"About 2.4 million years ago the build-up of polar ice triggered a cooling phase, plunging the world into a sustained ice age that ended only 10,000 years ago. Today its imprint shapes our modern landscape" – Graeme Stevens, *Prehistoric New Zealand*

*Fox Glacier, 2005*

Image Courtesy of Glacier Country Tourism Group
PHYLOGEOGRAPHY OF SOUTHERN NEW
ZEALAND: IMPLICATIONS OF PLIOCENE AND
PLEISTOCENE PROCESSES AND
EVOLUTIONARY CONCORDANCE ACROSS
INDEPENDENT LOCI

BY SHAY B. O’NEILL

A Thesis Submitted in Fulfillment of the Requirement for the Degree of
Master of Science at Victoria University
April 2008
Acknowledgements

I would like to thank my supervisor Peter Ritchie for all his support and encouragement throughout the past two years. Taking on a student, just completing an undergraduate degree in Biomedical Science, for a Masters project focusing on molecular ecology and evolution is a risky business but we made it in the end. Thanks Peter for always being there to answer my many questions. I would also like to thank my co-supervisor, Thomas Buckley, whose concepts, support and comments on drafts for my thesis proved invaluable. Thomas’s understanding of the concepts regarding phylogenetic analysis was an enormous help. I would also like to thank Thomas for organizing and sending me all the *Niveaphasma* samples from Auckland. I would like to thank Dave Chapple who acted as a supervisor throughout my research on the skinks. Having no prior knowledge of phylogeography, Dave’s knowledge on the subject and supervision got me through the difficult parts at the start. I would like to thank Dave also, for providing revisions on Chapter Two.

Thanks to Steph Greaves who was patient enough to run me through all the lab protocols required for this thesis. I had no prior knowledge of phylogenetic analysis before commencing this project and Steph taught me many of the basics for programs like PAUP* and MODELTEST. Thanks also Steph, for helping with the analysis in Chapter Two.

Thanks to Benno Kappers and Andrew Townsend for providing access to the Department of Conservation’s BioWeb Herpetofauna database. Raymond Coory provided access to ethanol preserved specimens housed at Te Papa (National Museum of New Zealand, Wellington), and Karen Britton and Sue Keall provided access to samples in the National Frozen Tissue Collection (NFTC) held at Victoria University of Wellington. Trent Bell provided additional *O. maccanni* tissue samples. Thanks to Tony Jewell for his assistance in collecting *Niveaphasma* samples. I would also like thank Lorraine Berry at the Allan Wilson Centre Genome Service (AWCGS) for her sequencing work. Thanks to Glacier County Tourism Group, Geographx (NZ) Ltd, Te Ara – Encyclopedia of New Zealand and Thomas Buckley who provided me with photos and maps.

Thanks to the Allan Wilson Centre (AWC), the Victoria University of Wellington University Research Fund, the Society for Research on Amphibians and Reptiles in New Zealand (SRARNZ), the New Zealand Marsden Fund and the George Mason Charitable Trust and Puke Ariki for funding.

I would like to thank everyone from the lab – Whitney, Mei Lin, Jen, Kim, Kristina, Mike, Hilary, Gemma, Lara, Leon and Libby – who provided comments on results and answered various other questions I raised over the last two years.

Finally I would like to thank my friends and family for all the support over the last two years. Thanks to my father Mike O’Neill and my Nana for their enthusiasm collecting
stick insects. Thanks also to Andrew for proofreading my thesis and to both Andrew and Claudia for making home such a comfortable place to be over the last two years.
My thesis is focused on two central research projects, presented in Chapter Two and Chapter Three. Both chapters are presented in a manuscript format intended for publication, with Chapter Two already submitted to *Molecular Phylogenetics and Evolution*. In order to maintain the concise manuscript format of Chapters Two and Three, I have presented background material in Chapter One and discussed future prospects for research in Chapter Four.

The contributions of each author to each manuscript must also be recognized. The concept and design of Chapter Two was prepared by Dave Chapple and Peter Ritchie. Sample collection was organized and carried out by Charles Daugherty. The concept and design of Chapter Three was prepared by Thomas Buckley and Peter Ritchie. Sample collection was organized by Thomas Buckley and carried out by Thomas Buckley and Tony Jewell. My contribution as main author was the collection and analysis of data and the writing of both manuscripts.
Abstract

The endemic fauna of the South Island has proven to be an ideal taxonomic group to examine the impact of climatic and geological processes on the evolution of New Zealand’s biota since the Pliocene. This thesis examines the phylogeography of McCann’s skink (*Oligosoma maccanni*) in order to provide insight into the relative contribution of Pliocene and Pleistocene processes on patterns of genetic structure in South Island biota. This thesis also investigates the phylogeography of the brown skink (*O. zelandicum*) to examine whether Cook Strait landbridges facilitated gene flow between the North and South Island in the late-Pleistocene. This thesis also investigates the presence of genealogical concordance across independent loci for the endemic alpine stick insect, *Niveaphasma*. I obtained mitochondrial DNA (mtDNA) sequence data (ND2 and ND4; 1284 bp) from across the range of both skink species and mtDNA (COI; 762 bp) and nuclear sequence data (EF1α; 590 bp) from across the range of *Niveaphasma*. I used DGGE in order to resolve nuclear EF1α alleles and examined phylogeographic patterns in each species using Neighbour-Joining, Maximum Likelihood and Bayesian methods. Substantial phylogeographic structure was found within *O. maccanni*, with divergences among clades estimated to have occurred during the Pliocene. Populations in the Otago/Southland region formed a well-supported lineage within *O. maccanni*. A genetic break was evident between populations in east and west Otago, while north-south genetic breaks were evident within the Canterbury region. There was relatively minor phylogeographic structure within *O. zelandicum*. Our genetic data supports a single colonization of the North Island by *O. zelandicum* from the South Island, with the estimated timing of this event (0.46 Mya) consistent with the initial formation of Cook Strait. There was substantial genetic structuring identified within *Niveaphasma*, with a well-supported lineage present in the Otago/Southland region. There was also a genetic break between populations in Canterbury and eastern Otago with those in central Otago and Southland. The genetic data provided strong genealogical concordance between mtDNA haplotypes and nuclear alleles suggesting an accurate depiction of the historical isolation identified between the major clades of *Niveaphasma*. This finding offers compelling evidence for the use of nuclear gene phylogeography alongside mtDNA for future evolutionary studies within New Zealand.
# Table of Contents

**Acknowledgements**  
Page iii

**Preface**  
Page v

**Abstract**  
Page vi

**Table of Contents**  
Page vii

**List of Tables**  
Page x

**List of Figures**  
Page xi

**List of Abbreviations**  
Page xiii

## 1 Chapter One: General Introduction  
Page 1

1.1 New Zealand’s Geological Past  
Page 1

1.2 Phylogeography  
Page 3

1.2.1 Mitochondrial DNA  
Page 4

1.2.2 Nuclear DNA  
Page 6

1.3 Phylogeographic Studies within New Zealand  
Page 9

1.4 Phylogenetic Inference  
Page 11

1.4.1 Alignment  
Page 11

1.4.2 Model Choice  
Page 12

1.4.3 Tree Building  
Page 14

1.4.4 Assessing Tree Reliability  
Page 17

1.4.5 Statistical Parsimony  
Page 19

1.5 New Zealand’s Endemic Taxa  
Page 20

1.5.1 Skinks  
Page 21

1.5.2 Stick Insects  
Page 22

1.6 The Significance of Phylogeography  
Page 23

1.7 References  
Page 24

## 2 Chapter Two: Phylogeography of Two New Zealand Lizards: McCann’s Skink (*Oligosoma maccanni*) and the Brown Skink (*O. zelandicum*)  
Page 37
2.1 Introduction

2.2 Materials and Methods

2.2.1 Taxonomic Sampling

2.2.2 DNA Extraction, Amplification and Sequencing

2.2.3 Phylogenetic Analysis

2.2.4 Estimating Divergence Times

2.2.5 Hypothesis Testing

2.3 Results

2.3.1 Oligosoma maccanni

2.3.2 Oligosoma zelandicum

2.4 Discussion

2.4.1 The Relative Contribution of the Pliocene and Pleistocene Processes in Shaping Genetic Structure in Oligosoma maccanni

2.4.2 Taxonomic Status of the ‘Garston’ Skink

2.4.3 Did Terrestrial Taxa Use Cook Strait Landbridges? Insights from Oligosoma zelandicum

2.5 References

2.6 Tables

2.7 Figure Legends

2.8 Figures

3 Chapter Three: Evolutionary Concordance across Independent Loci for the Niveaphasma Stick Insect Complex

3.1 Introduction

3.2 Materials and Methods

3.2.1 Sample Collection

3.2.2 DNA Extraction, Amplification and Sequencing

3.2.3 DGGE Analysis

3.2.4 Recovery of Bands from DGGE Gels and Sequence Analysis

3.2.5 Data Analysis

3.2.6 Phylogenetic Analysis

3.3 Results

3.4 Discussion

3.4.1 Mitochondrial DNA Indicates Pliocene and Pleistocene Processes have Shaped the Genetic Structuring of Niveaphasma Populations

3.4.2 The Distribution of Nuclear Gene Alleles Supports the Observations in Mitochondrial DNA

3.4.3 Niveaphasma shows Genetic Structuring within Canterbury and Eastern Otago
3.4.4 Evidence of Historical Isolation for Central Otago and Southland Populations of *Niveaphasma* 99

3.4.5 A Possible Geographical Intermediate Zone Observed for *Niveaphasma* in Eastern Otago 101

3.4.6 Genealogical Concordance across Independent Loci Indicates Substantial Genetic Structuring for *Niveaphasma* 103

3.5 References 104

3.6 Tables 114

3.7 Figure Legends 125

3.8 Figures 127

4 Chapter Four: General Discussion 133

4.1 General Discussion 133

4.1.1 Determining Genealogical Concordance 133

4.1.2 New Directions for Molecular Evolutionary Studies 135

4.2 References 139

Appendix A

Mismatch Distribution Graphs and a Mantels Test Plot for Genetic Distance versus Geographic Distance for *Niveaphasma* 143

Appendix B

Electropherograms of *Niveaphasma* Samples following PCR and a Photograph of Resolved *Niveaphasma* Alleles following DGGE and Silver Staining 145
List of Tables

Chapter 2

Table 1. Locality information and GenBank accession numbers for *O. maccanni* and *O. zelandicum* samples used in this study 65

Table 2. Oligonucleotide primers used for *Oligosoma* samples 70

Table 3. Estimates of genetic diversity and population statistics for *O. maccanni* and *O. zelandicum* 71

Table 4. Uncorrected and Tamura-Nei (TrN) corrected genetic distances among *O. maccanni* clades 72

Chapter 3

Table 1. Sample location, haplotype and allelic pair information for *Niveaphasma* samples used in this study 114

Table 2. Oligonucleotide primers used for *Niveaphasma* samples 120

Table 3. Sequence statistics for mitochondrial and nuclear loci of *Niveaphasma* 121

Table 4. Genetic distances estimated for the COI and EF1α data sets using HKY + I+Γ and TVM + I+Γ substitution models respectively 122

Table 5. Mismatch analysis results for the COI and EF1α data sets 123

Table 6. Position of nucleotide changes for the 19 alleles determined from the EF1α sequence 124
List of Figures

Chapter 1

Figure 1. Schematic diagram of the animal mitochondrial genome 4

Figure 2. EF1α during the initiation stage of translation 7

Figure 3. Photograph of the brown skink, Oligosoma zelandicum 21

Figure 4. Photograph of Niveaphasma 22

Chapter 2

Figure 1. Distributional range of O. maccanni and O. zelandicum obtained from the Department of Conservation BioWeb Herpetofauna database (2006). The inset shows the main provincial regions of the South Island and lower North Island 74

Figure 2. Locality of O. maccanni and O. zelandicum tissue samples used in this study, with the approximate distribution of the O. maccanni clades also shown. The inset shows the coastline of New Zealand during the last glacial maximum 75

Figure 3. Bayesian consensus tree for O. maccanni and O. zelandicum based on 1282 bp of combined ND2 and ND4 mitochondrial DNA 76

Chapter 3

Figure 1a. Bayesian consensus tree for Niveaphasma based on 762 bp of COI mitochondrial DNA 127

Figure 1b. Bayesian consensus tree for Niveaphasma based on 590 bp of EF1α nuclear DNA 128

Figure 2a. Location of Niveaphasma tissue samples used in this study, with the approximate distribution of the three major clades and subclades identified from Figure 1a. The inset shows the main provincial regions of the South Island 129
Figure 2b. Approximate distribution of the 21 EF1α alleles identified from Figure 1b. The inset shows some major geographical features of the South Island

Figure 3a. Statistical parsimony network for the EF1α alleles of Niveaphasma

Figure 3b. Statistical parsimony network for the COI haplotypes of Niveaphasma

Appendix A

Figure 1. Mismatch distribution graphs for Niveaphasma mtDNA COI pairwise differences

Figure 2. Mantels test plot of mtDNA COI genetic distance versus geographic distance for Niveaphasma

Appendix B

Figure 1. Gel electropherogram of Niveaphasma COI PCR products following DNA amplification

Figure 2. Gel electropherogram of Niveaphasma EF1α alleles following DGGE analysis

Figure 3. Photograph of resolved Niveaphasma EF1α alleles following DGGE analysis and silver staining
List of Abbreviations

A   adenine
AIC  Akaike Information Criterion
BIC  Bayesian Information Criterion
bp   base pairs
BSA  bovine serum albumin
C   cytosine
cm   centimeters
COI  Cytochrome Oxidase I
χ²   chi-squared
ddH₂O double distilled water
df   degrees of freedom
DGGE  denaturing gradient gel electrophoresis
dNTP deoxyribonucleotide triphosphate
EF1a  elongation factor 1 alpha
G   guanine
GD   genetic distance
GTP  guanosine triphosphate
Γ   gamma distribution
h   number of haplotypes
Hd   haplotype diversity
HKY  Hasagawa, Kishino and Yano model
hLRT  hierarchical likelihood ratio test
hrs  hours
I   invariant sites
invgamma  invariant sites and gamma distribution
ITS  internal transcribed spacers
JC69  Jukes and Cantor 1969 model
kb   kilobases
km   kilometers
km²   square kilometers
LGM  last glacial maximum
-ln L log likelihood
m   meters
M   molar
MCMC  Markov chain Monte Carlo
mg   milligrams
min  minutes
ML   Maximum Likelihood
ml   milliliters
mm   millimeters
mM   millimolar
MP   Maximum Parsimony
mtDNA  mitochondrial DNA
My   million years
Mya  millions of years ago
μL   microliters
μM   micromolar
n   number of mutations
N   number of individual sequences obtained
ND2  NADH dehydrogenase subunit 2
ND4  NADH dehydrogenase subunit 4
Nₑ  effective population size
Nₑ(mito) mitochondrial effective population size
Nₑ(nuclear) nuclear effective population size
NZAC  New Zealand Arthropod Collection
π nucleotide diversity
PCR polymerase chain reaction
r raggedness
rRNA ribosomal RNA
S number of segregating sites
sec seconds
spp. species
SSD sum of the squared deviation
SVL snout vent length
T thymine
TBR tree-bisection-reconnection
TrN Tamura-Nei model
tRNA transfer RNA
TS/TV transition/transversion ratio
TVMef transversion model with equal nucleotide frequencies
UV ultraviolet
Chapter One

General Introduction

1.1 New Zealand’s Geological Past

The ancestral land mass which eventually became New Zealand broke away from the super continent of Gondwana c. 80 million years ago (Mya). From that period onwards New Zealand has been in isolation, taking with it a unique range of animals and plants including the tuatara, peripatus and the kauri. Throughout the past 80 million years (My) New Zealand has been shaped by major geological processes (Stevens et al., 1995) with dramatic effects on the flora and fauna residing here.

Due to its archipelagic nature, fluctuations in worldwide sea levels have had a significant effect on New Zealand (Cooper and Millener, 1993). From the mid-Eocene to the mid-Oligocene (45 – 25 Mya) it is understood that much of New Zealand was underwater. Some suggest that the emergent part of New Zealand shrank from an area similar to that of today to only 18% of its present area. This
drastic reduction and reduced environmental diversity is thought to have resulted in a severe biological bottleneck within the country (Cooper and Cooper, 1995; Boyer et al., 2007).

During the Pliocene (5 – 2 Mya) the South Island underwent intense tectonism leading to the formation of the Southern Alps (Fleming, 1975). This prominent mountain range is a result of the thrusting of the Pacific Plate crust over the Australian Plate along the Alpine Fault. It is estimated that during the last 5 My convergence of the two plates has led to 50 km of shortening across the plate boundary (Batt and Braun, 1999), significantly altering the geography of the South Island and giving rise to new alpine flora and fauna (Wardle, 1963). Although tectonic uplift is responsible for the overall shape of the Southern Alps, much of the contemporary relief results from extensive glacial modification throughout the Pleistocene (2 – 0.1 Mya) (Cox and Findlay, 1995; Winkworth et al., 2005). The build up of ice on the Southern Alps led to increased rates of erosion, and ice flowing eastwards from Fiordland spilled out across inland Southland and Otago creating valleys as they passed (Stevens et al., 1995). These environmental processes are expected to have left a significant impression on species diversity, distributions and genetic variation on many of the South Island’s endemic taxa (Fleming, 1979).

Volcanism has been prominent in shaping the North Island over the last 2 My. The Central Plateau is a volcanic pile with three active mountain volcanoes (Ruapehu, Tongariro and Ngauruhoe), forming a dominant feature of the North Island. Catastrophic eruptions during the last 1.6 My created massive basin like depressions known as calderas, such as Lake Taupo (Worthy and Holdaway, 2002). The recent Taupo eruption about 1850 years ago deposited pyroclastic rock over 20,000 km² and covered 30,000 km² of New Zealand in ash (Wilson and Walker, 1985;
Healy, 1992). As a result, all animal and plant populations within 80 km of Lake Taupo are thought to have arrived over the last 1850 years (Morgan-Richards et al., 2000).

1.2 Phylogeography

Within the field of comparative biology, systematics is seen as relating to form; paleontology and embryology deal with time; and biogeography stresses space (Crisci et al., 2003). Phylogeography was introduced as a way of dealing with various aspects of these concepts so that time and space are jointly considered axes of phylogeography, onto which are mapped particular gene genealogies of interest (Avise, 2000). A more simple definition describes phylogeography as a scientific field concerned with the spatial distributions of genealogical lineages, including those within species (Avise, 2006). The role of phylogeographers is to interpret the extent and mode by which historical processes in population demographics have left evolutionary footprints on the contemporary geographic distributions of gene-based organismal traits (Avise, 2000). Like all scientific fields, phylogeography requires input and incorporates methods from many other fields, including ecology, molecular biology, population genetics, conservation biology, ethology, biogeography and geology (Avise, 2000; Crisci et al., 2003).

The current basis for phylogeography focuses strongly on the use of neutral molecular markers, highlighting the role of genetic drift and migration as the key forces that have shaped contemporary polymorphisms and the differences among populations (Jobling et al., 2004). Based on the neutral theory of evolution, the role of positive or balancing selection is considered minimal and the rate of substitution
used for molecular clocks is driven solely by the mutation rate. (Avise, 2000; Jobling et al., 2004).

1.2.1 Mitochondrial DNA

Just over 70% of phylogeographic studies utilize mitochondrial DNA (mtDNA) as the primary or exclusive molecular marker (Avise, 2000) and at times, the field has been criticized for being overly reliant on a single gene system (Bermingham and Moritz, 1998). However, to a certain degree, mtDNA meets all the criteria of an ideal molecular system to conduct phylogenetic analysis (Avise et al., 1987).

The animal mtDNA genome contains 37 genes whose products functions include energy production and mitochondrial protein synthesis (Jobling et al., 2004; Figure 1). These genes are packed within a mitochondrial genome varying in size from only 6 kb up to over 2000 kb in length and contain a very minute region of noncoding sequence known as the ‘D-loop’ or control region (Page and Holmes, 1998; Jobling et al., 2004). The mtDNA molecule is distinctive and ubiquitously distributed so comparisons can be made among a wide range of organisms (Avise et al., 1987). Mitochondrial DNA has a high substitution rate, giving it enough polymorphic characters in order to assay diversity efficiently in many samples. Mitochondrial DNA has a haploid genome and transmission is only possible through the female member of a species, thereby reducing the effective population size ($N_e$). Under a neutral model of evolution only genetic drift acts on mtDNA, thus enabling
measurable divergences between isolated populations to occur (Hare, 2001). The mtDNA genome is a relatively simple structure, free of repetitive DNA (except in the control region), transposable elements and introns and, for the most part, does not undergo recombination (Avise et al., 1987; Page and Holmes, 1998; Avise, 2000; Jobling et al., 2004). It also has a high copy number per cell, facilitating the analysis of degraded contemporary samples and ancient DNA (Jobling et al., 2004).

The stability of mtDNA gene arrangement is supported by studies showing the arrangement of the mitochondrial origin of replication and the small and large ribosomal RNA (rRNA) genes, are the same in organisms as divergent as humans, mice, rats and frogs (Attardi et al., 1976, Dawid et al., 1976, Kroon et al., 1977, Battey and Clayton, 1978, Brown et al., 1979). These studies show the heavy mtDNA strand of one species hybridizes exclusively with the light strand from a second species, even for species as distantly related as humans and frogs. Mitochondrial genome size is similar across species such as rodents and primates. Also, the number of genes within the mitochondrial genome is identical across many species, allowing the general assumption that mtDNA is free of additions or deletions (Brown et al., 1979). The high mutation rate of mtDNA is attributed to an inefficient editing function of the mtDNA replication complex and the high turnover rates of mtDNA in tissues provides more rounds of replication for errors to occur (Brown et al., 1979).

Recently paternal transmission and evidence for recombination has been observed in mtDNA, but the incidence is so low it is generally ignored in phylogenetic analysis (Gerber et al., 2001). It is also important to note that rates of mtDNA evolution are not uniform across different lineages, thereby requiring local calibrations for the correct use of molecular clocks (Gerber et al., 2001).
Mitochondrial DNA is also assumed to be a neutral marker adhering to the theory of neutral evolution (Cann et al., 1984; Avise et al., 1986). This assumption is based on the idea that the majority of substitutions in mtDNA are synonymous, with insertions or deletions occurring only at noncoding regions, thereby having no differential effect on an organism’s fitness (Avise et al., 1987; Gerber et al., 2001). This assumption is not entirely true with evidence of purifying selection coming from studies on disease, observing disorders in humans and mice resulting from mutations in mtDNA (Wallace, 1999). Another problem surrounding the assumption of neutral evolution in mtDNA is that mtDNA is believed to be susceptible to introgression (the movement of genes between species via hybridization and backcrossing) making it heavily dependant on the selective forces acting on linked sites (Machado and Hey, 2003). The concern is that selective sweeps, background selection and genetic hitchhiking act on mtDNA, reducing the levels of variability within a species (Page and Holmes, 1998).

1.2.2 Nuclear DNA

In the past, various technical and biological hurdles have hindered the use of nuclear genes in the retrieval of the evolutionary history of a species (Avise, 2000). Recently, the use of nuclear DNA in phylogeography has gained momentum with researchers’ identifying the need for agreement between unlinked markers in order to infer the evolutionary history of a population or species (Pamilo and Nei, 1988; Wu, 1991; Avise, 2000).

The nuclear genome contains the complete set of chromosomes found in each nucleus of a given species and includes all the genetic material (Lodish et al., 2000). Nuclear genome size can vary 300,000-fold among eukaryotic organisms, with the
largest genomes containing nearly all non-genic secondary DNA (Cavalier-Smith, 2005). The nuclear genome of a typical eukaryote can contain about 25,000 protein coding genes and can be made up of three billion nucleotide pairs, which is anywhere between 15,000 and 30,000 times that observed in the mitochondrial genome (Page and Holmes, 1998).

A high rate of recombination in nuclear DNA is a problem for phylogenetic studies. Nuclear haplotypes generally have more than one immediate ancestor, meaning different segments within that haplotype therefore have an independent evolutionary history (Hare, 2001). Ignoring this fact can ruin any molecular clock approach and undermine phylogenetic reconstruction (Hare, 2001). There are techniques used which are able to build nuclear gene trees despite recombination. One simple approach is to remove any detected recombinants, if they are small enough in number, before conducting a phylogenetic analysis (Harding et al., 1997). It has also been shown that most recombination exists at concentrated sites along a DNA sequence. Once identified, these recombination hotspots can be removed and phylogenetic analysis can be performed on partitioned sequences either side of these sites (Templeton et al., 2000). A third solution is to construct a network which is better suited to the reticulated relationships of haplotypes generated by recombination (Posada and Crandall, 2001). It has also been found that if a population is isolated for long enough, gene flow becomes restricted to within that
population, thus preventing recombination between populations and enabling phylogenetic structure to increase between those populations (Bernardi et al., 1993).

Nuclear gene phylogeographic structure is expected to be low because it has a higher effective population \( (N_e) \) than mtDNA. The understanding is that, in a random mating population with an equal ratio of males to females, four times more nuclear gene copies are passed on to the next generation because the nuclear genome is diploid and alleles are contributed by both sexes (Hoelzer, 1996). This means under a neutral theory of evolution, genetic drift causes divergence between isolated populations to evolve at least four times more slowly for nuclear loci than mtDNA loci (Moore, 1995; Avise, 2000; Hare, 2001). It has been found that departure from a random mating system has the potential to increase the ratio of \( N_e(\text{mito})/N_e(\text{nu}) \) to values greater than unity. For example, the increased variance in male reproduction success in polygnous mating systems reduces \( N_e(\text{nu}) \) without affecting \( N_e(\text{mito}) \) (Nunney, 1993). Another solution to the low phylogeographic structure found in nuclear DNA is to focus on noncoding regions where variation is higher, or focus on nuclear loci for which a great deal is already known about its structure and variability. Examples include \( \beta \)-globin or Elongation Factor 1 \( \alpha \) (EF1\( \alpha \)), which is involved in eukaryotic 43S pre-initiation complex formation (Figure 2).

Containing a diploid genome means the PCR products of nuclear genes have two copies and the presence of heterozygotes raises difficulties in resolving phylogenetic signals (not an issue for mtDNA because it is haploid). However, these issues can be overcome by using techniques such as Denaturing Gradient Gel Electrophoresis (DGGE), which allows double stranded PCR products that differ by one or more mutation to be separated on acrylamide gels with linear gradients of denaturants such as urea and formamide (Lessa, 1992; Yergeau et al., 2005). This is
possible because mutations or polymorphisms alter the DNA segment’s resistance to
denaturation, resulting in changes in mobility on gradient gels (Lessa, 1992). Ten
years ago, one of the major challenges for phylogeography was to utilize unlinked
molecular markers and improve analytical approaches testing for evolutionary
congruence between nuclear and organelle genes (Bermingham and Moritz, 1998).
Today this challenge is less daunting than previously thought.

1.3 Phylogeographic Studies within New Zealand

Since its introduction in 1987 by John C. Avise, phylogeography has had a huge
impact on evolutionary studies throughout the world. This is especially true of
evolutionary studies in New Zealand, where the country’s unique biota and dynamic
geological history make it an interesting and challenging system in which to conduct
phylogeographic study (Fleming, 1979; Trewick and Wallis, 2001).

The idea that areas within New Zealand are characterized by different levels
of biodiversity was proposed decades ago (Cockayne, 1926; Wardle, 1963;
McGlone, 1985). A study by Daugherty et al. (1990) was one of the first in New
Zealand to use genetic variation by way of allozyme analysis in tuatara (Sphenodon
spp.) to illustrate how genetic lineages were associated with different geographic
regions. Since then, with the aid of advances in molecular genetic techniques,
phylogeography in New Zealand has gathered a great deal of momentum.

Phylogeographic studies in New Zealand have helped shed light on some
extremely important questions in regards to alternative historical scenarios
accounting for the spatial arrangements of organisms (Avise, 2000). Many studies
have revolved around the glacial and Alpine Fault hypotheses which make quite

Another aspect of New Zealand which makes phylogeographic research such an interesting prospect relates more to the scientists doing the research than to the biota. Like many scientific fields, there can be much dispute between scientists over possible hypotheses and historical scenarios. Phylogeographic research is no different this regard. One recent example was the “beech gap” debate where it was observed that many endemic taxa are absent from the central narrow waist of the
South Island. This is typified by the absence of southern beech (*Nothofagus*) from the area – the “beech gap” (Wallis and Trewick, 2001). Phylogeographic studies on the issue gave two contrasting conclusions, one centering on glaciations as the cause of the disjunction (Burrows, 1965; Wardle, 1965) and the other claiming tectonic movements as the cause (Haase, 1999). The debate raised by such findings increases the interest towards phylogeographic study, bringing in more funding and increasing the standards of research conducted within the field. Such outcomes, combined with the large array of endemic taxa within New Zealand still to be investigated, is creating many opportunities to conduct phylogeographic research within the country.

1.4 Phylogenetic Inference

1.4.1 Alignment

The initial process of phylogenetic analysis begins with the collection of a data set containing a series of nucleotide sequences for a specified mitochondrial or nuclear gene of interest. At this point, utmost care is taken in regard to the sampling and collection of these sequences (Sanderson and Shaffer, 2002). Many studies will include nucleotide sequences taken directly from GenBank, but it must be emphasized that GenBank is an unmonitored repository for sequence data. For that reason, it may include specimen data that have been incorrectly identified, sequences of dubious quality, specimens of unknown origin or a combination of all three (Linton *et al.*, 2002).

Having obtained sequence data for a number of individuals, the next stage is identifying the evolutionary changes between those sequences. In such a case, a multiple alignment is performed specifying the homology between each nucleotide in
the set of sequences. Sequences must be aligned before any further comparative
analysis can be performed (Page and Holmes, 1998; Jobling et al., 2004).

Many researchers choose to manually align their sequences, but it is
recommended that such alignments are routinely checked with multiple sequence
alignment programs such as CLUSTALW (Thompson et al., 1994). Mistakes can arise
with manual alignments inferring incorrect gaps or polymorphisms into data sets,
producing false results (Sanderson and Shaffer, 2002). Too often it is taken for
granted that the sequence data consists of a set of homologous characters and that
they are of mitochondrial origin. This is a serious problem considering the discovery
of mitochondrial like sequences in the nuclear genome. The presence of nuclear
pseudogenes has the ability to confuse phylogenetic analysis (Zhang and Hewitt,
1996). The presence of pseudogenes may be discovered when unexpected deletions,
insertions, frameshifts or stop codons occur. Other indicators for the presence of
pseudogenes are nucleotide sequences, which are drastically different from what is
expected and also tree phylogenetic positioning significantly different from the
expected result. To help identify and prevent such instances, the use of target-
specific primers and enriched mtDNA is recommended. Another recommendation is
to translate nucleotide sequences to detect the presence of nonsense codons which
would indicate the occurrence of pseudogenes (Zhang and Hewitt, 1996).

1.4.2 Model Choice

The reconstruction of molecular phylogenetic relationships seems a relatively simple
exercise. However, the nature of DNA sequence evolution and culmination of
molecular forces acting on sequences can make phylogenetic inference a complex
matter (Bos and Posada, 2005). For this reason, the choice of evolutionary model is
always critical and is made even more difficult if multiple genes are involved. Genes
evolve under different evolutionary constraints, so a single model of evolution may
be unsuitable. Thus separate models and parameters may need to be generated for
each gene (Jones et al., 2006).

The evolutionary model presents a framework through which subsequent
phylogenetic analysis is based. Models vary in complexity according to the number
of parameters they use. Model parameters can determine differences in nucleotide
frequencies, substitution rates (C ↔ T, A ↔ G) and among site variation (Bos and
Posada, 2005). It is generally assumed that a more complex model improves the fit of
the data but parameter estimates become worse as there is less data information
available per parameter. Another drawback is that complex models tend to be very
computationally expensive (Posada and Buckley, 2004).

A commonly used model choice program is MODELTEST (Posada and
Crandall, 1998), which selects the model of nucleotide substitution that best fits the
data. Substitution models include the Tamura-Nei (TrN) and Hasagawa, Kishino and
Yano (HKY) distance models. Models correct for multiple hits, taking into account
the differences in substitution rate between nucleotides and the inequality of
nucleotide frequencies (Tamura et al., 2007). Models distinguish between
transitional substitution rates between purines and transversional rates between
pyrimidines (Tamura et al., 2007). Variations of these models include providing
gamma (Γ) shape and proportion of invariable site (I) parameters. The program
chooses among 56 substitution models and implements three different model
selection frameworks: hierarchical likelihood ratio tests (hLRTs); Akaike
Information Criterion (AIC) (Akaike, 1973); and Bayesian Information Criterion
(BIC). The program used in tandem with MODELTEST is PAUP* (Swofford, 2002)
which is used to calculate the likelihood scores of the data based on the 56 substitution models. MODELTEST also determines levels of model uncertainty and allows for model averaging and calculation of parameter importance (Posada and Crandall, 1998). Of the three model selection frameworks applied in MODELTEST, the most widely accepted is the hierarchical likelihood ratio test (hLRT) (Huelsenbeck and Crandall, 1997). The hLRT compares the ability of a more complex model against a previous simpler model, hence comparing two models in a ‘hierarchical’ fashion. The basic principle is to ensure that the increase in a model complexity significantly improves our ability to explain the data (Jobling et al., 2004). Recently it has been argued that the AIC offers considerable benefits over hLRT, including the ability to simultaneously compare nested or non-nested models, assess model uncertainty and estimate model parameters using all available models (Posada and Buckley, 2004). Despite this, hLRTs continue to be the most commonly used algorithm for model choice (Posada and Buckley, 2004). Furthermore, as model selection grows their use will continue to spread in ecology and evolution, expanding the set of statistical tools available to researchers (Johnson and Omland, 2004).

1.4.3 Tree Building

The task of molecular phylogenetics is to use an evolutionary model to convert the information given in the sequence data into an evolutionary tree (Page and Holmes, 1998). In nearly all experimental situations, the true phylogenetic tree underlying the evolutionary history of the observed sequences is unknown (Li et al., 2000). Many methods for reconstructing evolutionary relationships between taxa have been described; the challenge is selecting the method best suited to the data (Page and Holmes, 1998; Li et al., 2000). Given the rapid development of tree building
methods, it is not possible to cover all of them in this chapter. The following passage briefly describes the two main methods regarding the treatment of the data during the tree building process: distance versus discrete methods. It then details two of the more commonly used and accepted discrete methods: Maximum Likelihood and Bayesian analysis.

Distance methods, such as Neighbour Joining, initially convert aligned sequences into a pairwise distance matrix then input that matrix into a tree building algorithm. Distance methods are based on the idea that if the actual evolutionary distance between members of a set of sequences were known, it would be relatively easy to reconstruct their evolutionary history (Page and Holmes, 1998). While distance methods are computationally simple, they draw two major criticisms. One is that summarizing a set of sequences by a pairwise distance matrix loses phylogenetic information. The other criticism is that observed branch lengths are not always accurate reflections of the evolutionary distances between sequences. However, branch lengths can be corrected by using a substitution model in order to estimate the ‘true’ distance (Page and Holmes, 1998; Holder and Lewis, 2003).

Sometimes the use of discrete methods is preferred over distance methods because discrete methods consider each nucleotide site directly. This means they avoid the loss of information that occurs when sequences are converted into distances. One of the most common discrete methods used in phylogenetic reconstruction is Maximum Likelihood (ML). The ML model chooses the value of the parameter and the tree that maximizes the probability of observing the data (Huelsenbeck and Crandall, 1997). Likelihood methods always require a specified model of sequence evolution. The analysis starts with a specified tree derived from the dataset and swaps branches on the starting tree until the tree with the highest
likelihood score is determined. The likelihood is a function of both the tree topology and branch lengths. Maximum Likelihood therefore allows an explicit examination of the assumptions made about sequence evolution (Harrison and Langdale, 2006). One of the criticisms of ML is its dependence on an explicit model of evolution (although this can also be viewed as a benefit). The question, however, is which model one should choose (Page and Holmes, 1998). Programs such as MODELTEST are able to perform model selection, thereby alleviating this problem. Another drawback is that ML analysis is computationally time consuming (Page and Holmes, 1998; Harrison and Langdale, 2006). However, increased computer speed and improved computer programs have reducing this problem somewhat (Huelsenbeck and Crandall, 1997).

Bayesian statistics are closely aligned with ML and, while Bayesian approaches to phylogenetics are relatively new, they are generating a great deal of excitement because the primary analysis not only estimates a tree, it also measures the uncertainty for the groups on the tree. Another benefit of Bayesian analysis is that it offers a faster way to assess support for trees than ML methods (Holder and Lewis, 2003). Bayesian inference of phylogeny is based on a quantity called the posterior probability. The posterior probability of a tree is best described as the probability that the tree is correct given the data and the model. It is obtained by combining the prior probability and likelihood for each tree using Bayes’s formula. The prior probability of a tree represents the probability of the tree before any observations have been made. It is generally assumed that all trees must be considered equally probable, a priori. The likelihood is proportional to the probability of the observations (the DNA sequences), requiring making some assumptions on how the sequences were generated via an evolutionary model (Huelsenbeck et al., 2001).
Chapter One: General Introduction

Calculating the posterior probability of a tree involves the summation over all possible trees and integration over all combinations of branch lengths and substitution-model parameters, which becomes impossible to calculate analytically. Therefore, the only option is to approximate the posterior probability, and the most useful method to do this is Markov chain Monte Carlo (MCMC). The Markov chain is a sequence of random values whose probabilities at a time interval depend upon the number of these values at the previous time. Monte Carlo refers to independent random sampling of the posterior probability distribution of trees (Huelsenbeck et al., 2002). The MCMC algorithm involves creating a new tree by stochastically perturbing the current or starting tree. The new tree is either accepted or rejected based on its probability. If the tree is accepted, it undergoes further perturbation. If the Markov chain is adequately run, the proportion of time that a tree is visited is a valid approximation of that tree’s posterior probability (Huelsenbeck et al., 2001).

One of the drawbacks of MCMC is assessing whether the Markov chain has run long enough to provide reliable estimates of the posterior probability (Holder and Lewis, 2003). The problem occurs because it is virtually impossible to determine if the Markov chain has converged on the target distribution or is simply trapped in a local peak in parameter space. The best solution to this problem is to start independent runs with random trees and focus on the integration over tree space. If several chains started from random points in parameter space produce identical samples, there is a strong indication that convergence has occurred (Huelsenbeck et al., 2002).

1.4.4 Assessing Tree Reliability
Following phylogenetic analysis of the sequence data, a tree is produced and the next step is determining how strongly the data supports each of the relationships in the tree (Holder and Lewis, 2003). A common tool for obtaining such confidence levels on phylogenies is bootstrapping. Bootstrapping involves randomly removing and replacing individual characters (DNA sequences) in the data matrix and replacing it with another character from the same data set to produce a pseudoreplicate data set. This process is repeated for a specified number of pseudoreplications and the proportion of trees in which a clade is recovered is presented as the bootstrap value for any given node on that tree (Holder and Lewis, 2003; Soltis and Soltis, 2003; Harrison and Langdale, 2006).

The interpretation of bootstrap values has been controversial. It is considered that bootstrap values of 95% or greater be recognized as statistically significant or taken to indicate support of a tree (Felsenstein, 1985). Some, however, have suggested that a value of 70% is a better indication of support (Hillis and Bull, 1993). A concern over bootstrapping is that it provides assessments of support for specific clades rather than joint confidence for the entire tree. Therefore, although support for several clades may be high, joint confidence in the interrelationship among clades cannot be properly determined (Soltis and Soltis, 2003). It is important to note that bootstrap values estimate precision rather than accuracy. It is possible for a node to have a high bootstrap value but still be completely wrong in terms of its phylogenetic positioning on the tree. If, for instance, the tree-building method infers an incorrect tree from the given data, a robust bootstrap value will support the wrong tree (Page and Holmes, 1998).

As discussed previously, Bayesian analysis provides the posterior probability value of a node as the means of best describing that the tree is correct (Huelsenbeck
et al., 2001). An advantage of the Bayesian approach is that it obtains tree support quicker than bootstrapping. The Bayesian approach yields a much larger sample of trees in the same computational time as bootstrapping because it produces one tree for every posterior probability distribution cycle compared to one tree per tree search in ML bootstrapping (Holder and Lewis, 2003). Despite their differences and similarities many researchers choose both methods to assess branch support on a phylogenetic tree. It should also be highlighted that many journals focusing on phylogenetic study insist that bootstrap or posterior probability values are included in all submitted papers.

1.4.5 Statistical Parsimony

In some instances, it is not always relevant to represent intraspecific gene evolution on a bifurcating tree. Often, population genealogies are multifurcated and ancestral haplotypes can persist in the population, allowing them to be sampled together with their descendants. Combining this with recombination produces reticulated relationships. For these reasons, several networking approaches have been developed. Networks are able to account for the population level phenomena described and they also provide a way of representing more of the phylogenetic information present in a data set. For instance, the presence of loops in a network may be due to recombination or they may be the product of homoplasies (Posada and Crandall, 2001).

One networking method is Statistical Parsimony, which implements the statistical parsimony algorithm presented by Templeton et al (1992). The process begins by estimating the maximum number of differences among haplotypes (the parsimony limit) resulting from single substitutions with 95% statistical confidence.
From that point, haplotypes differing by one change are connected, followed by those differing by two, three or more changes. Eventually all the haplotypes become connected into a single network. Statistical Parsimony emphasizes what is shared among haplotypes rather than the differences among them (Templeton et al., 1992; Posada and Crandall, 2001).

Networks are distance based methods, meaning some phylogenetic information may be lost as sequence data is converted into a distance matrix. The rooting of a network is also a problem, as outgroups are nearly always separated from the ingroup by several mutations, creating a lack of control in deciding where rooting should take place. Generally, the oldest ancestral haplotype is considered to be the root of the phylogeny. The position and number of connections of a haplotype can also be used to determine the root of a network (Posada and Crandall, 2001).

1.5 New Zealand’s Endemic Taxa

The terrestrial New Zealand fauna have been able to develop in isolation on an ancient landmass of continental origin since the Mesozoic. New Zealand’s continental origins mean the country harbors survivors of many ancient lineages. However, New Zealand’s fauna also resembles that of an isolated archipelago: many higher taxa are missing; some have undergone extensive radiations in situ; and levels of endemism approach 100% in many groups (Daugherty et al., 1993). The level of endemism attained by New Zealand’s plants and animals is among the highest in the world, an observation shared with other Southern Hemisphere lands like New Caledonia, Australia and South Africa. Common features of small isolated islands are levels of endemism far greater than those of continents. Moreover, the more
isolated the islands, the higher the proportion of unique life forms present and New Zealand is no exception (Gibbs, 2006).

Some notable features of the endemic species diversity in New Zealand are over 178 species of earthworms, more than 1000 species of snails and slugs, 1567 species of moths, 4300 species of beetles, over 60 species of lizards and nearly 200 species of birds (Daugherty et al., 1993). These levels of endemism make New Zealand a fascinating place for the study of evolution of a unique biota influenced by the changing patterns of climate, sea level changes and the growth and decay of mountains (Cooper and Millener, 1993).

1.5.1 Skinks

Skinks are a large family of lizards widely distributed in tropical and temperate regions of the world. They have flat, shiny body scales with small eyes and, notably, all New Zealand species have moveable lower eyelids (Gill and Whitaker, 1996). There are at least 29 endemic skink species in New Zealand, which fall into two groups, *Oligosoma* and *Cyclodina* (Daugherty et al., 1994; Patterson and Daugherty, 1995). The *Oligosoma* skinks (23 species) (maximum snout vent length (SVL) = 70 – 140 mm) tend to have shallow pointed heads, bodies that are oval in cross-section and long fingers and toes (Figure 3). They prefer open habitats and are active by day (except for *O. suteri*) and enjoy sun-basking. *Cyclodina* skinks (6 species) (SVL = 60 mm – 145 mm) have deeper and blunter heads, bodies that are squarish in cross-section and shorter fingers and...
Chapter One: General Introduction

Toes. They enjoy more shaded habitats, are active at dawn and dusk, and do not sun-bask (Gill and Whitaker, 1996). All endemic New Zealand skinks give birth to live young (except *O. suteri*). They are carnivorous, feeding mainly on arthropods, such as insects, but some species eat soft berries if available (Gill, 1986). New Zealand’s skinks are found throughout the country and on offshore islands, such as the Chatham Islands and the Mercury Islands.

Skinks have poor dispersal abilities and several species are distributed across important biogeographic barriers within New Zealand (Hardy, 1977). Studies within New Zealand have revealed that skinks show substantial phylogeographic patterns and are sensitive to barriers of gene flow (Berry and Gleeson, 2005; Greaves *et al.*, 2007).

1.5.2 Stick Insects

Stick insect is the popular name given to relatively large (New Zealand species range from 4 – 15 cm in length), flightless, elongate insects that share a remarkable similarity to twigs and small sticks. Stick insects belong to the insect order Phasmatodea, containing two families: the Phylliidae and the Phasmatidae. New Zealand has nine genera of stick insects (all endemic to New Zealand) and 22 recognized species all belonging to the family Phasmatidae (Salmon, 1991; Buckley, 2008). Stick insects are found throughout New Zealand (including outlying islands such as Little Barrier and
Kapiti) in forests, scrubland and often appear on garden plants. They are found in lowland areas and two species inhabit alpine zones of between 1000 – 1500 m, such as *Niveaphasma* (Figure 4). Most species within New Zealand are relatively common but their nocturnal nature and resemblance to their host vegetation can make them difficult to find (Jewell and Brock, 2002; Buckley, 2008). When disturbed stick insects become cataleptic, releasing their hold on the foliage and falling to the ground where they can remain motionless for up to an hour. One of the more peculiar behaviors of stick insects is the ‘dance’ in which the body is swayed from side to side by a flexing motion of the legs; the purpose of this performance is unknown (Salmon, 1991; Buckley, 2008). Stick insects can be identified by the arrangement of body spines, body size (males are smaller than females of the same species), structure of genitalia and morphology of the eggs. Some New Zealand species contain no males and females reproduce by a process known as parthenogenesis (Salmon, 1991; Buckley, 2008).

There are currently several research studies involving stick insects being conducted within New Zealand. These studies are using morphological and genetic methods in an attempt to determine species boundaries in taxonomically problematic species, as well as determining the impact of environmental changes on species radiation and evolution (Buckley, 2008).

### 1.6 The Significance of Phylogeography

With the advent of molecular techniques such as DNA sequencing and PCR, the genetic information housed within cells is now freely available to scientists. The genetic data collected from their work has enabled researchers to illustrate the
evolutionary relationships or phylogenies of various organisms (Avise, 2006). However, scientists seek more than just family trees and the application of phylogenetics spans many scientific fields, including ecology, microbiology, biosystematics and conservation. Phylogenetics was used to predict which strains of influenza present in the current population would be progenitors of future generations of the virus (Bush et al., 1999a). Another important application of phylogenetics is the plotting of geographical arrangements of species onto species phylogenies determined from DNA sequence data (Avise, 2006). An example of this was a study by Cann et al. (1987) looking at mtDNA sequences from people around the world to examine the current geographical distributions of human matrilineal lineages. This study discovered that modern humans originated once, and only once, on the African continent, sometime during the last few hundred thousand years (Avise, 2006). This finding has come to be known as the famous “out of Africa” hypothesis.

The following research thesis begins by providing an example of the application of mtDNA and phylogenetics towards our understanding of the historical processes involved in the genetic patterns, distribution and structuring of two endemic New Zealand skinks species (O. maccanni and O. zelandicum). The thesis then moves away from a traditional phylogeographic approach in order to incorporate nuclear gene analysis to achieve phylogeographical agreement across two unlinked loci in the endemic New Zealand stick insect complex, Niveaphasma.

1.7 References


Chapter One: General Introduction


Chapter One: General Introduction


Chapter One: General Introduction


Chapter One: General Introduction


Chapter One: General Introduction


Chapter Two

Phylogeography of Two New Zealand Lizards: McCann’s Skink (*Oligosoma maccanni*) and the Brown Skink (*O. zelandicum*)

2.1 Introduction

Intense tectonic activity, glacial cycles and dramatic fluctuations in sea-level have shaped the New Zealand archipelago and its landscape throughout the Pliocene and Pleistocene (Cooper and Millener, 1993; Markgraf *et al.*, 1995; Worthy and Holdaway, 2002). The New Zealand skink fauna has proven to be an ideal taxonomic group in which to examine the impact of these climatic and geological processes on the evolution of the New Zealand biota since the Pliocene (Berry and Gleeson, 2005; Greaves *et al.*, 2007, in press; Hare *et al.*, 2008; Chapple *et al.*, in press). New
Chapter Two: Phylogeography of Two New Zealand Lizards

Zealand supports a diverse endemic skink fauna, comprising 29 described species and several undescribed species in two genera, Oligosoma (23 species) and Cyclodina (6 species) (Gill and Whitaker, 2001; Chapple and Patterson, 2007). Several New Zealand skink species have widespread distributions that span recognised biogeographic barriers (Hardy, 1977; Towns et al., 1985). Lizards typically exhibit substantial phylogeographic structuring due to their low dispersal abilities (Avise, 2000). Since skinks have been present in New Zealand since at least the early Miocene (Hickson et al., 2000; Worthy et al., 2006; Smith et al., 2007), they represent an ideal taxonomic group in which to examine the influence of historical processes throughout the Pliocene and Pleistocene on gene flow and patterns of genetic structuring.

Two ongoing debates in New Zealand biogeography have focused on: i) whether Pliocene tectonism was a more important factor than Pleistocene glaciation in shaping the distribution and patterns of genetic structure in the South Island biota (McGlone, 1985; Wardle, 1988; McGlone et al., 2001; Trewick and Wallis, 2001; Greaves et al., 2007); and ii) whether landbridges that formed intermittently during glacial maxima in the late-Pleistocene enabled the interchange of terrestrial taxa between the North and South Islands (Lewis et al., 1994; Worthy and Holdaway, 2002; Greaves et al., 2007). In this study we examine the phylogeography of McCann’s skink (Oligosoma maccanni) in order to gain insight into the relative contribution of Pliocene and Pleistocene processes on patterns of genetic structure in the South Island biota. We also investigate the phylogeography of the brown skink (O. zelandicum) to examine whether Cook Strait landbridges facilitated gene flow between the North and South Islands in the late-Pleistocene.
Chapter Two: Phylogeography of Two New Zealand Lizards

The Southern Alps that characterize the South Island of New Zealand were formed by tectonic activity along the Alpine Fault line (boundary of the Australasian and Pacific plates) that commenced during the Miocene, and intensified during the Pliocene and early-Pleistocene (Gage, 1980; Suggate, 1982; Stevens et al., 1995). The presence of mountainous regions in the South Island has facilitated extensive glaciation during the Pleistocene, creating an expansive alpine zone, and fundamentally altered climatic conditions and prevailing weather patterns (Suggate, 1990; Pillans, 1991; Worthy and Holdaway, 2002). During Pleistocene glacial maxima, the merging of glaciers resulted in large ice sheets along the Canterbury Plains (Figure 1). Ice from Fiordland spilled out across the landscape of Southland and Otago, scouring out valleys as they passed (Stevens et al., 1995). Glaciation was patchy in the northernmost region of the South Island, with the Nelson region consisting mainly of valley glaciers (Stevens et al., 1995). Early authors tended to emphasise the importance of Pleistocene processes on shaping the South Island biota (e.g. Hardy, 1977; Wardle, 1988). Recent molecular studies suggest that patterns of genetic structure in the South Island fauna are attributable to Pliocene mountain building (Trewick et al., 2000; Buckley et al., 2001; Trewick, 2001; Arensburger et al., 2004; Trewick and Morgan-Richards, 2005; Buckley and Simon, 2007). However, several studies highlight a role for both Pliocene and Pleistocene processes in shaping patterns of genetic diversity (Lloyd, 2003; Smissen et al., 2003; Chinn and Gemmell, 2004; Neiman and Lively, 2004; Winkworth et al., 2006).

*Oligosoma maccanni* (maximum snout-vent length (SVL), 73 mm) occurs in the southern South Island, east of the divide from Canterbury to Southland at altitudes from sea-level to 1500 m (Patterson and Daugherty, 1990; Gill and Whitaker, 2001; Figure 1). It lives in open habitats, usually occupying dry, rocky
environments and is especially abundant in montane grassland (Gill and Whitaker, 2001; Whitaker et al., 2002). Given the substantial levels of phylogeographic structure evident in other taxa in the southern South Island, we predict that *O. maccanni* will exhibit patterns of genetic structuring consistent with both Pliocene mountain building (e.g. an east-west split in Otago similar to that observed in *O. grande*; Berry and Gleeson, 2005) and Pleistocene glacial cycles. Since there is substantial morphological variation within *O. maccanni* (Daugherty et al., 1990), we also examine the potential for an unrecognised species within *O. maccanni*. In particular, we assess the taxonomic status of the ‘Garston’ skink, known only from a single sub-adult specimen collected from the Garston region of Otago (R. Hitchmough, pers. comm.). Initial allozyme and morphological analysis of this specimen suggests that it may be a melanistic colour morph of *O. maccanni* (R. Hitchmough, pers. comm.), but we further assess its taxonomic status using mitochondrial DNA sequence data and topology tests.

The lower North Island was inundated throughout the Pliocene, with the Manawatu Strait occurring between present-day Taranaki/Hawke’s Bay and Wellington. However, there is geological evidence for the presence of intermittent landbridges between the North Island and South Island during the Pliocene (Lewis et al., 1994; King, 2000; Worthy and Holdaway, 2002). Tectonic activity in the late-Pliocene and early-Pleistocene resulted in the uplift of the lower North Island and the draining of the Manawatu Strait (Lewis et al., 1994; Worthy and Holdaway, 2002). Present-day Cook Strait (separating the North Island and South Island; Figure 2) was formed during the late-Pleistocene (~0.45 million years ago (Mya); Lewis et al., 1994). Sea-level fluctuations (up to 130 m below present) associated with Pleistocene glacial cycles are believed to have resulted in intermittent landbridges.
across Cook Strait during glacial maxima of the late-Pleistocene (Lewis et al., 1994; Stevens et al., 1995; Worthy and Holdaway, 2002). However, it has been suggested that no landbridge existed (or no landbridge suitable for dispersal by terrestrial taxa) during the last glacial maximum (LGM; Figure 2), with the most recent landbridge occurring ~0.13 Mya (Worthy and Holdaway, 1994, 2002). Despite the availability of geological evidence for repeated Cook Strait landbridges during the late-Pleistocene, few molecular studies have revealed empirical evidence consistent with gene flow or dispersal across these landbridges. Divergences across Cook Strait have been dated between the Pliocene and mid-Pleistocene, prior to the formation of Cook Strait in the late-Pleistocene (Baker et al., 1995; Buckley et al., 2001; Efford et al., 2002; Lloyd, 2003; Greaves et al., 2007). These studies support the suggestion that Cook Strait landbridges were not sufficiently long-lived to facilitate dispersal of terrestrial species between the North Island and South Island (Worthy and Holdaway, 1994, 2002).

*Oligosoma zelandicum* (maximum SVL 73 mm) occurs in the west of the North Island from Taranaki to Wellington, and in the South Island in Marlborough, Nelson and northern Westland (Gill and Whitaker, 2001; Figure 1). It lives in a variety of densely-vegetated and, usually damp, habitats from supralittoral scrubs and grassland to lowland forest (Gill, 1976; Gill and Whitaker, 2001). Its present-day distribution either side of Cook Strait makes *O. zelandicum* an ideal species to investigate whether Cook Strait landbridges (between Nelson/Marlborough and Taranaki; Figure 2) facilitated gene flow and dispersal between the North and South Islands during the late-Pleistocene. Recent molecular studies have indicated that grassland (or open habitat) *Oligosoma* species (*O. lineoocellatum, O. infrapunctatum*) have not utilised late-Pleistocene Cook Strait landbridges (Greaves
et al., 2007, in press). Hardy (1977) hypothesized that some New Zealand skink species retreated to forested refugial areas in the Nelson/Marlborough region during glacial maxima. A recent allozyme study revealed only minor genetic divergence across the range of *O. zelandicum* (Miller, 1999). *Oligosoma zelandicum* lives in forested areas, therefore we hypothesise that it was present in Nelson/Marlborough refugia during glacial maxima, and dispersed across Cook Strait landbridges to the North Island during the late-Pleistocene.

Here we use mitochondrial DNA sequence data (ND2, ND4; 1282 bp) to examine the phylogeography of *O. maccanni* and *O. zelandicum*. We use patterns of genetic structuring and divergence time estimates (ND2 molecular clock) to test specific hypotheses for the impact of historical processes during the Pliocene and Pleistocene on *O. maccanni* and *O. zelandicum*. In addition, we complete Shimodaira-Hasegawa topology tests to examine the taxonomic status of the ‘Garston’ skink.

### 2.2 Materials and Methods

#### 2.2.1 Taxonomic Sampling

We obtained tissue samples from across the entire distribution of *O. maccanni* (32 samples) and *O. zelandicum* (except the Taranaki region; 17 samples) (Table 1, Figures 1, 2). Samples were obtained primarily from the National Frozen Tissue Collection (NFTC; Victoria University of Wellington, New Zealand) and ethanol-preserved specimens housed at Te Papa (National Museum of New Zealand, Wellington). Four additional *O. maccanni* samples were collected for this study.
Chapter Two: Phylogeography of Two New Zealand Lizards

We included two New Caledonian *Eugongylus*-lineage skinks (*Nannoscincus mariei*, *Marmorosphax tricolor*) as outgroups in our study (Table 1).

2.2.2 DNA Extraction, Amplification and Sequencing

Total genomic DNA was extracted from liver, toe or tail samples using a modified phenol-chloroform extraction protocol (Sambrook *et al.*, 1989). For each sample we targeted portions of the mitochondrial ND2 (~600 bp) and ND4 genes (~850 bp; incorporating most of the flanking 3’ tRNA cluster, including the histidine and serine tRNA genes). These regions were chosen because work at comparable taxonomic levels in other squamate reptile groups has indicated useful levels of variability (e.g. Chapple *et al.*, 2004, 2005; Greaves *et al.*, 2007; Hare *et al.*, 2008). The primers used to amplify and sequence ND2 and ND4 are provided in Table 2. PCR and sequencing were conducted as outlined in Greaves *et al.* (2007).

2.2.3 Phylogenetic Analysis

Sequence data were edited using CONTIGEXPRESS version 10.0.1 (Invitrogen), and aligned using the default parameters of CLUSTALX (Thompson *et al.*, 1997). The aligned sequences were translated into amino acid sequences using the vertebrate mitochondrial code. This was performed to determine whether these data were truly mitochondrial in origin. As no premature stop codons were observed, we conclude that all sequences obtained are true mitochondrial copies. GenBank accession numbers for all sequences are provided in Table 1.

MEGA version 3.1 (Kumar *et al.*, 2004) was used to calculate Tamura-Nei (TrN) corrected genetic distances for the concatenated dataset (ND2 + ND4) and pairwise uncorrected genetic distances for the ND2 dataset. The number and
frequency of all unique mtDNA haplotypes, as well as summary statistics for estimates of molecular diversity, were obtained using DNAsp 4.1 (Rozas et al., 2003).

Neighbour-Joining (NJ) analyses, using the TrN correction, were conducted in PAUP* version 4.0b10 (Swofford, 2002). Model selection for Maximum Likelihood (ML) analyses was performed with hierarchical likelihood ratio tests (hLRT) using MODELTEST version 3.7 (Posada and Crandall, 1998). Base frequencies, substitution rates, the proportion of invariant sites (I), gamma distribution (G), and the among-site substitution rate were estimated in MODELTEST, with these values implemented in PAUP* to generate a ML tree. Due to the size of the dataset, the ML tree was found by heuristic searching via the tree-bisection-reconnection (TBR) method, with random addition of sequences to the dataset. Ten replicate searches were conducted. Only trees within 5% of the target likelihood score were subject to reiteration.

Bayesian analyses were completed using the program MrBayes version 3.1.2 (Huelsenbeck and Ronquist, 2001). We implemented parameters determined from the hLRT used in MODELTEST and PAUP*. We used the default value of four Markov chains per run, and ran the analysis for five million generations. To ensure that the analyses obtained a sampling of the full tree space, rather than becoming trapped in local optima, the analysis was run twice. Chains were sampled every 100 generations to obtain 50,000 sampled trees. The program TRACER version 1.3 (Rambaut and Drummond, 2003) was used to check for chain convergence. The first 12,500 sampled trees were discarded as the burn-in phase, with the last 37,500 trees used to estimate the Bayesian posterior probabilities. We used both bootstrap values and Bayesian posterior probabilities to assess branch support. Maximum-Parsimony
(MP) bootstraps (1000 replicates) were generated in PAUP*. We consider branches supported by bootstrap values greater than or equal to 70% (Hillis and Bull, 1993), and posterior probability values greater than or equal to 95% (Wilcox et al., 2002) to be well-supported by our data. We also constructed a haplotype network for \textit{O. zelandicum} using the method of Templeton \textit{et al.} (1992) in TCS version 1.21 (Clement \textit{et al.}, 2000).

\subsection*{2.2.4 Estimating Divergence Times}

To estimate the divergence time of lineages, we calibrated the evolutionary rate of ND2 by re-analysing the data from Macey \textit{et al.} (1998) for the agamid genus \textit{Laudakia}. Specifically, we re-calculated the evolutionary rate for \textit{Laudakia} using only the 507 bp fragment of ND2 used in the present study (e.g. Smith \textit{et al.}, 2007). We calculated average between-group nucleotide differences across each of the calibrated nodes from Macey \textit{et al.} (1998) (1.5, 2.5, 3.5 Mya). We then plotted them against time and used the slope of the linear regression to calculate a rate of evolution for our 507 bp fragment of ND2. This resulted in an evolutionary rate of 1.3% per million years (My) (0.65% per lineage, per My), the same as calculated by Macey \textit{et al.} (1998).

\subsection*{2.2.5 Hypothesis Testing}

We completed Shimodaira-Hasegawa tests in PAUP* (Shimodaira and Hasegawa, 1999; Goldman \textit{et al.}, 2000) using full optimization and 1000 replicates to examine the taxonomic status of the Garston skink. Allozyme evidence previously suggested that the Garston skink was a melanistic colour morph (darker colouration and glossier scales) of \textit{O. maccanni} in the Garston region of Otago (CHD, unpublished...
data). We tested the significance of the log-likelihood difference between our optimal ML tree (using the ML –ln L) and the alternative hypothesis that the Garston skink (OMA4) represents a distinct species.

### 2.3 Results

The edited alignment comprised 1282 characters (507 bp ND2, 775 bp ND4 + tRNAs), of which 425 (33%) were variable and 306 (24%) were parsimony-informative. For the ingroup only, the alignment contained 304 (24%) variable characters, of which 264 (21%) were parsimony-informative. Base frequencies were unequal \( \text{A} = 0.3510, \text{T} = 0.2397, \text{C} = 0.3040, \text{G} = 0.1054 \), but a \( \chi^2 \) test confirmed the homogeneity of base frequencies among sequences \( (\text{df} = 150, P = 1.0) \).

The hLRT from ModelTest supported the TrN + I + \( \Gamma \) substitution model as the most appropriate for our dataset \( (-\ln L = 5505.2729) \). Parameters estimated under this model were: relative substitution-rates \( \text{A} \leftrightarrow \text{C} = 1.0, \text{A} \leftrightarrow \text{G} = 33.55, \text{A} \leftrightarrow \text{T} = 1.0, \text{C} \leftrightarrow \text{G} = 1.0, \text{C} \leftrightarrow \text{T} = 14.01, \text{G} \leftrightarrow \text{T} = 1.00 \), gamma shape parameter \( (1.9090) \) and proportion of invariable sites \( (0.5972) \). The topologies of the NJ, MP, ML (optimal tree, \(-\ln L = 5475.96442\)) and Bayesian trees were very similar, therefore we present the mean consensus Bayesian tree with MP bootstrap values and posterior probabilities indicating branch support (Figure 3). The Bayesian tree strongly supports the monophyly of both \( O. maccanni \) and \( O. zelandicum \) (100 bootstrap and 1.0 posterior probability in both cases).

#### 2.3.1 Oligosoma maccanni
Substantial genetic structuring was evident within *O. maccanni*, with a mean pairwise uncorrected ND2 genetic distance (GD) of 4.1% and seven clades identified (Figure 3). Estimates of genetic diversity for *O. maccanni* are shown in Table 3. Haplotypic diversity was extremely high, with virtually every sample having a different haplotype (Table 3). Clade 1 is represented by a single sample (OMA27) from Frankton, near Queenstown in western Otago (Figure 2). Clade 2 (100 bootstrap, 1.0 posterior probability, GD = 0.2%, 0.15 Mya) encompasses samples from western Otago and western Southland, from the Remarkables to just south of Lake Te Anau (Figure 2). The ‘Garston skink’ (OMA4) is part of Clade 2, with our Shimodaira-Hasegawa topology test clearly rejecting (*P* < 0.0001) the hypothesis that it is distinct from *O. maccanni*. Clade 3 is represented by a single sample (OMA31) from Lake Hawea in north-western Otago (Figure 2). Clade 4 (0.79 posterior probability, but well-supported in the NJ and ML trees, GD = 1.2%, 0.92 Mya) comprises samples from eastern Otago and eastern Southland, from the Ida Range south to the Hokonui Hills (Figure 2). Clade 5 (100 bootstrap, 1.0 posterior probability, GD = 0.8%, 0.62 Mya) incorporates samples from central Canterbury, from Lake Tekapo south to Lake Pukaki (Figure 2). Clade 6 (100 bootstrap, 1.0 posterior probability, GD = 1.6%, 1.2 Mya) is represented by two samples from Banks Peninsula in Canterbury (Figure 2). Clade 7 (100 bootstrap, 1.0 posterior probability, GD = 0.2%, 0.15 Mya) encompasses two samples from south Canterbury, north of the Waitaki River (Figure 2).

Considerable genetic differentiation is present among *O. maccanni* clades, with genetic divergences ranging from 0.2% (0.15 Mya) to 6.2% (4.77 Mya) (Table 4). There is support for a close relationship between Clade 1 and Clade 2 (56 bootstrap, 1.0 posterior probability), with relatively little genetic divergence evident
between these two clades (0.2%, 0.15 Mya; Table 4). Although the subdivision of clades is generally north-south, there is a substantial east-west split in Otago and Southland (Figure 2). Substantial genetic differentiation exists (~4.7%, 3.6 Mya) between populations in western Otago/Southland (Clades 1 and 2) and eastern Otago/Southland (Clade 4). There is strong support for the monophyly (82 bootstrap, 1.0 posterior probability) of the Otago/Southland clades (Clades 1-4), but there is no support for the Canterbury populations representing a well-supported group.

2.3.2 Oligosoma zelandicum

Haplotype diversity in *O. zelandicum* was extremely high, with 12 unique haplotypes (from 17 samples) distinguished by 34 segregating sites and 34 mutations (Table 3). Relatively minor levels of phylogeographic structuring were evident across the range of *O. zelandicum* (maximum pairwise uncorrected ND2 GD = 1.2%; Figure 3). Genetic divergence among North Island populations (GD = 0.2%) was lower than that of South Island populations (GD = 0.4%). However, there was strong support for the monophyly of the North Island populations (85 bootstrap, 1.0 posterior probability; Figure 3). This indicates the presence of several mtDNA lineages in the South Island, but only a single mtDNA lineage in the North Island, suggesting a single colonisation of the North Island from the South Island. The level of genetic differentiation between the North Island and South Island populations (GD = 0.6%), suggests that *O. zelandicum* reached the North Island ~0.46 Mya.

2.4 Discussion
2.4.1 The Relative Contribution of Pliocene and Pleistocene Processes in Shaping Genetic Structure of Oligosoma maccanni

Substantial phylogeographic structure is evident within *O. maccanni*, with seven distinct clades identified. The seven *O. maccanni* clades appear to have diverged during the early-Pliocene to late-Pleistocene (4.77-0.15 Mya), although the majority exhibit Pliocene divergence (Table 4). Our evidence for Pliocene divergence among *O. maccanni* clades is consistent with studies of other South Island taxa (e.g. Trewick *et al.*, 2000; Buckley *et al.*, 2001; Trewick, 2001; Arensburger *et al.*, 2004; Chinn and Gemmell, 2004; Trewick and Morgan-Richards, 2005; Buckley and Simon, 2007; Greaves *et al.*, 2007), and is generally attributed to increased tectonic activity along the Alpine Fault line during the Miocene (Gage, 1980; Suggate, 1982; Stevens *et al.*, 1995).

The four clades from the Otago and Southland regions of the South Island (Clades 1-4) form a well-supported lineage within *O. maccanni* (Figure 3). The Otago and Southland clades appear to have diverged from those in Canterbury (Clades 5-7) in the early- to mid-Pliocene (4.77-3.69 Mya; Table 4). The genetic break between the Otago/Southland lineage and the Canterbury clades occurs approximately 20 km north of Oamaru along the Waitaki River. The Waitaki River designates the border between the Canterbury and Otago regions of New Zealand (Figure 1), and also delineates the boundary between the central and southern biogeographic regions of the South Island (Emerson *et al.*, 1997; Wallis *et al.*, 2001). The Waitaki River valley has the potential to represent a significant biogeographic barrier, as it follows a set of active faults (Darby *et al.*, 2003), and is surrounded to the southwest by the Hawkdun range and Kakanui Mountains, and the Two Thumbs range, Hunters Hills and Canterbury Plains (grassland) in the northeast (Nordmeyer,
1981). In addition, the Waitaki River is the drainage for a large catchment area (~11,826 km²; including the three natural lakes of the Mackenzie Basin and several glaciers) in the Southern Alps (Nordmeyer, 1981; Darby et al., 2003). Indeed, several other South Island taxa exhibit genetic breaks concordant with the Waitaki River valley (weta, Deinacrida connectens, Trewick et al., 2000; cicadas, Maoricicada campbelli, Buckley et al., 2001; skinks, O. lineoocellatum-chloronoton, Greaves et al., 2007; O. nigriplantare polychroma, L. Liggins, DGC, CHD, PAR, unpublished data). Interestingly, O. maccanni populations either side of the Waitaki River valley exhibit colour pattern differences (e.g. mid-dorsal stripe, speckling under chin, ventral pattern, dorsal pattern, stripe on anterior surface of forelimbs; Daugherty et al., 1990), supporting the genetic results that there has been restricted gene flow between the Canterbury and Otago regions since the Pliocene.

A substantial genetic break was present between O. maccanni populations in east Otago (Clade 4) and west Otago (Clades 1 and 2). Oligosoma maccanni populations in west Otago occur in mountainous sub-alpine regions of the Southern Alps, while those in east Otago generally occur in grassland habitats that remained un-glaciated during Pleistocene glacial cycles (Stevens et al., 1995; Figure 2). Indeed, the position of the east-west Otago break in O. maccanni is concordant with the Nevis-Cardrona fault system that has been active since the Miocene and delineates a topographic boundary between east Otago (gently-deformed erosion surface) and west Otago (deeply-eroded rugged mountains) (reviewed in Waters et al., 2001). Similar east-west splits are evident in other skink species, with divergence time estimates in O. maccanni (ND2: 4.7%, 3.6 Mya), O. grande (control region: 11%, 3.8 Mya; Berry and Gleeson, 2005), O. otagense (control region: 10%, 3.5 Mya; Birkett, 2004), O. chloronoton (ND2: 7.4%, 5.3 Mya; Greaves et al., 2007),
and *O. inconspicuum* (ND2: 3.7%, 2.6 Mya; DGC, S. Greaves & CHD, unpublished data) indicating that divergence between populations in east and west Otago occurred during the late Miocene-Pliocene. However, the east-west Otago split in some other taxa (e.g. alpine weta, *Hemideina maori*, COII: 0.8-1.2 Mya; King *et al*., 2003) appears to have occurred more recently in the mid-Pleistocene.

Although there is an east-west split evident in the Otago region, the three clades in the Canterbury region (Clades 5, 6, and 7) are separated by north-south genetic breaks (Figure 2). Divergence time estimates indicate that the differentiation among Canterbury clades occurred during the Pliocene. Concordant north-south breaks in the Canterbury region (of Pliocene origin) are evident in the *O. lineoocellatum* - *O. chloronoton* species complex (Greaves *et al*., 2007). However, this pattern has not been documented in other taxa, as the majority of studies have focused on the alpine fauna of this region of the South Island (e.g. Buckley *et al*., 2001; Chinn and Gemmell, 2004). The Canterbury region is known to be depauperate in biodiversity as a result of repeated glaciation (periglacial), and glacial erosion/washout from the Southern Alps, during Pleistocene glacial cycles (Wardle, 1963; McGlone *et al*., 1985; Connor, 2002). It is possible that the north-south genetic breaks in the Canterbury region in *O. maccanni* originated as the result of Pliocene tectonism and have been maintained through the extirpation of intervening populations during Pleistocene glacial cycles.

Patterns of within-clade genetic divergences in *O. maccanni* appear to be the result of climatic and geological processes during the mid- to late-Pleistocene (1.2-0.15 Mya). While Pliocene processes appear to have driven the diversification of the seven clades within *O. maccanni*, Pleistocene processes (predominately glacial cycles) appear to have acted to maintain divergences among these clades and shaped
the pattern of within-clade divergences. Thus, although the pattern of genetic structure in *O. maccanni* is predominately shaped by Pliocene processes, the genetic imprint of Pleistocene processes is also evident. Studies on other New Zealand skink species have clearly indicated that both Pliocene and Pleistocene processes are responsible for patterns of genetic differentiation within species (Berry and Gleeson, 2005; Greaves *et al.*, 2007, in press; Hare *et al.*, 2008; Chapple *et al.*, in press).

### 2.4.2 Taxonomic Status of the ‘Garston’ Skink

The ‘Garston’ skink is known only from a single sub-adult specimen collected from the Garston region of Otago (R. Hitchmough, pers. comm.). It was initially suggested to be distinct from sympatric *O. maccanni* due to its darker colouration and glossier scales. However, our topology test clearly rejects the hypothesis that the ‘Garston’ skink is genetically distinct from *O. maccanni*. This result supports previous allozyme and morphological analyses suggesting that the ‘Garston’ skink represents a melanistic colour morph of *O. maccanni* (R. Hitchmough, pers. comm.; CHD, unpublished data).

### 2.4.3 Did Terrestrial Taxa use Cook Strait Landbridges? Insights from Oligosoma zelandicum

Although the distribution of *O. zelandicum* spans Cook Strait, a recognised biogeographic barrier, there was only relatively limited phylogeographic structure evident. This indicates that the divergences among the current *O. zelandicum* populations occurred during the mid- to late-Pleistocene. While there were several genetic lineages present in the South Island, the North Island populations formed a well-supported lineage within *O. zelandicum* (Figure 3). Since the lower North Island
was inundated throughout the Pliocene by the Manawatu Strait, it is possible that terrestrial taxa such as *O. zelandicum* only colonised the region from the South Island during the late-Pleistocene (~0.45 Mya) with the formation of Cook Strait landbridges (Lewis *et al*., 1994; Worthy and Holdaway, 2002). Our genetic data supports a single colonisation of the North Island by *O. zelandicum* from the South Island, with the estimated timing of this event (0.46 Mya) consistent with the initial formation of Cook Strait.

There has been some debate as to whether landbridges that formed intermittently during glacial maxima in the late-Pleistocene enabled the interchange of terrestrial taxa between the North Island and South Island (Lewis *et al*., 1994; Worthy and Holdaway, 1994, 2002). The majority of molecular studies, that have examined species whose distribution spans Cook Strait, have documented genetic divergences between the North Island and South Island that substantially pre-date the initial formation of Cook Strait landbridges: *O. lineoocellatum* (ND2: 4.6%, 3.3 Mya; Greaves *et al*., 2007), *O. nigriplantare polychroma* (ND2: 2.9%, 2.1 Mya; L. Liggins, DGC, CHD, PAR, unpublished data), land snails (*Wainui umula*, ~4 Mya; Efford *et al*., 2002), brown kiwi (*Apteryx australis*, 0.9 Mya; Baker *et al*., 1995), cicadas (*M. campbelli*, 0.9 Mya; Buckley *et al*., 2001), and short-tailed bats (*Mystacina tuberculata*, 0.89 Mya; Lloyd, 2003).

In contrast, our data suggests that *O. zelandicum* might have used a Cook Strait landbridge to disperse from the South Island to the North Island. Geological evidence indicates that Cook Strait landbridges connected the Nelson/Marlborough region of the South Island to the lower North Island (Lewis *et al*., 1994; King, 2000; Worthy and Holdaway, 2002) (Figure 2). Hardy (1977) hypothesized that several skink species persisted in refugia in the Nelson/Marlborough region during
Pleistocene glacial maxima. Thus, the presence of *O. zelandicum* in the Nelson/Marlborough region during glacial maxima may have enabled it to disperse into the North Island via a Cook Strait landbridge. However, it appears that *O. zelandicum* might have only used the initial Cook Strait landbridge and not landbridges in subsequent glacial maxima. This result supports the suggestion by Worthy and Holdaway (1994, 2002) that there was no Cook Strait landbridge (or no landbridge suitable for dispersal by terrestrial taxa) during the LGM, with the most recent landbridge occurring at least ~0.13 Mya.

### 2.5 References


Chapter Two: Phylogeography of Two New Zealand Lizards


Chapter Two: Phylogeography of Two New Zealand Lizards


Chapter Two: Phylogeography of Two New Zealand Lizards


Chapter Two: Phylogeography of Two New Zealand Lizards


### 2.6 Tables

**Table 1.** Locality information and GenBank accession numbers for samples used in this study. Samples with CD or FT codes were obtained from the National Frozen Tissue Collection (NFTC) housed at Victoria University of Wellington, New Zealand. Samples with RE codes were obtained from ethanol-preserved specimens housed at Te Papa, National Museum of New Zealand, Wellington (S codes refer to specimens from the former Ecology Division collection, now housed at Te Papa).

<table>
<thead>
<tr>
<th>Species</th>
<th>Museum Code</th>
<th>Tissue Code</th>
<th>GenBank Accession No.</th>
<th>Locality</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Oligosoma maccanni</em></td>
<td>CD930</td>
<td>OMA1</td>
<td><strong>EF081195</strong></td>
<td><strong>EF081222</strong></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD1261</td>
<td>OMA2</td>
<td><strong>EF447115</strong></td>
<td><strong>EF447148</strong></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>FT3213</td>
<td>OMA3</td>
<td><strong>EF447116</strong></td>
<td><strong>EF447149</strong></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>FT3039</td>
<td>OMA4</td>
<td><strong>EF447117</strong></td>
<td><strong>EF447150</strong></td>
</tr>
</tbody>
</table>
### Chapter Two: Phylogeography of Two New Zealand Lizards

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession</th>
<th>Accession</th>
<th>Location</th>
<th>Location Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>O. maccanni</em></td>
<td>-</td>
<td>OMA5</td>
<td>EF447118 EF447151</td>
<td>Alexandra, central Otago</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>-</td>
<td>OMA6</td>
<td>EF447119 EF447152</td>
<td>Alexandra, central Otago</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>-</td>
<td>OMA7</td>
<td>EF447120 EF447153</td>
<td>Macraes Flat, Otago</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>-</td>
<td>OMA8</td>
<td>EF447121 EF447154</td>
<td>Macraes Flat, Otago</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD956</td>
<td>OMA9</td>
<td>EF447122 EF447155</td>
<td>Edwards Creek, Canterbury</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD804</td>
<td>OMA10</td>
<td>EF447123 EF447156</td>
<td>Mount Hay Station, Canterbury</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD615</td>
<td>OMA11</td>
<td>EF447124 EF447157</td>
<td>Birdlings Flat, Canterbury</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD1254</td>
<td>OMA12</td>
<td>EF447125 EF447158</td>
<td>Gordon Peak, Knobby Range, Otago</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD1119</td>
<td>OMA13</td>
<td>EF447126 EF447159</td>
<td>Otematata, Canterbury</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD1106</td>
<td>OMA14</td>
<td>EF447127 EF447160</td>
<td>Gorge Burn, Eyre Mountains, Southland</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD778</td>
<td>OMA15</td>
<td>EF447128 EF447161</td>
<td>Burgan Stream Hut, Otago</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD1255</td>
<td>OMA16</td>
<td>EF447129 EF447162</td>
<td>Hills Creek, Otago</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD837</td>
<td>OMA18</td>
<td>EF447130 EF447163</td>
<td>Tara Hills, Canterbury</td>
</tr>
<tr>
<td>Species</td>
<td>Accession</td>
<td>Location Details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>-----------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD635</td>
<td>Mount John Station, Canterbury</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD634</td>
<td>Simons Hill, Canterbury</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD627</td>
<td>Mount Mary, Canterbury</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD626</td>
<td>Taieri Ridge, Otago</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD567</td>
<td>Conroys Dam, Otago</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD428</td>
<td>Remarkables, Otago</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD426</td>
<td>10 km North of Millers Flat, Otago</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD425</td>
<td>Frankton, Otago</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>CD10</td>
<td>Hokonui Hills, Southland</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>RE4990</td>
<td>Lake Tekapo, Canterbury</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>RE4890</td>
<td>Near outlet, Lake Hawea, Otago</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>RE4888</td>
<td>Tekowai Island, Lake Pukaki, Canterbury</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>RE4860</td>
<td>Port Hills, Lyttleton, Canterbury</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Chapter Two: Phylogeography of Two New Zealand Lizards

<table>
<thead>
<tr>
<th>Lizard Species</th>
<th>Accession Number (Location)</th>
<th>Accession Number (Location)</th>
<th>Location/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>O. maccanni</em></td>
<td>RE4449 (S807) OMA35 EF447144 EF447177</td>
<td></td>
<td>Ida Range, 13 km north of Naseby, Otago</td>
</tr>
<tr>
<td><em>O. maccanni</em></td>
<td>RE3917 (S275) OMA37 EF447145 EF447178</td>
<td></td>
<td>Te Anau, Southland</td>
</tr>
<tr>
<td><em>Oligosoma zelandicum</em></td>
<td>FT6516 OZE2 EF447181 EF447201</td>
<td></td>
<td>Pukerua Bay, Wellington</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>FT3814 OZE3 EF447182 EF033070</td>
<td></td>
<td>Mokihinui River Mouth, West Coast</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>CD331 OZE4 EF447183 EF447202</td>
<td></td>
<td>Maud Island, Nelson/Marlborough</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>CD542 OZE5 EF447184 EF447203</td>
<td></td>
<td>Stephens Island, Nelson/Marlborough</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>CD1951 OZE6 EF447185 EF447204</td>
<td></td>
<td>Outer Chetwode Island, Nelson/Marlborough</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>FT3743 OZE7 EF447186 EF447205</td>
<td></td>
<td>Gentle Annie, West Coast</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>FT3773 OZE8 EF447187 EF447206</td>
<td></td>
<td>Scotts Beach, Heaphy Track, West Coast</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>FT6525 OZE9 EF447188 EF447207</td>
<td></td>
<td>Johnsonville, Wellington</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>RE1594 OZE19 EF447189 EF447208</td>
<td></td>
<td>Foxton, Manawatu</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>CD335 OZE20 EF447190 EF447209</td>
<td></td>
<td>Maud Island, Nelson/Marlborough</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>CD543 OZE21 EF447191 EF447210</td>
<td></td>
<td>Stephens Island, Nelson/Marlborough</td>
</tr>
</tbody>
</table>
### Chapter Two: Phylogeography of Two New Zealand Lizards

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>O. zelandicum</em></td>
<td>CD544</td>
<td>OZE22 EF447192 EF447211</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>CD1952</td>
<td>OZE23 EF447193 EF447212</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>FT3744</td>
<td>OZE24 EF447194 EF447213</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>FT6517</td>
<td>OZE25 EF447195 EF447214</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>FT6518</td>
<td>OZE26 EF447196 EF447215</td>
</tr>
<tr>
<td><em>O. zelandicum</em></td>
<td>FT6526</td>
<td>OZE27 EF447197 EF447216</td>
</tr>
<tr>
<td><em>Nannoscincus mariei</em></td>
<td>NR9808</td>
<td>EUG1</td>
</tr>
<tr>
<td><em>Marmorosphax tricolor</em></td>
<td>NR9800</td>
<td>EUG2</td>
</tr>
</tbody>
</table>
Table 2. Oligonucleotide primers used in this study. The letters L and H refer to the light and heavy strands. Values in 5’ position refer to the position of the 5’ base of the primer in the complete *Eumeces egregious* mtDNA sequence (Kumazawa and Nishida, 1999).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Name</th>
<th>Sequence (5’ – 3’)</th>
<th>5’ Position</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND2</td>
<td>L4437</td>
<td>AAGCTTTCCGGGCCCATACC</td>
<td>3833</td>
<td>Macey et al. (1997)</td>
</tr>
<tr>
<td></td>
<td>ND2r102</td>
<td>CAGCCTAGGTTGGCGGATTG</td>
<td>4432</td>
<td>Sadlier et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>H4980</td>
<td>ATTTTTTCGTAGTTGGTTTGGRTT</td>
<td>4396</td>
<td>Macey et al. (1997)</td>
</tr>
<tr>
<td></td>
<td>L4221</td>
<td>AAGGATTACTTTTGATAGAGT</td>
<td>3691</td>
<td>Macey et al. (1997)</td>
</tr>
<tr>
<td>ND4</td>
<td>ND4I</td>
<td>TGACTACAAGAACGCTCATGTAGAGC</td>
<td>10796</td>
<td>Forstner et al. (1995)</td>
</tr>
<tr>
<td></td>
<td>ND4R-NZ</td>
<td>CCAAGRGTTTTTGGTCCTAAGACC</td>
<td>11670</td>
<td>Greaves et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>tRNA-Leu</td>
<td>TACCTTTACTGGATTGCAACCA</td>
<td>11691</td>
<td>Forstner et al. (1995)</td>
</tr>
</tbody>
</table>
Table 3. Estimates of mitochondrial DNA (ND2, ND4) genetic diversity for *Oligosoma maccanni* (including 7 clades) and *O. zelandicum* (North Island and South Island samples). $N$ is the number of individual sequences obtained. $h$ is the number of haplotypes in each clade/population. $Hd$ is the haplotype diversity. $S$ is the number of polymorphic (segregating) sites. $n$ is the number of mutations (substitutions) for each clade/population. $\pi$ represents the nucleotide diversity within each clade/population.

<table>
<thead>
<tr>
<th>Species/clade</th>
<th>N</th>
<th>$h$</th>
<th>$Hd$</th>
<th>S</th>
<th>n</th>
<th>$\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Oligosoma maccanni</em></td>
<td>32</td>
<td>31</td>
<td>0.998</td>
<td>198</td>
<td>210</td>
<td>0.0390</td>
</tr>
<tr>
<td>Clade 1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Clade 2</td>
<td>6</td>
<td>5</td>
<td>0.933</td>
<td>6</td>
<td>6</td>
<td>0.0023</td>
</tr>
<tr>
<td>Clade 3</td>
<td>1</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Clade 4</td>
<td>13</td>
<td>13</td>
<td>1</td>
<td>68</td>
<td>78</td>
<td>0.0150</td>
</tr>
<tr>
<td>Clade 5</td>
<td>7</td>
<td>7</td>
<td>1</td>
<td>23</td>
<td>23</td>
<td>0.0057</td>
</tr>
<tr>
<td>Clade 6</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>12</td>
<td>0.0094</td>
</tr>
<tr>
<td>Clade 7</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0.0016</td>
</tr>
<tr>
<td><em>Oligosoma zelandicum</em></td>
<td>17</td>
<td>12</td>
<td>0.956</td>
<td>34</td>
<td>34</td>
<td>0.0061</td>
</tr>
<tr>
<td>North Island</td>
<td>6</td>
<td>5</td>
<td>0.933</td>
<td>18</td>
<td>18</td>
<td>0.0062</td>
</tr>
<tr>
<td>South Island</td>
<td>11</td>
<td>7</td>
<td>0.909</td>
<td>15</td>
<td>15</td>
<td>0.0032</td>
</tr>
</tbody>
</table>
Table 4. Mean genetic distances among *Oligosoma maccanni* clades identified in Figure 3. ND2 uncorrected genetic distances are shown below the diagonal and Tamura-Nei (TrN) corrected genetic distances for the concatenated dataset (ND2 + ND4) are shown above the diagonal.

<table>
<thead>
<tr>
<th></th>
<th>Clade 1</th>
<th>Clade 2</th>
<th>Clade 3</th>
<th>Clade 4</th>
<th>Clade 5</th>
<th>Clade 6</th>
<th>Clade 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clade 1</td>
<td>–</td>
<td>0.027</td>
<td>0.042</td>
<td>0.027</td>
<td>0.053</td>
<td>0.050</td>
<td>0.066</td>
</tr>
<tr>
<td>Clade 2</td>
<td>0.002</td>
<td>–</td>
<td>0.047</td>
<td>0.048</td>
<td>0.052</td>
<td>0.051</td>
<td>0.063</td>
</tr>
<tr>
<td>Clade 3</td>
<td>0.055</td>
<td>0.057</td>
<td>–</td>
<td>0.034</td>
<td>0.056</td>
<td>0.051</td>
<td>0.061</td>
</tr>
<tr>
<td>Clade 4</td>
<td>0.046</td>
<td>0.047</td>
<td>0.038</td>
<td>–</td>
<td>0.050</td>
<td>0.052</td>
<td>0.062</td>
</tr>
<tr>
<td>Clade 5</td>
<td>0.054</td>
<td>0.056</td>
<td>0.060</td>
<td>0.048</td>
<td>–</td>
<td>0.045</td>
<td>0.060</td>
</tr>
<tr>
<td>Clade 6</td>
<td>0.053</td>
<td>0.055</td>
<td>0.059</td>
<td>0.052</td>
<td>0.046</td>
<td>–</td>
<td>0.062</td>
</tr>
<tr>
<td>Clade 7</td>
<td>0.062</td>
<td>0.062</td>
<td>0.058</td>
<td>0.054</td>
<td>0.059</td>
<td>0.056</td>
<td>–</td>
</tr>
</tbody>
</table>
2.7 Figure Legends

**Figure 1.** Distributional range of *Oligosoma maccanni* (black circles) and *O. zelandicum* (black triangles). Distributional data were obtained from the New Zealand Department of Conservation’s BioWeb Herpetofauna database (2006). (Figure adapted with permission from Geographx (NZ) Ltd). Inset: Major geographic regions in New Zealand.

**Figure 2.** Location of *O. maccanni* (black circles) and *O. zelandicum* (black triangles) tissue samples used in this study. The approximate distribution of *O. maccanni* clades identified in Figure 3 are shown. Note that clade 1 (OMA27) and clade 3 (OMA31) are each represented by a single sample. (Figure adapted with permission from Geographx (NZ) Ltd). Inset: Map indicating the coastline (dark shading) of New Zealand during the last glacial maximum (LGM). The map highlights the landbridge that connected the Taranaki region with the Nelson region during the LGM (Figure reproduced with permission from Te Ara, Encyclopedia of New Zealand).

**Figure 3.** Bayesian consensus tree for *Oligosoma maccanni* and *O. zelandicum* based on 1282 bp of mitochondrial DNA (507 bp ND2, 775 bp ND4). The topologies of the Neighbour-Joining, Maximum Parsimony (MP), and Maximum Likelihood trees were almost identical to the Bayesian tree shown. Two measures of branch support are indicated with MP bootstraps on the left and Bayesian posterior probabilities on the right (only values over 50 and 0.7, respectively, are shown). The seven main clades within *O. maccanni* are indicated.
2.8 Figures

Figure 1.
Chapter Two: Phylogeography of Two New Zealand Lizards

Figure 2.
Chapter Two: Phylogeography of Two New Zealand Lizards

Figure 3.
Chapter Three

Evolutionary Concordance across Independent Loci for the *Niveaphasma* Stick Insect Complex

3.1 Introduction

As the field of phylogeography progresses, more sophisticated models are being used to test for particular historical patterns. This demand for increased resolution requires phylogeography, to expand from a largely marker-specific discipline, to build a more general analytical approach that can be applied to a range of independent loci (Hare, 2001). The technical complications of heterozygosity and recombination associated with most nuclear sequences have probably inhibited their use in phylogeographic studies (Avise, 2000; Hare, 2001) However, many of these inherent complications have been overcome with the advances of techniques such as Denaturing Gradient
Gel Electrophoresis (DGGE). These analytical improvements mean variation between sequences and resolution of nuclear haplotypes from diploid tissues can easily be resolved.

Recombination in nuclear DNA is seen to hinder phylogenetic signals because it results in the mixing of evolutionary histories within the same gene. However, reduced gene flow between isolated populations should restrict the effect of recombination to within these populations and should not obscure strong phylogeographical patterns (Bernardi et al., 1993; Hare, 2001).

Doubts have also been raised in regard to the effectiveness of mtDNA for inferring species phylogenies. The use of mtDNA can be limited by lineage sorting because genes are inherited as a single linkage group and, as a result, mtDNA gene trees and species trees can be incongruent (Moore, 1995; Hoelzer, 1997; Page and Holmes, 1998). Another concern is that non-recombining loci, such as the mitochondrial genome, are susceptible to introgression (the movement of genes between species via hybridization) and are heavily dependant on selective forces acting on linked sites (Machado and Hey, 2003). This can lead to a reduction in polymorphisms due to selective sweeps and background selection (Kliman et al., 2000; Machado and Hey, 2003). This illustrates why caution must be taken when inferring the evolutionary history of recently diverged populations or species when using a single locus or non-recombining loci (Moore, 1995; Giannasi et al., 2001; Machado and Hey, 2003). For this reason, a species or population evolutionary history is best inferred from at least two independent loci (Pamilo and Nei, 1988; Wu, 1991; Moore, 1995).

In attempting to answer phylogeographic questions regarding New Zealand’s South Island biota, studies have sought genealogical concordance across co-
distributed species (Wagstaff and Garnock-Jones, 1998; McGlone et al., 2001; Trewick and Wallis, 2001), with other biogeographical data (Wallis et al., 2001; Smissen et al., 2003), within a single gene from a single species (Trewick, 2001; Chinn and Gemmell, 2004) or across linked genes of a species (Greaves et al., 2007). Many studies have relied solely on mtDNA as the preferred molecular marker (Buckley et al., 2001; Trewick, 2001; Boyer et al., 2007), while a few have used allozymes or isozymes (Wallis et al., 2001) or nuclear molecular markers such as ITS in the case of plants (Wagstaff and Garnock-Jones, 1998; Smissen et al., 2003). Concordance across independent loci is essential in demonstrating fundamental phylogenetic subdivisions at the population or species level (Avise, 2000). Recent studies have applied multiple markers at the speciation level (stick insects; Trewick et al., 2005; Buckley et al., 2008; cicada, *Maoriciada* Dugdale; Buckley et al., 2006; Buckley and Simon, 2007). But no phylogeographical study in New Zealand has addressed genealogical concordance at the intraspecific level, using both mitochondrial DNA and nuclear gene sequence data with resolution of alleles.

The South Island presents the ideal opportunity to observe phylogenetic divisions of its endemic taxa, due to the high degree of modification asserted on its landscape during the recent geological past. Throughout the Pliocene (5 - 2 million years ago (Mya); Stevens et al., 1995), the South Island experienced extensive periods of tectonic uplift along the Alpine Fault; a zone of subduction where the edge of the Indian-Australian Plate is being pushed over the oceanic Pacific Plate (Stevens et al., 1995). Tectonic uplift resulted in the formation of the Southern Alps which now dominate the South Island landscape. A period of dramatic climatic cooling occurred during the Pleistocene (2 – 0.1 Mya). This gave rise to the formation of
glaciers, which carved out massive valleys, in turn creating spectacular fiords such as those seen in Milford Sound (Stevens et al., 1995).

The expansion of the alpine zone during the Pliocene/Pleistocene presented a new habitat for many South Island taxa (Buckley and Simon, 2007). Molecular studies on endemic alpine plant species have linked repeated glaciations to the extinction, speciation and dispersal of species in the South Island (Hebe complex; Wagstaff and Garnock-Jones, 1998; vascular plant flora; McGlone et al., 2001; alpine cushion plants, Raoulia spp. Smissen et al., 2003; alpine flora; Winkworth et al., 2005). Similar findings have been attributed to speciation and dispersal events involving South Island vertebrate species (Galaxiidae; Waters and Wallis, 2000; Wallis et al., 2001; spotted and green skink, Oligosoma lineocellatum and O. chloronoton; Greaves et al., 2007; McCann’s skink, Oligosoma maccanni; O’Neill et al., submitted manuscript). Studies have also argued that glaciations throughout the Pleistocene caused genetic divergence, dispersal and speciation events among alpine invertebrate populations (alpine weta, Deinacrida connectins; Trewick, 2001; cicada, Maoricicada campbelli; Buckley et al., 2001, 2007; cockroaches, Celatoblatta spp; Chinn and Gemmell, 2004).

This study utilizes both mtDNA and nuclear DNA to investigate evolutionary concordance within the stick insect genus Niveaphasma. Niveaphasma (Jewell and Brock, 2002) belongs to the stick insect subfamily Pachymorphinae. Currently the only recognized species within Niveaphasma is N. annulata; however it has been hypothesized by Jewell and Brock (2002) that undescribed species may exist. Niveaphasma has a stout body with the female ranging from 49 – 56 mm in length and the male only reaching up to 40 mm (Jewell and Brock, 2002). Its colouring varies between shades of medium to dark brown, and the female often possesses
small spines on its posterior abdomen. *Niveaphasma* is found only in the South Island from Nelson to Southland and occurs from sea level up to sub-alpine zones of ~1360 m (Jewell and Brock, 2002). Its alpine habitat makes it of particular phylogeographical interest in the South Island. A previous study on the alpine cicada genus *Maoricicada* demonstrated the usefulness of the nuclear gene EF1α in detecting phylogenetic signals in endemic South Island invertebrates (Buckley and Simon, 2007). Successful PCR, molecular cloning and sequencing of EF1α alleles also proved useful investigating hybridization events between New Zealand stick insects *Acanthoxyla* spp. and *Clitarchus hookeri* (Buckley et al., 2008). Intraspecific nuclear and mitochondrial gene concordance has been recognized previously across vertebrates and invertebrates (teleost fish, *Fundulus heteroclitus*; Bernardi et al., 1993; copepod, *Tigriopus californics*; Burton and Lee, 1994) in Northern Hemisphere systems. Evidence also suggests that glacial periods in North America (2-3 Mya) coincide with extensive speciation within the stick insect genus *Timema* (Sandoval et al., 1998; Law and Crespi, 2002).

Using mitochondrial and nuclear data, *Niveaphasma* provides the opportunity to investigate evolutionary concordance across independent loci and the patterns of evolution of one of New Zealand’s endemic South Island species. Findings based on this study present the prospect to more fully understand the complex patterns evolving in genes as a response to environmental change (Hare, 2001).

### 3.2 Materials and Methods

#### 3.2.1 Sample Collection
Tissue samples were obtained from ethanol preserved specimens housed at the New Zealand Arthropod Collection (NZAC), Landcare Research, Auckland, New Zealand. The samples used cover the entire known range of the *Niveaphasma* species complex, except for the Nelson/Marlborough region. Based on a wider phylogeographic study on the relationships of all members of the New Zealand stick insect radiation (TRB, unpublished data) samples from *Asteliaphasma jucundus* (ASJ1) and *Tectarchus huttoni* (THE2) were included as outgroups (Table 1).

3.2.2 DNA Extraction, Amplification and Sequencing

Total genomic DNA was isolated from the right front leg of each specimen using a modified phenol-chloroform extraction protocol (Sambrook *et al*., 1989). Where juveniles or small specimens were used, more than one leg was required to ensure enough tissue was obtained. In some cases, the head was also used because legs were incomplete, damaged or missing altogether. For each sample a portion of the Cytochrome Oxidase I (COI) mitochondrial gene (~762 bp) and the Elongation Factor 1α (EF1α) nuclear gene (~590 bp) was targeted. Both these regions were targeted because work at comparable phylogeographic levels in other stick insects and endemic New Zealand invertebrates have shown the usefulness of these genes in determining levels of variability (Crespi and Sandoval, 2000; Law and Crespi, 2002; Morgan-Richards and Trewick, 2005; Trewick *et al*., 2005; Buckley *et al*., 2006; Buckley and Simon, 2007).

The COI sequences were obtained using primers originally described in Simon *et al*. (1994) and later used in Buckley *et al*. (2001). Primers for the amplification of the EF1α nuclear gene were designed for this study. The primers used are provided in Table 2. PCR reactions were 25 µL in total volume and
Chapter Three: Genealogical Concordance across Independent Loci

contained 0.4 µM of each primer, 5 mM of each dNTP, 0.4 mg/mL of Bovine serum albumin (BSA) (ICPbio Ltd), 2.5 µL 10× Reaction buffer (670 mM Tris pH 8.0, 160 mM (NH₄)₂SO₄, 1.5 mM MgCl₂ and 1 U BIOTAQ DNA polymerase (BIOLINE). In the case of EF1α amplification, 1 U BIOTAQ DNA polymerase was replaced with 1 U Platinum Taq DNA polymerase (INVITROGEN). PCR of double stranded template was carried out using an Eppendorf Mastercycler ep gradient S thermal cycler. Reactions were initially denatured at 95°C for 3 min. The cycling sequencing was carried out using the following profile for 35 cycles: 94°C for 30 sec, 57°C for 30 sec, 72°C for 1 min. A final extension step was performed for 10 min at 72°C. PCR products were purified either by a High Pure™ PCR Product Purification Kit (Roche Diagnostics) or adding 1 µL Exonuclease I and Shrimp Alkaline Phosphatase (ExoSAP-IT) (USB) followed by incubation at 37°C for 30 min and 80°C for 15 min. Samples were then directly sequenced using the BigDye Terminator v3.1 Reaction (Applied Biosystems) and analyzed on an ABI 3730 Capillary Sequencer (Applied Biosystems).

3.2.3 DGGE Analysis

A second round of PCR amplification was performed on the EF1α gene with the addition of a GC-clamp on the EF1for primer (Table 2). PCR products recovered were subjected to Denaturing Gradient Gel Electrophoresis (DGGE). DGGE analysis was performed using 8% acrylamide gels (ratio of acrylamide to bisacrylamide 37.5:1) with a 30-60% denaturant gradient in which 100% denaturant was defined as 7 M urea plus 40% deionized formamide. Approximately 1-3 µg of PCR product was loaded per sample in a final volume of 10 µL. The gels were electrophoresed at 60°C at 100 V for 14 hrs using the DCode Universal Mutation Detection System (Bio-Rad
Laboratories, Hercules, CA). The gels were stained for 20 min with ethidium bromide and washed in ddH$_2$O for 5 min prior to UV transillumination. The DGGE gels were digitized using the UVitec DBT-08 photo system (Total Lab Systems Ltd, Cambridge, UK). Gel images were examined with the UVipro software package version 12.5 (Total Lab Systems Ltd).

### 3.2.4 Recovery of Bands from DGGE Gels and Sequence Analysis

When bands were to be excised from gels, a silver stain replaced the ethidium bromide stain. The gel was initially soaked in fixer (7.5% glacial acetic acid) for 10 min and rinsed in ddH$_2$O. Gels were then silver stained (0.1% silver nitrate; 0.15% formaldehyde) for 20 min and rinsed in ddH$_2$O. Following silver staining gels were placed in a developer solution (3% sodium carbonate; 0.3% formaldehyde; 0.2% sodium thiosulphate) for 5-10 min and rinsed in ddH$_2$O.

Following silver staining gels were visualized on a light box. Prominent DGGE bands were selected and used for excision and nucleotide sequence determination. For each band selected, only the middle portion was excised with a sterile razor blade. The resultant slices (approximately 30 mg [wet weight]) were placed in 1.5 mL Eppendorf tubes containing 150 µL TE buffer. Tubes were incubated at 37°C for 4-5 hrs to elute the DNA from the bands. 1 µL of buffer containing DNA was used as a template for PCR, and subsequent PCR amplification was performed under the conditions described previously for the EF1α nuclear gene.

### 3.2.5 Data Analysis

Sequence data were edited using CONTIGEXPRESS version 10.0.1 (Invitrogen Corporation) and aligned using the default parameters of CLUSTALW (Thompson *et*
al., 1994) and edited by eye. The edited COI sequences were translated into amino acid sequences using the invertebrate mitochondrial genetic code with codon initiation from the 1st site. No premature stop codons were observed in the COI gene, allowing the conclusion that these data were indeed mitochondrial in origin.

Estimates of character variability of the edited COI and EF1α data sets were conducted using MEGA version 4.0 (Tamura et al., 2007). Genetic distances were determined in PAUP* version 4.0b10 (Swofford, 2002) using the parameters determined from the AIC analysis performed in MODELTEST version 3.7 (Posada and Crandall, 1998) and PAUP*. The number and frequency of unique mtDNA haplotypes and nuclear DNA alleles was determined using TCS version 1.21 (Clement et al., 2000) and DNAsp version 4.10.9 (Rozas et al., 2003).

A partition-homogeneity test conducted using PAUP* showed that the phylogenetic signal from both loci (COI and EF1α) was incongruent (100 replicates; $P > 0.01$), enabling a data set for each individual locus to be constructed and analyzed.

3.2.6 Phylogenetic Analysis

Model selection was performed using the Akaike Information Criterion (AIC) (Akaike, 1973; Posada and Buckley, 2004) as implemented in PAUP* and MODELTEST. AIC was used over the more commonly applied model selection approach of the hierarchical likelihood ratio test (hLRT) as it has been argued that AIC offers considerable benefits over hLRT. Such benefits include the ability to simultaneously compare multiple nested or non-nested models, assess model uncertainty and allow for the estimation of phylogenies and model parameters using all available models (Posada and Buckley, 2004). The best fit model was determined
by estimating the NJ tree under the JC69 model (Jukes and Cantor, 1969). This enabled the likelihood of the tree to be calculated by applying the 56 substitution models implemented in MODELTEST. The Maximum Likelihood (ML) trees under the resulting best-fit models were estimated using a heuristic search with TBR branch swapping from 10 random addition replicates. The purpose of this was to reduce the computational load encountered with ML analysis. A bootstrap analysis (Felsenstein, 1985) of 100 pseudoreplicates was performed on the two loci, with an initial NJ tree constructed followed by a heuristic search with TBR branch swapping.

Bayesian phylogenetic analysis was performed using MrBAYES version 3.1.2 (Huelsenbeck and Ronquist, 2001) on both data sets. A default value of four Markov chains per run was initiated from a random tree. Likelihood settings for Nst and Rates were set according to the best-fit model selected from PAUP* and MODELTEST i.e. for the HKY + I + Γ model Nst = 2 and Rates = invgamma. A full analysis was run twice for each data set, with each analysis run for 5,000,000 generations and every 1000th tree sampled to obtain 5000 sampled trees. The program TRACER version 1.3 (Rambaut and Drummond, 2003) was used to check for chain convergence. A burn-in phase of 2500 sampled trees was performed, leaving 5000 trees in total from both runs to estimate Bayesian posterior probabilities. Bootstrap values and Bayesian posterior probabilities were used to assess branch support. As a general guide, bootstrap values equal to or greater than 75% and posterior probability values equal to or greater than 95% are considered significantly supportive of branches (Hillis and Bull, 1993; Wilcox et al., 2002).

Population genetic analysis was estimated using ARLEQUIN version 3.01 (Excoffier et al., 2006) and DNAsp (Rozas et al., 2003). Mismatch frequency histograms were plotted in DNAsp to determine whether the number of pairwise
differences within *Niveaphasma* reflected a rapid population expansion model or a constant population history model (Harpending, 1994). For populations remaining stationary over long periods of time, mismatch distributions become ragged and erratic, whereas populations undergoing growth and expansion produce smooth bell shaped curves with a peak (Harpending, 1994; Ray *et al*., 2003). The raggedness ($r$, sum of the squared difference between neighboring peaks) and sum of the squared deviation (SSD) statistics are used to distinguish between the two mismatch models and were calculated in ARLEQUIN.

A Mantel test was conducted to determine matrix correlation between COI genetic distance and geographic distance using the ISOLATION BY DISTANCE, WEB SERVICE version 3.15 (Jensen *et al*., 2005). Geographic distances were measured in kilometers using MAPINFO PROFESSIONAL version 5.5 (MapInfo Corporation). A data matrix was then produced using COI genetic distances and the corresponding linear geographical distance. Statistical parsimony networks were constructed by TCS for both the COI mtDNA and EF1α nuclear DNA data sets using DNA pairwise differences.

### 3.3 Results

The edited COI *Niveaphasma* alignment contained 99 sequences (97 *Niveaphasma* sequences; two outgroup sequences) and 762 characters. The edited EF1α alignment contained 107 sequences (105 *Niveaphasma* sequences; two outgroup sequences) and 590 characters. [For full details of sequence statistics of both loci refer to Table 3]. Discrepancies in the number of sequences between data sets are due to missing data and the presence of heterozygotes in the EF1α data set. NIV56 and NIV72 are
missing from the COI data set, while NIV51 is missing from the EF1α data set. DGGE analysis confirmed the presence of nine heterozygotes within the EF1α data set (NIV37, NIV59, NIV60, NIV75, NIV77, NIV78, NIV85, NIV96, and NIV97) of which seven were resolved into individual alleles (NIV37a and b, NIV59a and b, NIV60a and b, NIV77a and b, NIV78a and b, NIV96a and b, NIV97a and b). Both NIV75 and NIV85 EF1α alleles could not be resolved because PCR amplification of these fragments was unsuccessful following elution of DNA from the excised DGGE bands. In the COI data, 48 haplotypes (h) were observed with a haplotype diversity (Hd) of 0.93, and a total of 19 alleles (labeled a – s) were identified within the EF1α data set.

Phylogenetic analysis produced a single rooted phylogram for each of the data sets (Figures 1a, 1b). The COI and EF1α data sets each produced trees of similar topology for the ML and Bayesian analysis (optimal tree, COI, \(-\ln L = 2950.21\); EF1α, \(-\ln L = 1231.10\)). Consequently the mean Bayesian consensus tree is presented for each data set, with ML bootstrap values and posterior probabilities indicating branch support (Figures 1a, 1b). It is apparent that considerable genetic variation exists within the *Niveaphasma* COI data, with an average HKY + I + Γ genetic distance (GD) of 2.91%. Three major clades (Clades 1, 2 and 3) were identified from the COI data (Figure 1a). NIV82 was isolated from each of the three major clades, as its phylogenetic position was not concordant across the COI and EF1α data sets.

Clade 1 (<50% bootstrap; 0.95 posterior probability; average GD = 0.70%) comprises samples from eastern Otago and Canterbury (Figure 2a). All samples collected from the Canterbury region are restricted to this clade. Clade 1a (57% bootstrap; 0.95 posterior probability) includes samples from Mt Cook Village and as far south as Dunedin. Within this clade are three COI haplotypes and seven EF1α
alleles (Table 1; Figure 1a). Clade 1b (<50% bootstrap; 1.00 posterior probability) occurs at Arthur’s Pass in the Southern Alps, with the samples collected from this location all sharing a unique COI haplotype and a common EF1α allele (Table 1; Figure 1a). Clade 1c (<50% bootstrap; 1.00 posterior probability) occurs at Waipori Falls, with samples collected from this site sharing a unique COI haplotype and a common EF1α allele (Table 1; Figure 1a). NIV 81 is the only sample taken from eastern Canterbury and it possesses a unique COI haplotype. NIV55 (highlighted in bold typeset; Figure 1a) lies within Clade 1a, but is located in central Otago near Cromel Stream (underlined typeset; Figure 2a).

Clade 2 (82% bootstrap; 1.00 posterior probability; average GD = 1.58%) includes all the Southland samples and also those from central Otago (Figure 2a). Clade 2 comprises ~60% of the samples from the COI data set and contains several clades. Clade 2a (66% bootstrap; <0.5 posterior probability) occurs in southern Southland, near Riverton and has three COI haplotypes which share a common EF1α allele (Table 1; Figure 1a). Clade 2b (86% bootstrap; 1.00 posterior probability) is situated near Fiordland, stretching from the Eglinton River Valley to Lake Monowai and contains two COI haplotypes, sharing the same EF1α allele (Table 1; Figure 1a). Clade 2c (53% bootstrap; <0.5 posterior probability) is located in central Otago and includes sites near Cromwell, Queenstown and Wanaka. It comprises of 11 COI haplotypes and five EF1α alleles (Table 1; Figure 1a). Clade 2d (<50% bootstrap; 1.00 posterior probability) occurs in Southland at Piano Flat, Waikaia Bush. It contains four COI haplotypes and two EF1α alleles, plus one undetermined EF1α allele for NIV48 (Table 1; Figure 1a). Clade 2e (79% bootstrap; 1.00 posterior probability) is situated in Southland at Piano Flat, Waikaia Bush, Hokonui Hills and Cromel Stream and comprises four COI haplotypes and three EF1α alleles (Table 1;
Figure 1a). Clade 2f (<50% bootstrap; <0.5 posterior probability) stretches from the Hokonui Hills to Daffodil Bay near Riverton. It contains four COI haplotypes which all share an identical EF1α allele (Table 1; Figure 1a). Clade 2g (89% bootstrap; 1.00 posterior probability) is situated at Lake Hauroko and contains two COI haplotypes and two EF1α alleles (Table 1; Figure 1a). Clade 2h (90% bootstrap; 1.00 posterior probability) is located at Windy Creek, near the Takitimu Mountains and has two COI haplotypes which share an identical EF1α allele (Table 1; Figure 1a). Clade 2i (<50% bootstrap; 0.98 posterior probability) is situated in Otago at Shingle Creek, near Old Man Range and contains one COI haplotype with a unique EF1α allele (Table 1; Figure 1a). NIV67, located near Lake Hauroko, was not placed in any of the minor clades, as its phylogenetic position was not concordant across the COI and EF1α tree.

Clade 3 (84% bootstrap; 1.00 posterior probability; average GD = 1.04%) is situated in eastern Otago, extending from Alexandra to the Catlins (Figure 2a). Clade 3a (60% bootstrap; <0.5 posterior probability) comprises samples from Alexandra and Rigney, which lies between Beaumont and Millers Flat. It consists of four COI haplotypes and seven EF1α alleles (Table 1; Figure 1a). Clade 3b (77% bootstrap; 1.00 posterior probability) includes all the samples collected from the Catlins, and contains two COI haplotypes and four EF1α alleles (Table 1; Figure 1a). Clade 3c (<50% bootstrap; 0.99 posterior probability) is a small clade from Lammerlaw Range, near Lake Mahinerangi, consisting of one COI haplotype and two EF1α alleles.

The range of genetic distances for the COI data are variable across the three major clades; Clade 1 (GD = 0.13 – 1.69%), Clade 2 (GD = 0.13 – 3.67%) and Clade 3 (GD = 0.27 – 2.35%) (Table 4). Average genetic distances between the three clades...
ranged from 3.05% (Clades 1 and 3), 4.31% (Clades 2 and 3) and 4.29% (Clades 1 and 2) (Table 4).

The EF1α data revealed low phylogeographic structuring (Figure 1b). There is virtually no bootstrap or posterior probability support for any nodes in the tree topology shown in Figure 1b. There is high posterior probability support for three very small clades, two located at the Catlins and one near Alexandra (Figure 1b). A mean genetic distance of 0.9% was estimated, using the TVMef + I + Γ model determined from PAUP* and MODELTEST; consistent with the low phylogenetic signal of the EF1α data set. Also, a Tajima’s D statistic of -2.389 (P <0.01) was calculated reflecting a relative excess of low frequency polymorphisms (Fischer et al., 2004).

Of the 19 alleles found in the EF1α data set (Figure 1b) the five most common (alleles a, b, c, d and e; Figure 1b; Figure 2a) were found in 83% of the samples taken from the EF1α data set. Allele a is found mainly in central Otago and Southland, while allele b is located in the west of Southland, near Fiordland (Figure 2b). Alleles a and b are substantially represented in Clade 2, found in 84% of the samples from that population (Figures 1a, 1b). Allele a is situated only once outside Clade 2 at Macraes Flat (NIV27; not shown in Figure 2b) and allele b exists outside Clade 2 only twice, once at Alexandra (NIV 18) and once in Dunedin (NIV29; not shown in Figure 2b). Alleles c and d occur primarily in Canterbury and eastern Otago (Figure 2b), and were found in 83% of the samples from Clade 1 (Figures 1a, 1b). Allele c occurred twice outside Clade 1, once near Cromwell (NIV26) and once in the Catlins (NIV77), while allele d was located once outside Clade 1, near Cromwell (NIV28). Allele e occurs at Piano Flat, Waikaia Bush and is also found at Rigney, the Blue Mountains and the Lammerlaw Range (Figure 2b). Allele f (NIV80, NIV88, and NIV98) is located at the Catlins, while Allele g (NIV58, NIV71) is located at
Chapter Three: Genealogical Concordance across Independent Loci

Alexandra. Allele h (NIV77a, NIV78b) is situated at the Catlins and Allele i (NIV8, NIV20) is located at Old Man Range. Allele j (NIV21) is located at Mt Cook Village and Allele’s k (NIV59a) and l (NIV59b) are located at Lake Ohau. Allele m (NIV60b) is located at Home Hills; Allele n (NIV78a) is located at the Catlins and Allele o (NIV91) is situated at Lammerlaw Range. Allele’s p (NIV96a), q (NIV96b), r (NIV97a) and s (NIV97b) are all located near Alexandra. Alleles f – s are presented in bold text in Figure 1b and underlined (NIV48 is also underlined) in Figure 2b. A complete list of Niveaphasma EF1α alleles from this study is shown in Table 6.

Statistical parsimony networks are shown for the EF1α and COI data sets in Figures 3a and 3b. Outgroups (ASJ1 and THE2) were included in the initial parsimony network analysis but they proved too distant from the ingroup for TCS to root them reliably. Hence they were not shown in the parsimony networks. There is a discrepancy in the number of EF1α alleles, with TCS determining only 18 alleles for the EF1α data. Due to caveats described in TCS, paradoxes may occur and the order of the sequence may change the results of the collapsing (Clement et al., 2000). As a result of this, TCS is unable to distinguish a difference between allele f and allele g (Figure 3a). But based on sequence data and phylogenetic analysis using PAUP* and MRBAYES, there is sufficient evidence to suggest they are two distinct alleles (Figure 1b). It is also apparent that EF1α allele d is too distant for TCS to connect reliably. As such, two ancestral EF1α alleles, alleles d and e are observed in Figure 3a. Similarly, COI haplotype one was too distant for TCS to connect reliably to Clade 1. Due to this, there are two ancestral haplotypes, haplotypes one and eight, represented in Clade 1 for Figure 3b. TCS analysis of the COI data confirms the presence of the three major clades determined from the ML and Bayesian analyses (Figure 1a). The
EF1α allele of NIV48 was left undetermined due to the occurrence of several ambiguous nucleotide sites along its 590 bp EF1α sequence (Figure 3b).

ARLEQUIN was used to perform a mismatch analysis, resulting in raggedness \((r)\) and sum of squared deviation \((SSD)\) values for the COI data set of 0.029 \((P = 0.41)\) and 0.025 \((P = 0.03)\) respectively, for the constant population model. Values under the rapid population expansion model are \(r = 0.029\) \((P = 0.01)\) and \(SSD = 0.219\) \((P = 0.01)\) (Table 5). A significant p-value indicates that we can reject the hypothesis that genetic diversity observed between Niveaphasma individuals is based on a long-term constant sized population demographic. Mismatch analysis was conducted for the EF1α data resulting in \(r\) and \(SSD\) values of 0.203 \((P = 0.80)\) and 0.007 \((P = 0.09)\) respectively, for the constant population model. Values under the rapid population expansion model were \(r = 0.203\) \((P = 0.78)\) and \(SSD = 0.009\) \((P = 0.38)\) (Table 5). P-values for the EF1α mismatch analysis indicates both the constant population and rapid expansion models can be rejected as explanations for the genetic diversity observed between individuals for this gene. The mismatch distribution for the expected and observed results of the COI data set is presented in graphical form in Appendix A. Mismatch results for each of the three major COI clades are given in Table 5.

The Mantel test for matrix correlation between COI genetic distance and linear geographic distance gave a positive correlation coefficient of 0.3829 \((P = 0.001,\) from 1000 randomizations). A plot of genetic distance versus geographic distance from this analysis is shown in Appendix A. This finding does suggest the Niveaphasma population does fit isolation by distance model.
3.4 Discussion

3.4.1 Mitochondrial DNA Indicates Pliocene and Pleistocene Processes have Shaped the Genetic Structuring of Niveaphasma

Based on the results presented in this study, it is apparent *Niveaphasma* has a high level of genetic structuring, similar to that revealed in other South Island invertebrate taxa (Trewick et al., 2000; Buckley et al., 2001; Trewick, 2001; Chinn and Gemmell, 2004; Boyer et al., 2007). The COI mtDNA data revealed three major clades (Figure 1a; Figure 3b), each restricted to a well defined geographical region (Figure 2a). There is a prominent level of divergence between populations from central Otago and Southland with populations from Canterbury (GD = 4.29%) and eastern Otago (GD = 4.31%). The focus of this study was genealogical agreement across independent loci and, for this, molecular clock or dating techniques was not considered a necessary component. However, several phylogeographic studies within the South Island involving invertebrates have employed various molecular dating techniques to estimate sequence divergence (in Mya) based on genetic distances in COI mtDNA. These studies indicate that genetic distances ranging from 1.8 – 3.3% suggest divergence times of 1 - 2 Mya (alpine weta, *Deinacrida connectins*; Trewick et al., 2000; Trewick, 2001; cicada, *Maoricicada campbelli*; Buckley et al., 2001; cockroaches, *Celatoblatta* spp; Chinn and Gemmell, 2004; mite harvestmen, *Aoraki denticulata*; Boyer et al., 2007). An estimated genetic distance of 3 – 6% indicates speciation events in the *Timena* stick insect genus of North America occurred 1.5 – 3 Mya (Sandoval et al., 1998).

In view of these rates of divergence, a conservative estimate for a split between central Otago and Southland populations with Canterbury and eastern Otago
populations would be nearly 2.5 Mya. The timing of this split is significant because it occurs at the Pliocene/Pleistocene temporal boundary, when the formation of glaciers followed on from extensive tectonic processes in the Southern Alps. These glaciers flowed eastwards from the Southern Alps accelerating erosion and building river courses (Fleming, 1979; Stevens et al., 1995), thus isolating populations of *Niveaphasma* in Central Otago, Canterbury and eastern Otago. Using similar rates of divergence implies populations in Canterbury and eastern Otago were split slightly later, about 1.75 Mya. The site of this split occurs at the Taieri Basin, a known tectonic depression influenced by active faults (Litchfield et al., 2002). Timing of this split could indicate tectonic processes within the Taieri Basin during the Quaternary have acted to isolate the Canterbury population from eastern Otago populations.

3.4.2 The Distribution of Nuclear Gene Alleles Supports Observations in Mitochondrial DNA

The EF1α results reveal a high level of genetic structuring within the *Niveaphasma* population, but genetic diversity was more homogenous (GD = 0.9%) when compared to the COI genetic diversity (GD = 2.91%). The low genetic diversity within EF1α is most likely due to a slower rate of evolution compared with the mitochondrial COI. Divergence between isolated populations occurs four times slower at nuclear loci than at cytoplasmic loci (Hoelzer, 1997; Page and Holmes, 1998; Avise, 2000; Hare, 2001) because the effective population size ($N_e$) of nuclear loci is four times the $N_e$ of the mitochondrial loci; the nuclear genome is diploid and alleles are contributed by both sexes. The negative Tajima’s $D$ statistic estimated from the EF1α (-2.389) is a possible indication of an excess of rare variants relative
to standard neutral expectations. One explanation is a departure from the neutral model (Fisher et al., 2004), due to a past population expansion or an ancient population bottleneck (Slatkin and Hudson, 1991; Fay and Wu, 1999; Wakeley and Aliacar, 2001). We were not able to reject the constant population size model using mismatch analysis statistics or a Bayesian approach, but this is possibly due to the low signal at this locus.

The EF1α data shared a significant degree of concordance with the patterns observed in the COI data, providing two lines of evidence for the population history of *Niveaphasma*. This genealogical concordance, emphasized by the geographic distribution of the major clades and EF1α alleles (Figures 2a, 2b), corresponds to a possible population split that occurred sometime in the past and, to a certain degree, these populations have remained separated from that time. This pattern is also evident in the distribution of COI haplotypes where each of the three major clades comprises haplotypes unique to those clades. The only exception is the occurrence of haplotype eight near Cromel Stream (NIV55) which is outside the estimated range of Clade 1 (Figure 2a). However, the extremely close proximity of these major clades and evidence from the EF1α allele distribution suggests that these *Niveaphasma* populations are in contact, especially in Central Otago. Evidence shows EF1α alleles predominately from the Central Otago region are now located in eastern Otago (NIV27 and NIV29; Figure 2b) and EF1α alleles from Canterbury and eastern Otago are moving west into Central Otago (NIV26 and NIV28; Figure 2b). This movement could be a response to the opening of habit after the retreating of glaciers after the last glacial maximum (LGM) within Central Otago (Stevens et al., 1995).

The major partition of the population observed at both loci occurs, for the most part, across several major geographical features of the lower South Island. Most
notably, these include the Hawkdown, Kakanui, Dunstan, Rock and Pillar and the Lammerlaw Ranges. These mountains appear to split *Niveaphasma* populations in eastern Otago. Similarly, the Garvie Mountains, Old Man Range, Hokonui Hills and Umbrella Range form a geographic boundary between *Niveaphasma* populations in Otago and Southland. These geographic features have been noted previously as being possible barriers in the distribution and evolution of South Island taxa (Emerson and Wallis, 1995; Trewick, 2000; Buckley et al., 2001, Wallis et al., 2001; King et al., 2003; Chinn and Gemmell, 2004; Trewick and Morgan-Richards, 2005; Trewick, 2007).

### 3.4.3 *Niveaphasma* shows Genetic Structuring within Canterbury and Eastern Otago

Clade 1 (from the COI data) and allele's c and d (from EF1α data) are almost exclusively restricted to the Canterbury region and Otago region, east of the Hawkdown, Kakanui, Dunstan, Rock and Pillar and the Lammerlaw Ranges. This population is interesting in that subclade 1a occurs across the Waitaki River system—the biogeographic boundary between Canterbury and Otago. Several phylogeographic studies within New Zealand have detected significant genetic breaks across the Waitaki River system (weta, *Deinacrida connectins*, Trewick et al., 2000; cicadas, *Maoricicada campbelli*, Buckley et al., 2001; skinks, *Oligosoma lineocellatum-chloronoton*, Greaves et al., 2007 and *O. maccanni*, O’Neill et al., submitted manuscript). The fact that similar genetic breaks have not occurred in Clade 1 is surprising, considering many of the taxa observed in other studies occupy alpine zones and have poor dispersal capabilities, similar to that of *Niveaphasma*. Taxa from other studies have also had some level of phylogeographic structuring.
attributed to the Pliocene or Pleistocene, similar to that observed with *Niveaphasma* (Trewick *et al*., 2000; Buckley *et al*., 2001; Greaves *et al*., 2007). However, considering the complex biological arrangements of organisms, there are many unforeseen factors which could have allowed *Niveaphasma* from Clade 1 to extend across the Waitaki river system, while at the same time inhibiting other taxa.

A common COI haplotype and shared EF1α allele occurs at Mt Cook (NIV21, 53), Lake Ohau (NIV15, 33, 34, 59 and 73), Ida Range (NIV99), Kakanui Range (NIV6), Macraes Flat (NIV27, 50) and Dunedin (NIV64; shared COI haplotype only). These sites follow a series of extensive water ways and associated low valleys, draining from Lake Ohau at the base of the Southern Alps. The formation of these valleys during the Pleistocene (Stevens *et al*., 1995) may have allowed *Niveaphasma* to spread east from the Southern Alps during this period. Conversely, *Niveaphasma* may have been allowed to expand northwest from eastern Otago into Canterbury as the glaciers retreated following the LGM (Stevens *et al*., 1995).

Evidence from studies of skinks has found that mtDNA haplotypes from eastern Canterbury are quite distinct from those in western Canterbury (Greaves *et al*., 2007; O’Neill *et al*., submitted manuscript). It would be interesting to sample *Niveaphasma* from eastern Canterbury for further evidence in support of recolonization towards the east. The low genetic diversity of Clade 1 (average GD = 0.70%; GD range = 0.13 – 1.69%; Table 4) is in keeping with other studies which have made similar observations regarding the Canterbury region, attributing genetically homogeneous taxa to the unstable and recent geological history of the region (Wallis *et al*., 2001; Waters and Wallis, 2001; Smith *et al*., 2005; Apte *et al*., 2007). There is the possibility that diversity was reduced in this area during the
Pleistocene from glacial outwash which flowed eastwards from the Southern Alps (Stevens et al., 1995) and recolonization has been a recent event (Wallis et al., 2001). Another possibility is that a high rate of gene flow exists within the region, made possible by the numerous unstable braided rivers of the Canterbury Plains preventing differentiation (Wallis et al., 2001; Waters and Wallis, 2001). It should be acknowledged that sampling of Niveaphasma within the Canterbury region occurred primarily in the west, whereas several other studies have focused on more easterly parts where organisms were more prone to the effects of glacial outwash during the Pleistocene (Wallis et al., 2001; Waters and Wallis, 2001; Smith et al., 2005; Apte et al., 2007; Greaves et al., 2007).

### 3.4.4 Evidence of Historical Isolation for Central Otago and Southland Populations of Niveaphasma

Only ~35 km west of the Canterbury population (Clade 1) is a large population (Clade 2) situated in central Otago and Southland (Figure 2a). Despite the close geographical positioning of these two populations, they remain extremely distinct, both in COI genetic diversity and in the nature of their EF1α alleles. The level of genetic diversity within clade 2 (average GD = 1.58%; GD range = 0.13 – 3.67%; Table 4) is much higher than that of the population in Canterbury and eastern Otago (Clade 1). The high genetic diversity observed within Otago and Southland is often attributed to the stable geology of the region allowing for genetically distinct populations to evolve. It has been suggested by some observers that Otago escaped marine inundation during the Oligocene (Fleming, 1979) and Pliocene mountain building played a significant role in the divergence of several other endemic insects of the region (Buckley et al., 2001; Trewick, 2001; Trewick and Wallis, 2001; Chinn
and Gemmell, 2004, Trewick and Morgan-Richards, 2005). There is strong genealogical concordance within the central Otago and Southland populations across the COI mtDNA locus and EF1α nuclear locus. This highlights the high degree of separation that has taken place between *Niveaphasma* populations within the lower South Island. This level of structuring is demonstrated by the split which occurs between the western Southland population with the central Otago and remaining Southland population (Figure 1a; Figure 2a). This is supported by the EF1α data, showing allele b dominant throughout the western Southland population and allele a found in nearly all the central Otago and remaining Southland population (Figure 2b).

Numerous studies have noted the high level of phylogeographic structuring and genetic diversity of taxa within this area of the South Island (alpine weta, *Deinacrida connectins*; Trewick *et al.*, 2000; Galaxiid fish, *Galaxias vulgaris*; Wallis *et al.*, 2001; upland bullies, *Gobiomorphus breviceps*; Smith *et al.*, 2005; freshwater crayfish, *Paranephrops* spp; Apte *et al.*, 2007; endemic skinks, *O. chloronotont*; Greaves *et al.*, 2007; *O. maccanni*; O’Neill *et al.*, submitted manuscript). All the studies have involved the mtDNA locus and have accredited phylogenetic patterns of the area to coincide with Pliocene mountain building leading to habitat diversification (Trewick *et al.*, 2000; Apte *et al.*, 2007). Phylogenetic patterns of the region have also been attributed to the effects of a cooler climate during the Pleistocene, offering alpine taxa a larger environment while at the same time glaciers maintaining genetic diversity by way of isolating populations (Trewick *et al.*, 2000). The major phylogeographic break splitting western Southland with eastern Southland and central Otago populations forms across the Takitimu and Eyre mountain ranges. This is well supported by both mitochondrial and nuclear data (although EF1α allele
b is situated across this supposed geographical boundary on three occasions; Figure 2b). Eastern Otago and the Southland region east of Fiordland escaped major glaciations during the Pleistocene. This presents the likelihood that the western Southland *Niveaphasma* population became isolated by glaciers moving west towards Fiordland (Stevens et al., 1995). More specifically, a glacier which cut Lake Wakatipu sometime during the Pleistocene was suggested as a cause for a similar phylogenetic break between west and east Southland populations of alpine weta (Trewick, 2001).

The COI genetic distance range within Clade 2 (GD = 0.13 – 3.67%; Table 4) is of a similar level found in other studies (alpine weta, *Deinacrida connectins*; Trewick et al., 2000; Trewick, 2001; cicada, *Maoricicada campbelli*; Buckley et al., 2001; cockroaches, *Celatoblatta* spp; Chinn and Gemmell, 2004; mite harvestmen, *Aoraki denticulata*; Boyer et al., 2007) which attribute Pleistocene glacial cycles as maintaining separation of divergent populations following the Pliocene. Mismatch results for Clade 2 ($r = 0.044; P = 0.00$; Table 5) show a smooth, bell-shaped distribution, supportive of a past population expansion which may be evidence indicating recolonization of *Niveaphasma* population Southland from Central Otago as glaciers retreated. This post-glacial recolonization event would lend support to the distribution of EF1α allele a, which is located almost exclusively throughout Central Otago and Southland (Figure 2b), and the close phylogenetic positioning of the Central Otago and eastern Southland population in the COI data (Figure 1a).

### 3.4.5 A Possible Geographical Intermediate Zone Observed for *Niveaphasma* in Eastern Otago
Clade 3 (occupying eastern Otago) comes within 20 km of Clade 2 and ~35 km of Clade 1 and despite this close proximity, there is enough phylogenetic concordance across the mitochondrial and nuclear data to indicate that this eastern Otago population has undergone a considerable amount of separation from other *Niveaphasma* populations. COI genetic diversity is higher within this population (average GD = 1.04%; GD range = 0.27 – 2.35%) compared to the Canterbury population even though the latter occupies an area three times the size, indicating the eastern Otago population is of an older lineage. There is a very high frequency of unique EF1α alleles within this population relative to the other populations (Figure 3a), suggesting genetic admixture may have taken place between two ancestral populations inhabiting this region. The results also indicate Clade 3 may be acting as an intermediate geographical boundary between eastern Otago, Canterbury and Southland, with evidence of this more apparent in the nuclear loci, where distribution boundaries become more ambiguous than those of mtDNA (Figure 2b).

There is some suggestion that overlap in the distribution ranges of the Canterbury (Clade 1) and eastern Otago (Clade 3) populations has taken place, indicated by the shared EF1α allele observed at the Catlins (Figure 2b). It is generally assumed that eastern Otago escaped major glaciations during the Pleistocene (Stevens *et al.*, 1995), however, as mentioned previously, the region is a known tectonic depression, the Taieri Basin, dominated by a series of active faults (Litchfield *et al.*, 2002). There is every possibility that past tectonic processes have reshaped the landscape, allowing secondary contact between populations in eastern Otago and Canterbury. There is also the occurrence of EF1α allele e within Southland indicating contact between east and west Southland populations may be increasing. Phylogenetic structuring was also found between east and west Otago
populations of alpine weta, *Hemideina maori* (King et al., 2003) and skinks, *O. grande* (Berry and Gleeson, 2005); *O. chloronoton* (Greaves et al., 2007). Trewick (2000) indicated the unique phylogeographical position of the Catlins with a study on *Peripatoides*. The Catlins has a series of faults, including the Settlement Fault, along its eastern coast (Hayward et al., 2007). It may be these faults perpetuating isolation between Catlins populations with other eastern Otago populations. This study also reveals unique mitochondrial haplotypes and unique nuclear alleles in *Niveaphasma* at the Catlins, consistent with Trewick’s finding in *Peripatoides*.

### 3.4.6 Genealogical Concordance across Independent Loci Indicates Substantial Genetic Structuring in *Niveaphasma*

Data presented in this study provides good evidence in support of genealogical concordance across two independent loci for the stick insect genus *Niveaphasma*. Significant genetic structuring is apparent within *Niveaphasma* across both genes, with the location of these population divisions also concordant across both genes. Based on conclusions drawn from other studies using molecular dating techniques, this study suggests that tectonic processes during the Pliocene, followed by the extreme glacial episodes of the Pleistocene, may have played a role in the historical isolation of these populations (Trewick et al., 2000; Buckley et al., 2001; Trewick, 2001; Chinn and Gemmell, 2004; Boyer et al., 2007). This study offers a novel approach to phylogeography by incorporating DGGE analysis in order to resolve the heterozygotes within the nuclear locus, and, as a consequence, strongly supports the use of appropriately selected nuclear genes for phylogeographic studies within New Zealand. Although mitochondrial data remains a priority in phylogeography, this
study demonstrates that, used in combination with nuclear phylogeography, many of
the obscurities between taxa at intraspecific levels will become clearer (Hare, 2001).

3.5 References

Akaike, H. 1973. Information theory as an extension of the Maximum Likelihood

Zealand freshwater crayfishes, Paranephrops spp. Molecular Ecology
16:1897-1908.

University Press, Cambridge, MA.

nuclear DNA phylogenies for populations of the Teleost fish Fundulus
heteroclitus. Proceedings of the National Academy of Sciences of the United
States of America 90:9271-9274.

Berry, O., and Gleeson, D. M. 2005. Distinguishing historical fragmentation from a
recent population decline - shrinking or pre-shrunk skink from New Zealand.

denticulate (Arachnida, Opiliones, Cyphophthalmi): a widespread ‘mite


Chapter Three: Genealogical Concordance across Independent Loci


Chapter Three: Genealogical Concordance across Independent Loci


### 3.6 Tables

**Table 1.** Sample location, haplotype and allelic pair information for *Niveaphasma* specimens (NIV#) and outgroup specimens used in this study. Samples designated NIA# refers to recognized *Niveaphasma annulata* specimens. All samples were housed at the New Zealand Arthropod Collection (NZAC), Landcare Research, Auckland following collection. Sample location codes: C – Canterbury; O – Otago; SL – Southland. Coordinates were taken from the NZGD2000 geodatum database.

<table>
<thead>
<tr>
<th>Sample name/Tissue Code</th>
<th>Haplotype (COI)</th>
<th>Allele Pair (EF1α)</th>
<th>Sample location</th>
<th>Latitude, Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIA1</td>
<td>1</td>
<td>d/d</td>
<td>Dunedin water works, Dunedin, O</td>
<td>45.845706, 170.499009</td>
</tr>
<tr>
<td>NIV29</td>
<td>1</td>
<td>b/b</td>
<td>The Cove, Dunedin, O</td>
<td>45.881052, 170.545793</td>
</tr>
<tr>
<td>NIV62</td>
<td>1</td>
<td>d/d</td>
<td>Dunedin water reservoir, Dunedin, O</td>
<td>45.845706, 170.499009</td>
</tr>
<tr>
<td>NIA2</td>
<td>2</td>
<td>a/a</td>
<td>Lake George, Riverton, SL</td>
<td>46.359449, 167.851819</td>
</tr>
<tr>
<td>NIV1</td>
<td>3</td>
<td>b/b</td>
<td>Lake Monowai Rd, Fiordland, SL</td>
<td>45.884674, 167.431919</td>
</tr>
<tr>
<td>NIV2, NIV51</td>
<td>4</td>
<td>a/a, -/-</td>
<td>Rastus Burn Rd, Remarkables, O</td>
<td>45.033332, 168.766673</td>
</tr>
</tbody>
</table>
### Chapter Three: Genealogical Concordance across Independent Loci

<table>
<thead>
<tr>
<th>Loci</th>
<th>Replicate</th>
<th>Location Description</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIV7, NIV52, NIV66</td>
<td>4</td>
<td>Carrick Range, O</td>
<td>45.066664, 169.050003</td>
</tr>
<tr>
<td>NIV10, NIV45</td>
<td>4</td>
<td>Crown Range, O</td>
<td>44.992181, 168.939461</td>
</tr>
<tr>
<td>NIV22, NIV54</td>
<td>4</td>
<td>Matukituki River, near MacPherson Creek, O</td>
<td>44.490772, 168.778357</td>
</tr>
<tr>
<td>NIV49</td>
<td>4</td>
<td>Leaning Rock, Dunstan Mts, O</td>
<td>45.066663, 169.350005</td>
</tr>
<tr>
<td>NIV60, NIV71, NIV75</td>
<td>4</td>
<td>Falls Dam, Home Hills, O</td>
<td>44.893267, 169.936192</td>
</tr>
<tr>
<td>NIV3</td>
<td>5</td>
<td>Piano Flat, Waikaia River Valley, SL</td>
<td>45.549999, 169.016673</td>
</tr>
<tr>
<td>NIV4, NIV46</td>
<td>6</td>
<td>Mt Iron, Wanaka, O</td>
<td>44.683333, 169.166667</td>
</tr>
<tr>
<td>NIV5</td>
<td>7</td>
<td>Bealey Bridge, Arthur’s Pass, C</td>
<td>43.016667, 171.583333</td>
</tr>
<tr>
<td>NIV17, NIV39, NIV42</td>
<td>7</td>
<td>Cass Field Station, Arthur’s Pass, C</td>
<td>43.033578, 171.761076</td>
</tr>
<tr>
<td>NIV6</td>
<td>8</td>
<td>Kakanui Range, O</td>
<td>45.116666, 170.050000</td>
</tr>
<tr>
<td>NIV15, NIV59</td>
<td>8</td>
<td>Freehold Creek, Lake Ohau, C</td>
<td>44.255933, 169.828036</td>
</tr>
<tr>
<td>NIV21, NIV53</td>
<td>8</td>
<td>Govenors Bush Track, Mt Cook Village, C</td>
<td>43.737618, 170.096496</td>
</tr>
<tr>
<td>NIV27</td>
<td>8</td>
<td>Emerald Creek, Macraes Flat, O</td>
<td>45.450000, 170.416667</td>
</tr>
<tr>
<td>NIV33, NIV34, NIV73</td>
<td>8</td>
<td>Temple Stream, Lake Ohau, C</td>
<td>44.107070, 169.816699</td>
</tr>
<tr>
<td>NIV55</td>
<td>8</td>
<td>Five Rivers Farm, Cromel Stream, SL</td>
<td>45.584216, 168.666412</td>
</tr>
</tbody>
</table>
### Chapter Three: Genealogical Concordance across Independent Loci

<table>
<thead>
<tr>
<th>Location Code</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIV64</td>
<td>8</td>
<td>d/d</td>
<td>The Cove, Dunedin, O</td>
</tr>
<tr>
<td>NIV99</td>
<td>8</td>
<td>c/c</td>
<td>Mt Buster, Little Kye Burn, Ida Range, O</td>
</tr>
<tr>
<td>NIV8, NIV20</td>
<td>9</td>
<td>i/i</td>
<td>Shingle Creek, Old Man Range, O</td>
</tr>
<tr>
<td>NIV9</td>
<td>10</td>
<td>a/a</td>
<td>Nicholas Saddle, Eyre Mts, O</td>
</tr>
<tr>
<td>NIV11</td>
<td>11</td>
<td>b/b</td>
<td>Windy Creek, Takitimu Mts, SL</td>
</tr>
<tr>
<td>NIV12</td>
<td>12</td>
<td>a/a</td>
<td>Mt Cardrona, Crown Range, O</td>
</tr>
<tr>
<td>NIV14, NIV40</td>
<td>13</td>
<td>a/a</td>
<td>Railway Bridge, Riverton, SL</td>
</tr>
<tr>
<td>NIV16</td>
<td>14</td>
<td>b/b</td>
<td>Windy Creek, Takitimu Mts, SL</td>
</tr>
<tr>
<td>NIV18, NIV58</td>
<td>15</td>
<td>b/b, g/g</td>
<td>Little Valley Rd, Alexandra, O</td>
</tr>
<tr>
<td>NIV19, NIV32, NIV61, NIV74</td>
<td>16</td>
<td>b/b</td>
<td>Cascade Creek, Eglinton River Valley, SL</td>
</tr>
<tr>
<td>NIV30, NIV57, NIV70</td>
<td>16</td>
<td>b/b</td>
<td>Monowai Flats, Lake Monowai Rd, SL</td>
</tr>
<tr>
<td>NIV35</td>
<td>16</td>
<td>b/b</td>
<td>Princhester Hut, Takitimu Mts, SL</td>
</tr>
<tr>
<td>NIV23</td>
<td>17</td>
<td>b/b</td>
<td>Five Rivers Farm, Cromel Stream, SL</td>
</tr>
<tr>
<td>NIV24</td>
<td>18</td>
<td>a/a</td>
<td>Pisa Range, near Cromwell, O</td>
</tr>
<tr>
<td>NIV63</td>
<td>18</td>
<td>a/a</td>
<td>Dunstan Mts, near Cromwell, O</td>
</tr>
</tbody>
</table>
### Chapter Three: Genealogical Concordance across Independent Loci

<table>
<thead>
<tr>
<th>NIV25</th>
<th>19</th>
<th>a/a</th>
<th>Mt Rosa, Hector Mts, O</th>
<th>45.061769, 168.963608</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIV26</td>
<td>20</td>
<td>c/c</td>
<td>Dunstan Mts, near Cromwell, O</td>
<td>45.050166, 169.389499</td>
</tr>
<tr>
<td>NIV28</td>
<td>21</td>
<td>d/d</td>
<td>Cromwell Gorge, O</td>
<td>45.112163, 169.308809</td>
</tr>
<tr>
<td>NIV36</td>
<td>22</td>
<td>a/a</td>
<td>Sandy Point, Daffodil Bay, SL</td>
<td>46.496071, 168.304730</td>
</tr>
<tr>
<td>NIV37</td>
<td>23</td>
<td>a/b</td>
<td>Lake Hauroko jetty, SL</td>
<td>45.990191, 167.383236</td>
</tr>
<tr>
<td>NIV38</td>
<td>24</td>
<td>a/a</td>
<td>Kawarau Gorge, near Cromwell, O</td>
<td>45.023872, 167.095707</td>
</tr>
<tr>
<td>NIV41</td>
<td>25</td>
<td>a/a</td>
<td>Sandy Point, Daffodil Bay, SL</td>
<td>46.496123, 168.007459</td>
</tr>
<tr>
<td>NIV43</td>
<td>26</td>
<td>a/a</td>
<td>Mt Cardrona, Crown Range, O</td>
<td>44.866667, 168.949999</td>
</tr>
<tr>
<td>NIV44</td>
<td>27</td>
<td>b/b</td>
<td>Lake Hauroko jetty, SL</td>
<td>45.990191, 167.383236</td>
</tr>
<tr>
<td>NIV47</td>
<td>28</td>
<td>a/a</td>
<td>Mt Rosa, Hector Mts, O</td>
<td>45.060799, 168.965009</td>
</tr>
<tr>
<td>NIV48</td>
<td>29</td>
<td>/-/</td>
<td>Piano Flat, Waikaia Bush, SL</td>
<td>45.554168, 169.018977</td>
</tr>
<tr>
<td>NIV50</td>
<td>30</td>
<td>c/c</td>
<td>Emerald Creek, Macraes Flat, O</td>
<td>45.450000, 170.416668</td>
</tr>
<tr>
<td>NIV67</td>
<td>31</td>
<td>b/b</td>
<td>Rd to Lake Hauroko, SL</td>
<td>45.066664, 169.050003</td>
</tr>
<tr>
<td>NIV68</td>
<td>32</td>
<td>b/b</td>
<td>Piano Flat, Waikaia Bush, SL</td>
<td>46.004040, 167.406911</td>
</tr>
<tr>
<td>NIV69, NIV90</td>
<td>33</td>
<td>c/e</td>
<td>Piano Flat, Waikaia Bush, SL</td>
<td>45.554168, 169.018977</td>
</tr>
</tbody>
</table>
### Chapter Three: Genealogical Concordance across Independent Loci

<table>
<thead>
<tr>
<th>NIV Code</th>
<th>Location</th>
<th>Notes</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIV76, NIV83</td>
<td>Dolamore Park, Hokonui Hills, SL</td>
<td>a/a</td>
<td>46.063066, 168.830324</td>
</tr>
<tr>
<td>NIV77, NIV78</td>
<td>Old Coach Rd, Tahakopa river mouth, O</td>
<td>c/h, h/n</td>
<td>46.552606, 169.473507</td>
</tr>
<tr>
<td>NIV98</td>
<td>MacLennan Range, Tahakopa River, Catlins, O</td>
<td>f/f</td>
<td>46.511863, 169.245922</td>
</tr>
<tr>
<td>NIV79, NIV84</td>
<td>Dunsdale, Hokonui Hills, SL</td>
<td>a/a</td>
<td>46.133897, 168.602799</td>
</tr>
<tr>
<td>NIV80, NIV88</td>
<td>Tautuku Bay, Catlins, O</td>
<td>f/f</td>
<td>46.577594, 169.444222</td>
</tr>
<tr>
<td>NIV81</td>
<td>Gunns Bush, Hunter Hills, C</td>
<td>c/e</td>
<td>44.664703, 170.972723</td>
</tr>
<tr>
<td>NIV82</td>
<td>Beaumont-Rongahere Rd, Blue Mts, O</td>
<td>e/e</td>
<td>45.957891, 169.471301</td>
</tr>
<tr>
<td>NIV85</td>
<td>Piano Flat, Waïkaia Bush, SL</td>
<td>e/-</td>
<td>45.554327, 169.018645</td>
</tr>
<tr>
<td>NIV86, NIV87</td>
<td>Waipori Falls, O</td>
<td>d/d</td>
<td>45.929105, 170.055108</td>
</tr>
<tr>
<td>NIV89</td>
<td>Croydon Bush, Hokonui Hills, SL</td>
<td>a/a</td>
<td>46.065613, 168.880689</td>
</tr>
<tr>
<td>NIV91, NIV100</td>
<td>Lammerlaw Range, near Lake Mahinerangi, O</td>
<td>o/o, e/e</td>
<td>45.783337, 169.816669</td>
</tr>
<tr>
<td>NIV92, NIV93</td>
<td>Rigney, between Beaumont &amp; Millers Flat, O</td>
<td>e/e</td>
<td>45.721631, 169.470006</td>
</tr>
<tr>
<td>NIV94</td>
<td>Hokonui Hills, SL</td>
<td>a/a</td>
<td>45.946566, 168.505012</td>
</tr>
<tr>
<td>NIV95</td>
<td>Hokonui Hills, SL</td>
<td>a/a</td>
<td>45.946566, 168.505012</td>
</tr>
<tr>
<td>NIV96</td>
<td>Butchers Dam, Alexandra, O</td>
<td>p/q</td>
<td>45.291801, 169.344677</td>
</tr>
</tbody>
</table>
### Chapter Three: Genealogical Concordance across Independent Loci

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Location</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIV97</td>
<td>Butchers Dam, Alexandra, O</td>
<td>45.291801, 169.344677</td>
</tr>
<tr>
<td>ASJ1</td>
<td>Parau Track, Nihotupu Ridge, Auckland</td>
<td>36.977500, 174.590300</td>
</tr>
<tr>
<td>THE2</td>
<td>Eastbourne, Wellington</td>
<td>41.294200, 174.883200</td>
</tr>
</tbody>
</table>
### Table 2. Oligonucleotide primers used in this study.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Name</th>
<th>Sequence (5’ – 3’)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>COI</td>
<td>TL2-N-3014</td>
<td>TCCAATGCACTAATCTGCCATATTA</td>
<td>Simon <em>et al.</em> (1994)</td>
</tr>
<tr>
<td></td>
<td>C1-J-2195</td>
<td>TTGATTTTTTTGGTCATCCAGAAGT</td>
<td>Simon <em>et al.</em> (1994)</td>
</tr>
<tr>
<td>EF1α</td>
<td>EF1for</td>
<td>CAAGGGTTGGAATGTGGAGAGG</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>EF1rev</td>
<td>CTCTTTGATCTCTGCGAACTTGC</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>EFforGC-clamp</td>
<td>GCGGCCCGCCCGCCCGCCCGCGCCGCCGCCGCCGCCAAGGGTTGGAATGTGGAGAGG</td>
<td>This study</td>
</tr>
</tbody>
</table>
**Table 3.** Sequence statistics for mitochondrial and nuclear loci of *Niveaphasma.*

<table>
<thead>
<tr>
<th></th>
<th>mtDNA</th>
<th>EF1α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sequences</td>
<td>99</td>
<td>107</td>
</tr>
<tr>
<td>Number of sites</td>
<td>762</td>
<td>590</td>
</tr>
<tr>
<td>Number of variable sites</td>
<td>168 (22.0%)</td>
<td>44 (7.5%)</td>
</tr>
<tr>
<td>Number of parsimony informative sites</td>
<td>92 (12.1%)</td>
<td>28 (4.7%)</td>
</tr>
<tr>
<td>%AT</td>
<td>72.5%</td>
<td>50.0%</td>
</tr>
<tr>
<td>TS/TV(^a)</td>
<td>5.339</td>
<td>1.568</td>
</tr>
<tr>
<td>Best AIC model</td>
<td>HKY + I + (\Gamma)^b</td>
<td>TVMeF + I + (\Gamma)^c</td>
</tr>
<tr>
<td>Likelihood score (-ln L)</td>
<td>2619.28</td>
<td>1231.10</td>
</tr>
<tr>
<td>Bayesian likelihood score (-ln L)</td>
<td>2950.21</td>
<td>1379.63</td>
</tr>
<tr>
<td>Proportion of invariable sites (I)</td>
<td>0.6229</td>
<td>0.8292</td>
</tr>
<tr>
<td>Gamma distribution shape parameter ((\Gamma))</td>
<td>0.6788</td>
<td>0.7787</td>
</tr>
<tr>
<td>Number of haplotypes ((h))</td>
<td>48</td>
<td>n/a</td>
</tr>
<tr>
<td>Haplotype diversity ((Hd))</td>
<td>0.93</td>
<td>n/a</td>
</tr>
<tr>
<td>Nucleotide diversity ((\pi))</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>Number of alleles</td>
<td>n/a</td>
<td>19</td>
</tr>
</tbody>
</table>

\(^a\)Transition/Transversion ratio. \(^b\)Hasagawa, Kishino and Yano model with invariable sites and gamma distributed rates. \(^c\)Transversion model with equal nucleotide frequencies, invariable sites and gamma distributed rates.
Table 4. Average and range of genetic distances estimated for the COI data set (762 bp) and EF1α data set (590 bp). Distances were determined using the HKY + I + Γ and TVM + I + Γ models respectively, using parameters determined from MODELTEST and PAUP*. COI mtDNA genetic distances were also determined for the three major clades identified in Figure 1a using the HKY + I + Γ model.

<table>
<thead>
<tr>
<th></th>
<th>All COI data</th>
<th>Clade 1</th>
<th>Clade 2</th>
<th>Clade 3</th>
<th>All EF1α data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average genetic distance</td>
<td>2.91%</td>
<td>0.70%</td>
<td>1.58%</td>
<td>1.04%</td>
<td>0.9%</td>
</tr>
<tr>
<td>Genetic distance range</td>
<td>0.13% - 6.04%</td>
<td>0.13% - 1.69%</td>
<td>0.13% - 3.67%</td>
<td>0.27% - 2.35%</td>
<td>0.18 - 2.68%</td>
</tr>
<tr>
<td>Average genetic distance between clades</td>
<td>Clade 1 and Clade 3 = 3.05%</td>
<td>Clade 1 and Clade 2 = 4.29%</td>
<td>Clade 2 and Clade 3 = 4.31%</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>
Table 5. Mismatch analysis results for the COI (762 bp) and EF1α (590 bp) data sets as well as the three major COI mtDNA clades identified from Figure 1a.

<table>
<thead>
<tr>
<th></th>
<th>All COI data</th>
<th>Clade 1</th>
<th>Clade 2</th>
<th>Clade 3</th>
<th>All EF1α data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant population model</td>
<td>Rapid expansion model</td>
<td>Constant population model</td>
<td>Rapid expansion model</td>
<td>Constant population model</td>
</tr>
<tr>
<td>Raggedness, $r$</td>
<td>0.029</td>
<td>0.029</td>
<td>0.167</td>
<td>0.167</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>$P=0.41$</td>
<td>$P=0.01$</td>
<td>$P=0.54$</td>
<td>$P=0.78$</td>
<td>$P=0.00$</td>
</tr>
<tr>
<td>Sum of squared deviation, $SSD$</td>
<td>0.025</td>
<td>0.219</td>
<td>0.047</td>
<td>0.089</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>$P=0.03$</td>
<td>$P=0.01$</td>
<td>$P=0.37$</td>
<td>$P=0.07$</td>
<td>$P=0.73$</td>
</tr>
</tbody>
</table>

To distinguish between constant population size and rapid population expansion models a raggedness statistic ($r$) is used. This is simply the squared difference between neighboring peaks (refer Figure 3) and is obtained by the equation $r = \sum_{i=1}^{d} (x_i - x_{i+1})^2$ where $d$ is the greatest number of differences between alleles or haplotypes and $x_i$ is the relative frequency of $I$ pairwise differences (Jobling et al., 2004).
Table 6. Position of the 19 nucleotide changes for the 19 alleles determined from the *Niveaphasma* EF1α data set. The number along the top row corresponds to the nucleotide position along the 590 bp sequence of the EF1α gene where the polymorphic site occurred.

<table>
<thead>
<tr>
<th>EF1α Nucleotide position</th>
<th>99</th>
<th>132</th>
<th>147</th>
<th>168</th>
<th>252</th>
<th>273</th>
<th>345</th>
<th>383</th>
<th>427</th>
<th>430</th>
<th>435</th>
<th>439</th>
<th>442</th>
<th>468</th>
<th>494</th>
<th>495</th>
<th>500</th>
<th>568</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele a</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>c</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>d</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>e</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>f</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>g</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>C</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>h</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>i</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>j</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>k</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>l</td>
<td>T</td>
<td>C</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>m</td>
<td>T</td>
<td>C</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>C</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>o</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>p</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>q</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>s</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
3.7 Figure Legends

Figure 1a. Bayesian consensus tree for *Niveaphasma* based on 762 bp from the cytochrome oxidase I (COI) mitochondrial gene. Three major clades are identified within this population; Clade 1 comprises three sub-clades (1a – 1c); Clade 2 comprises nine sub-clades (2a – 2i); Clade 3 comprises three sub-clades (3a – 3c). Shown in parenthesis is the EF1α allelic pair for each sample. Figure 1b. Bayesian consensus tree for *Niveaphasma* based on 590 bp from the Elongation Factor 1α nuclear (EF1α) gene. 19 alleles are identified within this population. Alleles f – s are highlighted in bold typeset. The COI clade to which each EF1α sample belongs is indicated by the circle taken from the Figure legend in Figure 2a. Maximum Likelihood trees were almost identical in topology to the trees shown. Two measures of branch support are indicated with ML bootstrap values on the left (Figure 1a only) and Bayesian posterior probability values on the right in bold typeset (only values over 50 and 0.95 are shown, respectively).

Figure 2a. Location of *Niveaphasma* tissue samples used in this study. The approximate distributions of the three major clades identified from Figure 1a are shown by solid black lines. The Clade legend indicates which of the three major clades each sample belongs to. The approximate distribution of the minor clades from Figure 1a are indicated by broken lines and shading. NIV55 is underlined as it belongs to Clade 1 but is situated near Cromel Stream. Source of map: LENZ hypsometric map 500m elevation. Inset: Main provisional regions of the South Island. Figure 2b. The approximate distributions of the 19 alleles determined from Figure 1b. The distribution of the alleles a – e is indicated by solid black lines. The approximate distribution of allele’s a and b are indicated by broken lines and
shading. The samples representing allele’s f – s are underlined (NIV48 is also underlined). The COI clade to which each EF1α allele belongs is indicated by the Figure legend. Inset: Major geographical features of the South Island.

Figure 3a. Statistical Parsimony network showing the relationship between the EF1α nuclear DNA alleles of *Niveaphasma*. The network was obtained using the Statistical Parsimony algorithm implemented in TCS (Clement *et al.*, 2000). Circles on branches represent the number of mutation events leading to each allele, while numbers on branches indicate the nucleotide position where the mutation occurred. TCS represents the most ancestral allele by a squared box. In some instances alleles are too distant from one another to connect reliably. TCS indicates this by placing more than one haplotype within a squared box. Figure 3b. Statistical Parsimony network for the COI mtDNA data. Circles on branches represent the number of mutation events leading to each haplotype. Unconnected networks for Figures 3a and 3b are represented by a dashed line. Sizes of the ovals in Figures 3a, 3b are not proportional to the number of samples contained within each EF1α allele or mtDNA haplotype. Similarly, branch lengths for both figures are not proportional to the genetic distance between each EF1α allele or mtDNA haplotype.
3.8 Figures

Figure 1a.
Figure 2a.
Chapter Three: Genealogical Concordance across Independent Loci

Figure 2b.
Figure 3a.
Chapter Three: Genealogical Concordance across Independent Loci

Figure 3b.
Chapter Four

General Discussion

4.1 General Discussion

4.1.1 Determining Genealogical Concordance

Recent studies have focused on the genetic structuring of the South Island’s endemic taxa with a particular interest in alpine radiation and diversification (Buckley et al., 2001; Trewick, 2001; Trewick and Morgan-Richards, 2005; Buckley and Simon, 2007). Many of these studies focus on various invertebrate groups within the South Island, due primarily to the very unique alpine biota inhabiting this region of New Zealand. A lack of mammals within New Zealand and many endangered vertebrate species means vertebrates are not always suitable when it comes to assessing genetic diversity (Lloyd, 2003; Berry and Gleeson, 2005). This study therefore provides a
valuable insight into the genetic patterns of the lesser known but equally important vertebrate fauna of this country. At the same time this study also presents the opportunity to investigate the genetic patterns within one of New Zealand’s more distinctive and to a certain extent, more anonymous invertebrate groups.

This study revealed prominent genetic patterns within two endemic skink species and a stick insect complex of which very little was previously known. However, these studies must be placed alongside work on other endemic taxa if we are to convincingly assess the evolutionary forces acting within New Zealand. Detecting statistical significance for the phylogeny of one species is only a starting point in determining accurate genealogical concordance among species. The next important steps are establishing phylogenetic patterns across multiple co-distributed species and across spatial boundaries between traditionally recognized biogeographic regions (Avise, 2000). To advance further, one must then determine genetic agreement at the population level across independent loci, ideally employing both mtDNA and nuclear DNA to do this (Avise, 2000). This study presents the first phylogeographical study within New Zealand to infer intraspecific genealogical concordance across independent loci using mtDNA and resolved nuclear alleles. The findings based on this study offer compelling evidence in support of the use nuclear gene based phylogeography within New Zealand.

Applying mtDNA to phylogeography offers high resolution genetic data, but reveals only some of the historical evolutionary complexities underlying the processes that have shaped genetic diversity within New Zealand (Hare, 2001). Nuclear phylogeography, used alongside mtDNA, presents the opportunity to understand more completely the mosaic of genealogical evolutionary patterns at work within the unique biota of this country.
4.1.2 New Directions for Molecular Evolutionary Studies

After undergoing a rapid rise over the past decade, there is a suggestion that biogeography is becoming hindered by studies too focused on single genes and species resulting in *ad hoc* hypotheses. Subsequently, researchers become too focused on attempting to estimate the timing of speciation or biogeographic events based on these hypotheses (Humphries, 2000; Brooks, 2006). The concern is that this approach is failing to take into account the full range of processes underlying speciation and diversification (Riddle, 2005). There is also the ongoing debate over the almost exclusive use of mtDNA in phylogenetic reconstruction, with many systematists voicing concern over its utility in estimating species evolutionary history (Brooks, 2006). Part of the criticism is directed at the high-rate of evolutionary turnover and maternal-only inheritance of mtDNA which may be distorting the true evolutionary history of populations. Another concern is that introgression plays a more influential role on the patterns of mtDNA heritability than initially predicted. This has created a growing need for phylogeneticists to accept methodologies recognizing introgression and identifying the role it has played on the genetic diversification of a species (Brooks, 2006). Fortunately molecular ecology is witnessing the development of a new generation of analytical methods and the accessibility of high quality molecular data, allowing the field to move to a more sophisticated level (Riddle, 2005).

Part of this new development was observed in recent studies, which found evidence for rapid evolution at the phenotypic and genetic level in response to global climate change (Reush and Wood, 2007). These studies suggest there may be more capability for diverse organisms to respond to natural selection than previously
thought (Grant and Grant, 2002). This idea is enabling molecular ecology to move away from a largely neutral-marker discipline and one reliant on mtDNA, to one characterizing selectively important genetic polymorphisms (Feder and Mitchell-Olds, 2003; Purugganan and Gibson, 2003; Reusch and Woods, 2007).

Studies have found as climate zones rapidly shift under global warming, selection favors the increased dispersal potential of many organisms. An elevated state of vegetative growth in *Sanguisorba minor*, a common European herb, was observed as a response to selection by increased CO$_2$ concentration (Wieneke et al, 2004). The frequency of long-winged forms in two cricket species (*Metrioptera roeselii* and *Conocephalus discolor*) increased in newly established populations at the expanding margin, indicating phenotypic evolution of increased dispersal ability (Thomas et al, 2001). Similar findings were observed in a species of speckled wood butterfly (*Pararge aegeria*), where populations colonizing new zones were larger and had larger thorax muscles compared to populations in the central region of its range (Hill et al, 1999).

These examples highlight the occurrence of evolution in response to environmental change. However, it is equally important to be able to isolate the genetic basis of traits under global-change induced selection (Reusch and Wood, 2007). One such study has found a naturally occurring mutant allele in the circadian clock gene *timeless* of *Drosophila melanogaster*. This mutant allele has risen recently, enhancing the fly’s ability to adapt to seasonal conditions. It is suggested this mutation coincided with post-glacial period colonization of Europe by *D. melanogaster* (Tauber et al, 2007). Another study found populations of European flounder (*Platichthys flesus* L.), were able to adapt to different local environmental conditions based on genetic differentiation in the heat-shock protein gene, *Hsc70*
Chapter Four: General Discussion

(Hemmer-Hansen et al., 2007). Both these studies highlight the usefulness of the candidate gene approach as a potential future direction for molecular evolutionary studies (Hemmer-Hansen et al., 2007).

One problem facing such an approach, however, is the difficulty in targeting new eukaryotic genomes. Very few laboratories are equipped to carry out such a procedure using traditional sequencing methods and only model organisms seem likely candidates (Ellegren, 2008). Regardless of these issues, progress is being made through new technologies, such as genome scans, which use moderate to large numbers of genetic markers to sample or ‘scan’ the desired regions of a selected genome. It is assumed that marker loci displaying a larger population genetic differentiation than expected, under a null neutral model, are under positive selection (Reusch and Woods, 2007). This method was employed in an AFLP-based genome scan in the coastal snail species *Littornia saxatilis*. Genetic differentiation suggested selection was acting across populations physically separated by desiccation and thermal stress levels (Wilding et al., 2001). There is also a new ultra-high throughput sequencing approach available called ‘454-sequencing technology’. Using this technology, the protein-coding sequences of the Glanville fritillary butterfly have been determined. A candidate gene approach then targeted coding regions, thought to infer some phenotypic variability, for microarray analysis (Ellegren, 2008). As observed with this study, there are several benefits of 454-sequencing, one being that it offers direct access to sequence information necessary for candidate gene targeting. Another is that it creates a large pool of sequence variants because large numbers of individuals are used for sequencing, thus providing several genetic markers for further use in population genetic analysis. Thirdly, it allows for the construction of species-specific microarray analysis (Ellegren, 2008). These new findings
surrounding the genetic basis of traits under selection through the onset of environmental change, are set to impact the study of molecular ecology in ways difficult to imagine only a few years ago (Ellegren, 2008).

Molecular evolutionary studies have also witnessed an increase in the number of analysis involving multiple loci per species. An early study by Bernardi et al. (1993), showed the usefulness of using unlinked loci in determining phylogenetic concordance in teleost fish (*Fundulus heteroclitus*) populations. Several recent studies have similarly used independent loci in identifying phylogenetic concordance in pigmy sunfishes (Quattro et al., 2001), phylogeographic concordance in Australian skinks (Dolman and Phillips, 2004) and founder events in sockeye salmon (Ramstad et al., 2007). Studies within New Zealand are also recognizing the importance of using independent loci to discover more accurate historical evolutionary patterns. Both mitochondrial and nuclear loci have been targeted to investigate the evolutionary radiation of alpine cicada (Buckley and Simon, 2007), the role of hybridization in the evolution of the parthenogenetic stick insect, *Acanthoxyla* (Buckley et al., 2008) and as observed with this study phylogenetic concordance within the alpine stick insect complex, *Niveaphasma*.

It is also important that molecular evolutionary studies incorporate concordance across multiple co-distributed species for addressing common biogeographic problems. This is essential within New Zealand, especially in the South Island, where so many studies are focusing on the shared evolutionary and ecological factors shaping the distribution and dispersal of fauna within the region (Avise, 2000). This study alone provides insight into the biogeographic events taking place across co-distributed species within the South Island, suggesting that intraspecific gene trees do map consistently to the same geographic regions across
several species (Avise, 2000). However, this study also identifies phylogeographical incongruence (the Waitaki river system) across these co-distributed species. This highlights the need for a better analytical approach towards incongruence, in order to find the full range of evolutionary events underlying speciation and diversification within New Zealand (Riddle, 2005).

Where biogeography and phylogeography were facing the risk of becoming too heavily focused on neutral markers and mtDNA, new technologies and a departure from traditional methods, are beginning to open up new opportunities for researchers keen to move away from these constraints, thereby, expanding evolutionary studies to more advanced levels.

### 4.2 References


European flounder (*Platichthys flesus* L.) (Advance online publication, doi.10.1038/sj.hdy.6801055).


Appendix A

Figure 1. Mismatch distribution graphs based on Niveaphasma COI mtDNA pairwise differences.

Mismatch distribution based on a long-term constant-sized population model. The open circles show the observed distribution of pairwise differences. The solid black line indicates the expected values (at equilibrium for no recombination) in a stable population.

Mismatch distribution based on rapid population expansion model. The open circles show the observed distribution of pairwise differences. The solid black line indicates the expected values (no recombination) in a growing and declining population.
Figure 2. A Mantels test plot for matrix correlation of COI mtDNA genetic distance versus geographic distance for *Niveaphasma*. A Reduced Major Axis Regression was used to calculate intercept and slope of genetic distance vs. geographic distance. A log genetic distance versus log geographic distance plot was unable to be produced because geographic distances of zero were encountered.
Appendix B

Figure 1. Gel electropherograms showing PCR products following DNA amplification of *Niveaphasma* COI and EF1α genes.

Gel electropherogram of COI PCR products following amplification of *Niveaphasma* samples NIV14 – NIV39

Gel electropherogram of EF1α PCR products following amplification of *Niveaphasma* samples
**Figure 2.** Gel electropherogram following DGGE analysis, showing the resolved alleles of *Niveaphasmas* EF1α genes. A single band indicates the presence of a homozygote, while two bands reveal the presence of a heterozygote.
**Figure 3.** Acrylamide gel photograph of resolved *Niveaphasma* EF1α alleles, following DGGE analysis and silver staining. The bands representing the nine heterozygotes were excised from the gel; the DNA was eluted and used for further PCR and sequencing. The samples with only one band are homozygotes and were included as negative controls. The bands are indicated with arrows.