A mtDNA study of aspects of the recent evolutionary history and phylogeographic structure of selected teleosts in coastal environments of south-western Australia.

By

Richard James Hoddell

B. Sc. (Biological Sciences)
James Cook University, Queensland, 1994

Grad. Dip. ASOS (with Hons.)
Institute of Antarctic and Southern Ocean Studies, Tasmania, 1996

This thesis was submitted for the degree of Doctor of Philosophy at Murdoch University

2003
Declaration

I declare that this thesis is my own account of my research and contains, as its main content, work that has not previously been submitted for a degree at any tertiary education institution. To the best of my knowledge, the thesis contains no material previously published or written by another person, except where due reference is made.

Richard James Hoddell
On the shoulders of giants....

“The affinities of all beings of the same class have sometimes been represented by a great tree. As buds give rise by growth to fresh buds and these, if vigorous, branch out and overtop on all sides many a feebler branch, so by generation I believe it has been with the great Tree of Life, which fills with its dead and broken branches the crust of the Earth and covers the surface with its ever branching and beautiful ramifications”

Charles Darwin, 1859

“In the study of dispersal and distribution of animals, it is important to see that the physical conditions lead and that in a more or less definite succession the flora and fauna follow. Thus, the fauna comes to fit the habitat as a flexible material does a mould. The time is passed when faunal lists should be the aim of faunal studies. The study must not only be comparative, but genetic, and much stress must be laid on the study of the habitat, not in a static, rigid sense, but as a fluctuating or periodical medium”

Charles Adams, 1901

“Ford! There’s an infinite number of monkeys outside who want to talk to us about this script for Hamlet they’ve worked out”

Douglas Adams, 1979
Abstract

At present, there is a general lack of information regarding the spatial genetic architecture and genetic diversity of estuarine and coastal freshwater fish in Australia or about the interacting intrinsic, extrinsic and historical influences responsible for sculpting these patterns. This thesis represented the first investigation of the phylogeographic structure and recent evolutionary histories of teleost fishes from the coastal and estuarine environments of south-western Australia, using the resolution afforded by mtDNA sequence data. Available evidence indicated that, to different degrees, these species have limited potential for dispersal amongst local assemblages from different water bodies. As this theoretically reduces the confounding effects of recent gene flow on extant genetic structure, these fishes were well suited to studying the influences of historical factors. Historical influences were expected to be particularly profound, given that these coastal environments underwent massive modifications during Late Quaternary eustatic fluctuations.

The thesis consists of four major components, which explored different aspects of interspecific and intraspecific phylogeny and phylogeography of three teleost species, based on mtDNA control region and cytochrome b fragments. First, the relationship between the endemic, ‘strictly estuarine’ Leptatherina wallacei (Atherinidae) and the more widespread, ‘estuarine & marine’ L. presbyteroides was examined, with a view to establishing whether L. wallacei represents a monophyletic or polyphyletic lineage and whether this species was derived recently (i.e. in Holocene estuaries). Second, the phylogeographic structure and genetic diversity of L. wallacei were
investigated and compared with data from *L. presbyteroides*, with a view to using this information to interpret the recent evolutionary histories of each congener. Third, the divergence between assemblages of *L. wallacei* inhabiting two isolated coastal lakes was used to estimate a maximal substitution rate for the control region, which was then used to infer general time frames for the divergence between the two *Leptatherina* species and between the major phylogeographic partitions within each species. Fourth, investigations were initiated into phylogeographic patterns and levels of genetic diversity within and among assemblages of *Pseudogobius olorum* (Gobiidae) from several coastal lakes and an estuary.

Phylogenetic analyses indicated that the two *Leptatherina* species were characterised by exclusive and reciprocally-monophyletic lineages of haplotypes from both mtDNA regions, supporting the monophyletic origins of *L. wallacei*. Both *L. wallacei* and *L. presbyteroides* exhibited high levels of genetic diversity and extensive overall subdivision (*e.g.* $\Phi_{ST} = 0.691$ & 0.644 respectively for control region data). There was a profound phylogeographic break in both species between all conspecific assemblages from the lower west coast (LWC phylogroup) and all those from the south coast (SC phylogroup), which suggested the influences of shared extrinsic and/or historical factors. There was limited genetic structuring within the two major phylogroups of either *Leptatherina* species, apparently reflecting recent connectivity amongst local assemblages, with subsequent fragmentation and insufficient time for lineage sorting. However, two major phylogeographic breaks distinguished monophyletic control region phylogroups of *L. wallacei*
from the isolated coastal Lake Clifton and Lake Walyungup, consistent with their independent evolution following lacustrine entrapment during the Holocene.

The divergence between these two isolated lacustrine assemblages of Leptatherina wallacei formed the basis for an estimate of the maximal substitution rate of the control region. While these data were unable to provide a precise estimate of the actual rate of molecular evolution, all the evidence suggested that it was proceeding very rapidly. The maximal rate estimate of 172.3% lineage\(^{-1}\) My\(^{-1}\) was among the fastest ever reported. Based on this rate, the two Leptatherina species diverged at least 19Kya, thus rejecting a Holocene origin for L. wallacei. The divergence between the LWC and SC phylogroups of L. wallacei has been ongoing for at least 6Kya, while the equivalent divergence in L. presbyteroides has been ongoing for at least 11Kya. As the time frames of these divergences were consistent with periods of massive environmental modifications associated with the end-Pleistocene fall in sea level and the HMT, it was likely that these factors have played important roles in sculpting the species’ divergence and intra-specific genetic structure. Although useful in temporally scaling genetic divergences within and between the two Leptatherina species, wider application of this rate estimate to questions regarding other taxa was limited. For example, evident rate heterogeneity between the genera precluded its use with even the relatively closely-related atherinid Atherinosoma elongata.
Phylogeographic analyses identified high levels of genetic diversity and extensive genetic subdivision (e.g. $\Phi_{ST} = 0.652$ for control region) amongst an estuarine and several lacustrine assemblages of *Pseudogobius olorum*, although phylogeographic structure was shallower than in either *Leptatherina* species. There was increased divergence between three assemblages from the lower west coast and two from the south coast, consistent with the profound break evident in the *Leptatherina*. One lacustrine assemblage appeared to represent a distinct lineage and a preliminary maximal rate estimate (~61.4% lineage$^{-1}$ My$^{-1}$) was calculated based on the minimum divergence of this assemblage from its nearest conspecifics. Although slower than the rate calculated for *L. wallacei*, this was still high for teleost fishes.

Overall, this study indicated that historical environmental factors, especially those related to Quaternary eustatic changes, have played important roles in sculpting the phylogeography and evolution of three teleost species from south-western Australia. Moreover, as these species have differential dependencies on estuarine environments (*i.e.* ‘strictly estuarine’ vs ‘estuarine & marine’) and represented two different taxonomic groups (*i.e.* Atherinoidei & Gobioidae), historical environmental factors may have exerted similar influences on other coastal species in the region.

**Key words:** phylogeography, teleost, estuaries, isolated coastal lakes, sea level changes, Pleistocene, Holocene Marine Transgression (HMT), mtDNA, control region, cytochrome b, local molecular clock, concordance
# Table of contents

Declaration  \( \text{ii} \)

Abstract  \( \text{iv} \)

Table of contents  \( \text{viii} \)

List of Figures & Tables  \( \text{xii} \)

Acknowledgements  \( \text{xvi} \)

## Chapter 1: General Introduction

1.1 Influences of coastal environments on fish phylogeography  \( \text{1} \)

1.2 Coastal environments & fishes of south-western Australia  \( \text{2} \)

   1.2.1 Physical environmental structure during the Quaternary  \( \text{2} \)

   1.2.2 Coastal ichthyofauna of south-western Australia  \( \text{6} \)

1.3 Molecular techniques in evolutionary studies  \( \text{14} \)

   1.3.1 Development & application of molecular techniques  \( \text{14} \)

   1.3.2 The use of mtDNA in phylogeny & phylogeography  \( \text{15} \)

   1.3.3 Estimating time scales for phylogenies  \( \text{16} \)

   1.3.4 A caution regarding molecular data  \( \text{18} \)

1.4 Aims, objectives & significance of the research  \( \text{18} \)

## Chapter 2: General Materials & Methods

2.1 Introduction  \( \text{21} \)

2.2 Sampling sites & methods  \( \text{21} \)

   2.2.1 Isolated coastal lakes  \( \text{21} \)

   2.2.2 Species collected  \( \text{23} \)

   2.2.3 Sampling methods  \( \text{29} \)

2.3 Laboratory methods  \( \text{30} \)

   2.3.1 DNA extraction  \( \text{30} \)

   2.3.2 PCR amplification  \( \text{30} \)

   2.3.3 mtDNA sequencing  \( \text{31} \)

2.4 General data analyses  \( \text{33} \)

   2.4.1 Consensus sequences & alignments  \( \text{33} \)

   2.4.2 Haplotype & nucleotide diversities  \( \text{35} \)
Chapter 3: Inferring the evolutionary relationships of the atherinids *Leptasterina wallacei* & *L. presbyteroides* from coastal water bodies in south-western Australia, based on mtDNA

3.1 Introduction

3.1.1 Role of estuaries in genetic structure & evolution of fish 36
3.1.2 Biology of the *Leptasterina* in south-western Australia 37
3.1.3 Aims & objectives 40

3.2 Materials & Methods 42

3.2.1 Sample Collection 42
3.2.2 DNA extraction, PCR amplification & sequence generation 42
3.2.3 Data analyses 42

3.3 Results 47

3.3.1 Sequence characteristics, haplotype identification & diversity 47
3.3.2 Genetic distance & interspecific divergence 58
3.3.3 Evolutionary relationships between the *Leptasterina* species 65
3.3.4 Proposed time frame for speciation 68

3.4 Discussion 73

3.4.1 Is *Leptasterina wallacei* monophyletic or polyphyletic? 73
3.4.2 Did *L. wallacei* diverge in the Holocene or the Pleistocene? 77

Chapter 4: Recent evolutionary history of the estuarine teleost *Leptasterina wallacei* in south-western Australia, inferred from phylogeographic partitioning & patterns of mtDNA diversity, with reference to the congeneric *L. presbyteroides*

4.1 Introduction 82

4.1.1 Evolutionary significance of coastal environments 82
4.1.2 Genetic structure of coastal teleosts 84
4.1.3 Historical influences on coastal teleosts of the region 88
4.1.4 Analysis & interpretation of phylogeographic data 88
4.1.5 Aims & objectives 90

4.2 Materials & Methods 92

4.2.1 Origins of fishes & sequences 92
4.2.2 Genetic diversity & phylogeography 92
4.2.3 Temporal scaling 96

4.3 Results 98
4.3.1 Overview 98
4.3.2 Variation within sampling locations 98
   4.3.2.1 Leptatherina wallacei 98
   4.3.2.2 Leptatherina presbyteroides 100
4.3.3 Variation &/or subdivision amongst locations 102
   4.3.3.1 Leptatherina wallacei 102
   4.3.3.2 Leptatherina presbyteroides 119
4.3.4 Phylogeographic structure 121
   4.3.4.1 Leptatherina wallacei 121
   4.3.4.2 Leptatherina presbyteroides 134
   4.3.4.3 Summary 139
4.3.5 Temporal scaling 142

4.4 Discussion 144
4.4.1 Overall subdivision & genetic distinctiveness of assemblages 144
4.4.2 Phylogeographic structure of both Leptatherina species 147
4.4.3 Proposed chronology for sculpting of phylogeography 155
4.4.4 Does L. wallacei represent an incipient species complex? 159

Chapter 5: Estimating the rate of molecular evolution of the mtDNA control region of Leptatherina wallacei, based on the genetic divergence between assemblages in two isolated coastal lakes

5.1 Introduction 161
5.1.1 Utility of lacustrine assemblages for rate estimation 161
5.1.2 Relevant aspects of the molecular clock theory 164
   5.1.2.1 Definition & limitations of the molecular clock 164
   5.1.2.2 Local molecular clocks & sources of error 165
   5.1.2.3 Relevance of published clocks for this study 167
5.1.3 Specific aims of this study 171

5.2 Materials & Methods 172
5.2.1 Origins of fishes & mtDNA sequences 172
5.2.2 Characterisation of haplotype suites 172
5.2.3 Estimating rates of molecular evolution 175

5.3 Results 176
5.3.1 Characteristics of haplotype suites 176
Chapter 6: Genetic structure & aspects of the recent evolutionary history of the goby *Pseudogobius olorum* from some isolated coastal lakes in south-western Australia

6.1 Introduction
6.1.1 Genetic consequences of inhabiting isolated coastal lakes
6.1.2 Genetic structure of teleosts in south-western Australia
6.1.3 Aspects of the biology of *Pseudogobius olorum*
6.1.4 Aims & objectives

6.2 Materials & Methods
6.2.1 Sampling design & sampling collection
6.2.2 Specimen processing to sequence generation
6.2.3 Data analyses

6.3 Results
6.3.1 Overview of goby sequence data
6.3.2 Variation within assemblages
6.3.3 Variation & subdivision amongst assemblages
6.3.4 Phylogeography & evolutionary relationships
6.3.5 Assessing potential for estimating rates of divergence

6.4 Discussion
6.4.1 Diversity & phylogeography of *Pseudogobius olorum*
6.4.2 Estimating the rate of divergence between lineages
6.4.3 Historical influences on phylogeography

Chapter 7: General Discussion & Conclusions

7.1 Overview of thesis
7.2 Divergence of the *Leptatherina* species 246
7.3 Phylogeography & recent evolutionary history 250
  7.3.1 Overall subdivision 250
  7.3.2 Regional-scale phylogeography 252
  7.3.4 Local scale phylogeography 255
  7.3.3 Isolated lacustrine populations & ‘clock’ rate estimates 258
7.4 Conclusions 261

**Appendices**

A GenBank Accession numbers 263
B Geographic distances between sampling locations 271

**References** 273

**List of Figures & Tables**

**Chapter One:**
**Figures**
Fig. 1.2.1 Approximate sea level curve for last 110Ky 3
Fig. 1.2.2 Effects of sea level changes on coastal morphology 4
Fig. 1.2.3 Photograph of *Leptatherina wallacei* 10
Fig. 1.2.4 Photograph of *Leptatherina presbyteroides* 10
Fig. 1.2.5 Distributions of both *Leptatherina* species 10
Fig. 1.2.6 Biological drawing of *Pseudogobius olorum* 13
Fig. 1.2.7 Distribution of *Pseudogobius olorum* 13

**Chapter Two:**
**Figures**
Fig. 2.2.1 Map of sampling locations across south-western Australia 26
Fig. 2.2.2 Maps of Lake Clifton & Lake Walyungup 27
Fig. 2.2.3 Maps of Lake Jasper, Donnelly River & Lake Richmond 28
Fig. 2.4.1 mtDNA showing sequenced fragments 33

**Chapter Three:**
**Figures**
Fig. 3.3.1 Nucleotide composition of cytochrome *b* haplotypes 62
Fig. 3.3.2 Nucleotide composition of control region haplotypes 62
Fig. 3.3.3  Sliding window analyses of control region substitutions  56
Fig. 3.3.4  Transitions vs transversions of cytochrome b  59
Fig. 3.3.5  Transitions vs transversions of control region  62
Fig. 3.3.6  Distance-based phylogeny of cytochrome b  69
Fig. 3.3.7  Parsimony-based phylogeny of cytochrome b  70
Fig. 3.3.8  Distance-based phylogeny of control region  71
Fig. 3.3.9  Parsimony-based phylogeny of control region  72

Tables
Table 3.3.1  Summary of sequenced atherinid individuals  58
Table 3.3.2  Variable nucleotides of Leptatherina cytochrome b  61
Table 3.3.3  Variable nucleotides of L. presbyteroides control region  63
Table 3.3.4  Variable nucleotides of L. wallacei control region  64
Table 3.3.5  Mean pairwise genetic distances & substitution types  61
Table 3.3.6  Summary of mean interspecific genetic distances  64

Chapter Four:
Figures
Fig. 4.3.1  MSN for L. wallacei control region haplotypes  123
Fig. 4.3.2  Nested design for L. wallacei control region (1)  124
Fig. 4.3.3  Nested design for L. wallacei control region (2)  125
Fig. 4.3.4  MSN & NCA for L. wallacei cytochrome b  130
Fig. 4.3.5  MSN & NCA for L. presbyteroides control region  135
Fig. 4.3.6  Nested design for L. presbyteroides control region (1)  136
Fig. 4.3.7  Nested design for L. presbyteroides control region (2)  137
Fig. 4.3.8  MSN & NCA for L. presbyteroides cytochrome b  140
Fig. 4.3.9  Geographic distributions of Leptatherina haplotypes  141

Tables
Table 4.3.1  Haplotype diversity of L. wallacei control region  104
Table 4.3.2  Pairwise differentiation of L. wallacei control region  106
Table 4.3.3  Haplotype diversity of L. wallacei cytochrome b  107
Table 4.3.4  Pairwise differentiation of L. wallacei cytochrome b  108
Table 4.3.5  Haplotype diversity of L. presbyteroides control region  109
Table 4.3.6  Pairwise differentiation of L. presbyteroides control region  110
Table 4.3.7  Haplotype diversity of L. presbyteroides cytochrome b  111
Table 4.3.8  Pairwise differentiation of *L. presbyteroides* cytochrome *b* 113
Table 4.3.9  AMOVA results for *L. wallacei* control region 114
Table 4.3.10 AMOVA results for *L. wallacei* cytochrome *b* 115
Table 4.3.11 AMOVA results for *L. presbyteroides* control region 116
Table 4.3.12 AMOVA results for *L. presbyteroides* cytochrome *b* 117
Table 4.3.13 NCA results for *L. wallacei* control region 127
Table 4.3.14 NCA inferences for *L. wallacei* control region 128
Table 4.3.15 NCA results for *L. wallacei* cytochrome *b* 131
Table 4.3.16 NCA inferences for *L. wallacei* cytochrome *b* 131

Chapter Five:
Figures
Fig. 5.2.1 Relative rates test 173
Fig. 5.3.1 Control region MSN for Lakes Clifton & Walyungup 179
Fig. 5.3.2 Sliding window analyses of control region substitutions 186

Tables
Table 5.1.1 Rates of evolution in various DNA markers 169
Table 5.3.1 Pairwise substitutions amongst both lake haplotype suites 183
Table 5.3.2 Summary of genetic distances 184
Table 5.3.3 Relative rates test results 187
Table 5.3.4 Tajima & Fu tests of selective neutrality 188
Table 5.3.5 Rate estimates for control region 189

Chapter Six:
Figures
Fig. 6.3.1 Phylogenetic trees for *P. olorum* control region 228
Fig. 6.3.2 MSN & nesting design for *P. olorum* control region 229
Fig. 6.3.3 Higher-level nested clades of *P. olorum* control region 230
Fig. 6.3.4 Phylogenetic trees for *P. olorum* cytochrome *b* 233
Fig. 6.3.5 MSN & nesting design for *P. olorum* cytochrome *b* 234

Tables
Table 6.3.1 Summary of sequenced goby individuals 214
Table 6.3.2 Variable nucleotides in control region of *P. olorum* 215
Table 6.3.3 Average pairwise genetic distances 216
Table 6.3.4 Variable nucleotides in cytochrome *b* of *P. olorum* 217
Table 6.3.5  Control region haplotype diversity  219
Table 6.3.6  Pairwise differences for control region  220
Table 6.3.7  Cytochrome b haplotypes diversity  221
Table 6.3.8  Pairwise differences for cytochrome b  221
Table 6.3.9  Control region AMOVA  225
Table 6.3.10 Cytochrome b AMOVA  226
Table 6.3.11 NCA inferences for control region  232

Chapter Seven:

Figures
Fig. 7.2.1 Late Quaternary MSL & ETD for Leptatherina  247

Tables
Table 7.2.1 Summary of genetic structure & phylogeography  249
Acknowledgements

Above all, I would like to thank my principal supervisor, Dr. Jennie Chaplin, for supplying invaluable advice and guidance throughout my research and the evolution of this thesis. I am grateful to Professor Ian Potter for his support during the project, which was conducted under his Directorship of the Centre for Fish & Fisheries Research, Murdoch University. I would also like to thank Dr. Howard Gill for his comments on some draft material and for various helpful discussions over the last few years.

This project could not have been completed without the assistance of various members of the Fish Research Group at Murdoch University. I especially thank Sorcha Gillen, Steve Beatty and Bryn Farmer for each accompanying me on the larger field trips; Dr. Dave Morgan and Steeg Hoeksema for providing certain important fish samples; Alex Hesp, Fiona Valesini and Ertug Seizmis for their assistance on some day trips; and Carina Marshall for letting me run week-long PAUP analyses on her computer.

I am also indebted to various other members of the academic and administrative staff in the School of Biological Sciences and the State Agricultural and Biotechnology Centre (SABC) at Murdoch University, most notably Ms. Frances Brigg, for ably commanding the SABC’s ABI sequencer.

In addition, I would like to thank my family and friends for their company, understanding, support and encouragement throughout: my mum Jenny, my sister Suzanne, Charles Barton, Byron & Bethany Tovey, Patrick Smith, Jeremy Nash, Bevan & Marija Clark, Ben & Michelle Lewis, Steve Schupp and the rest of the usual suspects.

This study was primarily supported through a Murdoch University Research Scholarship stipend and annual maintenance grant. In addition, Fishcare WA provided a small grant that supported much of the Lake Jasper and Donnelly River research.