevant (Livezey 2003). Rails demonstrate an important feature in the assembly of the NZ avifauna, whereby repeated colonisation rather than insular radiation, is responsible for the accumulation of species diversity. Indeed, flying species of *Fulica*, *Gallirallus* and *Porphyrio* appear to be recent colonists of New Zealand.

The genus *Gallinula* is represented by two species shared with Australia and also an extinct endemic *G. hodgenorum*. The European Coot *Fulica atra* arrived in New Zealand in the 1950's almost certainly from Australia, but at least one earlier colonisation yielded large endemic *Fulica*. The extant population of the flying species *Porphyrio porphyrio* is not separable genetically from populations elsewhere in the region, and judging by fossils and midden remains appears to have arrived after Polynesian colonisation (Trewick and Worthy 2001). However, earlier

and separate colonisations resulted in the evolution of two large flightless species (*P. hochstetteri* and *P. manteli*) (Trewick 1996), and morphologically similar species are evident in fossil remains on other islands (Steadman 2006). An analogous history explains the (extinct) flightless relatives of the ubiquitous Banded Rail *G. philippensis*. Although data for the NZ endemic *Capellirallus karamu* (=*G. karamu*) are not yet available, other species including *G. australis* (New Zealand), *G. dieffenbachii* (Chatham), *G. modestus* (Chatham) and *Diaphorapteryx hawkinsi* (=*G. hawkinsi*) (Chatham) are evidently products of separate colonisations by flying ancestors most probably similar to *G. philippensis* (Trewick 1997a, SAT unpublished data). The widespread flying species (*G. philippensis*) comprises low genetic diversity with shallow ancestry consistent with recent range expansion across Oceania including New Zealand. In contrast, flightless island endemics including the New Zealand Weka (*G. australis*) arise from older nodes in the *Gallirallus* tree (Trewick 1997a; Kirchman 2009).

The extinct robust flightless Adzebill *Aptornis* is thought to be a gruiform and has sometimes been interpreted as a relative of the Rallidae. Adaptations to terrestrial habitat and putatively predatory habit make placement of *Aptornis* on morphological evidence difficult. Analysis of a small amount of molecular data (Houde et al. 1997) indicates that *Aptornis* is closer to the rails than to a group of taxa including the New Caledonian Kagu *Rhynochetos jubata* with which it has also been allied on the basis of bone characteristics (Livezey 1998).

Charadriiformes (shorebirds)

Shorebirds comprise the second largest order in New Zealand, although a large proportion are nonbreeding migrant or vagrant species. The shorebirds may have started diversifying in the Cretaceous according to Baker et al. (2007), who estimate that 14 modern lineages extend to before the K/T boundary. Nevertheless, genus level diversification is much more recent, indicating that native shorebirds must have colonised the NZ archipelago in the Tertiary or later. Oystercatchers *Haematopus* and stilts *Himantopus* both have global distributions, although genus level molecular analyses are not yet available. Banks and Paterson (2007), found low levels of sequence variation among three NZ species of *Haematopus*, including a form on the Chatham Islands. *Himantopus* in New Zealand appear to have resulted from two separate arrivals but the taxa are so closely related that they

hybridise (Wallis 1999). Speciation of both these genera in the Australia-New Zealand region might be latest Miocene in age (Thomas et al. 2004). The genus *Charadrius,* which may have originated in South America (Joseph et al. 1999), and diversified in the early Miocene (Thomas et al. 2004) has a global distribution including a number of Australasian species. However, the Australasian taxa do not appear to be monophyletic indicating more than one colonisation. In addition to vagrants, the NZ fauna includes four endemic *Charadrius* species allied to Australian ones. The distinctive Wrybill is the sole member of an endemic New Zealand genus, which is sister to an Australian species (*Peltahyas australis* = *Charadrius australis*). Despite current taxonomic distinction, the *Anarhynchus/Peltahyas* clade is estimated to have originated as recently as the mid Pliocene (Thomas et al. 2004).

A number of gulls and terns are rare visitors to New Zealand. However, the masked gulls are an Australasian clade (*Chroicocephalus*) with New Zealand Black-billed Gull sister to the New Zealand Red-billed Gull and Australian Silver Gull. This is consistent with two colonisation events; either both to New Zealand or one each way (Given et al. 2005; Pons et al. 2005). The masked gull group appears to have radiated in the Pleistocene and much of the taxonomic diversity may have emerged in the last 600 Ky (Thomas et al. 2004; Given et al. 2005). The terns are also a young radiation and the *Chlidonias* group, which includes the NZ endemic *C. albostriatus* probably radiated during the Pliocene (Bridge et al. 2005).

Snipe (*Coenocorypha*) have reached New Zealand and offshore islands (two taxa seem to have existed on the Chatham Islands- Tennyson and Martinson 2006), and vagrants of the Japanese species (*Gallinago*) also reach New Zealand. New Zealand *Coenocorypha* on islands show strong population genetic structure (Baker et al. 2009). The relationship between these two genera is not well resolved but the clade that includes *Gallinago* and *Coenocorypha* may be early Miocene in age (Thomas et al. 2004).

Psitaciformes (parrots)

Three species (two genera) are endemic to New Zealand; the forest Kaka *Nestor meridionalis*, alpine Kea *N. notabilis*, and flightless, lek-breeding Kakapo *Strigops habroptilus*. A third species of *Nestor* (*N. productus*) is extinct on Norfolk Island

(750 km NE of New Zealand), and a fourth may have existed on the Chatham Islands (800 km east) (Millener 1999). Strikingly, NZ parrots *Strigops* and *Nestor* appear to form a clade that is sister to all other extant parrots (de Kloet and de Kloet 2005) and this has been used to calibrate parrot phylogenies on the assumption of vicariant isolation of New Zealand (Tavares et al. 2006; Tokita et al. 2007; Wright et al. 2008). Wright et al. (2008) also inferred a “basal split” between *Nestor* and *Strigops* implying an early separation of these two lineages. However, although the sister relationship of NZ large parrots to other parrots may be consistent with a vicariant history, molecular phylogenetics of Neoaves dated using penguin and galliform fossils (Pratt et al. 2009) found that *Strigops* probably arose after the K/T (i.e. < 65 Ma). The use of an 82 Ma age for the common ancestor of all parrots needs to be reassessed (see Discussion).

The *Cyanoramphus* parakeet radiation has most of its diversity in the NZ region with lineages on the main islands and Chathams, Subantarctics, Norfolk Island and New Caledonia. The New Caledonian lineage appears to be sister to a group including those in New Zealand and Norfolk Island (Boon et al. 2001a; Chambers et al. 2001). Molecular analysis with other representatives of the platycercine parakeets including the New Caledonian *Eunymphicus* supports the New Zealand/ New Caledonia relationship (Boon et al. 2008). Genetic exchange among *Cyanoramphus* populations is problematic for taxonomy and conservation, but is evidence of a youthful radiation associated with allopatry on islands (Chan et al. 2006). Boon et al. (2001) estimated the split between New Zealand and New Caledonian *Cyanoramphus* as mid-late Pleistocene in age. A clade including *Cyanoramphus, Eunymphicus* and other Australasian Platycercini started diversifying in the early Miocene if tree calibration with *Strigops* is reliable (Wright et al. 2008).

Cuculiformes (cuckoos)

The Long-tailed Cuckoo *Eudynamys taitensis* (= *Urodynamis taitensis*) is not closely related to other *Eudynamys*, but instead may group with another Australasian taxon *Scythrops novaehollandiae* (Sorenson and Payne 2005). The Long-Tailed Cuckoo breeds only in New Zealand but winters on islands of the western Pacific. In contrast there is shallow genetic diversity among Shining Cuckoo *Chrysococcyx lucidus* populations breeding in Australia, New Caledonia

and New Zealand, and this species belongs to a radiation distributed from Africa through southern Asia to Australasia. *C. lucidus* is sister to a tropical species *C. ruficollis* breeding in Papua New Guinea and Indonesia (Sorenson and Payne 2005).

Strigiformes (owls)

Of the two native owls, the extinct monotypic endemic *Sceloglaux albifacies* is of uncertain affinity, although Falla (1953) inferred it to be a close relative of *Ninox*. The extant native Owl *Ninox novaeseelandiae* also occurs in Australia along with three other *Ninox* species. Distributed from Asia to Australia, the species-rich *Ninox* has received little molecular phylogenetic treatment (Norman et al. 1998; Olsen et al. 2002). New Zealand *N. novaeseelandiae* are closely related to other populations of this species, though slightly closer to the Norfolk Island rather than Tasmanian population (Norman et al. 1998).

Coraciiformes (kingfishers)

The kingfishers are represented in New Zealand by a single native species, the Sacred Kingfisher *Halcyon sanctus* (= *Todiramphus sanctus*). The same species occurs in Indonesia, throughout Australia and on many other Pacific Islands, indicating recent dispersal. Low kingfisher diversity on islands contrasts with the high taxonomic and phylogenetic diversity of the group in the Australian region (Moyle 2006). Phylogeography indicates that this diversity and endemism is not due to an ancient history in the region but rather recent radiation and colonisation from Asia (Moyle 2006). Kingfishers therefore provide a nice counterpoint to assumptions about the relationship between centres of diversity and centres of origin.

Passeriformes (perching birds)

The native passerine fauna of New Zealand and offshore islands comprised between 27 and 38 species (depending on taxonomic treatment, see Table 2.1). These represent 13 families spread across global passerine diversity (Fig. 2.4). Two of these are NZ endemic families (Acanthisittidae with six species, Callaeidae with three or five species), and a third (Notiomystidae) has been proposed for the monotypic genus *Notiomystis* (Driskell et al. 2007). At other taxonomic levels, 13 of 21 genera and 23 of 27 species are endemic to New Zealand.

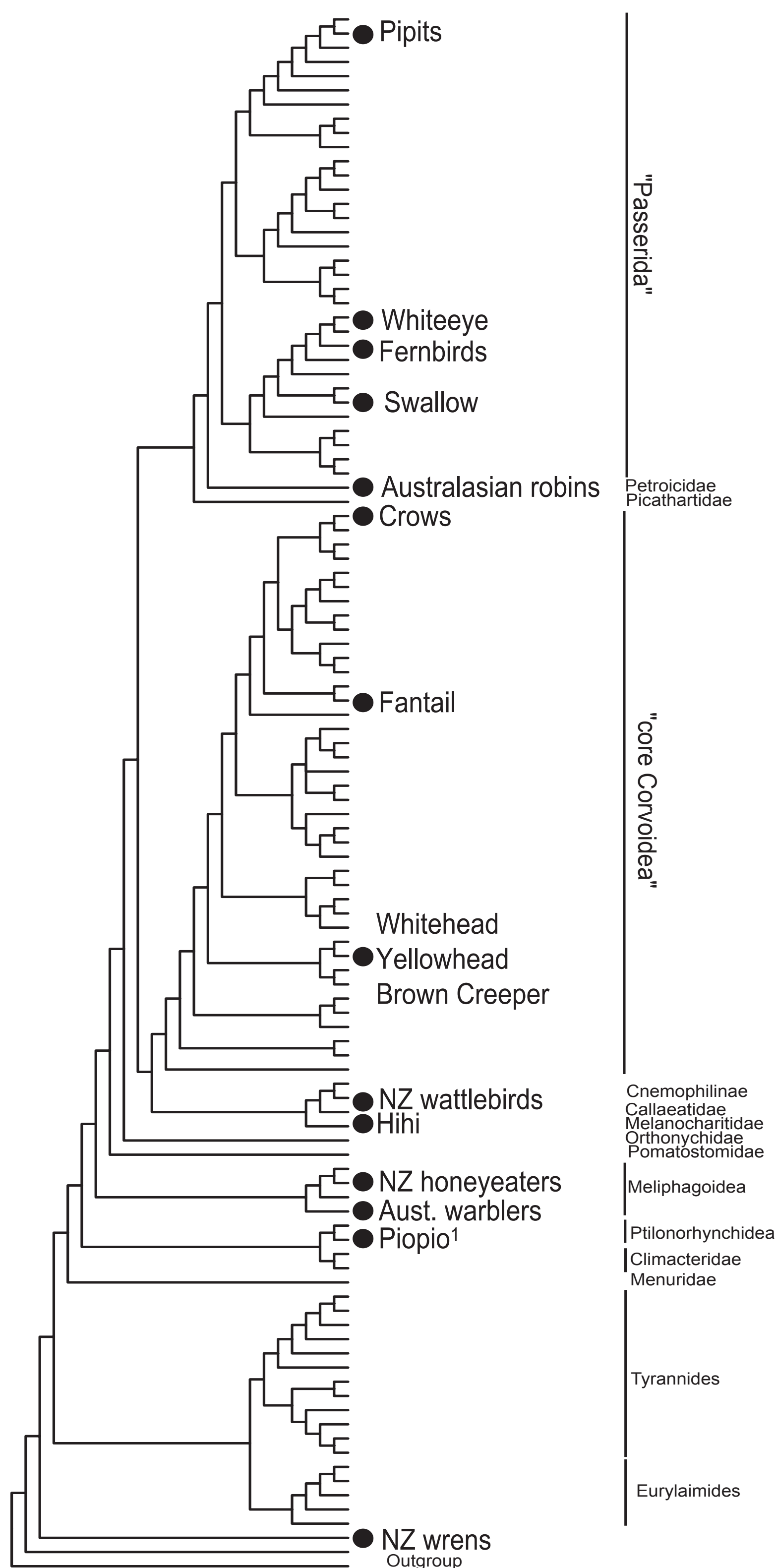

FIGURE 2.4—Phylogeny of passerine diversity (from Barker et al. 2004), indicating likely placement of native New Zealand taxa. Placements are based either on published molecular phylogenetic data for NZ taxa or close relatives of NZ taxa, or on current taxonomy. [1]Placement of Piopio based on Christidis *et al.* (1996) but unpublished data (GCG) indicates this is incorrect.

TABLE 2.1—New Zealand native passerine diversity and endemicity (in bold) across taxonomic levels, based on the NZRBN (Robertson and Medway 2003)

Family	Genus	Species	Subspecies	English Name	Location	Status
ACANTHISITTIDAE : New Zealand Wrens						
	Acanthisitta	***chloris***	***granti***	NI Rifleman	NI, L&GBar	
			chloris	SI Rifleman	SI, St, Cod	
	Xenicus	***longipes***	***stokesii***	NI Bush Wren	NI	Extinct
			longipes	SI Wren	SI	Extinct?
			variabilis	Stead's Bush Wren	St	Extinct?
			gilviventris*	Rock Wren	SI	
	Traversia	***lyalli***		Stephens I Wren	Step	Extinct
	Pachyplichas	***yaldwyni***		SI Stout-legged Wren	SI	Extinct
		jagmi ^		NI Stout-legged Wren	NI	Extinct
	Dendroscansor	***decurvirostris***		Long-billed Wren	SI	Extinct
HIRUNDINIDAE : Swallows and Martins						
	Hirundo	*neoxena*		Welcome Swallow	NI, SI Chat	
MOTACILLIDAE : Wagtails and Pipits						
	Anthus	*novaeseelandiae*	***novaeseelandiae***	NZ Pipit	NI, SI	
			chathamensis*	Chatham I Pipit	Chat	
			aucklandicus*	Auckland I Pipit	Auk	
			steindachneri	Antipodes I Pipit	Ant	
SYLVIIDAE : Old World Warblers						
	[1]*Megalurus*	***punctatus***	***vealeae***	NI Fernbird	NI	
			punctata	SI Fernbird	SI	
			stewartiana	Stewart I Fernbird	St	
			wilsoni	Codfish I Fernbird	Cod	
			caudata*	Snares I Fernbird	Snares	
		rufescens		Chatham I Fernbird	Chat	Extinct
PACHYCEPHALIDAE : Whistlers and Allies						
	Mohoua	***albicilla***		Whitehead	L&GBar, Kapiti	
		ochrocephala		Yellowhead	SI, St	
		novaeseelandiae		Brown Creeper	SI, St	
ACANTHIZIDAE : Australasian Warblers						
	Gerygone	***igata***		Grey Warbler	NI, SI, offsh	
		albofrontata		Chatham I Warbler	Chat	
MONARCHIDAE : Monarch Flycatchers						
	Rhipidura	*fulginosa*	***placabilis***	NI Fantail	NI	
			fulginosa	SI Fantail	SI, St	
			penita	Chatham I Fantail	Chat	
PETROICIDAE : Australasian Robins						
	Petroica	***macrocephala***	***toitoi****	NI Tomtit	NI, offsh	
			macrocephala	SI Tomtit	SI, St	
			chathamensis	Chatham I Tomtit	Chat	

Family	*Genus*	*Species*	Subspecies	English Name	Location	Status
			dannefaerdi*	Snares I Tomtit	Snares	
			marrineri*	Auckland I Tomtit	Auck	
		australis	***longipes****	NI Robin	NI, L&GBar, Kapiti	
			australis	SI Robin	SI	
			rakiura	Stewart I Robin	St	
			traversi*	Black Robin	Chat	
ZOSTEROPIDAE : White-eyes						
	Zosterops	*lateralis*	*lateralis*	Silvereye	NI, SI Chat	
MELIPHAGIDAE : Honeyeaters						
	[2]***Notiomystis***	***cincta***		Hihi	LBar	
	Anthornis	***melanura***	***obscura***	Three Kings Bellbird	Three Kings	
			oneho	Poor Knights Bellbird	Poor Knights	
			melanura	Bellbird	NI, SI, St	
			melanocephala*	Chatham I Bellbird	Chat	Extinct
	Prosthemadera	***novaeseelandiae***	***novaeseelandiae***	Tui	NI, SI, St	
			chathamensis	Chatham I Tui	Chat	
CALLAEATIDAE : New Zealand Wattlebirds						
	Callaeas	***cinerea***	***wilsoni****	NI Kokako	NI, GBar	
			cinerea	SI Kokako	SI, St	Extinct?
	Philesturnus	***carunculatus***	***rufusater****	NI Saddleback	NI, offsh	
			carunculatus	SI Saddleback	SI, St	
	Heteralocha	***acutirostris***		Huia	NI	Extinct
PARADISAEIDAE : Birds-of-Paradise, Bower-birds and Piopios						
	[3]***Turnagra***	***capensis***	***tanagra****	NI Piopio	NI	Extinct?
			capensis	SI Piopio	SI, offsh	Extinct?
CORVIDAE : Crows and Jays						
	Corvus	***moriorum***	**C***	Extinct NZ Crow	NI, SI, St, Chat	Extinct

Annotations indicate alternative taxonomic treatment: * and ^ are taxa treated as species or subspecies respectively by Worthy & Holdaway (2002), [1]*Megalurus* is otherwise placed in endemic genus *Bowdleria*, [2]*Notiomystis* was recently proposed as sole representative of New Zealand family Notiomystidae (Driskell et al. 2007), [3]*Turnagra* has been placed in Paradisaeidae (Christidis et al. 1996) but new data indicate otherwise (GCG. in prep), C *Corvus* from Chathams, but a second undescribed *Corvus* is known from New Zealand mainland. Location codes are in bold as follows: **N**orth **I**sland, **S**outh **I**sland, **Cod**fish I., **Cha**tham Is., **St**ewart I., **offsh**ore, **Kapiti**, **Snares**, **Ant**ipodes Is., **Auck**land Is., **L**ittle **Bar**rier I., **G**reat **Bar**rier I., **Steph**ens I.

The New Zealand wrens (Acanthisittidae) comprise five genera, of which three are extinct. *Xenicus* is now represented by the Rock Wren *X. gilviventris* with few individuals in alpine habitats of South Island, while the Rifleman *Acanthisitta chloris* is still fairly abundant in native forests. No thorough analysis of the phylogenetic relationships of the living and extinct species has been completed although Cooper (1994) found genetic distance data might be consistent with placement of *Xenicus* within *Acanthisitta*.

Petroicid robins predominate in the Austral-Papuan region and on Pacific Islands, and include two or three NZ species (*Petroica*). Molecular analysis is consistent with the hypothesis that New Zealand was colonised twice by *Petroica* (as proposed by Fleming 1950), and populations have managed to reach offshore islands (Miller and Lambert 2006). A melanic robin lineage *P. traversi* on the Chatham Islands appears to be sister to mainland Tomtit *P. megacephala* rather than Robin *P. australis* (Miller and Lambert 2006). Fine-scale mtDNA sequence variation shows strong spatial partitioning between North Island and South island *P. australis*, but more mixing among *P. megacephala* (Miller and Lambert 2006).

The New Zealand wattlebirds Callaeatidae comprise three endemic species in three genera that form a monophyletic group when compared to a range of other passerines (Shepherd and Lambert 2007). The Huia *Heteralocha acutirostris* is extinct as are most populations of Kokako *Callaeas cinerea* and Saddleback *Philesturnus carunculatus*. Wattlebirds appear to share a common ancestor with Australian Cnemopholine Birds of Paradise (Barker et al. 2004) around 39-34 Ma (Shepherd and Lambert 2007). The Hihi *Notiomystis*, formerly placed with the honeyeaters (Meliphagidae), is sister to the New Zealand wattlebirds (Ewen et al. 2006; Driskell et al. 2007). The split between Saddleback and Hihi is estimated at 39-28 Ma, which could indicate existence of the callaeid/notiomystid lineage in New Zealand since the Oligocene (Driskell et al. 2007). However, this and the study of Shepherd & Lambert (2007) used the assumed separation at 82 Ma of Acanthisittidae from all other passerines by continental drift for date calibration, and even if correct the presence of these birds in New Zealand now does not preclude their possible existence elsewhere in the past. A monotypic family has been proposed for Hihi mostly on the grounds of missing morphological characteristics typical of wattlebirds (Driskell et al. 2007). No near relatives of Hihi are available for comparison and so far only one New Zealand wattlebird has been

used to represent this group in phylogenetic analysis. A low level of DNA sequence variation exists between North Island and South Island forms of Kokako (Murphy et al. 2006).

The diverse Austral-Papuan honeyeaters, Meliphagidae, are represented in New Zealand by two species (Tui *Prosthemadera novaeseelandiae* and Bellbird *Anthornis melanura*). Tui and Bellbird lie within one of four large honeyeater clades (Driskell & Christidis 2004), and may have diverged from other members of the clade ~19-31 Ma, and from one another in the late Pliocene (Driskell and Christidis 2004; Driskell et al. 2007).

An analysis by Christidis et al. (1996) indicated that the New Zealand Thrush or Piopio *Turnagra capensis* might be closer to the bowerbirds than birds of Paradise, however node support was weak, and taxon sampling insufficient to test alternative placements.

DNA-DNA hybridisation indicates that the New Zealand Fernbird *Bowdleria punctatus* is sister to the Australian Grassbird *Megulurus gramineus* and thus belongs in *Megalurus* (Sibley and Ahlquist 1987), a relationship that had previously been proposed by Oliver (1955), although Olson (1990) has argued for retention of *Bowdleria*. Molecular data indicate that the three New Zealand putative pachycephalines, Yellowhead *Mohoua ochrocephala*, Whitehead *M. albicilla* and Brown Creeper *M. novaeseelandiae* do not belong to that group but are part of the so-called core Corvoidea that comprises many Austral-Papuan taxa (Norman et al. 2009).

Two recent arrivals to New Zealand have ranges beyond New Zealand and sister species in the region. The swallow *Hirundo neoxena* is closely related to the Pacific Swallow *H. tahitica* and has been considered a race of this species (Sheldon et al. 2005). The silvereye is one of three Australian races, and a member of a speciose genus distributed in Africa, Asia and Australasia (Clegg et al. 2002). Diversity in this genus over this geographic scale reflects dispersal and isolation during and since the Pleistocene (Warren et al. 2006). The New Zealand Pipit *Anthus novaeseelandiae* is the same species in Australia and belongs to a clade distributed in Africa, Asia and Australasia (Voelker 1999). Nevertheless, it has been suggested that allozyme data indicate NZ populations are currently reproductively isolated

from one another and might constitute distinct species (Foggo et al. 1997). The warblers *Gerygone* and fantail *Rhipidura* in New Zealand belong to speciose groups distributed in Australia and the Pacific, and both have close allies in Australia. Molecular analysis has resolved six main clades among *Rhipidura* species (Nyári et al. 2009). Limited genetic divergence in the clade that includes the NZ native, *R. fulginosa*, indicates recent radiation and dispersal (Nyári et al. 2009).

2.4 DISCUSSION

If New Zealand had been biologically isolated since the mid/late Cretaceous when modern birds were first evolving, we might expect it to have a very distinctive avifauna comprising phylogenetically deep endemic radiations of early bird lineages diverged from their cousins elsewhere in Gondwana. In fact the NZ avifauna is far from monophyletic and few families or genera are endemic (Fig. 2.3). None of the orders that are reasonably speciose in New Zealand (e.g. ratites, parrots, shorebirds, passerines) form monophyletic groups. Evidently, 'places' in New Zealand have repeatedly 'been seized on by intruders' (Darwin 1859 p81-82).

The NZ avifauna is thus composite in nature; its ancestry mixed in space and time. As Robert Falla (1953) and others observed, a large proportion of NZ species have their closest relatives in Australia. This is true for shallow species level (e.g. fantail, robin) and deeper family level (e.g. kiwi) relationships. Others are associated with Pacific or northern regions (Fleming 1979). That is not to say that the NZ avifauna is Australian in character; it is quite distinct in terms of behavioural, morphological, and ecological characteristics, family composition, and species diversity, which is generally lower than Australia but notably higher among ratites and penguins (Fig. 2.3). Falla (1953) thought the composition of the NZ avifauna typical of an oceanic island and this view is largely confirmed by molecular phylogenetic evidence. In addition, molecular data have demonstrated instances of unexpected taxonomic distinctiveness (e.g. Hihi), and shallower than expected connections with non-NZ taxa (e.g. Haast's Eagle), and have opened an important window on recently extinct species and populations (e.g. moa). Finally, molecular evidence has been instrumental in revealing the fantastic way that natural selection has yielded, often in a short time frame, extraordinary changes in behaviour and form (e.g. flightless rails, giant eagle, nocturnal parrot).

Though the affinity with the avifauna of Australia is consistent with a vicariant history involving Zealandian continental drift, the depth and extent of endemicity is not. The phylogenetic composition of the NZ avifauna shows that it is drawn from throughout the diversity of modern birds (Fig. 2.3) and it has been drawn repeatedly, not at a single event. In this phylogenetic sense, New Zealand's avifauna is a subset of Australia's, which is in turn a subset of global diversity (Fig. 2.3). The composite nature of the avifauna is evident at all levels, with for instance, representation of multiple families across passerine diversity (Barker et al. 2002) (Fig. 2.4). Some of these passerine families are Australasian (cis-Wallacean of Barker et al. 2002), but others are not (i.e. they occur on both sides of Wallace's Line). The observation that the Australasian passerine lineages tend to arise basally in phylogenetic reconstructions indicates that the Passeriformes might have originated in this southern region (Barker et al. 2002; Edwards and Boles 2002; Ericson et al. 2002; Barker et al. 2004). This is supported by presence in Australia of the oldest known songbirds fossils from the Eocene (Boles 1995, 1997) whereas northern hemisphere songbird fossils are younger (Cooper and Penny 1997).

2.4.1 Accumulation & Radiation

Island biogeography theory (MacArthur and Wilson 1967) assumes that island biotas develop primarily through colonisation balanced by extinction, but colonisers may modify their environment, influence evolution of competitors, and themselves evolve under natural selection or genetic drift. This combination of processes is evident among the NZ birds, and is exemplified by the rail fauna, which appears to have accumulated sequentially from separate founding populations (Trewick 1997a; Crisp 2008). Rates of dispersal by flying rails are apparently high, but this inference is derived from analysis of current distributions and relationships rather than direct observation of immigration. Similar rates of dispersal may exist in other taxa but be less apparent in the absence of rapid reproductive isolation of successive colonisations (e.g. by adaptive shifts associated with flightlessness). The rails also exemplify the influence of directional asymmetry in colonisation (Cook and Crisp 2005a). A speciose New Zealand clade with an Australian or cosmopolitan taxon at nested within it, might be seen as indicating that New Zealand is source of dispersal as suggested for other organisms (e.g. Winkworth et al. 2002). However, the same pattern is obtained if the likelihood of dispersal and colonisation in different directions is not equal. In the case of rails where NZ

endemics are flightless, the likelihood of dispersal out of rather than into New Zealand is intuitively considerably lower (Crisp 2008).

Flightlessness is not confined to rails, but is a feature of the NZ avifauna, as on many other islands (Roots 2006; Steadman 2006). Many NZ birds show a general naïveté of ground predators and some evidence of adaptation to avoid diurnal aerial predators (e.g. nocturnal Kakapo and kiwi). Rather than a dichotomy in flight ability, flightlessness is a continuum from moa, which lack any wing structure, through kiwi, some rails, and Kakapo, to Kokako, wrens and Owlet-nightjar that use(d) their wings to varying degrees. Many species lack defensive behaviour against terrestrial predators and are especially susceptible when nesting (e.g. Kaka, Saddleback, Hihi). Rather than being evidence of ancient isolation, these traits are likely to have been derived fairly recently in NZ birds as Zealandia has not always lacked a mammal fauna; fossils indicate that terrestrial mammals, whether descended from Zealandian stock or later colonists were present in Miocene New Zealand (Worthy et al. 2006). Modern experience in conserving NZ birds (King 1985) make it abundantly clear that species such as the large, flightless, ground-dwelling, strong-smelling and behaviourally defenceless Kakapo are unlikely to have evolved in the presence of ground hunting mammals. Rather than being evidence of an ancient origin, the peculiarities of Kakapo and other endemic birds are derived and relate to their recent evolutionary environment. The NZ wrens included the only confirmed flightless passerines in the world and they might be taken to represent an ancient biota. However, flightlessness has evolved uniquely in many bird groups on islands and sometimes in short geological time, including Henderson Island Fruit Dove, Seychelles Warbler, Australian Scrub-birds, New Caledonian Kagu, South American Tapaculos and Mauritius Dodo, whilst others such as the rails have representatives in many places (Roots 2006). Even among ratites the evidence for independent evolution of flightlessness is now strong (Harshman et al. 2008; Phillips et al. 2010).

The capacity for successful colonisation in the past can be understood to some degree by scrutiny of recent colonisation, which in turn provides some information on dispersal. Successful long distance dispersal that fails to result in colonisation will usually go unobserved, although the incidence of vagrants may provide some indication of actual dispersal rates that have not yet resulted in colonisation (Falla 1953). The ecological characteristics of the birds that have naturally colonised New

Zealand in recent times are varied. Recent dispersal yielding viable populations has occurred in small arboreal (e.g. Silvereyes ~11 g) and large terrestrial species (e.g. Pukeko ~1000 g). Neither the small body mass of Silvereyes, nor the ungainly flight and skulking habit of Pukeko indicate these species as candidates for distance travel. In contrast, some habitat tolerant and powerful fliers are notable by their absence (e.g. Peregrine Falcon, Caughley 1964). Thus, while dispersal might be a stochastic process, success in colonisation may be strongly influenced by availability of niche space (i.e. colonisation opportunity).

Only a few New Zealand lineages are represented by genuine radiations. Dominant among these were the moa, in which substantial variation in size within and among species may have facilitated sympatry (Baker et al. 2005; Turvey and Holdaway 2005; Wood et al. 2008). Wattlebirds comprised three species, each of which represent different ecotypes, whereas honeyeaters, large parrots, and wrens were represented by two, three and five species, respectively. However, five kiwi species appear to have occupied different geographic areas. Many taxa are represented by North and South Island forms, and sometimes also additional forms on offshore islands. Although the North Island and South Island are close to one another today, in the early Pleistocene and before, they were separated by a larger seaway that may have contributed to population isolation. Overall diversity is greatest when we consider the islands of the New Zealand region together, as many small islands provided sufficient isolation to allow the evolution of distinct morphological forms (e.g. *Nestor, Hemiphaga, Cyanoramphus* - Fleming, 1976). Sometimes these island populations are sufficiently distinct to be deemed separate species. In all, dispersal has been a key process in establishing isolation and in some instances dispersal is yielding a reticulate pattern of evolution (e.g. *Cyanoramphus* parakeets on the Chatham Islands Boon et al. 2001b, Chan et al. 2006). Fleming (1976) inferred dispersal of some bird taxa even further including to New Caledonia (e.g. *Cyanoramphus*). Although inferring direction of exchange from phylogenetic and distributional evidence can be misleading (Cook and Crisp 2005a; Goldberg et al. 2010), it is probable that *Cyanoramphus* came to New Zealand from New Caledonia (Boon et al. 2008).

Some New Zealand birds have distinctive morphology and behaviour (e.g. flightless Stout-Legged Wren, flightless Kakapo parrot, extreme sexual dimorphism of beak in Huia, small omnivorous kiwi, Giant Eagle), which is a common feature of island

species (Quammen 1997). However, there is no phylogenetic evidence to contradict the hypothesis that these forms have evolved in relatively recent times in response to novel ecological opportunities.

2.4.2 A Zealandian element?

New Zealand is essentially the product of tectonic activity since the Miocene (Fig. 2.2). It is just one part of a Zealandian system, which had a long 60 million year biotic history of its own after separation from Gondwana, but before New Zealand as we know it began to form. Part of the modern NZ fauna could have been derived from the Zealandian biota carried on that continental fragment. It is not known whether all parts of Zealandia were submerged prior to emergence of the current islands. Evidence for continuous existence of substantial land in New Zealand's position is equivocal, and current knowledge of Zealandian geology indicates that persistence of terrestrial life may be best envisaged in the context of a largely tectonic, ephemeral archipelago. Accumulating evidence indicates presence of Miocene islands north of New Zealand (Meffre et al. 2006), New Caledonia emergence in Eocene (Paris 1981), New Zealand emergence during Miocene, Chatham Islands emergence in late Pliocene plus volcanic edifices including Lord Howe Island and Norfolk Island to the north and west, and the islands on the Campbell plateau to the south (Campbell and Hutching 2007) (Figs. 1 & 2).

A putative Zealandian element might include currently unplaced lineages (e.g. *Aptornis*), and taxa on long branches with nodes basal to their respective group (e.g. large parrots, New Zealand wrens, New Zealand Pigeon). But a paradox exists, as at least some of these lineages include populations on offshore islands that indicate a retained capacity for dispersal and colonisation over substantial areas of ocean. For example parrots *Nestor* and pigeons *Hemiphaga* are also represented on offshore islands >600 km away (Norfolk I. and Chatham I.) that are just a few million years old and were never connected directly to mainland New Zealand (Fleming 1976; Goldberg et al. 2010). Their ancestors are likely to have been at least as mobile and therefore able to move in or out of Zealandia or New Zealand in earlier times.

The endemic NZ bird families might represent lineages with old ancestry in the region, although this need not be as old as Zealandia, nor geographically fixed in

"New Zealand". Identifying the mechanisms that have resulted in endemism of deep taxonomic lineages in New Zealand is difficult and perhaps impossible. Absence today of a given lineage (e.g. bird family) elsewhere is not proof of its absence in the past. Phylogenetically old lineages might not have evolved in the places in which their descendants are found today (Fig. 2.5). Furthermore, there is less chance of current distribution confounding inference of lineage origin when we are considering those that are speciose (see Cook and Crisp 2005b; Crisp and Cook 2005).

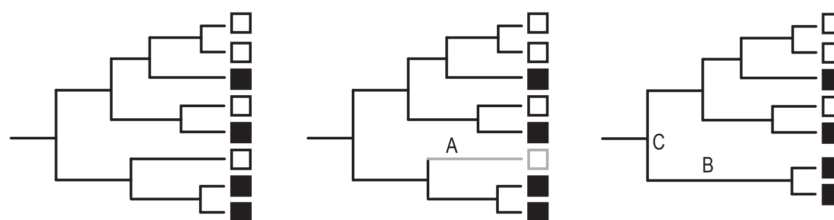

FIGURE 2.5—The effect of taxon sampling on tree interpretation. The rooted tree on the left depicts the relationships of a hypothetical set of taxa distributed in two geographic areas (black or white squares). Consider the situation if a single lineage (A) in the tree was missing because it has become extinct or was not sampled for some other reason (grey branch and box in centre tree). On the right is the tree that would in this case be inferred. Notice that branch B now extends to the most basal node (C) in the rooted tree and lineage B now appears to be sister to all other taxa. If the black and white areas had been separated by some vicariant event at a known time, molecular clock calibration that applied this date to node C would overestimate the age of the entire clade. It could also be inferred using parsimony that the geographic area (or morphology or behaviour) depicted by black boxes was the more likely ancestral location (or morphology or behaviour) of the entire group of taxa.

Endemic bird families exist on other landmasses of greater and lesser size and age than New Zealand (e.g. Caribbean Palmchat, New Caledonian Kagu, Greater Antilles todies, Hawaiian honeycreepers, mesites and ground-rollers of Madagascar, Dodo/Solitaire of Mascarene Islands). Endemic families might indicate local survival of relatively old lineages, or rapid evolution of novel and unusual traits (resulting in taxonomic splitting). For example, molecular phylogenetics has shown that the distinctive Raphidae (Dodo and Solitaire) endemic to the Mascarene Islands are actually embedded within the geographically widespread pigeon family, Columbidae (Pereira et al. 2007). Estimates using molecular phylogenies of Columbiformes calibrated with multiple fossils indicate the Raphidae lineage originated about 33 Ma, which is some 20 million years before the islands they lived on came into existence. In this case, morphology

misleads taxonomy, but the lineage cannot have originated where it was found. The same situation could easily apply to endemic families in New Zealand.

2.4.3 Disharmony & Extinction

A striking feature of the avifauna, also seen in the NZ biota as a whole, is its patchiness (Fig. 2.2), sometimes described as disharmonic (Carlquist 1974) or ill-balanced (Falla 1953). For the birds this is particularly evident in low species diversity, i.e. the tendency for genera and even deeper taxonomic levels to be represented by single or few species. In Carlquist's (1974) terms the NZ avifauna has a higher representation of dispersing species than would be expected from an old 'harmonic' biota. On human contact, the New Zealand region including offshore islands had some 245 species (Worthy and Holdaway 2002). The modern list extends to 305 (with introductions), compared to UK (599), New Guinea (764), Hawaii (303) and Galapagos (147) (Clements 2007 with updates from Avibase, 2009). New Zealand bird diversity is low considering its large area and wide latitudinal range (Neall and Trewick 2008). This condition has frequently been ascribed to a long period of isolation (e.g. Baker 1991). However, while geophysical isolation of the land is evident, it is obvious that even the modern NZ biota is not isolated; some species migrate in and out of New Zealand repeatedly in their lifetimes (e.g. cuckoo, godwit, albatross). Regardless, there is no obvious direct correlation between long physical isolation and low species diversity (Goldberg et al. 2008). Examples of the reverse are abundant, and are well recognised as having provided a vital stimulus for Charles Darwin's thesis on evolution (Trewick and Cowie 2008). While some islands, and especially the more distantly isolated ones may lack some lineages, they often include radiations within the lineages that are present (e.g. Hawaiian honeycreepers, Galapagos finches). Island biogeography theory (MacArthur and Wilson 1967) leads us to predict that low species diversity is best explained by extreme youth of island (insufficient time), small area and/or low habitat diversity (insufficient niche space).

The combination of broad phylogenetic representation (Fig. 2.3) and low species diversity in the NZ avifauna could be explained by a combination of fairly extensive extinction and stochastic arrival of lineages. Long naked branches in phylogenetic trees are indicative of extinction, but do not tell us where extinction or survival has taken place in space or time (Fig. 2.5). Recent colonists may be the sole

representatives of lineages that are speciose elsewhere, and are unlikely to have had time to diversify within New Zealand (e.g. swallow, Silvereye). Earlier colonists might have diversified after arrival either by adaptive radiation or allopatric isolation or might have been supplemented by additional colonisations as in the rails. The modern observation of numerous insular taxa on offshore islands around New Zealand indicates that we can expect equivalent diversification to have arisen in this manner in the past. However, phylogenetic and taxonomic information show that in general inter-island diversity within lineages is shallow, implying it is short lived. In some cases this is easily understood as the islands themselves are young, but at least two main islands of New Zealand have probably existed in some form at least since the Miocene.

Substantial climate cooling in Plio-Pleistocene time may also have reduced diversity, but fossils of terrestrial avian taxa that might provide evidence for early Pleistocene extinction are few. In contrast, extinction of Miocene plants and animals is evident from the fossil record (Lee et al. 2001, Worthy et al. 2006). Oligocene land area in the New Zealand region was small and possibly negligible. We can therefore infer that only a small fraction of the Zealandian biota could have survived, as the effects of small population size environmental stochasticity on species survival are well recognised (MacArthur and Wilson 1967; Saether et al. 2005). Not only does geology show most of the terrestrial environment of Zealandia was lost in the Oligocene (Graham 2008; Landis et al. 2008), molecular data indicate that only few extant taxa could plausibly represent lineages with long term existence in the region (Goldberg et al. 2008; Wallis and Trewick 2009). In addition, phylogenetic evidence so far fails to show consistent structure indicative of a shared evolutionary history of the avifauna on the remnants of Zealandia (principally New Zealand and New Caledonia).

Delving further back into the history of Zealandia and its biota, it is recognised that many species and higher taxonomic groups were lost globally around the K/T boundary (Fortey 1999). At that time (65 Ma) Zealandia was already a shrinking and low lying landmass (Graham 2008) and local extinction of angiosperms and most gymnosperms is evident (Vajda et al. 2001). Forests were destroyed globally and any fauna associated with forests and indeed other vegetation types must have been similarly devastated (e.g. Labandeira et al. 2002). While it is obvious that globally, many lineages persisted through the K/T, and these include one lineage of

dinosaurs (Neorthines) (Penny and Phillips 2004; Padian and de Ricqlès 2009) we can assume that many did so only locally and then expanded their ranges to occupy vacant ecological space, rather than surviving in all areas. Whilst it is possible that some lineages might have survived in Zealandia it would be a parochial bias to assume all did without evidence for that. Even in the absence of explicit details about biotic turnover it has to be recognised that, in the long history of Zealandia, extinction must have been a major force in shaping the recent biota.

2.4.4 Dating diversification

With the application of the molecular clock (Bromham and Penny 2003) it is possible to estimate the timing of lineage formation, although the reliability of this approach is dependent on the consistency of rates of molecular evolution among lineages or their appropriate accommodation, and the suitability and reliability of clock calibrations. In studies of bird evolution, molecular clock calibrations have used two main tools, fossils and vicariant events. The value of fossils is dependent on their accurate dating, identification as representing ancestors of a given lineage and completeness of the fossil record. In recent years, some important bird fossils have been found to supplement fossils of other vertebrates in molecular calibrations. Among them, the dating of the stem Anseriform at 66 Ma (*Vegavis* fossil, Clarke *et al.* 2005) and stem penguins (New Zealand, Waimanu fossils 64-61 Ma Slack *et al.* 2006), both lie close to the K/T boundary and are useful in dating the diversification of lineages in Neoaves (Ho and Phillips 2009). Fossils usually represent minimum ages for origin of lineages, as they may form anytime after but not before a taxon evolves. Similarly, the timing of vicariance events needs to be accurately assessed with a rigorous hypothesis developed for their influence on the lineages in question. The assumptions made to justify use of vicariance events in clock calculations are rarely well expressed and can lead to circular reasoning; an assumption of vicariance cannot be used to date the origin of a lineage that is then used to demonstrate a role of vicariance in lineage formation (Waters and Craw 2006). Molecular clock analyses have to be able to accommodate variation in rates of molecular evolution among lineages and modelling that meaningfully expresses confidence about the age and affinity of fossils or vicariance events used for calibration (Ho and Phillips 2009). The use of point values for calibrations can be especially misleading and implies a level of certainty about age and taxonomic placement of fossils that may be rarely if ever warranted. Studies that use point

calibrations or fixed rates of molecular evolution can yield point estimates of node age lacking confidence intervals and thus giving a false impression of certainty (Ho and Phillips 2009). What are needed when applying molecular clock calibrations are “rigidly defined areas of doubt and uncertainty” (Adams 1979).

Although vicariance and dispersal are often seen as polar opposites, especially in the context of New Zealand biogeography, they are both features that influence the general process of population exchange or gene flow (Trewick and Morgan-Richards 2009a). Principally because of the combination of geological and biological characteristics typical of both islands and continents, neither vicariance (via continental drift) nor long distance dispersal can be excluded *a priori* as potential influences on New Zealand biology (Daugherty et al. 1993; Goldberg et al. 2008; Trewick and Morgan-Richards 2009b). Thus, both biogeographic interpretations have been in favour at one time or another and the conundrum of New Zealand has been central to the consideration of many biogeographers and led one to suggest: ‘Explain New Zealand and the world falls into place around it’ (Nelson 1975). Molecular analyses have contributed to both sides of the debate, but paradoxically, a growing recognition of the considerable importance of dispersal in founding isolated populations of many kinds of organisms and thus biotas (Sanmartin and Ronquist 2004; Cowie and Holland 2006), has occurred at the same time as some phylogeneticists have turned to vicariance events (for want of alternatives) to date speciation and diversification.

Thus, the earliest time of separation of Zealandia from Gondwana has been used as a calibration point for avian phylogeny. A date of 82 Ma is widely applied in studies of birds including parrots (Tavares et al. 2006; Wright et al. 2008), passerines (Barker et al. 2002; Ericson et al. 2002; Irestedt et al. 2008), moa (Cooper et al. 2001; Baker et al. 2005) and honeyeaters (Driskell et al. 2007). This is despite the fact that many of these same analyses have required that dispersal events are hypothesised to reconcile phylogeny and geography. For example Barker et al. (2004) infer and date 10 major dispersal events for passerine taxa, largely on the assumption that one group, ancestors of the New Zealand wrens, never dispersed. Similarly Tavares et al. (2006) note the value of an independent geological calibration point in allowing rates of molecular evolution to be estimated. While true in concept, this approach relies on an assumption that the geological event is both correctly dated and correctly inferred as the cause of vicariance of some

parrots (in this case). Again, confidence in the correctness of this assumption is undermined because the same study infers dispersal for other parrots (Tavares et al. 2006).

If dispersal happens sometimes or at least has to be inferred sometimes, how do we know it did not happen following the vicariant event in question? A case in point is that of the New Zealand wrens (represented by the extant rifleman, *Acanthisitta*). Some analyses of DNA sequences indicate that *Acanthisitta* might be the phylogenetic sister of all other passerines (Barker et al. 2002; Ericson et al. 2002; Ericson et al. 2003; Barker et al. 2004), but (e.g. Slack *et al*. 2006). Because the Acanthisittidae are endemic to New Zealand, *Acanthisitta* potentially provides a useful tool for dating the evolutionary history of the largest living bird radiation (Passeriformes), but only if there is confidence in the history of the *Acanthisitta* lineage.

The New Zealand wren lineage either (1) evolved in Gondwana before separation of Zealandia, (2) evolved in Zealandia following separation, or (3) arrived in Zealandia after separation. As scenarios 1 and 3 both require subsequent extinction of the lineage everywhere else except New Zealand they appear to be equally plausible. Only scenario 2 provides a strong evolutionary basis for using the vicariance event to date evolution, because in this scenario, vicariance is the driver of evolution so that phylogenetic and geographic splitting were contemporaneous. Thus only scenario 2 provides a robust molecular clock calibration, whereas scenarios 1 and 3 could result in underestimation or overestimation respectively of the timing of avian diversification. The evidence for the age of the New Zealand wren lineage and correlation of its origin with Zealandia's formation is circumstantial, so it provides at best only tentative calibration.

It is increasingly evident that modern birds have a history dating at least to the late Cretaceous, prior to K/T boundary (Cooper and Penny 1997; Penny and Phillips 2004). Just how far back remains unresolved and is dependent on the quality of fossils and stringency of molecular calibration studies (for a discussion of fossil history see Padian and de Ricqlès 2009). When considering the assembly of the NZ avifauna where there exists the possibility of long isolation, the earlier diversification of modern birds opens the possibility that a number of distinct lineages were isolated by continental-drift vicariance. If we are to accept a stronger

influence of vicariance in the past and dispersal more recently, then we are assuming that birds of that early period had different dispersal capacities than today. Ericson et al. (2003 page 6) suggest that 'the basal members of the major clades of passerines are feebled-winged groups'[sic] and thus less capable of the type of dispersal that would result in significant biogeographic change. But there appears to be no strong evidence that older passerines were less capable fliers than modern ones. There is little logic in inferring weak flight as a condition among passerines or other modern birds, when it is likely or at least arguable that the achievement of effective flight was one of the key features enabling their success and radiation. In some cases at least, inference of feeble flight in 'old' lineages is confounded by what is probably the secondary evolution of that condition in a number of Zealandian lineages (e.g. New Zealand moa, wattle birds, wrens and New Caledonian Kagu) as reduction in flight capability is well documented in a broad diversity of avian lineages, especially on islands (Roots 2006). As already noted, large skulking rails and small passerines alike have shown themselves equal to the challenge of trans-oceanic dispersal.

2.5 CONCLUSION

Falla's proposal in 1953 of a series of invasions assumed dispersal was the driver of regional diversity. However, the radical shift in understanding of geophysical history of the earth that took place soon after, introduced a very plausible alternative. Continental drift provided a mechanism to sunder populations and migrate entire biotas, which led to the proposal of a vicariant origin of New Zealand's biota. The development even more recently of molecular phylogenetic tools provided the means to test the relative importance of these processes. The result has been confirmation of dispersal as the major process in the assembly of the avifauna and the biota in general (Wallis and Trewick 2009). However, the story is not simple and its interpretation is hindered by the gaps in understanding of the geological history of the Zealandian islands including New Zealand.

The New Zealand avifauna is most closely related to that of Australia, both in terms of phylogenetic relationships and overall composition. Many taxa are recent arrivals, whilst others appear to have evolved from lineages reaching New Zealand earlier in its history. Disharmony appears to reflect both extinction and stochastic

colonisation with most diversity derived by accumulation, sometimes accompanied by evolution of flightlessness and other distinctive adaptations, rather than in situ radiation. Some species found today only in New Zealand appear to represent lineages that originated quite early in the evolution of modern birds. Some of these lineages might be remnants of a Zealandian avifauna isolated by vicariance, but might equally have arrived in Zealandia (and latterly New Zealand) after geological separation from Gondwana. What is striking and very clear is that the vast majority of New Zealand bird lineages do not owe their presence merely to ancient vicariant isolation. Instead the development of the avifauna reflects a much more convoluted, dynamic and interesting evolutionary odyssey.

In addition to filling in missing phylogenetic information on placement of New Zealand species among the extant global avifauna and efforts to conserve New Zealand's distinctive taxa and assemblages, we predict that fruitful research about the way biotas develop and how susceptible they are to environmental change will come from the application of combined ecological and genomic methods to the evolution of New Zealand's birds, old and new.

2.6 ACKNOWLEDGEMENTS

Thanks to Walter Boles who provided a valuable summary of Australia's avifauna. Thanks to Shannon Hackett, Rebecca Kimball, and Sushma Reddy (Hackett et al. 2008), and Keith Barker (Barker et al. 2004) for assistance with figures originally published in their papers. Thanks to Mary Morgan-Richards who provided valuable comments on the manuscript at various stages, and to Leo Joseph who invited SAT to present a paper at the 2007 Australasian Ornithological Congress in Perth.

2.7 REFERENCES

Abbott, C., M. Double, R. Gales, G. Baker, A. Lashko, C. Robertson, and P. Ryan. 2006. Molecular provenance analysis for Shy and White-capped Albatrosses killed by fisheries interactions in Australia, New Zealand, and South Africa. Conserv Genet 7:531-542.

Abbott, C. L., and M. C. Double. 2003a. Genetic structure, conservation genetics and evidence of speciation by range expansion in Shy and White-capped Albatrosses. Mol Ecol 12:2953.

Abbott, C. L., and M. C. Double. 2003b. Phylogeography of Shy and White-capped Albatrosses inferred from mitochondrial DNA sequences: implications for population history and taxonomy. Mol Ecol 12:2747.

Adams, D. 1979. The Hitchhiker's Guide to the Galaxy. Pan Books.

Alderman, R., M. C. Double, J. Valencia, and R. P. Gales. 2005. Genetic affinities of newly sampled populations of Wandering and Black-browed Albatross. Emu 105:169-179.

Anderson, A. 2003. Prodigious Birds: Moas and Moa-hunting in New Zealand. Cambridge University Press.

Archey, A. 1941. The moa: a study of the Dinornithiformes. . Bulletin of the Auckland Institute and Museum, No. 1. Unity Press, Auckland.

Austin, J. 1996. Molecular phylogenetics of *Puffinus* shearwaters: preliminary evidence from mitochondrial cytochrome *b* gene sequences. Mol Phylogenet Evol 6:77-88.

Austin, J. J., V. Bretagnolle, and E. Pasquet. 2004. A global molecular phylogeny of the small *Puffinus* shearwaters and implications for systematics of the little-audubon's shearwater complex. Auk 121:847-864.

Baker, A. 1991. A review of New Zealand ornithology. Cur Biol 8:1-67.

Baker, A., S. Pereira, O. Haddrath, and K. Edge. 2006. Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. Proc R Soc B 273:11-17.

Baker, A. J., L. J. Huynen, O. Haddrath, C. D. Millar, and D. M. Lambert. 2005. Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: The giant moas of New Zealand. P Natl Acad Sci USA 102:8257-8262.

Baker, A. J., S. Pereira, and T. A. Paton. 2007. Phylogenetic relationships and divergence times of Charadriiformes genera: multigene evidence for the Cretaceous origin of at least 14 clades of shorebirds. Biol Lett 3:205-209.

Banks, J. C., A. D. Mitchell, J. R. Waas, and A. M. Paterson. 2002. An unexpected pattern of molecular divergence within the Little Blue Penguin *Eudyptula minor* complex. Notornis 49:29-33.

Banks, J. C., and A. M. Paterson. 2007. A preliminary study of the genetic differences in New Zealand oystercatcher species. N Z J Zool 34:141-144.

Barker, F. K., G. F. Barrowclough, and J. G. Groth. 2002. A phylogenetic hypothesis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. Proc R Soc Lond B Biol Sci 269:295-308.

Barker, F. K., A. Cibois, P. Schikler, J. Feinstein, and J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. P Natl Acad Sci USA 101:11040-11045.

Bertelli, S., and N. P. Giannini. 2005. A phylogeny of extant penguins (Aves: Sphenisciformes) combining morphology and mitochondrial sequences. Cladistics 21:209-239.

Boessenkool, S., J. J. Austin, T. H. Worthy, P. Scofield, A. Cooper, P. J. Seddon, and J. M. Waters. 2008. Relict or colonizer? Extinction and range expansion of penguins in southern New Zealand. Proc R Soc B.

Boles, W. E. 1995. The world's oldest songbird. Nature 374:21-22.

Boles, W. E. 1997. Fossil songbirds (Passeriformes) from the Early Eocene of Australia. Emu 97:43-50.

Boon, W. M., C. H. Daugherty, and G. K. Chambers. 2001a. The Norfolk Island Green Parrot and New Caledonian Red-crowned Parakeet are distinct species. Emu 101:113-121.

Boon, W. M., J. C. Kearvell, C. H. Daugherty, and G. K. Chambers. 2001b. Molecular systematics and conservation of the Kakariki (*Cyanoramphus spp.*). Sci for Cons 176:46.

Boon, W. M., O. Robinet, N. J. Rawlence, V. Bretagnolle, J. A. Norman, L. Christidis, and G. K. Chambers. 2008. Morphological, behavioural and genetic differentiation within the Horned Parakeet *Eunymphicus cornutus* and its affinities to *Cyanoramphus* and *Prosopeia*. Emu 108:251-260.

Bridge, E. S., A. W. Jones, and A. J. Baker. 2005. A phylogenetic framework for the terns (Sternini) inferred from mtDNA sequences: implications for taxonomy and plumage evolution. Mol Phylogenet Evol 35:459-469.

Bromham, L., and D. Penny. 2003. The modern molecular clock. Nature Reviews Genetics 4:216-224.

Brown, J. W., J. S. Rest, J. Garcia-Moreno, M. D. Sorenson, and D. P. Mindell. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol 6:6.

Bunce, M., M. Szulkin, H. R. L. Lerner, I. Barnes, B. Shapiro, A. Cooper, and R. N. Holdaway. 2005. Ancient DNA provides new insights into the evolutionary history of New Zealand's extinct giant eagle. PLoS Biol 3:44-46.

Bunce, M., T. H. Worthy, T. Ford, W. Hoppitt, E. Willerslev, A. Drummond, and A. Cooper. 2003. Extreme reversed sexual size dimorphism in the extinct New Zealand moa *Dinornis*. Nature 425:172-175.

Bunce, M., T. H. Worthy, M. J. Phillips, R. N. Holdaway, E. Willerslev, J. Haile, B. Shapiro, P. Scofield, A. J. Drummond, P. J. J. Kamp, and A. Cooper. 2009. The evolutionary history of the extinct ratite moa and New Zealand Neogene paleogeography. P Natl Acad Sci USA Early Edition.

Butlin, R. K. 2008. Population genomics and speciation. Genetica Advance Online Access.

Cain, A. J. 1953. Affinities of the fruit-pigeon *Ptilinopus perousii*. Ibis 96:104-110.

Campbell, H. J., and G. Hutching. 2007. In search of ancient New Zealand. Penguin, Auckland & GNS Science, Lower Hutt, New Zealand.

Carlquist, S. 1974. Island Biology. Columbia University Press, New York.

Caughley, G. 1964. Does the New Zealand vertebrate fauna conform to zoogeographic principles? Tuatara 12:50-57.

Chambers, G. K., W. M. Boon, T. R. Buckley, and R. A. Hitchmough. 2001. Using molecular methods to understand the Gondwanan affinities of the New Zealand biota: three case studies. Australian Journal of Botany 49:377-387.

Chan, C. H., K. N. Ballantyne, H. Aikman, D. Fastier, C. H. Daugherty, and G. K. Chambers. 2006. Genetic analysis of interspecific hybridisation in the

world's only Forbes' Parakeet *Cyanoramphus forbesi* natural population. Conserv Genet 7:493-506.

Christidis, L., and W. E. Boles. 2008. Systematics and taxonomy of Australian birds. CSIRO Publishing, Melbourne, Australia.

Christidis, L., P. R. Leeton, and M. Westerman. 1996. Were bowerbirds part of the New Zealand fauna? P Natl Acad Sci USA 93:3898-3901.

Clarke, J. A., C. P. Tambussi, J. I. Noriega, G. M. Erickson, and R. A. Ketcham. 2005. Definitive fossil evidence for the extant avian radiation in the Cretaceous. Nature 433:305-308.

Clegg, S. L., S. M. Degnan, J. Kikkawa, C. Moritz, A. Estoup, and I. P. F. Owens. 2002. Genetic consequences of sequential founder events by an island-colonizing bird. . P Natl Acad Sci USA 99:8127-8132.

Clements, J. F. 2007. The Clements Checklist of Birds of the World. Cornell University Press, Ithaca, New York.

Cook, L. G., and M. D. Crisp. 2005a. Directional asymmetry of long-distance dispersal and colonization could mislead reconstructions of biogeography. J Biogeo 32:741-754.

Cook, L. G., and M. D. Crisp. 2005b. Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. . Proc R Soc Lond B Biol Sci 272.

Cooper, A., C. Lalueza-Fox, S. Anderson, A. Rambaut, J. Austin, and R. Ward. 2001. Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. Nature 409:704-707.

Cooper, A., C. Mourer-Chauviré, G. K. Chambers, A. Vonhaeseler, A. C. Wilson, and S. Paabo. 1992. Independent origins of New Zealand Moas and Kiwis. P Natl Acad Sci USA 89:8741-8744.

Cooper, A., and D. Penny. 1997. Mass survival of birds across the Cretaceous-Tertiary boundary: Molecular evidence. Science 275:1109-1113.

Cowie, R. H., and B. S. Holland. 2006. Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. J Biogeo 33:193–198.

Cracraft, J. 1974. Phylogeny and Evolution of Ratite Birds. Ibis 116:494-521.

Crisp, M. D. 2008. Book Review: Ghosts of Gondwana: The History of Life in New Zealand. Syst Biol 57:329-332.

Crisp, M. D., and L. G. Cook. 2005. Do early branching lineages signify ancestral traits? Trends Ecol Evol 20.

Darwin, C. 1859. On the Origin of Species by Means of Natural Selection. John Murray, London.

Daugherty, C. H., G. W. Gibbs, and R. A. Hitchmough. 1993. Mega-island or micro-continent? New Zealand and its fauna. Trends Ecol Evol 8:437-442.

De Beer, G. R. 1956. The evolution of the ratites. Bull Brit Mus (Nat Hist) Zool Series 4:1-63.

de Kloet, R. S., and S. R. de Kloet. 2005. The evolution of the spindlin gene in birds: Sequence analysis of an intron of the spindlin W and Z gene reveals four major divisions of the Psittaciformes. Mol Phylogenet Evol 36:706-721.

Diamond, J. M. 1974. Colonization of exploded volcanic islands by birds: The supertramp strategy. Science 184:803-806.

Driskell, A. C., and L. Christidis. 2004. Phylogeny and evolution of the Australo-Papuan honeyeaters (Passeriformes, Meliphagidae). Mol Phylogenet Evol 31:943-960.

Driskell, A. C., L. Christidis, B. J. Gill, W. E. Boles, F. K. Barker, and N. W. Longmore. 2007. A new endemic family of New Zealand passerine birds: adding heat to a biodiversity hotspot. Aust J Zool 55:73-78.

Dumbacher, J. P., T. K. Pratt, and R. C. Fleischer. 2003. Phylogeny of the owlet-nightjars (Aves : Aegothelidae) based on mitochondrial DNA sequence. Mol Phylogenet Evol 29:540-549.

Edwards, S. V., and W. E. Boles. 2002. Out of Gondwana: the origin of passerine birds. Trends Ecol Evol 17:347-349.

Ericson, P., C. Anderson, T. Britton, A. Elzanowski, and U. Johansson. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett 2:543-547.

Ericson, P. G. P., L. Christidis, A. Cooper, M. Irestedt, J. Jackson, U. S. Johansson, and J. A. Norman. 2002. A Gondwanan origin of passerine birds supported

by DNA sequences of the endemic New Zealand wrens. Proc R Soc Lond B Biol Sci 269:235-241.

Ericson, P. G. P., M. Irestedt, and U. S. Johansson. 2003. Evolution, biogeography, and patterns of diversification in passerine birds. J Avian Biol 34:3-15.

Ewen, J. G., I. Flux, and P. G. P. Ericson. 2006. Systematic affinities of two enigmatic New Zealand passerines of high conservation priority, the Hihi or Stitchbird *Notiomystis cincta* and the Kokako *Callaeas cinerea*. Mol Phylogenet Evol 40:281-284.

Fain, M. G., and P. Houde. 2004. Parallel radiations in the primary clades of birds. Evolution 58:2558-2573.

Falla, R. A. 1953. The Australian element in the avifauna of New Zealand. Emu 53:36-46.

Fleischer, R. C., C. E. McIntosh, and C. L. Tarr. 1998. Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K–Ar based ages of the Hawaiian Islands to estimate molecular evolutionary rates. Mol Ecol 7:533-545.

Fleming, C. A. 1950. New Zealand flycatchers of the genus *Petroica* Swainson (Aves), Parts I and II. Trans Roy Soc N Z 78:14-47, 126-160.

Fleming, C. A. 1962. History of the New Zealand land bird fauna. Notornis 9:270-274.

Fleming, C. A. 1974. The coming of the birds. Nature Heritage 1:61-68.

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

Fleming, C. A. 1979. The geological history of New Zealand and its life. Auckland University Press & Oxford University Press, Auckland.

Foggo, M. N., R. A. Hitchmough, and C. H. Daugherty. 1997. Systematics and conservation implications of geographic variation in pipits (Anthus: Motacillidae) in New Zealand and some offshore islands. Ibis 139:366-373.

Fortey, R. 1999. Life: A Natural History of the First Four Billion Years of Life on Earth. Vintage.

Friesen, V. L., and D. J. Anderson. 1997. Phylogeny and evolution of the Sulidae (Aves: Pelecaniformes): A test of alternative modes of speciation. Mol Phylogenet Evol 7:252-260.

Gibb, G. C., O. Kardailsky, R. T. Kimball, E. L. Braun, and D. Penny. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol Biol Evol 24:269-280.

Gibb, G. C., and D. Penny. 2010. Two aspects along the continuum of pigeon evolution: a South-Pacific radiation and the relationship of pigeons within Neoaves. Emu:in review.

Gibbs, G. 2006. Ghosts of Gondwana - The History of Life in New Zealand. Craig Potton Publishing.

Gillespie, G. D. 1985. Hybridization, introgression, and morphometric differentiation between Mallard *Anas platyrhynchos* and Grey Duck *Anas superciliosa* in Otago, New Zealand. Auk 102:459-469.

Given, A. D., J. A. Mills, and A. J. Baker. 2005. Molecular evidence for recent radiation in southern hemisphere masked gulls. Auk 122:268-279.

Goldberg, J., R. G. Powlesland, and S. A. Trewick. 2010. Population structure in the New Zealand Wood Pigeon (*Hemiphaga*: Aves: Columbidae) - inferred from mitochondrial DNA sequences. Emu:in review.

Goldberg, J., S. A. Trewick, and A. M. Paterson. 2008. Evolution of New Zealand's terrestrial fauna: a review of molecular evidence. Philos Trans R Soc Lond B Biol Sci 363:3319-3334.

Graham, I. J. 2008. A Continent on the Move: New Zealand Geoscience into the 21st century. Geological Society of New Zealand & GNS Science, Wellington.

Griffiths, C. 1999. Phylogeny of the Falconidae inferred from molecular and morphological data. Auk 116:116-130.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Harshman, J., E. L. Braun, M. J. Braun, C. J. Huddleston, R. C. K. Bowie, J. L. Chojnowski, S. J. Hackett, K.-L. Han, R. T. Kimball, B. D. Marks, K. J. Miglia, W. S. Moore, S. Reddy, F. H. Sheldon, D. W. Steadman, S. J. Steppan, C. C. Witt, and T. Yuri. 2008. Phylogenomic evidence for multiple losses of flight in ratite birds. P Natl Acad Sci USA 105:13462-13467.

Ho, S. Y. W., and M. J. Phillips. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58:367-380.

Holdaway, R. N., T. H. Worthy, and A. J. D. Tennyson. 2001. A working list of breeding bird species of the New Zealand region at first human contact. N Z J Zool 28:119-187.

Houde, P., A. C. Cooper, E. Leslie, A. E. Strand, and G. A. Montano. 1997. Phylogeney and evolution of 12S rDNA in Gruiformes (Aves). Pp. 121-158 *in* D. P. Mindell, ed. Avian Molecular Evolution and Systematics. Academic Press, San Diego.

Huynen, L., C. D. Millar, R. P. Scofield, and D. M. Lambert. 2003. Nuclear DNA sequences detect species limits in ancient moa. Nature 425:175-178.

Irestedt, M., J. Fuchs, K. A. Jønsson, J. I. Ohlson, E. Pasquet, and P. G. P. Ericson. 2008. The systematic affinity of the enigmatic Lamprolia victoriae (Aves: Passeriformes)—An example of avian dispersal between New Guinea and Fiji over Miocene intermittent land bridges? Mol Phylogenet Evol 48:1218-1222.

Jadwiszczak, P. 2009. Penguin past: The current state of knowledge. Polish Polar Res 30:3-28.

Johnson, K. J., and M. D. Sorenson. 1999. Phylogeny and biogeography of dabbling ducks (genus: Anas): a comparison of molecular and morphological evidence. Auk 116:792-805.

Johnson, K. P., and M. D. Sorenson. 1998. Comparing molecular evolution in two mitochondrial protein coding genes (cytochrome *b* and ND2) in the dabbling ducks (Tribe: Anatini). Mol Phylogenet Evol 10:82-94.

Joseph, L., E. P. Lessa, and L. Christidis. 1999. Phylogeny and biogeography in the evolution of migration: shorebirds of the *Charadrius* complex. J Biogeo 26:329-342.

Keast, A. 1971. Continental drift and the evolution of the biota on southern continents. Quart Rev Biol 56:359.

Kennedy, M., R. D. Gray, and H. G. Spencer. 2000. The phylogenetic relationships of the shags and cormorants: Can sequence data resolve a disagreement between behavior and morphology? Mol Phylogenet Evol 17:345-359.

Kennedy, M., and H. G. Spencer. 2000. Phylogeny, biogeography, and taxonomy of Australasian teals. Auk 117:154-163.

King, C. 1985. Immigrant Killers: Introduced Predators and the Conservation of Birds in New Zealand. Oxford University Press, Auckland.

King, M. 2003. The Penguin history of New Zealand. Penguin Books, Auckland.

Kirchman, J. J. 2009. Genetic tests of rapid parallel speciation of flightless birds from an extant volant ancestor. Biological Journal of the Linnean Society 96:601-616.

Knapp, M., K. Stockler, D. Havell, F. Delsuc, F. Sebastiani, and P. Lockhart. 2005. Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (southern beech). PLoS Biol 3:38-43.

Ksepka, D. T., S. Bertelli, and N. P. Giannini. 2006. The phylogeny of the living and fossil Sphenisciformes (penguins). Cladistics 22:412-441.

Labandeira, C. C., K. R. Johnson, and P. Wilf. 2002. Impact of the terminal Cretaceous event on plant–insect associations. P Natl Acad Sci USA 99:2061-2066.

Landis, C., H. J. Campbell, J. G. Begg, D. C. Mildenhall, A. M. Paterson, and S. Trewick. 2008. The Waipounamu Erosion Surface: questioning the antiquity of the New Zealand land surface and terrestrial fauna and flora. Geol Mag 145:173-197.

Lawrence, H. A., G. A. Taylor, C. D. Millar, and D. M. Lambert. 2008. High mitochondrial and nuclear genetic diversity in one of the world's most endangered seabirds, the Chatham Island Taiko *Pterodroma magentae*. Conserv Genet 9:1293-1301.

Lepage, D. 2009. Avibase: The World Bird Database. Bird Studies Canada.

Lerner, H. R. L., and D. P. Mindell. 2005. Phylogeny of eagles, Old World vultures, and other Accipitridae based on nuclear and mitochondrial DNA. Mol Phylogenet Evol 37:327-346.

Livezey, B. C. 1998. A phylogenetic analysis of the Gruiformes (Aves) based on morphological characters, with an emphasis on the rails (Rallidae). Phil Trans R Soc B 353:2077-2151.

Livezey, B. C. 2003. Evolution of Flightlessness in Rails (Gruiformes: Rallidae): Phylogenetic, Ecomorphological, and Ontogenetic Perspectives. Ornithological Monographs no. 53. American Ornithologists' Union, Washington D.C.

Long, J. A. 1998. Dinosaurs of Australia and New Zealand and Other Animals of the Mesozoic Era. Harvard University Press.

MacArthur, R. H., and E. O. Wilson. 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, N.J.

Mathews, G. M., and T. Iredale. 1921. The nature of the New Zealand avifauna. Emu 20:210-221.

Mayr, E. 1963. Animal Species and Evolution. Belknap, Cambridge.

McDowall, R. M. 1968. Oceanic Islands and Endemism. Systematic Zoology 17:346-350.

McDowall, R. M. 1994. Gamekeepers for the Nation: The Story of New Zealand's Acclimatisation Societies, 1861–1990. Canterbury University Press, Christchurch.

McDowall, R. M. 2008. Process and pattern in the biogeography of New Zealand – a global microcosm? J Biogeo 35:197-212.

Meffre, S., A. J. Crawford, and P. G. Quilty. 2006. Arc-continent collision forming a large island between New Caledonia and New Zealand in the Oligocene. Extended Abstracts, AESC2006 Australian Earth Sciences Congress Melbourne, Australia.

Millener, P. R. 1999. The history of the Chatham Islands' bird fauna of the last 7000 years – a chronicle of change and extinction. Proceedings of the 4th International meeting of the Society of Avian Paleontology and Evolution (Washington, D.C., June 1996). Smithsonian Contrib Paleobiol 89:85-109.

Miller, H. C., and D. M. Lambert. 2006. A molecular phylogeny of New Zealand's *Petroica* (Aves: Petroicidae) species based on mitochondrial DNA sequences. Mol Phylogenet Evol 40:844-855.

Molnar, R., and M. Pole. 1997. A Miocene crocodilian from New Zealand. Alcheringa 21:65-70.

Morgan-Richards, M., S. Trewick, A. Bartosch-Härlid, O. Kardailsky, M. J. Phillips, P. A. McLenachan, and D. Penny. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol Biol 8:20.

Mortimer, N. 2004. New Zealand's geological foundations. Gondwanan Research 7:261-272.

Moyle, R. G. 2006. A molecular phylogeny of kingfishers (Alcedinidae) with insights into early biogeographic history. The Auk 123:487-499.

Murphy, S. A., I. A. Flux, and M. C. Double. 2006. Recent evolutionary history of New Zealand's North and South Island Kokako *Callaeas cinerea* inferred from mitochondrial DNA sequences. Emu 106:41-48.

Neall, V. E., and S. A. Trewick. 2008. The age and origin of the Pacific Islands: a geological overview. Phil Trans R Soc B 363:3293–3308.

Nelson, G. J. 1975. Reviews: Biogeography, the vicariance paradigm, and continental drift. Systematic Zoology 24:490-504.

Norman, J., P. Olsen, and L. Christidis. 1998. Molecular genetics confirms taxonomic affinities of the endangered Norfolk Island Boobook Owl *Ninox novaeseelandiae undulata*. Biol Cons 86:33-36.

Norman, J. A., P. G. P. Ericson, K. A. Jønsson, J. Fjeldså, and L. Christidis. 2009. A multi-gene phylogeny reveals novel relationships for aberrant genera of Australo-Papuan core Corvoidea and polyphyly of the Pachycephalidae and Psophodidae (Aves: Passeriformes). Mol Phylogenet Evol 52:488-497.

Nyári, Á. S., B. W. Benz, K. A. Jønsson, J. Fjeldså, and R. G. Moyle. 2009. Phylogenetic relationships of fantails (Aves: Rhipiduridae). Zoologica Scripta.

Oliver, W. R. B. 1949. The moas of New Zealand and Australia. Dominion Museum Bulletin 15, Wellington, New Zealand.

Oliver, W. R. B. 1955. New Zealand Birds. A.H. and A.W. Reed, Wellington.

Olsen, J., M. Wink, H. Sauer-Gurth, and S. Trost. 2002. A new *Ninox* owl from Sumba, Indonesia. Emu 102:223-231.

Olson, S. L. 1990. Comments on the osteology and systematics of the New Zealand passerines of the genus *Mohoua*. Notornis 37:157-160.

Overeem, R. L., A. J. Peuker, C. M. Austin, P. Dann, and C. P. Burridge. 2008. Contrasting genetic structuring between colonies of the World's smallest penguin, *Eudyptula minor* (Aves: Spheniscidae). Conserv Genet 9:893-905.

Padian, K., and A. de Ricqlès. 2009. L'origine et l'evolution des oiseaux: 35 années de progress. Comptes Renus Palevol 8:257-280.

Paris, J.-P. 1981. Géologie de la Nouvelle Caledonie. Bureau de Recherches Géologiques et Minières Mémoire 13:1-278.

Parkes, K. C., and G. A. Clark. 1996. An additional character linking ratites and tinamous, and an interpretation of their monophyly. The Condor 68:459-471.

Paterson, A. M., L. J. Wallis, and G. P. Wallis. 2000. Preliminary molecular analysis of Pelecanoides georgicus (Procellariiformes: Pelecanoididae) on Whenua Hou (Codfish Island): implications for its taxonomic status. N Z J Zool 27:415-423.

Paton, T., O. Haddrath, and A. J. Baker. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proc R Soc Lond B Biol Sci 269:839-846.

Penny, D., and M. J. Phillips. 2004. The rise of birds and mammals: are microevolutionary processes sufficient for macroevolution? Trends Ecol Evol 19:516-522.

Pereira, S. L., and A. J. Baker. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol Biol Evol 23:1731-1740.

Pereira, S. L., K. P. Johnson, D. H. Clayton, and A. J. Baker. 2007. Mitochondrial and nuclear DNA sequences support a cretaceous origin of Columbiformes and a dispersal-driven radiation in the Paleogene. Syst Biol 56:656-672.

Peucker, A. J., P. Dann, and C. P. Burridge. 2009. Range-wide phylogeography of the Little Penguin *Eudyptula Minor*: Evidence of long-distance dispersal. . Auk 126:397-408.

Phillips, M. J., G. C. Gibb, E. A. Crimp, and D. Penny. 2010. Tinamous and moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. Syst Biol 59:90-107.

Pons, J.-M., A. Hassanin, and P.-A. Crochet. 2005. Phylogenetic relationships within the Laridae (Charadriiformes: Aves) inferred from mitochondrial markers. Mol Phylogenet Evol 37:686-699.

Pratt, R. C., G. C. Gibb, M. Morgan-Richards, M. J. Phillips, M. D. Hendy, and D. Penny. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol 26:313-326.

Quammen, D. 1997. The Song of the Dodo: Island Biogeography in an age of extinctions. Simon & Schuster, New York.

Rhymer, J. M., M. J. Williams, and M. J. Braun. 1994. Mitochondrial analysis of gene flow between New Zealand Mallards *Anas platyrhynchos* and Grey Ducks *A. superciliosa*. Auk 111:970-978.

Rhymer, J. M., M. J. Williams, and R. T. Kingsford. 2004. Implications of phylogeography and population genetics for subspecific taxonomy of Grey (Pacific Black) Duck *Anas superciliosa* and its conservation in New Zealand. Pacific Conservation Biology 10:57-66.

Robertson, B., T. Steeves, K. McBride, S. Goldstien, M. Williams, and N. Gemmell. 2007a. Phylogeography of the New Zealand Blue Duck *Hymenolaimus malacorhynchos*: implications for translocation and species recovery. Conserv Genet 8:1431-1440.

Robertson, C. J. R., P. Hyvonen, M. J. Fraser, and C. R. Pickard. 2007b. Atlas of Bird Distribution in New Zealand 1999-2004. The Ornithological Society of New Zealand, Wellington.

Robertson, C. J. R., and D. G. Medway, eds. 2003. New Zealand Recognised Bird Names (NZRBN) Database. Ornithological Society of New Zealand.

Roots, C. 2006. Flightless Birds. Greenwood Publishing Group, Santa Barbara.

Saether, B. E., S. Engen, A. P. Moller, M. E. Visser, E. Matthysen, W. Fiedler, M. M. Lambrechts, P. H. Becker, J. E. Brommer, J. Dickinson, C. Du Feu, F. R. Gehlbach, J. Merila, W. Rendell, R. J. Robertson, D. Thomson, and J. Torok. 2005. Time to extinction of bird populations. Ecology 86:693-700.

Sanmartin, I., and F. Ronquist. 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst Biol 53:216 - 243.

Seabrook-Davison, M., L. Huynen, D. M. Lambert, and D. H. Brunton. 2009. Ancient DNA resolves identity and phylogeny of New Zealand's extinct and living quail (*Coturnix sp.*). PLoS ONE 4:E6400.

Shaffer, S. A., Y. Tremblay, H. Weimerskirch, D. Scott, D. R. Thompson, P. M. Sagar, H. Moller, G. A. Taylor, D. G. Foley, B. A. Block, and D. P. Costa. 2006. Migratory shearwaters integrate oceanic resources across the Pacific Ocean in an endless summer. P Natl Acad Sci USA 103:12799-12802.

Sheldon, F. H. 1987. Phylogeny of herons estimated from DNA-DNA hybridization data. The Auk 104:97-108.

Sheldon, F. H., C. E. Jones, and K. G. McCracken. 2000. Relative patterns and rates of evolution in heron nuclear and mitochondrial DNA. Mol Biol Evol 17:437-450.

Sheldon, F. H., L. Whittingham, R. G. Moyle, B. Slikas, and D. W. Winkler. 2005. Phylogeny of swallows (Aves: Hirundinidae) estimated from nuclear and mitochondrial DNA sequencing. Mol Phylogenet Evol 35:254-270.

Shepherd, L. D., and D. M. Lambert. 2006. Nuclear microsatellite DNA markers for New Zealand kiwi (*Apteryx spp.*). Mol Ecol Notes 6:227-229.

Shepherd, L. D., and D. M. Lambert. 2007. The relationships and origins of the New Zealand wattlebirds (Passeriformes, Callaeatidae) from DNA sequence analyses. Mol Phylogenet Evol 43:480-492.

Sibley, C. G., and J. E. Ahlquist. 1987. The relationships of four species of New Zealand passerine birds. Emu 87:63-66.

Skipworth, J. P. 1974. Continental drift and the New Zealand biota. N Z J Geog 57:1-13.

Slack, K. E., F. Delsuc, P. A. Mclenachan, U. Arnason, and D. Penny. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol Phylogenet Evol 42:1-13.

Slack, K. E., C. M. Jones, T. Ando, G. L. A. Harrison, R. E. Fordyce, U. Arnason, and D. Penny. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol 23:1144-1155.

Sorenson, M. D., E. Oneal, J. Garcia-Moreno, and D. P. Mindell. 2003. More taxa, more characters: The hoatzin problem is still unresolved. Mol Biol Evol 20:1484-1498.

Sorenson, M. D., and R. B. Payne. 2005. Molecular systematics: cuckoo phylogeny inferred from mitochondrial DNA sequences. Pp. 68-94 *in* R. B. Payne, ed. Bird Families of the World: Cuckoos. Oxford University Press, Oxford.

Steadman, D. W. 2006. Extinction and Biogeography of Tropical Pacific Birds. The University of Chicago Press, Chicago.

Sutherland, R. 1999. Basement geology and tectonic development of the greater New Zealand region: an interpretation from regional magnetic data. Tectonophysics 308:341-362.

Tarling, D. H., and M. P. Tarling. 1971. Continental Drift: A Study of the Earth's Moving Surface. G. Bells & Sons, London.

Tavares, E., A. J. Baker, S. L. Pereira, and C. Y. Miyaki. 2006. Phylogenetic relationships and historical biogeography of neotropical parrots (Psittaciformes: Psittacidae: *Arini*) inferred from mitochondrial and nuclear DNA sequences. Syst Biol 55:454-470.

Tennyson, A. J. D., and P. Martinson. 2006. Extinct Birds of New Zealand. Te Papa Press, Wellington, New Zealand.

Thomas, G. H., M. A. Wills, and T. Szekely. 2004. A supertree approach to shorebird phylogeny. BMC Evol Biol 4:14.

Thomson, G. M. 1922. The Naturalization of Plants and Animals in New Zealand. Cambridge University Press, Cambridge.

Tokita, M., T. Kiyoshi, and K. N. Armstrong. 2007. Evolution of craniofacial novelty in parrots through developmental modularity and heterochrony. Evol & Devel 9:590-601.

Trewick, S. A. 1996. Morphology and evolution of two takahe: Flightless rails of New Zealand. J Zool 238:221-237.

Trewick, S. A. 1997a. Flightlessness and phylogeny amongst endemic rails (Aves: Rallidae) of the New Zealand region. Phil Trans R Soc B 352:429-446.

Trewick, S. A. 1997b. Sympatric flightless rails *Gallirallus dieffenbachii* and *G. modestus* on the Chatham Islands, New Zealand; morphometrics and alternative evolutionary scenarios. J R Soc N Z 27:451-464.

Trewick, S. A., and R. H. Cowie. 2008. Introduction. Evolution on Pacific islands: Darwin's legacy. Phil Trans R Soc B 363:3289-3291.

Trewick, S. A., and M. Morgan-Richards. 2009a. Evolution in New Zealand: getting it in perspective. Pp. 34-36. Geology & Genes IV. Geological Society of New Zealand Miscellaneous Publication 126, Wellington.

Trewick, S. A., and M. Morgan-Richards. 2009b. New Zealand, Biology. Pp. 665-673 *in* G. D. Gillespie, and D. A. Clague, eds. Encyclopedia of Islands. University of California Press, Berkeley, CA.

Trewick, S. A., A. M. Paterson, and H. J. Campbell. 2007. Hello New Zealand. J Biogeo 34:1-6.

Trewick, S. A., and T. H. Worthy. 2001. Origins and prehistoric ecology of Takahe based on morphometric, molecular and fossil data. Pp. 31-48 *in* W. G. Lee, and I. G. Jamieson, eds. The Takahe, Fifty Years of Conservation Management and Research. Otago University Press, Dunedin.

Turvey, S. T., and R. N. Holdaway. 2005. Postnatal ontogeny, population structure and extinction of the Giant Moa *Dinornis*. J Morph 265:70-86.

Vajda, V., R. J. L., and C. J. Hollis. 2001. Indication of global deforestation at the Cretaceous-Tertiary boundary by New Zealand fern spike. Science 294:1700-1702.

van Tuinen, M., D. B. Butvill, J. A. W. Kirsch, and S. B. Hedges. 2001. Convergence and divergence in the evolution of aquatic birds. Proc R Soc Lond B Biol Sci 268:1345-1350.

Veltman, C. J., S. Nee, and M. J. Crawley. 1996. Correlates of introduction success in exotic New Zealand birds. Am Nat 147:542-557.

Voelker, G. 1999. Molecular evolutionary relationships in the avian genus *Anthus* (pipits: Motacillidae). Mol Phylogenet Evol 11:84-94.

Wallis, G. P. 1999. Genetic status of New Zealand Black Stilt *Himantopus novaezelandiae* and impact of hybridisation. Conservation Advisory Science Notes No. 239. Department of Conservation, Wellington.

Wallis, G. P., and S. A. Trewick. 2009. New Zealand phylogeography: evolution on a small continent. Mol Ecol 18:3548-3580.

Walrond, C. 2008. 'Acclimatisation', Te Ara - the Encyclopedia of New Zealand, (updated 20-Nov-2008) http://www.TeAra.govt.nz/TheSettledLandscape/IntroducedPlantsAndAnimals/Acclimatisation/en.

Warren, B. H., E. Bermingham, R. P. Prys-Jones, and C. Thébaud. 2006. Immigration, species radiation and extinction in a highly diverse songbird lineage: white-eyes on Indian Ocean islands. Mol Ecol 15:3769-3786.

Waters, J. M., and D. Craw. 2006. Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. Syst Biol 55:351-356.

Williams, G. R. 1953. The dispersal from New Zealand and Australia of some introduced European passerines. Ibis 95:676-692.

Winkworth, R. C., S. J. Wagstaff, D. Glenny, and P. J. Lockhart. 2002. Plant dispersal N.E.W.S from New Zealand. Trends Ecol Evol 17:514-520.

Wood, J. R., N. J. Rawlence, G. M. Rogers, J. J. Austin, T. H. Worthy, and A. Cooper. 2008. Coprolite deposits reveal the diet and ecology of the extinct New Zealand megaherbivore moa (Aves, Dinornithiformes). Quat Sci Rev 27:2593-2602.

Worthy, T. H. 2001. A giant flightless pigeon gen. et sp nov and a new species of Ducula (Aves: Columbidae), from Quaternary deposits in Fiji. J R Soc N Z 31:763-794.

Worthy, T. H., S. J. Hand, J. P. Worthy, A. J. D. Tennyson, and R. P. Scofield. 2009. A Large Fruit Pigeon (Columbidae) from the Early Miocene of New Zealand. Auk 126:649-656.

Worthy, T. H., and R. N. Holdaway. 2002. Lost World of the Moa: Prehistoric Life of New Zealand. Indiana University Press, Bloomington, Indiana.

Worthy, T. H., and S. L. Olson. 2002. Relationships, adaptations, and habits of the extinct duck *Euryanas finschi*. Notornis 49:1-17.

Worthy, T. H., A. J. D. Tennyson, M. Archer, A. Musser, S. Handi, C. Jones, B. J. Douglas, J. A. McNamara, and R. M. D. Beck. 2006. Miocene mammal reveals a Mesozoic ghost lineage on insular New Zealand, southwest Pacific. P Natl Acad Sci USA 103:19419-19423.

Worthy, T. H., A. J. D. Tennyson, C. Jones, J. A. McNamara, and B. J. Douglas. 2007. Miocene waterfowl and other birds from central Otago, New Zealand. J Syst Palaeontol 5:1-39.

Wright, T. F., E. E. Schirtzinger, T. Matsumoto, J. R. Eberhard, G. R. Graves, J. J. Sanchez, S. Capelli, H. Muller, J. Scharpegge, G. K. Chambers, and R. C. Fleischer. 2008. A multilocus molecular phylogeny of the parrots (Psittaciformes): Support for a Gondwanan origin during the Cretaceous. Mol Biol Evol 25:2141-2156.

TABLE S2.1—Latin and common names of birds mentioned in Introduction and Discussion.

Latin Name	Common Name
Acridotheres tristis	Common Mynah
Acrocephalus sechellensis	Seychelles Warbler
Alauda arvensis	Skylark
Anas superciliosa	Grey Duck
Anomalopteryx, Emeus, Euryapteryx, Megalapteryx, Pachyornis	Moa
Ardea novaehollandiae	White Faced Heron
Arenaria interpres	Ruddy Turnstone
Athene noctua	Little Owl
Atrichornis spp.	Australian Scrub-birds
Bubulcus ibis	Cattle Egret
Calidris canutus	Knot
Carduelis spp	finches and allies
Charadrius melanops	Black Fronted Dotterel
Cnemiornis	geese
Columba livia	Rock Pigeon
Corvus frugilegus	Rook
Coturnix novaezelandiae	Quail
Cygnus	swan
Diomedea epomophora	Royal Albatross
Drepanididae	Hawaiian honeycreepers
Egretta alba	White Egret
Fulica	coots
Fulica atra	Australian Coot
Gallirallus australis	Weka
Gallirallus philippensis	Banded Rail
Harpagornis moorei	Giant Eagle
Heteralocha acutirostris	Huia
Himantopus himantopus	Pied Stilt
Himantopus novaezelandiae	Black Stilt
Hirundo tahitica	Pacific Swallow
Ixobrychus novaezelandiae	Bittern
Limosa lapponica	Godwit
Megadyptes antipodes	Yellow-eyed Penguin
Mesitornithidae	mesites
Morus serrator	Australasian Gannet
Ninox novaeseelandiae	Morepork
Palmchat Dulidae	Caribbean
Passer domesticus	House Sparrow
Petroica australis	New Zealand Robin
Petroica macrocephala	Tomtit
Phalacrocorax sulcirostris	Little Black Shag
Porphyrio hochstetteri	Takahe
Porphyrio porphyrio	Pukeko
Porzana tabuensis	Spotless Crake
Ptilinopus insularis	Henderson Island Fruit Dove
Puffinus huttoni	Hutton's Shearwater
Raphidae	Dodo/Solitaire
Raphus cucullatus	Mauritius Dodo
Rhinocryptinae spp.	South American Tapaculos
Rhipidura fulginosa	Fantail
Rhynochetos jubatus	New Caledonian Kagu
Sceloglaux albifacies	Laughing Owl
Sturnus vulgaris	Starling
Thinornis novaeseelandiae	Shore Plover
Todidae	Greater Antilles todies
Turdus melura	Blackbird
Turdus philomelos	Song Thrush
Turnagra capensis	Piopio
Vanellus miles	Masked Plover
Xenicus longipes	Bush Wren

TABLE S2.2—Peer reviewed publications reporting application of molecular data to New Zealand birds, categorised by taxonomic depth of study.

Taxon Group	**Deep Phylogeny**	**Genus & Species level**	**Species & Population**
Palaeognathae	Cooper et al. (2001)	Baker et al. (2005) moa, many mt, radiation	Allentoft et al. moa micros. (2009)
	kiwi, moa mt genomes	Bunce et al. (2009) moa, CR phylogeog.	Baker et al. (1995) kiwi phylogeog. Alloz, mt
	Haddrath and Baker (2001)	Burbidge et al.(2003) brown kiwi CR, Atp6/8, Cyt*b*	Bunce et al. (2003) moa sexing
	kiwi, moa mt genomes	Cooper et al. (1992) ratite (kiwi, moa) 12S	Gemmell et al. (2004) moa CR, popn size
	Harshman et al. (2008)	Huynen et al. (2003) moa sexing	Huynen et al. (2002) moa sexing
	kiwi, 20 nuc	Huynen et al. (2008) moa ID, CR	Jensen et al. (2008) brown kiwi micros.
	Phillips et al. (2010)	Lambert et al. (2005) moa CoI barcoding	Shepherd and Lambert (2006) kiwi micros.
	kiwi, moa mt genomes	Shepherd and Lambert (2008) kiwi Cyt*b*, CR	Willerslev et al. (2003) DNA from soil
		van Tuinen et al. (1998) ratite (kiwi), tinamou 12S, 16S	Wood et al. (2008) moa diet
Anseriformes		Johnson and Sorenson (1998) dabbling ducks ND2, Cyt*b*	Abdelkrim et al. (2009) Blue Duck micros.
		Johnson and Sorenson (1999) dabbling ducks ND2, Cyt*b*	Rhymer et al. (1994) Grey Duck, re digest, hyb.
		Kennedy and Spencer (2000) teal ATPase6&8,12S	Rhymer et al. (2004) Grey Duck CR hyb.
		Sorenson et al. (1999) moa-nalos mt	Robertson et al. (2007) Blue Duck CR.
		Sraml et al. (1996) Australian waterfowl Cyt*b*	Worthy (1997) *Cnemiornis* CR
Galliformes			Seabrook-Davison et al. (2009) *Coturnix* COI, Cyt*b*
Columbiformes	Johnson (2004) Introns		Goldberg *et al* (2010) *Hemiphaga* Cyt*b*, CR.
	Pereira et al. (2007) many mt, nuc		
Apodiformes		Dumbacher et al. (2003) owlet nightjars Atp8, Cyt*b*	
Procellariformes		Austin (1996) *Puffinus*, Cyt*b*	Abbott and Double (2003a) albatross micros.
		Austin et al. (2004) *Puffinus*, Cyt*b*	Abbott and Double (2003b) albatross CR.
		Nunn et al. (1996) albatrosses, Cyt*b*	Abbott et al. (2005) albatross mtDNA gene order
		Nunn and Stanley (1998) rates, Cyt*b*	Abbott et al. (2006) albatross mtDNA SNP
		Paterson et al. (1995) petrel & penguin-alloz.,12S	Alderman et al. (2005) albatross CR,
			Burg (1999) albatross micros.
			Burg and Croxall (2001) albatross mtDNA, micros.
			Burg (2007) albatross RFLP, micros.
			Double et al. (2003) albatross mtDNA ID.
			Lawrence et al. (2008a) Taiko sex ratio
			Lawrence et al. (2008b) Taiko Cyt*b*
			Lawrence et al. (2008c) Taiko Cyt*b*, CR, micros.
			Lawrence et al. (2008c) Taiko Cyt*b*, CR, micros.
			Lawrence et al. (2009) Magenta Petrel=Taiko Cyt*b*
			Moore et al. (2001) albatross CR, sex ratios
			Paterson et al. (2000) Diving Petrel 12S
			Walsh and Edwards (2005) albatross Cyt*b*.

Taxon Group	**Deep Phylogeny**	**Genus & Species level**	**Species & Population**
Sphenisciformes		Paterson et al. (1995) petrel & penguin-alloz.,12S	Banks et al. (2002) Little Blue Penguin 12S, Cyt*b*
		Ksepka et al. (2006) penguins, many mt, RAG1	Boessenkool et al. (2008a) Yellow-eyed Penguin CR
		Baker et al. (2006) penuins, many mt, nuc.	Boessenkool et al. (2008b) Yellow-eyed Penguin micros.
			Boessenkool et al. (2009) Yellow-eyed Penguin CR, micros.
			Peucker et al. (2009) Little Blue Penguin CR
Pelecaniformes		Friesen and Anderson (1997) gannets Cyt*b*	Piertney et al. (1998) cormorant micros.
		Hughes et al. (2007) pelecaniform, many mt, lice co-phylogeny	
		Kennedy et al. (2000) shag/cormorant 12S, Atp6,8	
		Kennedy and Spencer (2004) frigate/tropicbird many mt	
		Kennedy et al. (2005) pelecaniform 12S, Atp6,8	
		Siegel-Causey (1997) pelecaniform 12S, 16S, Cyt*b*	
Ciconiiformes		Sheldon (1987) herons DNA-DNA hybrid.	
		Sheldon et al. (2000) herons Cyt*b*	
Falconiformes & Accipitriformes		Bunce et al. (2005) eagle Cyt*b*, ND2	
Gruiformes		Houde (1997) Gruiformes 12S	Gregory and Quinn (2006) *Porphyrio* micros.
		Trewick (1997) flighted/flightless rails 12S, Cyt*b*	Grueber et al. (2008) *Porphyrio* micros.
			Kirchman (2009) banded rail phylogeog. CR
Charadriiformes	Baker et al. (2007) many mt, nuc	Banks and Paterson (2007) *Haematopus* many mt	Baker et al. (2009) Snipe mt, micros.
		Bridge et al. (2005) terns Cyt*b*, ND2, 12S	Given et al. (2002) gulls micros.
		Given et al. (2005) masked gulls ND2, ND5, Atp6, 8	Liebers et al. (2004) Herring Gull Cyt*b*
		Joseph et al. (1999) *Charadrius* Cyt*b*	Steeves et al. (2008) *Himantopus* micros.
		Pons et al. (2005) Laridae Cyt*b*	Wallis (1999) *Himantopus* hyb
Psittaciformes	Wright et al. (2008) many mt, nuc	Boon et al. (2000) *Cyanoramphus* Cyt*b*	Boon et al. (2001) *Cyanoramphus* Cyt*b*
		Boon et al. (2001) *Cyanoramphus* Cyt*b*	Chan et al. (2005) *Cyanoramphus* micros.
		Boon et al. (2008) Cyanoramphus/ *Eunymphicus* Cyt*b*	Chan et al. (2006) *Cyanoramphus* micros.
		de Kloet and de Kloet (2005) parrot divisions-spindlin	Robertson et al. (2009) *Strigops* micros.
		Ribas et al. (2007) parrot ND2, Cyt*b*	Sainsbury et al. (2004) *Nestor* micros.
		Steiger et al. (2008) bird olfactory genes	Tokunaga et al. (2007) Forbes' Parakeet sexing
		Tokita et al. (2007) parrot phylogeny and development	
Cuculiformes		Sorenson and Payne (2005) cuckoo phylogeny	
Strigiformes			Norman et al. (1998) *Ninox* Cyt*b*
Coraciiformes		Moyle (2006) kingfishers Rag-1, ND2	

Taxon Group	**Deep Phylogeny**	**Genus & Species level**	**Species & Population**
Passeriformes	Barker et al.(2002) Rag-1 c-mos Barker et al. (2004) Rag-1 c-mos Chesser (2004) βFib-7 ND3, CoII Ericson et al. (2002) c-myc, Rag-1 Irestedt and Ohlson (2008) many nuc	Christidis et al. (1996) *Turnagra*-bower birds Cyt*b* Cooper (1994) Acanthisittidae 12S Driskell and Christidis (2004) Meliphagidae many mt, Fib-5 Driskell et al. (2007) *Notiomystis* ND2, Cyt*b*, 12S, Fib-5 Ewen et al. (2006) *Notiomystis Callaeas* Rag-1, c-mos Irestedt et al. (2008) passerine nuclear genes Miller and Lambert (2006) *Petroica* Cyt*b*, CR Norman et al. (2009) Corvoidea Rag-1, Cyt*b* Myo Nyári et al. (2009) *Rhipidura* ND2, ND3, TGFb2, Fib5 Sheldon et al. (2005) swallows ND2, Cyt*b*, βFib-7 Shepherd and Lambert (2007) Callaeatidae Rag-1, c-mos, Cyt*b*, ND2 Tebbutt and Simons (2002) *Heteralocha* 12S, c-mos Yuri et al. (2008) passerine gene duplication	Brekke et al. (2009) Hihi micros. Clegg et al. (2002) Estoup and Clegg (2003) *Zosterops* micros. Foggo et al. (1997) *Anthus* Alloz. Ma and Lambert (1997) *Petroica* minsat. hyb. Miller and Lambert (2004a) *Petroica* MHC Miller and Lambert (2006) *Petroica* Cyt*b*, CR Miller and Lambert (2004b) *Petroica* MHC Murphy et al. (2006) *Callaeas* CR Taylor et al. (2007) *Petroica, Philesturnus*
Multiple orders	Brown et al. (2008) many mt Ericson et al. (2006) 5 nuc Hackett et al. (2008) 19 nuc van Tuinen et al.(2001) many mt, nuc Complete mt genomes: Gibb et al. (2007) Harrison et al. (2004) Morgan-Richards et al. (2008) Pratt et al. (2009) Slack et al. (2003) Slack et al. (2006) Watanabe et al. (2006)	Carvalho et al. (2007) UV genes, vision analysis Chojnowski et al. (2008) Intron/exon comparison Cooper and Penny (1997) 12S, c-mos divergence timing Cooper and Cooper (1995) kiwi, moa NZ wren, ND6 Simon et al. (2004) 28 species, Preproinsulin gene Steiger et al. (2008) olfactory receptor genes, smell	Smith and Filardi (2007) island biogeo, ND2 Tavares and Baker (2008) CoI barcoding

Taxa studied are indicated and main types of data used are as follows: alloz. allozyme, biogeo. biogeography, CR control region, hyb. hybridisation, micros. microsatellite, mt mitochondrial, nuc nuclear, phylogeo. phylogeography, popn, population, SNP single nucleotide polymorphism, RE restriction enzyme, RFLP Restriction length polymorphism, MHC major histocompatibility complex, plus standard gene codes 12S, 16S, Atp6, Atp8, Cyt*b*, ND2,3,5, c-mos, c-myc, Rag-1, βFib-5, 7, TGFb2.

REFERENCES

Abbott, C., M. Double, R. Gales, G. Baker, A. Lashko, C. Robertson, and P. Ryan. 2006. Molecular provenance analysis for Shy and White-capped Albatrosses killed by fisheries interactions in Australia, New Zealand, and South Africa. Conserv Genet 7:531-542.

Abbott, C. L., and M. C. Double. 2003a. Genetic structure, conservation genetics and evidence of speciation by range expansion in Shy and White-capped Albatrosses. Mol Ecol 12:2953.

Abbott, C. L., and M. C. Double. 2003b. Phylogeography of Shy and White-capped Albatrosses inferred from mitochondrial DNA sequences: implications for population history and taxonomy. Mol Ecol 12:2747.

Abbott, C. L., M. C. Double, J. W. H. Trueman, A. Robinson, and A. Cockburn. 2005. An unusual source of apparent mitochondrial heteroplasmy: duplicate mitochondrial control regions in *Thalassarche* albatrosses. Mol Ecol 14:3605-3613.

Abdelkrim, J., B. C. Robertson, J. L. Stanton, and N. J. Gemmell. 2009. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. BioTechniques 46:185-192.

Alderman, R., M. C. Double, J. Valencia, and R. P. Gales. 2005. Genetic affinities of newly sampled populations of Wandering and Black-browed Albatross. Emu 105:169-179.

Allentoft, M. E., S. C. Schuster, R. N. Holdaway, M. L. Hale, E. McLay, C. Oskam, M. T. P. Gilbert, P. Spencer, E. Willerslev, and M. Bunce. 2009. Identification of microsatellites from an extinct moa species using high-throughput (454) sequence data. BioTechniques 46:185-200.

Austin, J. 1996. Molecular phylogenetics of *Puffinus* shearwaters: preliminary evidence from mitochondrial cytochrome *b* gene sequences. Mol Phylogenet Evol 6:77-88.

Austin, J. J., V. Bretagnolle, and E. Pasquet. 2004. A global molecular phylogeny of the small *Puffinus* shearwaters and implications for systematics of the little-audubon's shearwater complex. Auk 121:847-864.

Baker, A., C. Miskelly, and O. Haddrath. 2009. Species limits and population differentiation in New Zealand snipes (Scolopacidae: *Coenocorypha*) Conserv Genet.

Baker, A., S. Pereira, O. Haddrath, and K. Edge. 2006. Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. Proc R Soc B 273:11-17.

Baker, A. J., C. H. Daugherty, R. Colbourne, and J. L. McLennan. 1995. Flightless Brown Kiwis of New Zealand possess extremely subdivided population-structure and cryptic Sspecies like small mammals. P Natl Acad Sci USA 92:8254-8258.

Baker, A. J., L. J. Huynen, O. Haddrath, C. D. Millar, and D. M. Lambert. 2005. Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: The giant moas of New Zealand. P Natl Acad Sci USA 102:8257-8262.

Baker, A. J., S. Pereira, and T. A. Paton. 2007. Phylogenetic relationships and divergence times of Charadriiformes genera: multigene evidence for the Cretaceous origin of at least 14 clades of shorebirds. Biol Lett 3:205-209.

Banks, J. C., A. D. Mitchell, J. R. Waas, and A. M. Paterson. 2002. An unexpected pattern of molecular divergence within the Little Blue Penguin *Eudyptula minor* complex. Notornis 49:29-33.

Banks, J. C., and A. M. Paterson. 2007. A preliminary study of the genetic differences in New Zealand oystercatcher species. N Z J Zool 34:141-144.

Barker, F. K., G. F. Barrowclough, and J. G. Groth. 2002. A phylogenetic hypothesis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. Proc R Soc Lond B Biol Sci 269:295-308.

Barker, F. K., A. Cibois, P. Schikler, J. Feinstein, and J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. P Natl Acad Sci USA 101:11040-11045.

Boessenkool, S., J. J. Austin, T. H. Worthy, P. Scofield, A. Cooper, P. J. Seddon, and J. M. Waters. 2008a. Relict or colonizer? Extinction and range expansion of penguins in southern New Zealand. Proc R Soc B.

Boessenkool, S., T. M. King, P. J. Seddon, and J. M. Waters. 2008b. Isolation and characterization of microsatellite loci from the Yellow-eyed Penguin *Megadyptes antipodes*. Mol Ecol Resour 8:1043-1045.

Boessenkool, S., B. Star, J. M. Waters, and P. J. Seddon. 2009. Multilocus assignment analyses reveal multiple units and rare migration events in the recently expanded Yellow-eyed Penguin *Megadyptes antipodes*. Mol Ecol 18:2390-2400.

Boon, W. M., C. H. Daugherty, and G. K. Chambers. 2001. The Norfolk Island Green Parrot and New Caledonian Red-crowned Parakeet are distinct species. Emu 101:113-121.

Boon, W. M., J. C. Kearvell, C. H. Daugherty, and G. K. Chambers. 2000. Molecular systematics of New Zealand *Cyanoramphus* parakeets: conservation of Orange-fronted and Forbes' Parakeets. Bird Cons Int 10:211-239.

Boon, W. M., O. Robinet, N. J. Rawlence, V. Bretagnolle, J. A. Norman, L. Christidis, and G. K. Chambers. 2008. Morphological, behavioural and genetic differentiation within the Horned Parakeet *Eunymphicus cornutus* and its affinities to *Cyanoramphus* and *Prosopeia*. Emu 108:251-260.

Brekke, P., D. A. Dawson, G. J. Horsburgh, and J. G. Ewen. 2009. Characterization of microsatellite loci in the Hihi *Notiomystis cincta* (Notiomystidae, Aves). Mol Ecol Resour 9:1255-1258.

Bridge, E. S., A. W. Jones, and A. J. Baker. 2005. A phylogenetic framework for the terns (Sternini) inferred from mtDNA sequences: implications for taxonomy and plumage evolution. Mol Phylogenet Evol 35:459-469.

Bunce, M., M. Szulkin, H. R. L. Lerner, I. Barnes, B. Shapiro, A. Cooper, and R. N. Holdaway. 2005. Ancient DNA provides new insights into the evolutionary history of New Zealand's extinct giant eagle. PLoS Biol 3:44-46.

Bunce, M., T. H. Worthy, T. Ford, W. Hoppitt, E. Willerslev, A. Drummond, and A. Cooper. 2003. Extreme reversed sexual size dimorphism in the extinct New Zealand moa *Dinornis*. Nature 425:172-175.

Bunce, M., T. H. Worthy, M. J. Phillips, R. N. Holdaway, E. Willerslev, J. Haile, B. Shapiro, P. Scofield, A. J. Drummond, P. J. J. Kamp, and A. Cooper. 2009. The evolutionary history of the extinct ratite moa and New Zealand Neogene paleogeography. P Natl Acad Sci USA Early Edition.

Burbidge, M. L., R. M. Colbourne, H. A. Robertson, and A. J. Baker. 2003. Molecular and other biological evidence supports the recognition of at least three species of brown kiwi. Conserv Genet 4:167-177.

Burg, T. M. 1999. Isolation and characterization of microsatellites in albatrosses. Mol Ecol 8:337.

Burg, T. M. 2007. Genetic analysis of Wandering Albatrosses killed in longline fisheries off the east coast of New Zealand. Aquatic Conservation: Marine and Freshwater Ecosystems 17:S93-S101.

Burg, T. M., and J. P. Croxall. 2001. Global relationships amongst Black-browed and Grey-headed Albatrosses: analysis of population structure using mitochondrial DNA and microsatellites. Mol Ecol 10:2647-2660.

Carvalho, L. S., J. A. Cowing, S. E. Wilkie, J. K. Bowmaker, and D. M. Hunt. 2007. The molecular evolution of avian ultraviolet- and violet-sensitive visual pigments. Mol Biol Evol 24:1843-1852.

Chan, C., K. N. Ballantyne, D. M. Lambert, and G. K. Chambers. 2005. Characterization of variable microsatellite loci in Forbes' Parakeet *Cyanoramphus forbesi* and their use in other parrots. Conserv Genet 6:651-654.

Chan, C. H., K. N. Ballantyne, H. Aikman, D. Fastier, C. H. Daugherty, and G. K. Chambers. 2006. Genetic analysis of interspecific hybridisation in the world's only Forbes' Parakeet *Cyanoramphus forbesi* natural population. Conserv Genet 7:493-506.

Chesser, R. T. 2004. Molecular systematics of New World suboscine birds. Mol Phylogenet Evol 32:11-24.

Christidis, L., P. R. Leeton, and M. Westerman. 1996. Were bowerbirds part of the New Zealand fauna? P Natl Acad Sci USA 93:3898-3901.

Clegg, S. L., S. M. Degnan, J. Kikkawa, C. Moritz, A. Estoup, and I. P. F. Owens. 2002. Genetic consequences of sequential founder events by an island-colonizing bird. . P Natl Acad Sci USA 99:8127-8132.

Cooper, A. 1994. Ancient DNA sequences reveal unsuspected phylogenetic relationships within New Zealand Wrens (Acanthisittidae). Experientia 50:558-563.

Cooper, A., and R. A. Cooper. 1995. The Oligocene bottleneck and New-Zealand biota - genetic record of a past environmental crisis. Proc R Soc Lond B Biol Sci 261:293-302.

Cooper, A., C. Lalueza-Fox, S. Anderson, A. Rambaut, J. Austin, and R. Ward. 2001. Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. Nature 409:704-707.

Cooper, A., C. Mourer-Chauviré, G. K. Chambers, A. Vonhaeseler, A. C. Wilson, and S. Paabo. 1992. Independent origins of New Zealand Moas and Kiwis. P Natl Acad Sci USA 89:8741-8744.

Cooper, A., and D. Penny. 1997. Mass survival of birds across the Cretaceous-Tertiary boundary: Molecular evidence. Science 275:1109-1113.

de Kloet, R. S., and S. R. de Kloet. 2005. The evolution of the spindlin gene in birds: Sequence analysis of an intron of the spindlin W and Z gene reveals four major divisions of the Psittaciformes. Mol Phylogenet Evol 36:706-721.

Double, M. C., R. Gales, T. Reid, N. Brothers, and C. L. Abbott. 2003. Morphometric comparison of Australian Shy and New Zealand White-capped Albatrosses. Emu 103:287-294.

Driskell, A. C., and L. Christidis. 2004. Phylogeny and evolution of the Australo-Papuan honeyeaters (Passeriformes, Meliphagidae). Mol Phylogenet Evol 31:943-960.

Driskell, A. C., L. Christidis, B. J. Gill, W. E. Boles, F. K. Barker, and N. W. Longmore. 2007. A new endemic family of New Zealand passerine birds: adding heat to a biodiversity hotspot. Aust J Zool 55:73-78.

Dumbacher, J. P., T. K. Pratt, and R. C. Fleischer. 2003. Phylogeny of the owlet-nightjars (Aves : Aegothelidae) based on mitochondrial DNA sequence. Mol Phylogenet Evol 29:540-549.

Ericson, P., C. Anderson, T. Britton, A. Elzanowski, and U. Johansson. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett 2:543-547.

Ericson, P. G. P., L. Christidis, A. Cooper, M. Irestedt, J. Jackson, U. S. Johansson, and J. A. Norman. 2002. A Gondwanan origin of passerine birds supported by DNA sequences of the endemic New Zealand wrens. Proc R Soc Lond B Biol Sci 269:235-241.

Estoup, A., and S. M. Clegg. 2003. Bayesian inferences on the recent island colonization history by the bird *Zosterops lateralis lateralis*. Mol Ecol 12:657-674.

Ewen, J. G., I. Flux, and P. G. P. Ericson. 2006. Systematic affinities of two enigmatic New Zealand passerines of high conservation priority, the Hihi or Stitchbird *Notiomystis cincta* and the Kokako *Callaeas cinerea*. Mol Phylogenet Evol 40:281-284.

Foggo, M. N., R. A. Hitchmough, and C. H. Daugherty. 1997. Systematics and conservation implications of geographic variation in pipits (Anthus: Motacillidae) in New Zealand and some offshore islands. Ibis 139:366-373.

Fonseca, M. M., D. Posada, and D. J. Harris. 2008. Inverted Replication of Vertebrate Mitochondria. Mol Biol Evol 25:805-808.

Friesen, V. L., and D. J. Anderson. 1997. Phylogeny and evolution of the Sulidae (Aves: Pelecaniformes): A test of alternative modes of speciation. Mol Phylogenet Evol 7:252-260.

Garcia-Moreno, J., and D. P. Mindell. 2000. Rooting a Phylogeny with Homologous Genes on Opposite Sex Chromosomes (Gametologs): A Case Study Using Avian CHD. Mol Biol Evol 17:1826-1832.

Gemmell, N. J., M. K. Schwartz, and B. C. Robertson. 2004. Moa were many. Proc R Soc Lond B Biol Sci 271:S430-S432.

Gibb, G. C., O. Kardailsky, R. T. Kimball, E. L. Braun, and D. Penny. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol Biol Evol 24:269-280.

Given, A. D., J. A. Mills, and A. J. Baker. 2002. Isolation of polymorphic microsatellite loci from the Red-billed Gull *Larus novaehollandiae scopulinus* and amplification in related species. Mol Ecol Notes 2:416-418.

Given, A. D., J. A. Mills, and A. J. Baker. 2005. Molecular evidence for recent radiation in southern hemisphere masked gulls. Auk 122:268-279.

Goldberg, J., R. G. Powlesland, and S. A. Trewick. 2010. Population structure in the New Zealand Wood Pigeon (*Hemiphaga*: Aves: Columbidae) - inferred from mitochondrial DNA sequences. Emu:in review.

Grueber, C. E., T. M. King, J. M. Waters, and I. G. Jamieson. 2008. Isolation and characterization of microsatellite loci from the endangered New Zealand Takahe (Gruiformes; Rallidae; *Porphyrio hochstetteri*). Mol Ecol Resour 8:884-886.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Haddrath, O., and A. J. Baker. 2001. Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis. Proc R Soc Lond B Biol Sci 268:939-945.

Harrison, G. L., P. A. McLenachan, M. J. Phillips, K. E. Slack, A. Cooper, and D. Penny. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late Cretaceous. Mol Biol Evol 21:974-983.

Harshman, J., E. L. Braun, M. J. Braun, C. J. Huddleston, R. C. K. Bowie, J. L. Chojnowski, S. J. Hackett, K.-L. Han, R. T. Kimball, B. D. Marks, K. J. Miglia, W. S. Moore, S. Reddy, F. H. Sheldon, D. W. Steadman, S. J. Steppan, C. C. Witt, and T. Yuri. 2008. Phylogenomic evidence for multiple losses of flight in ratite birds. P Natl Acad Sci USA 105:13462-13467.

Houde, P., A. C. Cooper, E. Leslie, A. E. Strand, and G. A. Montano. 1997. Phylogeny and evolution of 12S rDNA in Gruiformes (Aves). Pp. 121-158 *in* D. P. Mindell, ed. Avian Molecular Evolution and Systematics. Academic Press, San Diego.

Hughes, J. M., M. Kennedy, K. P. Johnson, R. L. Palma, and R. Page. 2007. Multiple cophylogenetic analyses reveal frequent cospeciation between Pelecaniform birds and Pectinopygus lice. Syst Biol 26:232-251.

Huynen, L., C. D. Millar, and D. M. Lambert. 2002. A DNA test to sex ratite birds. Mol Ecol 11:851-856.

Huynen, L., C. D. Millar, R. P. Scofield, and D. M. Lambert. 2003. Nuclear DNA sequences detect species limits in ancient moa. Nature 425:175-178.

Irestedt, M., and J. Ohlson. 2008. The division of the major songbird radiation into Passerida and 'core Corvoidea' (Aves: Passeriformes) — the species tree vs. gene trees. Zoologica Scripta 37:305-313.

Johnson, K. J., and M. D. Sorenson. 1999. Phylogeny and biogeography of dabbling ducks (genus: Anas): a comparison of molecular and morphological evidence. Auk 116:792-805.

Johnson, K. P. 2004. Deletion bias in avian introns over evolutionary timescales. Mol Biol Evol 21:599-602.

Johnson, K. P., and M. D. Sorenson. 1998. Comparing molecular evolution in two mitochondrial protein coding genes (cytochrome *b* and ND2) in the dabbling ducks (Tribe: Anatini). Mol Phylogenet Evol 10:82-94.

Joseph, L., E. P. Lessa, and L. Christidis. 1999. Phylogeny and biogeography in the evolution of migration: shorebirds of the *Charadrius* complex. J Biogeo 26:329-342.

Kennedy, M., R. D. Gray, and H. G. Spencer. 2000. The phylogenetic relationships of the shags and cormorants: Can sequence data resolve a disagreement between behavior and morphology? Mol Phylogenet Evol 17:345-359.

Kennedy, M., B. R. Holland, R. D. Gray, and H. G. Spencer. 2005. Untangling long branches: Identifying conflicting phylogenetic signals using spectral analysis, neighbor-net, and consensus networks. Syst Biol 54:620-633.

Kennedy, M., and H. G. Spencer. 2000. Phylogeny, biogeography, and taxonomy of Australasian teals. Auk 117:154-163.

Kennedy, M., and H. G. Spencer. 2004. Phylogenies of the frigatebirds (Fregatidae) and tropicbirds (Phaethonidae), two divergent groups of the traditional order Pelecaniformes, inferred from mitochondrial DNA sequences. Mol Phylogenet Evol 31:31-38.

Kirchman, J. J. 2009. Genetic tests of rapid parallel speciation of flightless birds from an extant volant ancestor. Biological Journal of the Linnean Society 96:601-616.

Kitamura, S., S. Suzuki, T. Yumoto, P. Poonswad, P. Chuailua, K. Plongmai, T. Maruhashi, N. Noma, and C. Suckasam. 2006. Dispersal of *Canarium euphyllum* (Burseraceae), a large-seeded tree species, in a moist evergreen forest in Thailand. J Trop Ecol 22:137-146.

Ksepka, D. T., S. Bertelli, and N. P. Giannini. 2006. The phylogeny of the living and fossil Sphenisciformes (penguins). Cladistics 22:412-441.

Lambert, D. M., A. Baker, L. Huynen, O. Haddrath, P. D. N. Hebert, and C. D. Millar. 2005. Is a large-scale DNA-based inventory of ancient life possible? J Hered 96:279-284.

Lawrence, H. A., C. D. Millar, M. J. Imber, D. E. Crockett, J. H. Robins, P. Scofield, G. A. Taylor, and D. M. Lambert. 2009. Molecular evidence for the identity of the Magenta petrel. Mol Ecol Resour 9:458-461.

Lawrence, H. A., C. D. Millar, G. A. Taylor, L. D. Macdonald, and D. M. Lambert. 2008a. Excess of unpaired males in one of the World's most endangered seabirds, the Chatham Island Taiko *Pterodroma magentae*. J Avian Biol 39:359-363.

Lawrence, H. A., R. P. Scofield, D. E. Crockett, C. D. Millar, and D. M. Lambert. 2008b. Ancient genetic variation in one of the world's rarest seabirds. Heredity 101:543-547.

Lawrence, H. A., G. A. Taylor, C. D. Millar, and D. M. Lambert. 2008c. High mitochondrial and nuclear genetic diversity in one of the world's most endangered seabirds, the Chatham Island Taiko *Pterodroma magentae*. Conserv Genet 9:1293-1301.

Liebers, D., P. d. Knijff, and A. J. Helbig. 2004. The Herring Gull complex is not a ring species. Proc: Biol Sci 271:893-901.

Ma, W., and D. Lambert. 1997. Minisatellite DNA markers reveal hybridisation between the endangered Black Robin and Tomtit. Electrophoresis 18:1682-1687.

Mayr, G. 2008. Avian higher-level phylogeny: well-supported clades and what we can learn from a phylogenetic analysis of 2954 morphological characters. Journal of Zoological Systematics and Evolutionary Research 46:63-72.

Miller, H. C., and D. M. Lambert. 2004a. Gene duplication and gene conversion in class II MHC genes of New Zealand robins (Petroicidae). Immunogenetics 56:178-191.

Miller, H. C., and D. M. Lambert. 2004b. Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). Mol Ecol 13:3709-3721.

Miller, H. C., and D. M. Lambert. 2006. A molecular phylogeny of New Zealand's *Petroica* (Aves: Petroicidae) species based on mitochondrial DNA sequences. Mol Phylogenet Evol 40:844-855.

Moore, P. J., T. M. Burg, G. A. Taylor, and C. D. Millar. 2001. Provenance and sex ratio of Black-browed Albatross *Thalassarche melanophrys* breeding on Campbell Island, New Zealand. Emu 101:329-334.

Morgan-Richards, M., S. Trewick, A. Bartosch-Härlid, O. Kardailsky, M. J. Phillips, P. A. McLenachan, and D. Penny. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol Biol 8:20.

Morrison, D. A. 2007. Increasing the efficiency of searches for the maximum likelihood tree in a phylogenetic analysis of up to 150 nucleotide sequences. Syst Biol 56:988-1010.

Moyle, R. G. 2006. A molecular phylogeny of kingfishers (Alcedinidae) with insights into early biogeographic history. The Auk 123:487-499.

Murphy, S. A., I. A. Flux, and M. C. Double. 2006. Recent evolutionary history of New Zealand's North and South Island Kokako *Callaeas cinerea* inferred from mitochondrial DNA sequences. Emu 106:41-48.

Norman, J., P. Olsen, and L. Christidis. 1998. Molecular genetics confirms taxonomic affinities of the endangered Norfolk Island Boobook Owl *Ninox novaeseelandiae undulata*. Biol Cons 86:33-36.

Norman, J. A., P. G. P. Ericson, K. A. Jønsson, J. Fjeldså, and L. Christidis. 2009. A multi-gene phylogeny reveals novel relationships for aberrant genera of Australo-Papuan core Corvoidea and polyphyly of the Pachycephalidae and Psophodidae (Aves: Passeriformes). Mol Phylogenet Evol 52:488-497.

Nunn, G., and S. Stanley. 1998. Body size effects and rates of cytochrome b evolution in tube-nosed seabirds. Mol Biol Evol 15:1360-1371.

Nyári, Á. S., B. W. Benz, K. A. Jønsson, J. Fjeldså, and R. G. Moyle. 2009. Phylogenetic relationships of fantails (Aves: Rhipiduridae). Zoologica Scripta.

Nylander, J. A. A., J. C. Wilgenbusch, D. L. Warren, and D. L. Swofford. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24:581-583.

Paterson, A. M., G. P. Wallis, and R. D. Gray. 1995. Penguins, petrels and parsimony: does cladistic analysis of behavior reflect seabird phylogeny? Evolution 49:974-989.

Pereira, S. L., K. P. Johnson, D. H. Clayton, and A. J. Baker. 2007. Mitochondrial and nuclear DNA sequences support a cretaceous origin of Columbiformes and a dispersal-driven radiation in the Paleogene. Syst Biol 56:656-672.

Peucker, A. J., P. Dann, and C. P. Burridge. 2009. Range-wide phylogeography of the Little Penguin *Eudyptula Minor*: Evidence of long-distance dispersal. . Auk 126:397-408.

Phillips, M. J., G. C. Gibb, L. Crimp, and D. Penny. 2010. Tinamous and moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. Syst Biol:in press.

Piertney, S. B., A. Goostrey, J. F. Dallas, and D. N. Carss. 1998. Highly polymorphic microsatellite markers in the Great Cormorant *Phalacrocorax carbo*. Mol Ecol 7:133-140.

Pons, J.-M., A. Hassanin, and P.-A. Crochet. 2005. Phylogenetic relationships within the Laridae (Charadriiformes: Aves) inferred from mitochondrial markers. Mol Phylogenet Evol 37:686-699.

Pratt, R. C., G. C. Gibb, M. Morgan-Richards, M. J. Phillips, M. D. Hendy, and D. Penny. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol 26:313-326.

Rhymer, J. M., M. J. Williams, and M. J. Braun. 1994. Mitochondrial analysis of gene flow between New Zealand Mallards *Anas platyrhynchos* and Grey Ducks *A. superciliosa*. Auk 111:970-978.

Rhymer, J. M., M. J. Williams, and R. T. Kingsford. 2004. Implications of phylogeography and population genetics for subspecific taxonomy of Grey (Pacific Black) Duck *Anas superciliosa* and its conservation in New Zealand. Pacific Conservation Biology 10:57-66.

Ribas, C. C., R. G. Moyle, C. Y. Miyaki, and J. Cracraft. 2007. The assembly of montane biotas: linking Andean tectonics and climatic oscillations to independent regimes of diversification in *Pionus* parrots. Proc R Soc B 274:2399-2408.

Riskin, D. K., S. Parsons, W. A. Schutt, Jr, G. G. Carter, and J. W. Hermanson. 2006. Terrestrial locomotion of the New Zealand short-tailed bat *Mystacina tuberculata* and the common vampire bat *Desmodus rotundus*. J Exp Biol 209:1725-1736.

Robertson, B., T. Steeves, K. McBride, S. Goldstien, M. Williams, and N. Gemmell. 2007. Phylogeography of the New Zealand Blue Duck *Hymenolaimus malacorhynchos*: implications for translocation and species recovery. Conserv Genet 8:1572-9737.

Robertson, B. C., N. Frauenfelder, D. K. Eason, G. Elliott, and R. Moorhouse. 2009. Thirty polymorphic microsatellite loci from the critically endangered Kakapo *Strigops habroptilus*. Mol Ecol Resour 9:664-666.

Ross, H. A., S. Murugan, and W. L. S. Li. 2008. Testing the Reliability of Genetic Methods of Species Identification via Simulation. Syst Biol 57:216-230.

Sainsbury, J. P., E. S. Macavoy, and G. K. Chambers. 2004. Characterization of microsatellite loci in the Kaka, *Nestor meridionalis*. Mol Ecol Notes 4:623-625.

Seabrook-Davison, M., L. Huynen, D. M. Lambert, and D. H. Brunton. 2009. Ancient DNA resolves identity and phylogeny of New Zealand's extinct and living quail (*Coturnix sp.*). PLoS ONE 4:E6400.

Sheldon, F. H. 1987. Phylogeny of herons estimated from DNA-DNA hybridization data. The Auk 104:97-108.

Sheldon, F. H., C. E. Jones, and K. G. McCracken. 2000. Relative patterns and rates of evolution in heron nuclear and mitochondrial DNA. Mol Biol Evol 17:437-450.

Sheldon, F. H., L. Whittingham, R. G. Moyle, B. Slikas, and D. W. Winkler. 2005. Phylogeny of swallows (Aves: Hirundinidae) estimated from nuclear and mitochondrial DNA sequencing. Mol Phylogenet Evol 35:254-270.

Shepherd, L. D., and D. M. Lambert. 2007. The relationships and origins of the New Zealand wattlebirds (Passeriformes, Callaeatidae) from DNA sequence analyses. Mol Phylogenet Evol 43:480-492.

Shepherd, L. D., and D. M. Lambert. 2008. Ancient DNA and conservation: lessons from the endangered kiwi of New Zealand. Mol Ecol 17:2174-2184.

Siegel-Causey, D. 1997. Phylogeny of the Pelecaniformes: Molecular systematics of a privative group. Pp. 159-171 *in* D. P. Mindell, ed. Avian Molecular Evolution and Systematics. Academic Press, San Diego, CA.

Simon, J., S. Laurent, G. Grolleau, P. Thoraval, D. Soubieux, and D. Rasschaert. 2004. Evolution of preproinsulin gene in birds. Mol Phylogenet Evol 30:755-766.

Slack, K. E., A. Janke, D. Penny, and U. Arnason. 2003. Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. Gene 302:43-52.

Slack, K. E., C. M. Jones, T. Ando, G. L. A. Harrison, R. E. Fordyce, U. Arnason, and D. Penny. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol 23:1144-1155.

Smith, C. E., and C. E. Filardi. 2007. Patterns of molecular and morphological variation in some Solomon Island land birds. Auk 124:479-493.

Sorenson, M. D., A. Cooper, E. E. Paxinos, T. W. Quinn, H. F. James, S. L. Olson, and R. C. Fleischer. 1999. Relationships of the extinct Moa-nalos, flightless Hawaiian waterfowl, based on ancient DNA. Proc R Soc Lond B Biol Sci 266:2187-2193.

Sorenson, M. D., and R. B. Payne. 2005. Molecular systematics: cuckoo phylogeny inferred from mitochondrial DNA sequences. Pp. 68-94 *in* R. B. Payne, ed. Bird Families of the World: Cuckoos. Oxford University Press, Oxford.

Sraml, M., L. Christidis, S. Easteal, P. Horn, and C. Collet. 1996. Molecular relationships within Australasian waterfowl (Anseriformes). Aust J Zool 44:47-58.

Steeves, T., M. L. Hale, and N. J. Gemmell. 2008. Development of polymorphic microsatellite markers for the New Zealand Black Stilt *Himantopus novaezelandiae* and cross-amplification in the Pied Stilt *Himantopus himantopus leucocephalus*. Mol Ecol Resour 8:1105-1107.

Steiger, S. S., A. E. Fidler, M. Valcu, and B. Kempenaers. 2008. Avian olfactory receptor gene repertoires: evidence for a well-developed sense of smell in birds? Proc R Soc Lond B Biol Sci 275:2309-2317.

Tavares, E., and A. Baker. 2008. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. BMC Evol Biol 8:81.

Taylor, S. S., I. G. Jamieson, and G. P. Wallis. 2007. Historic and contemporary levels of genetic variation in two New Zealand passerines with different histories of decline. J Evol Biol 20:2035-2047.

Tebbutt, S. J., and C. Simons. 2002. Gene sequences from New Zealand's extinct Huia. J R Soc N Z 32:327-335.

Tokunaga, C., W. Lam, D. Fastier, H. Aikman, C. H. Chan, and G. K. Chambers. 2007. Morphological sexing of Forbes' Parakeet *Cyanoramphus forbesi* validated by molecular data. Notornis 54:38-41.

Trewick, S. A. 1997. Flightlessness and phylogeny amongst endemic rails (Aves: Rallidae) of the New Zealand region. Phil Trans R Soc B 352:429-446.

van Tuinen, M., and S. B. Hedges. 2001. Calibration of avian molecular clocks. Mol Biol Evol 18:206-213.

van Tuinen, M., C. G. Sibley, and S. B. Hedges. 1998. Phylogeny and biogeography of ratite birds inferred from DNA sequences of the mitochondrial ribosomal genes. Mol Biol Evol 15:370-376.

Wallis, G. P. 1999. Genetic status of New Zealand Black Stilt *Himantopus novaezelandiae* and impact of hybridisation. Conservation Advisory Science Notes No. 239. Department of Conservation, Wellington.

Walsh, H. E., and S. V. Edwards. 2005. Conservation genetics and Pacific fisheries bycatch: Mitochondrial differentiation and population assignment in Black-footedAalbatrosses *Phoebastria nigripes*. Conserv Genet 6:289-295.

Watanabe, M., M. Nikaido, T. T. Tsuda, T. Kobayashi, D. Mindell, Y. Cao, N. Okada, and M. Hasegawa. 2006. New candidate species most closely related to penguins. Gene 378:65-73.

Weir, J. T., and D. Schluter. 2008. Response to Comment on "The Latitudinal Gradient in Recent Speciation and Extinction Rates of Birds and Mammals". Science 319:901d-.

Willerslev, E., A. J. Hansen, J. Binladen, T. B. Brand, M. T. P. Gilbert, B. Shapiro, M. Bunce, C. Wiuf, D. A. Gilichinsky, and A. Cooper. 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science 300:791-795.

Wood, J. R., N. J. Rawlence, G. M. Rogers, J. J. Austin, T. H. Worthy, and A. Cooper. 2008. Coprolite deposits reveal the diet and ecology of the extinct New Zealand megaherbivore moa (Aves, Dinornithiformes). Quat Sci Rev 27:2593-2602.

Worthy, T. H. 2001. A giant flightless pigeon gen. et sp nov and a new species of Ducula (Aves: Columbidae), from Quaternary deposits in Fiji. J R Soc N Z 31:763-794.

Worthy, T. H., R. N. Holdaway, M. D. Sorenson, and A. C. Cooper. 1997. Description of the first complete skeleton of the extinct New Zealand Goose *Cnemiornis calcitrans* (Aves: Anatidae), and a reassessment of the relationships of *Cnemiornis*. J Zool 243:695-723.

Wright, T. F., E. E. Schirtzinger, T. Matsumoto, J. R. Eberhard, G. R. Graves, J. J. Sanchez, S. Capelli, H. Muller, J. Scharpegge, G. K. Chambers, and R. C. Fleischer. 2008. A multilocus molecular phylogeny of the parrots (Psittaciformes): Support for a Gondwanan origin during the Cretaceous. Mol Biol Evol 25:2141-2156.

Yuri, T., R. T. Kimball, E. L. Braun, and M. J. Braun. 2008. Duplication of accelerated evolution and growth hormone gene in passerine birds. Mol Biol Evol 25:352-361.

CHAPTER THREE

3 MITOCHONDRIAL GENOMES AND AVIAN PHYLOGENY: COMPLEX CHARACTERS AND RESOLVABILITY WITHOUT EXPLOSIVE RADIATIONS[1]

3.1 ABSTRACT

We improve the taxon sampling for avian phylogeny by analyzing seven new mitochondrial genomes (a toucan, woodpecker, osprey, forest falcon, American kestrel, heron and a pelican). This improves inference of the avian tree, and it supports three major conclusions. The first is that some birds (including a parrot, a toucan and the osprey) exhibit a complete duplication of the control region meaning that there are at least four distinct gene orders within birds. However it appears that there are regions of continued gene conversion between the duplicate control regions, resulting in duplications that can be stable for long evolutionary periods. Because of this stable duplicated state, gene order can eventually either revert to the original order, or change to the new gene order. The existence of this stable duplicate state explains how an apparently unlikely event (finding the same novel gene order) can arise multiple times. Although rare genomic changes have theoretical advantages for tree-reconstruction, they can be compromised if these apparently rare events have a stable intermediate state. Secondly, the toucan and woodpecker improve the resolution of the six-way split within Neoaves that has been called an “explosive radiation”. An explosive radiation implies that normal

[1] This chapter published as: Gibb, G. C., O. Kardailsky, R. T. Kimball, E. Braun, and D. Penny. 2007. Mitochondrial genomes and avian phylogeny: complex characters and resolvability without explosive radiations. Mol Biol Evol 24:269-280

microevolutionary events are insufficient to explain the observed macroevolution. By showing the avian tree is, in principle, resolvable we demonstrate that the radiation of birds is amenable to standard evolutionary analysis. Thirdly, and as expected from theory, additional taxa breaking up long branches stabilize the position of some problematic taxa (like the falcon). In addition, we report that within the birds of prey and allies we did not find evidence pairing New World vultures with storks, or Accipitrids (hawks, eagles and osprey) with Falconids.

Keywords—gene order, forest falcon, osprey, kestrel, woodpecker, heron, toucan, pelican, explosive radiation

3.2 INTRODUCTION

Our primary interest here is using the phylogenies of birds to test questions such as, whether the processes of microevolution are sufficient to explain macroevolution, or how frequently major changes occur in the ecological niche a group occupies. In practice we need to distinguish between the five models of Penny and Phillips (2004; see also Cooper and Penny 1997) on the extent that ecological, physiological and taxonomic diversification occurred prior to, or after, the Cretaceous/Tertiary (K/T) boundary. Such a program of inquiry needs to be broken down into many testable steps that can be examined using specific data sets. Here we use seven new mitochondrial genomes to consider three main aspects of the questions. The first is to understand why a particular change in mitochondrial gene order appears to have occurred several times during avian evolution, and therefore why (in this case) gene order may not be a useful phylogenetic character. The next aspect is that the additional taxa make it appear that resolution of the basal 6-way split among Neoaves (Cracraft 2001) will be possible, eliminating the need to postulate an 'explosive radiation' (e.g. Poe and Chubb 2004). Finally, breaking up some long branches increases the stability of the tree as predicted from theory.

Over the past 30 years the use of DNA or protein sequence data has increasingly become the main data type used to recover phylogeny in general, but there are fundamental limits on how far back sequence data will allow reliable recovery of evolutionary history (Mossel and Steel 2005). In principle, 'rare genomic changes' (Rokas and Holland 2000; Boore 2006), or more evocatively, 'sequence-characters

uniquely-derived' (SCUDs), such as changes in gene order, can retain information for long periods of time. When the number of character states is so high that the same change is unlikely ever to be repeated then simple parsimony is a maximum likelihood estimator (Steel and Penny 2004; 2005). Such rare DNA changes can, in principle, retain phylogenetic information even when primary sequence data must have become randomized due to the long time periods involved (see Mossel and Steel 2005). With mammals, the identification of retrotransposon insertions have been extremely valuable (Nishihara et al. 2005; 2006), a fact highlighted in the recent resolution of the placental mammal tree, including the position of the root, using only rare genomic changes (Kriegs et al. 2006). This was equivalent to giving DNA sequence data 30 years start, and catching and overtaking them in a single study. Although the particular repetitive elements used to study mammalian evolution may not be so useful in birds there is considerable potential for the use of these types of rare events in phylogenetic studies (see Snel et al. 2005). Two recent studies in birds have used the chicken repeat 1 (CR1) retrotransposon to determine relationships among closely related groups of birds, and this is promising for the future use of insertions in elucidating deeper avian phylogenetic relationships (St. John et al. 2005; Watanabe et al. 2006).

Differences in mitochondrial gene order have been useful for phylogenetic resolution of some groups of species, for example Arthropoda being monophyletic, and within this Crustacea grouping with Hexapoda to the exclusion of Myriapoda and Onychophora (see Boore 2006 and references therein). Birds also have a different mt gene order compared to other vertebrates, and this reinforced their already accepted monophyly (Desjardins and Morais 1990). The difficulty in general is ensuring that the rare genomic changes are genuinely unique events. Several different arrangements of mitochondrial gene order have been observed in birds, including the likely ancestral avian gene order first found in the chicken (*Gallus gallus* Desjardins and Morais, 1990), corresponding to cyt*b*/tThr/tPro/ND6/tGlu/Control Region/tPhe/12SrRNA (Figure 3.1b) and the alternative gene order reported by Mindell et al. (1998) that is cyt*b*/tThr/ Control Region/tPro/ND6/tGlu/Non-coding region/tPhe/12S rRNA (Figure 3.1e). However, Mindell et al. (1998) pointed out all plausible avian phylogenies implied that the alternative gene order arose independently several times within birds. Despite these parallel changes, they suggested that gene order may still be useful in certain cases, for example, to distinguish oscine from suboscine passerines; though

Bensch and Härlid (2000) later reported exceptions within oscines. Species with the alternative gene order typically had a short non-coding region between tRNA Glu and tRNA Phe. However, a control region duplication has been observed in Amazona parrots (Eberhard et al. 2001, Figure 3.1d), and another alternative (Figure 3.1c) in albatrosses was reported by Abbott et al. (2005). It is important to understand the reasons for the multiple origins of an alternative gene order because in many cases gene order has potential for being excellent markers for phylogeny (Snel, et al. 2005; Steel and Penny 2005; Boore 2006).

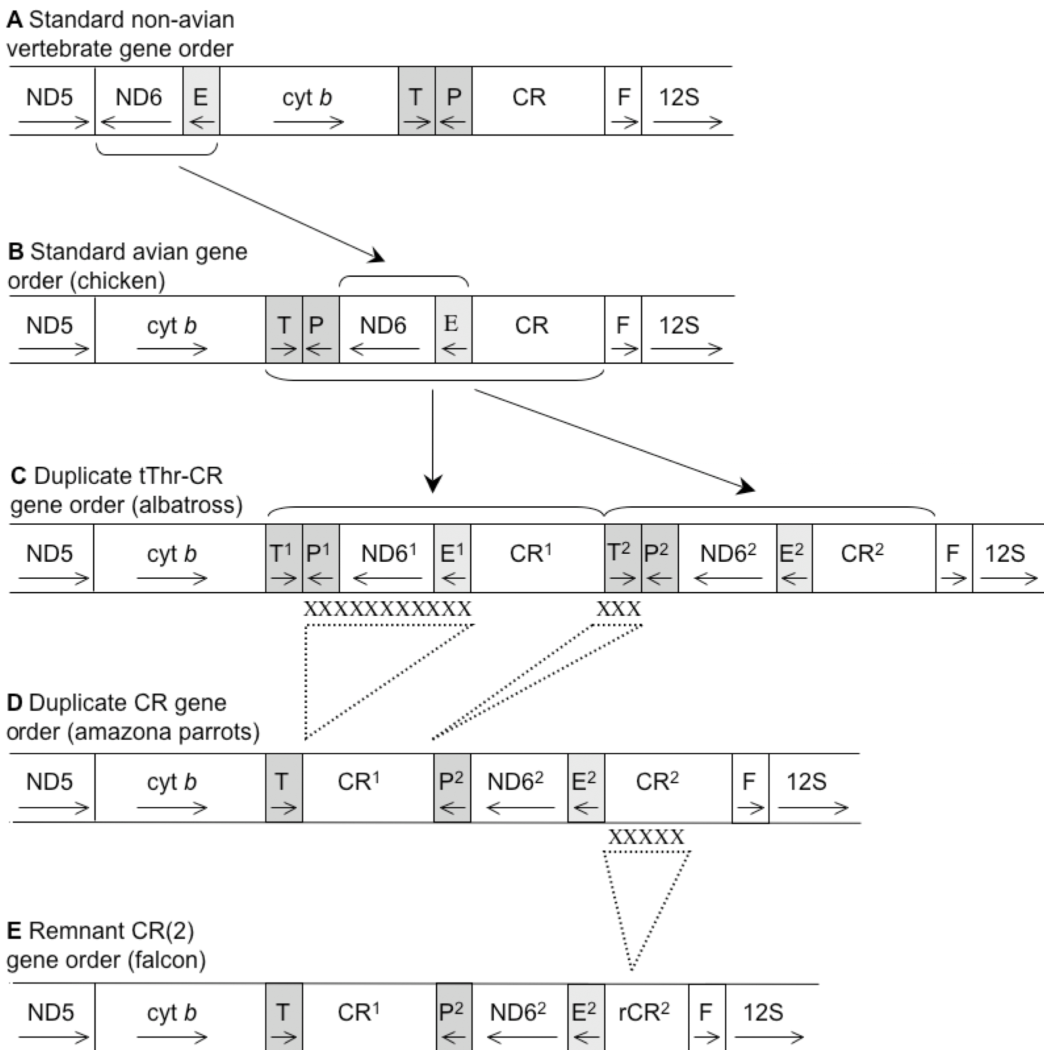

FIGURE 3.1—Gene orders found in avian mitochondrial genome control regions. Arrows between figure parts show one scenario for conversion between the different gene orders. Arrows underneath gene names represent gene directionality. X's and dotted triangles represent possible gene loss or reduction.

Turning to avian phylogeny in particular, there is support for a basal split into paleognaths (ratites and tinamou) and neognaths, with neognaths then being

further split into Galloanseres (chickens and ducks) and Neoaves (a group containing 95% of avian species). This three-way split is now found on morphological, nuclear and mitochondrial data (for example, Groth and Barrowclough 1999; Cracraft et al. 2004; Slack et al. 2007). In contrast, the basic divisions within Neoaves are not clear, and Cracraft (2001) suggests a 6 way split between:

Passerines (or perching birds, Passeriformes),

Parrots (Psittaciformes),

Cuckoos (Cuculiformes),

Woodpeckers/toucans, rollers/bee-eaters/kingfishers, jacamars/puffbirds and mousebirds (Coliiformes, Coraciiformes and Piciformes, jacamars and puffbirds sometimes split from Piciformes and placed in Galbuliformes),

Owls, nightjars, swifts and turacos (Strigiformes, Caprimulgiformes, Apodiformes, Musophagiformes)

Seabirds, shore birds, doves, cranes, raptors, rails, penguins, storks, loons and grebes (a very diverse group including the traditional orders Charadriiformes, Ciconiiformes, Columbiformes, Falconiformes, Gaviformes, Gruiformes, Pelicaniformes, Procellariformes and Sphenisciformes).

Even though this division is modified somewhat in Cracraft et al. (2004), we use these six groups as an informal prior for evaluating results, first to see how well those groups stand up to further analysis, and then to see how much resolution can be found in their branching order (if the six are supported). Of the six Neoaves groups, only four (passerines, an owl, a parrot and the seabird/shorebird/raptor alliance) are currently represented in the complete mitochondrial dataset. Two members of another group (a toucan and woodpecker) are added here, leaving only cuckoos unrepresented in the avian tree from mitochondrial genomes. The largest group represented (seabird/ shorebird/raptors/gruiformes) is very diverse and has various informal names such as Cracrafti or Conglomerati (Slack et al. 2007) or simply ‘water carnivores’ - because it includes the main carnivorous birds (raptors – buzzards, hawks, eagle, osprey, etc) and a large group of aquatic birds (shorebirds, seabirds and penguins) that are carnivorous.

This 6-way split of Neoaves has been called an 'explosive radiation' (see for example, Poe and Chubb 2004). This expression raises concerns regarding the sufficiency of microevolutionary processes to explain macroevolution (Penny and Phillips 2004). As commonly used, the term 'explosive radiation' implies that there are major examples where microevolution is unable to explain macroevolution. With the present example, we take 'explosive radiation' to imply both;

an unresolvable 6-way split, and

simultaneous (geologically very fast) morphological and ecological radiation of the six Neoavian lineages.

Such major morphological changes would be difficult to explain using known microevolutionary processes. However, as yet we have no information regarding the rate of morphological change during the radiation. Using parrots to illustrate this example, there are two quite separate issues;

when the parrot lineage diverged from other Neoaves, and

when the mix of morphological and ecological features arose by which we define modern parrots (the crown group).

In practice, the mix of parrot features could have occurred significantly after the divergence of the lineage. Therefore, we do not agree with the use of terms such as 'explosive radiations' just because phylogeny is difficult to resolve (Poe and Chubb 2004), when evidence regarding the speed of morphological and ecological adaptation is unavailable. Instead of using explosive radiation', we use 'adaptive radiation' when the divergences may be fast (in geological time), thus leading to short, difficult to resolve internodes. However, in such cases normal microevolutionary processes are sufficient to account for any adaptive component of the radiation. There are many examples of well-studied adaptive radiations (e.g. Lockhart et al. 2001). Thus it is important when testing the five hypotheses of Penny and Phillips (2004) to determine the times of divergence of the Neoavian groups (see Slack et al. 2006).

The third topic studied here is testing for increased stability of the avian tree by breaking up some long branches. Theoretical (Hendy and Penny 1989; Mossel and Steel 2005) and simulation based (Hillis et al. 1994), as well as empirical studies (Anderson and Swofford 2004) show that breaking up long branches is important

to increase the stability of a tree. Our experience with both mammalian (Lin et al. 2002; Phillips and Penny 2003) and avian (Slack et al. 2007) mitochondrial genomes has strongly supported this conclusion –increased taxon sampling has increased the agreement between nuclear and mitochondrial datasets. Thus it is important to improve taxon selection to get a reasonably stable tree. As mentioned above, only four of Cracraft's (2001) six Neoaves lineages are currently represented in the complete mitochondrial genome dataset. A fifth proposed lineage corresponding to the group containing woodpeckers, rollers, bee-eaters, kingfishers, jacanas and mousebirds is added here. The two species added are an ivory-billed aracari (a toucan, *Pteroglossus azara*) and a pileated woodpecker (*Dryocopus pileatus*). These two species are expected to be quite distantly related to each other, but sufficiently close to lessen the effects of long-branch attraction from having just one member from this proposed group.

The other five new taxa are from the Conglomerati/Cracrafti/water carnivore group, namely osprey, forest falcon, kestrel, a pelican and a heron. The novel raptors (osprey and forest falcon and American kestrel) were selected because the peregrine falcon has been difficult to place on the avian tree. Although predicted to be members of the water carnivore group related to the other seabirds and shorebirds (Cracraft 2001), the falcon tended to come out basal to the passerines when few mitochondrial genome sequences were available (see discussion in Slack et al. 2003). With additional sequences, especially another raptor (e.g. buzzard), the falcon usually shifts into the water-carnivore group (Slack et al. 2007). However, when a single parrot and/or owl sequence are included the falcon can join with one of these groups, even when the buzzard is in the dataset (Harrison et al. 2004). By contrast, the buzzard has never come outside the water-carnivores (Slack et al. 2007).

The reason for the instability of the placement of the falcon has not been identified. It could reflect some form of compositional bias (Phillips and Penny 2003) or a covarion-like shift like that reported in primates (Schmidt et al. 2005; see also Ane et al. 2005). However, given the instability of the falcon, we believe that it is important to add additional raptors into the dataset. The osprey (*Pandion haliaetus*, Accipitridae) is often placed in the same family as the buzzard but is not a close relative. Similarly, a forest falcon (*Micrastur gilvicollis*, Falconidae) is expected to be deep on the falcon lineage (see Sibley and Ahlquist 1990), again

breaking up a long branch. In addition, the American kestrel (*Falco sparvarius*) and hawk eagle (*Spizaetus alboniger,* Asai et al. unpublished) fall within the falcon/forest falcon and osprey/buzzard groupings respectively, and they would be expected to further stabilize this part of the tree.

The position of storks on the avian tree has also been uncertain (see Slack et al. 2006; 2007). As expected, they are within the Conglomerati, but have come closest to penguins, even when a turkey vulture (*Cathartes aura*) was included in the tree (Slack et al. 2007). Based upon morphological/behavioral characters (Ligon 1967) and DNA-DNA hybridization (Sibley and Ahlquist 1990) it had been suggested New World (or cathartid) vultures like the turkey vulture should be grouped with the storks, rather than raptors. We have added a white-faced heron (*Ardea novaehollandiae*) and an Australian pelican (*Pelecanus conspicillatus*) to the mitochondrial dataset to further examine the stork/penguin association.

Progress is made on each of the three questions discussed here. It appears as if a duplicated control region can be maintained for relatively long periods of time (tens of millions of years) by gene conversion between the two copies, and the maintenance of a duplicated control region has the potential to explain the apparent homoplasy in mitochondrial gene order. We find evidence that the radiation of Neoaves is potentially resolvable. While there is still methodological difficulty in resolving some parts of the tree due to short internodes, this does not mean unknown forces must be at work. Although these results need to be supported by nuclear data, unless further evidence were to come to light showing simultaneous morphological and ecological radiation also occured, it is not necessary to postulate an 'explosive radiation' (e.g. Poe and Chubb 2004), that would involve unknown evolutionary forces. Finally, adding additional taxa does seem to increase the stability of the avian tree.

3.3 MATERIALS AND METHODS

The forest falcon (*Micrastur gilvicollis*) and aracari (*Pteroglossus azara*) samples were provided by the Louisiana State University Museum of Natural Science Collection of Genetic Resources and are samples B-10720 and B-9081 respectively. The osprey (*Pandion haliaetus*) was provided by the Australian Museum (Sydney),

sample EBU 37010, and the Australian pelican (*Pelecanus conspicillatus*) by the Museum of Victoria, sample number MV 1883. The white-faced heron (*Ardea novaehollandiae*) was provided by the Department of Conservation (Waikanae). The pileated woodpecker (*Dryocopus pileatus*) and the American kestrel (*Falco sparverius*) were collected in North Central Florida near Gainsville and were part of the Braun/Kimball laboratory tissue collection.

For the forest falcon, aracari, osprey, pelican and heron, extractions of genomic DNA were from 25-50 mg of liver tissue using the High Pure™ PCR Template Preparation Kit (Protocol Vb; Boehringer Mannheim) according to the manufacturers instructions. To minimize the possibility of obtaining nuclear copies of mitochondrial (mt) genes (NUMTS), mitochondrial genomes were amplified in 2-3 long overlapping fragments (3.5 – 12kb in length) using the Expand ™ Long template PCR System (Roche). The woodpecker and kestrel were also amplified in 2 long overlapping segments, although Eppendorf Triple Master Taq was used for long PCR. The products were excised from agarose gel using an Eppendorf gel extraction kit and the long-range products were then used as templates for subsequent short range PCR of overlapping fragments 0.6 – 3 kb in length. Primers were found by searching an electronic database maintained in our laboratory (described in Slack et al. 2007) or by examining a list maintained by the Braun/Kimball group. Sequencing was performed using BigDye® Terminator Cycle Sequencing reagents v3.1 according to the manufacturers instructions (Applied Biosystems), and the reactions sequenced on ABI 3730 automated sequencers (Applied Biosystems). Sequences were aligned in Sequencher™ 4.2.2 (Gene Codes Corp.) and manually edited and checked for complete agreement between sequences.

In some cases (for example, length heteroplasmy in control regions from short nucleotide sequence repeats), PCR products were cloned using standard techniques with Promega pGemT Easy Vector system and Invitrogen Max efficiency DH5α competent cells. At least three clones were sequenced for each region to check for any PCR error. In all cases, overlaps between sequences were sufficient to ensure homology. Sequence identity was confirmed through a combination of BLAST searches of the NCBI database (http://www.ncbi.nlm.nih.gov/blast/), confirmation of amino acid translation in coding regions, and alignment with other species.

In addition to the seven new avian sequences reported in this paper, 33 other complete avian mt genomes were included in the analyses (26 neognaths and seven paleognaths). The 26 neognath taxa are: chicken (*Gallus gallus*; GenBank accession number AP003317), Japanese quail (*Coturnix japonica*; AP003195), Australian brush-turkey (*Alectura lathami*; AY346091), magpie goose (*Anseranas semipalmata*; AY309455), redhead (duck, *Aythya americana*; AF090337), greater white-fronted goose (*Anser albifrons*; AF363031), rifleman (NZ wren, *Acanthisitta chloris*; AY325307), gray-headed broadbill (*Smithornis sharpei*; AF090340), fuscous flycatcher (*Cnemotriccus fuscatus*; AY596278), superb lyrebird (*Menura novaehollandiae*; AY542313), village indigobird (*Vidua chalybeata*; AF090341), rook (*Corvus frugilegus*; Y18522), morepork (NZ owl, *Ninox novaeseelandiae*; AY309457), kakapo (NZ parrot, *Strigops habroptilus*; AY309456), peregrine falcon (*Falco peregrinus*; AF090338), Eurasian buzzard (*Buteo buteo*; AF380305), Blythe's hawk eagle (*Spizaetus alboniger*; AP008239), turkey vulture (*Cathartes aura*, AY463690), blackish oystercatcher (*Haematopus ater*; AY074886), ruddy turnstone (*Arenaria interpres*; AY074885), southern black-backed gull (*Larus dominicanus*, AY293619), Oriental stork (*Ciconia boyciana*; AB026193), red-throated loon (*Gavia stellata*; AY293618), little blue penguin (*Eudyptula minor*; AF362763), black-browed albatross (*Diomedea melanophris*; AY158677) and Kerguelen petrel (*Pterodroma brevirostris*; AY158678). The seven paleognath taxa are: emu (*Dromaius novaehollandiae*; AF338711), southern cassowary (*Casuarius casuarius*; AF338713), great spotted kiwi (*Apteryx haastii*; AF338708), greater rhea (*Rhea americana*; Y16884), ostrich (*Struthio camelus*; Y12025), great tinamou (*Tinamus major*; AF338707) and elegant crested tinamou (*Eudromia elegans*; AF338710). The New Zealand moa (Cooper et al. 2001, Haddrath and Baker 2001) were omitted from the analyses for reasons discussed in Slack et al. (2007), but do not affect this study. The issue of fine-tuning paleognath interrelationships will be readdressed once additional kiwi sequences become available (G. C. Gibb in preparation).

3.3.1 Phylogenetic analysis

Sequences were aligned in SeAl v2.0a11 (Rambaut 1996), at the amino acid level for protein-coding genes, and based on stem and loop secondary structure for RNA genes. The dataset has 12 protein-coding genes, two rRNAs and 21 tRNAs (lacking tRNA Phe). Gaps, ambiguous sites adjacent to gaps, the ND6 (light-strand encoded), and stop codons (often incomplete in the DNA sequence), were excluded

from the alignment. The full dataset had 13,139 base pairs, and the Neoaves-only dataset had 13,323 base pairs.

In previous work (Phillips and Penny 2003; Delsuc et al. 2003; Harrison et al. 2004; Phillips et al. 2006) we found that RY-coding of the most variable partitions of the nucleotide data (specifically the 3rd codon position) was advantageous. The recoding increases the proportion of observable changes on internal branches of the tree (treeness) and decreases the differences in nucleotide composition (Relative Compositional Variability, RCV). It also increases concordance between mitochondrial and nuclear datasets. RY-coding does increase the ML scores, but because RY-coding has amalgamated some nucleotide categories, the data is now different and it is not valid to compare directly the RY and nucleotide ML scores (M. A. Steel, pers. comm.). However, because of the better fit of the data to the model (higher treeness, and less variability in nucleotide composition (lower RCV), this is our preferred method of analysis of vertebrate mitochondrial data. Thus the trees reported here have the third codon positions recoded as R or Y. The full data set is available from our website http://awcmee.massey.ac.nz/downloads.htm.

Maximum likelihood analysis in PAUP*4.0b10 (Swofford 1998) used likelihood settings from the best-fit model (Transversional model, TVM+I+G, both transition classes are treated equally) selected by both hierarchical and AIC tests in Modeltest 3.7 (Posada and Crandall 1998). Preliminary results have shown current species fall into the expected three groups paleognaths, Galloanseres and Neoaves (data not shown). Therefore for further analyses we constrained the tree to these three groups as this drastically reduces analysis time (329 hours reduced to 186 hours for ML analysis of 40 birds). MP bootstrap analysis with 1000 bootstraps was also carried out (data not shown). For MrBayes (Huelsenbeck and Ronquist 2001) analysis, the data was partitioned into five character sets (1st codon, 2nd codon, 3rd codon with RY coding, RNA stems and RNA loops, as in Harrison et al. [2004]), unlinked (except for topology) and run for 10^7 generations. Sampling of the MCMC chain was assessed with Tracer v1.4 (Rambaut and Drummond 2003) and consensus networks (Holland et al. 2005) of MrBayes results were constructed with SplitsTree v4.3 (Huson and Bryant 2006).

3.4 RESULTS

The seven new mitochondrial genome sequences have been deposited in GenBank under the following accession numbers: Ivory billed Aracari (*Pteroglossus azara*: DQ780882; 18,736bp); pileated woodpecker (*Dryocopus pileatus*: DQ780879; 16,832bp); osprey (*Pandion haliaetus*: DQ780884); forest falcon (*Micrastur gilvicollis*: DQ780881; 17,344bp); American kestrel (*Falco sparverius*: DQ780880; 17,507bp); white faced heron (*Ardea novaehollandiae*: DQ780878; 17,511bp); and Australian pelican (*Pelecanus conspicillatus*: DQ780883; >16,846bp [incomplete]).

Because of the potential utility of 'rare genomic changes' we first describe the gene orders in these seven birds, and give a model for the mode of transition between them. We will return to the significance of these findings in the discussion and provide explanation for the apparent high frequency of the mitochondrial gene order changes, which reduce the phylogenetic utility of this potentially highly informative data. The heron has the gene order first identified in the chicken (Desjardins and Morais 1990), while the forest falcon, kestrel and woodpecker all have the alternative gene order where the control region (CR) lies between tRNAs Thr and Pro, and a second, unalignable and often shorter non-coding region lies between tRNAs Glu and Phe (Figure 3.1e) that was first identified by Mindell et al. (1998). The non-coding regions in the forest falcon and kestrel, much like the peregrine falcon (*Falco peregrinus,* Mindell et al. 1998), is mostly repeats of a short sequence - a 4bp repeat in the forest falcon and a 9bp sequence in the kestrel. Both birds also have longer repeat sequences at the end of the first control region. The short woodpecker non-coding region has neither discernable repeats nor any similarity to the woodpecker control region.

The osprey and aracari both have the gene order previously described only in *Amazona* parrots (Eberhard et al. 2001) where the control region is duplicated and the repeated control regions lie between tRNAs Thr and Pro, and Glu and Phe (Figure 3.1d). This is different to the gene order found in the falcon, as the two control regions are clearly duplicates and are easily alignable to each other. In both species, it is striking that the two control regions are nearly identical; differing only in the 5' and 3' ends (Figure 3.2). In the aracari, 1230bp are 100% identical between the two control regions, including a 90bp repeat sequence at the 5'end. This 90bp repeat occurs 6 or 7 times in the first control region, followed by a 71bp truncated

repeat (different clones contained different numbers of repeats). In the second control region, the 90bp repeat sequence is repeated 4 times followed by a 15bp truncated repeat; then a 14bp sequence repeated 7 or 8 times. The long repeats contribute to the genome length of 18,736bp. In contrast, the osprey contains no repeat sequences, but still has 99.2% similarity over 929bp between the two control regions. Neither the aracari, nor the osprey has identifiable remnants from ND6, nor tRNA Glu repeats, as was found in *Amazona* parrots (Eberhard et al. 2001).

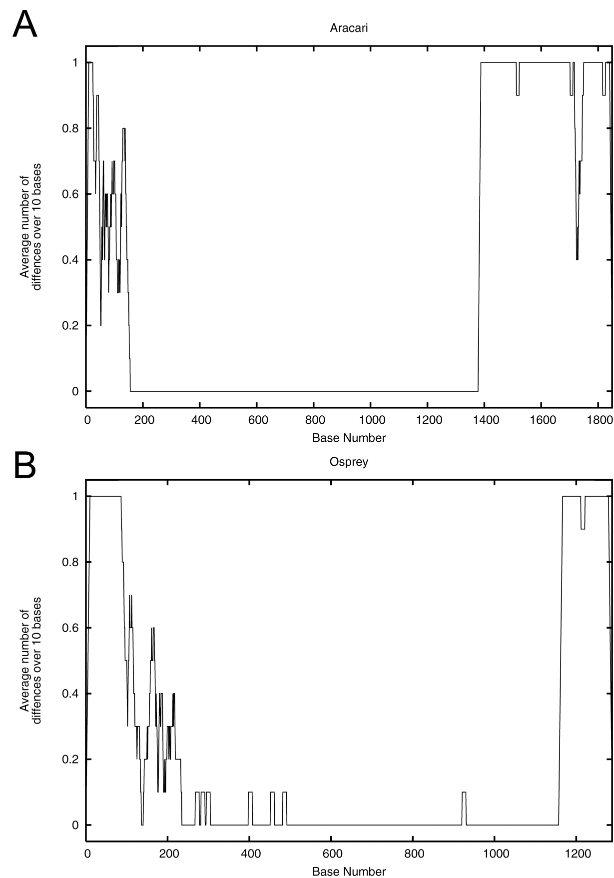

FIGURE 3.2—Comparison between CR(1) and CR(2) in Aracari (*A*) and Osprey (*B*) using the average number of differences in a sliding window of 10 bases.

Abbott et al. (2005) reported that *Thalassarche* albatrosses have a duplicated region from tRNA Thr to the control region (Figure 3.1c). We have re-checked the *Diomedea* albatross sequence (*Diomedea melanophris*; AY158677) reported in Slack et al. (2007), and have identified a duplicate region in this species as well. Because three tRNA's plus ND6 are duplicated, as well as the control region, it is possible to miss the duplicated region using standard primer pairs (see Figure 3.3 and Discussion). The revised genome length for the albatross is now 18,967bp, the longest avian mitochondrial genome reported so far. The duplicated segments are

nearly identical, beginning with a 100% match for the last 51 bases of cyt*b,* followed by Thr/Pro/ND6/Glu/CR. The control regions differ by 21 mismatched bases near the start, and the last 114 bases of CR(1) are unalignable, as CR(2) ends with a 22bp sequence repeated 15 times. The pelican may also have the duplicate tThr-CR gene order (based on a sequence fragment containing CR/Thr/Pro). However, the region between CR(1) and tRNA Phe is currently incomplete, so we cannot rule out a nuclear mitochondrial copy (numt), or another gene order for this region.

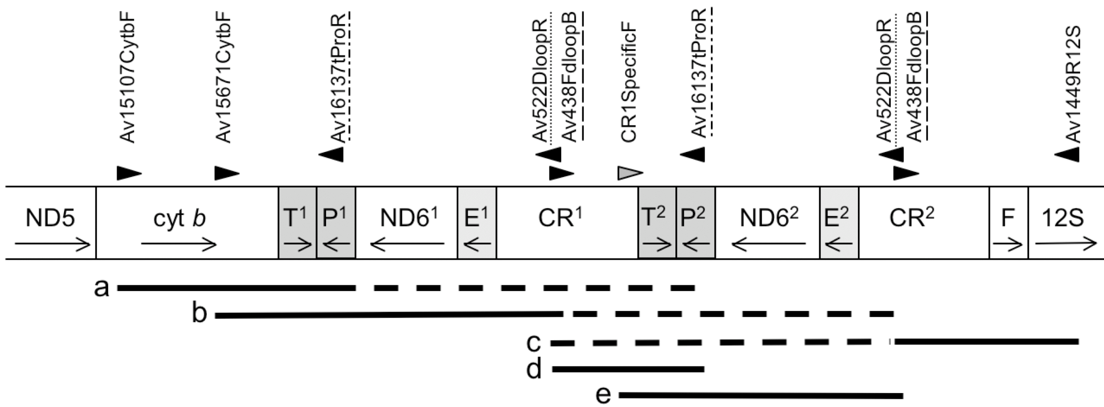

FIGURE 3.3—Primers used to sequence duplicate tThr-CR gene order. Arrows show the direction of the primer, and numbers in primer names refer to location relative to the chicken genome. Primers that bind twice are underlined with matching lines. Black lines labeled a-e are PCR products; dashed lines indicate longer unseen PCR products. Products a, b and c can be aligned to completely miss the gene duplication found by d and e.

The number of different mitochondrial gene order rearrangements in birds currently stands at four. At this point, names such as ' standard', 'normal', 'alternative', 'novel' and 'albatross' gene order loose their meaning, so a new naming system is required. Currently only the standard avian gene order exists for paleognaths and galloanseres (Figure 3.1b), so it is logical to assume this was the ancestral gene order at the root of the Neoaves. This order is only one rearrangement away from the presumably ancestral gene order found in many reptiles (Figure 3.1a). The three other orders require at least two rearrangements from the ancestral reptilian gene order. We refer to the order first found in the chicken (Desjardins and Morais 1990) as 'ancestral avian', the order first described in the falcon (Mindell et al. 1998) as 'remnant CR(2)', the order first described in the Amazona parrots (Eberhard et al. 2001) as 'duplicate CR' and the order first described in the albatross (Abbott et al. 2005) as 'duplicate tThr-CR'. This proposal provides a systematic framework that allows the naming of any additional gene

orders that might be discovered, e.g. 'duplicate ND6-CR', or 'remnant CR(1)'. Using the term 'remnant' can imply either a reduction from a full control region, or a left-over part when the control region moved from one location to another. Either scenario is possible for the falcon (for example), so the name should not imply one over the other. We prefer this to 'pseudo' or 'non-coding' region, as the remnant control region has been called in the past (Mindell, et al. 1998; Haring et al. 2001). In addition, duplicate sequences are labeled (1) and (2) from heavy strand 5′ to 3′ for ease of notation, even if this may imply CR(1) duplicated from CR(2) (see Discussion). These different arrangements have been found in all parts of the Neoavian tree, and do not uniquely define specific clades in the tree (see Figure 3.4).

3.4.1 Phylogenetic Analysis

We will return later to the significance of the gene order finds, but next report the maximum likelihood tree for the 40 bird sequences. Slack et al. (2007) reported that the improved taxon sampling has stabilized the root of the avian tree; as predicted earlier (Braun and Kimball 2002; Garcia-Moreno et al. 2003) there is now agreement between nuclear, mitochondrial and morphological data. We have run the current dataset with a reptilian outgroup, and the position of the root is again between neognaths and paleognaths (data not shown). However, excluding the more distantly related reptile species allows us to increase the total number of nucleotide positions from approximately 11,500 to 13,000. Therefore we use the paleognaths as the outgroup to root the neognath tree. Figure 3.4 clearly resolves into the three main groups: paleognaths, galloanseres and Neoaves.

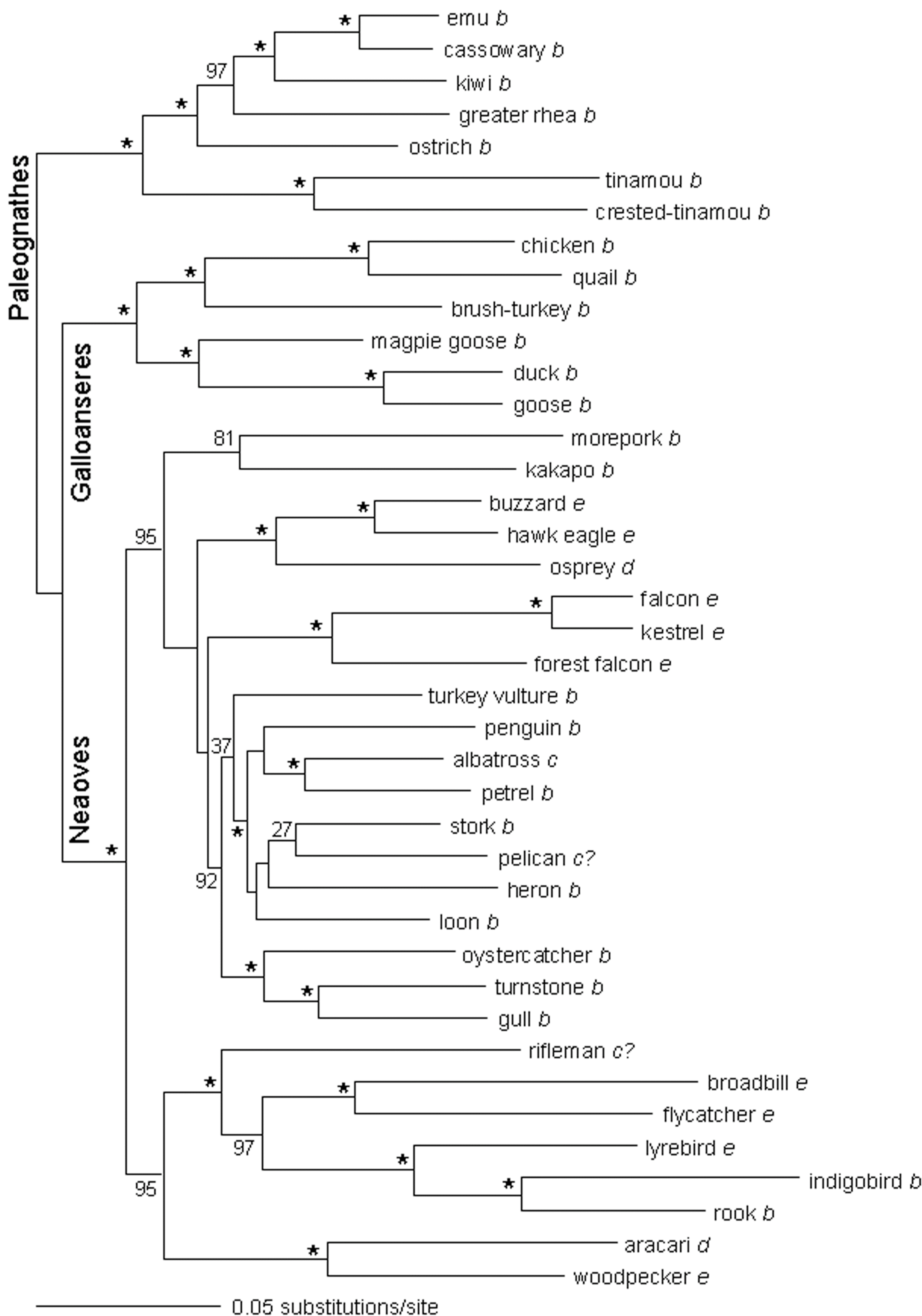

FIGURE 3.4—Maximum Likelihood tree of 40 birds. Asterisks represent 100% support in Bayesian MCMC analysis in MrBayes, with support over 25% shown. The letters in italics after each species name refer to gene orders shown in Figure 3.1.

Having confirmed that including the seven new Neoavian species does not lead to any unexpected effects, we are able to address the question of resolution within Neoaves. The groupings of paleognaths, galloanseres and of Neoaves were constrained for further analyses in PAUP*. We add two members of the fifth Cracraft (2001) group; an aracari and a woodpecker, intending to reduce problems of long-branch attraction that have hindered the placing of the morepork, kakapo and falcon. As can be seen from Figure 3.4, the aracari and the woodpecker pair together, and join basal to the Passeriformes. The fifth Cracraft (2001) group has long been placed in a hypothetical grouping called the "higher land birds" that contains the passerines (e.g. Mayr et al. 2003). Although the exact set of avian taxa that should be included in this higher land bird group is unclear and some suggestions conflict with the Cracraft (2001) six-way split, the proposed higher land bird group suggests that finding Piciformes sister to passerines should not be viewed as unexpected. In fact, these results could be viewed as validating earlier morphological work suggesting that Piciformes and passerines are related (e.g. Shufeldt 1900; Livezey and Zusi 2001). Neither maximum likelihood nor Bayesian analyses show any conflict in this placing. Given the congruence with other data, this result provides evidence at least some of Cracraft's (2001) six-way split can be resolved, and is initial evidence against any 'explosive radiation' hypothesis (e.g. Poe and Chubb 2004).

The kakapo and the morepork (a parrot and an owl) fall between the aracari/woodpecker/passerine grouping and the 'Conglomerati'. While they currently come together, we expect this is in part because of long-branch attraction, because both are long isolated branches. We are currently sequencing a lovebird and a barn owl to help resolve this grouping. It is interesting to note, that when the parrot is omitted, the owl moves more towards the aracari/woodpecker/passerine grouping (MrBayes analysis, data not shown). Conversely, in the absence of the owl, the parrot (kakapo) is closer to the Conglomerati. There has been speculation based on morphological data that owls are closely related to the Falconiformes (for example, Mayr and Clarke 2003; though not in Livezey and Zusi 2001), however our molecular evidence does not appear to support this. The limited stability of the morepork and kakapo in the present tree combined with the very long branches involved suggests their current position may not reflect their phylogenetic position, so we feel it is best to defer judgment regarding these taxa until taxon sampling has been improved further.

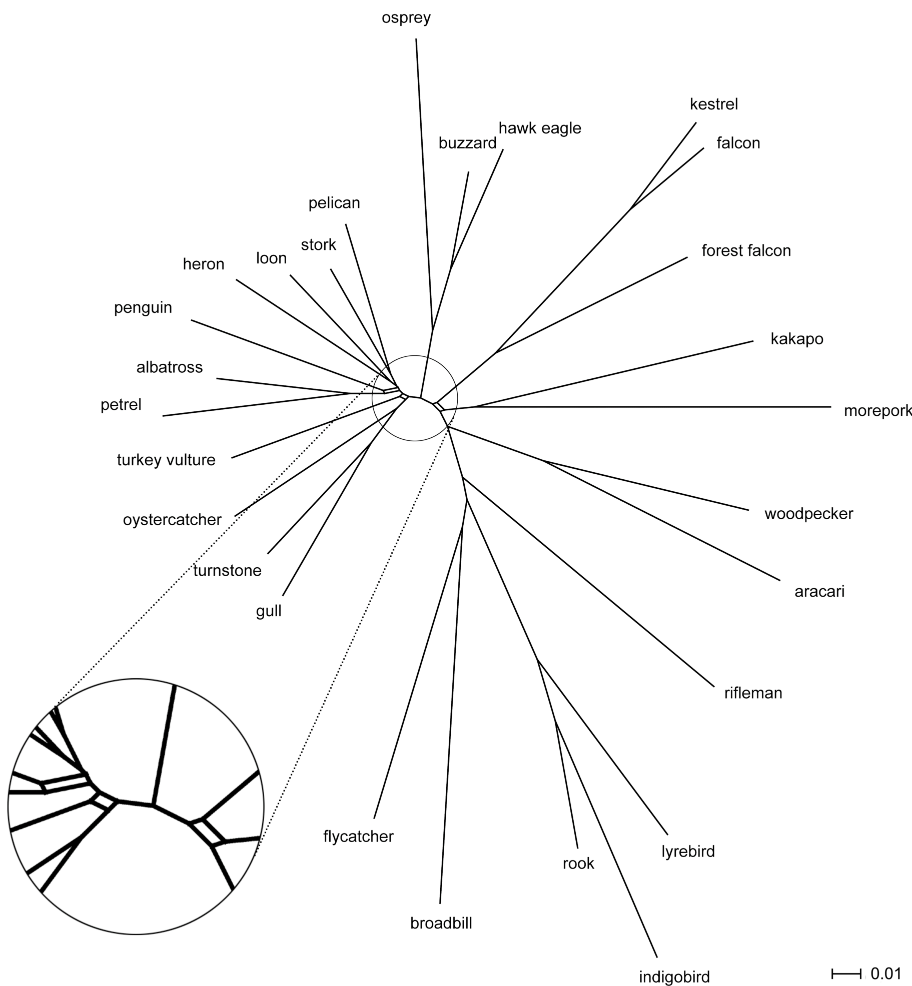

FIGURE 3.5—Consensus network for 27 Neoavian birds based on trees sampled by the Bayesian MCMC analysis using MrBayes. There is no conflict in splits with over 25% support; this network shows support greater than 20%.

Turning to the raptors, the falcon joins to the kestrel at a fairly shallow node, with the forest falcon joining deep on this branch, forming the Falconidae. The hawk eagle joins the buzzard with the osprey joining deeper on this branch, forming the Accipitridae. Both groups are well supported in our dataset, and are expected from previous studies (e.g. Mayr and Clarke 2003; Cracraft et al. 2004). The addition of these four species has helped to stabilize the position of the falcon and the buzzard, but interestingly have not joined the two groups into a strictly monophyletic group, though they may be paraphyletic. Considering the consensus network from MrBayes analysis (Figure 3.5), it appears the kakapo and morepork are still creating conflict for the Falconidae, pulling them out of the Conglomerati towards the passerine/Piciformes. When both the morepork and kakapo are removed from MrBayes analysis, releasing the Falconidae, they move further into the Conglomerati, and shorebirds (Charadriiformes) actually fall between Falconidae

and Accipitridae (data not shown). Although the positions of raptors and shorebirds are variable on a local scale on the tree, they all support the conclusion of Paton, Haddrath, and Baker (2002) that shorebirds are not deep in the avian tree – they are not 'transitional' to modern birds. At present we favor the suggestion that the tree is indicating an ancestral raptor group, members of which diverged towards new marine habitats.

The position of the turkey vulture is not completely resolved either. It falls within the seabird/shorebird part of the Conglomerati, rather than joining deep on the Accipitridae branch as predicted by some studies (Cracraft et al. 2004). However, the addition of further species in the seabird/shorebird group has shown that the turkey vulture also does not come within the Ciconiiformes (storks), as is sometimes suggested by current taxonomy (e.g. The American Ornithologists' Union Check-list, www.aou.org/checklist/index), and Sibley and Ahlquist (1990). Even though MrBayes analysis places the turkey vulture within the seabird/shorebird grouping with some conflict regarding its location, it does not fall closer to the stork than other species (Figure 3.5). Again it is a single isolated branch on the tree.

We have added two further members to the seabirds/shorebirds portion of the Conglomerati: a heron and a pelican. The most significant change is that the penguin/stork association (see Slack et al. 2006; 2007) was not seen in our analyses. The penguin (Sphenisciformes) and the albatross/petrel (Procellariiformes) are now united, though the relationship of these taxa to the loons (Gaviiformes) and other groups is not well defined (Figures 3.4 and 3.5). While it may seem surprising that the pelican and storks are relatively close, examining the positions of these species in Sibley and Ahlquist (1990) shows these were in fact probably the closest relatives present in the dataset (excluding the turkey vulture). In standard taxonomies, the heron and the stork should group together in Ciconiiformes, perhaps with the turkey vulture (discussed earlier), but separate from the loon, penguin and pelican. The MrBayes analysis still reveals conflict in the placement of these species, so it is unsurprising that sometimes these species group in slightly different conformations, depending on the analyses conducted. Now the heron has been added, the loon does not pair with the stork or penguin in either maximum likelihood or Bayesian analyses. However, there are very short internal edges on these branches suggesting that additional species,

improved methods and nuclear sequences could all help here. Despite discrepancies in positioning within this group, the group as a whole (albatross, petrel, penguin, stork, loon, heron and pelican) has 100% Bayesian support (Figure 3.4). As a caveat to this it is important to note that high Bayesian or bootstrap support does not necessarily mean the tree is 'right', and phylogenetic inferences need to take this into account (Phillips et al. 2004). There is still work to be done in untangling the deeper phylogenetic resolution within seabirds and shorebirds.

3.5 DISCUSSION

The most commonly suggested model for gene rearrangement involves duplication, followed by the reduction or loss of one copy (e.g. Bensch and Harlid 2000; Sano et al. 2005). While it is impossible to recreate exactly how the gene rearrangements took place, one scenario is that gene rearrangements began as the result of a duplication of tThr/tPro/ND6/tGlu/CR (to give Figure 3.1c) and that the duplicate tThr-CR, duplicate CR and remnant CR(2) are intermediates in the reduction of one of each of these duplicates (see Figure 3.1d and 1e). For example duplicate tThr-CR gene order shows little reduction in either copy; the duplicate CR order still has two control regions, but only one tRNA Thr, and small pseudo fragments of ND6(1) and tRNA Glu(1), and finally the remnant CR(2) gene order has only one copy of each gene-coding region, with CR(2) reduced to a short non-coding fragment. Continuing with this scenario, it would also be possible that the first duplication reduced again instead of the second, returning to the ancestral avian order. A different model for gene rearrangement is that gene region duplication is sloppy; sometimes duplicating the whole fragment tThr-CR, at other times just the control region, or something in between. This method would make each gene order the result of a different type of gene duplication. This second scenario requires the duplication to be re-inserted between tRNA Thr and tRNA Pro, rather than adjacent to the first copy. However gene rearrangement has occurred, it has still happened more than once. Each time a gene duplication or rearrangement has occurred, it could (in principle) have been by a different pathway.

Even more so than the *Amazona* parrots (Eberhard et al. 2001), the duplicate control regions are very similar in both the aracari and osprey. The aracari is

particularly striking, as the first 142bp of the two CRs do not align (about 50% of bases are mismatched), and then the following 1230bp are 100% identical (Figure 3.2). The last 90bp of CR(1) similarly does not align to CR(2), which has a 14bp sequence repeated 7 or 8 times. It is possible that this main central region reflects very recent gene duplication, with no time for mutations in one of the control regions. More likely, given our knowledge of concerted evolution, it is concerted evolution that has kept both copies identical. This could be tested by examining additional toucans, since the existence of multiple taxa with virtually identical control regions would be more parsimoniously explained by concerted evolution rather than multiple independent duplications, each without sufficient time for the duplicated sequences to diverge. However, even when our attention is limited to the aracari we observed that the first scenario (recent duplication) does not explain the few hundred mismatched bases at start and end of the control regions given that the average size of the intergenic spacer between tRNAs Thr and Pro in Neoaves is 6-14 bases (Slack et al. 2003). The second scenario (concerted evolution) would explain the unalignable nature of the ends of the control regions, which could be remnants accumulating as one control region replaced the other. Eberhard et al. (2001) also suggested concerted evolution when they showed that paralogous control regions were more alike than orthologous copies with nearest phylogenetic neighbors (the scenario we predict for toucans).

Many authors have suggested mechanisms that would cause gene rearrangements in a circular genome, although the exact mechanism is unknown. These include recombination, slipped-strand mispairing, errors in synchronizing the points of initiation and termination, and illicit priming of replication by tRNAs near the replication origin. (e.g. Mindell et al. 1998; Mueller and Boore 2005). The gene regions that are rearranged in birds are also predominantly coded by the heavy strand, and are near the origin of heavy strand replication. It is possible any combination of these mechanisms could be responsible for the gene orders seen in birds (Mueller and Boore 2005).

Bensch and Härlid (2000) also discuss how the control region gene rearrangements may have occurred, and suggest duplication of a region followed by multiple deletions. In their Figure 3.1c they show a hypothetical reconstruction of an intermediate stage between ancestral avian gene order and the remnant CR(2) gene order found in the willow warbler. This hypothetical intermediate shows a

duplication of tPro/ND6/tGlu/CR that is reduced to the remnant CR(2) gene order. This is very similar to the gene order that has now been found in albatross and possibly pelicans (except without the duplication of tRNA Thr). Even more interestingly, in preliminary work we have found this gene order in the Hihi (New Zealand stitchbird, *Notiomystis cincta*), probably a basal corvid (Meliphagidae - Honeyeater) (G. C. Gibb, in prep.).

Because of the duplicate nature of the tThr-CR repeat, it is possible to completely miss the 2nd duplication using standard primer pairs (see Figure 3.3). PCR primers will be more likely to amplify the shorter fragment, with the longer fragment going undetected. For example Cyt *b*-CR primer pairs may preferentially amplify CR(1), rather than all the way to CR(2) (Figure 3.3b), and CR-12S primer pairs may preferentially amplify CR(2), not the longer CR(1)-12S (Figure 3.3c). Additionally, because CR(1) and CR(2) can be nearly identical, it is possible to align the first part of CR(1) to the second half of CR(2), and miss an entire duplication. To correctly determine whether a duplication exists, primer pairs that have not traditionally been used are required, for example CR forward with tRNA Pro reverse (Figure 3.3d). A positive sequence result, crossing gene boundaries, will indicate the existence of a duplicated gene region (rather than just primers binding incorrectly). Of course, a negative PCR result does not completely disprove the existence of a duplicated region! We recommend adding this diagnostic primer combination when sequencing any avian genome. Currently we are checking all bird species sequenced in this lab for the possibility of previously undetected gene duplications.

Mindell et al. (1998) tested 137 bird species representing 13 orders for the two gene orders known in 1998 (ancestral avian and remnant CR(2)). It is not clear whether all DNA regions mentioned in their paper were analyzed, or just the gene regions shown with positive results in their Table 1. If all regions shown in their Figure 3.1 were tested in all their species, then we can discount the presence in those species of the other gene orders discovered since 1998. However, if DNA regions were tested sequentially, stopping after encountering a positive result, then gene orders such as duplicate tThr-CR or duplicate CR may have been overlooked, for the reasons discussed above (see Figure 3.3).

The significance of multi-state characters (such as gene order) in phylogeny is still being developed. Elementary logic indicates that shared character-states that are

genuinely unique must be informative for phylogeny and indeed, under such models parsimony is a maximum likelihood estimator (see Steel and Penny 2004; 2005). The positions of insertions/deletions (indels) in genes (as used in Fain and Houde 2004 and others, e.g. Kimball et al. 2001; Allen and Omland 2003; Kawakita et al. 2003) are potentially such unique characters. Others include gene order (Henz et al. 2005), gene fusions (for example, Stechmann and Cavalier-Smith 2002), the presence or absence of repetitive elements (e.g. SINES - short interspersed nuclear elements, Shedlock and Okada 2000, or LINES Kriegs et al. 2006). The use of such characters is reviewed in Rokas and Holland (2000), and Boore (2006) under the grouping of 'rare genomic changes'. In contrast, for primary sequence data the state-space is four for nucleotides (or two after RY coding) and 20 for amino acids characters, and it is both expected with such a small state-space that the same character state will arise multiple times and the standard maximum average likelihood is the preferred estimator (Steel and Penny 2000).

Thus there is a good theoretical basis for using rare genomic changes in the resolution of phylogeny, if the number of character states is so large that parallel changes and/or reversions are unlikely. However, these multistate characters are fraught with difficulty, and care will always be required to check for reversals. For example, Stechmann and Cavalier-Smith (2002) proposed an alternative with the rooting between (animals, fungi and choanazoa/choanaflagellates) and all other eukaryotes. This is based on a gene fusion between dihydrofolate reductase and thymidylate synthase genes. These genes are fused in most eukaryotes, but not in the animal/fungi/choanazoa group mentioned above. However, it is also known that reversals (fissions, when two genes end up separate) do occur (Snel et al. 2000) and so relying on a single gene fusion is risky. With SINEs, sequences around the insertion can be checked that the insertion is at precisely the same place (Shedlock and Okada 2000). We refer to such genomic characters with an extremely large state space as SCUDS (sequence-characters, uniquely-derived). They may be extremely effective, but used carelessly that can be highly damaging to the user! If there is a genuine very large character space, then they can be highly effective for phylogenies. Waddell et al. (2001) concluded that 3 SINEs were able to give 95% confidence limits to polytomies of three taxa, even when groups were so closely related that lineage sorting was the difficulty with sequence data. In the future it will be very interesting to integrate SCUDS with likelihood values from sequence data (Steel and Penny 2005).

On a more positive note, although mitochondrial gene order in birds is not yet useful as a multi-state character, it appears that the Neoavian radiation will be resolvable, ending the suggestion that it represents an 'explosive radiation' (e.g. Poe and Chubb 2004). Currently we are not getting the raptors as a monophyletic group. Although this may be unfortunate for a taxonomist who might like organisms in neatly arranged boxes, it may be of more importance to an ecologist or evolutionary biologist. The implication from Figures 4 and 5 is that there was an early group of raptors (in the Late Cretaceous to fit with the timing from Slack et al. 2007) from which a variety of other carnivore groups have adapted to a more aquatic environment. The present dataset, with seven raptors and ten sea and shorebirds, shows this clearly. Similarly, grouping the woodpecker/aracari clade (Piciformes) with the passerines is consistent with expectation based upon prior data (e.g. Mayr et al. 2003), and indicates that the grouping reflects evolutionary history. This suggests resolution of Cracraft's six Neoaves groups and rigorous testing of the monophyly of those groups is possible. This is strong evidence that there is not an irresolvable six way split at the base of Neoaves. The addition of the sixth group (cuckoos) along with taxa that can subdivide the long terminal edges corresponding to the morepork (owl) and kakapo (parrot) are likely to further resolve the avian tree. Since an evolutionary tree is not an end in itself but a guide toward answering biologically significant questions, we assert that the present tree provides evidence against an 'explosive radiation' at the base of the Neoaves, and suggests that birds with a terrestrial raptor phenotype may be ancestral to a wide range of other carnivores, especially marine carnivores. This tree just represents a starting point for biological studies, and further resolution of the tree will increase the insights from those biological studies.

3.6 ACKNOWLEDGMENTS

We thank Donna Dittmann for the forest falcon and aracari samples, Denis O'Meally for the osprey, Dick Gill for the heron, Les Christidis and Martyn Kennedy for the pelican and Laura Patterson for work on the woodpecker and kestrel genome sequences. Work at the Allan Wilson Centre for Molecular Ecology and Evolution was supported by a Marsden grant to David Penny. Gillian Gibb is supported by an Allan Wilson Centre Doctoral scholarship. Work at the University

of Florida was facilitated by National Science Foundation funding (grant DEB-0228682 to E.L.B., Rebecca T. Kimball and David W. Steadman).

3.7 REFERENCES

Abbott C, Double MC, Trueman JWH, Robinson A, Cockburn A. 2005. An unusual source of apparent mitochondrial heteroplasmy: duplicate mitochondrial control regions in Thalassarche albatrosses. Mol. Ecol. 14:3605–3613.

Allen ES, Omland KE. 2003. Novel intron phylogeny supports plumage convergence in orioles (Icterus). Auk 120:961-970.

Anderson FE, Swofford DL. 2004. Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rRNA. Mol. Phylogenet. Evol. 33:440-451.

Ane C, Burleigh JG, McMahon MM, Sanderson MJ. 2005. Covarion structure in plastid genome evolution: a new statistical test. Mol. Biol. Evol. 22:914-924.

Bensch S, Härlid A. 2000. Mitochondrial genomic rearrangements in songbirds. Mol. Biol. Evol. 17:107-113.

Boore JL. 2006. The use of genome-level characters for phylogenetic reconstruction. Trends Ecol. Evol. 21:439-446

Braun EL, Kimball RT. 2002. Examining basal avian divergences with mitochondrial sequences: Model complexity, taxon sampling, and sequence length. Syst. Biol. 51:614–625.

Cooper A, Lalueza-Fox C, Anderson S, Rambaut A, Austin J, Ward R. 2001. Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. Nature 409:704–707.

Cooper A, Penny D. 1997. Mass survival of birds across the Cretaceous-Tertiary boundary: Molecular evidence. Science 275:1109–1113.

Cracraft J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proc. Roy. Soc. Lond. B 268:459–469.

Cracraft J, Baker FK, Braun M, et al. (14 co-authors) 2004. Phylogenetic relationships among modern birds (Neornithes): toward an avian tree of life.

In: Cracraft J and Donoghue MJ, editors. Assembling the Tree of Life. New York: Oxford University Press. p 468-489.

Delsuc F, Phillips MJ, Penny D. 2003. Comment on "Hexapod origins: monophyletic or paraphyletic?" Science 301:1482d.

Desjardins P, Morais R. 1990. Sequence and gene organization of the chicken mitochondrial genome - a novel gene order in higher vertebrates. J. Mol. Biol. 212: 599-634.

Eberhard JR, Wright TF, Bermingham E. 2001 Duplication and Concerted Evolution of the Mitochondrial Control Region in the Parrot Genus Amazona. Mol. Biol. Evol. 18:1330–1342.

Fain MG, Houde P. 2004. Parallel radiations in the primary clades of birds. Evolution 58:2558-2573.

Garcia-Moreno J, Sorenson MD, Mindell DP. 2003. Congruent avian phylogenies inferred from mitochondrial and nuclear DNA sequences. J. Mol. Evol. 56:1–11.

Groth JG, Barrowclough GF. 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. Mol. Phylogenet. Evol. 12:115-123.

Haddrath O, Baker AJ. 2001. Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis. Proc. R. Soc. Lond. B 268:939–945.

Haring E, Kruckenhauser L, Gamauf A, Riesing MJ, Pinsker W. 2001. The complete sequence of the mitochondrial genome of *Buteo buteo* (Aves, Accipitridae) indicates an early split in the phylogeny of raptors. Mol. Biol. Evol. 18:1892–1904.

Harrison GL, McLenachan PA, Phillips MJ, Slack KE, Cooper A, Penny D. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the Late Cretaceous. Mol. Biol. Evol. 21:974-983.

Hendy MD, Penny D. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38:297–309.

Henz SR, Huson DH, Auch AF, Nieselt-Struwe K, Schuster SC. 2005. Whole-genome prokaryote phylogeny. Bioinformatics 21:2329-2335.

Hillis DM, Huelsenbeck JP, Cunningham CW. 1994. Application and accuracy of molecular phylogenies. Science 264:671-677.

Holland BR, Delsuc F, Moulton V. 2005. Visualizing conflicting evolutionary hypotheses in large collections of trees: Using consensus networks to study the origins of placentals and hexapods. Syst. Biol. 54:66-76.

Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference on phylogenetic trees. Bioinformatics 17:754–755.

Huson DH, Bryant D, 2006. Application of Phylogenetic Networks in Evolutionary Studies, Mol. Biol. Evol., 23:254-267.

Kawakita A, Sota T, Ascher JS, Ito M, Tanaka H, Kato M. 2003. Evolution and phylogenetic utility of alignment gaps within intron sequences of three nuclear genes in bumble bees (*Bombus*). Mol. Biol. Evol. 20:87-92.

Kimball RT, Braun EL, Ligon JD, Lucchini V, Randi E. 2001. A molecular phylogeny of the peacock-pheasants (Galliformes: Polyplectron spp.) indicates loss and reduction of ornamental traits and display behaviors. Biol. J. Linn. Soc. 73:187–198.

Kriegs JO, Churakov G, Kiefmann M, Jordan U, Brosius J, Schmitz J. 2006. Retroposed elements as archives for the evolutionary history of placental mammals. PloS Biology 4:e91.

Ligon JD. 1967. Relationships of the cathartid vultures. Occ. Pap. Univ. Michigan Mus. Zool. 651:1-26.

Lin Y-L, McLenachan PA, Gore AR, Phillips MJ, Ota R, Hendy MD, Penny D. 2002. Four new mitochondrial genomes, and the stability of evolutionary trees of mammals. Mol. Biol. Evol. 19:2060–2070.

Livezey BC, Zusi RL. 2001. Higher-order phylogenetics of modern Aves based on comparative anatomy. Neth. J. Zool. 51:1179–1205.

Lockhart PJ, McLenachan PA, Havell D, Glenny D, Huson D, Jensen U. 2001. Phylogeny, radiation, and transoceanic dispersal of New Zealand alpine buttercups: Molecular evidence under split decomposition. Ann. Missouri Bot. Gard. 88:458-477.

Mayr G, Clarke J. 2003. The deep divergences of neornithine birds: a phylogenetic analysis of morphological characters. Cladistics 19:527-553.

Mayr G, Manegold A, Johansson US. 2003. Monophyletic groups within 'higher land birds' – comparison of morphological and molecular data. J. Zool. Syst. Evol. Research 41:233–248.

Mindell DP, Sorenson MD, Dimcheff DE. 1998. Multiple independent origins of mitochondrial gene order in birds. Proc. Natl Acad. Sci. USA 95:10693–10697.

Mossel E, Steel MA. 2005. How much can evolved characters tell us about the tree that generated them? In: Gascuel O, editor. Mathematics of Evolution and Phylogeny. Oxford University Press. p 384-412.

Mueller RL, Boore JL. 2005. Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders. Mol. Biol. Evol. 22:2104-2122.

Nishihara H, Satta Y, Nikaido M, Thewissen JGM, Stanhope MJ, Okada N. 2005. A retroposon analysis of Afrotherian phylogeny. Mol. Biol. Evol. 22:1823-1833.

Nishihara H, Hasegawa M, Okada N. 2006. Pegasoferae, an unexpected mammalian clade revealed by tracking ancient retroposon insertions. Proc. Natl Acad. Sci. USA 103:9929–9934.

Paton T, Haddrath O, Baker AJ. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proc. Roy. Soc. Lond. B 269:839–846.

Penny D, Phillips MJ. 2004. The rise of birds and mammals: Are microevolutionary processes sufficient for macroevolution. Trends Ecol. Evol. 19:516-522.

Phillips MJ, Delsuc F, Penny D. 2004. Genome-scale phylogeny and the detection of systematic biases. Mol. Biol. Evol. 21:1455-1458.

Phillips MJ, Penny D. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. Mol. Phylogen. Evol. 28:171–185.

Phillips MJ, McLenachan PA, Down C, Gibb GC, Penny D. 2006. Combined mitochondrial and nuclear DNA sequences resolve the interrelations of the major Australasian marsupial radiations. Syst. Biol. 55:122-137.

Poe S, Chubb AL. 2004. Birds in a bush: five genes indicate explosive evolution of avian orders. Evolution 58:404-415.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Rambaut A. 1996. Se-Al: Sequence Alignment Editor. Available at http://evolve.zoo.ox.ac.uk

Rambaut A, Drummond AJ. 2003. Tracer version 1.2. Available at http://evolve.zoo.ox.ac.uk

Rokas A. Holland PWH. 2000. Rare genomic changes as a tool for phylogenetics. Trends Ecol. Evol. 15:454-459.

Sano N, Kurabayashi A, Fujii T, Yonekawa H, Sumida M. 2005. Complete nucleotide sequence of the mitochondrial genome of Schlegel's tree frog *Rhacophorus schlegelii* (family Rhacophoridae): duplicated control regions and gene rearrangements. Genes Genet. Syst. 80:213-224.

Schmidt TR, Wildman DE, Uddin M, Opazo JC, Goodman M, Grossman LI. 2005. Rapid electrostatic evolution at the binding site for cytochrome c on cytochrome c oxidase in anthropoid primates. Proc. Natl Acad. Sci USA 102:6379-6384.

Shedlock AM, Okada N. 2000. SINE insertions: powerful tools for molecular systematics. BioEssays 22:148-160.

Shufeldt RW. 1900. On the osteology of woodpeckers. Proc. Am. Philos. Soc. 39:578-622.

Sibley CG, Ahlquist JE. 1990. Phylogeny and classification of birds. Yale University Press, New Haven, Connecticut.

Slack KE, Janke A, Penny D, Arnason U. 2003. Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. Gene 302:43–52.

Slack KE, Delsuc F, McLenachan PA, Arnason U, Penny D. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol. Phylogenet. Evol. 42:1-13

Slack KE, Jones CM, Ando T, Harrison GL, Fordyce E, Arnason U, Penny D. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol. Biol. Evol. 23:1144-1155.

Snel B, Bork P, Huynen M. 2000. Genome evolution: gene fusion versus gene fission. Trends Gen. 16:9-11.

Snel B, Huynen MA, Dutilh BE. 2005. Genome trees and the nature of genome evolution. Ann. Rev. Microbiol. 59:191-209.

St. John J, Cotter J-P, Quinn TW. 2005. A recent chicken repeat 1 retrotransposition confirms the Coscoroba-Cape Barren goose clade. Mol. Phylogenet. Evol. 37:83-90.

Stechmann A, Cavalier-Smith T. 2002. Rooting the eukaryote tree by using a derived gene fusion. Science 297:89-91.

Steel MA, Penny D. 2005. Maximum parsimony and the phylogenetic information in multistate characters. In: V. Albert, editor. Parsimony, Phylogeny and Genomics. Oxford University Press. p 163-178.

Steel MA, Penny D. 2004. Two further links between MP and ML under the Poisson model. Appl. Math. Lett. 17:785-790.

Steel MA, Penny D. 2000. Parsimony, likelihood and the role of models in molecular phylogenetics. Mol. Biol. Evol. 17:839-850.

Swofford DL. 1998. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer Associates Sunderland, Massachusetts.

Waddell PJ, Kishino H, Ota R. 2001. A phylogenetic foundation for comparative mammalian genomics. Genome Informatics 12:141-154.

Watanabe M, Nikaido M, Tsuda TT, Inoko H, Mindell DP, Murata K, Okada N. 2006. The rise and fall of the CR1 subfamily in the lineage leading to penguins. Gene 365:57-66.

CHAPTER FOUR

4 TWO ASPECTS ALONG THE CONTINUUM OF PIGEON EVOLUTION: A SOUTH-PACIFIC RADIATION AND THE RELATIONSHIP OF PIGEONS WITHIN NEOAVES[1]

4.1 ABSTRACT

Phylogenetics reveals a continuum of shallower to deeper divergences, and along this continuum varying lengths of DNA sequence can be used to answer questions about biological relationships. We use shorter, and then longer mitochondrial DNA sequences to address two aspects of pigeon evolution along this continuum. Firstly, we examine the phylogenetic relationships of the eight genera within the South Pacific *Ducula–Ptilinopus* radiation, and compare this radiation with other pigeon clades. Within *Ducula*, taxa are closely related, whereas *Ptilinopus* is very diverse, and paraphyletic. One third of all pigeon species are within the *Ducula–Ptilinopus* radiation, however all are very similar ecologically. Secondly, we study the deeper question regarding the relationship of pigeons to other birds. To this end, we report the complete mitochondrial genome of *Hemiphaga novaeseelandiae*, a member of the *Ducula–Ptilinopus* radiation. We use it, along with additional sandgrouse (*Pterocles namaqua*) mitochondrial genes to assess various candidates for the closest relative of pigeons. Of parrots, shorebirds, and sandgrouse, we find highest support for the sandgrouse-pigeon grouping. Furthermore in these analyses the pigeon and sandgrouse group closer to the falcons than any other included taxon.

[1] This chapter submitted as: Gibb, G. C., and D. Penny. Two aspects along the continuum of pigeon evolution: a South-Pacific radiation and the relationship of pigeons within Neoaves. Mol Phylogenet Evol

The finding that pigeons and sandgrouse may be more closely related to falcons than to previous candidates such as shorebirds or parrots invites further investigation.

Keywords—Ptilinopus, *Ducula*, *Hemiphaga,* Columbiformes, sandgrouse, mitochondrial genomes, phylogeny

4.2 INTRODUCTION

This study addresses two main questions about pigeon phylogeny, using mitochondrial DNA sequences. Firstly we evaluate the phylogenetic radiation of pigeons in the Asia/Pacific region, dominated by the speciose genera *Ducula* (imperial pigeons) and *Ptilinopus* (fruit doves). Secondly we examine more closely the question of where pigeons fit within Neoaves. We especially investigate the closest purported relatives of pigeons, for which there have been many candidates. This is part of a larger study investigating broader questions of avian phylogeny (Gibb et al. 2007; Pratt et al. 2009).

4.2.1 A South East Asian/Pacific radiation of pigeons.

Ducula and *Ptilinopus*, together with six other near-monotypic genera *(Hemiphaga, Lopholaimus, Gymnophaps, Cryptophaps, Alectroenas,* and *Drepanoptila*), form a radiation of pigeons and doves that stretches from the Himalayas to Polynesia (Gibbs et al. 2001; Johnson 2004). This *Ducula–Ptilinopus* radiation includes approximately one third of all pigeon species (99 out of 312 species, Gibbs et al. 2001), all of which are arboreal and frugivorous. Almost all are island dwellers: only nine of the 99 species breed in mainland Asia or Australia.

Both *Ducula* and *Ptilinopus* have been successful in dispersing across the Pacific, with most islands having one or more species of each genus before human contact (Steadman 1997). Although many pigeon species have become locally or totally extinct since the arrival of humans (Steadman 2006), at least one (*Ducula pacifica*) appears to have successfully expanded its range through the Pacific since human colonization (Steadman 1997).

There are a few recent molecular studies of pigeons. None have included more than four *Ducula* or *Ptilinopus* species, and some only included one species per genus. Shapiro et al. (2002) considered particularly the relationships of the extinct dodo and solitaire to other pigeons. Johnson (2004) used short mitochondrial (mt) and nuclear sequences to examine the utility and evolution of indels in pigeon phylogeny. Pereira et al. (2007) used longer nuclear and mt sequences from just one species per genus to study biogeography and times of divergence. From the latter analysis, the huge radiation of *Ducula* and *Ptilinopus* can appear insignificant when compared to the six monotypic lineages that surround them.

4.2.2 The closest relative of Pigeons.

At a deeper phylogenetic level, the closest relative of pigeons among the Neoaves has been debated for decades, and is still unresolved. The favoured hypothesis based on morphology is sandgrouse (Pteroclidae) as the nearest relative of pigeons (e.g. Livezey and Zusi 2007), and there is some molecular support for this (e.g. Paton et al. 2003). However parrots (Psittaciformes) and shorebirds (Charadriiformes), among others, have also been proposed (see Sibley and Ahlquist 1990 for discussion). The issue is particularly difficult because the sandgrouse's specialized adaptation to arid environments may be confusing morphological comparisons (Sibley and Ahlquist 1990).

Two analyses of shorebirds using the nuclear gene RAG-1 favoured a sandgrouse/ pigeon pairing (Ericson et al. 2003; Paton et al. 2003), however support was weak (38–67% Bayesian support) and only a few other species were included for comparison. Analysis by Shapiro et al. (2002) investigating potential pigeon outgroups found highest support for shorebirds as sister to pigeons, although again support was not strong. Sandgrouse, pigeons and shorebirds are thought to be part of the radiation we refer to as the 'Conglomerati' (e.g. Cracraft 2001; Gibb et al. 2007). In contrast, a recent morphological study (Livezey and Zusi 2007) places the sandgrouse inside Columbiformes, in the superorder Psittacimorphae with parrots.

4.2.3 Nuclear versus Mitochondrial DNA and the problem of introns in deep phylogeny

Even the large avian study of Hackett et al. (2008) has not solved the relationship of pigeons. That tree places pigeons near the base of Neoaves with a group of other

unsupported lineages including sandgrouse, mesites, tropicbirds and the flamingo/grebe clade. They did not assess which genes in their dataset provided signal that grouped these species together. 74% of the Hackett dataset is comprised of introns (2008), which at this phylogenetic depth can be quite variable in length, and are therefore hard to align accurately (Wong et al. 2008; Liu et al. 2009). Morgan-Richards et al. (2008) showed that at deep divergences, introns with many gaps can provide overriding signal for an incorrect tree, such as the Metaves/Coronaves split found by Fain and Houde (2004).

An advantage of using mitochondrial DNA for deep phylogeny is that a high proportion of it is coding regions that are straightforward to align. A conservative alignment of deeply divergent birds containing no gaps can be 13 Kb in length. In contrast, the 32 Kb alignment of Hackett et al. (2008) reduced to 9.5 Kb when gaps were removed, and the 2 Kb β-fibrinogen intron 7 alignment of Fain and Houde (2004) reduces to zero when our stricter alignment criteria are used (Morgan-Richards et al. 2008). In the longer term, we need to compare the results of studies using diverse datasets such as nuclear, mitochondrial and morphological characters so that we can evaluate where they agree and use sources of disagreement for further study. This allows us to test quantitatively the agreement between trees (Pratt et al. 2009). Far more confidence can be placed on results when there is agreement between datasets—in our case nuclear and mitochondrial. In general, as taxon sampling improves there is good agreement between nuclear (exon) and mitochondrial trees (Slack et al. 2007).

To summarise, this study combines the available short pigeon mt DNA sequences from previous diverse studies with additional new sequences, particularly focusing on the Asia/Pacific *Ducula–Ptilinopus* radiation. We examine the phylogenetic relationships, ecology and success of this radiation compared to other pigeons, and in the context of their South East Asia/Pacific environment. We report the first complete mitochondrial genome of a pigeon (*Hemiphaga novaeseelandiae*—from the *Ducula–Ptilinopus* group) and use this longer mt sequence, along with the newly completed sandgrouse mt genome, *Pterocles namaqua*, as a further step in more extensive and ongoing studies of avian phylogeny (see Gibb et al. 2007; Pratt et al. 2009). The intention of the present analysis is not to resolve Neoaves in its entirety, but rather to shed further light on the relationship of pigeons with other birds, particularly the sandgrouse. This allows us to select new taxa and genes for

further analyses. We evaluate the various hypotheses for the closest relative of pigeons, using complete mitochondrial genome analyses and quantify the increase in knowledge provided by this study.

4.3 MATERIALS AND METHODS

4.3.1 Taxon sampling

Columba vitiensis, Ptilinopus victor and *P. luteovirens* were feather samples from the University of the South Pacific, Suva, Fiji. *P. solomonensis* (EBU13397L), *P. rivoli* (EBU30251L), *P. regina* (EBU47818M), *P. magnificus* (EBU10237L), *Ducula melanochroa* (EBU30294L), *D. rubricera* (EBU30289L), and *Lopholaimus antarcticus* (EBU45523M) were tissue samples from the Australian Museum, Sydney Australia. The Waikanae office of the New Zealand Department of Conservation provided the *Hemiphaga novaeseelandiae* tissue sample.

4.3.2 DNA extraction and PCR

DNA extraction from feather tips was by standard SDS/proteinase K digestion and phenol-chloroform extraction (Sambrook et al. 1989). DNA extraction from all other tissue samples used the High Pure PCR template preparation kit (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instructions. All short PCR products (between 0.5 kb and 3 kb) were amplified using primers identified from our database (available from the authors on request) employing standard protocols (e.g., see Gibb et al. 2007).

The complete mitochondrial genome of the New Zealand pigeon was initially amplified in two overlapping segments 10kb and 12kb long using the Expand Long Template PCR System (Roche). These long-range PCR products were used as templates for subsequent PCR of short overlapping fragments. This process is described in more detail in Gibb et al. (2007) and references therein. All protein coding, 12S and 16S sandgrouse (*Pterocles namaqua*) sequences used in our analyses were previously published in Paton and Baker (2006). Additional unpublished tRNA sequences were kindly provided by Tara Paton (Toronto), and had been sequenced according to the methods described in Paton and Baker (2006).

Sequences were aligned in SeAl v2.0a11 (Rambaut 1996), at the amino acid level for protein-coding genes, and based on stem and loop secondary structure for RNA genes. Gaps, ambiguous columns adjacent to gaps, the ND6 gene (light-strand encoded), non-coding regions and stop codons (often incomplete in the DNA sequence), were excluded from the alignments. The alignments are available from our website http://awcmee.massey.ac.nz/downloads.htm.

4.3.3 Phylogenetic analyses

All available mitochondrial sequences from pigeons in the *Ducula–Ptilinopus* clade were used. In addition, further mt sequences representing all general pigeon groupings identified in previous studies were selected from GenBank. This was done to confirm predictions about where the *Ducula–Ptilinopus* clade sits within pigeons, and where the root of the tree within pigeons joins with other birds. The final dataset of 49 pigeon species was assembled from five previously published sources, and includes sequences from eleven additional species. Novel sequences from this study are deposited in GenBank under accession numbers GU230684-731, and accession numbers for all species and genes used in these analyses are listed in Table 4.1.

TABLE 4.1—Accession numbers and mt genes included in analyses of pigeon phylogeny.

Species	12S	16S	ND1	ND2	COI	Cyt *b*
Alectroenas madagascariensis	AF483307					AF483344
Caloenas nicobarica	EF373289			EF373326	EF373363	AF483336
Chalcophaps stephani	EF373293			EF373328	EF373365	AY443673
Claravis pretosia	EF373294			EF373329	EF373366	AF182682
Columba livia	EF373295			AF353433	DQ432860	AF182694
Columba vitiensis	GU230684	GU230684	GU230685	GU230688	GU230686	GU230687
Columbina squammata	EF373296			EF373330	EF373368	AF483347
Drepanoptila holosericea	AF483308					AF483345
Ducula aenea	AF483294					AF483331
Ducula bicolor						AF182705
Ducula melanochroa	GU230689	GU230689	GU230689	GU230691	GU230692	GU230690
Ducula pacifica						AY443667
Ducula pinon	AF483295					AF483332
Ducula pistrinaria						AY443669
Ducula rubricera	GU230697	GU230697	GU230696	GU230693	GU230695	GU230694
Ducula rufigaster	EF373297			EF373331	EF373369	EF373277
Ducula zoeae	AF483296					AF483333
Gallicolumba jobiensis	EF373298			EF373332	EF373370	EF373278
Geopelia striata	EF373299			EF373333	EF373371	EF373279
Geotrygon montana	EF373301			EF373335	EF373373	AF182696
Goura cristata	EF373302			EF373336	EF373374	AF182709

Species	12S	16S	ND1	ND2	COI	Cyt *b*
Gymnophaps albertisii	EF373303			EF373337	EF373375	EF373280
Hemiphaga chathamensis	GXXXXXX	GXXXXXX	GXXXXXX	GXXXXXX	GXXXXXX	GXXXXXX
Hemiphaga novaeseelandiae	EU725864	EU725864	EU725864	EU725864	EU725864	EU725864
Leptotila rufaxilla	EF373306			EF373340	EF373378	AF182698
Lopholaimus antarcticus	GU230702	GU230702	GU230701	GU230698	GU230700	GU230699
Macropygia amboinensis	EF373309			EF373343	EF373381	EF373283
Oena capensis	EF37331			EF373345	EF373383	AF483353
Otidiphaps nobilis	EF373312			EF373346	EF373384	AF483352
Patagioenas leucocephala	AY274023		AY274070	AY274070		AY274041
Patagioenas speciosa	EF373313			EF373347	EF373385	AF279711
Ptilinopus leclancheri						AF182708
Ptilinopus luteovirens	GU230707	GU230707	GU230706	GU230703	GU230705	GU230704
Ptilinopus magnificus	GU230712	GU230712	GU230711	GU230708	GU230710	GU230709
Ptilinopus melanospila	AF483291					AF483328
Ptilinopus occipitalis	AF483293				AF279740	AF483330
Ptilinopus pulchellus	EF373317			EF373351	EF373389	EF373285
Ptilinopus rarotongensis						AY443663
Ptilinopus regina	GU230713	GU230713	GU230713	GU230716	GU230714	GU230715
Ptilinopus richardsii						AY443664
Ptilinopus rivoli	GU230721	GU230721	GU230720	GU230717	GU230719	GU230718
Ptilinopus solomonensis	GU230722	GU230722	GU230723	GU230726	GU230724	GU230725
Ptilinopus superbus	AF483292					AF483329
Ptilinopus victor	GU230731	GU230731	GU230730	GU230727	GU230729	GU230728
Streptopelia capicola	EF373319			EF373353	EF373391	AF279709
Streptopelia chinensis	AF483304			AF353431	DQ434172	AF483341
Treron calva	EF373320			EF373354	EF373392	AY443674
Treron sieboldii	AY274024		AY274071	AY274071		AY274042
Zenaida macroura	EF373325	U55844		EF373359	EF373397	AF182703

Not all genes are available for all species, thus the total sequence lengths vary, ranging from 1,426 bp to 6,013 bp. To accommodate this, the data was analysed using a partitioned MrBayes analysis (Huelsenbeck and Ronquist 2001). The three partitioned character sets were A: characters present for the fewest species (mostly 16S, NADH1 and 3′ COI, 2,788 bp), B: characters present for most species (part of 12S, ND2 and 5′ COI, 1,799 bp), and C: characters present for all species (remainder of 12S and cyt *b* 1,426 bp). Each was analysed using the GTR model with a proportion of invariable sites and four-category gamma-distributed rate variation (GTR+G+I). Data were unlinked except for topology and run for 10^7 generations. Sampling of the MCMC chain was assessed with Tracer v1.4 (Rambaut and Drummond 2003), and consensus networks (Holland et al. 2005) of MrBayes

results were constructed with SplitsTree v4.8 (Huson and Bryant 2006). The pigeon dataset was also analysed using maximum likelihood with 100 bootstrap replicates in GARLI (Zwickl 2006).

TABLE 4.2—Accession numbers and species names for birds used in complete mt genome analyses.

Scientific Name	Common Name	Accession number
Alectura lathami	Australian Brush-Turkey	AY346091
Thalassarche melanophris	Black-Browed Albatross	AY158677
Haematopus ater	Blackish Oystercatcher	AY074886
Melopsittacus undulatus	Budgerigar	EF450826
Gallus gallus	Chicken	AP003317
Apus apus	Common Swift	AM237310
Micrastur gilvicollis	Forest Falcon	DQ780881
Podiceps cristatus	Great Crested Grebe	AP009194
Phoenicopterus ruber roseus	Greater Flamingo	EF532932
Geococcyx californianus	Greater Roadrunner	EU410488
Pteroglossus azara	Ivory-Billed Aracari (Toucan)	DQ780882
Strigops habroptilus	Kakapo	AY309456
Pterodroma brevirostris	Kerguelen Petrel	AY158678
Eudynamys taitensis	Long-Tailed Cuckoo	EU410487
Anseranas semipalmata	Magpie Goose	AY309455
Pterocles namaqua	Namaqua Sandgrouse	See note
Hemiphaga novaeseelandiae	New Zealand Pigeon	EU725864
Falco peregrinus	Peregrine Falcon	AF090338
Dryocopus pileatus	Pileated Woodpecker	DQ780879
Aythya americana	Redhead Duck	AF090337
Archilocus colubris	Ruby-Throated Humming Bird	EF532935
Arenaria interpres	Ruddy Turnstone	AY074885
Larus dominicanus	Southern Black-Backed Gull	AY293619

NOTE—The sandgrouse mitochondrial genome was submitted to GenBank in individual genes. Their accession numbers are DQ385284 16S ribosomal RNA gene, DQ385267 12S ribosomal RNA gene, DQ385250 ATP synthase F0 subunit 8 (ATP8) gene, DQ385233 ATP synthase F0 subunit 6 (ATP6) gene, DQ385216 cytochrome b (cytb) gene, DQ385199 cytochrome oxidase subunit III (COIII) gene, DQ385182 cytochrome oxidase subunit II (COII) gene, DQ385165 cytochrome oxidase subunit I (COI) gene, DQ385148 NADH dehydrogenase subunit 5 (ND5) gene, DQ385131, NADH dehydrogenase subunit 4 (ND4) gene, DQ385114 NADH dehydrogenase subunit 4L (ND4L) gene, DQ385097 NADH dehydrogenase subunit 3 (ND3) gene, DQ385080 NADH dehydrogenase subunit 2 (ND2) gene, and DQ385063 NADH dehydrogenase subunit 1 (ND1) gene. The sandgrouse tRNA sequences used in this study are GXXXXXX-Y

The complete mitochondrial genome dataset had 23 birds (listed with their accession numbers in Table 4.2) and is 12,892 base pairs long. The species included in the analyses were selected as potential closest relatives to the pigeon, as well as more distant outgroups for rooting the tree. Where possible, pairs of species were chosen to help minimize the effect of long-branch attraction (Slack et al. 2007). The most variable sites in the dataset were RY coded or removed by site stripping as

described in Pratt et al. (2009). A Bayesian analysis comprising two 10^7 runs was performed in MrBayes, partitioned as in Gibb et al. (2007). Runs were checked for convergence and the first 10^6 trees were discarded as burnin. Maximum Likelihood (ML) analyses were performed using GARLI (Zwickl 2006) assuming the GTR+G+I model on an unconstrained dataset and also on a selection of constraints for the suggested closest relatives of pigeon and sandgrouse. The Bayesian tree with highest posterior probability, and all trees from GARLI were also optimised in PAUP* (Swofford 2002) in order to directly compare the likelihood scores. The tree scores were compared using the Shimodaira-Hasegawa (SH) test (Shimodaira and Hasegawa 1999) implemented in PAUP* (RELL, one-tailed test, 1,000 bootstrap replicates).

4.4 RESULTS

4.4.1 A South East Asian/Pacific radiation within Pigeons

We begin by examining the phylogenetic relationships within pigeons. Figure 4.1 shows an unrooted network of the Bayesian analysis, depicting all splits present in more than 20% of trees. The network representation has the advantage of showing where conflicting signal is present in internal branches. Overall, there is a well-supported clade of South East Asian/Pacific pigeons. The *Ducula–Ptilinopus* group as a whole is monophyletic and the branching structure of the three internal clades is highly supported by both bootstrap and BPP.

The genus *Ducula* is strictly monophyletic and internal branching has high support. Three of the four subgroups described in Gibbs et al. (2001) are represented in our analyses (see Roman numerals in Fig. 4.1). However, these subgroups do not form monophyletic groups in our phylogeny, and are largely intermingled. This clearly identifies an area where additional species information from DNA and morphology is required to resolve the phylogeny.

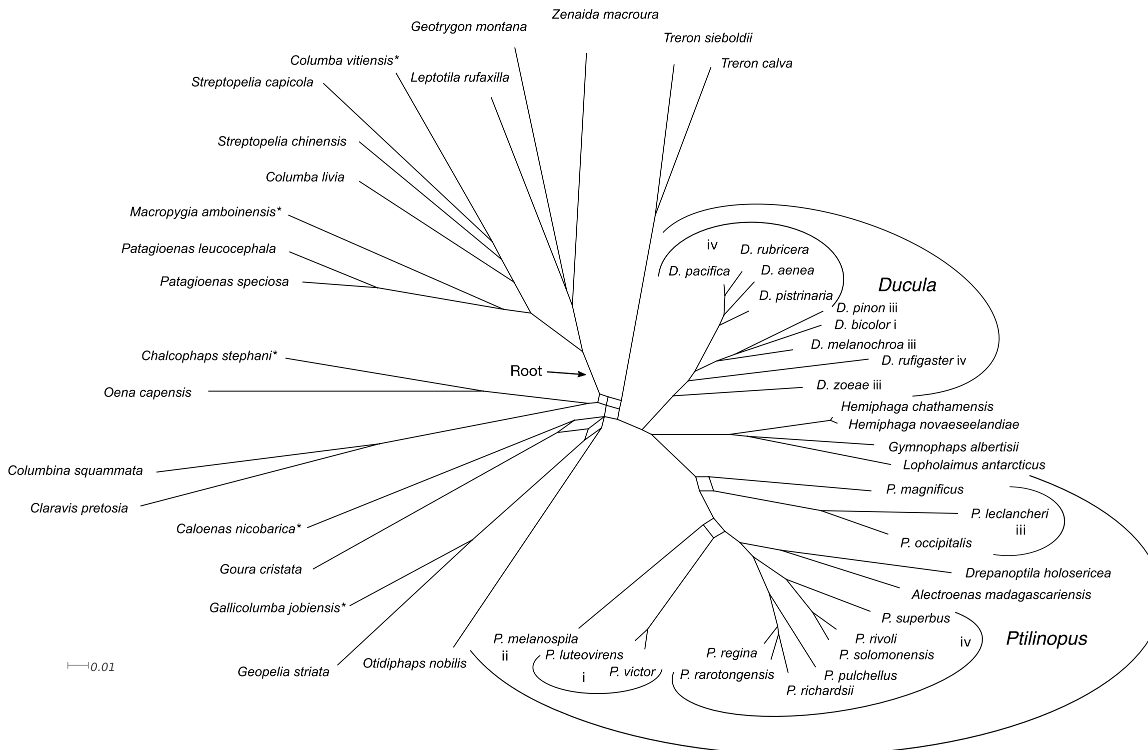

Figure 4.1—Phylogeny of pigeons with emphasis on the South East Asia/Pacific radiation. The MrBayes network shows splits present in over 25% of trees; the sequences are from 6 mitochondrial genes and are up to 6,013 bp long. Subgroups of *Ptilinopus* and *Ducula* as detailed by Gibbs et al. (2001) are indicated by roman numerals. Lineages outside the *Ducula-Ptilinopus* part of the tree that contain species also found in the South Pacific region are marked with an asterisk.

Gibbs et al. (2001) also splits *Ptilinopus* into four subgroups. Unlike *Ducula*, each subgroup is quite distinct on our network (Fig. 4.1), and most have high support. Additionally, *Ptilinopus* is paraphyletic on current taxonomy, because *Alectroenas* and *Drepanoptila* fall within the genus (Shapiro et al. 2002). This is confirmed here. The exact position of *Alectroenas* and *Drepanoptila* within *Ptilinopus* is not completely stable, and differs by one branch between the unrooted and rooted analyses (Figs. 4.1 and 4.2). Nevertheless this level of agreement is still highly significant (Pratt et al. 2009). *Alectroenas* and *Drepanoptila* are represented by only two of the six genes analysed, which will lead to lower resolution, and this again identifies an area for future work.

Hemiphaga, Lopholaimus and *Gymnophaps* group together with high bootstrap and BPP support (Fig. 4.2) and join the tree between *Ptilinopus* and *Ducula*. Thus the three genera have a strong affinity with the other South East Asian/Pacific pigeons. The difference between the two *Hemiphaga* species in the New Zealand

region is very shallow, as can be evidenced by the short distance separating *H. novaeseelandiae* and *H. chathamensis* (Fig. 4.1, and J. Goldberg pers. comm.).

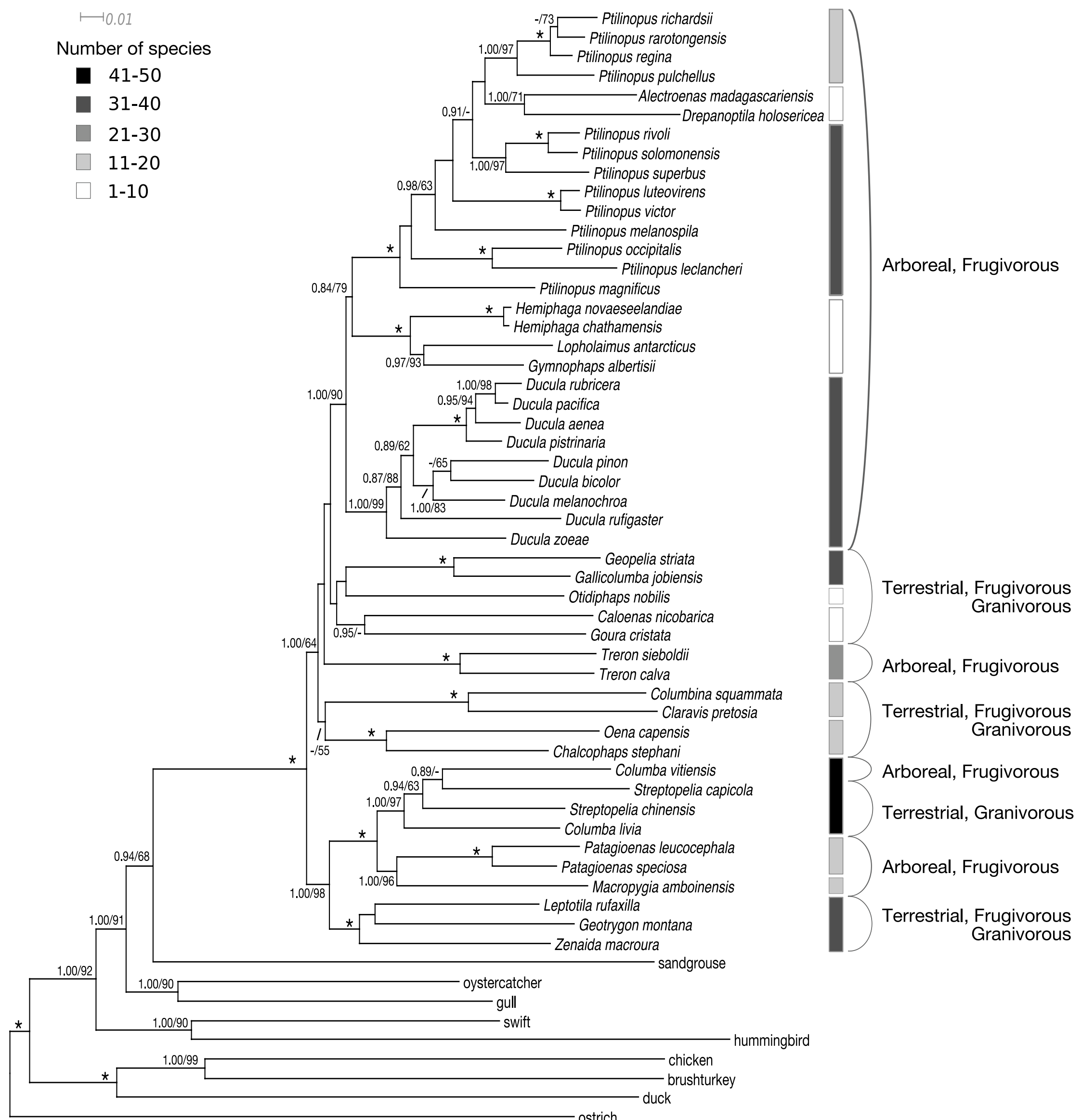

Figure 4.2—Bayesian rooted consensus tree of pigeons. The vertical bars to the right of the tree indicate the number of species represented by this phylogeny. Feeding habits (summarised from Gibbs et al. 2001) are also illustrated. Note the feeding habits are a generalisation, and many individual species may have more complex habits, e.g. some terrestrial granivorous pigeons also eat snails and insects. Branch support shown with Bayesian posterior probabilities > 0.8 and Maximum Likelihood bootstraps > 50%. Support of 1.0/100 is indicated by *.

In addition, there are several pigeons in the Pacific region outside of the *Ducula–Ptilinopus* clade. One of these, *Columba vitiensis*, does not group with *C. livia*. While not highly supported by bootstrapping, it appears *Columba* and *Streptopelia* are paraphyletic with respect to each other. There was only 29% bootstrap support

for *C. livia* and *C vitiensis* grouping together (data not shown), compared with 63% bootstrap and 0.94 BPP for the grouping of *C. vitiensis* with *Streptopelia* (Fig. 4.2).

When considering the root of the pigeon tree, we have found that it is best practice to first run an unrooted tree, and then add the root taxa in a separate run (Holland et al. 2003; Harrison et al. 2004). This helps detect one source of error where adding the outgroup may disrupt an otherwise correct ingroup tree. In the present pigeon tree, the root falls between a clade containing New World pigeons and allies (clade A in Pereira et al. 2007) and all other pigeons (Fig. 4.2). After addition of the outgroup, the ingroup branching pattern is broadly consistent with the unrooted analysis (Figs. 4.1 and 4.2), so adding the outgroup has made no major rearrangements to the ingroup tree.

4.4.2 Phylogenetic relationships of Pigeons within Neoaves

The complete mitochondrial genome of the New Zealand pigeon is 17,264 bp long (GenBank accession EU725864), and has the standard avian mitochondrial gene order found in the chicken (see Gibb et al. 2007). The gene order of the sandgrouse is not known, because the area including the control region has not been sequenced. Many avian groups have the standard gene order, so positional information is not informative in this situation.

The maximum likelihood tree from Garli (Zwickl 2006) of potential pigeon relatives is shown in Figure 4.3, with posterior probabilities over 0.8 from MrBayes, and maximum likelihood bootstrap support over 50%. Firstly, the species groups outlined with grey boxes have very high support. Secondly it can be seen that the pigeon and sandgrouse, while near neighbours (with one branch separating them on the tree), do not actually group together, nor do they have high support with respect to the falconiform clade they are grouped with. Because pigeon and sandgrouse are both single long branches, they are potentially unstable and cannot be placed confidently in the tree. Having said that, it does appear they may be closer to each other than to the alternative closest-relative hypotheses (such as shorebirds and parrots, see below and Table 4.3).

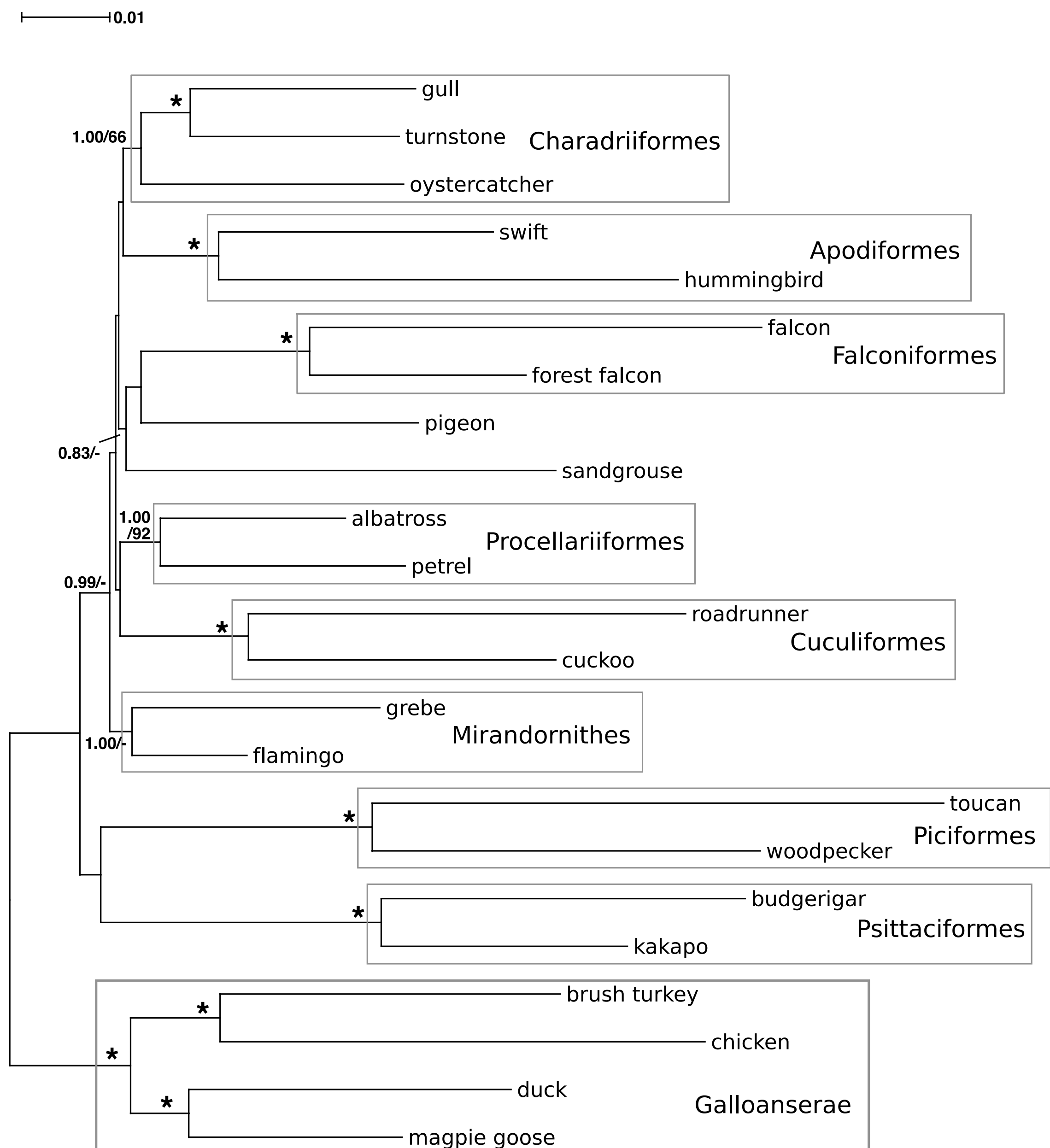

Figure 4.3—Maximum likelihood tree of birds from complete mitochondrial genomes generated using Garli. Branch support is shown as indicated in Figure 4.2.

Constraint groups were also considered. Constraining pigeons to group either with parrots, or shorebirds, or sandgrouse plus shorebirds (Sibley and Ahlquist 1990), all yielded trees with worse likelihoods than the unconstrained ML tree, and although the likelihood values differ by several units this is not statistically significant as assessed by the conservative SH test (Shimodaira and Hasegawa 1999, see Table 4.3). Because the internal branches on the tree are so short, it is difficult to have high statistical significance, however the best and worst trees are 20 log likelihood units apart. Analysis of trees produced by the different constraints showed the only differences in branching patterns related to the position of the

constrained groups. Additionally when not specifically constrained to group elsewhere, both the sandgrouse and the pigeon grouped with the falcons (data not shown).

Table 4.3—Testing alternative hypotheses about pigeon relatives using the SH test.

Constraint groups	Log-Likelihood	Difference in Log-Likelihood	P
Maximum Likelihood (Garli)	-51042.13126	(best)	
pigeon+sandgrouse	-51043.09900	0.96774	0.903
sandgrouse+shorebirds	-51044.83032	2.69906	0.789
Bayesian (MrBayes)	-51044.36383	2.23257	0.763
pigeon+shorebirds	-51047.72791	5.59665	0.647
pigeon+sandgrouse+shorebirds	-51051.24656	9.11530	0.498
pigeon+parrots	-51061.81728	19.68602	0.274

We can assess the information content of this result. If the tree contains n=11 generally accepted groups (the nine grey boxes plus pigeon and sandgrouse—Fig. 4.3), there are $(2n–5)!!$ possible unrooted binary trees, that is 34,459,425 trees. Of course, only a small proportion of these will be likely. Consider the probability that three groups (here pigeon, sandgrouse and falcons) form a single group, relative to the others. The probability (P_R) of observing a predefined clade of k=3 taxa in a binary rooted tree of n taxa is $P_R(n,k) = 3/(2n\text{-}3)(2n\text{-}5)$ (see Pratt et al. 2009 Appendix 1 for details). So the probability of finding the pigeon, sandgrouse and falcons grouping together by chance is 0.0093, a highly significant result. In summary, if we had begun with a flat prior (i.e no prior knowledge about the 11 groupings) then there are 3.4×10^7 possible trees, and a 0.9% probability of finding pigeons, sandgrouse and falcons grouping together by chance.

As an alternative approach, we can use our prior knowledge and assess our best hypothesis: that the pigeon and the sandgrouse group together. The probability that 2 out of n=11 groups are sister is $1/(2n\text{-}5)$, that is ~0.059. Our best tree found just one branch between the pigeon and sandgrouse. The probability of finding pigeon and sandgrouse either together or one node apart is 0.1294.

4.5 DISCUSSION

4.5.1 The Ducula–Ptilinopus radiation in a biogeographic context

This study examines the *Ducula–Ptilinopus* radiation in the South Pacific—it is not a complete taxonomic phylogeny of the *Ducula–Ptilinopus* radiation. It should be noted, however, that from morphological descriptions of species in these genera (summarised in Gibbs et al. 2001) it is expected that all other species will fall within the bounds of those analysed here. It is interesting that the molecular results are generally in excellent agreement with those groupings suggested by morphology. Although there is a strong tendency to 'believe' whatever class of data a researcher uses, it is more reassuring when there is agreement between different data sets (Pratt et al. 2009).

The widely distributed *D. pacifica* appears to be recently derived from *D. rubricera*. We would anticipate the very closely related allopatric *D. oceanica* (found throughout Micronesia) to also belong to this recent radiation. Interestingly, *D. rufigaster* and allospecies have been thought of as 'an offshoot of *D. aenea* stock' (Goodwin 1960; Gibbs et al. 2001), whereas it is clearly quite distinct. Overall, *Ducula* is a relatively compact monophyletic genus, with all species studied being fairly closely related. The location of this genus across many islands probably does reduce gene flow and has increased the number of described species found within it.

In contrast, *Ptilinopus* is more complex and, as currently prescribed it is paraphyletic with respect to *Alectroenas* (Madagascan) and *Drepanoptila* (New Caledonian). This is in agreement with the results of Shapiro et al. (2002). Therefore we synonymise *Alectroenas* Gray, 1840 and *Drepanoptila* Bonaparte, 1855 with *Ptilinopus* Swainson, 1825. Continuing with generic boundaries, one subgroup, the Fijian *P. luteovirens* and *P. victor* have sometimes been placed in their own genus, *Chrysoenas* Layard, 1876, whereas here we confirm they group well within *Ptilinopus*, and are also closely related to each other (although not as closely as the two *Hemiphaga* species). The present results strongly support the Fijian species being included with *Ptilinopus*.

Gymnophaps and *Lopholaimus* are found to be sister taxa, with *Hemiphaga* basal. This is in agreement with Pereira et al. (2007), but was not previously predicted by taxonomy (see Gibbs et al. 2001). On the basis of genetic similarity alone it might be reasonable to combine these three genera into one. The remaining genus in this group, *Cryptophaps,* contains a single rare species found in the highlands of Sulawesi. No *Cryptophaps* sample was available for this, or any previous molecular study. While the genetic affinity of this species is unknown, based on morphology it is expected to group with these three genera.

Given the high number of pigeon species in Australia, it is perhaps surprising that mainland New Zealand has only one native pigeon (*Hemiphaga*). It is less surprising that this pigeon would be part of the *Ducula–Ptilinopus* radiation clearly successful at island colonization. Distance from Australia is unlikely to be a factor, as pigeons have been highly successful at colonizing other Pacific islands, reaching all but Hawaii and Easter Island (Steadman 1997, 2006). New Zealand is, however more temperate than the mainly tropical distribution of most pigeons.

While *Hemiphaga* has not speciated within the large mainland islands of New Zealand, a very closely related species, *H. chathamensis,* is found on the Chatham Islands (J. Goldberg pers. comm., Millener and Powlesland 2001). This close relationship is consistent with the young age of the Chatham Islands (Campbell and Hutchings 2007). Additionally an extinct (1800's) subspecies of *H. novaeseelandiae* existed on Norfolk Island (Tennyson and Martinson 2006), which is also a recent radiation (J Goldberg, pers. comm.). This emphasises that *Hemiphaga* is capable of differentiating into new species as space arises; however it has not evolved into multiple species within mainland New Zealand. Contrast this with some islands of the Pacific, which have had 2-3 species of *Ducula* or *Ptilinopus* (for example, Fiji, and see Steadman 2006). Pigeons have, however, been part of the New Zealand biota for some time. Although fossils are rare, a columbid is part of the New Zealand St Bathans Fauna (16-18Ma, Worthy et al. 2007), and recent work suggests it is part of the lineage leading to *Hemiphaga* (Worthy et al. 2009).

The *Ducula–Ptilinopus* group is not a recent radiation. From the dating analysis of Pereira et al. (2007), this South East Asia/Pacific radiation may be at least 30Ma. The ecological niche filled by the species in *Ducula* and *Ptilinopus*, and the overall characters defining and separating these from other species of pigeons and doves

have remained consistent over a long period of time. Generally these birds are arboreal frugivores (see Fig. 4.2), often nomadic, and strong fliers capable of colonizing isolated islands. *Ducula–Ptilinopus* clade fruit-pigeons only digest the soft part of the fruit, leaving the seeds intact (Gibbs et al. 2001). A related pigeon genus (*Treron*) is also arboreal and frugivorous, but is still capable of grinding and digesting hard seeds, more like *Columba*.

It should be noted that some pigeons from other genera are also found in parts of this region (for example, *Columba vitiensis* and *Gallicolumba spp.,* see Fig. 4.1) but may fill niches outside those of *Ducula* and *Ptilinopus* (see Fig. 4.2). New Zealand, Tokelau, Tuvalu, Niue, Cook Islands, Pitcairn and the Marshall Islands only have extant pigeons from the *Ducula–Ptilinopus* group, although *Columba livia* and *Streptopelia chinensis* have been introduced to some islands (Lepage 2009). There is evidence of extinct pigeons from other clades (Steadman 2006)

It is interesting that within the *Ptilinopus* radiation principally found in the South East Asia/Pacific region one genus (*Alectroenas*) lives in the Madagascan region. This emphasises the capacity for dispersal of many pigeons. It is not obvious why some genera are not more widely distributed; we need to consider possible long-term stability of some lineages in some ecological roles (Poole et al. 2003). Explanations may involve some form of competitive exclusion, and/or it may require a reasonable size of founding population for a species to establish itself (e.g. Veltman et al. 1996).

Although beyond the scope of the present study, it does appear that the remaining pigeons outside the Asia/Pacific grouping may be over-split at the generic level, and that some species could be combined into more appropriate genera. For example, the radiation represented here by *Gallicolumba* and *Geopelia* (Fig. 4.1) also contains another five genera with one to four species each. Some species within *Ptilinopus* are more distantly related to each other than the seven genera represented by *Gallicolumba* and *Geopelia*. While we do not support a literal translation of evolutionary trees into a taxonomic classification, there may be room for reduction of genera here. The *Columba*/*Streptopelia* complex is more complicated in that *Columba vitiensis* is not closely related to *C. livia*. Because *C. vitiensis* is an arboreal frugivore, unlike *C. livia* and *Streptopelia,* it is likely *Columba* may have multiple distinct subgroups. The instability of the speciose

Columba and lack of monophyly is consistent with previous subdivisions into the old world *Columba* and the new world *Patagioenas* (Johnson et al. 2001).

There is a need for studies about maximizing the information content of a classification and what number of genera is the most informative for a given number of species. Within pigeons, there is an asymmetric distribution of species number per genus; six of the 43 genera account for half of all pigeon species, and a third of the genera are monophyletic (numbers calculated from Gibbs et al. 2001). Is this because of extinction removing diversity in many genera, as suggested by the recent discovery of extinct species of *Didunculus, Caloenas,* and near relatives like *Bountyphaps*, (Steadman 2006)? Or is it a coincidence of success of a small number of genera, or perhaps over-classification of island variation, such as with the large flightless dodo *Raphus cucullatus* (Shapiro et al. 2002)? Now that the phylogenetic diversity of pigeons is being resolved, it is questions like these we can begin to answer.

4.5.2 Closest relative of Pigeons

Our phylogeny is in basic agreement with previous mitochondrial and exon nuclear analyses (e.g., Hackett et al. 2008; Pratt et al. 2009) and we do not see the Metaves/Coronaves split (Fain and Houde 2004) found with some nuclear intron data (see Morgan-Richards et al. 2008 for discussion). In our analyses (both Garli and MrBayes) the pigeon and sandgrouse join to the base of the falcon branch. This is an intriguing result, albeit one with low support. It is possibly the result of attraction of the two single branches (pigeon and sandgrouse) to the relatively long internal falcon branch. A similar result was also seen by Brown et al. (2008) who were using more species but shorter sequences (only 5 Kb mt data) to study the origin of modern avian lineages. We expect the addition of further mitochondrial genomes in the future to influence our result, and future examination is warranted.

The sandgrouse may be the closest relative of pigeons, however it is not close enough to stably group with pigeons. Having shown that none of the putative close relatives are close enough to stabilise either pigeons or sandgrouse, we assess what can be done next. To help resolve their placement, mt genomes from species to pair with the pigeon and sandgrouse may have to be sequenced. Unfortunately ideal taxa that are not too closely or too distantly related may not exist. A second pigeon

will almost certainly help, preferably from the opposite side of the root to *H. novaeseelandiae*. This would decrease the length of the internal branch to other avian species. From our results, the common rock dove, *Columba livia* is a good choice. It lies on the opposite side of the root of pigeons, relative to *Hemiphaga*, and is also relatively slowly evolving (Fig. 4.2). Candidates outside Columbiformes and *Pterocles* may be difficult to find. The intron-heavy study of Hackett et al. (2008) found Mesitornithidae (mesites) to group with pigeons and sandgrouse, although with low support and, once again joining very deeply on the pigeon branch. Additionally a sequence from the other sandgrouse genus, *Syrrhaptes* may help stabilise *Pterocles*.

In summary, the relationships found in this study, both within pigeons (specifically within the *Ducula–Ptilinopus* group) and between pigeons and other birds builds on the results of past studies, identifies areas for improving taxon sampling, and sets up hypotheses to be tested with nuclear (exon) data and morphological datasets.

4.6 ACKNOWLEDGEMENTS

We thank Tara Paton for the additional tRNA sand grouse sequences; Walter Boles for tissue samples from the Australian Museum, Sydney, Australia; Craig Morley for feather samples from Fiji; Dick Gill, Waikanae Department of Conservation, New Zealand for *Hemiphaga novaeseelandiae*. Thanks to Julia Goldberg and Steve Trewick for providing *H. chathamensis* sequences ahead of publication, and for helpful discussion on the manuscript. Thanks also to Trevor Worthy for comments, and for suggesting that *Alectroenas* and *Drepanoptila* be synonymised with *Ptilinopus*. The Allan Wilson Centre for Molecular Ecology and Evolution supported this research.

4.7 REFERENCES

Brown, J. W., J. S. Rest, J. Garcia-Moreno, M. D. Sorenson, and D. P. Mindell. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol 6:6.

Campbell, H. J., and Hutchings. 2007. In search of ancient New Zealand. Penguin, Auckland & GNS Science, Lower Hutt, New Zealand.

Cracraft, J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proc R Soc Lond B Biol Sci 268:459-469.

Ericson, P. G. P., I. Envall, M. Irestedt, and J. A. Norman. 2003. Inter-familial relationships of the shorebirds (Aves: Charadriiformes) based on nuclear DNA sequence data. BMC Evol Biol 3:16.

Fain, M. G., and P. Houde. 2004. Parallel radiations in the primary clades of birds. Evolution 58:2558-2573.

Gibb, G. C., O. Kardailsky, R. T. Kimball, E. L. Braun, and D. Penny. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol Biol Evol 24:269-280.

Gibbs, D., E. Barnes, and J. Cox. 2001. Pigeons and Doves: A Guide to Pigeons and Doves of the World. Yale University Press, New Haven.

Goodwin, D. 1960. Taxonomy of the genus *Ducula*. Ibis 102:526-535.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Harrison, G. L., P. A. McLenachan, M. J. Phillips, K. E. Slack, A. Cooper, and D. Penny. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late Cretaceous. Mol Biol Evol 21:974-983.

Holland, B., F. Delsuc, and V. Moulton. 2005. Visualizing Conflicting Evolutionary Hypotheses in Large Collections of Trees: Using Consensus Networks to Study the Origins of Placentals and Hexapods. Syst Biol 54:66-76.

Holland, B. R., D. Penny, and M. D. Hendy. 2003. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock - A simulation study. Syst Biol 52:229-238.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference on phylogenetic trees. Bioinformatics 17:754-755.

Huson, D. H., and D. Bryant. 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol 23:254-267.

Johnson, K. P. 2004. Deletion bias in avian introns over evolutionary timescales. Mol Biol Evol 21:599-602.

Johnson, K. P., S. De Kort, K. Dinwoodey, A. C. Mateman, C. Ten Cate, C. M. Lessells, and D. H. Clayton. 2001. A molecular phylogeny of the dove genera *Streptopelia* and *Columba*. Auk 118:874-887.

Lepage, D. 2009. Avibase: The World Bird Database. Bird Studies Canada.

Liu, K., S. Raghavan, S. Nelesen, C. R. Linder, and T. Warnow. 2009. Rapid and Accurate Large-Scale Coestimation of Sequence Alignments and Phylogenetic Trees. Science 324:1561-1564.

Livezey, B. C., and R. L. Zusi. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool J Linn Soc 149:1-95.

Millener, P. R., and R. G. Powlesland. 2001. The Chatham Islands pigeon (Parea) deserves full species status; *Hemiphaga chathamensis* (Rothschild 1891); Aves: Columbidae. J R Soc N Z 31:365-383.

Morgan-Richards, M., S. Trewick, A. Bartosch-Härlid, O. Kardailsky, M. J. Phillips, P. A. McLenachan, and D. Penny. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol Biol 8:20.

Paton, T. A., and A. J. Baker. 2006. Sequences from 14 mitochondrial genes provide a well-supported phylogeny of the Charadriiform birds congruent with the nuclear RAG-1 tree. Mol Phylogenet Evol 39:657-667.

Paton, T. A., A. J. Baker, J. G. Groth, and G. F. Barrowclough. 2003. RAG-1 sequences resolve phylogenetic relationships within Charadriiform birds. Mol Phylogenet Evol 29:268-278.

Pereira, S. L., K. P. Johnson, D. H. Clayton, and A. J. Baker. 2007. Mitochondrial and nuclear DNA sequences support a cretaceous origin of Columbiformes and a dispersal-driven radiation in the Paleogene. Syst Biol 56:656-672.

Poole, A. M., M. J. Phillips, and D. Penny. 2003. Prokaryote and eukaryote evolvability. Biosystems 69:163-185.

Pratt, R. C., G. C. Gibb, M. Morgan-Richards, M. J. Phillips, M. D. Hendy, and D. Penny. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol 26:313-326.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor.

Rambaut, A., and A. J. Drummond. 2003. Tracer: MCMC Trace Analysis Tool. University of Oxford.

Sambrook, J., E. F. Fritsch, and T. Maniatas. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbour Laboratory Press, New York.

Shapiro, B., D. Sibthorpe, A. Rambaut, J. Austin, G. M. Wragg, O. R. P. Bininda-Emonds, P. L. M. Lee, and A. Cooper. 2002. Flight of the dodo. Science 295:1683-1683.

Shimodaira, H., and M. Hasegawa. 1999. Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. Mol Biol Evol 16:1114-1116.

Sibley, C. G., and J. E. Ahlquist. 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale University Press, New Haven, Connecticut.

Slack, K. E., F. Delsuc, P. A. Mclenachan, U. Arnason, and D. Penny. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol Phylogenet Evol 42:1-13.

Steadman, D. W. 1997. The historic biogeography and community ecology of Polynesian pigeons and doves. J Biogeo 24:737-753.

Steadman, D. W. 2006. Extinction and Biogeography of Tropical Pacific Birds. The University of Chicago Press, Chicago.

Swofford, D. M. 2002. PAUP* 4.0b10 Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinaur Associates, Sunderland, Massachusetts.

Tennyson, A. J. D., and P. Martinson. 2006. Extinct Birds of New Zealand. Te Papa Press, Wellington, New Zealand.

Veltman, C. J., S. Nee, and M. J. Crawley. 1996. Correlates of introduction success in exotic New Zealand birds. Am Nat 147:542-557.

Wong, K. M., M. A. Suchard, and J. P. Huelsenbeck. 2008. Alignment uncertainty and genomic analysis. Science 319:473-476.

Worthy, T. H., S. J. Hand, J. P. Worthy, A. J. D. Tennyson, and R. P. Scofield. 2009. A Large Fruit Pigeon (Columbidae) from the Early Miocene of New Zealand. Auk 126:649-656.

Worthy, T. H., A. J. D. Tennyson, C. Jones, J. A. McNamara, and B. J. Douglas. 2007. Miocene waterfowl and other birds from central Otago, New Zealand. J Syst Palaeontol 5:1-39.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD dissertation. University of Texas, Austin, Texas.

CHAPTER FIVE

5 BEYOND PHYLOGENY: PELECANIFORM AND CICONIIFORM BIRDS, AND LONG-TERM NICHE STABILITY[1]

5.1 ABSTRACT

Phylogenetic trees are just a starting point for the study of further evolutionary and ecological questions. We show that for avian evolutionary relationships, improved taxon sampling, longer sequences and additional data sets are giving stability to the prediction of a pelecaniform and ciconiiform grouping, thus allowing inferences to be made about long-term niche occupancy. Here we report the phylogeny of the pelecaniform birds and their allies using complete mitochondrial genomes, and show that the basic groupings are consistent with nuclear sequence phylogenies, even though many short branches are not yet fully resolved. We show that the Pelecaniformes (minus the tropicbird) and the core Ciconiiformes (storks, herons and allies) form a natural group (the P&C) within a seabird water-carnivore clade. We find pelicans are the closest relatives of the hammerkop, in a clade with the shoebill, and we confirm that tropicbirds are definitely not pelecaniforms. The P&C group appears to be an adaptive radiation into an 'aquatic carnivore' niche that the group has occupied for 70-80 million years. From an ecological and life history perspective there is a unity to the combined pelecaniform–ciconiform grouping, which appears to be more natural than focusing on details of morphology. These findings allow us to start integrating molecular evolution and macroecology.

[1] This chapter submitted as: Gibb, G. C., Kennedy, M. and Penny, D. Beyond Phylogeny: Pelecaniform and Ciconiiform Birds, and Long-Term Niche Stability. Systematic Biology

Keywords—adaptive radiation, life history, mitochondrial DNA, seabirds, shorebirds, water-carnivores

5.2 INTRODUCTION

What do we do when we have solved the tree of life? While a phylogeny is informative in itself, the interpretations that can be drawn from a stable phylogeny are what is most interesting, and focus our goals for future work – moving beyond phylogeny. When phylogenetic agreement has been reached by multiple datasets, it is possible to use this framework to ask broader questions about macroecology and long-term ecological niche stability (Poole et al. 2003). Here, we ask whether phylogenetic agreement has been reached, and use this agreement to investigate whether the pelecaniform group has occupied an essentially consistent niche space, with species turnover within it, for a very long time (many tens of millions of years). In this context, we are more concerned with general niche stability that can operate at several levels, rather than more specific species-level niche occupancy. If, as we hypothesize (and it is consistent with nuclear intron data, Hackett et al. 2008), the Pelecaniformes and Ciconiiformes (P&C) form a natural group within a larger framework of water-carnivores, what can be said about the long-term life history of the whole group? Looking to the future, our longer-term goal is the integration of evolutionary trees with a range of life history traits (see Paterson et al. 1995; Kennedy et al. 1996; Slikas 1998).

5.2.1 Consilience of Induction

To have confidence in our findings it is important in general to have agreement among multiple datasets. Datasets are more likely to agree as taxon sampling and information content increases. For example Paton et al. (2006) showed that relationships within the Charadriiformes were resolvable with long mitochondrial sequences whereas shorter mitochondrial sequences had been unable to resolve the groupings. Additionally, their results agreed with independent findings from nuclear data (Paton et al. 2003). At deeper phylogenetic levels we have also found good agreement between mitochondrial and nuclear exon datasets (Slack et al. 2007; Hackett et al. 2008; Pratt et al. 2009). We expect this will likely be the case for the Pelecaniformes and their allies as well.

Earlier molecular studies that included the Pelecaniformes and their allies used short DNA sequences that provided insufficient resolution of the deeper branches (for example, Hedges and Sibley 1994; van Tuinen et al. 2001; Fain and Houde 2004). Other studies focused on resolving relationships within families, and as such did not have a wide enough sampling to analyze the group as a whole (e.g. Slikas 1997; Kennedy et al. 2005). More recent studies of deeper avian phylogeny have often had only one or two members of an order as representatives (e.g. Brown et al. 2008; Pratt et al. 2009). The problem of deep avian phylogeny resolution is especially compounded because many internodes in the deeper neoavian tree branches are relatively short (Hackett et al. 2008; Pratt et al. 2009). Nevertheless, if there is an adaptive radiation (see Gibb et al. 2007) we may find that a group as a whole may be stable, even if the group has short internal branches within it that are not yet stably resolvable.

Recent nuclear-based studies have increased in sequence length and taxon sampling for the Pelecaniformes and their allies (Ericson et al. 2006; Hackett et al. 2008). However, some caution must accompany the results of these studies, as a high proportion of their nuclear data comes from introns, which are problematic to align at deeper divergences (Morgan-Richards et al. 2008; Wong et al. 2008; Liu et al. 2009). These studies showed that at deep avian divergences, intron alignments with many indels can provide overriding signal for an incorrect tree, such as the Metaves/Coronaves split found by Fain and Houde (2004). Introns are, however, very useful for resolving within-family level phylogenies (e.g. de Kloet and de Kloet 2005). In general the findings of these nuclear-based studies provide phylogenetic predictions we can test with whole mitochondrial genomes, and vice versa.

5.2.2 History of the Pelecaniformes

Although one could focus solely on the Pelecaniformes, however from assessing the results of earlier studies (morphological, mitochondrial and nuclear) it appears that some pelecaniform families may be closely associated with ciconiiform families (storks, herons and ibises), within a larger adaptive radiation of aquatic and semi-aquatic carnivores (Cracraft 2001; van Tuinen et al. 2001; Hackett et al. 2008; Pratt et al. 2009). This wider seabird water-carnivore clade may also include the tubenoses, loons and penguins, but probably not the shorebirds nor flamingoes and grebes. This clade would represent impressive niche stability, especially as some

groups have moved from the sea to freshwater (or the reverse), including the ciconiiform birds, darters and some cormorants. Deeper nodes in this clade are often poorly resolved (e.g. Ericson et al. 2006; Brown et al. 2008). We expect that longer sequences and improved taxon sampling will improve this resolution.

There is an extensive history of debate surrounding the Pelecaniformes, leading Sibley and Ahlquist (1990) to suggest that the Pelecaniformes 'may present the most complex and controversial questions in the avian phylogeny'. For a long time they were considered a monophyletic group comprising the frigatebirds (Fregatidae), tropicbirds (Phaethontidae), pelicans (Pelecanidae), darters (Anhingidae), cormorants/shags (Phalacrocoracidae), and gannets/boobies (Sulidae). Despite its 'divergency of structure' (Beddard 1898), this group was apparently united by morphological characters including totipalmate feet (all four toes being connected by a web), gular pouch, lack of a brood patch, and the salt gland being completely enclosed within the orbit (rather than being in a cavity on top of the skull) (Cracraft 1985; Hedges and Sibley 1994; Nelson 2005; summarised in Table 5.1).

The monophyly of the 'core pelecaniforms' (this includes the cormorants, darters and sulids, though ironically not the pelicans) has never really been in question (Cracraft 1985; Kennedy and Spencer 2004; Nelson 2005). While molecular studies have had varying degrees of coverage with taxon sampling and resolution, a molecular consensus is forming regarding the pelicans being closely allied with the ciconiiform shoebill and hammerkop, and the tropicbirds being very distantly related to the other pelecaniform birds (e.g. Hackett et al. 2008 with nuclear data).

TABLE 5.1—Characters previously used to group pelecaniform families. For additional behavioral characters that have been used to group taxa within the families see Kennedy et al. (1996).

Character	Phaethontidae	Fregatidae	Pelecanidae	Sulidae	Anhingidae	Phalacrocoracidae
Totipalmate foot [1]	✓	✓	✓	✓	✓	✓
Gular pouch [1]	✓	✓	✓	✓	✓	✓
Method of feeding chicks [2]	✓	✓	✓	✓	✓	✓
Pre-landing call [2]	✓	✓	✓	✓	✓	✓
Presentation of nest material [2]	-	✓	✓	✓	✓	✓
Bowing [2]	-	-	✓	✓	✓	✓
Hop [2]	-	-	-	✓	✓	✓
Sky-pointing & Wing-waving [2]	-	-	-	✓	✓	✓
Pointing [2]	-	-	-	-	✓	✓
Kink-throating [2]	-	-	-	-	✓	✓

[1] From Siegel-Causey (1997) Table 1

[2] From van Tets (1965) Figure 47

The morphological findings of Livezey and Zusi (2007), however, still place the tropicbirds and pelicans within Pelecaniformes, with the shoebill as a monotypic sister order (but see discussion in Mayr 2008). This result is apparently supported by four diagnostic apomorphies grouping the shoebill as the sister order to the Pelecaniformes, four diagnostic apomorphies grouping the Pelecaniformes, and seven and nine supportive apomorphies for these groupings respectively (see Livezey and Zusi's Table 2). We argue the phylogenetic signal in morphological characters can be obfuscated by homoplasy. Perhaps more informative is to use the genetic evidence to establish the phylogeny, and then to map the ecological life history/morphological characters (Table 5.1) onto this tree, and then to understand the life history changes that have occurred.

In this study, we use complete mitochondrial DNA sequences to test hypotheses about the phylogenetic relationships within the Pelecaniformes and their close allies, particularly the Ciconiiformes (P&C). Our analysis includes seven additional mitochondrial genomes within the P&C. Four genomes (cormorant, gannet, darter and the shoebill) are first reported here, plus three ciconiiforms not previously analyzed (egret, spoonbill and ibis). Previous studies have published/analyzed mitochondrial genomes of the pelican, frigatebird and tropicbird. Of the ciconiiform taxa, only two mitochondrial genomes (heron and stork) have previously been analyzed. With these analyses, we demonstrate the extent to which the mitochondrial, nuclear and morphological results agree or disagree, and the implications the combined results have for interpreting the morphology and life history of this seabird water-carnivore group. We find that at least the pelecaniform-ciconiiform group appears to have occupied a 'water-carnivore' role for many tens of millions of years.

5.3 METHODS

5.3.1 Taxon Sampling, DNA Extraction, PCR and Sequencing

Species sampled were *Morus serrator* (Australasian gannet), *Anhinga rufa* (African darter), *Phalacrocorax chalconotus* (Stewart Island shag), and *Balaeniceps rex* (shoebill). Total genomic DNA was extracted from the samples using standard phenol/chloroform extraction and ethanol precipitation. Mitochondrial genomes were amplified in 2-3 overlapping long-range fragments,

which were subsequently amplified in shorter 0.5-3 Kb fragments using primers from our database. This reduces the possibility of sequencing nuclear copies of mtDNA (numts). Primer sequences are available from the authors on request. For more details see Gibb et al. (2007) and references therein. Within each genome, sequences were aligned and manually checked using Sequencher 4.7 (Gene Codes Corp., Ann Arbor, MI).

5.3.2 Species for Analysis

In addition to the four new sequences, mitochondrial genomes from all P&C families available on GenBank were included in this study. Within the two orders, including the new genomes reported here, all but one of the families (the hammerkop, Scopidae) are represented by complete mitochondrial genomes. Shorter mitochondrial DNA sequences (3.5 – 5.7 Kb) are available for four additional P&C taxa (including the hammerkop), and these were included in some analyses.

As well as members of the P&C group, our dataset contains representatives from groups that have been suggested to be their close relatives, and that could potentially disrupt the P&C group. Therefore our dataset also includes seabirds (Procellariiformes), penguins (Sphenisciformes) and loons (Gaviiformes). There are indications that the tropicbird may not group with any of these families, therefore shorebirds (Charadriiformes), flamingo and grebe (Mirandorniths), hummingbird, swift and owlet nightjar (Apodiformes and Caprimulgiformes) were also added as potential relatives for the tropicbird (see Hackett et al. 2008). Four Galloanseres species were also used as a known outgroup to all these taxa (Slack et al. 2007; Hackett et al. 2008).

The following genomes were included in the analyses: frigatebird (*Fregata* sp., AP009192), Australian pelican (*Pelecanus conspicillatus*, DQ780883), red-tailed tropicbird (*Phaethon rubricauda*, AP009043), white-faced heron (*Ardea novaehollandiae*, DQ780878), Chinese egret (*Egretta eulophotes*, EU072995), black-faced spoonbill, (*Platalea minor*, EF455490), crested ibis (*Nipponia nippon*, AB104902), Oriental white stork (*Ciconia boyciana*, AB026193), black browed albatross (*Diomedea melanophris*, AY158677), Kerguelen petrel (*Pterodroma brevirostris*, AY158678) Australasian little grebe (*Tachybaptus novaehollandiae*,

EF532936), greater flamingo (*Phoenicopterus ruber roseus*, EF532932), little blue penguin (*Eudyptula minor*, AF362763), rockhopper penguin (*Eudyptes chrysocome*, AP009189), red-throated loon (*Gavia stellata*, AY293618), Pacific loon (*Gavia pacifica*, AP009190) blackish oystercatcher (*Haematopus ater*, AY074886), ruddy turnstone (*Arenaria interpres*, AY074885), southern black-backed gull (*Larus dominicanus*, AY293619), ruby-throated hummingbird (*Archilochus colubris*, EF532935), common swift (*Apus apus,* AM237310), Australian owlet-nightjar (*Aegotheles cristatus cristatus*, EU344979) magpie-goose (*Anseranas semipalmata*, AY309455), redhead duck (*Aythya americana*, AF090337), chicken (*Gallus gallus*, AP003317), and Australian brush-turkey (*Alectura lathami*, AY346091).

Shorter mtDNA sequences were included for four additional pelecaniform and ciconiiform species. They were wood stork (*Mycteria americana*, AY274076, DQ433030, AF082066, DQ485797), fasciated tiger-heron (*Tigrisoma fasciatum*, EU167034, EU166937, EU166980), hammerkop (*Scopus umbretta,* AF339360, EU372682, U08936), and masked booby (*Sula dactylatra*, EU372670, EU372686, DQ433230, AY941806, EU372677). The sequences mostly overlap between species, and total 3.5-5.7 Kb for each species. Sequences included are from 12S—COI, Atp6/8 and cyt*b*.

5.3.3 Alignment

Sequences were aligned and checked in Se-Al v2a11 (Rambaut 1996) at the amino acid level for protein coding genes and based on stem loop secondary structure for RNA genes. A conservative alignment procedure was used where gaps, ambiguous alignments next to gaps, NADH6 (light strand encoded), non-coding regions and stop codons are excluded from the alignment. The final dataset is 12,978 nucleotides long and contains 31 taxa. Third codon positions in protein coding regions are coded as RY (as explained in Gibb et al. 2007; Phillips et al. 2010). The dataset is available from our website http://www.allanwilsoncentre.ac.nz/ downloads.

5.3.4 Phylogenetic Analyses and Tree Building

Data was analyzed by a range of methods. Maximum Likelihood with bootstrapping was implemented in Garli (Zwickl 2006) using a general time reversible model with

estimated invariable sites and gamma distribution (GTR+i+γ), and rapid bootstrapping in RAxML (Stamatakis et al. 2008) with estimated invariable sites. The dataset was also analyzed by Bayesian methods using MrBayes (Huelsenbeck and Ronquist 2001). For RAxML and MrBayes analyses the dataset was partitioned into codon positions (protein) and stems and loops (RNA). When the four additional taxa (with shorter mtDNA sequences) were added the data was further partitioned to separate complete sections from those missing data. Trees were viewed using Dendroscope (Huson et al. 2007) and SplitsTree 4.8 (Huson and Bryant 2006).

5.3.5 Molecular Dating

Dated analyses were performed using BEAST v1.4.8 (Drummond and Rambaut 2007) with the 31-taxon dataset partitioned as for Bayesian analyses. An uncorrelated relaxed clock model was used with rates among branches distributed according to a lognormal distribution (Drummond et al. 2006). Nucleotide partitions used an estimated GTR+i+γ model; RY partitions a modified Kimura two-state model. The following dates and calibration priors were used following the recommendations of Ho and Phillips (2009). Root: normal distribution of 66-121 Ma (98% range). Galloanserae: normal distribution, 95% range from 66-86 Ma, with the minimum based on *Vegavis* (Clarke et al. 2005). For the penguin-loon split (Sphenisciformes and Gaviiformes), both normal distribution of 61.5-65.5 Ma (95% range), and lognormal distribution with a hard minimum of 61 Ma (based on the penguin – Waimanu, Slack et al. 2006), a mean of 65 Ma and a 97.5% soft maximum at 74Ma were trialed with similar results. Runs totaling 40,000,000 MCMC generations ensured ESS values > 200 (as estimated in Tracer v 1.4 Rambaut and Drummond 2003). Chains were sampled every 5,000th generation after removing the burnin of at least 10% (4,000,000 generations). Trees were viewed using Figtree v1.2.3 (http://tree.bio.ed.ac.uk/software/figtree/).

5.4 RESULTS AND DISCUSSION

5.4.1 Details of New Mitochondrial Genomes and Gene Orders

We report four new pelecaniform and ciconiiform mitochondrial genomes. The shoebill mitochondrial genome (accession number GU071053) is 15,752+ bp long, and has the standard avian gene order found in the chicken (and in the pelican).

The three core pelecaniforms (gannet [19,285 bp, GU071056], darter [19,385 bp, GU071055] and cormorant, [19,073 bp, GU071054]), have a novel mitochondrial gene order (see Mindell et al. 1998; Gibb et al. 2007), with the duplicated region spanning half of cyt*b*, and all of tRNA Thr, tRNA Pro, NADH 6, tRNA Glu and the control region (Fig. 5.1). These three genomes are all over 19 Kb in length, and are the longest avian mitochondrial genomes reported to date. All three genomes show evidence of concerted evolution, in that the duplicated coding regions are nearly identical within each individual, but very different between the three genera (shown in Fig 5.1). It is unknown if this gene duplication is also present in the frigatebird, because the frigatebird mitochondrial genome retrieved from GenBank (AP009192) does not have the full control region.

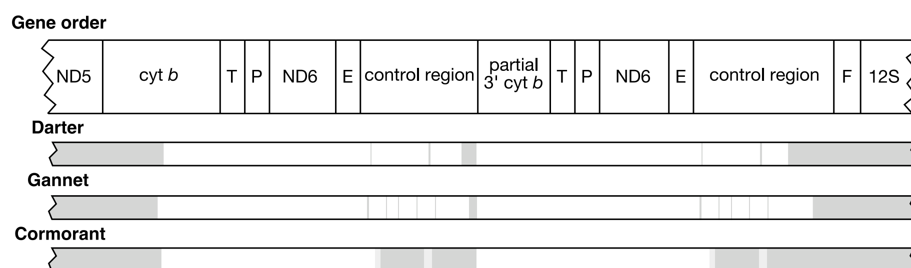

FIGURE 5.1—Comparison of duplicated regions in the three core Pelecaniformes. Within each individual, unique regions are shaded grey; duplicated regions that are 100% identical are white. Pale grey regions in the cormorant are alignable but not identical—the two cormorant control regions are about 90% different from each other. In contrast, the gannet and darter control regions are virtually identical.

5.4.2 Implications of Core Pelecaniform Gene Duplications for Phylogenetic Studies

The gene duplication found in the core pelecaniforms is a good character for core pelecaniform monophyly, because in situations of rare genomic changes such as gene rearrangements, maximum parsimony can be regarded as a maximum likelihood estimator (Steel and Penny 2000; 2005). These mitochondrial gene duplications (Fig. 5.1) will need to be taken into consideration for pelecaniform population level studies, which often use cyt*b* and the control region in their analyses. It is possible that structure within a population could be inferred incorrectly because of differential amplification of the two control regions.

There is a chance that other taxa in the P&C group also have the gene duplication around the control region. The pelican and the heron (sequenced previously by our group) showed tantalizing hints of a potential duplication (Gibb et al. 2007) but numts (nuclear copies of mtDNA) could not be excluded, and the quality and quantity of the DNA available precluded extensive testing. The shoebill did not show evidence of duplication but again DNA quality was poor. Other groups have sequenced all remaining mitochondrial genomes in the P&C, and we do not know if a gene-duplication was tested for. In some cases only the coding regions of the mitochondrial genome have been published (e.g. Watanabe et al. 2006). The nature of the duplication means that if it is not deliberately excluded by testing otherwise incompatible PCR primers pairs, a duplication may not be detected (Gibb et al. 2007). Outside the P&C, a duplication has not been detected in penguins and loons, but a similar duplication is present in albatross (with only 50 bp of cyt *b* duplicated, Abbott et al. 2005), and may be present in petrels (Lawrence et al. 2008). Alternative duplication/rearrangement patterns have also been detected in a number of other avian lineages (Mindell et al. 1998; Gibb et al. 2007).

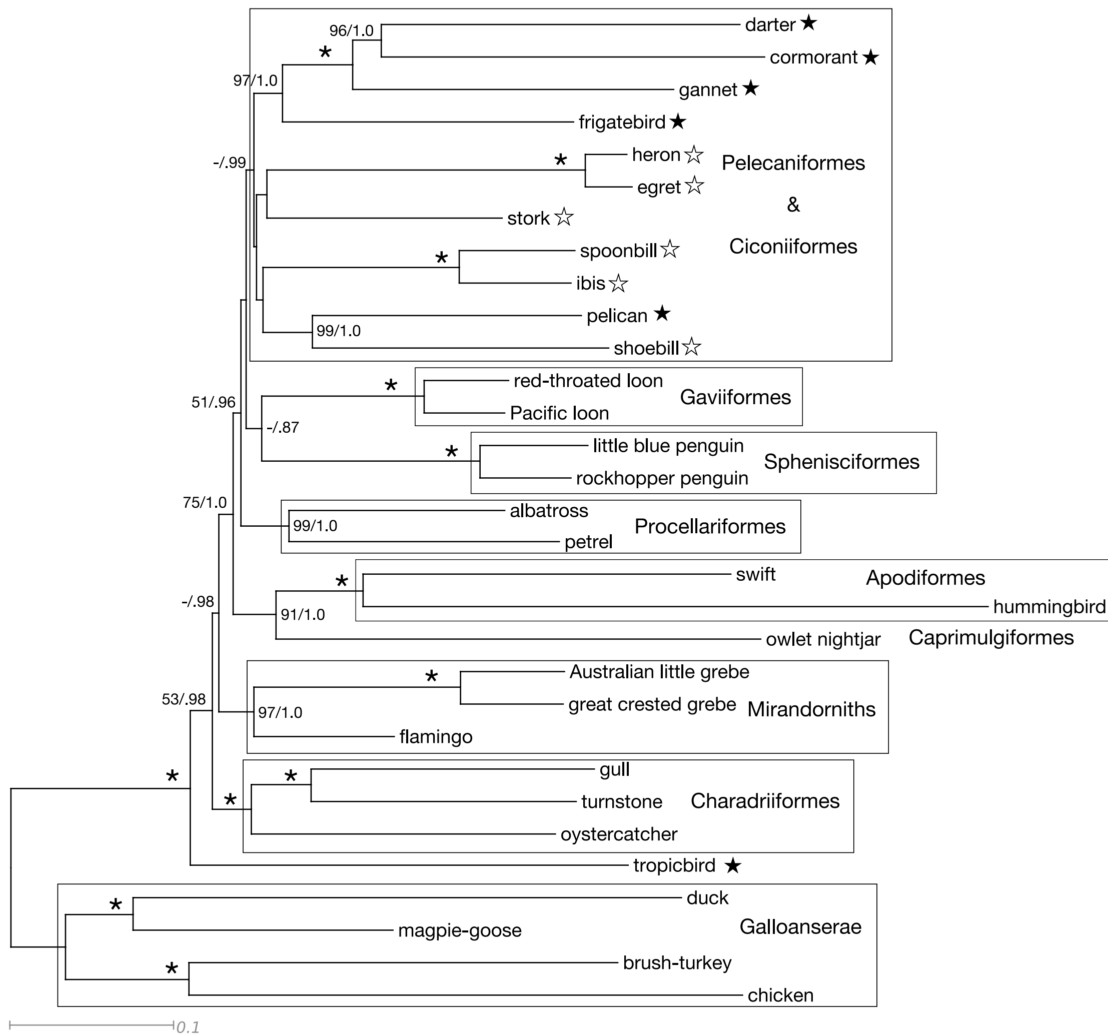

FIGURE 5.2—Maximum likelihood tree from analysis using RAxML, with 31 taxa. Branch support is shown as ML bootstraps over 50%, and Bayesian posterior probabilities over 0.8. Results of 100/1.0 are indicated with an asterisk. The pelecaniform taxa are marked with a black star, and the ciconiiform taxa with a white star.

5.4.3 Phylogenetic Analyses

We begin by analyzing the full mitochondrial dataset (31 taxa, Fig. 5.2), without addition of the partial mitochondrial genome sequences. All the analyses we performed returned essentially the same basic tree, which enhances our confidence in the stability of the output. The analyses were both ML and Bayesian, using the programs Garli, RAxML, MrBayes and BEAST. The latter three used partitioned datasets (as outlined in the methods section). Areas of incomplete resolution between the different trees can be seen in Figure 5.3 (which is a consensus network of the four analytical methods used), and these regions of incomplete resolution also had low support (see Fig. 5.2). It is clear that all four methods find the P&C as

a monophyletic group, and most of the disagreement between the different methods of analysis relate to some of the basal relationships *within* the P&C. Thus these local rearrangements do not affect our discussion of long-term niche stability; indeed they reinforce the conclusion of an adaptive radiation.

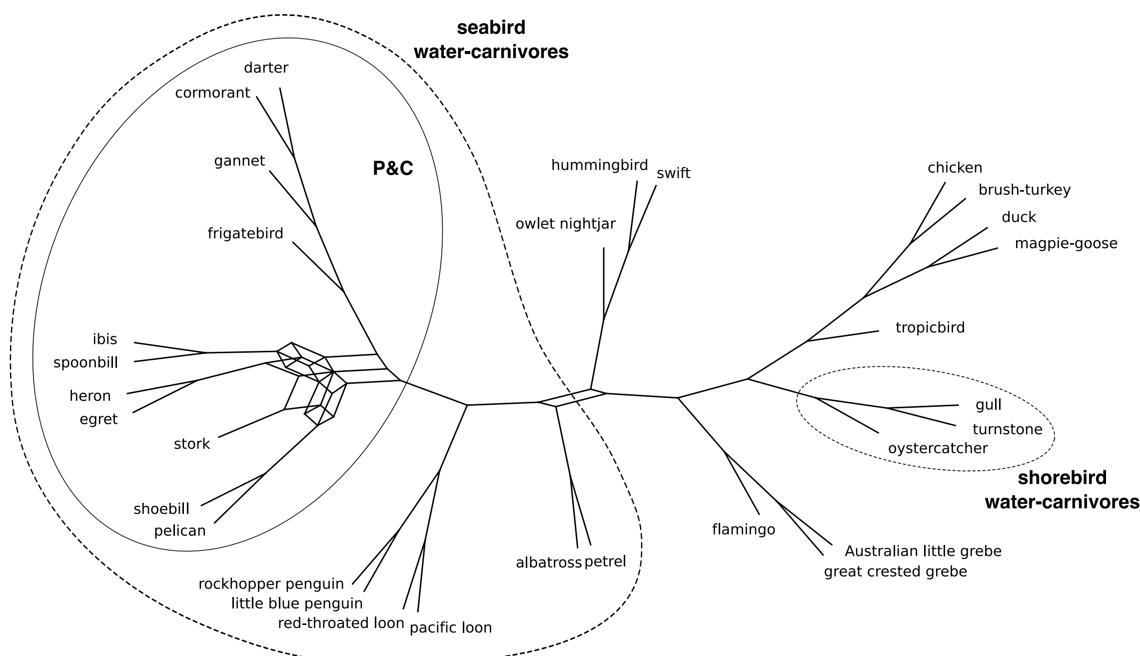

FIGURE 5.3—Consensus network of the four analytical methods used (Garli, RAxML, MrBayes and BEAST) for the 31 taxa dataset, showing areas of agreement and disagreement. Branch lengths are a maximum of four units long, representing agreement between all four methods.

We then included the incomplete mtDNA sequences from the four additional species into our analyses (Fig. 5.4). Because these shorter sequences introduce substantial amounts of missing data into the dataset, we must be cautious when analyzing this dataset. There is a problem in the way missing data is treated in many phylogenetic programs; missing data can be handled as A, T, C or G (for nucleotides). This can mean that rates are over-estimated if only G's and C's are observed at a site (for example). After addition of the four shorter sequences the basic tree topology is consistent with the previous trees (Figs. 5.2 and 5.3), and the four new species grouped with their expected sister taxa (e.g. the booby grouped with the gannet, the tiger-heron with the heron etc.). Many of the basal internal branches within the P&C still have low bootstrap and Bayesian posterior probability support, however, it is important to note that three methods now return the same tree topology (shown in Fig. 5.4). The fourth method, BEAST (Drummond and Rambaut 2007) did not converge on an optimal tree and was discarded (see molecular dating section for further details).

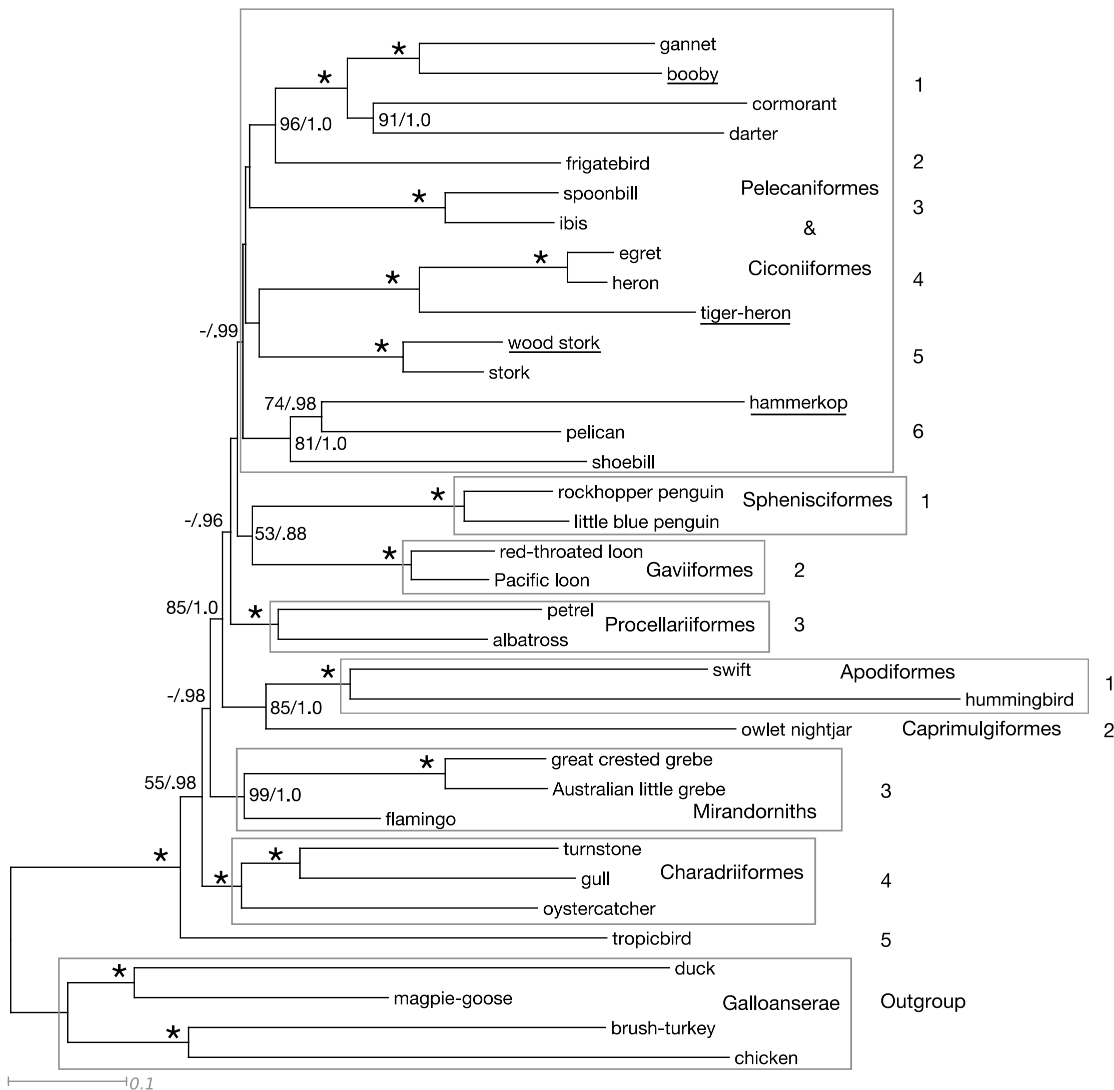

FIGURE 5.4—Maximum likelihood tree from analyses using RAxML of 35 taxa, including the four incomplete mitochondrial genomes (underlined). This tree topology was found by RAxML, Garli and MrBayes analyses. Support at nodes is indicated as in Figure 5.2. The numbers to the right of the tree refer to the 14 general groups described in Appendix 1.

We have also confirmed that removal of the outgroup (Galloanserae) did not greatly affect the ingroup topology (Holland et al. 2003); the only difference related to Mirandorniths (grebe and flamingo) grouping with Charadriiformes (results not shown). Thus, we are confident that neither the addition of the incomplete sequences, nor the outgroup has significantly biased the tree topology. Having assessed the stability of the trees, we can now examine the phylogenies.

5.4.4 Resolution of the Core Pelecaniforms

The darters, gannets, and cormorants group together with 100% bootstrap support and with a Bayesian posterior probability of 1.0; forming the core Pelecaniformes.

Additionally, the frigatebird is always sister to the core Pelecaniformes (Figs. 5.2 and 5.3). Addition of 5 Kb of booby mtDNA sequence confirmed that, as predicted (e.g. Friesen and Anderson 1997), it joins deep on the gannet branch (Fig. 5.3). Perhaps not surprisingly, because a substantial amount of missing data was incorporated, including the booby slightly reduced support for the darter-cormorant group, however, that support is still high.

We find agreement between nuclear and mitochondrial datasets for the grouping of the core pelecaniforms, including the sister frigatebird group. An alternative topology (cormorants basal to darters+sulids) found by Kennedy et al. (2005) using only 1.13 Kb of mtDNA was not supported by our longer mtDNA results, although bootstrap and BPP support for cormorants+darters decreased slightly after addition of the shorter mitochondrial sequences of the booby. This decreased support is possibly an artifact of the missing data introduced into the dataset, and longer sequences for the booby are needed to investigate this. Neither Hackett et al. (2008) nor Ericson et al. (2006) included two sulids (e.g. a gannet and a booby) in their analyses, so we cannot say if this alternate topology has support with nuclear data.

5.4.5 The Pelican is Not a Pelecaniform

There is now an agreement from nuclear and mitochondrial studies that the pelican, shoebill and hammerkop form a monophyletic group and are not sister to the core Pelecaniformes (Fig. 5.4). The pelican groups with the shoebill with very high bootstrap and Bayesian posterior probability support (99% and 1.0 respectively, Fig. 5.2), and this was consistent in all our analyses. When the hammerkop was included, all three clearly group together (81% bootstrap support and a BPP of 1.0), with the hammerkop closest to the pelican (Fig. 5.4, 74% bootstrap support and a BPP of 0.98). This internal group arrangement has not been seen in previous nuclear or shorter mitochondrial analyses. It is possible the grouping is influenced by the hammerkop missing substantial amounts of data when incorporated into our dataset, and full mitochondrial sequences will be needed to investigate this. Analysis of a shorter dataset (5 Kb) with no missing data still returned the same pelican + hammerkop grouping (not shown). There is a small amount of signal for an alternative pelican-shoebill grouping (21% bootstrap support), however, there is almost no support for the grouping found by Hackett et

al. (2008) with nuclear data, where the pelican is sister to both the shoebill and hammerkop. Nevertheless, the three taxa are 'locally stable' (the trees are simply the alternative arrangements around a single internal branch, Cooper and Penny 1997), and this variation does not affect our discussion.

Many morphologists have investigated the possibility that the shoebill is closely related to the pelicans (e.g. Cottam 1957; Saiff 1978; Cracraft 1985). However, in doing so, most suggested that the shoebill is an aberrant pelecaniform, rather than pelicans being 'aberrant long-legged waders' (van Tuinen et al. 2001). The morphological characters linking pelicans to cormorants, darters and sulids (primarily totipalmate feet, but see Table 5.1), appeared to preclude any alternative placements of pelicans, thereby bringing the shoebill into Pelecaniformes. When Cracraft (1985) examined the characters linking the pelicans and shoebills he concluded they were convergences that arose as mechanical responses to similarities in feeding behavior. In the light of the modern DNA-based consensus, we can now say that characters linking pelicans and shoebills represent a shared history, and those tying pelicans to the core pelecaniforms may be either convergent responses to the mechanical stresses of feeding behavior, or else basal to the whole P&C group.

The totipalmate foot, for example, would usually be argued to have evolved in the common ancestor of the pelecaniforms (perhaps as an adaptation for foraging in water). However, given the relationships of the P&C group it appears that the totipalmate foot is either convergent, or has been lost multiple times in the Ciconiiformes. Possible explanations for the evolution of the totipalmate foot in the supposed pelecaniform common ancestor include foraging mode, but these modes have evolved to differ substantially in the extant taxa. Tropicbirds, gannets and boobies are plunge divers, pelicans are mostly surface filterers (though brown pelicans also plunge dive), frigatebirds are entirely aerial feeders that specialize in dipping, and cormorants in foot propelled pursuit diving (Schreiber and Burger 2002). While all of these modes of foraging may benefit from a fully webbed foot in some way, the totiplamate condition cannot have simply evolved just once – it must have evolved independently in at least the tropicbirds, and possibly the pelicans. If it evolved independently in the tropicbirds and the ancestor of the P&C group it would imply several independent losses in the ciconiiforms, whereas if it evolved

independently in the tropicbirds, pelicans and the ancestor of the remaining pelecaniforms no losses are required.

5.4.6 The predicted P&C Group

As expected, the other ciconiiform families are each monophyletic (the storks, the herons and the ibises), and these all have high support. The shorter tiger-heron and wood-stork sequences were added to the dataset specifically to help break the long branches leading to the heron/egret and stork respectively, and theoretically reduce the conflict seen in this area of the tree. This does appear to be the case, as there is less topological uncertainty after the addition of the four extra species (underlined in Fig. 5.4). Bootstrap and Bayesian posterior probability support in the basal branches of the P&C group are still low, and the deeper branches are not yet clearly resolved, supporting the assertion that there has been an adaptive radiation. We did not find the stork basal to all other P&C, as was seen by the intron-laden dataset of Hackett et al. (2008). However, all the pelecaniform taxa (except the tropicbird) do group with the ciconiiform taxa (with a BPP of 0.99).

In Figure 5.4, with four incomplete mitochondrial genomes, the spoonbill/ibis group is basal to the core pelecaniforms. This grouping implies that either the 'wading lifestyle' evolved several times within the P&C group (at least twice, perhaps more often in our phylogeny, depending on the classification of the pelican), or alternatively the core pelecaniforms regained flying/diving sea-foraging after a (possibly freshwater) wading phase. The latter scenario with the re-gain of a lost trait seems less plausible, whereas a wading lifestyle has also evolved in a number of other more distantly related groups, for example, rails, cranes, and flamingoes. In relation to the long legged wading lifestyle, however, it should be noted that 'ground-foraging' appears to be the ancestral state in modern birds. This state occurs in the Palaeognathae and the Galloanserae, and in several deeply diverging Neoaves. Our results also suggest that the pelicans are aberrant waders, rather than the shoebill being an aberrant pelecaniform (Saiff 1978). The alternate topologies inferred by analyses without the shorter sequences would imply fewer shifts to a 'wading lifestyle' (see Figs. 5.2 and 5.5). In the future, full mitochondrial genome sequences, plus additional nuclear data (particularly exon data) are needed for the booby, tiger-heron, wood-stork and hammerkop, to test if the associations seen here are still supported. We predict full-length sequences for these species will

increase the stability of these groups. Unfortunately, there are no deeper extant relatives in any of the ciconiiform families that could further break those long internal branches.

5.4.7 The Nature of the Tropicbird

Our findings (together with those of Hackett et al. 2008) clearly show that the tropicbird does not group with the other pelecaniform taxa, nor even within the greater pelecaniform/ciconiiform group (P&C, Fig. 5.3). The very deep position of the tropicbird on the branch to the outgroup (see Fig. 5.4) is quite likely caused by long-branch attraction, and therefore the exact position of the tropicbird is still unresolved. Removal of the tropicbird did not affect the remaining tree topology (results not shown), though omitting it did increase bootstrap and BPP for the outgroup with the deeper Neoaves.

Thus, the totipalmate foot condition must be discounted as a unifying morphological character for phylogeny. Siegel-Causey (1997) argued that the foot webbing of the tropicbird *and* frigatebird was not the same as the other Pelecaniformes (including the pelican), so this is still not completely in agreement with the molecular results. We have shown the tropicbird does not group with the Caprimulgiformes, nor the flamingo-grebe clade, as was proposed (albeit with low support) by Hackett et al. (2008). Because we are not confident yet of where the tropicbird is placed, all we can currently conclude is that it is a deeply diverging Neoavian lineage, definitely outside the P&C group, and an independent lineage of water-carnivores.

Outside the P&C group, all ordinal groups have very high support, as predicted. Additionally, we find high support for the grouping of flamingoes with grebes (Mirandorniths; >95% bootstrap support and a Bayesian posterior probability of 1.0). This analysis also (confidently) places loons and penguins together with medium support (53% bootstrap support, 0.88 BPP), something not seen in previous phylogenetic studies, mitochondrial or nuclear.

The inclusion of two penguins and two loons has undoubtedly helped increase the support for these two orders grouping together. In previous studies, both nuclear and mitochondrial, often only one of each of penguins and loons has been included

(e.g. Ericson et al. 2006; Watanabe et al. 2006; Hackett et al. 2008; Morgan-Richards et al. 2008; Pratt et al. 2009), which can lead to long-branch attraction (Bergsten 2005). This potential problem has often been compounded by the inclusion of only a few unpaired P&C taxa (especially just one stork or heron). When only some taxa are included there appears to be considerable intermingling of the P&C and their nearest neighbors – loons, penguins and tubenoses (e.g. Pratt et al. 2009). In light of this, many of the local differences between our phylogeny and Hackett et al.'s (2008) findings may relate to the single unpaired stork, ibis, loon and penguin included in their analyses, as well as the predominance of introns in their study (Morgan-Richards et al. 2008).

5.4.8 Molecular Dating

Moving to the dated analyses, Figure 5.5 shows the results of the BEAST analysis with the smaller 31 taxa dataset. This dataset successfully converged on a stable tree topology, with high support and ESS values over 200, as determined using Tracer v1.4 (Rambaut and Drummond 2003). The larger 35 taxa dataset was also analyzed, but successful modeling of missing data in the RY partition in BEAST was challenging. The four additional taxa (with substantial amounts of missing data) led to unstable tree topologies, poor support and low ESS values in the larger dataset, yet dates on nodes in the trees with and without the four incomplete taxa were very similar (results not shown). This finding is in part because many nodes are highly constrained by those around them and therefore have little room to vary.

Evolution of water carnivory appears to have happened relatively quickly and early in Neoaves (approximately 70-80 MA from Fig. 5.5). By 60 MA, already seven extant lineages within the seabird water-carnivores were present. The timing of neoavian water-carnivory evolution (70-80 MA) is consistent with independent evidence suggesting a decline in pterosaur diversity (potential non-avian competitors), and an increase in 'web-footed birds' during the Campanian/Maastrichtian (Kim et al. 2003; Slack et al. 2006). It is pleasing to see 'consilience of induction', in this case similar results from molecular trees and fossil footprint data.

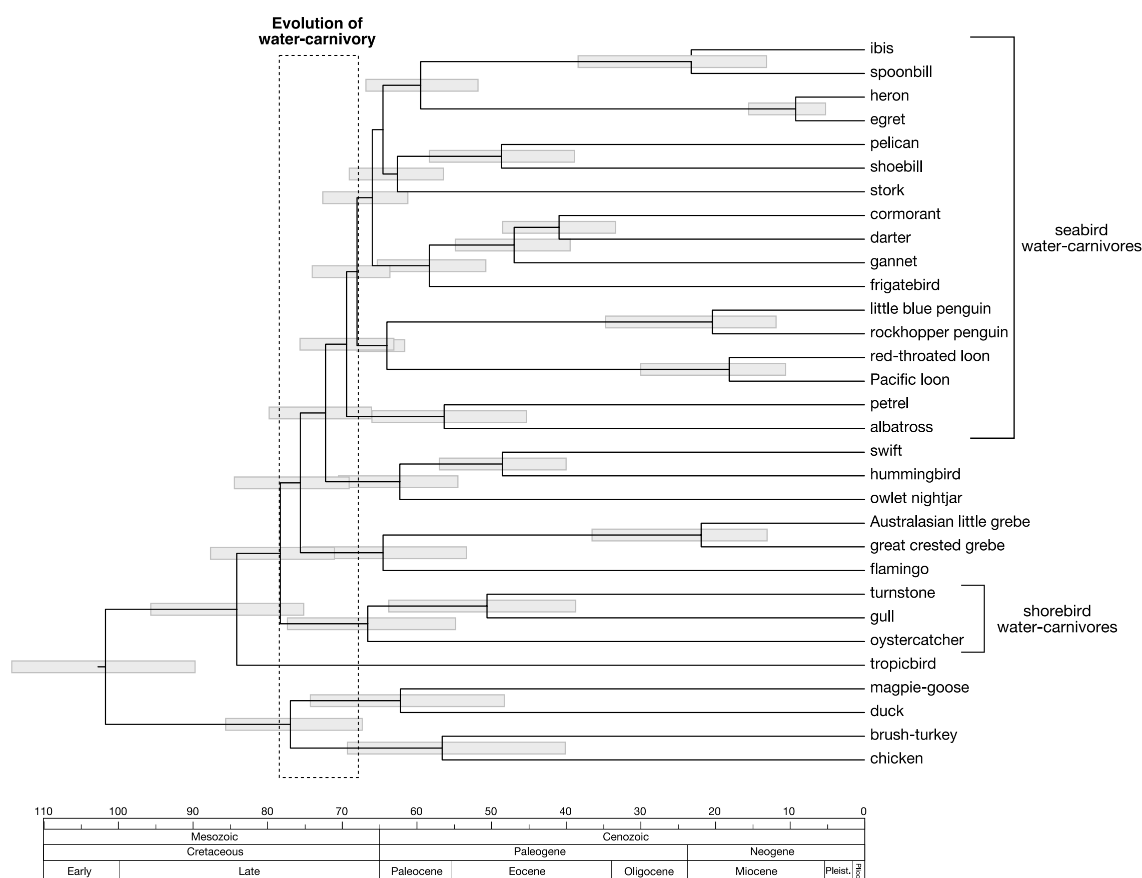

FIGURE 5.5—Maximum clade credibility chronogram of the 31 taxa dataset inferred using BEAST. The time scale is given in millions of years ago (Ma). Error bars represent 95% posterior credibility intervals and are displayed for nodes present in more that 50% of trees sampled.

The core pelecaniform birds (including the frigatebird) began diversifying a short time later, around 58-60 MA. Within the core Pelecaniformes, Friesen and Anderson (1997) estimated the boobies, *Sula*, are a fairly deep relative (>20 million years ago) of the gannets, *Morus*, based on cyt *b* sequences and an estimated transversion rate of 0.2 base pairs per 100 million years. Our preliminary results are consistent with this (results not shown). Additionally, dates are consistent with Oligocene cormorant fossils from Australia (T. Worthy, pers. comm.).

5.4.9 Water-Carnivory in Neoaves, Adaptive Radiations and Niche Stability

Consistent with our predictions from previous findings, we find both the grouping of nine 'seabird water-carnivores' nested within the other Neoavian taxa, plus the six P&C within this nine (described in Appendix 1). Thus, the seabird water-carnivores taxa arrangement from mitochondrial genomes is in agreement with the prediction from Hackett et al (2008). The probability of finding these two pre-

specified groups on new data is ≈1.69×10^{-06} (from Appendix 1). In addition, the optimal tree (Fig. 5.4) has a group of land birds (Apodiformes, and Caprimulgiformes) between the seabird water-carnivores and the shorebirds, and grebes and flamingos. This result was also predicted from our reference tree (Hackett et al. 2008), but is not taken into account during the above calculations. Our test is therefore conservative in this respect.

The P&C are part of the major water-carnivore radiation, which also includes penguins, loons and tubenoses (loosely, seabirds, ~320 species). Distinct from this clade, it appears the other major water-carnivore group (shorebirds – Charadriiformes, ~350 species) evolved separately, and perhaps began radiating slightly later (Fig. 5.5). Among other Neoaves, a few individual species or clades have also adopted some kind of water-carnivory (for example, tropicbird, osprey, fish owl, sea eagles, cranes and kingfishers) but no group is as successful in terms of number of species, or breadth of water-carnivory niches as either the seabird or shorebird groups. Additionally, having evolved water-carnivory, there is a low probability within the seabird water-carnivores of transition out of this niche, even though a number of groups have moved from the sea to freshwater (e.g. ciconiiform families, darters, some cormorants).

We can conclude that the core pelecaniforms, the P&C group and the seabird water-carnivore group are well supported by both nuclear and mitochondrial data. However, within the seabird water-carnivore group, the deep branching is not certain, supporting an early adaptive radiation of this group as a whole. Our results show these lineages diverged from each other beginning at least 70 MA, over a period of about five million years. While it should be possible to resolve branching over a five million year period from the recent past (say, within the Miocene), it stands to reason that resolution becomes more challenging the longer ago this radiation occurred. While it will be hard to resolve short branching times over five Myr, 70-80 MA, an early adaptive radiation into this general niche is probably much more interesting.

Why would there have been long-term niche stability within water-carnivores? The present study is not an in-depth consideration of methods of competition, however, we comment generally on niche stability. Niche conservation at the levels of genus, family and order may be buffered against local variation such as environmental

changes, and may be more affected by intrinsic life history traits (Hadly et al. 2009). Consider the two fitness peaks of the two-state model of Poole et al. (2003) as representing the life history traits 'water-carnivory' and 'higher land birds' (sensu Olson 1985). If, for example, an organism deviates from its optimal fitness in a water-carnivory role and shifts towards a more land bird based niche it will have become less fit than its water carnivore relatives, while not yet approaching optimal fitness in the land bird niche. Thus, in many scenarios, the organism is likely out-competed on both fronts. Additionally, Pond (1977) suggested the ability of birds (and similarly, mammals) to feed their young has allowed birds to specialize more completely in just one niche. Contrast this with reptiles, where a number of different niches are required for juvenile and adult individuals within the one species (Pond 1977), requiring delicate balance to maintain optimal fitness over multiple niches. This is perhaps where we can go 'after phylogeny'.

5.5 CONCLUSION

Increasingly, we see it is essential to integrate the results of phylogenetic research with life history, developmental and behavioral information (or more generally, the biology of organisms). As a start to this longer-term aim we have discussed our findings regarding the Pelecaniformes and their allies. It is now clear from nuclear and mitochondrial data that we have a consensus on pelecaniform phylogeny, as well as an established grouping of pelecaniform and ciconiiform birds within a larger seabird water-carnivore clade.

We see that within the long term stability of the seabird water-carnivore niche there have been local changes in lifestyle and ecology, but no water-carnivores have had ecologically diverse radiations out of this niche, for example becoming 'higher land birds', nor vice versa. Evolutionary-stable niche-discontinuity (ESND) is a theoretical framework for exploring this concept (Poole et al. 2003). Generally, any group diversifying into a new niche must compete not only with those already specialized for that niche, but also with its relatives in the ancestral niche. This competition can lead to long-term stability of general niche boundaries, such as seen here in the water-carnivores. With this work, we can begin to move beyond phylogeny. There is life beyond trees!

5.6 ACKNOWLEDGEMENTS

We would like to thank Blair Hedges for the *Balaeniceps rex* DNA sample, Chris Wood and Burke Museum for the *Anhinga rufa* sample (Burke Museum accession number 779), Clinton Duffy and the Department of Conservation, Hawke's Bay, New Zealand for the *Morus serrator* sample, and Lyndsey Chadderton and the Department of Conservation, Southland, New Zealand for the *Phalacrocorax chalconotus* sample. Thank you to Mike Hendy for assistance with the calculation of probabilities. Thanks also to Klaus Schliep, Simon Hills and Matt Phillips for helpful discussions involving BEAST. The Allan Wilson Centre for Molecular Ecology and Evolution supported this work.

5.7 REFERENCES

Abbott, C. L., M. C. Double, J. W. H. Trueman, A. Robinson, and A. Cockburn. 2005. An unusual source of apparent mitochondrial heteroplasmy: duplicate mitochondrial control regions in *Thalassarche* albatrosses. Mol Ecol 14:3605-3613.

Beddard, F. E. 1898. The structure and classification of birds. Longmans, Green and Co., London.

Bergsten, J. 2005. A review of long-branch attraction. Cladistics 21:163-193.

Brown, J. W., J. S. Rest, J. Garcia-Moreno, M. D. Sorenson, and D. P. Mindell. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol 6:6.

Clarke, J. A., C. P. Tambussi, J. I. Noriega, G. M. Erickson, and R. A. Ketcham. 2005. Definitive fossil evidence for the extant avian radiation in the Cretaceous. Nature 433:305-308.

Cooper, A., and D. Penny. 1997. Mass survival of birds across the Cretaceous-Tertiary boundary: Molecular evidence. Science 275:1109-1113.

Cottam, P. A. 1957. The Pelecaniform characters of the skeleton of the Shoebill Stork, Balaeniceps rex. Bull Brit Mus (Nat Hist) Zool Series 5:49-72.

Cracraft, J. 1985. Monophyly and phylogenetic relationships of the pelecaniformes: a numerical cladistic analysis. Auk 102:834-853.

Cracraft, J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proc R Soc Lond B Biol Sci 268:459-469.

de Kloet, R. S., and S. R. de Kloet. 2005. The evolution of the spindlin gene in birds: Sequence analysis of an intron of the spindlin W and Z gene reveals four major divisions of the Psittaciformes. Mol Phylogenet Evol 36:706-721.

Drummond, A., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed Phylogenetics and Dating with Confidence. PLoS Biol 4:e88.

Ericson, P., C. Anderson, T. Britton, A. Elzanowski, and U. Johansson. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett 2:543-547.

Fain, M. G., and P. Houde. 2004. Parallel radiations in the primary clades of birds. Evolution 58:2558-2573.

Friesen, V. L., and D. J. Anderson. 1997. Phylogeny and evolution of the Sulidae (Aves: Pelecaniformes): A test of alternative modes of speciation. Mol Phylogenet Evol 7:252-260.

Gibb, G. C., O. Kardailsky, R. T. Kimball, E. L. Braun, and D. Penny. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol Biol Evol 24:269-280.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Hadly, E. A., P. A. Spaeth, and C. Li. 2009. Niche conservatism above the species level. P Natl Acad Sci 106:19707-19714.

Hedges, S. B., and C. G. Sibley. 1994. Molecules vs. morphology in avian evolution: the case of the "pelecaniform" birds. P Natl Acad Sci USA 91:9861-9865.

Ho, S. Y. W., and M. J. Phillips. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58:367-380.

Holland, B. R., D. Penny, and M. D. Hendy. 2003. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock - A simulation study. Syst Biol 52:229-238.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference on phylogenetic trees. Bioinformatics 17:754-755.

Huson, D., D. Richter, C. Rausch, T. Dezulian, M. Franz, and R. Rupp. 2007. Dendroscope: An interactive viewer for large phylogenetic trees. BMC Bioinformatics 8:460.

Huson, D. H., and D. Bryant. 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol 23:254-267.

Kennedy, M., B. R. Holland, R. D. Gray, and H. G. Spencer. 2005. Untangling long branches: Identifying conflicting phylogenetic signals using spectral analysis, neighbor-net, and consensus networks. Syst Biol 54:620-633.

Kennedy, M., and H. G. Spencer. 2004. Phylogenies of the frigatebirds (Fregatidae) and tropicbirds (Phaethonidae), two divergent groups of the traditional order Pelecaniformes, inferred from mitochondrial DNA sequences. Mol Phylogenet Evol 31:31-38.

Kennedy, M., H. G. Spencer, and R. D. Gray. 1996. Hop, step and gape: do the social displays of the Pelecaniformes reflect phylogeny? Animal Behaviour 51:273-291.

Kim, C. B., M. Huh, C. S. Cheong, M. G. Lockley, and H. W. Chang. 2003. Age of the pterosaur and web-footed bird tracks associated with dinosaur footprints from South Korea. The Island Arc 12:125-131.

Lawrence, H. A., G. A. Taylor, C. D. Millar, and D. M. Lambert. 2008. High mitochondrial and nuclear genetic diversity in one of the world's most endangered seabirds, the Chatham Island Taiko *Pterodroma magentae*. Conserv Genet 9:1293-1301.

Liu, K., S. Raghavan, S. Nelesen, C. R. Linder, and T. Warnow. 2009. Rapid and Accurate Large-Scale Coestimation of Sequence Alignments and Phylogenetic Trees. Science 324:1561-1564.

Livezey, B. C., and R. L. Zusi. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool J Linn Soc 149:1-95.

Mayr, G. 2008. Avian higher-level phylogeny: well-supported clades and what we can learn from a phylogenetic analysis of 2954 morphological characters. Journal of Zoological Systematics and Evolutionary Research 46:63-72.

Mindell, D. P., M. D. Sorenson, and D. E. Dimcheff. 1998. Multiple independent origins of mitochondrial gene order in birds. P Natl Acad Sci USA 95:10693-10697.

Morgan-Richards, M., S. Trewick, A. Bartosch-Härlid, O. Kardailsky, M. J. Phillips, P. A. McLenachan, and D. Penny. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol Biol 8:20.

Nelson, B. 2005. Pelicans, cormorants and their relatives: Pelecanidae, Sulidae, Phalacrocoracidae, Anhingidae, Fregatidae, Phaethontidae. Oxford University Press, Oxford.

Olson, S. L. 1985. The fossil record of birds. Pp. 79-252 *in* D. S. Farner, J. R. King, and K. C. Parkes, eds. Avian Biology, Vol. 8. Academic Press, New York.

Paterson, A. M., G. P. Wallis, and R. D. Gray. 1995. Penguins, petrels and parsimony: does cladistic analysis of behavior reflect seabird phylogeny? Evolution 49:974-989.

Paton, T. A., and A. J. Baker. 2006. Sequences from 14 mitochondrial genes provide a well-supported phylogeny of the Charadriiform birds congruent with the nuclear RAG-1 tree. Mol Phylogenet Evol 39:657-667.

Paton, T. A., A. J. Baker, J. G. Groth, and G. F. Barrowclough. 2003. RAG-1 sequences resolve phylogenetic relationships within Charadriiform birds. Mol Phylogenet Evol 29:268-278.

Phillips, M. J., G. C. Gibb, L. Crimp, and D. Penny. 2010. Tinamous and moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. Syst Biol:in press.

Pond, C. M. 1977. The Significance of Lactation in the Evolution of Mammals. Evolution 31:177-199.

Poole, A. M., M. J. Phillips, and D. Penny. 2003. Prokaryote and eukaryote evolvability. Biosystems 69:163-185.

Pratt, R. C., G. C. Gibb, M. Morgan-Richards, M. J. Phillips, M. D. Hendy, and D. Penny. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol 26:313-326.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor.

Rambaut, A., and A. J. Drummond. 2003. Tracer: MCMC Trace Analysis Tool. University of Oxford.

Saiff, E. I. 1978. The middle ear of the skull of birds: the Pelecaniformes and Ciconiiformes. Zool J Linn Soc 63:315-370.

Schreiber, E. A., and J. Burger, eds. 2002. Biology of Marine Birds. CRC Press, Boca Raton, Florida.

Sibley, C. G., and J. E. Ahlquist. 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale University Press, New Haven, Connecticut.

Siegel-Causey, D. 1997. Phylogeny of the Pelecaniformes: Molecular systematics of a privative group. Pp. 159-171 *in* D. P. Mindell, ed. Avian Molecular Evolution and Systematics. Academic Press, San Diego, CA.

Slack, K. E., F. Delsuc, P. A. Mclenachan, U. Arnason, and D. Penny. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol Phylogenet Evol 42:1-13.

Slack, K. E., C. M. Jones, T. Ando, G. L. A. Harrison, R. E. Fordyce, U. Arnason, and D. Penny. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol 23:1144-1155.

Slikas, B. 1997. Phylogeny of the avian family Ciconiidae (storks) based on cytochrome b sequences and DNA-DNA hybridization distances. Mol Phylogenet Evol 8:275-300.

Slikas, B. 1998. Recognizing and Testing Homology of Courtship Displays in Storks (Aves: Ciconiiformes: Ciconiidae). Evolution 52:884-893.

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Syst Biol 57:758-771.

Steel, M., and D. Penny. 2000. Parsimony, Likelihood, and the Role of Models in Molecular Phylogenetics. Mol Biol Evol 17:839-850.

Steel, M. A., and D. Penny. 1993. Distributions of Tree Comparison Metrics-Some New Results. Syst Biol 42:126-141.

Steel, M. A., and D. Penny. 2005. Maximum parsimony and the phylogenetic information in multistate characters. Pp. 163-178 *in* V. Albert, ed. Parsimony, phylogeny and genomics. Oxford University Press, Oxford (UK).

van Tets, G. F. 1965. A comparative study of some social communication patterns in the Pelecaniformes. Ornithological Monographs 2:1-88.

van Tuinen, M., D. B. Butvill, J. A. W. Kirsch, and S. B. Hedges. 2001. Convergence and divergence in the evolution of aquatic birds. Proc R Soc Lond B Biol Sci 268:1345-1350.

Watanabe, M., M. Nikaido, T. T. Tsuda, T. Kobayashi, D. Mindell, Y. Cao, N. Okada, and M. Hasegawa. 2006. New candidate species most closely related to penguins. Gene 378:65-73.

Wong, K. M., M. A. Suchard, and J. P. Huelsenbeck. 2008. Alignment uncertainty and genomic analysis. Science 319:473-476.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD dissertation. University of Texas, Austin, Texas.

5.8 APPENDIX 1

5.8.1 Predicting the probabilities of observing splits in a new dataset.

As our reference point we take the tree of Hackett et al. (2008), which is based on nuclear sequences (though mostly introns). We then test whether the P&C grouping and seabird water-carnivores (P&C plus penguins/loons/albatrosses) grouping occurs in the optimal tree from the mitochondrial data. In particular, we calculate the probability of finding these two groupings.

The species in the two datasets (nuclear and mitochondrial) are not identical, but we consider 14 groups that are in common (see below, and the numbered groups shown in Fig. 5.4). More than one taxon can be in a group, as long as the membership of the group is generally accepted. For example, the two penguins form a group, as do the two loons. Thus, our test is the prediction that the same six basic clusters of P&C taxa (see below) is also found in the mitochondrial data, and with a further three groups (penguins/loons/albatross) immediately outside (or sister) to that group. Then we have five other Neoaves as the immediate outgroup to these, and with four Galloanserae (chickens and ducks and their relatives) to root the overall tree.

The six groups of P&C are (core pelecaniforms; frigate bird; ibis/spoonbill; pelican/hammerkop/shoebill; herons; storks). Then, as mentioned above, we expect the (loons; penguins; petrel/albatross) grouping outside the P&C. Finally we expect the other Neoaves groups (hummingbird/swift; owlet nightjar; Charadriiformes (shorebirds); tropicbird; Mirandorniths (flamingo/grebe)) and with the root coming within this latter group. In addition, we expect that the shorebirds, and the flamingo/grebe clade are not adjacent to the seabird water-carnivores and we consider the pelican, hammerkop and shoebill to be a single group; both of these expectations make our quantitative test conservative. If, for example, we allowed pelican, hammerkop and shoebill to be three separate groups, we would have 16 groups, lowering the probabilities still further. However, we prefer the more cautious approach.

This approach gives a subgroup of nine taxa in the seabird water-carnivore group (consisting of six P&C subgroups, and the three subgroups of loons; penguins; and

petrel/albatross) on one side of the primary split. These nine groups are predicted to be a monophyletic group within the other five Neoaves groups (giving a total of 14). In other words, we have 14 taxa, and we calculate the probabilities of a finding a tree, by chance, with the specified groups of five and nine taxa, with the nine then having an additional split separating the specified subgroups of three and six taxa. The root must occur within the group of five.

We do the probability calculation in two steps, based on all trees being equally likely (Steel and Penny 1993). The first is calculating the probability of finding the primary division of nine taxa within the rooted tree of 14 taxa, which is then multiplied by the probability of six taxa monophyletic within the subtree of nine. We follow the probability calculations from Appendix A of Pratt et al. (2009), where they show the probability of k taxa being monophyletic within a rooted tree of n taxa is

$$R(n,k) = R(k) \times R(n - k + 1) / R(n)$$

So the probability of 9 monophyletic taxa within 14 is

$$R(14,9) = R(9) \times R(14 - 9 + 1) / R(14)$$
$$= 0.000242$$

and the probability of 6 monophyletic taxa within the 9 is

$$R(9,6) = (R(6) \times R(4)) / R(9)$$
$$= 0.00699301$$

These are independent, and so the overall probability of observing the two groupings in a tree on new data is the product,

$$= 0.000242 \times 0.00699301 = 1.69436 \text{ x } 10^{-06}$$

Future work could see the development of an equivalent Bayesian analysis, perhaps using Bayes factors, from the proportion of trees in a converged MrBayes run with the predicted splits (Alexei Drummond, pers. comm.).

CHAPTER SIX

6 NEW ZEALAND PASSERINES HELP RESOLVE BASAL OSCINE PHYLOGENY

6.1 ABSTRACT

Despite the increasing amount of sequence information and improving phylogenetic methods, the basal nodes of Passeriformes, the largest avian radiation, are still poorly resolved. The basal phylogeny of oscine passerines is important for interpreting the radiation of this group out of the Australasian region. We report three new complete mitochondrial genomes from basal oscine lineages (tui, saddleback and hihi). Our well-resolved phylogeny is in agreement with recent analyses of certain nuclear exons and introns, but partly disagrees with phylogenies based on Rag-1, one of the most commonly used nuclear exons in passerine phylogenetics. We find several so-called 'basal corvid' taxa are more likely basal to Passerida. This study also reports 1.8 Kb of mitochondrial DNA from the extinct New Zealand piopio. We identify that the previously published piopio cyt *b* is incorrect, and show the piopio is not a basal bowerbird, but falls within the core Corvoidea.

Keywords—ancient DNA, *Turnagra*, Passeriformes, mitochondria, oscine biogeography

6.2 INTRODUCTION

Passerines (perching birds) are the most speciose avian lineage, and oscine songbirds (the major clade within passerines), account for almost half of all bird species (Clements 2007). Many have attempted to resolve oscine phylogeny, but the basal branches of the group are still incompletely resolved. A well-resolved oscine phylogeny will provide a framework for understanding the biogeography and ancestral ecology of this diverse, worldwide group.

6.2.1 Oscine biogeography

The DNA-DNA hybridization studies of Sibley and Ahlquist (1990) found oscines formed two sister groups (suborders Passerida and Corvida) that are broadly European and Australasian respectively. Further resolution from short nuclear sequences have determined that Corvida is both basal and paraphyletic with respect to Passerida (Barker et al. 2002; Ericson et al. 2002; Barker et al. 2004). The preponderance of Australian and New Zealand taxa in these basal corvid lineages, and the discovery in Australia of the oldest passerine fossils (Boles 1995) have strengthened the hypothesis that oscines evolved in the Australo-Papuan region.

Sister to oscines, suboscines form two clades, broadly New World and Old World in composition. The species-rich New World clade originated in South America, and the relatively species-poor Old World lineages are distributed through Africa, Madagascar and Asia (Barker et al. 2004). Basal to all passerines is the tiny clade of New Zealand wrens (Acanthisittidae). Biogeographic hypotheses regarding the worldwide distribution of passerines often begin by assuming the isolation of the NZ wrens on the Zealandian landmass since it rifted from Australia 82 million years ago (Barker et al. 2004). Vicariance and the breakup of Gondwana are then often posited as the major influence on the diversity of Passerines found today (reviewed in Ericson et al. 2003), although dispersal must then be inferred for a number of more recent distributions. Because the NZ wrens are a single isolated lineage (both phylogenetically and geographically), it is actually very hard to infer anything about their distribution and composition through time (Trewick and Gibb 2010), which makes it particularly hard to test passerine biogeography hypotheses. Greater understanding and phylogenetic resolution of the modern New Zealand

element in the basal branches of the oscine lineage will help in understanding both oscine and passerine extinction and biogeography.

6.2.2 Limitations of current deep oscine phylogenetic resolution

A recent supertree summary of available genetic phylogenies (mitochondrial and nuclear) has created a basic framework hypothesis for oscine passerines (Jønsson and Fjeldså 2006). Irestedt and Ohlson (2008) noted that the vast majority of signal in studies comprising this supertree comes from just one nuclear gene, the 3 Kb Rag-1 exon. Removal of Rag-1 revealed underlying signal for alternative trees from short nuclear introns and exons (Irestedt and Ohlson 2008). Generally, exons would be considered better for deep phylogenetic resolution because they are slower evolving and more conserved, thus easier to align and analyse. Introns can be difficult to align when deeply divergent, due to the large number of indels (Wong et al. 2008; Liu et al. 2009). This can lead to signal for an incorrect tree, such as the metaves/coronaves split proposed by Fain and Houde (2004) (Morgan-Richards et al. 2008). However, at the medium depth of oscine resolution (between families within an order), it may be that introns will provide important resolution.

Whether or not introns or exons will prove more useful in oscine phylogeny, successful resolution of basal oscine nodes likely requires more sequence information (longer sequences) than has been used to date. There are no studies of deep oscine phylogeny that use more than 5 Kb of data. Most avian studies that do use longer sequences were focused more on the resolution of neoaves than passerines, and so have included only a few oscine taxa (e.g. Slack et al. 2007; Hackett et al. 2008; Pratt et al. 2009).

It is also important to have agreement between independent datasets such as mitochondrial, nuclear and morphological datasets. Mitochondrial genomes bridge the gap between exons (which perhaps evolve too slowly) and introns (which can be difficult to align at deeper divergences). It is relatively straightforward to sequence complete genomes from highly divergent avian species using conserved primers. A conservative alignment of protein and RNA genes provides 13 Kb of sequence after the removal of indels and non-coding sequence. Because the mitochondrion is a single molecule, all genes on it can be considered to have the same gene tree, hopefully reducing the incidence of incongruence seen with Rag-1 above.

Passerines have also been difficult to resolve morphologically, because convergent evolution in life history traits is common. Recent genetic studies have discovered convergence in a range of traits, such as adaptation for nectarivory of both New Zealand and Hawaiian honeyeaters (Ewen et al. 2006; Fleischer et al. 2008), evolution of toxicity in the paraphyletic *Pitohui* (Dumbacher et al. 2008; Jønsson et al. 2008) and behavior and plumage in birds of paradise (Paradisaeidae, Cracraft and Feinstein 2000).

Only three oscines have been included in complete mitochondrial genome analyses before (rook, indigobird and lyrebird). To improve oscine phylogeny, we analyse six additional oscine mitochondrial genomes, three from basal corvid New Zealand genera (hihi, saddleback and tui), and three previously published, but unanalysed, passerid taxa (zebrafinch, warbler and blackcap). These analyses help in understanding the makeup of the New Zealand avifauna, and strengthen basal oscine phylogenetic resolution. In addition, we report novel mtDNA sequence from the extinct New Zealand piopio, and show that previously published piopio cyt*b* (Christidis et al. 1996) has been misidentified. In light of this, we also analyse a range of available mtDNA genes to reexamine where the piopio fits in the passerine tree.

6.3 METHODS

6.3.1 DNA extraction, PCR and sequencing of modern samples.

The tui (*Prosthemadera novaeseelandiae*) was sourced from Waikanae Department of Conservation, New Zealand. The saddleback (*Philesturnus carunculatus*) and hihi (*Notiomystis cincta*) blood samples were from birds on Tiritiri Matangi Island. DNA was extracted from blood using Roche High Pure PCR template preparation kit (Roche Applied Science, Mannheim, Germany) as per the manufacturers instructions.

Complete mitochondrial genomes were initially amplified in 2-3 overlapping segments using the Expand Long Template PCR System (Roche). These long-range PCR products were used as templates for subsequent PCR of short overlapping fragments. This process is described in more detail in Gibb et al. (2007) and references therein. All short PCR products were amplified using primers identified

from our database (available from the authors on request) employing standard protocols (e.g. Slack et al. 2007).

6.3.2 Ancient DNA protocols

The piopio (*Turnagra capensis*) was from the South Australian Museum (sample number B12353). Collection details record the location as South Island New Zealand. A photo of the specimen is included as Supplementary Material. DNA extraction and PCR were performed using strict ancient DNA procedures at the Australian Centre for Ancient DNA, University of Adelaide, South Australia.

Total genomic DNA was extracted from piopio toepad using proteinase K digestion followed by Phenol Chloroform extraction. Fragments 100-350bp in length were amplified using a combination of both standard short PCR with Platinum *Taq* (Invitrogen) and multiplex PCR following Römpler et al. (2006). Duplicate PCR of cyt*b* from the same DNA extraction was also carried out in the ancient DNA laboratory of the Allan Wilson Centre, Massey University, Palmerston North, New Zealand, with identical sequence results. Primer sequences and exact protocol conditions are available from the authors on request.

In addition to the three new mt genomes reported here, 18 published genomes were included in the analysis (nine passerines and nine other neognathes). They are rifleman (*Acanthisitta chloris*; GenBank accession number AY325307), grey-headed broadbill (*Smithornis sharpei*; AF090340), fuscous flycatcher (*Cnemotriccus fuscatus*; AY596278), superb lyrebird (*Menura novaehollandiae*; AY542313), village indigobird (*Vidua chalybeata*; AF090341), rook (*Corvus frugilegus*; Y18522), zebrafinch, (*Taeniopygia guttata*; DQ422742), Eurasian reed warbler (*Acrocephalus scirpaceus*; AM889139), blackcap (*Sylvia atricapilla*; AM889140), chicken (*Gallus gallus*; AP003317), Australian brush-turkey (*Alectura lathami*; AY346091), magpie goose (*Anseranas semipalmata*; AY309455), redhead duck (*Aythya americana*; AF090337), dollarbird (*Eurystomus orientalis*, EU344978) white-tailed trogon (*Trogon viridis,* EU410490), New Zealand kingfisher (*Todyramphus sancta vagans*, EU410489), ivory billed toucan (*Pteroglossus azara*, DQ780882), and pileated woodpecker (*Dryocopus pileatus*, DQ780879).

6.3.3 Alignment

Individual mitochondrial genomes were assembled from sequenced PCR products using Sequencher 4.8 (Gene Codes Corp, Ann Arbour, MI) and checked by eye. Multiple genomes were manually aligned and checked in Se-Al v2.0a11 (Rambaut 1996), at the amino acid level for protein-coding genes, and based on stem and loop secondary structure for RNA genes. Gaps, ambiguous sites adjacent to gaps, the ND6 (light-strand encoded), non-coding regions and stop codons (often incomplete in the DNA sequence), were excluded from the alignment. Protein coding regions have 3rd position coded as RY (as explained in Gibb et al. 2007) and are separated into five partitions, three by codon position, and then RNA stems and loops. The complete mitochondrial genome dataset is 13,548 base pairs long, and is available from our website http://www.allanwilsoncentre.ac.nz/downloads.htm.

TABLE 6.1—Genbank accessions of additional samples used in piopio analyses, in addition to the complete mitochondrial genomes listed in Methods. Common names are from Clements (2007).

Latin Name	Common Name	12S	16S	ND1	ND2	CoI	Atp6/8	Cytb
Acanthiza chrysorrhoa	Yellow-rumped Thornbill	AY488247	AF129183		AY488317	AF197852		FJ821109
Ailuroedus crassirostris	Green Catbird				AY064750			U10371
Callaeas cinerea	Kokako	DQ469308			DQ469298			DQ469302
Cnemophilus macgregorii	Crested Satinbird					AF197842		AF197841
Colluricincla megarhyncha	Rufous Shrike-thrush				EF592276			FJ821114
Coracina melaschistos	Black-winged Cuckoo-shrike	AF386464	AF391229		AY529948			AF376894
Erpornis zantholeuca[1]	White-bellied Erpornis	AF376921	AF376901					DQ092886
Grallaria squamigera	Undulated Antpitta	AY139636		AY139637 (ND1-WANCY)				AF127188
Hemispingus frontalis	Oleaginous Hemispingus	AF447235		AY139640 (ND1-WANCY)		AF383098	AF447335	AF383020
Malurus lamberti	Variegated Fairy-wren	AY488256			AY488326		AY192106	AY488402
Meliphaga gracilis	Graceful Honeyeater	AY488215			AY488288			AY353241
Melipotes fumigatus	Smoky Honeyeater	AY488216			AY488289	AF197864		AY488361
Mohoua albicilla	Whitehead							FJ821127
Oreoica gutturalis	Crested Bellbird				EF592285		EF592440	FJ821130
Oriolus xanthornus	Black-hooded Oriole		AF094645		AY529964			AF094615
Pachycephala pectoralis	Golden Whistler				EF592288		EF592443	FJ821134
Paradisaea raggiana	Raggiana Bird-of-Paradise				EF592295	AF197828		EF592241
Pardalotus striatus	Yellow-tipped Pardalote	AY488252			AY488322	AF197848		AF197847
Petroica rosea	Rose Robin				EF592297			EF592242
Philemon corniculatus	Noisy Friarbird	AY488226			AY488299	AF197856		AY488372
Pitohui nigrescens	Black Pitohui				EF592321		EF592477	FJ821138

Latin Name	Common Name	12S	16S	ND1	ND2	CoI	Atp6/8	Cytb
Ptilonorhynchus violaceus	Satin Bowerbird				EU341431	AF197833	AY621550	X74256
Ptiloris magnificus	Magnificent Riflebird				AY064761			AY228092
Rhipidura rufifrons	Rufous Fantail				DQ468898			FJ821142
Sayornis phoebe	Eastern Phoebe	U83765		AF536747 (ND1-WANCY)				AF536744
Sericornis frontalis	White-browed Scrubwren	AY488253	AF129171		AY488323	AF197850		FJ821144
Sphecotheres vieilloti	Australian Figbird							FJ821145
Turnagra capensis	Piopio[2]							U51734
Turnagra capensis	Piopio[3]	GAXXXXX		GAXXXXX	GAXXXXX	GAXXXXX		GAXXXXX
Vireo olivaceus	Red-eyed Vireo	AY136584			AY136614	AY666271	AY115280	X74260
Vireolanius leucotis	Slaty-capped Shrike-Vireo				AY030130	AY030175		AF081959

[1]Listed as *Yuhina* on GenBank

[2]Christidis et al. (1996)

[3]This publication

To more fully examine the phylogenetic relationship of the piopio, available short mitochondrial sequences from a range of passerines were added to the complete mt genome alignment. Previous analysis suggests the piopio is sister to bowerbirds (Ptilonorhynchidae, Christidis et al. 1996), but other contrasting hypotheses include whistlers (Pachycephalidae, Sibley and Monroe 1991) and satinbirds (Cnemophilidae, Olson et al. 1983), necessitating inclusion of taxa to test these hypotheses. These sequences and accession numbers are recorded in Table 6.1. Of course, not all mt genes are available for all passerine species, and so the data matrix includes missing data. Importantly, all gaps in this alignment represent missing data, and are not the result of indels or ambiguous alignments. This piopio dataset was further partitioned to separate the longer 'backbone' regions with less sequence coverage from the shorter 'detail' regions with fuller coverage.

6.3.4 Phylogenetic Analyses

Data was analyzed using Maximum Likelihood methods implemented in Garli (Zwickl 2006) and RAxML (Stamatakis et al. 2008), using the GTR model with estimated gamma and invariant sites, the RAxML analyses were also partitioned. Bayesian analyses were performed using MrBayes (Huelsenbeck and Ronquist 2001). Partitions were unlinked except for branch lengths within the 'backbone' and 'detail' partition groups of the piopio alignment. Datasets were run for five million generations, and sampled every 5,000 generations after a burnin of 500,000 generations. Independent runs were checked for convergence, and trace files analysed using Tracer (Rambaut and Drummond 2003) to ensure ESS values greater than 200. Trees were viewed using SplitsTree 4.8 (Huson and Bryant 2006).

6.4 RESULTS & DISCUSSION

6.4.1 Mitogenomic features

The tui (16,850 bp, accession number GYXXXXX) and the saddleback (16,822 bp, GYXXXX) both have the standard gene order as found in the chicken (Desjardins and Morais 1990). The hihi mitochondrial genome (GYXXXXX) is 18,645 bp long and contains a novel gene order, involving duplication of the region tRNA Pro-CR. The duplicated regions are completely identical for 1,802 bp including the genes tRNA Pro, NADH6, and tRNA Glu. The only differences between the two duplicates

are a different number of C's in the goose hairpin and 15 bp at the end of CR (1) that is unalignable with the last 115 bp of CR (2) before the start of tRNA Phe. This gene order differs from previously reported avian mitochondrial gene orders, as tRNA Thr is not part of the duplicated region (see Fig. 1, Gibb et al. 2007).

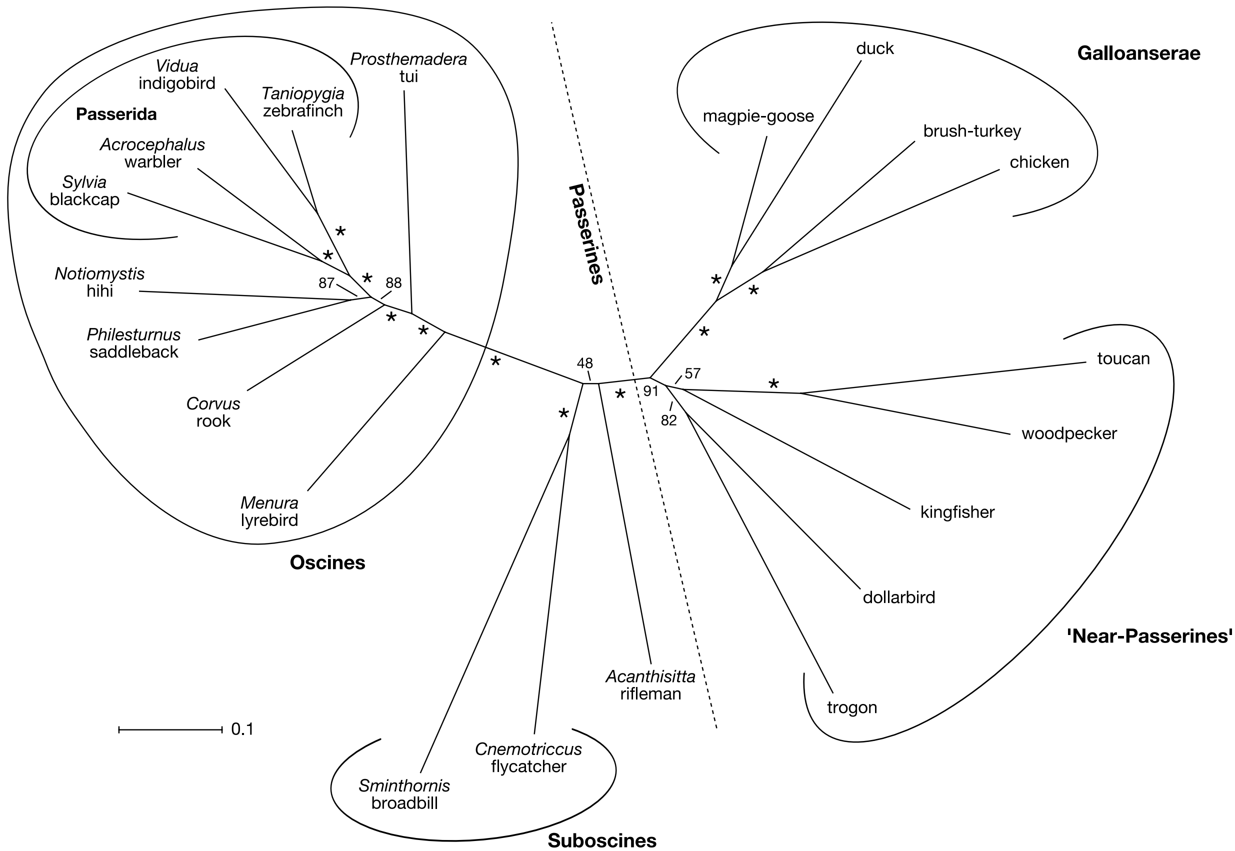

FIGURE 6.1—Unrooted maximum likelihood analysis of complete mitochondrial genome dataset, with bootstrap support, * indicates 100% support. Latin genera are included for species also shown in Figures 6.2 and 6.3, full species names are shown in methods.

6.4.2 *Phylogenetic analyses*

We begin by reporting the results of the full mt genome analysis. The full mtDNA tree (Fig. 6.1) has high bootstrap support, and the same tree was found by all three methods. As expected, passerines form three groups, oscines, suboscines and the basal *Acanthisitta* rifleman. Within oscines, the lyrebird is basal, and Passerida is monophyletic and nested within Corvida, which are therefore paraphyletic. This is as expected from previous nuclear datasets (Barker et al. 2004; Hackett et al. 2008). The warbler, blackcap and zebrafinch have not been included in previous mitochondrial phylogenetic analyses. As expected, the warbler and blackcap pair,

and are sister to the zebrafinch and indigobird. These four taxa represent two of the three broad Passerida lineages, Sylvioidea, Passeroidea and Muscicapoidea (Sibley and Ahlquist 1990). There are no Muscicapoidea mitochondrial genomes available yet.

The hihi and saddleback are sister taxa, as predicted from nuclear analyses (Ewen et al. 2006; Driskell et al. 2007). Interestingly, the pair are sister to the Passerida. This is contra to the results of studies including the Rag-1 gene (e.g. Barker et al. 2004) where the Callaeatidae are sister to the core Corvoidea, here represented by the rook. That Callaeatidae are sister to the Passerida was predicted by Irestedt and Ohlsen (2008), who analysed short introns and exons from a mixture of kokako and saddleback (both Callaeatidae). The different results seen with some intron, exon and mitochondrial data highlights the danger of simply combining different datasets in search of 'total evidence' without first checking for congruence between datasets. In the case of the Callaeatidae, it appears Rag-1 overpowered the signal for an alternative topology now found independently by intron, exon and mitochondrial data (Fig. 1 and Irestedt and Ohlson 2008). This is akin to the case of the difficult-to-align β-fibrinogen intron 7 that provided overriding signal supporting the Metaves/Coronaves split (Fain and Houde 2004) even in the presence of additional genes with alternate signal (Ericson et al. 2006).

6.4.3 Ancient DNA sequence verification

DNA fragments of the extinct piopio were successfully amplified from six mitochondrial gene regions (12S, ND1 (two parts), WANCY tRNAs, CoI, ND5 and cyt*b*). The total sequence length amplified is 1,783 bp. All sequences successfully align with other passerine mt genes, and are consistent for amino acid translation and RNA stem-loop structure. We compared our new cyt*b* sequence to U51734 previously published by Christidis et al. (1996) and found that of the 307 bp that overlap between them, 45 are different (14.7%). This is a comparable number of differences as between this study's piopio cyt*b* and any other oscine cyt*b* in our alignment. Both sequences translate successfully, as many differences are third codon position changes, however six amino acids are different. All alternative amino acids occur at variable positions within our alignment of passerines. A BLAST search of GenBank found neither cyt*b* sequence is closely related to any other published passerine cyt*b* (checked November 2009).

6.4.4 Phylogenetic analysis of the piopio with broader oscine sampling

To date we have only sequenced a few small regions of piopio mtDNA, and plan to sequence more. The discovery that our piopio sequence is very different from that previously published prompted us to determine its phylogenetic affinity relative to the published *Turnagra* cyt *b*, as much as is possible from the sequence information we have. Beginning with the backbone of complete mt genomes, additional mitochondrial gene sequences were included from a selection of passerine species selected to cover the range of putative piopio relatives (Table 6.1).

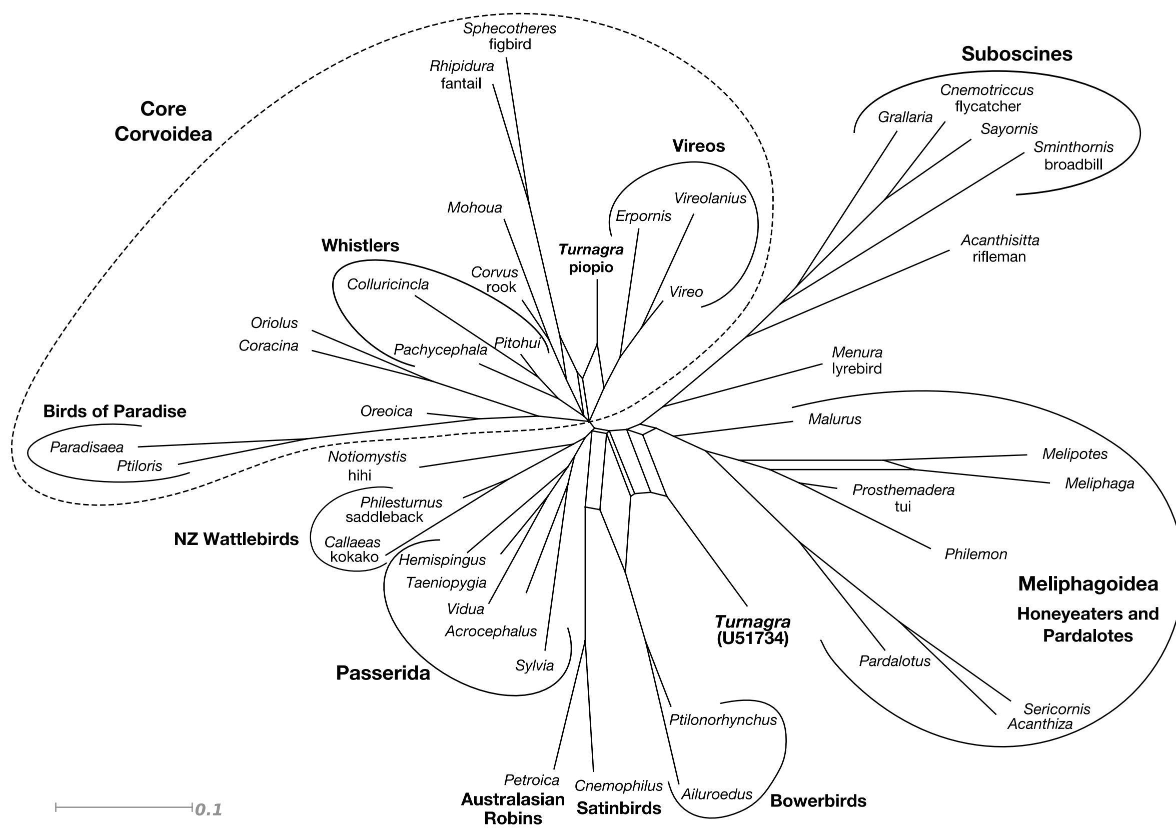

FIGURE 6.2—Network of Bayesian analysis showing splits present in more that 25% of trees from the 43 passerine dataset including *Turnagra* cyt*b* U51734. Species and groups referred to in text by common names are indicated; full species names are in Table 6.1.

Initially the published *Turnagra* cyt *b* sequence was included in the analyses (Fig 6.2). In both maximum likelihood and Bayesian analyses this sequence was clearly problematic, showing no obvious affinity, and causing serious problems with convergence between MrBayes runs. The Bayesian network in Figure 6.2 shows some of the variability introduced into the analyses by the *Turnagra* U51734

sequence, while also identifying that our new piopio sequences are affiliated with a different part of the tree. Because the provenance of the published *Turnagra* sequence is uncertain, and may not be correct, further analyses were undertaken without inclusion of this sequence.

The current piopio work has been undertaken by one person from removal of the toepad to phylogenetic analysis. All laboratory procedures have been carried out under strict aDNA protocols, and no other modern or ancient passerine work was being undertaken in any laboratory at the same time. The Museum of Victoria *T. capensis* accession number B19053 recorded in Christidis et al. (1996) is correctly identified as a piopio (Wayne Longmore, MoV ornithology collection manager, pers. comm.). Discussion between T. Worthy and L. Christidis have strengthened our conclusion that the cyt*b* sequence published in Christidis et al. (1996) is not a genuine *Turnagra capensis* B19053 sequence.

Analyses of our new sequences clearly show that the piopio falls within the core Corvoidea (Fig. 6.3). Resolution within the core Corvoidea is low, partially because many species in this clade have little available mitochondrial sequence data (Table 6.1). However, the piopio tends to group with the vireos and there is also some signal for affiliation with *Rhipidura* (fantails) and *Sphecotheres* (figbirds). At this level of resolution, the piopio is not showing affiliation with many previously predicted groups, such as bowerbirds, birds of paradise and satinbirds (cnemophiline birds of paradise, Olson et al. 1983), so this is a novel, but perhaps not unexpected finding.

The conclusions drawn from this dataset (Fig. 6.3), which contains a large amount of missing data, are only preliminary. Encouragingly, many of the groups we would expect to find from prior knowledge (nuclear, mitochondrial and morphological) were also recovered in these analyses. For example within the oscines there is clear separation of the lyrebird, Meliphagoidea, core Corvoidea, and a group containing both Passerida and the remaining so-called ‘basal corvids’ (sensu Barker et al. 2004). The relationships found with only complete mitochondrial sequences (in Fig. 6.1) are still recovered in this larger analysis (Fig. 6.3). This increases the confidence we can place on the position of the piopio.

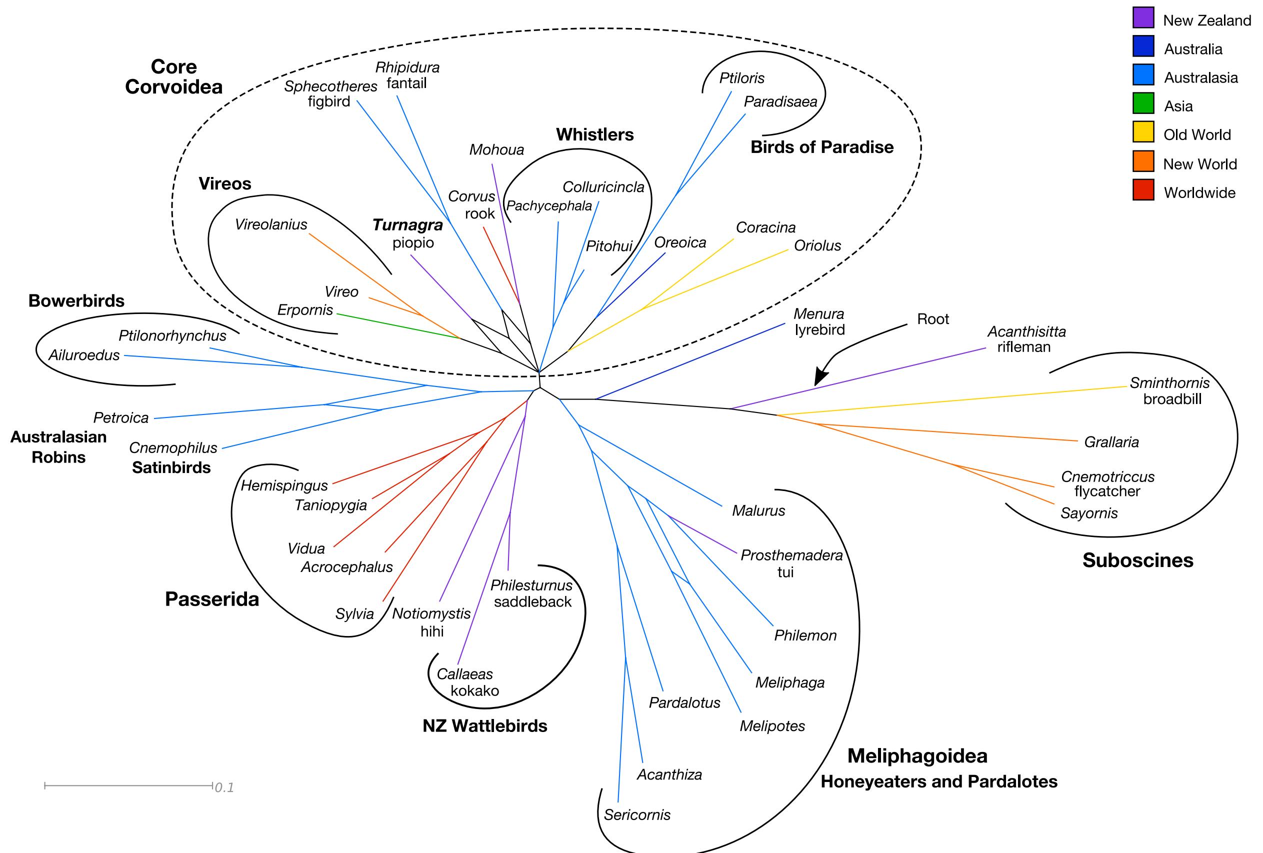

FIGURE 6.3—Network of Bayesian analysis without *Turnagra* U51734. Splits present in more that 25% of trees are shown, species and groups referred to in text by common names are indicated. Full species names are listed in Table 6.1.

Some of the results of this analysis differ to previous analyses, and warrant further testing with more complete mitochondrial datasets. For example here the figbird tends to group with the fantail with 22% bootstrap support (this increases to 64% nearest-neighbour bootstrap support, indicating 'local stability', Penny et al. 1999), whereas analyses including Rag-1 group the figbird with orioles (Barker et al. 2004). Norman et al. (2009) found figbirds and orioles grouped in Rag-1 and myo-2 analyses, but were unresolved in their 50% majority-rule cyt*b* phylogeny. Here as well, it will be interesting to see which associations are predominantly found by Rag-1, and which by multiple different datasets.

The grouping of bowerbirds, satinbirds and NZ robins, although weakly supported at the moment, is another interesting case requiring further resolution. Bowerbirds and satinbirds are two more groups of 'basal corvids' that were found much deeper in the passerine tree with Rag-1 analyses. Other nuclear genes (Irestedt and Ohlson

2008) have also suggested a Passerida affinity for satinbirds, however bowerbirds were still found between the Meliphagoidea and lyrebird.

We have confirmed many deep oscine groups have an Australian and New Zealand distribution. As the phylogeny continues to be more fully resolved, we will be able to examine how many times oscines have 'escaped' out of Australasia. From just our current analyses, there are three separate core Corvoidea clades with distribution outside Australasia, as well as the Passerida (Fig. 6.3). While outside the scope of this study, with regards to Passerida, it will be interesting to determine if this lineage originally escaped just once (with later distribution around the world, including back into Australasia) or multiple times.

6.5 FUTURE WORK

Complete mitochondrial genome sequencing is underway for three further New Zealand oscines. The grey warbler is a relative of thornbills and pardalotes, expected to pair with the tui, and the fantail (*Rhipidura*) is a core corvid, which will hopefully help piopio resolution (Fig. 6.3). Finally, the North Island robin (*Petroica*) was found to be sister to the Passerida by the Rag-1 study of Barker et al. (2004) but our current short mitochondrial sequence studies show the NZ wattlebird/hihi group may be closer to Passerida than the robins (Fig. 6.3). With these additional taxa, all New Zealand basal corvid groups will have been sampled.

6.5.1 Dating Passeriformes

Our longer-term goal is to be able to confidently date the passerine radiation, however dating lineages within the passerines is a difficult challenge. There are no good deep fossils for calibration, because most passerine fossils are found in Europe, belong to crown groups, and are relatively recent (Mourer-Chauviré et al. 1989). The only deep passerines are early Eocene fossils from Australia (~54 Ma) reported by Boles (1995, 1997). Biogeography supports an Australasian origin with at least four lineages (Passerida and three corvid) migrating to the rest of the world. Unfortunately, these Eocene fossils cannot confidently be assigned to any lineage within passerines (such as oscines), so the confidence interval in using this date for any node would be large. Nevertheless, because they are found in Australia, this is consistent with the oscine radiation in that region

Acanthisittidae are basal to all other passerines, and the extant members have limited flight capabilities (Heather and Robertson 2005). It has been very tempting to conclude they have been isolated on the New Zealand landmass since Zelandia separated from Gondwana 82 Ma, and almost all dating of passerines has used this vicariant calibration (e.g. Ericson et al. 2002; Barker et al. 2004; Shepherd and Lambert 2007). Recent dating of large avian datasets, using fossil calibrations external to passerines and not the 82 Ma calibration point have varied wildly in estimates of passerine origins. Dates have ranged from 107 Ma (4.6 kb mtDNA Brown et al. 2008) to 60 Ma (5 kb nuclear data, but no Acanthisittidae, Ericson et al. 2006). Recent studies with more sequence length include Hackett et al. (2008) with 32 kb of nuclear data, who did not date their analysis and Pratt et al. (2009) with whole mitochondrial genomes who found Acanthisitta diverged about 61 Ma.

It is difficult to date passerines using external fossils, because passerines are a relatively fast evolving group. The best avian fossil calibrations are in Galloanserae (Clarke et al. 2005) and penguins (Slack et al. 2006), which are both relatively slowly evolving lineages. A dataset with calibrations on a few slowly evolving outgroups could lead to incorrect tree topology, with passerines drawn to the root of the tree, and overestimation of passerine node ages. It is probable any dating of passerine lineages using the above calibrations will need to have broad avian sampling.

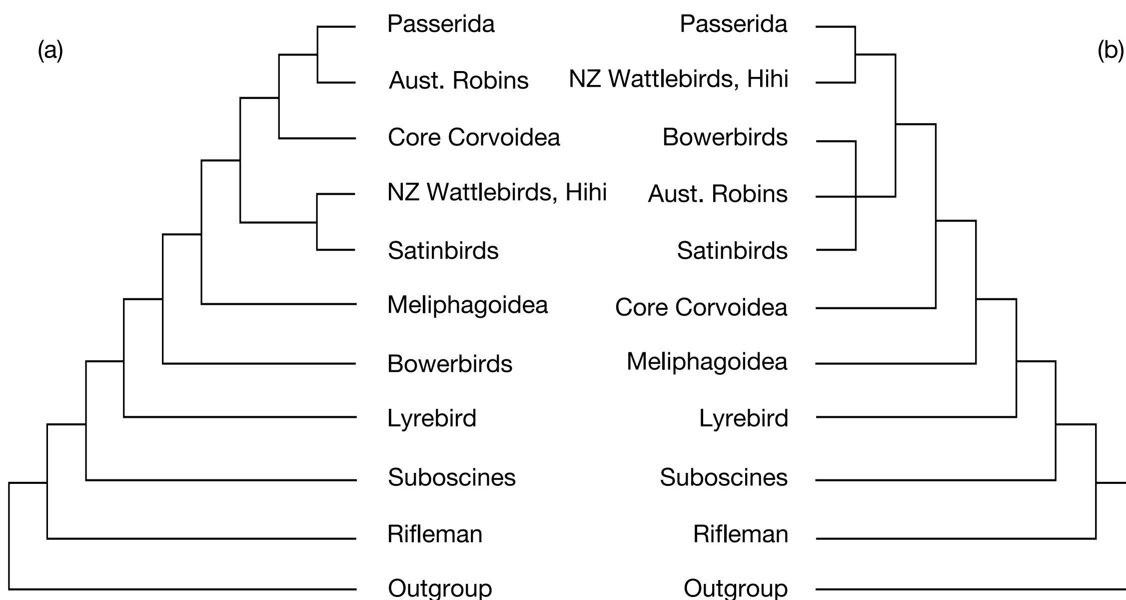

FIGURE 6.4—Consensus cladograms of passerine phylogeny highlighting the different arrangement of basal oscine groups found by (a) Rag-1 (Barker et al. 2004) and (b) preliminary mitochondrial analyses.

6.6 CONCLUSION

Here we begin improving the resolution of basal oscines, highlighting the importance of testing phylogenetic hypotheses with multiple datasets. We find congruence in tree topology between our mitochondrial datasets and short nuclear datasets that exclude the exon Rag-1 (Fig. 6.4). These preliminary results suggest basal nodes within oscines are resolvable with mitochondrial genome sequences. Additionally we find previously published *Turnagra* cytb sequence is incorrect, and the piopio *Turnagra capensis* is likely a core corvid species, not a basal bowerbird.

6.7 ACKNOWLEDGEMENTS

Thanks to John Ewan and Phil Casey for supplying hihi and saddleback blood samples, and Dick Gill, Waikanae Department of Conservation for the tui sample. Thanks to Philippa Horton and Trevor Worthy for access to the piopio specimen at the South Australian Museum. Thank you to Alan Cooper and Jeremy Austin at the Australian Centre for Ancient DNA for advice, assistance, and use of their facilities for ancient DNA work. Finally thank you to Wayne Longmore, Trevor Worthy and Les Christidis for help verifying the Piopio specimen at the Museum of Victoria. The Allan Wilson Centre for Molecular Ecology and Evolution supported this work.

6.8 REFERENCES

Barker, F. K., G. F. Barrowclough, and J. G. Groth. 2002. A phylogenetic hypothesis for passerine birds: taxonomic and biogeographic implications of an analysis of nuclear DNA sequence data. Proc R Soc Lond B Biol Sci 269:295-308.

Barker, F. K., A. Cibois, P. Schikler, J. Feinstein, and J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. P Natl Acad Sci USA 101:11040-11045.

Boles, W. E. 1995. The world’s oldest songbird. Nature 374:21-22.

Boles, W. E. 1997. Fossil songbirds (Passeriformes) from the Early Eocene of Australia. Emu 97:43-50.

Brown, J. W., J. S. Rest, J. Garcia-Moreno, M. D. Sorenson, and D. P. Mindell. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol 6:6.

Christidis, L., P. R. Leeton, and M. Westerman. 1996. Were bowerbirds part of the New Zealand fauna? P Natl Acad Sci USA 93:3898-3901.

Clarke, J. A., C. P. Tambussi, J. I. Noriega, G. M. Erickson, and R. A. Ketcham. 2005. Definitive fossil evidence for the extant avian radiation in the Cretaceous. Nature 433:305-308.

Clements, J. F. 2007. The Clements Checklist of Birds of the World. Cornell University Press, Ithaca, New York.

Cracraft, J., and J. Feinstein. 2000. What is not a bird of paradise? Molecular and morphological evidence places Macgregoria in the Meliphagidae and the Cnemophilinae near the base of the corvoid tree. Proc R Soc Lond B Biol Sci 267:233-241.

Desjardins, P., and R. Morais. 1990. Sequence and gene organization of the chicken mitochondrial genome : A novel gene order in higher vertebrates. Journal of Molecular Biology 212:599-634.

Driskell, A. C., L. Christidis, B. J. Gill, W. E. Boles, F. K. Barker, and N. W. Longmore. 2007. A new endemic family of New Zealand passerine birds: adding heat to a biodiversity hotspot. Aust J Zool 55:73-78.

Dumbacher, J. P., K. Deiner, L. Thompson, and R. C. Fleischer. 2008. Phylogeny of the avian genus *Pitohui* and the evolution of toxicity in birds. Mol Phylogenet Evol 49:774-781.

Ericson, P., C. Anderson, T. Britton, A. Elzanowski, and U. Johansson. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett 2:543-547.

Ericson, P. G. P., L. Christidis, A. Cooper, M. Irestedt, J. Jackson, U. S. Johansson, and J. A. Norman. 2002. A Gondwanan origin of passerine birds supported by DNA sequences of the endemic New Zealand wrens. Proc R Soc Lond B Biol Sci 269:235-241.

Ericson, P. G. P., M. Irestedt, and U. S. Johansson. 2003. Evolution, biogeography, and patterns of diversification in passerine birds. J Avian Biol 34:3-15.

Ewen, J. G., I. Flux, and P. G. P. Ericson. 2006. Systematic affinities of two enigmatic New Zealand passerines of high conservation priority, the Hihi or Stitchbird *Notiomystis cincta* and the Kokako *Callaeas cinerea*. Mol Phylogenet Evol 40:281-284.

Fain, M. G., and P. Houde. 2004. Parallel radiations in the primary clades of birds. Evolution 58:2558-2573.

Fleischer, R., H. James, and S. L. Olson. 2008. Convergent evolution of Hawaiian and Australo-Pacific honeyeaters from distant songbird ancestors. Cur Biol 18:1927-1931.

Gibb, G. C., O. Kardailsky, R. T. Kimball, E. L. Braun, and D. Penny. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol Biol Evol 24:269-280.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Heather, B., and H. A. Robertson. 2005. The Field Guide to the Birds of New Zealand. Penguin Books, Auckland.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference on phylogenetic trees. Bioinformatics 17:754-755.

Huson, D. H., and D. Bryant. 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol 23:254-267.

Irestedt, M., and J. Ohlson. 2008. The division of the major songbird radiation into Passerida and 'core Corvoidea' (Aves: Passeriformes) — the species tree vs. gene trees. Zoologica Scripta 37:305-313.

Jønsson, K. A., R. C. K. Bowie, J. A. Norman, L. Christidis, and J. Fjeldså. 2008. Polyphyletic origin of toxic Pitohui birds suggests widespread occurrence of toxicity in corvoid birds. Biol Lett 4:71-74.

Jønsson, K. A., and J. Fjeldså. 2006. A phylogenetic supertree of oscine passerine birds (Aves: Passeri). Zoologica Scripta 35:149-186.

Liu, K., S. Raghavan, S. Nelesen, C. R. Linder, and T. Warnow. 2009. Rapid and Accurate Large-Scale Coestimation of Sequence Alignments and Phylogenetic Trees. Science 324:1561-1564.

Morgan-Richards, M., S. Trewick, A. Bartosch-Härlid, O. Kardailsky, M. J. Phillips, P. A. McLenachan, and D. Penny. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol Biol 8:20.

Mourer-Chauviré, C., M. Hugueney, and P. Jonet. 1989. Découverte de Passeriformes dans l'Oligocène Supérieur de France. Comptus Rendus Académie des Sciences Paris (series 2) 309:843-849.

Norman, J. A., P. G. P. Ericson, K. A. Jønsson, J. Fjeldså, and L. Christidis. 2009. A multi-gene phylogeny reveals novel relationships for aberrant genera of Australo-Papuan core Corvoidea and polyphyly of the Pachycephalidae and Psophodidae (Aves: Passeriformes). Mol Phylogenet Evol 52:488-497.

Olson, S. L., K. C. Parkes, M. H. Clench, and S. R. Borecky. 1983. The Affinities of the New Zealand Passerine Genus *Turnagra*. Notornis 30:319-336.

Penny, D., M. Hasegawa, P. J. Waddell, and M. D. Hendy. 1999. Mammalian evolution: Timing and implications from using the LogDeterminant transform for proteins of differing amino acid composition. Syst Biol 48:76-93.

Pratt, R. C., G. C. Gibb, M. Morgan-Richards, M. J. Phillips, M. D. Hendy, and D. Penny. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol 26:313-326.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor.

Rambaut, A., and A. J. Drummond. 2003. Tracer: MCMC Trace Analysis Tool. University of Oxford.

Römpler, H., P. H. Dear, J. Krause, M. Meyer, N. Rohland, T. Schöneberg, H. Spriggs, M. Stiller, and M. Hofreiter. 2006. Multiplex amplification of ancient DNA. Nature Protocols 1:720-728.

Shepherd, L. D., and D. M. Lambert. 2007. The relationships and origins of the New Zealand wattlebirds (Passeriformes, Callaeatidae) from DNA sequence analyses. Mol Phylogenet Evol 43:480-492.

Sibley, C. G., and J. E. Ahlquist. 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale University Press, New Haven, Connecticut.

Sibley, C. G., and B. L. Monroe. 1991. Distribution and Taxonomy of Birds of the World. Yale University Press, New Haven, CT.

Slack, K. E., F. Delsuc, P. A. Mclenachan, U. Arnason, and D. Penny. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol Phylogenet Evol 42:1-13.

Slack, K. E., C. M. Jones, T. Ando, G. L. A. Harrison, R. E. Fordyce, U. Arnason, and D. Penny. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol 23:1144-1155.

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Syst Biol 57:758-771.

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

Wong, K. M., M. A. Suchard, and J. P. Huelsenbeck. 2008. Alignment uncertainty and genomic analysis. Science 319:473-476.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD dissertation. University of Texas, Austin, Texas.

6.9 SUPPLEMENTARY MATERIAL

(a)

(b)

FIGURE S6.1—Piopio specimen *Turnagra capensis* (labeled here *T. crassirostris*, a junior synonym of *T capensis* Sparrman) from the South Australian Museum. (a) Ventral, (b) Dorsal view.

CHAPTER SEVEN

7 THESIS SUMMATION

This thesis set out to improve deep resolution of birds generally, and answer some specific questions regarding avian phylogenetics in the New Zealand context. Real progress has been made in all of these areas.

The review of New Zealand avifauna in **Chapter Two** laid out the current understanding from molecular phylogenetics with respect to New Zealand birds. For some taxa we have a lot of information on closest relatives, genetic diversity and age of radiation in New Zealand and around the World, for others we have almost nothing, and there are still large holes in our knowledge. It is now possible to see where assumptions may have been made with almost no molecular work to support them, and situations where molecular analyses have highlighted just how malleable species are. The work presented in this thesis has helped fill some of those gaps, and highlighted other areas where more work is required.

In **Chapter Three**, we set out a hypothesis for why the apparently rare genomic change of control region gene rearrangement appears to occur frequently and can be maintained for long evolutionary periods. This mechanism means gene order is not currently a useful phylogenetic tool for resolving deep Neoaves. Further work since the publication of Gibb et al. (2007) has continued to show examples where gene rearrangement and duplication around the control region is widespread and maintainable over long evolutionary periods. This is exemplified by the arrangements seen in cuckoos (Pratt et al. 2009), the core pelecaniforms (**Chapter Five**), and in some passerines (**Chapter Six**, and Singh et al. 2008). As more work

is done regarding mitochondrial gene rearrangements, the processes involved in creating and maintaining them will become better understood, refining the model put forward in **Chapter Three**. We hope that gene order, when all the alternatives are better known, will be a useful subsidiary phylogenetic criterion.

Chapter Three and **Appendix One** have shown that by increasing taxon sampling, breaking long branches and enhancing the signal in the data, in principle Neoaves is resolvable, although a lot more work needs to be done. There has been a great increase in DNA sequence data since this project was begun, and the advent of second-generation sequencing technology will only increase this. The research of Hackett et al. (2008) published 32 Kb of nuclear DNA from 169 avian species, a significant increase in data. This research highlights areas of Neoaves that look like they are now becoming resolved, and also those that need more analysis and testing with other datasets. It is interesting to see where their results agree and disagree with the findings of our work and others.

While many deep divergences are becoming clear with comparison of mitochondrial and nuclear datasets (e.g. **Chapters Five**, **Six** and **Appendix Two**), we may need longer nuclear exon sequences and rare genomic changes (or SCUDs) to successfully resolve some basal Neoaves divergences (e.g. pigeons, **Chapter Four**). It is likely that introns are more successful than exons at resolving shallower divergences (for example within and between families). The use of techniques to increase signal on internal versus external branches, such as RY coding of third codon positions, and site stripping (Pratt et al. 2009) will also help resolution.

The largest published avian exon dataset is the 7.4 Kb portion of Hackett et al. (2008), and a detailed analysis of this dataset by itself has not been published. In the longer term, 7 Kb may not be enough and more exon sequences may be needed to successfully resolve Neoaves. Next generation sequencing will greatly accelerate data collection (e.g. B. McComish and S. Hills submitted), and one way to greatly increase exon sequencing by direct targeting using the second-generation sequencing technology. For example, by targeting highly expressed exons using mRNA from a single tissue extraction, such as blood or liver.

We are now approaching one hundred published avian mitochondrial genomes (even after removing duplicate zebrafinches and chickens). No more than 45 have ever been included in one analysis (Pratt et al. 2009). With the addition of those presented in this thesis, and only a few more key genomes (for example another rail to pair with the Takahe), there is a significant dataset available for study. We can examine the phylogenies found independently by mitochondrial, exon and intron datasets, and significantly explore how well we are able to resolve the phylogeny of birds, and particularly the basal branches of Neoaves.

As a very preliminary introduction to this, I have done a quick initial analysis of available mitochondrial genomes (Fig. 7.1). The dataset contains 81 avian taxa, and is a partitioned RAxML (Stamatakis et al. 2008) bootstrapped analysis. As in previous work, the data is partitioned into first, second and third protein codon positions, and RNA stems and loops. The third codon has been RY coded. The potoo, takahe, kagu, turkey vulture and tropicbird have been omitted, as all are known to currently cause problems such as long-branch attraction, and will require further work, such as taxa to pair with for better resolution (Chapter Seven, and Gibb et al. 2007; Morgan-Richards et al. 2008; Pratt et al. 2009). Taxa with close relatives were also omitted, such as additional *Gallus spp*.

The dataset has not yet been site stripped, nor other measures to help increase signal on internal branches (except RY coding of the third position). The network shown in Figure 7.1 however gives an indication of the current signal in the data. Areas that are resolved, and those with the most conflict, where the most work will be required for resolution are obvious (polytomies and boxes).

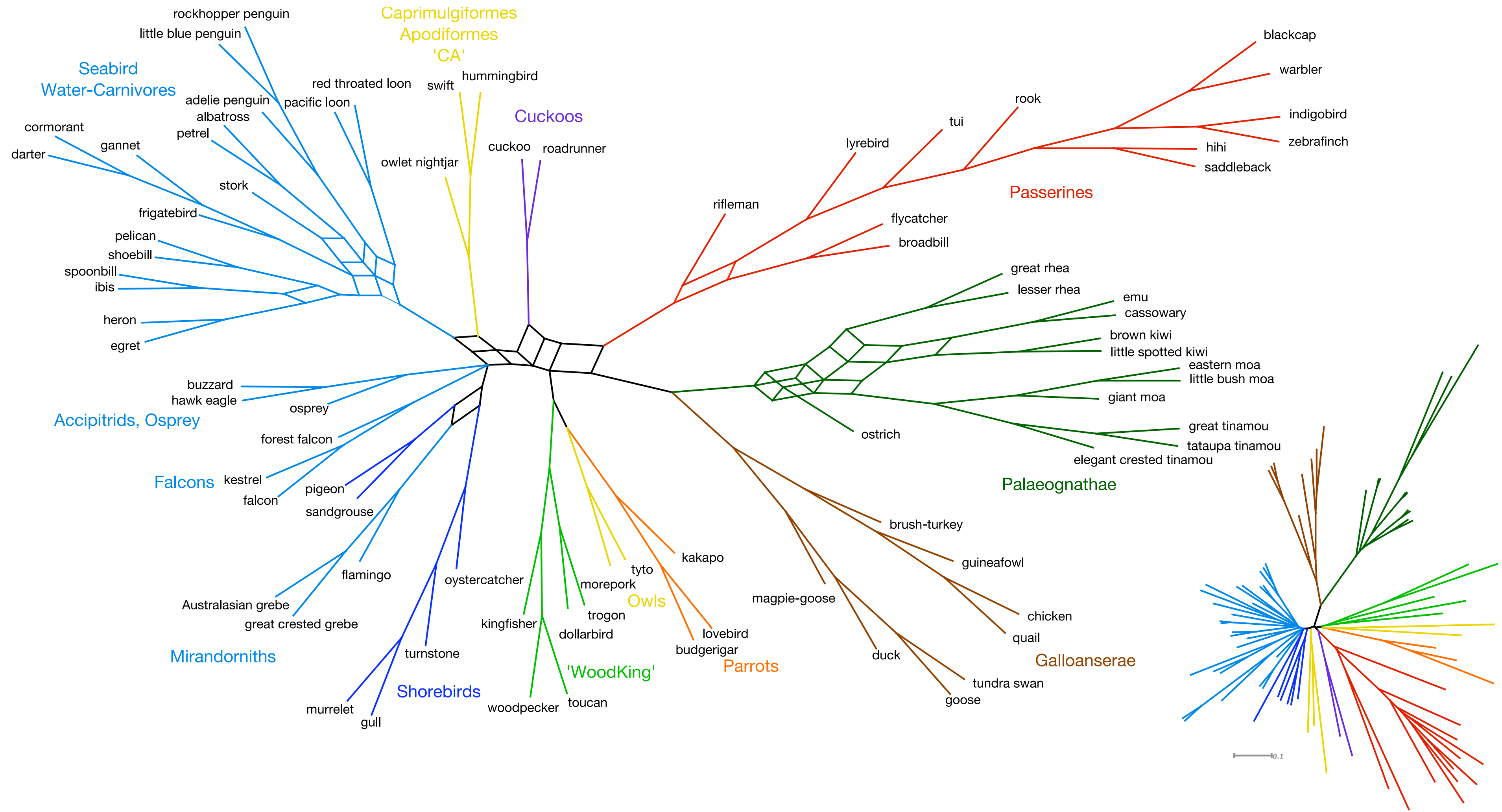

FIGURE 7.1—Network showing splits over 20% of RAxML bootstrap analysis. Branch lengths are a relative count of split support. The insert shows the best RAxML tree, indicating relative branch lengths. Colours indicate the Palaeognathae, Galloanserae and the six-way split of Neoaves (Cracraft 2001). Light and dark blue is used to indicate two of the three subgroups within the Conglomerati.

Interestingly, the six groups of Cracraft (2001) are surprisingly well defined, save for the owls, which we showed in **Appendix One** (Pratt et al. 2009) do not group with the Caprimulgiformes and Apodiformes. The large 'Conglomerati' will likely be the most difficult to resolve, as the deep internal branches are very short, suggesting an adaptive radiation. We have made good progress in resolving the seabird water-carnivore clade within the 'Conglomerati' (**Chapter Five**), and highlight where more work is needed. Future challenges include successfully resolving members of the missing third Conglomerati subgroup, to which the takahe and kagu may belong (as a beginning, sequencing of the weka and sandhill crane to respectively pair with each of these is underway, S. Pilkington, P. McLenachan pers. comm.).

The land bird clade of Hackett et al (2008) is not strictly supported by this mtDNA analysis (Fig. 7.1), nor in our more detailed analysis of fewer species (Pratt et al. 2009). The main difference in this respect is the placement of the accipitrids and falcons, which fall within the Conglomerati in our analysis (as predicted by Cracraft 2001), but nearer the passerines in Hackett et al (2008). In our analyses, the root of Neoaves (where the Palaeognathae and Galloanserae join the tree) is amongst the faster evolving lineages where internal branches are longer (see Fig. 7.1 insert), and is it possible this is an artifact. The root of Neoaves in Hackett et al. (2008) is more approximately near the base of where the 'CA' are in our tree. We certainly hope that more exon sequences, and techniques such as down-weighting will help resolve the root of the tree.

Here, as in the more detailed analyses in **Appendix Two** (Phillips et al. 2010), the tinamous clearly fall within the ratites, and are sister to the moa (100% bootstrap support, not shown). Within the Palaeognathae, signals for the alternative groupings outlined in the discussion of Phillips et al. (2010) are seen as boxes, including the ostrich grouping with Casuariidae and kiwi (an artifact of low cytosine relative to thymine), kiwi sister to moa and tinamous, and rhea basal within Palaeognathae. The isolation of New Zealand features prominently in many biogeographic hypotheses for modern birds (reviewed in **Chapter Two**). Our discovery that moa and tinamous are sister groups nested within Palaeognathes, and that therefore flight has been lost up to six times within the group (Phillips et al. 2010) will necessitate a reevaluation of the ratite vicariance hypotheses.

Some relationships still require more work for resolution. In this analysis, encouragingly the pigeon still groups with the sandgrouse, as it does in nuclear sequence analyses (Hackett et al. 2008). As discussed in **Chapter Four**, a second pigeon and sandgrouse will hopefully help improve resolution here, as support is weak and the divergences are very deep. In the longer term, more nuclear exon sequences and rare genomic changes may be required for resolution of the 'Conglomerati', including pigeons.

The study of Hackett et al. (2008) finds evidence for the separation of Neoaves into well supported land birds, water birds and then a selection of terrestrial and arboreal miscellaneous clades that are less well supported. Ecologically, this is interesting but still requires further testing with datasets less reliant on intron data. We have explored the seabird water-carnivore clade (**Chapter Five**), and find good agreement between nuclear and mitochondrial datasets. Our calculations suggest this seabird water-carnivore clade began diversifying 70-80 Ma. This is consistent with independent lines of evidence showing webbed bird footprints and a decline of pterosaurs around the same time (Kim et al. 2003; Slack et al. 2006). Of Penny and Phillips' five models for the divergence of birds (2004, and reprinted in **Chapter One**), these results suggest support for the right hand of the continuum (models D to E) for this group, supporting the competitive displacement of pterosaurs by birds, independent of the major impact that marks the end of the Cretaceous.

In our preliminary analyses of oscine passerines, we have already shown many NZ 'basal corvid' groups are actually 'basal Passerida'. With the mitochondrial genomes published here (tui, saddleback, hihi, piopio, **Chapter Six**), and the three in production (grey warbler, fantail and NZ robin) we cover the spectrum of NZ basal oscine groups. Now our attention turns to the Australian deep oscines, which will complement the New Zealand ones. They will help resolve the basal branches of oscines generally, and help test the 'out of Australasia' theory of oscine biogeography. In particular, catbirds and bowerbirds (Ptilonorhynchidae, now that we have shown the piopio does not belong here), treecreepers (Climacteridae), logrunners (Orthonychidae), and babblers (Pomatostomidae) are basal Australian groups that should be targeted, along with Australian Meliphagoidea to pair with the New Zealand tui and grey warbler.

There are many interesting questions to ask as oscine phylogeny becomes resolved. Within oscine passerines, the basal groups are all Australasian. Many of these are frugivorous and nectarivorous, while the large successful radiations of granivorous passerines are all within the Passerida (Ericson et al. 2003). When did the Passerida radiation occur? Did 'escape' from Australasia occur simultaneously, before or after the evolution of seed eating? What other modern birds and relatives existed outside Australasia at or before that time (for example the seed-eating Jeholornis of the early Cretaceous, Zhou and Zhang 2002), were they in competition with basal oscines, or were they out-competed by the Passerida?

7.1 FUTURE DIRECTIONS: TO PHYLOGENY AND BEYOND

Phylogenetic resolution has already helped progress in answering some of the bigger questions. For example, the assertion of Feduccia (1995) that modern birds are descended from 'transitional shorebirds' around the period of the K/T boundary has now been refuted by multiple lines of evidence (for example Paton et al. 2002; Hackett et al. 2008; Pratt et al. 2009). The question of how many avian lineages crossed the K/T boundary also seems to be resolving in favour of more, rather than less (Brown et al. 2008; Pratt et al. 2009), although opposing arguments still continue (e.g. Ericson et al. 2006). In addition, the more studies that are undertaken, the more support we find for the sufficiency of microevolutionary changes to explain the macroevolution observed through history (Penny and Phillips 2004).

It turns out phylogenetic resolution is a time consuming problem, and much of this thesis has focused on the problems of phylogeny. This thesis has generated many more questions and directions for future work. Ranging through specific and general, these include:

1. A formal comparison of avian phylogenies from mitochondrial and nuclear datasets, following on from the brief beginning made in this chapter.

2. More data (always more data!), specifically more mitochondrial sequences for those taxa mentioned in this thesis, and more nuclear exon data. Next generation sequencing offers potential avenues for finding good exon sequences. The recently published proposal to sequence the genomes of

10, 000 vertebrate species (Genome 10K Community of Scientists 2009) is an indication of the huge increase in data that will occur in the next few years.

3. Has the origin of flight been ground up or tree down? We are seeing evidence that see almost all birds in the basal branching groups of tree are ground-foraging (palaeognaths, galloanseriforms), and it is the more derived groups that are both better at flight and also forage on the wing or in trees (Fig. 7.1). This would suggest birds most likely evolved flight from the ground up. It will be interesting to see where rails and their relatives fit in Neoaves (also ground foraging, and early indications suggest they may be quite deep).

4. Tied to the last question, how easy is it to lose flight, and what predictors are there? McCall et al. (1998) suggested loss of flight has occurred in at least 11 extant avian families, suggesting relative wing length is a predictor of flight loss. Steadman (2006) reported more than 100 losses of flight on the Pacific Islands among the Rallidae alone. We have also suggested six losses of flight in the Palaeognathes (Phillips et al 2010). Maynard-Smith (1968, pp. 13-14) showed the cost of powered flight increases with the linear dimension (l) in proportion to $l^{3.5}$, whereas the power from the muscle (or the area of the wing), only increases in proportion to l^2. Shen et al (2009) found relaxation on selective constraints in mitochondrial DNA after loss of flight. Taking all of this into consideration, the gain of flight in the first place seems a very surprising occurrence!

Birds have proven to be an endless source of interesting questions and observations, and I hope that this thesis contributes to our understanding of them.

7.2 REFERENCES

Brown, J. W., J. S. Rest, J. Garcia-Moreno, M. D. Sorenson, and D. P. Mindell. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol 6:6.

Cracraft, J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proc R Soc Lond B Biol Sci 268:459-469.

Ericson, P., C. Anderson, T. Britton, A. Elzanowski, and U. Johansson. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett 2:543-547.

Ericson, P. G. P., M. Irestedt, and U. S. Johansson. 2003. Evolution, biogeography, and patterns of diversification in passerine birds. J Avian Biol 34:3-15.

Feduccia, A. 1995. Explosive Evolution in Tertiary Birds and Mammals. Science 267:637-638.

Genome 10K Community of Scientists. 2009. Genome 10K: A Proposal to Obtain Whole-Genome Sequence for 10 000 Vertebrate Species. J Hered 100:659-674.

Gibb, G. C., O. Kardailsky, R. T. Kimball, E. L. Braun, and D. Penny. 2007. Mitochondrial genomes and avian phylogeny: Complex characters and resolvability without explosive radiations. Mol Biol Evol 24:269-280.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320:1763-1768.

Kim, C. B., M. Huh, C. S. Cheong, M. G. Lockley, and H. W. Chang. 2003. Age of the pterosaur and web-footed bird tracks associated with dinosaur footprints from South Korea. The Island Arc 12:125-131.

Maynard-Smith, J. 1968. Mathematical ideas in biology. Cambridge University Press, Cambridge.

McCall, R. A., S. Nee, and P. H. Harvey. 1998. The role of wing length in the evolution of avian flightlessness. Evolutionary Ecology 12:569-580.

Morgan-Richards, M., S. Trewick, A. Bartosch-Härlid, O. Kardailsky, M. J. Phillips, P. A. McLenachan, and D. Penny. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol Biol 8:20.

Paton, T., O. Haddrath, and A. J. Baker. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proc R Soc Lond B Biol Sci 269:839-846.

Penny, D., and M. J. Phillips. 2004. The rise of birds and mammals: are microevolutionary processes sufficient for macroevolution? Trends Ecol Evol 19:516-522.

Phillips, M. J., G. C. Gibb, E. A. Crimp, and D. Penny. 2010. Tinamous and moa flock together: mitochondrial genome sequence analysis reveals independent losses of flight among ratites. Syst Biol 59:90-107.

Pratt, R. C., G. C. Gibb, M. Morgan-Richards, M. J. Phillips, M. D. Hendy, and D. Penny. 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol Biol Evol 26:313-326.

Shen, Y.-Y., P. Shi, Y.-B. Sun, and Y.-P. Zhang. 2009. Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. Genome Research 19:1760-1765.

Singh, T. R., O. Shneor, and D. Huchon. 2008. Bird mitochondrial gene order: Insight from 3 warbler mitochondrial genomes. Mol Biol Evol 25:475-477.

Slack, K. E., C. M. Jones, T. Ando, G. L. A. Harrison, R. E. Fordyce, U. Arnason, and D. Penny. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol 23:1144-1155.

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Syst Biol 57:758-771.

Steadman, D. W. 2006. Extinction and Biogeography of Tropical Pacific Birds. The University of Chicago Press, Chicago.

Zhou, Z., and F. Zhang. 2002. A long-tailed, seed-eating bird from the Early Cretaceous of China. Nature 418:405-409.

Toward Resolving Deep Neoaves Phylogeny: Data, Signal Enhancement, and Priors

Renae C. Pratt, Gillian C. Gibb,* Mary Morgan-Richards,* Matthew J. Phillips,† Michael D. Hendy,* and David Penny**

*Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, Palmerston North, New Zealand; and †Centre for Macroevolution and Macroecology, School of Botany and Zoology, Australian National University, Canberra ACT, Australia

We report three developments toward resolving the challenge of the apparent basal polytomy of neoavian birds. First, we describe improved conditional down-weighting techniques to reduce noise relative to signal for deeper divergences and find increased agreement between data sets. Second, we present formulae for calculating the probabilities of finding predefined groupings in the optimal tree. Finally, we report a significant increase in data: nine new mitochondrial (mt) genomes (the dollarbird, New Zealand kingfisher, great potoo, Australian owlet-nightjar, white-tailed trogon, barn owl, a roadrunner [a ground cuckoo], New Zealand long-tailed cuckoo, and the peach-faced lovebird) and together they provide data for each of the six main groups of Neoaves proposed by Cracraft J (2001). We use his six main groups of modern birds as priors for evaluation of results. These include passerines, cuckoos, parrots, and three other groups termed "WoodKing" (woodpeckers/rollers/kingfishers), "SCA" (owls/potoos/owlet-nightjars/hummingbirds/swifts), and "Conglomerati." In general, the support is highly significant with just two exceptions, the owls move from the "SCA" group to the raptors, particularly accipitrids (buzzards/eagles) and the osprey, and the shorebirds may be an independent group from the rest of the "Conglomerati". Molecular dating mt genomes support a major diversification of at least 12 neoavian lineages in the Late Cretaceous. Our results form a basis for further testing with both nuclear-coding sequences and rare genomic changes.

Introduction

Perhaps, the greatest current challenge of avian systematics for molecular evolutionists and systematists alike is the resolution of the polytomy at the base of the Neoaves. The basic paleognath (tinamous and ratites)–neognath division (all other modern birds) is supported by studies of morphology (Cracraft and Clarke 2001), nuclear-coding DNA (Groth and Barrowclough 1999; García-Moreno and Mindell 2000), and mitochondrial (mt) genomes (Sorenson et al. 2003; Harrison et al. 2004; Slack et al. 2007). Within the Neognathae, the Galloanserae (chickens, ducks, and their relatives) represent the earliest divergence, leaving the large majority (all remaining orders) of birds in the Neoaves. Again, coding regions of both mt genomes and nuclear DNA, together with morphological data, agree with the Galloanserae division. However, resolving the relationships within Neoaves is still elusive.

Thus, resolution of the basal Neoavian polytomy could be seen as the "last frontier" for resolving deep-level systematics among modern birds. There are a range of views best illustrated by the two ends of a spectrum—first, the theory that the basal polytomy is due to an "explosive radiation" after the Cretaceous–Paleogene (K–Pg, formerly K–T) boundary. That is, birds and mammals "inherited the earth" only after the demise of the dinosaurs and pterosaurs (Feduccia 2003; Chubb 2004; Poe and Chubb 2004; Ericson et al. 2006). The other end of the spectrum are hypotheses that basal avian lineages were diversifying in an "adaptive radiation" long before the asteroid impact that marks the K–Pg boundary (Cooper and Penny 1997; Cracraft 2001; van Tuinen and Hedges 2001; Penny and Phillips 2004; Pereira and Baker 2006; van Tuinen et al. 2006; Brown et al. 2007, 2008). This latter approach represents mainstream evolutionary theory in that it attempts to explain the past by reference to known mechanisms—to "causes now in operation" (Penny and Phillips 2004).

Poe and Chubb (2004) suggested that the large polytomy at the base of Neoaves represents a rapid radiation that "might be considered essentially simultaneous." If a lack of resolution is not caused by truly short times between divergences, then ultimately relationships should be resolvable (Whitfield and Lockhart 2007). We have already commented (Gibb et al. 2007) that an explosive radiation implies both short divergence times between avian orders and also that the ecological and morphological differences that identify the crown groups of orders within Neoaves must have occurred over the same short timescale. It would scarcely be an explosive radiation if the lineages diverged quickly, but it then took tens of millions of years for genetic changes to occur leading to the ecological and morphological characters that distinguish crown group Neoavian orders today. Apart from being real, short branch lengths in phylogenies can result for a number of reasons. For example, different characters or data sources that provide support for conflicting trees (rather than from the absence of support) can result in short branches; under these conditions, even standard maximum likelihood (ML) can seriously underestimate branch lengths (Penny et al. 2008). Additionally, use of inappropriate genes or analysis methods can return short branches. To improve divergence time estimates, Brown et al. (2008) recommend longer sequences (they used 4,594 bp of mtDNA) and the use of multiple independent nuclear loci. In addition to incorporating longer sequences, improved analytic methods such as networks allows visualization of conflict between essentially equally good phylogenies (Holland et al. 2004).

A greater understanding of the evolutionary history of Neoaves is still needed. We have found Cracraft's (2001) six groupings within Neoaves to be useful as "informal priors" for recent studies on passerines and for the group termed "Conglomerati" (e.g., Gibb et al. 2007). With

Key words: Neoaves, mitochondrial genomes, site-stripping, down-weighting, hypothesis testing, avian evolution.

E-mail: r.c.pratt@massey.ac.nz.

Mol. Biol. Evol. 26(2):313–326. 2009
doi:10.1093/molbev/msn248
Advance Access publication November 3, 2008

© The Author 2008. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org

additional data presented here, we have representatives (and test) his six prior groupings within Neoaves:

(i) Passeriformes (passerines, perching birds);
(ii) Psittaciformes (parrots);
(iii) Cuculiformes (cuckoos);
(iv) Coliiformes, Coraciiformes, Trogoniformes, and Piciformes (mousebirds [not included in the present study], rollers/bee-eaters/kingfishers, trogons, woodpeckers/toucan, and jacamars). Informally, and for ease of communication, we call this the "WoodKing" grouping;
(v) Strigiformes, Caprimulgiformes, Apodiformes, and Musophagiformes (owls, nightjars, potoos, owlet-nightjars, hummingbirds, swifts, and turacos). We have sampled three of the four groups, so we will refer to these as "SCA"(or "CA" when omitting the owls [Strigiformes]); and
(vi) a diverse group dubbed the "Conglomerati" (Slack et al. 2007) that includes three main subgroups: 1) Falconiformes, Pelecaniformes, Ciconiiformes, Procellariiformes, Gaviiformes Sphenisciformes, and Podicipediformes (raptors, pelicans and relatives, storks, seabirds, loons, penguins, and grebes); 2) Charadriiformes, Columbiformes, and Pteroclidae (shorebirds, pigeons, and sandgrouse); and 3) Gruidae, Rallidae, Otididae, Eurypygidae, and Turnices (cranes, rails, bustards, kagu, sun bitterns, and button quails).

In recent papers (Slack et al. 2006, Slack et al. 2007), we concentrated on the relationships within (i) and (vi) but have recently extended this to include members of groups (iv) and (v) (Gibb et al. 2007; Morgan-Richards et al. 2008). Because we previously only had a single representative for the orders Psittaciformes (parrots) and Strigiformes (owls) (Harrison et al. 2004), we omitted them from recent analyses because they are long branches that are known to be problematic in phylogeny generally (Hendy and Penny 1989), including birds (Harrison et al. 2004).

At this stage, we are particularly concerned as to whether the members within each of the above six groups come together—this will help evaluate whether the deep Neoavian lineages are resolvable. As such, we are not especially concerned if some of Cracraft's six groups have paraphyletic lineages within them—that is, a taxonomic question, not a question about the resolvability of the deepest lineages. For example, we are interested in whether the combined Coliiformes/Coraciiformes/Piciformes ("WoodKing") form a natural group within Neoaves—even if one of the subgroups turns out to be paraphyletic within this "WoodKing" grouping.

After this work was submitted, Hackett et al. (2008) published what is probably the most comprehensive report on bird evolution since Cracraft (2001). Although the scope of data is impressive, it is largely based on noncoding intron sequences. Nuclear intron data, in combination with current mt and fossil data, have the potential to be extremely useful as long as we can be confident in alignments that span the phylogenetic depth of the avian clade. Some authors have suggested that intron sequences are not appropriate for deeper divergences (Shapiro and Dumbacher 2001) due to alignment ambiguities resulting from multiple insertions and deletions (for general comments on alignments see Löytynoja and Goldman 2008). For example, Morgan-Richards et al. (2008) showed that the alignment of β-fibrinogen intron 7 (which supports the controversial metaves–coronaves split) has no constant sites across the wide taxon sampling required for determining deep avian divergences. Introns are potentially well suited to resolving rapid radiations as they evolve fast enough to accumulate changes during this time (of divergence), while being slow enough to not become random and therefore lose signal (Matthee et al. 2007).

In general, the main avian orders found in both Hackett et al. (2008) and Cracraft (2001) are the same (e.g., Passeriformes, Psittaciformes, and Cuculiformes etc.). However, the relationships among the orders are different, with the deep branches of Neoaves receiving low support (<80% bootstrap support) in Hackett et al. (2008). One difference is their "land birds." Within land birds, Hackett et al. (2008) found Passeriformes (passerines) sister to Psittaciformes (parrots) and suggested a sister relationship between these and Falconidae. These groupings only have support when the intron data are included. Ericson et al. (2006) also inferred this relationship but only when all genes were combined, including β-fibrinogen intron 7 (see their Supplementary Material figs. ESM-1–8). If correct, this placement would be very interesting; however, intron alignment and/or long-branch attraction may be a factor here. The long internal branch to passerines may be attracting the long internal branch of parrots (see Hackett et al. 2008, fig. 3). At this point, support for such a grouping has not been found with mtDNA (e.g., Gibb et al. 2007; Brown et al. 2008). Cracraft (2001) by comparison included morphological, geographical, and early molecular data in support of his groupings. We therefore feel it is appropriate to use Cracraft (2001) as the basis for testing deep divergences within Neoavian birds rather than any one molecular data set.

As a step toward increasing the taxon sampling of coding sequences, we add nine new mt genomes: the dollarbird (*Eurystomus orientalis*) and New Zealand kingfisher (*Halcyon sancta vagans*) as representatives of the Coraciiformes, together with the white-tailed trogon (*Trogon viridis*) from Trogoniformes, are suggested to group with Piciformes; the great potoo (*Nyctibius grandis*) and Australian owlet-nightjar (*Aegotheles cristatus cristatus*) as representatives of Caprimulgiformes; barn owl (*Tyto alba*), expected to pair with the New Zealand owl (morepork, *Ninox novaeseelandiae*) to form Strigiformes; the roadrunner (a ground cuckoo, *Geococcyx california*) expected to pair with the New Zealand long-tailed cuckoo (*Eudynamys taitensis*) to form Cuculiformes; and the peach-faced lovebird (*Agapornis roseicollis*) from Psittaciformes expected to join with the budgerigar (*Melopsittacus undulatus*) and the ground parrot (kakapo, *Strigops habroptilus*). Thus, we have reduced the number of long branches in our data set by the addition of representatives from each of the six Neoavian lineages described in Cracraft (2001). For the third subgroup of the "Conglomerati", only a rail (takahe) and the kagu are published. It is unclear whether these really are a natural group (Morgan-Richards et al.

2008), and this again leaves us with two long isolated branches. In accord with our previous practice, these species have temporarily been omitted until sequences from more closely related species are available for each of them.

We now change from the birds to the analysis. Perhaps, the most fundamental problem occurring while reconstructing deep-level phylogeny is substitution saturation (Curole and Kocher 1999; Phillips et al. 2006). Phylogenetic signal can be eroded by factors including superimposed substitutions and "nonhistorical" biases (such as from compositional nonstationarity)—which accumulate more rapidly at faster evolving sites. Attempts to limit these problems have been made in recent studies by identifying fast-evolving sites at which signal erosion is expected to be high (Morgan-Richards et al. 2008).

In previous work (Delsuc et al. 2003; Phillips and Penny 2003; Phillips et al. 2004), we found standard RY coding (Honeycutt and Adkins 1993), especially the third-codon positions, to be advantageous for the most variable partitions of nucleotide data. This recoding both increases the proportion of changes on internal branches of the tree (i.e., a "treeness" measure) and decreases the differences in nucleotide composition (relative compositional variability). This latter is important in reducing nucleotide composition effects because they have been long known to bias tree reconstruction (Lockhart et al. 1992). Because of the better fit of the data to the model (higher treeness and less variability in nucleotide composition), this has been our preferred method of analysis for vertebrate mt data.

Down-weighting the faster evolving sites or grouping faster evolving nucleotides (or amino acids) into a single category has been quite widely used (see Honeycutt and Adkins 1993; Philippe et al. 2000; Jeffroy et al. 2006)—although the theoretical aspects have not been well developed in phylogenetics. Rodriguez-Ezpeleta et al. (2007) report that omitting the fastest evolving sites, grouping amino acids into functional categories, and some mixture models, all enhanced the phylogenetic signal for deeper divergences. Susko and Roger (2007) similarly report improvements from down-weighting. However, some approaches may not be optimal if valuable sites are excluded simply because they are grouped under some prior definition (e.g., codon positions that have many fast-evolving sites). Conversely, some saturated sites may be retained because they are in a category that, on average, does not have site saturation. Thus, we can also group the justification for down-weighting into those using "a priori" categories (such as third position, stems vs. loops in RNA or amino acid groups) and "conditional" categories (down-weighting of each site independently).

We have used both in the past, the RY coding (a priori weighting category) (Phillips et al. 2004), and also a conditional down-weighting (Penny and Hendy 1986), based on the numbers of observed and expected incompatibilities. In general, all the methods mentioned here are examples of a standard statistical approach of "noise reduction/signal enhancement" (Proakis et al. 2002). Here, we implement a conditional noise reduction technique in which the information retained from the sequence is determined on a site-by-site basis. The Materials and Methods section has more detail on this conditional recoding (down-weighting) of sites, an approach that we call site-stripping (Morgan-Richards et al. 2008).

Along with data partitioning/down-weighting and fossil calibrations (see Supplementary Material S1 online), our additional sequences mean that we can, in principle, calculate the probabilities that prior hypotheses are supported. In other words, we can calculate the proportion of trees that will have a split (or clade) that has been predicted. Unfortunately, there appears to be little use in phylogenetics for specifying a priori hypotheses and then testing the probability of finding them with new data. Rather, results are treated somewhat "post hoc," looking at the trees after they are built and then trying to explain the results. In principle, a Bayesian approach allows alternative hypotheses to be given different weightings, but it appears that a "flat prior" is the norm; and this does not really differentiate between trees or hypotheses. Based on nuclear-coding sequences, Lin et al. (2002) took the four-way split within eutherian mammals and calculated the probability of finding the same split from mt data. The result was certainly very highly significant; $P \approx 2.1 \times 10^{-7}$. It is very important in phylogeny, though perhaps seldom carried out, to give quantitative estimates of the increase in information from a phylogeny. Penny et al. (1991) demonstrated that it is simple to calculate probabilities (see Appendix and table 1) of a single pair of taxa predicted to come together on the tree. That is, there is one chance in $2n - 5$ of them coming together "by chance" on an unrooted tree and one in $2n - 3$ for a rooted tree (given n taxa). It is even more improbable that a predefined grouping of three or more taxa will come together, and here, we develop measures in order to evaluate quantitatively the priors from Cracraft (2001).

Our approach here is to infer evolutionary trees from complete mt genomes as a start on resolving the deep Neoavian splits. We investigate the robustness of the priors set out by Cracraft (2001) by using a novel site-stripping method and exploring networks (see Supplementary Material S1 online for fossil calibrations). In addition, we assess quantitatively the usefulness of these priors by calculating clade probabilities. By finding resolution in the basal node of the Neoaves, hypotheses regarding the number of lineages present before the K–Pg boundary can be tested.

Materials and Methods

Taxon Sampling

The dollarbird (*E. orientalis*) and barn owl (*T. alba*) were supplied from the Australian Museum, Sydney, Australia, under sample numbers EBU 11118 and EBU 2564, respectively. The roadrunner (*G. california*), great potoo (*N. grandis*), and the white-tailed trogon (*T. viridis*) were provided by the Louisiana State University Museum of Natural Science Collection of Genetic Resources under sample numbers LSUMZ B-8504, LSUMZ B-8954, and LSUMZ B-28495. The Australian owlet-nightjar (*A. c. cristatus*) was provided by Fritz Geiser, University of New England, Armidale, Australia. The New Zealand long-tailed cuckoo (*E. taitensis*) and the New Zealand kingfisher (*H. s. vagans*) were obtained from Dick Gill, NZ Department of Conservation, Waikanae, New Zealand.

Table 1
Probabilities of a Predefined Clade Joining the Tree as a Single Group or as Two Subclades

k Taxa in clade	Unrooted Trees Minimal	Rooted Trees Minimal
2	$1/(2n-5)$	$1/(2n-3)$
3	$3/(2n-5)(2n-7)$	$3/(2n-3)(2n-5)$
4	$15/(2n-5)(2n-7)(2n-9)$	$15/(2n-3)(2n-5)(2n-7)$
5	$105/(2n-5)(2n-7)(2n-9)(2n-11)$	$105/(2n-3)(2n-5)(2n-7)(2n-9)$

The peach-faced lovebird (*A. roseicollis*) was obtained locally from commercial breeders.

Molecular Methods

Extractions of genomic DNA from each of the newly sampled birds were performed at the Allan Wilson Centre from 25 to 50 mg of liver tissue using the High Pure PCR Template Preparation Kit (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instructions. To minimize the chance of obtaining nuclear copies of mt genes (NUMTs), 2–4 overlapping long-range polymerase chain reaction (PCR) fragments (3.5–12 kb in length) were first amplified using the Expand Long template PCR System (Roche Applied Science). The products were excised from 1% agarose gels and purified using a QIAquick Gel extraction kit (Qiagen GmbH, Hilden, Germany) as per the manufacturer's instructions. These long-range products were subsequently used as template DNA for following short-range PCRs (overlapping fragments 0.5–3 kb in length). Short-range primer combinations were found using our laboratory database as described in Slack et al. (2006), and any new primers required were designed using Oligo 4.03 (National Biosciences, Inc., Plymouth, MN). Sequencing was performed using BigDye Terminator Cycle Sequencing reagents according to the manufacturer's instructions (Applied Biosystems, Foster City, CA) and then sequenced on an ABI 3730 automated sequencer (Applied Biosystems). Sequences were aligned using Sequencher 4.7 (Gene Codes Corp., Ann Arbor, MI) and then manually edited and checked for complete concurrence between overlapping sequences.

Where necessary (e.g., with length heteroplasmy in control regions [CRs] from microsatellite repeats), PCR products were cloned using the TOPO TA cloning kit for sequencing (Invitrogen, Carlsbad, CA). For each region, at least three clones were sequenced to safeguard against PCR errors. In all cases, overlaps between sequences were sufficient to ensure synonymy and sequence identity was confirmed through Blast searches (http://www.ncbi.nlm.nih.gov/blast/), confirmation of amino acid translation in coding regions, and alignment with other species.

In addition to the nine new bird mt genomes reported in this paper, 36 other complete avian mt genomes from NCBI GenBank were included in the analyses: 31 Neoaves and 5 Galloanserae. Paleognath taxa were not included in this data set; although their overall placement is now well established (Gibb et al. 2007; Slack et al. 2007), there are still important but unresolved issues around the placement of tinamous (Hackett et al. 2008; Harshmann et al. 2008; Phillips MJ, Gibb GC, Crimp EA, Penny D, in preparation). Instead, we rooted our Neoaves trees with the Galloanserae sequences (Gibb et al. 2007; Morgan-Richards et al. 2008). The full data set is available from the authors on request.

The Galloanserae taxa are Japanese quail (*Coturnix japonica*, AP003195), magpie goose (*Anseranas semipalmata*, AY309455), redhead duck (*Aythya americana*, AF090337), greater white-fronted goose (*Anser albifrons*, AF363031), and Australian brush turkey (*Alectura lathami*, AY346091). The 31 Neoaves taxa (modern birds) are rifleman (New Zealand wren, *Acanthisitta chloris*, AY325307), gray-headed broadbill (*Smithornis sharpei*, AF090340), fuscous flycatcher (*Cnemotriccus fuscatus*, AY596278), superb lyre bird (*Menura novaehollandiae*, AY542313), rook (*Corvus frugilegus*, Y18522), ivory billed toucan (*Pteroglossus azara*, DQ780882), pileated woodpecker (*Dryocopus pileatus*, DQ780879), morepork (a New Zealand owl, *N. novaeseelandiae*, AY309457), kakapo (flightless parrot *S. habroptilus*, AY309456), budgerigar (*M. undulatus*, EF450826), ruby-throated hummingbird (*Archilochus colubris*, EF532935), common swift (*Apus apus*, AM237310), peregrine falcon (*Falco peregrinus*, AF090338), forest falcon (*Micrastur gilvicollis*, DQ780881), Eurasian buzzard (*Buteo buteo*, AF380305), osprey (*Pandion haliaetus*, DQ780884), Blyth's hawk eagle (*Spizaetus alboniger*, AP008239), blackish oystercatcher (*Haematopus ater*, AY074886), ruddy turnstone (*Arenaria interpres*, AY074885), southern black-backed gull (*Larus dominicanus*, AY293619), red-throated loon (*Gavia stellata*, AY293618), little blue penguin (*Eudyptula minor*, AF362763), rockhopper penguin (*Eudyptes chrysocome*, AP009189), black-browed albatross (*Diomedea melanophris*, AY158677), Kerguelen petrel (*Pterodroma brevirostris*, AY158678), frigatebird (*Fregate* sp., AP009192), Australian pelican (*Pelecanus conspicillatus*, DQ780883), Australasian little grebe (*Tachybaptus novaehollandiae*, EF532936), greater flamingo (*Phoenicopterus ruber roseus*, EF532932), great crested grebe (*Podiceps cristatus*, AP009194), and the Oriental white stork (*Ciconia boyciana*, AB026193).

Phylogenetic Analysis

Sequences were aligned in Se-Al v2.0a11 at the amino acid level for protein-coding genes and based on secondary structure for RNA genes. The data set has 12 protein-coding genes, 2 ribosomal RNAs (rRNA), and 22 transfer RNAs (tRNA). Gaps, ambiguous sites adjacent to gaps, NADH6 (light-strand encoded), and stop codons (often incomplete in the DNA sequence) were excluded from the alignment. The 12 protein-coding genes were separated into first-, second-, and third-codon positions, whereas rRNA and tRNA genes were partitioned into stems (S) and loops (L).

Protein-coding genes were checked for NUMTs by translating into amino acids.

Previous studies from birds (Slack et al. 2003), mammals (Lin et al. 2002), and simulations (Holland et al. 2003) have all shown that the addition of outgroups can disrupt the ingroup tree. However, in such cases (from theory and with simulated data), the ingroup tree (i.e., with the outgroup omitted) is more likely to be correct (Holland et al. 2003). We therefore ran separate analyses either including or excluding the outgroup (five birds from the Galloanserae). A combined total of 13,412 nucleotides (excluding gaps) were used for the basis of further analyses (see below for number of characters per data set). As mentioned earlier, we partitioned the data: codons 1 and 2, codon 3, RNA stems, and RNA loops (Slack et al. 2007) for site-stripping. ML analyses were carried out using standard programs including PAUP* 4.0b10 (Swofford 2001) and GARLI v0.95 (Zwickl 2006). Bayesian analysis was carried out in MrBayes (Huelsenbeck and Ronquist 2001), and consensus networks were implemented in SplitsTree version 4 (Holland et al. 2004; Huson and Bryant 2006) and BEAST (Drummond et al. 2006; Drummond and Rambaut 2007). Optimal parameters for the ML models were determined using Modeltest 3.7 (Posada and Crandall 1998) and the AIC values used. The hierarchical and AIC tests were in agreement for Modeltest. Initial results from ML analyses were consistent with Bayesian analysis and were not used for site-stripping. The best model for the full data set was GTR + I + G. Bayesian analyses used default parameters and were run for 10 million generations or until convergence was obtained. In addition to Bayesian posterior probabilities (BPPs), we ran analyses in PhyML (Guindon and Gascuel 2003) and RAxML (Stamatakis 2006) to carry out 100 bootstrap replicates on the data sets both with and without the outgroup (see Supplementary Material S2 online for results).

Noise Reduction by Down-weighting (Site-Stripping)

Site-stripping compares sites based on the actual number of mutations required on the tree ("tree steps") versus the maximum possible number of mutations for that site (max). The calculation is the limit (L) = tree steps × tree steps/max steps. If the threshold "strictness" (s) is, for example, 4, then sites for which $L \leq 4$ remain unchanged and sites for which $L > 4$ are RY coded. If after RY coding L is still > 4, the site is excluded. The higher the threshold (the larger the value of s), the more sites that are included and fewer are RY coded. Conversely, the lower the threshold (lower values of s), the more sites RY coded or excluded. Therefore, the weighting of each site is a function of that site and is not predetermined by being a member of a class. This allows, for example, some hypervariable sites from first and second positions to be RY coded or omitted. A range of s values were used and resulted in the following:

$s = \infty$, 0 sites RY coded, 0 sites excluded (all 13,412 sites included);
$s = 6.0$, 20 sites RY coded, 0 sites excluded;
$s = 4.5$, 91 sites RY coded, 1 site excluded;
$s = 3.5$, 254 sites RY coded, 8 sites excluded;
$s = 2.28$, 578 sites RY coded, 63 sites excluded; and
$s = 2.0$, 891 sites RY coded, 159 sites excluded.

Bayesian inference analyses were carried out on each of the data matrices, including the fully weighted data. Note that RY coding increases the ML scores as it amalgamates some nucleotide categories; thus, the data are now different, and it is not valid to compare directly ML scores from RY and nucleotide coding (Steel MA, personal communication). Similarly, consistency index (CI) values are not directly comparable between nucleotide and RY coded data sets. That is, for an unrooted four-taxon tree, for which A, C, T, and G states are random (i.e., no signal remains), the expected average CI for nucleotide data is 0.949, whereas for the RY coded data, it is 0.778.

In order to identify fast-evolving sites and so facilitate noise reduction by site-stripping, nine additional close relatives were added to the alignment but were removed before phylogenetic analysis as they do not break up long branches or add phylogenetic signal and would only increase analysis times. The nine taxa included were chicken (*Gallus gallus*, AP003317); green junglefowl (*Gallus varius*, AP003324); gray junglefowl (*Gallus sonneratii*, AP006741); white stork (*Ciconia ciconia*, AB026818); Canadian goose (*Branta canadensis*, DQ019124); tundra swan (*Cygnus columbianus*, DQ083161); mountain hawk eagle (*Nisaetus nipalensis*, AP008238); Pacific loon (*Gavia pacifica*, AP009190); and American kestrel (*Falco sparverius*, DQ780880) (see Supplementary Material S3 online for the consensus tree of the above nine taxa and their close relatives).

Probabilities of Observing Predefined Clades

It is straightforward to calculate the probability of observing a prespecified clade in a tree from a new data set, that is, a tree using additional data not used to predict the clade (see Appendix A for details). In general, for n taxa, there are $B(n) = (2n - 5)!!$ unrooted binary trees, where the double factorial notation (!!) is the product of every second number, that is, $1 \times 3 \times 5 \times \ldots \times 2n - 5$ (table 1). Thus, for example, there are $\approx 3 \times 10^{20}$ possible unrooted binary trees for 20 taxa. In addition, even if a predicted clade for three or more taxa ends up as two subclades on the tree, we can calculate the probability of observing this; and it can indicate that there is still high information content in the data. The calculation can be extended further to three or more subclades, but here, we concentrate mainly on our prior clades being found. We need to define the composition of the prior clades carefully; they alter with the question being considered. For example, we may be interested whether the three parrots do form a natural group, in which case we would calculate the probabilities of three parrots coming together with $n = 40$ taxa in the data set—this is close to a flat prior. Alternatively, we may accept the grouping within orders and consider just $n = 21$ deep Neoavian groups (see later, fig. 2); this is testing the groupings ("priors") of Cracraft (2001). Other tests are also possible.

Results

The first of our three approaches to improving the Neoavian tree was the inclusion of more sequences. The

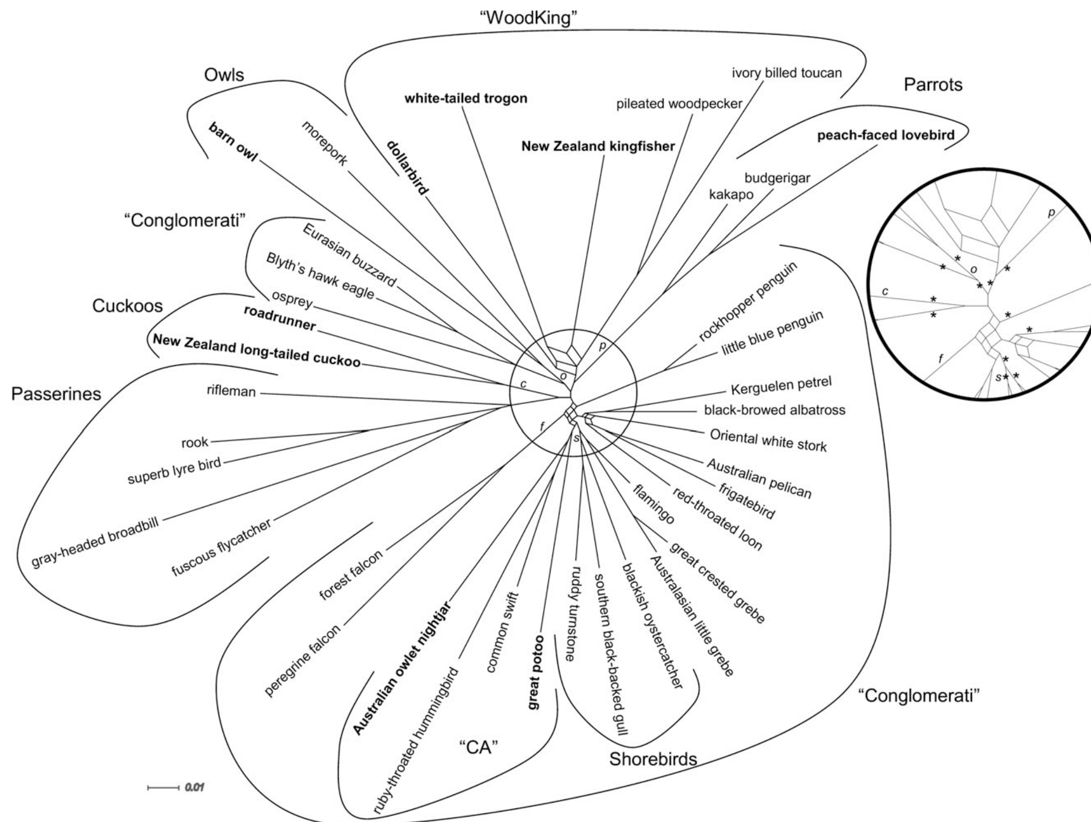

FIG. 1.—Unrooted Bayesian consensus network of Neoaves (modern birds), based on whole mtDNA genomes with "minimal" down-weighting (threshold $s = 6.0$; 20 sites RY coded, 0 sites excluded). Only splits occurring in >25% of trees are included in the network. Inset shows the central portion (indicated by the circle) expanded. Selected branches are labeled to ease comparison: c, cuckoos; f, falcons; o, owls; p, parrots; and s, shorebirds. New taxa included are highlighted in bold. Splits indicated by asterisk have 99+ Bayesian posterior support (BPP).

nine new mt genome sequences are deposited in GenBank under the following accession numbers: dollarbird (EU344978, 17,774 bp); barn owl (EU410491, >16,148 bp, incomplete ND6 and CR); roadrunner (EU410488, 17,091 bp); great potoo (EU344977, >14,396 bp, incomplete ND6 and CR); white-tailed trogon (EU410490, 17,751 bp); Australian owlet-nightjar (EU344979, 18,607 bp); peach-faced lovebird (EU410486, 16,732 bp); New Zealand long-tailed cuckoo (EU410487, 17,559 bp); and the New Zealand kingfisher (EU410489, 17,549 bp). Following the gene-order nomenclature from Gibb et al. (2007), the cuckoo and roadrunner were found to have the remnant CR(2) gene order first described in the falcon (Mindell et al. 1998). All other birds have the standard avian gene order as in the chicken (Desjardins and Morais 1990) with the possible exception of the great potoo and barn owl; because their sequences are currently incomplete in the CR and adjacent genes, their gene order is unknown.

Site-stripping (noise reduction) was our second approach to improving the tree. Bayesian analyses were carried out on a range of down-weighting values, initially excluding the outgroup, then including it. Figure 1 shows the result for the unrooted Neoavian data set with a high threshold (minimum down-weighting, strictness $s = 6.0$), whereas figure 2 shows results from the maximum down-weighting used (a low threshold, strictness $s = 2.0$). Both figures are networks showing splits occurring in at least 25% of Bayesian phylogenies (Holland et al. 2004).

Our third approach was a quantitative test of predefined groupings. There are several clades that were predicted and were returned both with different down-weightings and with or without the outgroup. Perhaps, the most straightforward example is that of the three parrots; this is just a trial calculation because we really had no doubt that the parrots would come together, as predicted from previous DNA sequence analyses (de Kloet RS and de Kloet SR 2005). The New Zealand ground parrot (kakapo) was the first to diverge. This is just a trial calculation with flat priors, and the probability of three taxa coming together on a tree is $P = 0.0005$ ($P = 3/(2n - 5)(2n - 7)$, with $n = 40$ (see Appendix A). There is considerable rate variation within parrots, with the kakapo being slower than the others, whereas the peach-faced lovebird is the

fastest (which is evident by the long edge in both figs. 1 and 2).

Again as expected from our informal priors, the New Zealand long-tailed cuckoo and the roadrunner (a ground cuckoo) always paired and do in prior morphological and molecular studies; thus, we considered them to form one independent "taxonomic group." Of the 40 taxa analyzed, we consider only 21 groups to be independent (see later and Appendix A for details of the groupings); consequently, the probability of two taxa forming a clade in the tree is $P = 0.027$ ($P = 1/(2n - 5)$, with $n = 21$ see Appendix A). Perhaps unexpectedly, cuckoos then group as sister to the five passerines in our data set, a result also observed by Mayr et al. (2003) using combined molecular and morphological data; however, bootstrap support was relatively low (see their figs. 5 and 7). As in previous analyses, the passerines always group together, with the New Zealand wrens (rifleman in this case), basal to the oscines and suboscines. The cuckoo–passerine pairing was found with all down-weightings, both with the ingroup alone and with the outgroup included (fig. 3). This grouping of the cuckoo–passerine clades is an interesting hypothesis and requires testing with both nuclear-coding sequences and rare genomic changes (Boore 2006).

The first real test of Cracraft's (2001) priors, that is, testing groups above the order level, stems from the clade we refer to as "WoodKing". In this case, all five taxa were always found as a clade, irrespective of the down-weighting, and with or without the outgroup (i.e., with both unrooted and rooted trees). As expected, the Piciformes (pileated woodpecker and ivory billed toucan) were always paired. Thus, if it is assumed that the woodpecker and toucan are sufficiently close, then really there are only four independent taxa, and $P = 7.12 \times 10^{-6}$ is a highly significant result (see Appendix A for details). Note that the calculation of the probabilities allows all possible ways of observing the four taxa on the tree, including any paraphyletic groups within it. As shown in figures 1 and 2, there is conflicting signal linking the kingfisher with either the white-tailed trogon or with the dollarbird. Being able to show both signals is a major advantage of networks (Holland et al. 2004) because it helps to prevent premature conclusions. With increased down-weighting of the faster sites (a stricter threshold), we again observed variation in the position of the dollarbird, which tended to be deeper in the clade, though still within the "WoodKing" group. Additional taxon sampling should resolve the splits fully, but our main conclusion is that the predicted grouping of Piciformes, Trogoniformes, and Coraciiformes (Cracraft 2001) is found (though not necessarily reciprocally monophyletic). Hackett et al. (2008) sampled more widely and found strong support for Coraciiformes + Piciformes (see their fig. 2, clade C), although Trogoniformes fell outside this and had less support (see our Supplementary Material S2 online).

However, the next result was not in our informal priors. We found that the parrots and the WoodKing group, irrespective of down-weighting extent, are always adjacent clades on the unrooted tree (supported by high BPP values but not by bootstrapping, see Supplementary Material S2 online). This result will need further investigation as the parrot lineage has considerable rate variation and there is a long internal branch from the three parrots to the rest of the tree. Because the grouping was not part of our priors, we cannot calculate the increase in support, but if additional data types support this relationship, then the probabilities could be calculated.

Next, we consider the "SCA" group (Strigiformes, Caprimulgiformes, and Apodiformes) predicted by Cracraft (2001). Four of the six taxa available form a monophyletic clade, to the exclusion of the two owls. The great potoo, Australian owlet-nightjar, common swift, and ruby-throated hummingbird formed a group of four ("CA"). The swift and hummingbird pairing was highly supported by BPP and bootstrapping (see Supplementary Material S2 online) as predicted from previous studies (Johansson et al. 2001; van Tuinen and Hedges 2001; Cracraft et al. 2004; Hackett et al. 2008; Morgan-Richards et al. 2008), the Australian owlet-nightjar came deeper, and finally the great potoo, which was always basal. It should be noted that the potoo did move slightly with bootstrapping, and we found more support for the Apodiformes and the owlet-nightjar to the exclusion of the potoo (see Supplementary Material S2 online) (see also Mayr 2002a; Barrowclough et al. 2006). Hackett et al. (2008) also observed high support (98%) for Apodiformes + *Aegotheles*, as have previous nuclear and morphological studies (see also Mayr 2002b; Barrowclough et al. 2006). The position of this group of four was variable in the tree. With lesser down-weighting (a higher threshold, $s = 6.0$), the group was found within the informal "Conglomerati" group (fig. 1), but with the maximum down-weighting, it was outside this group (apart from the shorebirds). Similarly, the barn owl and the morepork always paired, joining together quite deep in the tree (i.e., although both are "owls," they represent old divergences; indeed some short preliminary runs did not even join them together, see Supplementary Material S2 online for support values). We need to be cautious here as both owls have some of the highest rates of sequence evolution among the Neoaves. Although the owls did not group with the other members of the "SCA" clade, they were always found to group with the buzzard, hawk eagle (Accipitridae), and osprey. Only with the highest down-weighting used (figs. 2 and 3) did the falcons unite with the owls/Accipitridae/osprey; with the falcons basal, however, we found no bootstrap support and low BPP support for this grouping (see Supplementary Material S2 online). Grouping of the Accipitridae, osprey, and owls is certainly interesting. Relationships within the birds of prey are controversial as speculation over convergence and raptorial specialization has been raised (e.g., Livezey and Zusi 2007). This relationship was not found by Hackett et al. (2008) and needs to be tested with nuclear-coding data. We return to the raptors and owl question later.

Even though the proposed "SCA" clade came out as two groups in the tree, this in itself still has high information content. The simplest calculation is to assume the alternative prediction of the two owls being independent of the other four, and in this case, the probability of observing the "two plus four" grouping is simply the product of the probability of observing a pair ($P = 1/(2n - 5) = 0.013$ for $n = 40$) and the probability of observing a group of four ($P = 2.5 \times 10^{-6}$, see Appendix). This gives the combined

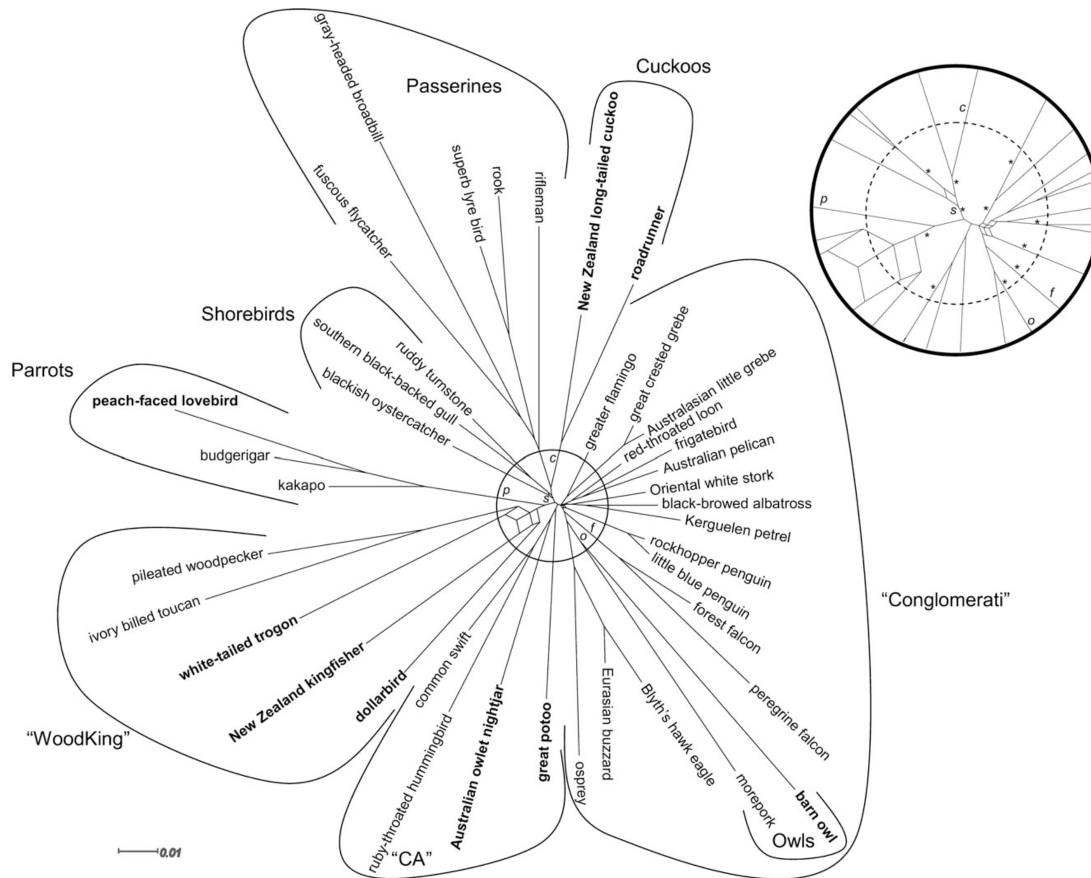

FIG. 2.—Unrooted Bayesian consensus network of Neoaves with the "maximum" down-weighting used (threshold $s = 2.0$; 891 sites RY coded, 159 sites excluded). Otherwise, the conventions are the same as in figure 1.

probability of 3.4×10^{-8} (about 1 chance in 300 million of observing the pattern). Strictly speaking, we need not make the assumption that the two owls (in particular) separate from the other four taxa; there are a total of 21 combinations ($^6C_2$) of pairs from six taxa; only one has the two owls paired. In general, we would multiply the probability by 21 (see Appendix A), still giving around 1 chance in 10 million of observing these two groups on the tree. Similarly, it could have been just a single taxon (six choices) that separated from the other five or two groups of three taxa (10, or $^6C_3$ combinations) but halved because each triplet of taxa is found twice. Appendix A and figure 4 show the general calculation, but in this case (because the owls join with the raptors, as an alternative prediction see Mayr 2005), it is reasonable to use the probability of finding just the two plus four grouping.

Our conclusion at this point is that excellent progress is being made in understanding the deeper levels of phylogeny of the Neoaves. If the predictions from Cracraft (2001) are, in general, being well supported with new data, then this implies that the basal polytomy is resolvable. However, our next step is to check that there are no major changes when the outgroup is added; this has been a major problem when a smaller number (24 ingroup) of taxa were sequenced (see Harrison et al. 2004).

Rooted Tree

Figure 3 shows our tree rooted with five Galloanserae taxa for the maximum down-weighting value (threshold $s = 2$). The two main points from this figure are that

1. there is only one local change to the unrooted tree when the root is added, and
2. the root comes between parrots and all other Neoaves.

The first point refers to the difference in the position of the flamingo/grebe clade between the unrooted and rooted trees. In the unrooted tree, they are basal to the loon/albatross/Pelecaniformes/stork grouping. In the rooted tree, we find them basal to the same group as before plus the Accipitridae/osprey/owl/falcon clade. In other words, the tree is "locally stable" in the terminology of Cooper and Penny (1997). Finding just a single edge (branch) different between two trees is very highly significant, about 5.7×10^{-54} for 40 taxa in the ingroup (Penny et al. 1982).

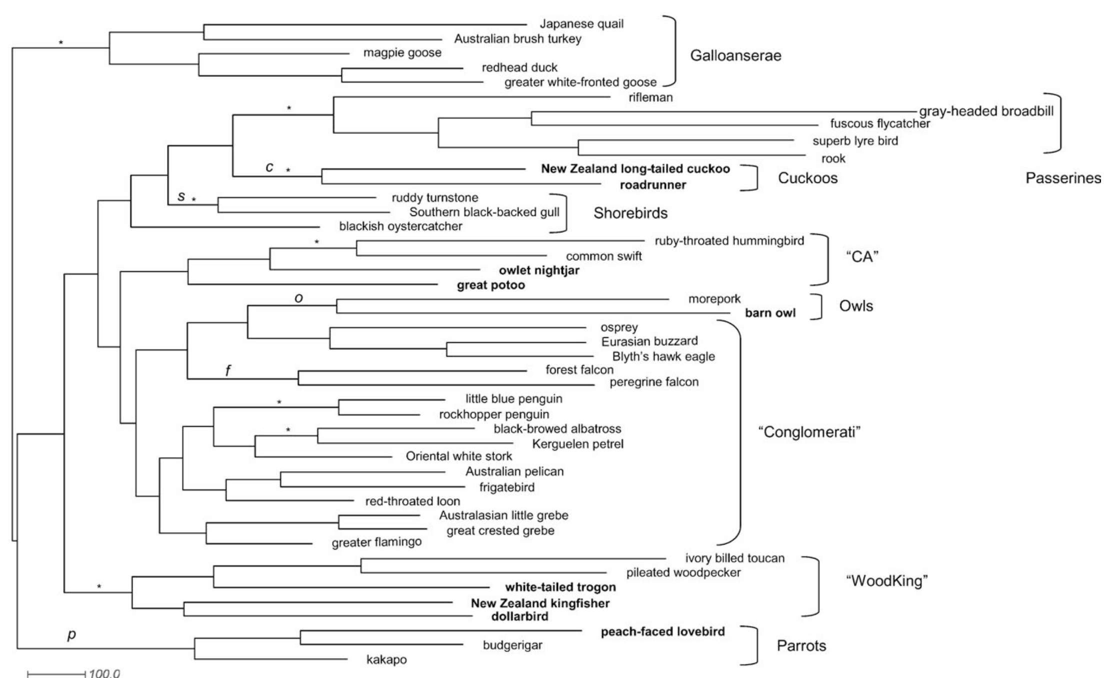

FIG. 3.—Rooted Bayesian phylogram of Neoaves with maximum down-weighting and including the five Galloanserae as the outgroup. Conventions are the same as in figure 1.

On the surface, this latter finding (the root between parrots and other Neoaves) could be suspicious because the branch at the base of the parrots is the longest internal branch on the tree! However, with lesser down-weighting ($s = 6.0$), the root joins one step away on the short branch at the base of the parrot/"WoodKing" group. The movement of the root from a shorter branch is not expected unless parrots are really the first subdivision of Neoaves. Without any prior information as to the root of the Neoaves, it is not possible to give any quantitative statement of confidence in this rooting.

Other aspects of the rooted tree are also interesting and noteworthy. We again find the same owl/raptor clade appearing only with the strongest down-weighting—that is, only the Accipitridae/osprey (but not falcons) unite with the owls with lesser down-weighting. The tree produced from BEAST resulted in the falcons joining the "CA" group (Supplementary Material S1 online). This latter observation, if real, is interesting as both falcons and the "CA" taxa are in-flight foragers typically specialized for bill-capture of prey; in comparison, the owls are specialized for inflight talon capture (like hawks). However, there was only conflicting support for this relationship shown in the network figure 1, and therefore, we are unable to comment further without additional data. It is worth noting that some of the deeper groupings within this large raptor/shore bird/water bird ("Conglomerati") still vary somewhat depending on taxon sampling (Morgan-Richards et al. 2008), and it is not clear yet whether further taxon sampling will resolve these issues. It may well be that the three subgroups in the "Conglomerati" (group (vi) of Cracraft 2001, see Introduction) should be considered independently.

Discussion

Resolving the evolutionary relationships within the modern birds (Neoaves) has been both problematic and controversial, with some suggesting that it will never be fully resolved (Poe and Chubb 2004). Here, we have shown that with more and longer DNA-coding sequences, along with improved noise reduction techniques, relationships within Neoaves are expected to be resolvable. This should occur relatively quickly with the addition of data from nuclear coding and rare genomic changes as they become available. Our approach to resolving this issue has been 3-fold: presenting additional data, improving noise reduction/signal enhancement techniques, and getting beyond flat priors, where it is assumed (sometimes correctly!) that there is no useful prior knowledge.

We consider it important that prior hypotheses can be evaluated quantitatively, and thus, the formulae developed in Appendix A will be useful for a wide range of studies. However, there is still more work required in developing these analyses. For example, the calculation for prespecified groups is for the optimal placement of that clade on the tree (even if the bootstrap or Bayesian posterior priors are less than 100% support). If these support values are indeed higher, then this gives even more confidence in the clades, so in that respect our probabilities are conservative. Thus, more thought is required on how to combine the

calculations developed here with the strength of support for branches in the tree from new data. Similarly, the calculation allows for any subtree within the clades (or subclades). However, if we prespecified that a particular grouping and subtree is expected, then the probability of finding this arrangement on the tree is even lower. For example, the calculation allows 15 ways (5!!) that a group of four could join the larger tree. This is because there are three unrooted trees for four taxa, each with five edges for joining to the rest of the tree—and the prediction does not specify which of the 3×5 (15) trees would be observed. In contrast, if we predict precisely how the group will join (forming a clade), then the number of possible trees is reduced 15-fold. Overall, it is important that we make better use of well-considered prior hypotheses when studying trees based on new data. Even though we are a long way from having the "one tree" (in this case) for Neoaves, we can be confident that the issue is resolving and that the data sets have, in a formal sense, high information content.

A related question is estimating how many trees within Neoaves are still likely—a "confidence set" of trees. At this higher taxonomic level, it is not yet clear which groupings are stable and which may be subject to change. For our Neoaves data set, there appears to be around 21 major groupings (shown as crossing into the inner dashed line in the insert of fig. 2 and in Appendix A). These groupings include Passeriformes, cuckoos, parrots, two shore bird lineages, three raptor clades (falcons, buzzard/osprey, and owls), rollers, kingfishers, woodpecker/toucan/trogon, potoos, owlet-nightjars, hummingbirds/swifts, flamingoes, grebes, Pelecaniformes, tubenoses, storks, penguins, and loon.

In principle, there are $R(21) \approx 3 \times 10^{23}$ possible rooted binary trees of which only a vanishingly small proportion are realistic. With eutherian mammals, there were initially 19 orders identified and therefore $R(19) \approx 2 \times 10^{20}$ possible rooted binary trees. But it quickly became apparent that no more than about 10^2 trees were likely (Lin et al. 2002)—an improvement of 18 orders of magnitude. The next step for birds is an equivalent analysis for Neoaves and thus getting beyond the debilitating view of the flat priors—that all trees were equally likely and that there is no information in previous studies.

The results from down-weighting the faster evolving sites are interesting, and these techniques need to be developed and tested further. With increased down-weighting, we find closer agreement between earlier predictions and the actual tree found. From first principles, we expect that reducing the influence of the saturated sites will help, and in general, it appears that the predefined groups are found more strongly. It is for this type of reason that we would like to see further development and evaluation of the noise reduction techniques including their application to nuclear-coding data. Although it is outside the range of this study, a simulation study is now an important next step.

Turning now from the more general issues to the Neoaves in particular, our current study assesses the stability and probability of the six groups proposed by Cracraft (2001) using a novel analysis method to down-weight sites of whole mt genome sequences. These Cracraftian priors were found to be robust with four of the six groups within Neoaves being recovered, the other two having relatively small changes; the owls moving to the raptors and possibility that the raptor/water carnivores ("Conglomerati") may be diphyletic. In this latter respect, it appears preferable, at least in the short term, to treat the three subgroups of Cracraft's group (vi) independently. Our resulting phylogenies appear relatively stable, differing little in overall topology both with and without the addition of the outgroup (Galloanserae).

The placement of the root of Neoaves needs additional support. Our analyses put the root in one of two possible locations: either with parrots (which have a higher mutation rate in mtDNA) as the most basal lineage or, with lesser down-weighting, the parrots plus the "WoodKing" grouping as the basal clade. Because the separation of parrots from the rest of the Neoaves occurred with the highest down-weighting, we cannot easily dismiss this possible rooting. Morphologically, parrots are distinct (for review, see Dyke and Cooper 2000; Waterhouse 2006) and a fragment of a mandible from the Maastrichtian, latest Cretaceous (65–70 Ma Lance Formation, North America) has been described (Stidham 1998). However, the identification of this fossil is contentious (Dyke and Mayr 1999; Mayr 2002a), though previous molecular work suggests a Cretaceous diversification for each of the African, Australian, and South American parrots (Miyaki et al. 1998). Dating carried out on our current data set suggests that the most basal parrot in our analysis, the kakapo, split from the other parrots sometime after the K–Pg boundary. The lineage as a whole however predates the K–Pg boundary with a mean date of ~85 Ma (see Supplementary Material S1 online). Hackett et al. (2008) suggested that the root of Neoaves be placed with the sister grouping of the Podicipediformes/Phoenicopteriformes/Phaethontidae/Pteroclididae/Mesitornithidae/Columbiformes to Gruiformes/Caprimulgiformes/Apodiformes (see their fig. 2, clades K, L, M, and N). However, they do not give support values for this lineage, and they state that their rooted tree only occurs when the β-fibrinogen intron 7 data are included. We have already demonstrated (Morgan-Richards et al. 2008) that by our more rigorous standards (deleting columns around gaps back to a constant column) that the intron sequences of this locus are not informative for deep divergences. Clearly, the root of Neoaves is still under debate; however, we now have a number of possibilities to be tested by future analyses.

Conclusion

The basal split within Neoaves and its timing are resolvable issues. If modern birds radiated over a short period (say 2–5 Ma) after the K–Pg extinction, then it will be very difficult to resolve the polytomy at the base of modern birds. However, using whole mtDNA coding sequences gives us a solid point from which to build. By the addition of more taxa, nuclear-coding sequences, and rare genomic changes, we expect resolution at the ordinal level to be achievable. In addition, the further development of noise reduction techniques for coding sequences (both organellar and nuclear) will enable more robust trees to be produced. We estimate that at least 12 Neoavian lineages had evolved prior to the K–Pg boundary, similarly, van Tuinen et al. (2006) and Brown et al. (2008) support pre-K–Pg origins

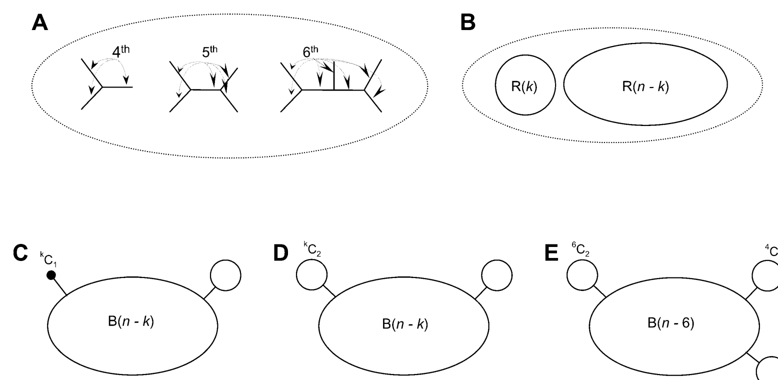

FIG. 4.—The basis for calculations for the probability of finding predefined clades (or subclades) on a tree. (*A*) A procedure for counting the number of trees. There is only one unrooted tree for three taxa, and there are three edges (branches) to add the fourth taxon—giving three trees for four taxa. Each of these three trees has five edges where the fifth taxon can be added giving $1 \times 3 \times 5 = 15$ trees. Similarly, each of these 15 trees has 7 edges for the sixth taxon to be added, leading to the formula $B(n) = (2n - 5)!!$, for the number of unrooted binary trees. (*B*) Calculating the probability of prespecified clade of k taxa on a new tree. There are $R(k)$ rooted trees for the clade of k taxa and $B(n - k)$ for the remaining $n - k$ taxa, leading to the calculation for the probability of observing a prespecified clade of k taxa forming a clade in each tree with n taxa. (*C*) The proposed grouping of k taxa with two subgroups; a single taxon in one and $k - 1$ in the other. (*D*) A similar case with two taxa on one group and $k - 2$ in the other. There are ${}^{k}C_2$ (k choose 2) ways of selecting the two taxa. (*E*) An example where the group of $k = 6$ ends up as $m = 3$ subgroups of 2 taxa each. There are ${}^{6}C_2$ for selecting the first pair of taxa and ${}^{4}C_2$ for the second pair and 3! ways for ordering the three pairs on a given tree.

for multiple modern lineages. In addition, Clarke et al. (2004) estimate a minimum of five Anseriformes lineages (duck, chicken, and ratite bird relatives) before this time supporting the presence of a diverse array of modern bird lineages prior to the extinction event. Lastly, with regards to the search for the one tree, we feel our data have made significant progress with support for four of the six Cracraftian groups. Given the very low probability of observing groupings by chance, the data are highly informative and should stimulate future work incorporating data from all facets of avian evolution.

Supplementary Material

Supplementary Material S1 (fossil calibrations and molecular dating with BEAST), Supplementary Material S2 (tabulated results from Bayesian analysis [BPP] and bootstrap analyses [bs] using RAxML and PhyML), and Supplementary Material S3 (consensus tree including all outgroup taxa used for alignment before for site-stripping) are available at *Molecular Biology and Evolution* online (http://www.mbe.oxforjournals.org/).

Acknowledgments

We thank the Australian Museum, Sydney; the Louisiana State University Museum of Natural Science Collection of Genetic Resources; Fritz Geiser, University of New England, Armidale, Australia; and Dick Gill, Department of Conservation, Waikanae, for supplying tissue samples. Thanks to Trish McLenachan for help with alignments and sequences and to Olga Kardailsky and Steve Trewick for additional mitochondrial genomes. Thanks to Simon Hills and Patrick Biggs for computer time and assistance with BEAST analyses. Work carried out at the Allan Wilson Centre for Molecular Ecology and Evolution was supported by the Marsden Fund (Royal Society of New Zealand) to D.P. The manuscript was improved by constructive comments from Associate Editor M. Gouy and three anonymous referees.

Appendix A

The Probability of Observing a PreSpecified Clade

We calculate the probability of observing a prespecified clade of k taxa in a binary tree on n taxa, as the proportion of all binary trees containing that clade. This can be extended to the probabilities of the clade being found as $m = 2$, 3, or even more subclades. We see below (that for all but small values of n and k) that these probabilities are very low so that finding a prespecified clade on a tree formed from new data is highly significant.

It is well known that for n taxa there are

$$B(n) = (2n - 5)!! = 1 \times 3 \times 5 \times \ldots \times (2n - 5) \quad (1)$$

unrooted binary trees where each tip (leaf) of the tree is labeled by a unique taxon (see Penny et al. 1991). Similarly, the number of rooted binary trees is

$$R(n) = (2n - 3)!! = B(n + 1). \quad (2)$$

A simplified approach to deriving the formulae is indicated in figure 4*A*, and the calculations are straightforward in an Excel spreadsheet.

Probability of a Specific Subset Forming a Single Clade (i.e., $m = 1$)

The probability (P) of observing a predefined clade of k taxa in a binary tree of n taxa is

$$P(n,k) = R(k) \times B(n - k + 1)/B(n), \quad (3)$$

where the numerator is the number of rooted subtrees for the clade ($R(k)$), multiplied by the number of trees on the remaining taxa ($B(n - k + 1)$, including a leaf for the clade). Dividing by the number of unrooted trees ($B(n)$) gives the proportion of trees having that clade of k taxa.

For rooted trees, the probability (P_R) of observing a predefined clade of k taxa in a rooted binary tree of n taxa is similarly

$$P_R(n,k) = R(k) \times R(n - k + 1)/R(n) = P(n + 1, k). \quad (4)$$

For this question, we consider all binary trees as equally likely; the trees are derived from a Markov model where there is no prior information about the distribution of tree shapes (Steel and Penny 1993). For two taxa in the predefined clade, the equation simplifies (see table 1) to

$P(n,2) = 1/(2n-5)$ for unrooted trees and

$P_R(n,k) = 1/(2n-3)$ for rooted trees.

And for three taxa, it simplifies to

$P(n,3) = 3/(2n-5)(2n-7)$ for rooted trees and

$P_R(n,3) = 3/(2n-3)(2n-5)$ for rooted trees.

In our analyses, we consider the 40 taxa to account for 20 independent taxonomic groupings: passerines (five taxa), cuckoos (two taxa), parrots (three taxa), shorebirds (three taxa), owls (two taxa), dollarbird, kingfisher, trogon, woodpecker + toucan (two taxa), potoo, owlet-nightjar, Apodiformes (two taxa), Accipitriformes (three taxa), falcons (two taxa), flamingo + grebes (three taxa), Pelecaniformes (two taxa), tubenoses (two taxa), stork, penguins (two taxa), and the loon. For rooted trees, there are 21 independent groupings, the above plus Galloanserae (five taxa). If, for example, $k = 5$ and $n = 40$ and the probability of observing a prespecified clade on new data where all taxa are included is

$$P_R(40,5) = R(5) \times R(36)/R(40) \approx 3.5 \times 10^{-6}$$

or where only the independent taxonomic groups are included,

$$P_R(21,5) = R(5) \times R(17)/R(21) \approx 6.3 \times 10^{-5}.$$

Predefined Clade Found as $m \geq 2$ Subgroups

The calculations can be extended to cases where the predicted clade is partitioned into $m = 2$ or more subclades on a tree. Figure 4*C* and *D* shows two cases where a predefined clade appears in two separate areas of the tree ($m = 2$). For a large number of taxa, it is still most unlikely that a predefined clade will be in just two locations on a new tree. In the case shown here, there are $k = 4$ taxa in the clade, and for $m = 2$, they can occur as either a single taxon and a group of three (fig. 4*C*) or as two groups, each with two taxa (fig. 4*D*). When the clade of k taxa is split into m subclades, with $k_1, k_2, \ldots, k_m$ taxa, respectively, then we must consider each combination of the m subclades separately and we find that the probability is

$$P(n,k,m) = \frac{k!\,B(n - k + m)}{m!\,B(n)} \sum_{\Sigma k_i = k} \prod_i \frac{R(k_i)}{k_i!}. \quad (5)$$

Literature Cited

Barrowclough GF, Groth JG, Mertz LA. 2006. The RAG-1 exon in the avian order Caprimulgiformes: phylogeny, heterozygosity and base composition. Mol Phyl Evol. 41:238–248.

Boore JL. 2006. The use of genome-level characters for phylogenetic reconstruction. Trends Ecol Evol. 21:439–446.

Brown JW, Payne RB, Mindell DP. 2007. Nuclear DNA does not reconcile 'rocks' and 'clocks' in Neoaves: a comment on Ericson et al. Biol Lett 3:257–259.

Brown JW, Rest JS, García-Moreno J, Sorenson MD, Mindell DP. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol. 6:6.

Chubb AL. 2004. New nuclear evidence for the oldest divergence among neognath birds: the phylogenetic utility of ZENK (i). Mol Phyl Evol. 30:140–151.

Clarke JA, Tambussi CP, Noriega JI, Erikson GM, Ketcham RA. 2004. Definitive fossil evidence for the extant avian radiation in the Cretaceous. Nature. 433:305–308.

Cooper A, Penny D. 1997. Mass survival of birds across the Cretaceous-Tertiary boundary: molecular evidence. Science. 275:1109–1113.

Cracraft J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proc Roy Soc Lond B Biol Sci. 268:459–469.

Cracraft J, Barker FK, Braun M, et al. 2004. (13 co-authors). 2004. Phylogenetic relationships among modern birds (Neornithes). In: Cracraft J, Donoghue MJ, editors. Assembling the Tree of Life. New York: Oxford University Press. p. 468–489.

Cracraft J, Clarke J. 2001. The basal clades of modern birds. In: Gauthier J, Gall LF, editors. New perspectives on the origin and early evolution of birds. New Haven (CT): Peabody Museum of Natural History, Yale University. p. 143–156.

Curole JP, Kocher TD. 1999. Mitogenomics: digging deeper with complete mitochondiral genomes. Trends Ecol Evol. 14:394–398.

de Kloet RS, de Kloet SR. 2005. The evolution of the spindlin gene in birds: sequence analysis of an intron of the spindlin W and Z gene reveals four major divisions of the Psittaciformes. Mol Phyl Evol. 36:706–721.

Delsuc F, Phillips MJ, Penny D. 2003. Comment on "Hexapod origins: monophyletic or paraphyletic?" Science. 301:1482.

Desjardins P, Morais R. 1990. Sequence and gene organization of the chicken mitochondrial genome. J Mol Evol. 32:153–161.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:699–710.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 7:214.

Dyke GJ, Cooper JH. 2000. A new psittaciform bird from the London clay (Lower Eocene) of England. Palaeontology. 43:271–285.

Dyke GJ, Mayr G. 1999. Did parrots exist in the Cretaceous period? Nature. 399:317–318.
Ericson PGP, Anderson CL, Britton T, Elzanowski A, Johansson US, Källersjö M, Ohlson JI, Parsons TJ, Zuccon D, Mayr G. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol Lett. 2:543–547.
Feduccia A. 2003. 'Big bang' for Tertiary birds? Trends Ecol Evol. 18:172–176.
García-Moreno J, Mindell DP. 2000. Using homologous genes on opposite sex chromosomes (gametologs) in phylogenetic analysis: a case study with avian CHD. Mol Biol Evol. 17:1826–1832.
Gibb GC, Kardailsky O, Kimball RT, Braun E, Penny D. 2007. Mitochondrial genomes and avian phylogeny: complex characters and resolvability without explosive radiations. Mol Biol Evol. 24:269–280.
Groth JG, Barrowclough GF. 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. Mol Phyl Evol. 12:115–123.
Guindon D, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 52:696–704.
Hackett SJ, Kimball RT, Reddy S, et al. (17 co-authors). 2008. A phylogenomic study of birds reveals their evolutionary history. Science. 320:1763–1768.
Harrison GL, McLenachan PA, Phillips MJ, Slack KE, Cooper A, Penny D. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late Cretacous. Mol Biol Evol. 21:974–983.
Harshmann J, Braun EL, Braun MJ, et al. (17 co-authors). 2008. Phylogenomic evidence for multiple losses of flight in ratite birds. Proc Natl Acad Sci USA. 105:13462–13467.
Hendy MD, Penny D. 1989. A framework for quantitative study of evolutionary trees. Syst Biol. 38:297–309.
Holland BR, Huber KT, Moulton V, Lockhart PJ. 2004. Using consensus networks to visualize contradictory evidence for species phylogeny. Mol Biol Evol. 21:1459–1461.
Holland BR, Penny D, Hendy MD. 2003. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock—a simulation study. Syst Biol. 52:229–238.
Honeycutt RL, Adkins RM. 1993. Higher-level systematics of eutherian mammals—an assessment of molecular characters and phylogenetic hypotheses. Annu Rev Ecol Syst. 24: 279–305.
Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference on phylogenetic trees. Bioinformatics. 17:754–755.
Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol. 23: 254–267.
Jeffroy O, Brinkmann H, Delsuc F, Philippe H. 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22:225–231.
Johansson US, Parsons TJ, Irestedt M, Ericson PGP. 2001. Clades within the 'higher land birds', evaluated by nuclear DNA sequences. J Zool Syst Evol Res. 39:37–51.
Lin Y-H, McLenachan PA, Gore AR, Phillips MJ, Ota R, Hendy MD, Penny D. 2002. Four new mitochondrial genomes and the increased stability of evolutionary trees of mammals from improved taxon sampling. Mol Biol Evol. 19:2060–2070.
Livezey BC, Zusi RL. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool J Linn Soc. 149:1–95.
Lockhart PJ, Penny D, Hendy MD, Howe CJ, Beanland TJ, Larkum AWD. 1992. Controversy on chloroplast origins. FEBS Lett. 301:127–131.
Löytynoja A, Goldman N. 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science. 320:1632–1635.
Matthee CA, Eick G, Willows-Munro S, Montgelard C, Pardini AT, Robinson TJ. 2007. Indel evolution of mammalian introns and the utility of non-coding nuclear markers in eutherian phylogenetics. Mol Phylogenet Evol. 42:827–837.
Mayr G. 2002a. On the osteology and phylogenetic affinities of the Pseudasturidae—lower Eocene stem-group representatives of parrots (Aves, Psittaciformes). Zool J Linn Soc. 136: 715–729.
Mayr G. 2002b. Osteological evidence for paraphyly of the avian order Caprimulgiforms (nightjars and allies). J Ornithol. 143:82–97.
Mayr G. 2005. The postcranial osteology and phylogenetic position of the Middle Eocene Messelastur gratulator Peters, 1994—a morphological link between owls (Strigiformes) and falconiform birds? J Vertebr Paleontol. 25:635–645.
Mayr G, Manegold A, Johansson US. 2003. Monophyletic groups within 'higher land birds'—comparison of morphological and molecular data. J Zoolog Syst Evol Res. 41: 233–248.
Mindell DP, Sorenson MD, Dimcheff DE. 1998. Multiple independent origins of mitochondrial gene order in birds. Proc Natl Acad Sci USA. 95:10693–10697.
Miyaki CY, Matioli SR, Burke T, Wajntal A. 1998. Parrot evolution and paleogeographical events: mitochondrial DNA evidence. Mol Biol Evol. 15:544–551.
Morgan-Richards M, Trewick SA, Bartosch-Härlid A, Kardailsky O, Phillips MJ, McLenachan PA, Penny D. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol Biol. 8:20.
Penny D, Foulds LR, Hendy MD. 1982. Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. Nature. 297:197–200.
Penny D, Hendy MD. 1986. Estimating the reliability of evolutionary trees. Mol Biol Evol. 3:403–417.
Penny D, Hendy MD, Steel MA. 1991. Testing the theory of descent. In: Miyamoto MM, Cracraft J, editors. Phylogenetic analysis of DNA sequences. Oxford: Oxford University Press. p. 155–183.
Penny D, Phillips MJ. 2004. The rise of birds and mammals: are microevolutionary processes sufficient for macroevolution? Trends Ecol Evol. 19:516–522.
Penny D, White WT, Hendy MD, Phillips MJ. 2008. A bias in ML estimates of branch lengths in the presence of multiple signals. Mol Biol Evol. 25:239–242.
Pereira SL, Baker AJ. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol Biol Evol. 23: 1731–1740.
Philippe H, Lopez P, Brinkmann H, Budin K, Germot A, Laurent J, Moreira D, Muller M, Le Guyader H. 2000. Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. Proc Roy Soc Lond B Biol Sci. 267:1213–1221.
Phillips MJ, Delsuc F, Penny D. 2004. Genome-scale phylogeny: sampling and systematic errors are both important. Mol Biol Evol. 21:1455–1458.
Phillips MJ, McLenachan PA, Down C, Gibb GC, Penny D. 2006. Combined mitochondrial and nuclear DNA sequences resolve the interrelations of the major Australasian marsupial radiations. Syst Biol. 55:122–137.
Phillips MJ, Penny D. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. Mol Phyl Evol. 28:171–185.
Poe S, Chubb AL. 2004. Birds in a bush: five genes indicate explosive evolution of avian orders. Evolution. 58:404–415.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics. 14:817–818.
Proakis JG, Rader CM, Ling F, Nikias CL, Moonen M, Proudler IK. 2002. Algorithms for statistical signal processing. Upper Saddle River (NJ): Prentice Hall.
Rodriguez-Ezpeleta N, Brinkmann H, Roure B, Lartillot N, Lang BF, Phipippe H. 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. Syst Biol. 56:389–399.
Shapiro LH, Dumbacher JP. 2001. Adenylate kinase intron 5: a new nuclear locus for avian systematics. Auk. 118:248–255.
Slack KE, Delsuc F, McLenachan PA, Arnason U, Penny D. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol Phyl Evol. 42:1–13.
Slack KE, Janke A, Penny D, Arnason U. 2003. Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. Gene. 302:43–52.
Slack KE, Jones CM, Ando T, Harrison GL, Fordyce E, Arnason U, Penny D. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol. 23:1144–1155.
Sorenson MD, Oneal E, García-Moreno J, Mindell DP. 2003. More taxa, more characters: the Hoatzin problem is still unresolved. Mol Biol Evol. 20:1484–1499.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.
Steel MA, Penny D. 1993. Distributions of tree comparison metrics—some new results. Syst Biol. 42:126–141.
Stidham TA. 1998. A lower jaw from a Cretaceous parrot. Nature. 396:29–30.
Susko E, Roger AJ. 2007. On reduced amino acid alphabets for phylogenetic inference. Mol Biol Evol. 24:2139–2150.
Swofford DL. 2001. PAUP* phylogenetic analysis using parsimony (* and other methods). Sunderland (MA): Sinauer Associates.
van Tuinen M, Hedges SB. 2001. Calibration of avian molecular clocks. Mol Biol Evol. 18:206–213.
van Tuinen M, Stidham TA, Hadly EA. 2006. Tempo and mode of modern bird evolution observed with large-scale taxonomic sampling. Hist Biol. 18:209–225.
Waterhouse DM. 2006. Parrots in a nutshell: the fossil record of Psittaciformes (Aves). Hist Biol. 18:223–234.
Whitfield JB, Lockhart PJ. 2007. Deciphering ancient rapid radiations. Trends Ecol Evol. 22:258–265.
Zwickl DJ. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion [PhD dissertation]. [Austin (TX)]: The University of Texas at Austin.

Manolo Gouy, Associate Editor

Accepted October 27, 2008

Syst. Biol. 59(1):90–107, 2010
© The Author(s) 2009. Published by Oxford University Press, on behalf of the Society of Systematic Biologists. All rights reserved.
For Permissions, please email: journals.permissions@oxfordjournals.org
DOI:10.1093/sysbio/syp079
Advance Access publication on November 13, 2009

Tinamous and Moa Flock Together: Mitochondrial Genome Sequence Analysis Reveals Independent Losses of Flight among Ratites

MATTHEW J. PHILLIPS[1]*, GILLIAN C. GIBB[2], ELIZABETH A. CRIMP[2], AND DAVID PENNY[2]

[1]*Centre for Macroevolution and Macroecology, School of Botany and Zoology, Research School of Biology, Australian National University, Canberra, ACT 0200, Australia; and*
[2]*Allan Wilson Center and Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand;*
**Correspondence to be sent to: Centre for Macroevolution and Macroecology, Research School of Biology, Australian National University, Gould Building, Canberra, ACT 0200, Australia; E-mail: matt.phillips@anu.edu.au.*

Received 19 January 2009; reviews returned 20 April 2009; accepted 14 October 2009
Associate Editor: Adrian Paterson

Abstract.—Ratites are large, flightless birds and include the ostrich, rheas, kiwi, emu, and cassowaries, along with extinct members, such as moa and elephant birds. Previous phylogenetic analyses of complete mitochondrial genome sequences have reinforced the traditional belief that ratites are monophyletic and tinamous are their sister group. However, in these studies ratite monophyly was enforced in the analyses that modeled rate heterogeneity among variable sites. Relaxing this topological constraint results in strong support for the tinamous (which fly) nesting within ratites. Furthermore, upon reducing base compositional bias and partitioning models of sequence evolution among protein codon positions and RNA structures, the tinamou–moa clade grouped with kiwi, emu, and cassowaries to the exclusion of the successively more divergent rheas and ostrich. These relationships are consistent with recent results from a large nuclear data set, whereas our strongly supported finding of a tinamou–moa grouping further resolves palaeognath phylogeny. We infer flight to have been lost among ratites multiple times in temporally close association with the Cretaceous–Tertiary extinction event. This circumvents requirements for transient microcontinents and island chains to explain discordance between ratite phylogeny and patterns of continental breakup. Ostriches may have dispersed to Africa from Eurasia, putting in question the status of ratites as an iconic Gondwanan relict taxon. [Base composition; flightless; Gondwana; mitochondrial genome; Palaeognathae; phylogeny; ratites.]

Modern birds have long been taxonomically divided on the basis of palatal characters (e.g., Huxley 1867) into Neognathae, which make up over 99% of all extant avian species, and Palaeognathae, which includes ratites and tinamous. Analyses of nuclear genes and complete mitochondrial (mt) genomes strongly support this primary avian division (e.g., García-Moreno and Mindell 2000; Hugall et al. 2007; Slack et al. 2007). The ratites, which are all flightless, are the most familiar palaeognaths and are generally large herbivores/omnivores. Extant members include the ostrich, rheas, emu, cassowaries, together with the recently extinct (post-human) moa of New Zealand and elephant birds of Madagascar. In addition, there are the smaller and primarily invertebrate-feeding kiwi. Bertelli and Porzecanski (2004) recognized 9 genera and 47 species of tinamou, all in South America. They are ground foraging birds that fly but are not considered strong flyers.

Although some early workers (e.g., Mayr and Amadon 1951) questioned ratite monophyly, recent morphological studies support tinamous and ratites being reciprocally monophyletic sister taxa but provide little consensus on affinities within ratites (see Lee et al. 1997; Livezey and Zusi 2007). Early molecular studies involving immunological distances (Prager et al. 1976), DNA–DNA hybridization (Sibley and Ahlquist 1990), and short DNA sequences (e.g., van Tuinen et al. 2000) are similar in recovering a ratite/tinamou division, while not clearly resolving relationships between ratite families, except for Casuariidae (cassowaries plus emu). Sequencing complete mt genomes, including several extinct moa, provided a substantial leap in statistical power for resolving ratite relationships (Cooper et al. 2001; Haddrath and Baker 2001). These papers appear to have founded an mt consensus among molecular studies on ratite phylogeny (see Fig. 1a), which has since been followed by numerous studies of molecular dating (e.g., Pereira and Baker 2006; Brown et al. 2008), biogeography (e.g., Sanmartín and Ronquist 2004; Karanth 2006), and phylogenetic inference (e.g., Paton et al. 2002).

Ratites, along with southern beaches (*Nothofagus*) and cichlid and galaxiid freshwater fish, are considered both to be quintessential Gondwanan taxa (e.g., Briggs 2003; Waters and Craw 2006) and to provide substantive evidence for vicariance models of biogeography (e.g., Cracraft 1974). Indeed, among these, only ratites are known to have been distributed across all the major Gondwanan landmasses. The early molecular studies of Prager et al. (1976) and Sibley and Ahlquist (1990) claimed good agreement with the vicariance hypothesis: Ostrich (Africa) basal, then Rheas (South America) as sister to the Kiwi (New Zealand), and Casuariidae (Australia-New Guinea).

The mt consensus (Fig. 1a) provides a challenge for interpreting Gondwanan biogeography. First, molecular estimates for the divergence between the Kiwi and Casuariidae (mean/point estimates from 45 Ma [Härlid et al. 1998] to 77 Ma [Pereira and Baker 2006]) postdate the separation of New Zealand and Australia, which began (and was initially rapid) before 80 Ma (Lawver et al. 1992). Paton et al. (2002) estimated older divergences for kiwi, although their results may be compromised by the *Emuarius* calibration date being increased

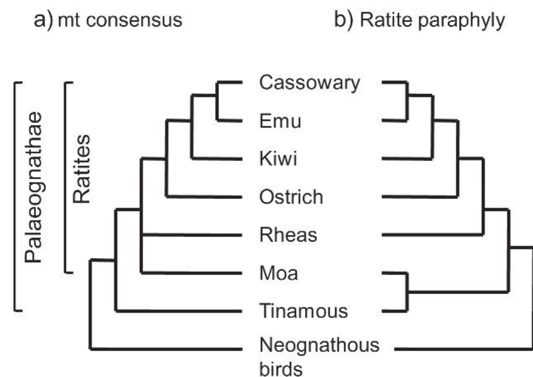

FIGURE 1. The "mt consensus" of palaeognath phylogeny (a) showing ratite monophyly and (b) the alternative topology of ratite paraphyly supported by our reanalysis of the mt data set of Cooper et al. (2001), in which a tinamou–moa grouping is favored.

by 40% without biological or geological explanation. Extended persistence of the Norfolk Rise and Lord Howe Ridge has been proposed to permit nonvolant (i.e., nonflying) dispersal between Australia and New Zealand, although these routes may have been submerged by 75 Ma (Cooper and Millener 1993).

A second biogeographic issue concerns the origins of African ratites. Geotectonic reconstructions show that Africa and South America were disconnected by at least 102 Ma (Veevers 2004), whereas fish fossil records indicate that open marine conditions in fact existed between these continents before 110 Ma (Maisey 2000). Explanations involving either vicariance or dispersal via the northern hemisphere both predict that the ostrich would be deep among ratites. Instead, rheas and moa fall to the base of the tree. Some geotectonic reconstructions (e.g., Hay et al. 1999) allow the possibility of the Kerguelen Plateau connecting Australia/Antarctica with Indo-Madagascar until around 80 Ma. This could have provided a staging post for ratite dispersals into Africa and Eurasia, whereas the ancestors of elephant birds remained on Madagascar. The presence of ostrich-like ratites in Early-Mid Tertiary Eurasia and flying palaeognathous birds (lithornithids) in Early Tertiary North America and Eurasia were cited in an alternative proposal for explaining discordance between ratite phylogeny and biogeography, namely that flight has been lost independently among ratites (e.g., Houde 1986).

The most serious challenge yet to ratite monophyly and the associated single origin of flightlessness and Gondwanan vicariance implications has come from the "Early Bird" Tree of Life project. In two recent papers (published after this work was first submitted) based on 20 or more nuclear loci, Hackett et al. (2008) and Harshman et al. (2008) find tinamous to be nested within ratites and the ostrich to be "basal" among extant palaeognaths, thus favoring ratite paraphyly over monophyly. Nevertheless, noncoding sequences that are often difficult to align dominate the "Early Bird" studies, and questions have been raised over the validity of other findings based primarily on the difficulty of getting good alignments of intronic sequences for deep avian divergences (see Morgan-Richards et al. 2008; Pratt et al. 2009). It is noteworthy here that partitioned maximum likelihood (ML) bootstrap support for tinamous grouping within ratites to the exclusion of the ostrich falls to 62% in Harshman et al. (2008) when only the exonic sequences are included. In light of this, reevaluation of the mt evidence is timely.

Identifying mt signal for tinamous grouping within ratites would lend important confirmation to the nuclear results. Equally, it is critical to understand why previous mitogenomic phylogenies have (apparently) incorrectly supported ratite monophyly, especially given the reliance on mt data for most ancient DNA studies and burgeoning mt genome availability from shotgun sequencing projects (e.g., Gilbert et al. 2007). We are also able to address a number of other questions the Early Bird studies leave unanswered: the sister group of the tinamous within ratites, the placement of the extinct moa and elephant birds, the timescale of palaeognath evolution, and statistical support for alternative scenarios for both dispersal and loss of flight among palaeognaths.

Previous analyses of palaeognath phylogeny based on molecular sequences have revealed a high rate of evolution among the tinamous relative to the ratites (e.g., Sibley and Ahlquist 1990; Paton et al. 2002). Hence, in the absence of other "long-branch" taxa among the palaeognaths, there is an expectation that the tinamous will tend to be attracted toward the deeper outgroup taxa (see Hendy and Penny 1989), so artifactually reinforcing ratite monophyly. Long-branch attraction artifacts depend on "unobserved" substitutions not being sufficiently accounted for and thus, the amount of parallel change being underestimated (see Felsenstein 1978). Modeling variation in substitution rates across sites (RAS) is critical in correcting for such unobserved substitution.

Ratite monophyly has been enforced in all previous molecular phylogenetic analyses that modeled rates-across-sites heterogeneity and included both tinamous and moa (e.g., Cooper et al. 2001; Paton et al. 2002; Pereira and Baker 2006). In order to reduce the influence of branch-length biases relative to earlier studies, we incorporated among-site rate heterogeneous models across separately modelled protein codon and RNA structural data partitions. Further to this end we have increased outgroup taxon sampling relative to previous studies and report newly completed mt genome sequences for 2 species of kiwi.

Our analyses strongly support moa grouping with tinamous and these together most likely being sister to a group that includes kiwi, cassowaries, emu, and elephant birds. The implication that the ostrich is the sister to all other extant palaeognaths is further supported by base frequency (BF) distance trees that were employed to examine the topological nature of base composition nonstationarity. We use relaxed molecular clock

methods to provide a temporal scale for this revised palaeognath phylogeny and consider its implications for independent origins of flightlessness and, in turn, for biogeography.

MATERIAL AND METHODS

Polymerase Chain Reaction and Sequencing

Two kiwi species were sequenced as part of this study. The brown kiwi (*Apteryx australis mantelli*) sequence is from the same specimen (K86) and DNA extraction as used in Cooper et al. (2001) and completes the mt genome reported therein. The little spotted kiwi (*Apteryx owenii*) sample was from the Otorohanga Kiwi House. Whole mitochondria were isolated from blood using red blood cell isolation followed by cell disruption, differential centrifugation, and DNase I digestion. The method is based on that of Higuchi and Linn (1995). Polymerase chain reaction (PCR) was used to confirm no nuclear DNA remained in the mt extraction.

Brown kiwi DNA was amplified in one 12-kb long-range product spanning the incomplete region from Cyt *b* to NADH1 using the Expand Long Template PCR System (Roche, Auckland, New Zealand). Both the long-range PCR product (brown kiwi) and mtDNA extraction (little spotted kiwi) were used as templates for subsequent PCR of short 1- to 2-kb overlapping fragments, using primers described in the supplementary appendix (available from http:/www.sysbio.oxforjounals.org/). This process is described in more detail in Gibb et al. (2007) and references therein. The complete mt genome of the little spotted kiwi is 17,020 bp long (GU071052) and the brown kiwi is 17,058 bp long (GU071057, AY016010). Both kiwi have the standard gene order found in all ratites.

Data Matrices

The data set includes complete mt genome protein, ribosomal RNA and transfer RNA coding sequences, totalling 14,190 nucleotides (upon exclusion of sequences with ambiguous homology, after alignment in Se-Al 2.0a9; Rambaut 1996). In addition to the new kiwi sequences, 12 other palaeognathous birds, 8 neognathous birds, and 2 crocodilians were sampled in the present study. These included great-spotted kiwi (*Apteryx haasti*, NC_002782), cassowary (*Casuarius casuarius*, NC_002778), emu (*Dromaius novaehollandiae*, NC_002784), giant moa (*Dinornis giganteus*, NC_002672), eastern moa (*Emeus crassus*, NC_002673), little bush moa (*Anomalopteryx didiformis*, NC_002779), greater rhea (*Rhea americana*, AF090339), lesser rhea (*Pterocnemia pennata*, NC_002783), ostrich (*Struthio camelus*, NC_002785), giant tinamou (*Tinamus major*, NC_002781), elegant crested tinamou (*Eudromia elegans*, NC_002772), tataupa tinamou (*Crypturellus tataupa*, AY016012), chicken (*Gallus gallus*, NC_001323), brush turkey (*Alectura lathami*, NC_007227), magpiegoose (*Anseranas semipalmata*, NC_005933), redhead duck (*Aythya Americana*, NC_000877), blackish oystercatcher (*Haematopus ater*, NC_003713), ruddy turnstone (*Arenaria interpres*, NC_003712), little blue penguin (*Eudyptula minor*, NC_004538), red-throated loon (*Gavia stellata*, NC_007007), American alligator (*Alligator mississippiensis*, AF069428), and caiman (*Caiman crocodilus*, NC_002744). Data sets and phylogentic trees are available from TreeBASE (SN4712).

Neognath birds and crocodilians are the closest living outgroups to palaeognaths. Furthermore, base frequency (BF) distances per variable site ($|A_i - A_j| + |C_i - C_j| + |G_i - G_j| + |T_i - T_j|$) from the palaeognaths for the combined protein-coding and RNA-coding alignment are smaller for the neognaths (average 0.088) and crocodilians (average 0.063) than for the non-archosaurs (average 0.136), green turtle, eastern painted turtle, blue-tailed mole skink, and iguana that were included in the studies of Harrison et al. (2004) and Slack et al. (2006).

Numerous studies have revealed compositional heterogeneity to be of particular concern for mitogenomic phylogenetics (e.g., Delsuc et al. 2003; Gibson et al. 2005). A recent examination of compositional heterogeneity among birds (Harrison et al. 2004) includes several relevant findings: 1) The influence of compositional heterogeneity on phylogenetic reconstruction was exacerbated among data partitions for which saturation has greatly eroded phylogenetic signal (e.g., third codon positions); 2) Compositional χ^2 tests are poor indicators of potential for phylogenetic bias and they are not comparable between data sets because their statistical power depends on factors such as the number of variable sites; and 3) Coding nucleotide data as purines and pyrimidies (RY-coding) was more efficient (in terms of phylogenetic signal retention) than using the amino acid sequence for excluding sources of compositional heterogeneity.

We follow the recommendation of Harrison et al. (2004) to RY-code protein third codon positions. Two metrics described in that paper and originally in Phillips et al. (2001) provide further justification for this RY-coding. First, the stemminess (the proportion of internal branch length contributing to tree length) of minimum evolution trees on *p*-distances inferred from the present data set third codon positions increases from 0.170 to 0.230 upon RY-coding. Simultaneously, relative composition variability (RCV) among third positions falls from 0.093 to 0.054. RCV is the average variability in composition between taxa; for nucleotides this is

$$RCV=\sum_{i=1}^{n}(|A_i-A^*|+|T_i-T^*|+|C_i-C^*|+|G_i-G^*|)/nt,$$

A_i, T_i, C_i, and G_i are the frequencies of each nucleotide for the *i*th taxon; A*, T*, C*, and G* are averages across the *t* taxa; and *n* is the number of sites. Uninformative sites effectively dilute apparent nonstationarity so were excluded (along with gapped sites) from RCV calculations.

Lower stemminess indicates greater phylogenetic signal erosion, and as noted in Phillips and Pratt (2008),

this compounds the potential for higher composition variability to mislead phylogenetic reconstruction. The improvements in both RCV and stemminess upon RY-coding the third positions provide considerable encouragement. Nevertheless, the possible influence of remnant compositional heterogeneity on phylogenetic reconstruction is also examined using BF distance trees (see below).

Phylogenetic Analyses

Palaeognath phylogeny was inferred from the mt genome data set as a single concatenation or partitioned by structure (stem and loop sites) for the RNA-coding data and by codon positions for the protein-coding data. Under the Akaike information criterion (AIC), this partitioning scheme was preferred over concatenation and alternative gene or gene-by-codon partitioning schemes (see Table 1). Bayes factor analyses within Tracer v1.4 (Rambaut and Drummond 2007) on Bayesian inference (BI) analyses (see below for details) performed in MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001) further support the conclusions based on the AIC results.

Given concerns for the influence of the long tinamou branches, phylogenetic analyses were repeated without the tinamous and the relationships among the ratites examined for consistency. Substitution model categories for each partition were assigned according to ModelTest 3.6 (Posada and Crandall 1998) AIC recommendations. In each case, these were GTR+I+Γ_4 for standard nucleotide partitions and F81+I+Γ_4 (equivalent to CF87+I+Γ_4; Cavender and Felsenstein 1987) for the RY-coded mt third codon partitions.

BI (MrBayes 3.1.2) analyses were run with the full substitution model and branch-length rate multipliers unlinked among the protein codon and RNA structural partitions. Within this framework, the GTR+I+Γ_4 version of the doublet model was employed for RNA stem pairs. Three Markov chain Monte Carlo (MCMC) chains for 2 independent runs proceeded for 3,000,000 generations with trees being sampled every 2000 generations. The burn-in for each MrBayes run (250,000) ensured that $-\ln L$ had plateaued, clade frequencies had converged between runs, and estimated sample sizes for substitution parameter estimates were above 200 (using Tracer v1.4).

ML analyses were performed within PAUP*4.0b10. ML bootstrapping (500 replicates) applied heuristic searches to random starting trees for the data sets as single concatenations. Following Phillips and Penny (2003), the TN93 (Tamura and Nei 1993) substitution model was applied to these concatenations such that the transversions in the standard nucleotide and RY-coded data are effectively weighted equally. As an alternative to optimizing substitution parameters on a Neighbor-Joining distance tree (as per ModelTest), these were optimized on the maximum parsimony (MP) tree, employed in an ML heuristic search and reoptimized on the resulting tree for use in the bootstrap analysis. In order to ensure computational feasibility for these bootstrap analyses, clades that are uncontroversial in all recent molecular and morphological classifications and also received posterior probabilities of 1.00 in the Bayesian analysis were constrained. These include Alligatorinae, Aves, Galloanserae, Galliformes, Anseriformes, Charadriiformes, Rheidae, Apterygidae, Casuariidae, and Dinornithidae.

Support among alternative topologies was further examined with KH (Kishino and Hasegawa 1989) and approximately unbiased (AU; Shimodaira 2002) tests, using the RELL method (100,000 replications) within CONSEL (Shimodaira and Hasegawa 2001). The AU test is related to the SH test (Shimodaira and Hasegawa 1999) and has been developed in order to overcome tree selection biases that affect the latter test when multiple topologies are being simultaneously compared.

The ML significance tests were applied with the sequences treated as separately modeled process partitions among codon positions (1, 2, 3) and RNA structures (stems, loops). Partitioning the data allows for more accurate models of sequence evolution that address differential influences on mutation and selection across the sequence (e.g., Yang 1996; Caterino et al. 2001; Buckley et al. 2002). Substitution model categories again followed the ModelTest AIC recommendations for both the standard nucleotide-coded partitions (GTR+I+Γ_4) and the RY-coded mt third codon partitions (CF87+I+Γ_4). All substitution parameters and branch lengths were ML optimized for each partition, for each tree hypothesis.

TABLE 1. Evaluation of partitioning using the AIC

Partitioning scheme[a]	df[b]	$-\ln L$	AIC	Bayes factor[c]
1a. All data (concat)	55	−95, 137.51	190, 385.02	382.7
1b. Ptn + RNA (par)[d]	110	−94, 232.91	188, 685.82	
2a. RNA (concat)	55	−18, 559.93	37, 229.86	116.3
2b. Stems + loops (par)[d]	110	−18, 261.55	36, 743.10	
3a. Ptn (concat)	55	−75, 672.98	151, 455.96	3005.2
3b. Ptn (par-gene)	715	−74, 564.38	150, 558.76	2574.5
3c. Ptn (par-codon)[d]	158	−72, 672.06	145, 660.12	
3d. Ptn (par-gene-by-codon)	2054	−70, 817.28	145, 742.56	

[a]Data are partitioned into (1) protein (Ptn) and RNA, (2) stems and loops for the RNA alone, and (3) individual genes, codon positions, or codon positions for each gene, for protein data alone.

[b]ML analysis (in PAUP* 4.0b10) of the 24-taxon data set provides 55 degrees of freedom (df) for each GTR+I+Γ modeled partition and 48 df for each CF87+I+Γ modeled RY-coded partition (third codon positions).

[c]ln Bayes factor differentials from the favored hypothesis except for the extremely parameter-rich 3D scheme, for which independent MrBayes runs of 100,000,000 generations showed no indications of converging or reaching stationarity.

[d]The favored scheme in each evaluation.

BF Distance Trees

Minimum evolution BF distance trees were constructed in PAUP* from matrices of pairwise BF

distances to assess the potential for compositional bias to affect phylogenetic inference. The basic idea is to compare BF distance tree-length differences between alternative topologies. In this way, Phillips and Penny (2003) showed that composition bias likely explains incorrect rooting of the mammalian tree in earlier mt genome studies. BF distances are half the sum of absolute frequency differences between taxon pairs for each nucleotide category. So the pairwise BF distance between taxa i and j is

$$\text{BF distance} = (|Ai - Aj| + |Ti - Tj| + |Ci - Cj| + |Gi - Gj|)/2,$$

Ai, Ti, Ci, and Gi and Aj, Tj, Cj, and Gj are the frequencies of each nucleotide for the ith and jth taxa, respectively. Dividing by 2 is necessary for minimum evolution (ME) distances between BF distance trees to be comparable with ME differences on standard (absolute) distance trees. This is because a substitution at a site in taxon i that previously had the same base as for taxon j will result in 1 unit of standard distance but 2 units of BF distance. Parsimony-uninformative characters were excluded from BF distance calculations as these cannot explain ME differences between standard distance trees. For all the ME trees, any negative branch lengths were treated as absolute values for computing tree length.

Molecular Dating

We estimated a timescale for palaeognath evolution using BEAST v.1.4.8 (Drummond and Rambaut 2007) with the 24-taxon data set partitioned as per the phylogenetic analyses. Among molecular dating programs BEAST is unique in incorporating a combination of characteristics that are desirable for analysis of the present data set. These include 1) separate GTR+I+Γ model allocation across the protein codon and RNA structure data partitions, including the equivalent model for the RY-coded third codon positions, 2) soft-bound calibration prior distributions, and 3) relaxation of the molecular clock without assuming rate correlation among branches.

An uncorrelated relaxed clock model was used with rates among branches distributed according to a lognormal distribution, which can provide greater flexibility than the exponential distribution option (Drummond et al. 2006). Note that a strict clock was rejected by a likelihood ratio clock test (using ML in PAUP*) at $P < 0.0001$, and in BEAST the null hypothesis of no rate autocorrelation among branches could not be rejected even at $P \leqslant 0.2$. Five independent runs totaling 90,000,000 MCMC generations ensured estimated sample size values >100 (as estimated in Tracer v1.4; Rambaut and Drummond 2007) for all node height, prior, posterior, $-\ln L$, tree, and substitution parameters. Chains were sampled every 5,000$^{\text{th}}$ generation after burn-ins of 2,000,000 generations.

Difficulties with calibrating molecular dating analyses for the Palaeognathae have not been fully appreciated. Two "internal" ratite calibration dates have commonly been used. One is geotectonic, the divergence between Australia and New Zealand, at ≈82 Ma (Cooper et al. 2001). The present finding of a tinamou–moa grouping undermines the basis of this calibration, which assumes that the divergence between moa and Casuariidae from a flightless ancestor predates the geotectonic divergence. The second internal calibration, the divergence between cassowaries and emus being >25 Ma, is not influenced by the finding of multiple losses of flight among ratites. However, this relies on accepting that the Late Oligocene/Early Miocene casuariiform, *Emuarius*, shares a closer phylogenetic relationship with emus than with cassowaries (Boles 1992). The tibiotarsus (lower leg bone) suggests *Emuarius* affinities with emus. This hypothesis depends partly on the assumption that the most recent common ancestor (MRCA) of Casuariidae was more cassowary-like such that skull and femur similarities between *Emuarius* and cassowaries are symplesiomorphic.

In order to provide temporal calibration, we have employed prior height distributions for 5 nodes external to ratites. We follow Barnett et al. (2005) in our usage of calibration bounds. The minimum marks the first appearance of a generally agreed upon member of the crown group, and the maximum covers the time back until relatively well-sampled fossil assemblages in potential geographic regions of origin contain no putative crown group members but contain stem members or ecological equivalents. Selection of uniform, normal, or lognormal distributions for calibration priors follows Ho and Phillips (2009).

Root (Archosauria): Normal distribution, 95% range from 235 to 250 Ma (Benton and Donoghue 2007).

Aves: Normal distribution, 98% range from 66 to 121 Ma. The normal 98% range acknowledges that both the minimum and the maximum are extremely conservative, and prior expectations for the actual value to be well within the given range are higher than for the other nodes. The minimum is based on *Vegavis* (Clarke et al. 2005) which branches at least 3 nodes internal to the root of modern birds. The maximum is based on the age of the younger of the avian bearing beds within the Jehol biota (Zhou 2006), from which (and before) only Enantiornithes and other "primitive" birds are known. Although the maximum allows for Early Cretaceous birds such as *Gansus* (110 Ma; You et al. 2006) being within the avian crown group, it is far more conservative than the 86 Ma of Benton and Donoghue (2007).

Galloanserae: Uniform distribution, range from 66 to 86 Ma (Benton and Donoghue 2007).

Seabirds (penguin vs. loon): Lognormal distribution, hard minimum 61 Ma (based on the penguin, *Waimanu*; Slack et al. 2006). Mean at 65.5 Ma and 97.5% soft maximum at 74 Ma, respectively, reflect expectations for a K/T boundary radiation after the extinction of numerous avian stem seabirds (Feduccia 1996) and the possibility of seabirds evolving in the Southern Hemisphere during the relative hiatus in that region's late Campanian to late Maastrichtian fossil record.

Alligatorinae: Lognormal distribution, hard minimum (64 Ma), mean expectation (70 Ma), and 97.5% soft maximum (80 Ma) follow Brochu (2004).

Two internal palaeognath calibrations were also included as uniform priors for a second BEAST analysis.

Casuariidae (emu vs. cassowary): Uniform distribution, range from 25 to 35 Ma (see Haddrath and Baker 2001; Phillips 2009).

Rheas versus Casuariidae/kiwi/tinamou–moa: Uniform distribution, range from 56 to 83 Ma. The minimum age is provided by *Diogenornis* (Alvarenga 1983), and the maximum covers the absence of even putative members of this clade in well-sampled Campanian faunas from South America and the Northern Hemisphere.

These internal palaeognath calibrations were not employed in our primary analysis because of the reservations stated above regarding *Emuarius* as a crown cassuariid and also some uncertainty remains over the status of *Diogenornis* as a stem rheiid.

Ancestral State Reconstruction

Histories for dispersal and loss or gain of flight among palaeognaths were inferred under MP and ML criteria in PAUP* and BayesTraits (Pagel and Meade 2006), respectively. The 24 taxa from the mitogenomic data set were coded as flighted or flightless. Character coding for the dispersal analysis was more complicated. First, in order to help reduce geographic sampling biases, regional assignments were considered at the family level, rather than the species level. This did not affect coding for the palaeognaths but resulted in several outgroup taxa being polymorphic. Second, due to landmasses shifting, the rate matrix should not be expected to be homogenous across the tree, particularly given the break-up of South Gondwana (South America, Australia–Antarctica, New Zealand) from about 80 Ma. We take this nonhomogeneity into consideration by using a nested design for our ML dispersal analyses.

For the overall tree, regions were coded as Northern Hemisphere, Africa, and South Gondwana. Within palaeognaths, the most inclusive clade originating after 80 Ma has modern members only on the former South Gondwana landmasses and its MRCA was assigned to South Gondwana in the overall tree ML analysis with $P > 0.99$. Hence, further analyses were performed on the South Gondwana clade (Rheidae, Tinamidae, Dinornithidae, Apterygidae, Casuariidae) assigned to South America, Australasia, and New Zealand. This nested design also has the benefit of reducing parameterization relative to having a single analysis with 5 regional character states. State transitions were symmetrically reversible in the dispersal analyses, in accord with likelihood ratio tests rejecting the asymmetrical alternatives at $P < 0.05$.

In considering heterogeneity in dispersal probabilities over time, our analyses have a partial analog with the more sophisticated dispersal–extinction–cladogenesis (DEC; e.g., Ree and Smith 2008). However, our method has much reduced parameterization (important for smaller data sets). Second, BayesTraits assumes that ancestral individuals are localized to one "area state," as befitting ratite mt evolution over timescales of nearly 100 Ma. In contrast, DEC is specialized for population biogeography and allows ancestors to be distributed over 2 or more "area states" for substantial proportions of lineage time even when organisms at all the tips are localized.

The BayesTraits analyses used the multistate ML option with 2 and 3 state categories, respectively, for flight gain/loss and regions, with characters evolving along the BEAST (median node height) dated tree. This design gives us the options of including extinct taxa and artificially altering sampling strategies, while maintaining comparability among these treatments (i.e., negating differential influences from priors and integration across alternative phylogenetic trees).

RESULTS

Preliminary Phylogenetic Analyses

As noted in Introduction, the mt consensus on ratite phylogeny (Fig. 1a) is based on analyses for which either ratite monophyly has been enforced or rates-across-sites variation has not been allowed for. Cooper et al. (2001) was the first of the complete mt genome studies to include the extinct moa. We reemployed their data set and ML methodology (see Appendix) and only by enforcing ratite monophyly were we able to replicate their tree, in which rheas are sister to all other ratites (consistent with Fig. 1a). Relaxing that monophyly enforcement results in the tinamous shifting from a basal position among palaeognaths to be sister to moa (Fig. 1b). This alternative tree is 15.365 $-\ln L$ units better than the tree for which ratite monophyly is enforced, which in a pairwise comparison is rejected by KH and SH tests at $P = 0.136$.

With our own data set, both ingroup (Palaeognathae) and outgroup (Neognathae and Crocodilia) taxon sampling have been increased; RY-coding for protein third codon positions introduced; and substitution models are partitioned across protein codon positions and RNA stems and loops. Each of these changes is expected to reduce the influence of biases on phylogeny reconstruction. Our BI and ML bootstrap results reveal 2 differences from the mt consensus. The first is that the tinamou–moa grouping that was found in our reanalysis of the data set of Cooper et al. (2001) is now very strongly supported (99% for ML and 1.00 for BI, see Fig. 2). The second difference involves the placement of the ostrich. In Figure 1, the ostrich is shown to group with emu, cassowaries, and kiwi, regardless of ratite monophyly being enforced. In our new analyses, Figure 2 shows the ostrich diverging from the basal node among palaeognaths.

Higher substitution rates among the tinamous (see Fig. 3) than among the ratites should in theory be well accounted for by the partitioned rates-across-sites models. However, in order to check for consistency, analyses

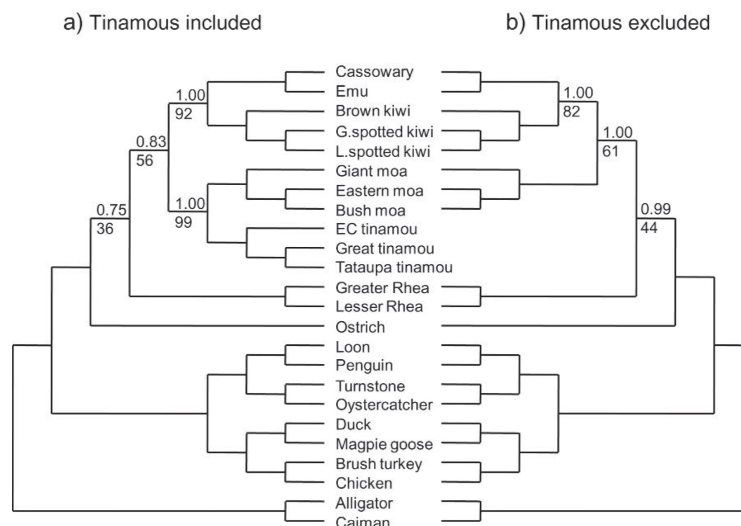

FIGURE 2. Avian phylogenies inferred from the complete mt protein- (third codons RY-coded) and RNA-coding DNA sequences (a) with and (b) without the inclusion of the tinamous. Support values are shown for Bayesian posterior probability (top) and ML nonparametric bootstrap (bottom). Values are not shown for nodes that receive maximum support in both analyses.

were also run without tinamous and it is encouraging that the favored tree (Fig. 2b) is congruent with the tree that includes tinamous (Fig. 2a). In fact, the ML bootstrap and BI support for all other palaeognaths grouping to the exclusion of the ostrich strengthens from 36% to 44% and 0.75 to 0.99, respectively.

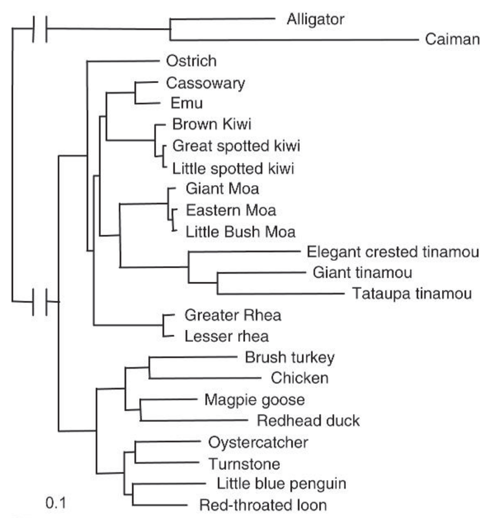

FIGURE 3. BI phylogram from the complete mt protein- (third codons RY-coded) and RNA-coding DNA sequences. The substitution model is partitioned across protein codons and RNA stems and loops, whereas rate multipliers were employed for branch-length estimation. Scale is substitutions per site.

ML Hypothesis Testing

Phylogenetic hypothesis testing using ML with partitions modeled separately circumvents both the concatenation problem associated with the ML bootstrapping in PAUP* and the restriction to parametric error estimation in MrBayes. Shimodaira and Hasegawa (1999) showed that such AU and KH hypothesis tests can more closely reflect sampling error than do bootstrap values, which in turn are far more faithful than typically overconfident Bayesian posterior probability (BPP) values (Suzuki et al. 2002; Gontcharov et al. 2004).

Among tinamous, moa, ostrich, rheas, kiwi, and Casuariidae, there are 945 possible rooted topologies. The combined ML scores for the separately modeled partitions and associated AU (and KH) P values are shown in Table 2(a) for the best 5 trees and several other trees of interest. The favored tree is the same as was found in the Bayesian and ML bootstrap analyses (Fig. 2) and groups moa with tinamous. The best tree containing the traditional split between tinamous and ratites (Tree 6) is consistent with Haddrath and Baker (2001), however, is over 37 $-\ln L$ units adrift from the overall best tree and is rejected at $P < 0.05$. Additionally, the Cooper et al. (2001) tree (7) is rejected at $P < 0.01$, as is a sister grouping of the 2 New Zealand taxa, moa and kiwi (Tree 8).

Beyond the sister groupings of tinamous and moa and of emus and cassowaries (Casuariidae), the relationships of the palaeognaths remain difficult to resolve. Although Table 2(a) shows the ostrich to be favored as the "basal-most" member, the second through fourth best trees include alternative basal-most taxa, tinamou–moa (2), rheas (3), and ostrich–rheas (4). None

TABLE 2. Log-likelihood differences between trees and their statistical significance under AU and KH tests (a) with tinamous included and (b) without tinamous

Alternative groupings[a]	$-\ln L^{b}$	P values	
		AU	KH
a) Among Palaeognathae[c]			
1. (Out, (Ost, (Rhe, ((Moa, Tin), (Cass, Kiwi)))))	< 90, 905.8 >	—	—
2. (Out, ((Moa, Tin), (Rhe, (Ost, (Cass, Kiwi)))))	+1.9	0.597	0.448
3. (Out, (Rhe, (Ost, ((Moa, Tin), (Cass, Kiwi)))))	+4.2	0.462	0.307
4. (Out, ((Rhe, Ost), ((Moa, Tin), (Cass, Kiwi))))	+7.5	0.177	0.175
5. (Out, (Ost, (Rhe, (Cass, ((Moa, Tin), Kiwi)))))	+7.9	0.237	0.167
6. (Out, (Tin, (Moa, (Rhe, (Ost, (Cass, Kiwi))))))	+37.1	0.011*	0.047*
7. (Out, (Tin, (Rhe, (Moa, (Ost, (Cass, Kiwi))))))	+56.0	0.004*	0.008*
8. (Out, (Ost, (Rhe, (Cass, (Tin, (Moa, Kiwi))))))	+67.8	<0.001*	<0.001*
b) Among ratites only—tinamous excluded			
1. (Out,(Ost, (Rhe, (Moa, (Cass, Kiwi)))))	< 78, 982.2 >	—	—
2. (Out, (Ost, (Rhe, (Cass, (Moa, Kiwi)))))	+5.8	0.345	0.208
3. (Out, (Ost, (Rhe, (Kiwi, (Moa, Cass)))))	+7.0	0.231	0.158
4. (Out, (Rhe, (Ost, (Moa, (Cass, Kiwi)))))	+7.4	0.369	0.236
5. (Out, (Moa, (Rhe, (Ost, (Cass, Kiwi)))))	+10.4	0.365	0.282

Note: *significant at $P \leq 0.05$.
[a]The best 5 trees among the 945 and 105 possible trees for (a) and (b) are shown, respectively. Additionally for (a), the best tree in which ratites are monophyletic is Tree 6, which is congruent with Haddrath and Baker (2001). Tree 7 is congruent with Cooper et al. (2001), and in Tree 8 the New Zealand ratites are monophyletic.
[b]ML models are partitioned for proteins (by codon) and for RNA (stems, loops).
[c]Taxon abbreviations: Out, outgroup neognaths and crocodilians; Ost, ostrich; Rhe, rheas; Moa, moa; Tin, tinamous; Cass, Casuariidae.

of these alternative trees can be rejected at $P < 0.175$ for either the KH or the AU tests. Indeed, even a tree (5) in which tinamou–moa group with kiwi is rejected only at $P = 0.167$ (KH) and $P = 0.237$ (AU). In Table 2(b), the best 5 trees are shown for equivalent analyses but with the tinamous excluded. The results are consistent with those in Table 2(a) such that the ostrich diverging first among the ratites is not an artifact of the inclusion of the tinamous. Furthermore, these 5 trees in Table 2(b) are sufficient to show that each grouping within the favored ratite tree (ostrich, (rheas, (moa, (Casuariidae, kiwi)))) has an alternative that cannot be rejected at $P \leqslant 0.15$ for the KH test or at $P \leqslant 0.20$ for the AU test.

Cooper et al. (2001) sequenced a portion of the elephant bird (*Mullerornis agilis*) mt genome and found its affinities to lie with kiwi and Casuariidae effectively as an unresolved trichotomy. We examined whether the novel phylogenetic context (tinamou–moa and ostrich basal among palaeognaths) and RY-coded protein third codon positions would affect the placement of elephant birds. The available elephant bird sequences align against only 880 sites from our primary data set. These sites were analyzed alone and under a single GTR+I+Γ model in accord with concerns for overparameterization. ML scores were calculated in PAUP* for each possible placement for the elephant bird on the palaeognath tree that is shown in Figure 2a. The best placement was as sister to kiwi ($-\ln L = 6007.35$). Alternative placements as sister to either Casuariidae, kiwi + Casuariidae, tinamous + moa, or kiwi + Casuariidae + tinamous + moa could not be rejected at $P \leqslant 0.25$. Only elephant bird placements with the ostrich or as sister to all other palaeognaths could be rejected at $P \leqslant 0.10$.

Exploration for Nonphylogenetic Biases

Without RY-coding the protein third positions, BF differences divide ratites into 2 groups, kiwi, emu, cassowary, and ostrich with low-cytosine/thymine relative frequencies (average 1.59) and rheas and moa with high-cytosine/thymine relative frequencies (average 2.79). The non-palaeognath archosaurs also have high-cytosine/thymine relative frequencies (average 3.15) and so for the BF distance trees, kiwi, emu, cassowary, and ostrich group together strongly (Fig. 4). The tinamous have relatively low-cytosine/thymine relative frequencies (average 1.79). With their inclusion, the same relationships among ratites are recovered as in Figure 4, although tinamous are placed as sister to kiwi.

With our primary (RY-coded) data set, BF distances favor tinamous and rheas grouping on one side of the palaeognath root, whereas on the other side ostrich groups with Casuariidae and kiwi, then moa diverge progressively deeper from these. The groupings recovered in these optimal BF distance trees are only phylogenetically relevant when compared between tree topologies that are potential candidates for representing evolutionary history. Accordingly, Table 3 shows the ME differences (based on BF distances) between the alternative tree candidates from the partitioned ML analyses in Table 2(a). These comparisons reveal that similarity in base composition clearly favors ratite monophyly (Trees 6 and 7) for both the standard and the RY-coded nucleotide data. Importantly though, the RY-coding greatly reduces the potential for compositional heterogeneity to bias phylogenetic reconstruction. For example, the optimal tree from our phylogenetic analyses (Tree 1) is disadvantaged in terms of compositional heterogeneity relative to the favored BF distance tree (Tree 7) by 128.4 changes before RY-coding but only by 24.4 changes after RY-coding. As such, these results are consistent with compositional heterogeneity contributing toward "apparent" phylogenetic signal for ratite monophyly and more so in studies that have not applied RY-coding to protein third positions. Notably, the favored BF distance trees 7 and 6 are consistent with the favored trees from Cooper et al. (2001) and Haddrath and Baker (2001), respectively.

High rates of substitution among the tinamous and outgroup taxa relative to the ratites should provide

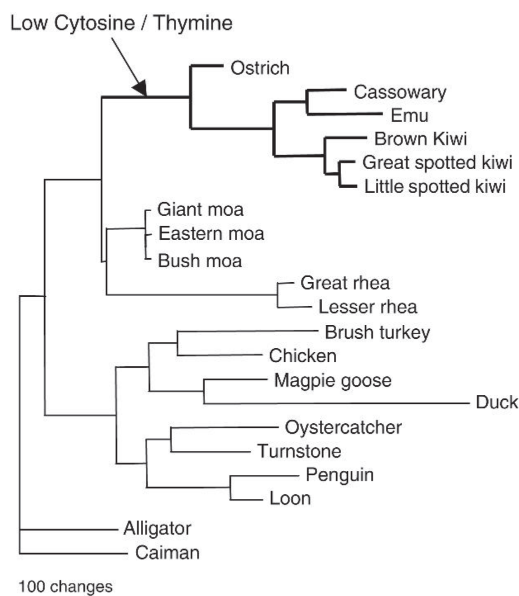

FIGURE 4. Minimum evolution tree on BF distances from the complete mt protein- and RNA-coding DNA sequences. Compositional bias favors ostrich grouping with kiwi and Casuariidae. Average cytosine/thymine (C/T) among these taxa is 1.59. Average C/T among the moa and rheas is 2.79, close to the average among the outgroup taxa (3.15).

artifactual "long-branch-attraction" signal for ratite monophyly if hidden substitutions are undercorrected for. This is indeed the case; for both the standard and the RY-coded data sets, minimum evolution on uncorrected

TABLE 3. BF distance minimum evolution differences between the tree topologies that were compared in Table 2(a) for ML scores

Alternative palaeognath trees[a]	a. Ptn123rna (all nucleotide)[b]	b. Ptn123rna (3RY)[c]
1. (Out, (Ost, (Rhe, ((Moa, Tin), (Cass, Kiwi)))))	+128.4	+24.4
2. (Out, ((Moa, Tin), (Rhe, (Ost, (Cass, Kiwi)))))	+85.7	+21.7
3. (Out, (Rhe, (Ost, ((Moa, Tin), (Cass, Kiwi)))))	+93.1	+3.0
4. (Out, ((Rhe, Ost), ((Moa, Tin), (Cass, Kiwi))))	+122.9	+66.8
5. (Out, (Ost, (Rhe, (Cass, ((Moa, Tin), Kiwi)))))	+174.0	+36.7
6. (Out, (Tin, (Moa, (Rhe, (Ost, (Cass, Kiwi))))))	+65.3	+2.0
7. (Out, (Tin, (Rhe, (Moa, (Ost, (Cass, Kiwi))))))	$< 3008.3 >$	$< 1175.9 >$
8. (Out, (Ost, (Rhe, (Cass, (Tin, (Moa, Kiwi))))))	+207.8	+29.1

[a]The first 5 trees are the best ML trees. Trees 6 and 7 are the best supported by BF distances and are consistent with ratite monophyly. Tree 8 groups the New Zealand ratites (moa and kiwi). See Table 2 for taxon abbreviations.
[b]Standard nucleotide coding.
[c]Protein third codon positions RY-coded.

TABLE 4. The influence of the gamma shape parameter (α) on ML support for the placement of tinamous either as in the favored (ratite paraphyly) tree in Table 2 (Tree 1) or as in the best tree with ratite monophyly constrained (Tree 6)

Gamma shape parameter (α)	ln*L* units favoring ratite paraphyly over monophyly[a]	KH *P* value[b]
Infinity	32.1	<0.001
2.000	37.1	<0.001
1.000	38.4	<0.001
0.750	38.7	<0.001
0.500	39.1	<0.001
0.358[c]	40.0	<0.001
0.300	41.0	<0.001
0.250	42.6	<0.001
0.200	44.5	<0.001

[a]The ML model used is identical to that used for the ML bootstrap analyses, except that α is varied.
[b]For pairwise comparisons, the KH test is equivalent to the SH test.
[c]The ML value for α was 0.358.

(*p*) distances drew tinamous toward the outgroup, leaving ratite monophyly (not shown). Similarly, rate heterogeneity that is unaccounted for in our models for the ML and Bayesian analyses might lead to the present support for a tinamou–moa grouping in fact being underestimated relative to ratite monophyly. Conversely, if our I+Γ models overestimate hidden substitutions (and gamma models do not always provide a good fit, see Susko et al. 2003), then long-branch repulsion (see Siddall 1998) may be providing artifactual signal for grouping the tinamous among the far-slower evolving ratites.

The ML model optimized for the bootstrap analysis on the concatenated data (with 3RY) had a gamma shape parameter of 0.358. Under this model, the tinamou–moa grouping is favored over ratite monophyly by 40.0 $-\ln L$ units. By varying only the gamma shape parameter, we were able to get an appreciation for the influence of lesser or greater correction for hidden substitutions, which should encourage long-branch attraction and repulsion, respectively. Hence, Table 4 shows that as expected, the support in favor of tinamous grouping with moa increases upon lowering the shape parameter and decreases upon increasing the shape parameter. The results suggest that it is possible that our primary analyses slightly underestimate or overestimate the support for tinamou–moa. Importantly though, even with the shape parameter (miss)specified at infinity, ratite monophyly is still rejected at $P < 0.001$.

Molecular Dating and Ancestral State Reconstruction

The median node heights from the BEAST molecular dating analysis provide the scale for the chronogram in Figure 5. Note that only the divergence of the ostrich at the basal palaeognath node occurs before the break-up of South Gondwana involved landmass separation with New Zealand rifting from Antarctica and starting to "unzip" from Australia. In terms of the terrestrial dispersal history of modern faunas, this division of South Gondwana is essentially complete with the isolation of

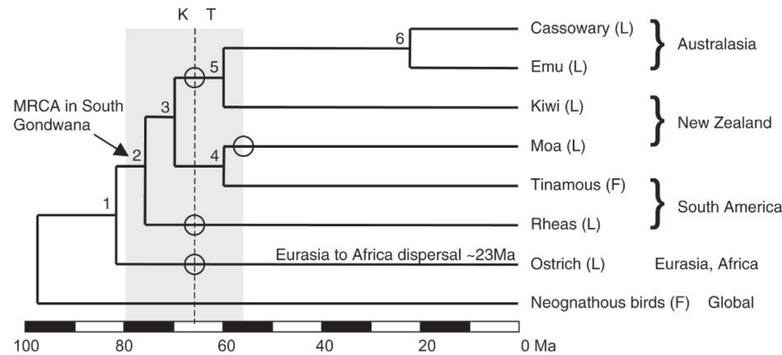

FIGURE 5. Timescale of palaeognath evolution with relative placements indicated for inferred losses of flight (circles), the Cretaceous-Tertiary boundary (dotted line), and the major period of fragmentation of South Gondwana (shaded). Inferred median dates from the BEAST analysis of the 24-taxon tree for palaeognath divergences at the numbered nodes in millions of years before present: (1) 83.4, (2) 76.4, (3) 70.6, (4) 60.0, (5) 59.9, (6) 20.7. 95% HPDs are shown in Table 5. Elephant birds of Madagascar have not been included due to the uncertainty of their phylogenetic placement. F and L in parentheses after taxon names denote flighted and flightless, respectively.

Australia and South America by about 55 Ma (Woodburne and Case 1996). Median estimates and 95% HPDs for palaeognath divergences are shown in Table 5(a). The dates noted below and which were subsequently used for the ancestral character state and dispersal analyses employed only the 5 calibration ranges that are external to Palaeognathae. The results with the additional inclusion of 2 preliminary internal calibrations (emu vs. cassowary and rhea vs. kiwi/Casuariidae/tinamou–moa) are also shown in Table 5(b).

After the Ostrich diverges at 83 Ma, the origin of the South Gondwanan clade of palaeognaths is dated at 76 Ma. Two of the deep divergences among palaeognaths, kiwi from Casuraiidae and tinamous from moa, are estimated to be post-K/T, at 60 Ma. The analyses with the internal palaeognath calibrations provide slightly younger dates still, which encouragingly, have far narrower 95% HPDs.

Dispersal analyses were conducted for both the overall rooted 24-taxon tree (O_{24}) tip-labeled as Northern Hemisphere, Africa, and South Gondwana and the rooted 13-taxon South Gondwanan clade (SG_{13})

TABLE 5. Palaeognath divergence estimates in Ma from BEAST analyses

Node[c]	(a) external calibrations[a]		(b) external and internal calibrations[b]	
	Median	95% HPD	Median	95% HPD
1	83.4	55.2–103.6	78.1	60.4–94.0
2	76.4	50.3–97.8	70.0	57.0–82.4
3	70.6	46.6–92.5	63.8	50.4–79.1
4	60.0	38.3–81.6	53.2	39.1–71.5
5	59.9	36.5–82.7	53.5	36.9–72.1
6	20.7	7.6–50.9	28.6	25.0–34.3

[a]With only 5 calibration ranges for nodes external to palaeognaths.
[b]With the addition of 2 calibration ranges within Palaeognathae.
[c]Node numbers refer to Figure 5.

tip-labeled as South America, Australasia, and New Zealand. In order to examine how robust the results were to incorporating information from the fossil record, 3 key fossil palaeognaths were included. These were the 42 Ma European *Paleotis* as sister to ostrich (Houde 1986), the ≈56 Ma Brazilian *Diogenornis* as sister to Rhea (Alvarenga 1983), and the ≈70 Ma oldest known lithornithid (Parris and Hope 2002) from North America. Sampling in the fossil record is too sparse to accurately infer how long it was because each of these fossils diverged from their sister taxa. For the chronogram used in the BayesTraits analyses, we nominally placed the divergences of these taxa from their sister taxon 10 Ma prior to the fossil age (of *Paleotis* and *Diogenornis*) or 10 Ma prior to the crown palaeognath root (for the lithornithid). Such timing allows the fossil taxa to inform the ML analyses without constraining their common ancestor with the modern taxa to share the same state. With the fossil taxa included, the overall and South Gondwanan dispersal analyses are referred to as O_{27} and SG_{14}. Note that among the fossil taxa only *Diogenornis* falls within the SG clade.

The results of the ML dispersal analyses are shown in Table 6. In both the O_{24} and O_{27} analyses, a South Gondwanan origin is preferred for the MRCA of modern palaeognaths. However, with the fossil taxa included the probability of a Northern Hemisphere origin is 0.18 and in fact becomes favored under MP (not shown). South Gondwanan origins are very strongly supported for the non-ostrich palaeognaths (Node 2 in Table 6) in all analyses, and more specifically favored to be in South America, particularly for SG_{14}, although New Zealand origins cannot be rejected at $P \leqslant 0.05$. None of the analyses (including MP) clearly distinguish between South American origins for kiwi/Casuariidae/tinamou–moa (Node 3 in Table 6) with independent dispersals of kiwi and moa to New Zealand or, alternatively, an earlier dispersal of Node 3 ancestors from

TABLE 6. ML inference of palaeognath dispersal history

Node[c]	Overall tree O_{27} (O_{24})[a]			South Gondwana clade SG_{14} (SG_{13})[b]		
	Northern Hemisphere	Africa	South Gondwana	South America	Australia	New Zealand
1	0.18 (0.00)	0.00 (0.09)	0.82 (0.90)	—	—	—
2	0.00 (0.00)	0.00 (0.00)	1.00 (0.99)	0.90 (0.55)	0.01 (0.03)	0.09 (0.42)
3	—	—	—	0.34 (0.28)	0.09 (0.09)	0.55 (0.62)
4	—	—	—	0.67 (0.69)	0.02 (0.01)	0.31 (0.30)
5	—	—	—	0.09 (0.05)	0.52 (0.55)	0.38 (0.39)

[a]Probabilities for the overall tree analyses (O_{24}, O_{27}) are provided for MRCAs at Nodes 1 and 2 being in the Northern Hemisphere, Africa, or South Gondwana. In each case, the probabilities are given both from analyses in which information from fossil taxa was included (and excluded).
[b]South Gondwana is divided into South America, Australasia, and New Zealand only for Nodes 2–5 (for SG_{13}, SG_{14}). These analyses are nested within the overall tree in recognition of dispersal probabilities being nonhomogenous across the tree and that dividing South Gondwana is only meaningful after approximately 80 Ma.
[c]Node numbers refer to Figure 5.

South America to New Zealand (or Australasia) and a later return for tinamou ancestors to South America. Support for the origination of the kiwi/Casuariidae clade (Node 5) in either Australasia or New Zealand is similarly difficult to tease apart.

Without obvious a priori reasons to expect the fragmentation of Southern Gondwana to influence the evolution of flight/flightlessness, we return to single analyses with the O_{24} data set. Loss of flight in the ancestors of palaeognaths and regain in the ancestors of tinamous are favored under standard MP. Ancestral character state reconstruction for O_{24} under ML is similar, except that slightly earlier regain of flight in the ancestors of tinamou–moa is favored at $P = 0.78$ (Table 7c). Lithornithids were flying birds (Houde and Olson 1981), whereas the flight status of *Paleotis* and *Diogenornis* is uncertain. Incorporating lithornithids and fixing the archosaur and avian roots to be flightless and flying, respectively (in agreement with essentially all fossil evidence and phylogenetic inference), make little difference to the ML character state reconstruction (Table 7d). A further modification involved artificially biasing the ML analyses in favor of flight loss over gain by assuming that we had instead sampled flightless duck and megapode species. As shown in Table 7, the ML estimates for the expected ratio of loss to gain of flight increases substantially and flightlessness cannot be rejected even at $P \leqslant 0.20$ at any of the palaeognath nodes 1–4. Dollo Parsimony in which flight cannot be regained once it is lost favors flight being lost independently in ostrich, rheas, moa, and kiwi/Casuariidae.

DISCUSSION

Palaeognath Phylogeny

Monophyly of ratites and their sister relationship with tinamous are two of the best agreed upon hypotheses of avian supraordinal phylogeny that derive originally from morphological data (e.g., Bock 1963; Cracraft 1974) and then from nuclear data (Prager et al. 1976; Sibley and Ahlquist 1990; van Tuinen et al. 2000). Analyses of complete mt genome sequences (Cooper et al. 2001; Haddrath and Baker 2001) appeared to offer substantive confirmation of both, as well as a "mt consensus" on relationships among palaeognaths that has been widely adopted (Fig. 1a). However, one of two fundamental problems exist for all previous phylogenetic analyses of palaeognath mt genomes for which tinamous, moa, and at least one other ratite were included. Either rates-across-sites heterogeneity was not accounted for, or when it was (e.g., under a gamma distribution) ratite monophyly was enforced. Upon relaxing this phylogenetic constraint with the data set of Cooper et al. (2001), tinamous and moa group together, rendering ratites paraphyletic.

Our primary analyses incorporate 5 functionally distinct partitions (protein codons 1, 2, 3 and RNA stems and loops) and RY-coded third positions. These retain

TABLE 7. Inference of the evolutionary history of flight (F) and flightlessness (L) among palaeognaths and the probability of being flighted

Node[b]	Ancestral state reconstruction models[a]				
	a. MP	b. MP-Dollo	c. ML (mod)	d. ML (lith)	e. ML (lith/alt)
1. Palaeognathae	L	F	0.02	0.04	0.24
2. Node 2	L	F	0.05	0.08	0.31
3. Node 3	L	F	0.17	0.22	0.46
4. Tinamou–moa	L	F	0.78	0.79	0.81
5. Kiwi/Casuariidae	L	L	0.02	0.04	0.13
Expected P(loss)/P(gain)	1.00	1/0	1.46	1.58	2.34

[a](a) Standard and (b) Dollo MP and the probability of being flighted according to ML analyses in which (c) only modern taxa were included, (d) a lithornithid was added, and in addition (e) a flightless duck and megapode were artificially sampled.
[b]Node numbers refer to Figure 5, such that Node 2 includes all modern palaeognaths except ostrich and Node 3 includes kiwi/Casuariidae/tinamou–moa.

one aspect of the mt consensus with solid, though not irrefutable support, emu and cassowary (Casuariidae) grouping with kiwi. Being able to confidently reject the closest alternative (Table 2a, Tree 4), in which tinamou–moa group with kiwi, will require further examination. Nuclear data may be important here. Harshman et al. (2008) were also unable to confidently resolve the relative affinities of kiwi, rheas, tinamous, and Casuariidae, although encouragingly their analyses favored the same topology we present for mt data (see Fig. 2).

Determining the earliest divergence among the ratites also requires further study. However, our analyses are consistent with the ostrich diverging from the MRCA node of ratites (and indeed palaeognaths). Support for this hypothesis increases further (from 1.9 to 7.4 ln$L$ units) with the exclusion of the tinamou sequences from the analyses (cf. Table 2a,b). The conflicting signal that supports the alternative rooting of ratites in the mt consensus, which places rheas and moa deepest, appears to derive largely from compositional bias. This is evidenced by rhea and tinamou–moa falling to the base of Palaeognathae if RY-coding is not used to reduce compositional bias at third positions in our Bayesian analyses (under both DNA and codon models, not shown).

The BF distance tree in Figure 4 shows that a long branch groups the ostrich with Casuariidae and kiwi. As noted in that figure, these ratite mt genomes have low cytosine relative to thymine when compared with the rheas, moa, and outgroup taxa. With the inclusion of the tinamous and compared among the best topologies from the partitioned ML analyses, the topologies that are best supported by standard (nucleotide) BF distances (Table 3a, Trees 6 and 7) are in fact the topologies supported by Haddrath and Baker (2001) and Cooper et al. (2001). These 2 trees remain the best supported by BF distances on the RY-coded data (Table 3b), although the recoding greatly reduces the magnitude of this artifactual support for placing tinamous outside of ratites and rheas/moa deepest among ratites.

Haddrath and Baker (2001) recognized the problem of compositional heterogeneity for inferring palaeognath phylogeny. However, the potential solutions they employed (LogDet distances [Lockhart et al. 1994] and nonhomogenous ML, NHML [Galtier and Gouy 1998]) did not allow for RAS variation. As noted earlier, accounting for RAS variation is critical for reducing branch-length artifacts. Moreover, ignoring RAS variation leads to the influence of compositional nonstationarity being underestimated (see Phillips et al. 2004). More recent versions of NHML incorporate RAS, although still assume that compositional heterogeneity conforms to the AT/GC pattern that is common for nuclear DNA and that among the BFs, A = T and G = C. Instead, the ML optimized "equilibrium" frequencies among the 2 purines and 2 pyrimidines differ considerably and compositional heterogeneity among the avian mt sequences primarily concerns variation among pyrimidines (C,T) specifically.

Our analyses favor the ostrich diverging from the base of Palaeognathae, in line with some previous mt studies that did not include moa (e.g., van Tuinen et al. 1998; Harrison et al. 2004). Importantly, compositional nonstationarity provides a "smoking gun" for explaining the deeper placement of moa and rheas in most previous studies. Hence, on both mt and nuclear data (e.g., Harshman et al. 2008) an early ostrich divergence must now be considered the best phylogenetic estimate on which to base evolutionary and biogeographical inferences. Another result, the tinamou–moa grouping is more salient; all ML bootstrap and BPP support values are $\geqslant$99% (Fig. 2) and partitioned ML hypothesis testing rejects the best tree with ratite monophyly (Table 2a, Tree 6) at $P < 0.05$.

It is encouraging that with the tinamou sequences excluded, moa maintain a concordant placement in the tree, as sister to Casuariidae and kiwi. Curiously, with moa excluded, tinamous tend to fall to a basal placement among palaeognaths (not shown, but as for Harrison et al. 2004). This situation is reminiscent of earlier efforts to root the tree of placental mammals (e.g., Kretteck et al. 1995). Analogous to tinamous, hedgehog sequences evolved at very high rates and were artifactually attracted to the base of Placentalia. It was not until the hedgehogs could be bound to a sufficiently close (and slower evolving relative), a shrew, that hedgehog affinities were shown to be nested well within Placentalia (Lin et al. 2002). One might speculate that for both the hedgehogs and the tinamous, the branch-length artifacts are partly associated with heterotachy, for which standard stationary models can be inconsistent (see Lockhart et al. 1998). If this is the case, then the importance of the shorter branch-length moa and shrews for binding their longer branch-length cousins while retaining deeper signals of ancestry is not surprising.

The first studies to include complete mt sequences from tinamous, moa, and other ratites in 2001 could have recovered the tinamou–moa grouping (see Fig. 1b). However, either ratite monophyly was enforced or RAS ML results were overlooked. It may be significant that in the lead-up to these papers several highly anomalous phylogenetic relationships were proposed on the basis of mt genome analyses, including passerines being sister to all other extant birds (e.g., Mindell et al. 1999) and the egg-laying monotremes grouping with marsupials among mammals (Janke et al. 1996). Ancient DNA studies were also in their infancy and were viewed with some suspicion (see Cooper and Poinar 2000). Indeed, we too were guilty of these mistrusts, having left moa out of an avian mt genome phylogeny (Harrison et al. 2004) partly because their inclusion induced the "wrong" tree, in which tinamous fell inside ratites.

Additional taxon sampling and advances in analytical methods have since resolved earlier anomalies for mt genome phylogenies, including for rooting the avian and mammalian trees (see Paton et al. 2002; Phillips and Penny 2003). Furthermore, ancient DNA is now a respectable and thriving industry. Having ruled out compositional and long-branch artifacts (Tables 3 and 4), we consider the present mt genome analyses to provide

strong and unambiguous support for tinamous and moa being sister taxa and hence ratites being paraphyletic.

Independent Evolution of Flightless Palaeognaths

Multiple losses of flight among ratites have often been suggested, typically under the influence of a need to explain apparent requirements for long-distance dispersal over water (e.g., Houde 1988; Briggs 2003). Harshman et al. (2008) suggested at least 3 losses of flight among palaeognaths, as inferred from ratite paraphyly in their phylogenetic tree and a belief that flight is more likely lost than regained among birds. Alternatively, they recognized that a flightless MRCA of palaeognaths and subsequent regain of flight in tinamous was the most parsimonious option. Standard MP is essentially no-common-mechanism ML (Tuffley and Steel 1997), and so reconstruction at any given node is not informed by patterns at other nodes that indicate that loss of flight is more likely than gain.

In Table 7(c–e), it is shown that as the expected ratio of flight loss to gain increases from 1.46 to 2.34 through various ML ancestral state reconstruction analyses, the probability of each of the palaeognath MRCAs being flighted (and hence, of subsequent flight losses) increases. Indeed, 4 losses of flight among palaeognaths cannot be rejected even at $P \leqslant 0.2$ in the latter scenario that assumes we had sampled 2 flightless galloanserae. Even here a flight loss/gain ratio of 2.34 is far too conservative given hundreds of reported flight losses among birds (as discussed below) and no clear evidence for any regains. Hence, our prior expectation approaches the Dollo parsimony situation with regains not permitted, such that losses are required along each of the ostrich, rhea, moa, and kiwi/Casuariidae stem lineages (as shown in Fig. 5). If the MRCA of Casuariidae and kiwi (for which our best estimate of 60 Ma postdates New Zealand–Australia separation) was volant and so was their MRCA with elephant birds, then up to 6 losses of flight among palaeognaths are required.

Loss of flight has occurred very frequently among birds; McCall et al. (1998) suggested loss of flight occurred in at least 11 extant avian families and Steadman (2006) reported more than 100 losses of flight on the Pacific Islands among the Rallidae alone. Indeed, Maynard-Smith (1968), Feduccia (1996), Bautista et al. (2001), and others have argued that theory predicts that flight will tend to be lost in the absence of direct selection for its maintenance (catching food on the wing, predator avoidance, etc.). Wings and associated pectoral apparatus are costly to maintain, and walking is more economical for ground-feeding birds if food sources are not widely distributed. Moreover, the "cost" of powered flight increases with the linear dimension (l) in proportion to $l^{3.5}$, whereas the power from the muscle (or the area of the wing) only increases in proportion to l^2 (Maynard-Smith 1968, pp. 13–14). Hence, loss of flight is not surprising given that larger size has clearly been selected for among ratites, relative to the typically chicken-sized, volant and probably para- or polyphyletic assemblage of lithornithids they are thought to derive from (see Leonard et al. 2005). By contrast, regaining flight in tinamous after at least 20 Ma of selection for flightless locomotion and erosion of previously evolved genetic architecture associated with all aspects of flight is implausible, especially in the presence of already flight-adapted competitors.

Morphological support for ratite monophyly to the exclusion of tinamous is typically strong in cladistic analyses (e.g., Livezey and Zusi 2007). However, many of the ratite synapomorphies are directly associated with flightlessness or have previously been suggested in various groups to be developmentally correlated through paedomorphosis (Livezey 1995; Härlid and Arnason 1999). Interestingly, Elzanowski (1995) focused his morphological study primarily on the basicranuim, which he believed to have relatively little functional/developmental association with flightlessness and as a result found tinamous falling within ratites and the ostrich deepest among palaeognaths. Even so, Cubo and Arthur (2001) showed that numerous cranial (not basicranial) characteristics appear to be developmentally correlated with the pelvic peramorphosis that occurs in flightless birds, particularly the more cursorial taxa. A number of authors (e.g., Jollie 1977; Härlid and Arnason 1999) have proposed such heterochronous developmental mechanisms to have been of fundamental importance for understanding the evolution of ratite morphologies. Additionally, differential character scaling with size (allometry) also disguises developmental correlations that can pose as phylogenetic signal (Szalay 1994).

If character correlations associated with allometric scaling and cursoriality contribute substantially to morphological signal for ratite monophyly, then it is predictable that kiwi, the smallest and least cursorial ratite, might be attracted toward a basal position in morphological cladistic analyses. Indeed, this has most often been the case. Livezey and Zusi (2007) grouped the more cursorial ratites to the exclusion of kiwi with 92% MP bootstrap support, whereas Houde (1986) advocated kiwi arising independently of other ratites, from a different lineage among lithornithids. Hopefully, the present molecular results will encourage some reconsideration of avian morphological character analysis. Utilizing the tinamou–moa grouping within a molecular phylogenetic scaffold for morphological studies may benefit inferences of character state polarities.

One of the curiosities of flight being lost independently among several ratite lineages is the implication that flying lithornithid-like lineages likely became extinct independently in North America, Eurasia, Madagascar, New Zealand, and Australia–Antarctica, with only the tinamous surviving in South America. All palaeognaths appear to have become extinct during the Tertiary in the northern continents, until the later rearrival of ostriches. The avian fossil records of the abovementioned Gondwanan landmasses are very poor or entirely missing from the Late Cretaceous through to

the Early Tertiary and contain no records of lithornithid-like birds. However, we might imagine scenarios similar to the more recent and transparent case of swamphens (*Porphyrio sp.*) flying to New Zealand, where ecological conditions differed from those for their ancestral population (an absence of mammalian predators in this case). Apparently selection for larger, flightless birds simultaneously resulted in the evolution of the Takahe (*Porphyrio hochstetteri*) and the elimination of flying swamphens—until humans and their commensals arrived (Trewick 1996).

Vicariance biogeography provides the most familiar hypothesis for the diversification of ratites as ancestrally flightless cousins set adrift upon the break-up of Gondwana (Cracraft 1974). Much of the Gondwanan break-up occurred too early for this to be plausible (see below). Here we suggest a new hypothesis that flying palaeognaths accessed similar novel niche opportunities that became available on different landmasses with the K/T boundary extinction of dinosaurs and in the absence of previously overwhelming predation pressures, independently became flightless. Consistent with this proposal, Figure 5 shows that each of the branches along which flight was apparently lost either crosses or originates after the K/T boundary. Relevant fossil evidence is sparse, but nonetheless fits. Flying lithornithids are known from the Late Cretaceous (Parris and Hope 2002), whereas the earliest fossil records for putatively flightless palaeognaths, which were somewhat larger than lithornithids, are known from just a few million years after the K/T boundary in Palaeocene deposits from Argentina (Alvarenga 1993) and possibly France (Martin 1992). Perhaps, the apparent mimicry between ratites such as the ostrich (*Struthio*) and dinosaurs such as *Struthiomimus* is more closely linked than previously recognized.

Palaeognath Biogeography and the Question of Gondwanan Origins

The geographic distribution of extant palaeognaths and moa is shown in Figure 5. Our proposal of 4–6 losses of flight among the ancestors of recent palaeognaths invalidates a major underpinning of the Gondwanan vicariance model for explaining this distribution, the assumption that the MRCA of ratites was flightless. ML probabilities for the geographic location of MRCAs at the 5 deepest palaeognath nodes are shown in Table 6.

Interpretations based on vicariance and the mt consensus both place the last common ancestor of ratites and potentially palaeognaths in Gondwana. In contrast, we are unable to give preference to either a Northern Hemisphere origin with a common ancestor of non-ostrich palaeognaths dispersing to South America or a Gondwanan origin with an ostrich ancestor dispersing from South America to the Northern Hemisphere. With fossil information included in our analyses, ML favors a South Gondwanan origin, though does not clearly reject a Northern Hemisphere origin ($P = 0.18$), which is in fact favored under MP. Fundamental to both reconstructions is a late migration of ostriches to Africa. Most studies (e.g., Cooper et al. 2001; Slack et al. 2006), including ours, have placed the divergence of ostriches from an ancestor with other ratites that was too recent to catch the splitting of Africa from South Gondwana. A sister relationship for the ostrich with the NZ–Australian clade in results influenced by compositional biases fitted better with the possible use of (still uncertain) connections between Antarctica, the Kerguelen plateau, and Indo-Madagascar as a dispersal route to Africa. Instead, the basal placement of the ostrich favors a Late Cretaceous proto-Antilles dispersal route between South America and North America as has been suggested by van Tuinen et al. (1998). Notably, this temporal window for traversing the proto-Antilles land bridge also appears to have been used for the northerly migration of titanosaurid dinosaurs and vice versa for marsupials (see Pascual 2006).

Eurasian origins for ostriches have previously been proposed on the basis of both the possibility of separate origins of ratites from among lithornithids and the earliest fossil evidence for the lineage being from Eocene deposits in Europe (*Paleotis*, Houde 1986). Furthermore, the first ostriches (or indeed ratites) known from Africa are not revealed in the fossil record until the Miocene (Leonard et al. 2006), just subsequent to the first major Eurasian-African biotic interchange (Kappelman et al. 2003). In either case, the evidence does not hold up the status of ratites as the iconic Gondwanan taxon—with origins that predate the supercontinent's major break-up events and that later dispersed in accordance with vicariance and remnant ridges/microcontinents (of uncertain temporal and geographic continuity between the major landmasses). Furthermore, we confirm the view of Cooper et al. (2001) that elephant birds are not closely related to the ostrich. These Madagascan ratites appear to derive from among the Casuariidae/kiwi/tinamou–moa clade, with which they would share a South Gondwanan MRCA. The oceanic barrier between these lands and Indo-Madagascar had opened up by 110 Ma (Hay et al. 1999), such that either flight or extensive emergence of the Kerguelan Plateau would appear to necessitate dispersal.

A South Gondwanan origin for the MRCA of the "non-ostrich" palaeognaths is supported unequivocally by the nested ML dispersal analyses (Node 2, Fig. 5). More specifically, with the fossil information included for just this South Gondwanan clade, a South American origin is strongly supported ($P = 0.90$, Table 6, Node 2).

Our ML analyses do not clearly favor either of 2 scenarios for dispersal of palaeognaths out of South America into Antarctica–Australia–New Zealand (see Table 6, Nodes 3–4). These being a single dispersal, followed by a back migration to South America for tinamous or 2 dispersals, one for the ancestors of kiwi/Casuariidae and another for the ancestors of moa. We prefer the latter scenario because the presumed longer history of

palaeognaths in South America and that continent's present richness of tinamous hint at greater potential as a dispersal source. Admittedly, the paucity of Late Cretaceous/Early Tertiary fossil evidence from Antarctica–Australia–New Zealand ensures that this reasoning is tentative. Nevertheless, the best estimates from either of our molecular timescales for any of these dispersals out of South America are contained within 76 and 53 Ma (see Table 5, between Nodes 2 to 3 and 2 to 4/5). These match the dates for marsupials entering Australia via trans-Antarctic dispersal from South America between 72 and 55 Ma (Woodburne and Case 1996; Beck 2008). The finding by Tambussi et al. (1994) of ratite fossil material from Eocene deposits on Seymour Island, Antarctica is consistent with similar timing for ratite dispersal. Hence, palaeognath dispersals between Australia–Antarctica and between the southern and northern hemispheres (and ultimately Africa) may have been associated with major faunal migration intervals. In contrast, a plausible explanation for the origins of palaeognaths in New Zealand and Madagascar is provided by retention of flight well beyond the MRCA of the ratite birds.

CONCLUSIONS

Previous mt genome analyses of palaeognath birds have either constrained ratites and tinamous to be reciprocally monophyletic or have relied on analyses in which rate variation among variable sites is ignored. Relaxing these topological and methodological constraints leads to moa and tinamous grouping together. Increased taxon sampling and RY-coding the rapidly evolving third codon positions reduces tendencies for long-branch attraction between the tinamous and the outgroup taxa and provides strong statistical support for a tinamou–moa grouping. The implication that flight has been lost multiple times among palaeognaths calls into question the use of continental break-up dates for calibrating molecular clocks. In addition, the central role ratites have played in arguments concerning Gondwanan biogeography and the proposed Oligocene drowning of New Zealand (Waters and Craw 2006) need to be reevaluated.

Our phylogenetic reconstructions and examination of nucleotide composition bias suggest that ostriches diverged from the root node of Palaeognathae, in agreement with recent nuclear studies (Hackett et al. 2008; Harshman et al. 2008). Our biogeographic reconstructions are consistent with early palaeognath dispersal to or even origins in the northern continents. In turn, this helps explain the absence of African ratites until after the End Oligocene biotic interchange with Eurasia, which substantially postdates Eocene fossil records for apparent ostrich relatives in Eurasia.

Inferences from taxonomic/phylogenetic diversity alone have been unable to uncover remnant signal among modern terrestrial vertebrates for an evolutionary response to the end Cretaceous events (e.g., Bininda-Emonds et al. 2006). Multiple losses of flight coincident with size increases among palaeognaths independently on different landmasses following the end Cretaceous extinction event is suggestive of such a signature. We propose that large size and cursoriality and consequently loss of flight were selected for among Early Tertiary ancestors of modern ratites in filling parts of the ecospace vacated upon the K/T boundary extinction of mid–large-sized terrestrial vertebrates, including dinosaurs. Further resolution of palaeognath relationships and a more precise timescale will be crucial for testing this hypothesis and further unravelling the evolutionary history of flightless palaeognaths.

SUPPLEMENTARY MATERIAL

Supplementary material can be found at http://www.sysbio.oxfordjournals.org/.

FUNDING

This work was funded by an Australian Research Council Discovery grant to M.J.P. and by the New Zealand CORE fund to D.P.

ACKNOWLEDGMENTS

We would like to thank Andrew Rambaut for providing an alignment from Cooper et al. (2001) and Walter Boles for advice on the extinct casuariiform genus, *Emuarius*. We have had several helpful discussions on palaeognath evolution with Alan Cooper and Nic Rawlence. *Systematic Biology* Editor-in-Chief Jack Sullivan, Associate Editors Adrian Paterson and Todd Oakley, and 4 anonymous reviewers provided constructive criticism.

REFERENCES

Alvarenga H.M.E. 1983. Uma ave ratita do paleoceno brasileiro: bacia caladria de Itaborai, estato do rio de Janeiro, Brasil. Boletim do Museu Nacional (Rio de Janeiro). Geologia. 37:1–194.

Barnett R., Barnes I., Phillips M.J., Martin L.D., Harington C.R., Leonard J.A., Cooper A. 2005. Evolution of the extinct Sabretooths and the American cheetah-like cat. Curr. Biol. 15:R589–R590.

Bautista L.M., Tinbergen J., Kacelnik A. 2001. To walk or to fly? How birds choose among foraging modes. Proc. Natl. Acad. Sci. USA. 98:1089–1094.

Beck R.M.D. 2008. A dated phylogeny of marsupials using a molecular supermatrix and multiple fossil constraints. J. Mammal. 89:175–189.

Benton M.J., Donoghue P.C.J. 2007. Paleontological evidence to date the tree of life (vol 24, pg 26, 2007). Mol. Biol. Evol. 24:889–891.

Bertelli S., Porzecanski A.L. 2004. Tinamou (Tinamidae) systematics: a preliminary combined analysis of morphology and molecules. Ornitol. Neotrop. 15 (Suppl.):1–8.

Bininda-Emonds O.R.P., Cardillo M., Jones K.E., MacPhee R.D.E., Beck R.M.D., Grenyer R., Price S.A., Vos R.A., Gittleman J.L., Purvis A. 2007. The delayed rise of present-day mammals. Nature. 446: 507–512.

Bock W.J. 1963. The cranial evidence for ratite affinities. In: Sibley C.G., editor. Proceedings of the XIII International Ornithological Congress. Washington (DC): American Ornithologists' Union. p. 39–54.

Boles W.E. 1992. Revision of *Dromaius gidju* Patterson and Rich 1987 from Riversleigh, northwestern Queensland, with a reassessment of its generic position. Nat. Hist. Mus. Los Angel. Cty. Sci. Ser. 36: 195–208.

Briggs J.C. 2003. Fishes and birds: Gondwana life rafts reconsidered. Syst. Biol. 52:548–553.

Brochu C.A. 2004. Patterns of calibration age sensitivity with quartet dating methods. J. Paleontol. 78:7–30.

Brown J.W., Rest J.S., Garcia-Moreno J., Sorenson M.D., Mindell D.P. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biol. 6:6.

Buckley T.R., Arensburger P., Simon C., Chambers G.K. 2002. Combined data, Bayesian phylogenetics, and the origin of the New Zealand cicada genera. Syst. Biol. 51:4–18.

Caterino M.S., Reed R.D., Kuo M.M., Sperling F.A.H. 2001. A partitioned likelihood analysis of swallowtail butterfly phylogeny (Lepidoptera: Papilionidae). Syst. Biol. 50:106–127.

Cavender J.A., Felsenstein J. 1987. Invariants of phylogenies in a simple case with discrete states. J. Classif. 4:57–71.

Clarke J.A., Tambussi C.P., Noriega J.I., Erickson G.M., Ketcham R.A. 2005. Definitive fossil evidence for the extant avian radiation in the Cretaceous. Nature. 433:305–308.

Cooper A., Lalueza-Fox C., Anderson S., Rambaut A., Austin J., Ward R. 2001. Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. Nature. 409:704–707.

Cooper A., Poinar H.N. 2000. Ancient DNA: do it right or not at all. Science. 289:1139.

Cooper R.A., Millener P. 1993. The New Zealand biota: historical background and new research. Trends Ecol. Evol. 8:429–433.

Cracraft J. 1974. Phylogeny and evolution of the ratite birds. Ibis. 116:494–521.

Cubo J., Arthur W. 2001. Patterns of correlated character evolution in flightless birds: a phylogenetic approach. Evol. Ecol. 14:693–702.

Delsuc F., Phillips M.J., Penny D. 2003. Comment on "Hexapod Origins: Monophyletic or Paraphyletic?". Science. 301:1482.

Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:699–710.

Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Elzanowski A. 1995. Cretaceous birds and avian phylogeny. Cour. Forsch. Inst. Senckenb. 181:37–53.

Feduccia A. 1996. The origin and evolution of birds. New Haven (CT): Yale University Press.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

Galtier N., Gouy M. 1998. Inferring pattern and process: maximum likelihood implementation of a non-homogeneous model of DNA sequence evolution for phylogenetic analysis. Mol. Biol. Evol. 15:871–879.

García-Moreno J., Mindell D.P. 2000. Using homologous genes on opposite sex chromosomes (gametologs) in phylogenetic analysis: a case study with avian CHD. Mol. Biol. Evol. 17:1826–1832.

Gibb G.C., Kardailsky O., Kimball R.T., Braun E.L., Penny D. 2007. Mitochondrial genomes and avian phylogeny: complex characters and resolvability without explosive radiations. Mol. Biol. Evol. 24:269–280.

Gibson A., Gowri-Shankar V., Higgs P.G., Rattray M. 2005. A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. Mol. Biol. Evol. 22:251–264.

Gilbert M.T.P., Tomsho L.P., Rendulic S., Packard M., Drautz D.I., Sher A., Tikhonov A., Dalen L., Kuznetsova T., Kosintsev P., Campos P.F., Higham T., Collins M.J., Wilson A.S., Shidlovskiy F., Buigues B., Ericson P.G.P., Germonpre M., Gotherstrom A., Iacumin P., Nikolaev V., Nowak-Kemp M., Willerslev E., Knight J.R., Irzyk G.P., Perbost C.S., Fredrikson K.M., Harkins T.T., Sheridan S., Miller W., Schuster S.C. 2007. Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. Science. 317:1927–1930.

Gontcharov A.A., Marin B., Melkonian M. 2004. Are combined analyses better than single gene phylogenies? A case study using SSU rDNA and rbcL sequence comparisons in the Zygnematophyceae (Streptophyta). Mol. Biol. Evol. 21:612–624.

Hackett S.J., Kimball R.T., Reddy S., Bowie R.C.K., Braun E.L., Braun M.J., Chojnowski J.L., Cox W.A., Han K.L., Harshman J., Huddleston C.J., Marks B.D., Miglia K.J., Moore W.S., Sheldon F.H., Steadman D.W., Witt C.C., Yuri T. 2008. A phylogenomic study of birds reveals their evolutionary history. Science. 320: 1763–1768.

Haddrath O., Baker A.J. 2001. Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis. Proc. R. Soc. Lond. B. 268:939–945.

Härlid A., Arnason A. 1999. Analyses of mitochondrial DNA nest ratite birds within the Neognathae: supporting a neotenous origin of ratite morphological characters. Proc. R. Soc. Lond. B. 266: 305–309.

Härlid A., Janke A., Arnason U. 1998. The complete mitochondrial genome of *Rhea americana* and early avian divergences. J. Mol. Evol. 46:669–679.

Harrison G.L., McLenachan P.A., Phillips M.J., Slack K.E., Cooper A., Penny D. 2004. Four new avian mitochondrial genomes help get basic evolutionary questions in the late cretaceous. Mol. Biol. Evol. 21:974–983.

Harshman J., Braun E.L., Braun M.J., Huddleston C.J., Bowie R.C.K., Chojnowski J.L., Hackett S.J., Han K.L., Kimball R.T., Marks B.D., Miglia K.J., Moore W.S., Reddy S., Sheldon F.H., Steadman D.W., Steppan S.J., Witt C.C., Yuri T. 2008. Phylogenomic evidence for multiple losses of flight in ratite birds. Proc. Natl. Acad. Sci. USA. 105:13462–13467.

Hay W.W., DeConto R.M., Wold C.N., Wilson K.M. Voigt S., Schulz M., Wold A.R., Dullo W.-C., Ronov A.B., Balukhovsky A.N., Söding E. 1999. Alternative global Cretaceous paleogeography. Geol. Soc. Am. Spec. Pap. 332:1–47.

Hendy M.D., Penny D. 1989. A framework for quantitative study of evolutionary trees. Syst. Biol. 38:297–309.

Higuchi Y., Linn S. 1995. Purification of all forms of HeLa cell mitochondrial DNA and assessment of damage to it caused by hydrogen peroxide treatment of mitochondria or cells. J. Biol. Chem. 270:7950–7956.

Ho S.Y.W., Phillips M.J. 2009. Accounting for calibration uncertainty in phylogenetic estimation of divergence times. Syst. Biol. 58:367–380.

Houde P., Olson S.L. 1981. Paleognathous carinate birds from the early tertiary of North-America. Science. 214:1236–1237.

Houde P.W. 1986. Ostrich ancestors found in the northern hemisphere suggest new hypothesis of ratite origins. Nature. 324:563–565.

Houde P.W. 1988. Palaeognathous birds from the early tertiary of the northern hemisphere. Publ. Nuttall Ornithol. Club. 22:1–148.

Huelsenbeck J., Ronquist F. 2001. MrBayes: Bayesian inference on phylogenetic trees. Bioinformatics. 17:754–755.

Hugall A.F., Foster R., Lee M.S. 2007. Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. Syst. Biol. 56:543–563.

Huxley T.H. 1867. On the classification of birds; and on the taxonomic value of the modifications of certain of the cranial bones observable in that class. Proc. Zool. Soc. Lond. 1867:415–472.

Janke A., Gemmell N.J., Feldmaier-Fuchs G., von Haeseler A., Paabo S. 1996. The mitochondrial genome of a monotreme—the platypus (Ornithorhynchus anatinus). J. Mol. Evol. 42:153–159.

Jollie M.A. 1977. Contribution to the morphology and phylogeny of Falconiformes. Evol. Theory. 3:209–300.

Kappelman J., Rasmussen D.T., Sanders W.J., Feseha M., Bown T., Copeland P., Crabaugh J., Fleagle J., Glantz M., Gordon A., Jacobs B., Maga M., Muldoon K., Pan A., Pyne L., Richmond B., Ryan T., Seiffert E.R., Sen S., Todd, L., Wiemann M.C., Winkler A. 2003. Oligocene mammals from Ethiopia and faunal exchange between Afro-Arabia and Eurasia. Nature. 426:549–552.

Karanth K.P. 2006. Out-of-India Gondwanan origin of some tropical Asian biota. Curr. Sci. 90:789–792.

Kishino H., Hasegawa M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data and the branching order Hominidae. J. Mol. Evol. 29: 170–179.

Kretteck A., Gullburg A., Arnason A. 1995. Sequence analysis of the complete mitochondrial DNA molecule of the hedgehog, Erinaceus europaeus, and the phylogenetic position of the Lipotyphla. J. Mol. Evol. 41:952–957.

Lawver L.A., Gahagan L.M., Coffin M.F. 1992. The development of palaeoseaways around Antarctica. Ant. Res. Ser. 56:7–30.

Lee K., Felsenstein J., Cracraft J. 1997. The phylogeny of ratite birds: resolving conflicts between molecular and morphological data sets.

In: Mindell D.P., editor. Avian molecular systematics and evolution. New York: Academic Press. p. 173–211.

Leonard L., Dyke G.J., van Tuinen M. 2005. A new specimen of the fossil palaeognath *Lithornis* from the Lower Eocene of Denmark. Am. Mus. Nov. 3491:1–11.

Leonard L., Dyke G.J., Walker C.A. 2006. New specimens of a fossil ostrich from the Miocene of Kenya. J. Afr. Earth Sci. 45:391–394.

Lin Y.-H., McLenachan P.A., Gore A.R., Phillips M.J., Ota R., Hendy M., Penny D. 2002. Four new mitochondrial genomes, and the increased stability of evolutionary trees of mammals from improved taxon sampling. Mol. Biol. Evol. 19:2060–2070.

Livezey B.C. 1995. Heterochrony and the evolution of avian flightlessness. In: McNamara K.J., editor. Evolutionary change and heterochrony. Chichester: John Wiley and Sons Ltd. p. 169–193.

Livezey B.C., Zusi R.L. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool. J. Linn. Soc. 149:1–95.

Lockhart P.J., Steel M.A., Barbrook A.C., Huson D.H., Charleston M.A., Howe C.J. 1998. A covariotide model explains apparent phylogenetic structure of oxygenic photosynthetic lineages. Mol. Biol. Evol. 15:1183–1188.

Lockhart P.J., Steel M., Hendy M., Penny D. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11:605–612.

Maisey J.G. 2000. Continental break up and the distribution of fishes of Western Gondwana during the Early Cretaceous. Cretaceous Res. 21:281–314.

Martin L.D. 1992. The status of the late Paleocene birds *Gastornis* and *Remiornis*. Nat. Hist. Mus. Los Angel. Cty. Sci. Ser. 36:97–108.

Maynard-Smith J. 1968. Mathematical ideas in biology. Cambridge: Cambridge University Press.

Mayr E., Amadon D. 1951. A classification of recent birds. Am. Mus. Novit. 1496:1–42.

McCall R.A., Nee S., Harvey P.H. 1998. The role of wing length in the evolution of avian flightlessness. Evol. Ecol. 12:569–580.

Mindell D.P., Sorenson M.D., Huddleston C.J., Hasegawa M., Ast J.C., Yuri T. 1999. Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. Syst. Biol. 48: 138–152.

Morgan-Richards M., Trewick S.A., Bartosch-Haerlid A., Kardailsky O., Phillips M.J., McLenachan P.A., Penny D. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. BMC Evol. Biol. 8:20.

Pagel M.D., Meade A. 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. Am. Nat. 6:808–825.

Parris D.C., Hope S. 2002. New interpretations of birds from the Navesink and Hornerstown formations, New Jersey, USA (Aves: Neornithes). In: Zhou Z., Zhang F., editors. Proceedings of the 5th Symposiuonm of the Society of Avian Paleontology and Evolution. Beijing: Science Press. p. 113–124.

Pascual R. 2006. Evolution and geography: the biogeographic history of south American land mammals. Ann. Mo. Bot. Gard. 93: 209–230.

Paton T., Haddrath O., Baker A.J. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. Proc. R. Soc. Lond. B. 269:839–846.

Pereira S.L., Baker A.J. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. Mol. Biol. Evol. 23:1731–1740.

Phillips M.J. 2009. Branch-length estimation bias misleads molecular dating for a vertebrate mitochondrial phylogeny. Gene. 441:132–140.

Phillips M.J., Delsuc F., Penny D. 2004. Genome-scale phylogeny: sampling and systematic errors are both important. Mol. Biol. Evol. 21:1455–1458.

Phillips M.J., Lin Y.-H., Harrison G.L., Penny D. 2001. Mitochondrial genomes of a bandicoot and a brushtail possum confirm the monophyly of Australiadelphian marsupials. Proc. R. Soc. Lond. B. 268:1533–1538.

Phillips M.J., Penny D. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. Mol. Phylogenet. Evol. 28: 171–185.

Phillips M.J., Pratt R.C. 2008. Family-level relationships among the Australasian marsupial "herbivores" (Diprotodontia: koala, wombats, kangaroos and possums). Mol. Phylogenet. Evol. 46:594–605.

Posada D., Crandall K.A. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics. 14:817–818.

Prager E.M., Wilson A.C., Osuga D.T., Feeney R.E. 1976. Evolution of flightless land birds on southern continents—transferrin comparison shows monophyletic origin of ratites. J. Mol. Evol. 8:283–294.

Pratt R.C., Gibb G.C., Morgan-Richards M., Phillips M.J., Hendy M.D., Penny D. 2009. Towards resolving deep Neoaves phylogeny: data, signal enhancement and priors. Mol. Biol. Evol. 26:313–326.

Rambaut A. 1996. Sequence Alignment Editor. Version v1.0 alpha1 [Internet]. Available from http://tree.bio.ed.ac.uk/software/seal/

Rambaut A., Drummond A.J. 2007. Tracer v1.0 [Internet]. Available from: http://evolve.zoo.ox.ac.uk.

Ree R.H., Smith S.A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57:4–14.

Sanmartín I., Ronquist F. 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst. Biol. 53:216–243.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51:492–508.

Shimodaira H., Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16:1114–1116.

Shimodaira H., Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics. 17: 1246–1247.

Sibley C.G., Ahlquist J.E. 1990. Phylogeny and classification of birds: a study in molecular evolution. New Haven (CT): Yale University Press.

Siddall M.E. 1998. Success of parsimony in the four-taxon case: long-branch repulsion by likelihood in the Farris zone. Cladistics. 14:209–220.

Slack K.E., Delsuc F., McLenachan P.A., Arnason U., Penny D. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. Mol. Phylogenet. Evol. 42:1–13.

Slack K.E., Jones C.M., Ando T., Harrison G.L., Fordyce E., Arnason U., Penny D. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol. Biol. Evol. 23:1144–1155.

Steadman D.W. 2006. Extinction and biogeography of tropical Pacific birds. Chicago (IL): University of Chicago Press.

Susko E., Field C., Blouin C., Roger A.J. 2003. Estimation of rates-across-sites distributions in phylogenetic substitution models. Syst. Biol. 52:594–603.

Suzuki Y., Glazko G.V., Nei M. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. USA. 99:16138–16143.

Szalay F.S. 1994. Evolutionary history of the marsupials and an analysis of osteological characters. New York: Cambridge University Press.

Szatmari P., Milani E.J. 1999. Microplate rotation in north-east Brazil during South Atlantic rifting: analogies with the Sinai microplate. Geology. 27:1115–1118.

Tambussi C.P., Noriega J.I., Gazdicki A., Tatur A., Reghuero M., Vizcaino S.F. 1994. Ratite bird from the Paleogene La Meseta formation, Seymour Island, Antarctica. Pol. Polar Res. 15:15020–15026.

Tamura K., Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10:512–526.

Trewick S.A. 1996. Morphology and evolution of two takahe: flightless rails of New Zealand. J. Zool. Lond. 238:221–237.

Tuffley C., Steel M. 1997. Links between maximum likelihood and maximum parsimony under a simple model of site substitution. Bull. Math. Biol. 59:581–607.

van Tuinen M., Sibley C.G., Hedges S.B. 1998. Phylogeny and biogeography of ratite birds inferred from DNA sequences of the mitochondrial ribosomal genes. Mol. Biol. Evol. 15:370–376.

van Tuinen M., Sibley C.G., Hedges S.B. 2000. The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. Mol. Biol. Evol. 17:451–457.

Veevers J.J. 2004. Gondwanaland from 650–500 Ma assembly through 320 Ma merger in Pangea to 185–100 Ma breakup: supercontinental tectonics via stratigraphy and radiometric dating. Earth Sci. Rev. 68:1–124.

Waters J.M., Craw D. 2006. Goodbye Gondwana? New Zealand biogeography, geology, and the problem of circularity. Syst. Biol. 55:351–356.

Woodburne M.O., Case J.A. 1996. Dispersal, vicariance, and the late cretaceous to early tertiary land mammal biogeography from South America to Australia. J. Mammal. Evol. 3:121–161.

Yang Z. 1996. Maximum-likelihood models for combined analyses of multiple sequence data. J. Mol. Evol. 42:587–596.

You H.-L., Lamanna M.C., Harris J.D., Chiappe L.M., O'Connor J., Ji S.-A., Lu J.-C., Yuan C.-X., Li D.-Q., Zhang X., Lacovara K.J., Dodson P., Ji Q. 2006. A nearly modern amphibious bird from the early Cretaceous of Northwestern China. Science. 312:1640–1643.

Zhou Z. 2006. Evolutionary radiation of the Jehol Biota: chronological and ecological perspectives. Geol. J. 41:377–393.

APPENDIX

Reanalysis of the mt Data Set from Cooper et al. (2001)

The data set of Cooper et al. (2001) includes 10,767 aligned mt protein-coding sites for the chicken and 9 palaeognaths, including 2 tinamous, rhea, ostrich, cassowary, emu, kiwi, and 2 moa. Note that the suffix -s is not added to kiwi and moa in their plural forms, in accordance with the Maori language from which they are derived. The copy of this data set that we obtained did not include the chicken sequence. We added the chicken mt sequence and in order to check that its aligned inclusion closely matched that used in the original analyses, we replicated the ML bootstrap analysis from Cooper et al. (2001). This analysis employed a single GTR+Γ model for the concatenated protein-coding DNA sequences, which was run in PAUP*4.0b10 for 200 nonparametric pseudoreplicates. With ratite monophyly enforced, the same topology was recovered as in the original paper and bootstrap support was similar. Differences in bootstrap support from the published analysis averaged only 2.3% across all nodes, which may largely be accounted for by stochastic variation in bootstrap sampling. When ratite monophyly was not enforced, tinamous and moa grouped together as shown in Figure 1b.