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.

Evolution in a Marine Gastropod:
Rocks, Clocks, DNA and Diversity

Simon Francis Kahu Hills

2010

A thesis presented in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Evolutionary Biology

Allan Wilson Centre for Molecular Ecology and Evolution
Institute of Natural Resources
Massey University
Palmerston North
New Zealand

ABSTRACT

Comprehensive integration of paleontological and molecular data remains a sought-after goal of evolutionary research. This thesis presents a dataset unlike any previously studied to document changes over time in the evolutionary history of the New Zealand marine mollusc genus *Alcithoe*.

In order to study evolutionary relationships in the *Alcithoe,* DNA sequence of approximately 8Kb of mitochondrial DNA was generated using universal and newly developed PCR primers. The gene composition of the resulting sequences has been thoroughly analysed, using a novel splits-based approach, to gain a clear understanding of the underlying phylogenetic signals in the data. Refinement of the phylogeny was achieved by considering subsets of both the taxa and genes. Taking these analyses into account the combined a robust phylogeny for the *Alcithoe* is presented for use in subsequent analyses.

The *Alcithoe* genus includes species that are exemplars of the problem of correctly identifying species by morphological traits, in both the living and extinct taxa. Taxonomic assignments were explored in a population level analysis of the highly morphologically variable species *A. wilsonae*. Analyses revealed that the various recognised forms of *A. wilsonae* are genetically indistinguishable and that the previously recognised species *A. knoxi* is a synonym of *A. wilsonae*. This result has significant implications for the interpretation of the paleontological data, as *A. knoxi* specimens are known from the Tongaporutuan stage (10.92 – 6.5 Ma) of the New Zealand geological timescale. Therefore, this finding also has implications on the assignment of calibration data in molecular clock analysis.

To ensure accurate estimation of divergence times and rates of molecular evolution, extensive explorations of parameter space in molecular-clock analyses were carried out. These analyses identified the most appropriate models and calibration settings for *Alcithoe* the dataset. The fossil data used to calibrate this analysis is amongst the most robust applied to molecular clock analyses to date. Statistical sampling uncertainty derived from the paleontological data was included in the calculation of prior distributions. Divergence dates inferred for the extant *Alcithoe* are largely consistent with the fossil record. However, the root of the tree was consistently inferred to be younger then expected. Rates of evolution in the species of *Alcithoe* included in this analysis are broadly consistent. However, some small rate differences are observed in some branches, for example, *Alcithoe fusus* appears to have a faster rate then the rest of the genus. This rate increase is the likely cause of topological inconsistencies observed for four closely related taxa, including *A. fusus*, and indicates that slight rate differences can cause phylogenetic instability when small genetic distances are involved.

Direct comparison of diversification rates between the molecular and paleontological data for the *Alcithoe* illustrated that modern *Alcithoe* species have origins that are around 13 millions years younger then the oldest known *Alcithoe* fossils. The suggestion that *A. fusus* is descended from a series of fossil *Leporemax* species is directly contradicted by the molecular tree. In light of the molecular evidence this result highlights the problem of morphological convergence in the interpretation of fossil *Alcithoe* species. Comparison of the molecular and paleontological datasets was difficult for absolute speciation and extinction rates, as errors inherent to each dataset led to disparate estimates. For example, the fossil record clearly fails to record most recent speciation events observed in the molecular phylogeny, but the molecular data cannot sufficiently account for the amount of extinction evident in the fossil record. It is clear that the assumption of a constant and equal probability of speciation and extinction for all lineages is violated in the *Alcithoe*. However, the general long-term trends estimated for both datasets are concordant, and demonstrate an increase in both speciation and extinction rates over the Cenozoic era.

The research described in this thesis represents significant progress toward the goal of more thorough integration of molecular and paleontological in the study of evolution. I have shown that reconciliation of molecular and paleontological data is not only possible, but can substantially improve the resulting interpretation of evolution. This study is the broadest analysis of the evolution in a single genus using combined molecular and paleontological data that the author is currently aware of. It illustrates the advantage of having quality paleontological data to compare to emerging molecular data, and how the molecular data can further inform the paleontological data. Furthermore, it adds support to the shift in perspective from an adversarial to a complementary approach to the consideration of molecular and paleontological data. This thesis is a comprehensive first step in the synthesis of molecular and paleontological data in the study of evolution of the New Zealand mollusc fauna, and alludes to many promising avenues for future study.

ACKNOWLEDGEMENTS

Firstly to Mary, thank you for your supervision and support throughout the period of my thesis work. I’m sure I have been an ‘entertaining’ charge.

Many thanks to my co-supervisors Steve, James, and David for your continuing support and inspiration.

Thanks also to my collaborators Bruce, Alan, and Roger, your willingness and expertise have been instrumental in getting me to where I am now.

To the summer students, Melissa and Logan, who had been enlisted into various molluscan related projects, thank for your contributions and enthusiasm.

Klaus, Barbara, Tim, Bennet and Patrick, thanks to you guys for all your help over the years.

Thanks Joy, Susan and Trish, without you none of us would get anything done

Thanks to every one in the AWC, the work environment and camaraderie in the centre is without par.

Mum and Dad, it’s been a while, but it looks like I might finally have finished being a student. Thanks for everything you’ve done to get me here.

Special thanks must be extended to the makers of fine single malt whisky, without whom levels of stress would have been much greater. Thanks also to Warren, for supplying said whisky, amongst other libations.

Finally, Gillian at the risk of being terribly clichéd, there are no words that adequately express my thanks to you. We have shared this arduous journey, and it has been all the richer for sharing it with you.

TABLE OF CONTENTS

ABSTRACT i
ACKNOWLEDGEMENTS iii
LIST OF FIGURES x
LIST OF TABLES xii

1. Chapter One

Introduction

1.1. THE GREAT DIVIDE IN EVOLUTION 1
1.1.1. Molluscs: A Tool 3
1.1.2. Molluscs and molecules 3
1.1.3. Comparative Analysis 5
1.1.4. Species Delimitation 5
1.1.5. Molecular Clocks and Rates of Molecular Evolution 6
1.1.6. The New Zealand marine mollusc fauna 7
1.1.7. A summary of the New Zealand Cenozoic fossil record 8
1.1.8. Volutes in New Zealand 12
1.1.9. The taxonomic context of *Alcithoe* 14
1.1.10. A history of Volutes in New Zealand 14
1.2. THESIS STRUCTURE 15
1.3. REFERENCES 17

2. Chapter Two

Phylogenetic Informativeness of Genes; Illustrated With Mitochondrial Data From a Genus of Volute Mollusc

2.1. INTRODUCTION 25
2.2. MATERIALS AND METHODS 29
2.2.1. Taxon Sampling 29
2.2.2. DNA Extraction and Amplification 30
2.2.3. Sequence Analysis and Phylogenetic Reconstruction 32
2.3. RESULTS 34
2.3.1. Out-group selection 34
2.3.2. Sequence data 35

2.3.3. Phylogeny from the complete nucleotide dataset 37
2.3.4. Summary statistics 38
2.3.5. Partition heterogeneity 38
2.3.6. Gene trees 39
2.3.7. Spectral analysis 40
2.3.8. A reduced taxon set clarifies six closely related species 45
2.3.9. Refinement of analysis 47
2.4. DISCUSSION 49
2.4.1. Taxon subsets exclude unnecessary noise 50
2.4.2. Marker selectivity 50
2.4.3. *Alcithoe* Systematics 51
2.5. REFERENCES 52

3. Chapter Three

The importance of correct identification of fossil species

3.1. INTRODUCTION 57
3.1.1. Species delimitation 57
3.1.2. Fossil calibration of molecular clock analysis 58
3.1.3. Taxonomic history of *Alcithoe wilonsae* and *Alcithoe knoxi* 59
3.2. MATERIALS AND METHODS 62
3.2.1. Samples 62
3.2.2. DNA extraction 64
3.2.3. PCR amplification and sequencing 64
3.2.4. Phylogenetic reconstruction and population structure 65
3.2.5. Dated analysis 66
3.3. RESULTS 67
3.3.1. Sequence Data 67
3.3.2. Population Genetic Structure 69
3.3.3. Key dates in the history of *Alcithoe wilsonae* 70
3.3.4. Demographic History of *Alcithoe wilsonae* 71
3.4. DISCUSSION 72
3.4.1. Species identification in the fossil record. 74
3.5. REFERENCES 75

4. Chapter Four

Molecular Clock Analysis of Alcithoe

4.1. INTRODUCTION 79
- 4.1.1. Calibrating the molecular clock 80
- 4.1.2. Molecular clock models 81
- 4.1.3. Nucleotide substitution models 82
- 4.1.4. Speciation or Tree prior 82
- 4.1.5. Model discrimination 83
- 4.1.6. Rates of evolution 83
- 4.1.7. The New Zealand Cenozoic mollusc fossil record 84

4.2. MATERIALS AND METHODS 85
- 4.2.1. Fossil species calibration priors 85
- 4.2.2. Sequence data 88
- 4.2.3. Molecular clock analysis 88

4.3. RESULTS 89
- 4.3.1. Fossil species calibration priors 89
- 4.3.2. Outgroup calibration 89
- 4.3.3. Tree root calibration 89
- 4.3.4. Speciation priors 92
- 4.3.5. Nucleotide substitution models 94
- 4.3.6. Molecular clock model 96
- 4.3.7. Tree topology 98
- 4.3.8. Analysis of divergence times in *Alcithoe* 101
- 4.3.9. Rates of molecular evolution in *Alcithoe* 103

4.4. DISCUSSION 105
- 4.4.1. Parameter testing 105
- 4.4.2. Divergence patterns in *Alcithoe* 106
- 4.4.3. The tempo of molecular evolution in *Alcithoe* 108

4.5. REFERENCES 108

5. Chapter Five

The evolution of New Zealand Volutes: comparison of molecular and paleontological evidence

5.1. INTRODUCTION 113
5.1.1. *Alcithoe* 115
5.1.2. Systematics of the fossil Alcithoini 116
5.1.3. Species patterns relevant to modern taxa 117
5.2. METHODS 118
5.3. RESULTS 119
5.3.1. Patterns in molecular and paleontological datasets 119
5.3.2. Rates from different datasets 122
5.3.3. Trends 125
5.4. DISCUSSION 127
5.4.1. Patterns in the Paleontological and Molecular Data 127
5.4.2. Disparate absolute rate estimates 128
5.4.3. Long-term trends 129
5.4.4. Concluding remarks 129
5.5. REFERENCES 130

6. Chapter Six

Reconciling paleontology and molecular biology

6.1. HOW HAVE THE *ALCITHOE* HELPED? 133
6.1.1. Species identification 133
6.1.2. Phylogenetic inference 134
6.1.3. Rocks and clocks 135
6.1.4. Biodiversity patterns 135
6.2. WHERE TO FROM HERE? 136
6.2.1. Expanding the *Alcithoe* dataset 136
6.2.1.1. Complete sampling of *Alcithoe* 136
6.2.1.2. Refinement of morphological analysis 136
6.2.1.3. From population rates to species rates 137
6.2.2. Markers 138
6.2.2.1. Development of nuclear markers 138
6.2.2.2 Complete mitochondrial genome sequences 138

6.2.2.3. Splits to assess markers 139
6.2.2.4. Multiscale analysis 139
6.2.3. More illuminative answers through more refined questions 140
6.2.3.1. What effect do extinction patterns have on divergence date estimates? 140
6.2.3.2. What exactly is a different rate of evolution? 140
6.2.3.3. Getting more information with more specific questions 141
6.2.4. The way forward in the synthesis of molecular evolution and paleontology 142
6.2.4.1. More than just age calibrations 142
6.2.4.2. Use fossil data a test of the quality of molecular clock data 142
6.3. REFERENCES 143

7. Appendix 1

High-throughput sequencing and de novo assembly of mitochondrial genomes: a mollusc example 147

LIST OF FIGURES

FIGURE 1.1 The New Zealand geological timescale 9
FIGURE 1.2 Summarised trends of biodiversity and environmental change in the New Zealand marine habitat during the Cenozoic 10
FIGURE 1.3 Examples of New Zealand Volutidae 13

FIGURE 2.1 Mitochondrial gene arrangement in the New Zealand marine mollusc genus *Alcithoe* 32
FIGURE 2.2 Phylogeny to establish the molecular context of *Alcithoe* within Volutidae 35
FIGURE 2.3 Molecular phylogeny of 11 species of the New Zealand marine mollusc genus *Alcithoe* 37
FIGURE 2.4 Splits network of *Alcithoe* species based on the complete dataset (7822 bp) 41
FIGURE 2.5 Summed split support from partitioned data reveals the relative amounts of signal and noise contained in each partition of mitochondrial DNA sequence from *Alcithoe* species 42
FIGURE 2.6 Refinement of phylogenetic inference for New Zealand *Alcithoe* species by consideration of a sub-tree only 46
FIGURE 2.7 Alternative phylogenetic topologies in a quality controlled dataset show that the relationships of *A. fusus* and *A. larochei* cannot be resolved with this data 78
FIGURE 2.8 Molecular phylogeny of the gastropod genus *Alcithoe* based on a reduced gene dataset with maximised signal to noise 49

FIGURE 3.1 Morphological forms of *Alcithoe wilsonae* and *Alcithoe knoxi* 60
FIGURE 3.2 Approximate geographic range of *A. wilsonae* and *A. knoxi* 61
FIGURE 3.3 A Neighbor-Joining consensus tree based on 701 bp of mitochondrial *nad2* recovers the monophyly of *A. wilsonae* and *A. knoxi* 68
FIGURE 3.4 Haplotype network of 35 *Alcithoe wilsonae* and *Alcithoe knoxi* specimens 69
FIGURE 3.5 Molecular clock analysis indicates that *A. wilsonae* diverged around 9 millon years ago 71
FIGURE 3.6 Bayesian skyline analysis reveals a population expansion in *A. wilsonae* 72

FIGURE 4.1 Construction of calibration prior distributions for molecular-clock analysis of *Alcithoe* 87
FIGURE 4.2 Posterior distributions for joint-prior tests (no DNA data) of alternative calibration regimes for internal fossil calibrated nodes 90
FIGURE 4.3 A zero-offset lognormal prior distribution on the root-node calibration is more robust than other root-node calibrations 91
FIGURE 4.4 Divergence time estimation under a Yule prior, a birth/death process and a uniform prior 94
FIGURE 4.5 The proportion of invariable site is the most important parameter of the nucleotide substitution model for accurate estimation of divergence times for the *Alcithoe* 95
FIGURE 4.6 A relaxed exponential clock model generates a high variance on inferred node ages 97
FIGURE 4.7 Bayesian phylogeny of the *Alcithoe* recovers high posterior probability support for all nodes 99
FIGURE 4.8 Inferred rates on branches in alternative topological arrangements of *A. fusus* and *A. larochei* 100
FIGURE 4.9 Time scaled molecular phylogeny of *Alcithoe* 101
FIGURE 4.10 Inferred rate ranges for branches of the *Alcithoe* tree exhibit considerable variance 104

FIGURE 5.1 Revised generic relations for the New Zealand Alcithoini 117
FIGURE 5.2 Direct comparison of the molecular and paleontological data for the tribe Alcithoini 120
FIGURE 5.3 A lineage through time plot reveals an increasing speciation rate from the dated molecular phylogeny of the living *Alcithoe* 123
FIGURE 5.4 A dynamic survivorship analysis indicates the extinction rates for the fossil Alcithoini and the fossil *Alcithoe* 124
FIGURE 5.5 Rates derived from a sliding scale molecular analysis show that both extinction and speciation rates in *Alcithoe* have increased during the Cenozoic 125
FIGURE 5.6 Origination and extinction rate trends derived from within stage rates indicates increasing rate during the Cenozoic 126

LIST OF TABLES

TABLE 2.1 Volute species used to study the phylogenetic information in 9 mitochondrial genes 30
TABLE 2.2 Primers used to amplify mtDNA of volute gastropods 31
TABLE 2.3 Summary of sequenced DNA fragments from 13 marine molluscs of the family Volutidae 36
TABLE 2.4 Summary statistics from alignments of mitochondrial DNA sequence data partitions for 13 volute species 39

TABLE 3.1 Samples 63
TABLE 3.2 Results of AMOVA 70

TABLE 4.1 Bayes factors for all model comparisons carried out 93
TABLE 4.2 Posterior probabilities that the substitution rate on the *A. fusus* branch is faster then other branches of the *Alcithoe* phylogeny 104