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.

PATTERNS AND PROCESSES IN ANIMAL EVOLUTION: MOLECULAR PHYLOGENETICS OF SOUTHERN HEMISPHERE FAUNA

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

At the Allan Wilson Centre for Molecular Ecology and Evolution,
Institute of Molecular BioSciences, Massey University, Palmerston North, New Zealand

Renae Pratt
2008

There is no better high than discovery

-Edward O. Wilson.

ABSTRACT

Three kinds of processes are known to modify the geographical spatial arrangement of organisms: dispersal, extinction and vicariance. The Southern Hemisphere has an intriguing and complicated geological history that provides an ideal backdrop to study these processes. This thesis focuses on three historical events that illustrate these processes: the proposed marine inundation of New Zealand in the Oligocene, the asteroid impact at the K – Pg boundary, and the continental breakup of Gondwana. It investigates what impact these events had on species diversification by studying the phylogenetic relationships of two groups of taxa – the family Anostostomatidae (insects), and Neoaves (birds). Anostostomatidae were studied in relation to the Oligocene drowning and the break up of Gondwana as they have a wide southern distribution, found on all "Gondwanan" fragments with the exception of Antarctica, and are thought represent an ancient lineage that predates the Gondwanan breakup. Birds, in particular Neoaves, were studied in relation to the asteroid impact at the K – Pg boundary. Although birds are mobile and many circumnavigate the globe between seasons, they are suggested to have originated in the Southern Hemisphere in Gondwanan times, and subsequently undergone range expansion and diversification around the world. In order to address the relationship (if any) between modern biotic diversity and historical geological events, phylogenetic relationships were determined and where possible, molecular clock analysis carried out. Timing information provided by molecular clock analysis is important as it enables distinction between opposing hypotheses such as vicariance and dispersal.

In *Chapter Two*, the phylogenetic relationships within the family Anostostomatidae are investigated. One of the most controversial times in New Zealand's geological history is during the Oligocene. Some suggest that the lack of fossils and evidence for recent dispersal of numerous taxa support the notion that all modern biota reached the region during the last 25 million years. Anostostomatidae were chosen as they represent a group of insects that are thought to be ancient and there is little published data in the literature. Previous studies focused on the relationships within *Hemideina* and *Deinacrida* suggesting that these groups diversified in the early Miocene. The data presented here are from mitochondrial (COI and 12S) and nuclear (18S and 28S) sequences.

Molecular dating using a relaxed clock as implemented in BEAST suggest that in fact some lineages were present at or shortly after continental breakup and could have survived throughout this turbulent time. As there were no definitive fossils to use for calibration points, geological events were used as calibration points for the molecular clock. Mutation rates obtained from the different analyses were compared to those published for other insects in an attempt to identify the most likely model. Both maximum likelihood and Bayesian analyses support the presence of three distinct ecological groups in New Zealand; *Hemiandrus* (ground weta), *Anisoura/Motuweta* (tusked weta) and *Hemideina–Deinacrida* (tree–giant weta). With regards to their Australasian relatives (taxa from Australia and New Caledonia) it appears that the family is divided with the most northern New Zealand taxa (tusked weta) more closely related to New Caledonian taxa while all other New Zealand taxa are more closely related to Australian taxa. There does not appear to be any link between the Australian and New Caledonian taxa studied here. Results should be viewed with caution however as an increased mutation rate was observed in the New Caledonian-tusked weta lineage, something future studies will have to address.

Chapter Three presents new sequence data and phylogenetic analyses that go towards resolving the apparent basal polytomy of neoavian birds. This chapter includes analyses carried out on previously published data with the addition of nine new mitochondrial genomes. My contribution to this larger project was to perform the phylogenetic analysis and to sequence three of the nine mitochondrial genomes. The genomes I sequenced were the Southern Hemisphere species: dollar bird (*Eurystomus orientalis*), Owlet nightjar (*Aegotheles cristatus cristatus*) and great potoo (*Nyctibius grandis*). The inclusion of these nine new genomes allows assessment and comparison of the six hypothesised groups reported in Cracraft (2001). First an improved conditional down-weighting technique is described reducing noise relative to signal, which is important for resolving deeper divergences. Second, a formula is presented for calculating probabilities of finding predefined groupings in the optimal tree. Maximum likelihood and Bayesian based phylogenetic analyses were carried out and in addition, dating using a relaxed molecular clock was performed in BEAST. Results suggested that the six groups suggested by Cracraft (2001) represent robust lineages.

The results suggested that one group, the owls, are more closely related to other raptors, particularly accipitrids (buzzards/eagles) and the osprey rather than the Caprimulgiformes, which could indicate morphological convergent evolution.

In addition, a group termed shorebirds appears to be distinct from the large group referred to as 'Conglomerati' to which previous publications have suggested they belong. The 'Conglomerati' is the least well studied group and may actually comprise of at least three subgroups (as suggested by Cracraft). Within the three suggested groups, Cracraft grouped shorebirds with pigeons and sandgrouse, neither of which (pigeons or sandgrouse) were analysed here. So although the shorebirds are at least close to the 'Conglomerati' and may be within that group, their exact position is still not clear. The molecular dating reported here utilised two fossil calibrations (*Vegavis* and *Waimanu*), for which there is relatively little dispute as to age or the lineage to which they belong. Calibrations resulting from BEAST analyses suggest that at least 12 distinct lineages were present prior to the K – Pg boundary, a finding supported by previous studies. Robust phylogenies will allow future studies to investigate not only the relationships within Neoaves, but look more closely at the biological and ecological evolution of the group.

Chapter Four for the first time investigates whether the phylogenetic relationships within the family Anostostomatidae follow the conventionally accepted order and timing of Gondwanan breakup. Following the initial restults for taxa studied in Australasia (*Chapter Two*) an attempt to resolve family relationships in a wider spatial (geographic) context was carried out to determine if Australasian taxa are monophyletic when other members of the family are included. Again both maximum likelihood and Bayesian phylogenetic analyses were carried out on both mitochondrial (COI and 12S) and nuclear (18S and 28S) sequences. In this chapter, datasets included samples from across the geographic range of Anostostomatidae (South Africa, Madagascar, South America, Australia, New Caledonia and New Zealand), and two clades were observed, congruent with earlier findings. Sequence divergence within geographic regions was found to be relatively high in the mitochondrial genes (COI and 12S) while low in the nuclear ribosomal RNA genes (18S and 28S) as expected given their relative mutation rates. Under the vicariance paradigm, phylogenetic relationships should follow the order of continental breakup, but this was not found.

Further, if dispersal and colonisation were continuous, no geographic substructure is expected, however distinct geographic substructure within clades was consistently observed. This interesting phylogenetic pattern may be a case of convergent evolution or paraphyletic sampling which highlights taxonomic issues of the group.

Future studies need to include not only molecular data but information on morphology, ecology and behaviour along with the implementation of biogeographic programs that can test alternative hypotheses (such as dispersal and vicariance) directly. Also, the inclusion of the recently reported fossil from the subfamily Euclydesinae (Martins-Neto 2007) should allow for more accurate date estimates within the family.

Taken as a whole the results presented in this thesis suggest that microevolutionary processes are sufficient to explain modern diversity without the need to invoke abiotic events. The three cases investigated here - marine inundation, asteroid impact and continental drift - all appear to have had only a limited effect on the diversity of taxa studied. To reach even stronger conclusions future studies should incorporate different data (for instance nuclear genes, intron position, and genome structure) and use biogeographic software capable of including ecological, morphological and habitat information.

ACKNOWLEDGMENTS

The important thing is never to stop questioning

- Albert Einstein

No thesis can be finished without the support of friends and family no matter how far away they are. For this, I thank my family and friends 'across the ditch' for their continuous support throughout this stage in my life. Thanks for picking up the pieces when all seemed to be going wrong and for the 'motivational' pep talks when needed. An extra special thank you to Carol Scarpaci. You are my mentor, my friend, an inspiration.

A big thanks to; Tim White, Warwick Allen, Jing Wang and Napat Kantanon (Boffins) for assistance with anything and everything computer related. To members of the Farside lab; Angela McGaughran, Simon Hills, Gillian Gibb, Olga Kardialsky, and Steve Pilkington and from the Pleb lab; Michael Knapp and Richard Carter. I thank you all for advice and direction to varying degrees and to those who showed interest, regardless of whether you were actually interested! Ultimately, I thank you all for your friendship and support.

I am indebted to Andrew Clarke, Barbara Holland, Trish "It'll be fine!" McLenachan, Melanie Pierson, Sarah Bray, Gail Ashton, Yasmin Pearce, Chris Lane and Corina Jordan for the many counselling sessions and positive encouragement. Also to Anouck de Bonth and Judith Stepper and Dora for their friendship and support and the ability to endure at times a grumpy, stressed-out housemate. To Gemma Henderson, Sheralee Cleland, Gerry Ryan, Melle Steringa, Michelle Turner, Marijn Kool, Bouke de Jong, Nikki Kenters, Tobias Van Den Broek and Lindsay King for the entertainment at pub quizzes, Singstar and Buzz nights, and New Zealand adventures. Laughter really is the best medicine.

Kim McBreen, Matt Phillips and Barbara Holland, I will forever be grateful for the time you spent answering my endless questions about analyses. Thank you for explaining things to me like a 2 year old when necessary and making the time even when you didn't have it. I would still be reading an introduction to phylogenetic analysis if it weren't for your patience.

To our AWC administrators Joy, Karen and Susan for all your assistance with financial matters, booking flights and sorting accounts and in general making things run smoothly. I know at times, this can be a thankless job.

Thank you to Ang for proof reading the thesis in one day! To Judith for reading something so far out of your expertise and to Corina for giving up her weekend to proof one last time. Thanks to everyone who has helped over the years with writing and presentation preparation.

Lastly I would like to thank my supervisors without who I would never have had the opportunity to achieve a dream. To my co-supervisors, Steve Trewick and Mary Morgan-Richards, without your treasure trove of samples, this work would not have been possible. I also thank you for you for advice, assistance with writing, and thought provoking discussions. I know I have learnt a lot from this experience. To my chief supervisor, David Penny, I am enormously appreciative for advice and guidance throughout this process. All your stories, sidetracks and motivational support…'you're on the downhill slide, just one more little bump to overcome first!' and of course the financial support that no student can do without.

TABLE OF CONTENTS

ABSTRACT ..... ***i***

ACKNOWLEDGMENTS ..... ***v***

LIST OF FIGURES ..... ***xii***

LIST OF TABLES ..... ***xiii***

1. CHAPTER ONE **1**

THESIS STRUCTURE AND SCOPE ***2***

1.1 FACTORS THAT INFLUENCE OUR THINKING (BACKGROUND) ***4***

1.1.1 Drift theory 4

1.1.2 Biogeography 5

1.1.3 Getting a date in the modern era - Fossils Vs. DNA 6

1.1.4 Problems with comparisons 8

1.2 CASE STUDIES ***9***

1.2.1 Testing the depths of the Oligocene (Chapter Two) 10

1.2.2 What impact did the asteroid have? (*Chapter Three*) 16

1.2.3 Dispersal or drift? Getting around on Gondwana (*Chapter Four*) 18

1.2.4 Patterns and predictions 21

1.3 CONTRIBUTION OF OTHERS TO THE THESIS ***23***

1.4. REFERENCES ***24***

2. CHAPTER TWO **31**

Diversification of New Zealand weta (Orthoptera: Ensifera: Anostostomatidae) and their relationships in Australasia ***32***

2.1 ABSTRACT ***32***

2.2 INTRODUCTION ***32***

2.3 MATERIALS AND METHODS ***36***

(a) Sampling 36

(b) Molecular Methods 37

(c) Phylogenetic Analysis 38

(d) Tree comparisons 39

(e) Divergence-time estimation 40

2.4 RESULTS ***41***

(a) Phylogeny 41

(b) Summary of phylogenetics 47
(c) Divergence estimation 48
2.5 DISCUSSION ***49***
2.6 ACKNOWLEDGEMENTS ***51***
2.7 REFERENCES ***52***
2.8 APPENDIX ***56***
3. CHAPTER THREE **58**
Towards resolving deep Neoaves phylogeny: data, signal enhancement and priors ***59***
3.1 ABSTRACT ***59***
3.2 INTRODUCTION ***59***
3.3 MATERIALS AND METHODS ***66***
(a) Taxon Sampling 66
(b) Molecular Methods 67
(c) Phylogenetic analysis 69
(d) Noise reduction by down-weighting (Site-Stripping) 69
(e) Probabilities of observing predefined clades 71
3.4 RESULTS ***72***
3.5 DISCUSSION ***83***
3.6 CONCLUSION ***86***
3.7 ACKNOWLEDGEMENTS ***86***
3.8. REFERENCES ***87***
3.9 APPENDICES ***91***
3.9.1 Fossil calibrations and molecular dating with BEAST 91
3.9.2 The probability of observing a pre-specified clade 95
3.9.3 Tabulated results from Bayesian (BPP) and bootstrap (bs) analyses using RAxML and PhyML 97
3.9.4 Consensus tree including all outgroup taxa used for alignment before site stripping 98
4. CHAPTER FOUR **100**
Return to Gondwana: exploring biogeographical patterns of the family Anostostomatidae (Orthoptera: Ensifera) ***101***
4.1 ABSTRACT ***101***
4.2 INTRODUCTION ***102***

4.3 MATERIALS AND METHODS ***105***
(a) Sampling 105
(b) Molecular Methods 105
(c) Phylogenetic analysis 106
(d) Tree comparisons 107
4.4 RESULTS ***108***
(a) Sequence divergence 108
(b) Phylogenetic analysis 110
(c) Stemminess and Relative Compositional Variability 115
(d) Summary phylogenetics 116
4.5 DISCUSSION ***117***
4.6 ACKNOWLEDGEMENTS ***121***
4.7 REFERENCES ***121***
4.8 APPENDICES ***123***
4.8.1 Orthoptera Species File Online: taxa hierarchy as of 250908 (Eades and Otte) 124
4.8.2 Sample data including name, location, accession number, and genes used. 126
5. CHAPTER FIVE **128**
Summation ***129***
5.1 EVIDENCE FOR ABIOTIC EFFECTS? ***129***
5.2 CONCLUDING REMARKS ***136***
5.3 FUTURE WORK ***136***
5.4 REFERENCES ***138***
A. APPENDIX A **141**
Mitogenomics – exploring mitochondrial genomes from New Zealand weta (Orthoptera: Ensifera: Anostostomatidae) ***142***
A.1 ABSTRACT ***142***
A.2 INTRODUCTION ***142***
A.3 MATERIALS AND METHODS ***145***
(a) Sampling 145
(b) DNA extraction 145
(c) Long Range PCR 145
(d) Short Range PCR 146

(e)Restriction digestion 147
(f) Solexa sequencing 147
(g) Sequence alignment and annotation 147
A.4 RESULTS ***148***
(a) Genome organisation 148
(b) True sequence Vs. NUMTS 148
(c) Solexa results 149
(d) Heteroplasmy or NUMTS 150
A.5 DISCUSSION ***150***
A.6 ACKNOWLEDGEMENTS ***152***
A.7 REFERENCES ***152***
A.8 RAW DATA ***155***
A.8.1 Contigs that translate 155
A.8.2 Contigs that don't translate - possible pseudogenes 159
B. APPENDIX B **165**
General Materials ***166***
B.1 LABORATROY EQUIPMENT ***166***
B.1.1 General 166
B.1.2 Glassware 166
B .1.3 Consumables 166
B.2 CHEMICALS AND REAGENTS ***167***
B.2.1 Laboratory chemicals and reagents 167
B.2.2 Enzymes 168
B.2.3 DNA Markers 168
B.3 KITS AND READY-TO-USE PRODUCTS ***168***
B.4 SOLUTIONS AND DYES ***169***
B.5 SOFTWARE AND PHYLOGENETIC PROGRAMS ***170***
C. APPENDIX C **171**
C.1 PRIMERS ***172***
C.2 EXTRACTION PROTOCOLS ***172***
C.2.1 Salting Out 172
C.2.2 Rapid mtDNA Extraction (Alkaline Lysis) 173
C.2.3 Bird Kits 174

C.3 PCR PROTOCOL ***175***
C.3.1 Long Range 175
C.3.2 Short Range 178
C4 CLONING PROTOCOL ***184***
C.4.1 Restriction Enzyme Digestion and Cloning 184
C.4.2 Precipitation of digestion and vector 185
C.4.3 Shotgun cloning of digested long range fragments 185
C.4.4 Transformation of Competent Cells 186
C.4.5 Selection of Transformed Cells 186
C.5 SEQUENCING PROTOCOL ***187***
C.5.1 Purification 187
C.5.2 Quantification 188
C.5.3 Sequencing 188
D. APPENDIX D 189

LIST OF FIGURES

Figure 1.1 Schematic diagram to illustrate different changes being compared 9
Figure 1.2 Topography of Zealandian continent 11
Figure 1.3.Was there land in New Zealand? 12
Figure 1.4. Riding the West Wind Drift 20
Figure 1.5. The breakup of Gondwana 22
Figure 2.1 Timing of Gondwanan breakup and Australasian distribution map 33
Figure 2.2 New Zealand taxa distribution map 36
Figure 2.3 Ensifera cladogram 42
Figure 2.4 Bayesian phylogram using nuclear DNA 43
Figure 2.5 Maximum likelihood phylogram using mitochondrial DNA 45
Figure 2.6 Bayesian phylogram for mitochondrial DNA for the genus *Hemiandrus* 46
Figure 3.1 Unrooted Bayesian consensus network with minimal down-weighting 73
Figure 3.2 Unrooted Bayesian consensus network with maximum down-weighting 75
Figure 3.3 Rooted Bayesian phylogram with maximum down-weighting 77
Figure 3.4 Probability of finding predefined clades on a tree 81
Figure 3.5 BEAST output tree for Neoaves 94
Figure 3.6 Consensus tree of Neoaves (modern birds) before site-stripping 98
Figure 4.1 Rooted Bayesian phylogram using nuclear DNA 112
Figure 4.2 Rooted Bayesian phylogram using mitochondrial DNA 114
Figure 4.3 Distribution map of Anostostomatidae used in this study 115
Figure A.1 Complete mitochondrial genomes 143
Figure A.2 SEQUENCHER alignments for *Hemiandrus pallitarsis*. 149

LIST OF TABLES

Table 2-1 Sample information 56
Table 3-1 Probabilities of a predefined clade joining the tree as a single group, or as two subclades. 66
Table 3-2 Support values 97
Table 4-1 Amino acid changes 110
Table 4-2 RCV and stemminess values 116
Table 4-3 OSF online Anostostomatidae heirarcy 124
Table 4-4 Sample information 126
Table A-1 *Hemiandrus pallitarsis* fragments 148
Table B-1 General laboratory equipment used 166
Table B-2 General glassware used 166
Table B-3 List of general consumables used 166
Table B-4 List of chemicals and reagents used 167
Table B-5 List of Enzymes used 168
Table B-6 DNA markers used 168
Table B-7 Kits and ready-to-use products 168
Table B-8 Software and programs used 170
Table C-1 Long range PCR mix 176
Table C-2 Long range primers used for Anostostomatidae 176
Table C-3 Long range primers used for the dollar bird 177
Table C-4 Long range primers used for the great potoo 177
Table C-5 Long range primers used for the owlet nightjar 178
Table C-6 Short range primers used for Anostostomatidae 179
Table C-7 Short range primers used for the dollar bird 181
Table C-8 Short range primers used for the great potoo 182
Table C-9 Short range primers used for the owlet nightjar 183