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

Fine scale population structure through space and time.

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Zoology

Massey University

New Zealand

Elizabeth Emma Daly

2016

Abstract

Terrestrial snails, with their diversity of interspecific forms, have provided biologists with fantastic material to study the evolution of ecotypes and the process of speciation. Snails have the advantage of shells that preserve well and exhibit trait variation readily perceived by taxonomists. Endemic to New Zealand is the genus of giant carnivorous *Powelliphanta* snails and three species of giant herbivorous *Placostylus*. Both genera display a range of phenotypic variation of shells within comparatively small geographic distances. The diversity within these snails has become a matter of high conservation interest, as many lineages occupy small or highly fragmented ranges that render them vulnerable to ongoing habitat loss, and predation by exotic pests. Combining mitochondrial sequence data and genotypes of microsatellite loci I documented the genetic structure within a species complex (*Powelliphanta* Kawatiri). Improved understanding of the distribution of this complex and the level of genetic diversity provided a picture of a naturally fragmented lineage, restricted to a particular ecological zone.

To investigate the evolution of *Placostylus ambagiosus* its mitochondrial genome and that of its sister species *P. hongii* were assembled and annotated. Gene order was consistent between the two *Placostylus* species although it varies slightly within the wider Sigmurethra suborder due to minor tRNA rearrangements. To distinguish the shell shape of spatially distinct populations of *Placostylus ambagiosus* two-dimensional geometric morphometric methods were used. This tool was used to study shell shape evolution through time. Stasis was found to be the most common evolutionary mode, however shell size followed a different model, in one population, an observation which would not be expected if gene flow was preventing local divergence. Investigation into the genetic structure of *Placostylus ambagiousus* (using RADseq) revealed a single admixed population illustrating gene flow had occurred between populations in the recent past. The formation and maintenance of locally adapted populations (ecotypes) within *Placostylus ambagiosus* does not seem to be prevented by gene flow within species.

Acknowledgements

Mary and Steve, I could not wish for better academic supervisors. Your open door policy and supportive, enthusiastic attitudes and belief in me have made this PhD happen. STARY, thank you for your friendship, I hope I'm still welcome to join you for coffee after this. Thanks to my friends in the Phoenix group (past and present), especially my number one Phoenix, Briar Taylor- Smith.

I am grateful to Eddie for advice with analysis, lab work (microsatellite development, primer design and ddRAD wet lab work) but thank you especially for your friendship. Mike and Felix, thanks for doing PhDs at the same time as me sharing ideas, R script, troubles and food with you has kept me going. Cila, thank you for your mentoring and friendship during the earlier parts of my PhD.

Kath, thank you for sharing not only your love of snails with me but also ideas and the extensive *Powelliphanta* tissue collection which you have invested so much time and energy into collecting. Your passion and knowledge of the *Powelliphanta* and your willingness to share it with me has been amazing.

James Crampton, thank you for your help explaining various morphometric methods and models, inviting us on conferences and field trips and for your positivity towards my study. Thank you to Bruce Marshall at Te Papa and the staff at the Auckland War Memorial Museum particularly Heidi Schlumpf and Wilma Blom for your friendly assistance in accessing collections.

I am grateful to the help from Vaughan Symonds for the microsatellite protocol, Trish McLenachan for advice and lab support; Lara Shepherd for all her microsatellite advice; and Marrin Conrady, for installing and testing the stacks pipeline.

Simon and Gillian –thank you for making yourselves available to teach me the quirks of genome assembly, and answer my seemingly endless (and sometimes pointless questions).

Special thanks to office 1.42 and the various configurations of occupants. *Muchas gracias a Julia por ser tan maravillosa y especial.* Also to the Fitzherbert foxes who are an amazing group of inspired woman. I am thankful to have you all on my team.

And to my parents whom I owe so much: Mum and Rex, who provide encouragement, childcare and unwavering support. I am eternally grateful. To my Dad who was unable to see me get to this point but knew I could do it. And finally to my husband Elton, whose support, encouragement and understanding have been constant throughout. I am fortunate to have him as my partner in life. Our two beautiful children Sophia and Oliver bring me joy every day. I love you guys.

This research was supported in part with funding from the New Zealand Department of Conservation, the Julie Alley Bursary and Marsden funding (11-MAU-030).

Contents

Abstract ... i

Acknowledgements ... ii

Chapter 1. Thesis Introduction ... 1

Species variation, range, and structure ... 3
Morphological Evolution in the Fossil Record ... 5
Snail Evolutionary Examples ... 6
Thesis structure ... 9
References ... 11

Chapter 2. Fine scale conservation genetics of the *Powelliphanta* Kawatiri complex land snails. ... 15

Introduction ... 17
Methods ... 20
Results ... 27
Discussion ... 34
References ... 39
Appendix ... 43

Chapter 3. Mitochondrial genome and 45S RNA cassette sequence of New Zealand land snails, *Placostylus ambagiosus* and *Placostylus hongii* (Mollusca: Gastropoda: Orthalicoidea: Bothriembryontidae: Placostylinae) ... 47

Introduction ... 49
Methods ... 54
Results and Discussion ... 56
Conclusion ... 67
References ... 68

Chapter 4. Optimising landmark selection for geometric morphometric analysis of the New Zealand land snail *Placostylus ambagiosus*. ... 73

Introduction ... 75
Methods ... 78
Results ... 83
Discussion ... 91
References ... 92

Chapter 5. Understanding morphological stasis in a heterogenous environment. ... 95

Introduction ... 97
Methods ... 103
Results ... 110
Discussion ... 121
References ... 124
Appendix ... 128

Chapter 6. Investigating population structure and gene flow within the New Zealand land snail *Placostylus ambagiosus* using anonymous nuclear markers. ... 133

Introduction ... 135
Methods ... 141

Results ... 148
Discussion ... 156
References ... 163
Appendix ... 167

Chapter 7. Thesis Discussion ... 173
Species concepts and conservation ... 175
Stasis, gene flow and the maintenance of ecotypes ... 179
References ... 182

Chapter 1. Thesis Introduction

Species variation, range, and structure

Individuals within a species are not identical and it is populations that evolve (Darwin 1859). Thus it is important to understand how geographical patterns of morphological variation and allele frequency differences among populations arise and are maintained. Examining the variation within species, the division and connections between populations, and the differences among individuals will aid our understanding of the principals and processes of evolution (Mayr 1963; Nei, 1987; Slatkin 1987).

> "*Hence I look at individual differences, though of small interest to the systematist, as of high importance for us, as being the first step towards such slight varieties as are barely thought worth recording in works on natural history.*"
> Darwin (1860, Chapter 2, page 51)

The geographic range of a species is regarded as one of the most significant factors influencing genetic diversity and its distribution (Charlesworth 2009; Hamrick et al. 1989). A species' geographic range is determined by a combination of biotic, abiotic and historical factors (MacArthur 1972; Udvardy 1969). The range a species occupies depends not only on environmental variation but also on when and where it originated and what barriers to dispersal it encountered. Habitats are ephemeral and shifting and barriers emerge, change, and disappear via a range of processes, including climatic cycling, sea level changes, ecological succession, competition, natural disasters, and major geological events. These can result in extinction, adaptation and/or a change in species' geographic range (Berg et al. 2010; Bulgarella et al 2014; Wisz et al. 2013).

On a smaller scale, the way that species are distributed within their geographic range is mostly determined by ecological factors (Hutchinson 1957) including climate, predators, competitors and usable resources. These factors reduce an organism's fundamental niche to the fraction that may be exploited – the realised niche of a species (Hutchinson 1957). As a result population structure may appear

essentially continuous over large geographic areas or be patchy with areas of relatively high abundance separated by areas in which a species is rarely or never found.

Range margins are set by the limits of the species' ecological tolerances (Woodward & Kelly 2003) and adaptation at the periphery of a species' range might be prevented by gene flow from the centre of the species' range where individuals are more numerous (Haldane 1956). The flow of alleles from populations in optimal habitat (source populations) into populations in suboptimal habitat (sink populations) could similarly prevent populations from adapting to different environments and thus limit the species' distribution (Case et al 2000; Holt 1996; Kawecki 1995; Kirkpatrick & Barton 1997; Lenormand 2002). This same process results in species cohesion that might result in morphological stasis when viewed through time. However, if selection on populations experiencing different conditions is strong it can overcome the constraining nature of gene flow (Slatkin 1987).

Gene flow between populations is no longer considered a barrier to speciation but part of the process of speciation (Abbott et al. 2013). Earlier theory predicted that gene flow inhibited divergence by preventing natural selection and genetic drift from maintaining local differentiation (Mayr 1963). However gene flow is neither uncommon nor necessarily an impediment to divergence and natural selection can cause populations to diverge and adapt regardless of species distributions (Mallet 2001; Nosil 2008; Wu 2001). Strong selection on particular traits can produce adaptations to local conditions in the presence of substantial gene flow (Ehrlich & Raven 1969).

Natural selection can be much more effective than genetic drift in either preventing or establishing local differences. Selection in favour of the same alleles or the same traits in different populations would produce geographic uniformity regardless of any gene flow. Selection favouring different alleles or traits in different locations will succeed in producing local differences reflecting adaptations to local conditions increasing an individual's fitness (Haldane 1930; Nagylaki 1975). If selection is strong enough, divergence and the formation of ecotypes will occur, which many

now regard as a stage in the speciation process (Mallet 2008; Räsänen & Hendry 2008; Rundle & Nosil 2005). Natural selection is again considered the driving force of speciation (Fiorentino et al. 2013). Alternative models of speciation without selection have been based on random genetic drift and sexual selection (Nosil 2012) and while these may be important mechanisms, it is likely very rare that they are the sole driver of speciation (Fiorentino et al. 2013). Genetic drift causes correlation between reproductive isolation and time, absence of ecological divergence and population bottlenecks (e.g. founder effect). Theoretical and empirical studies suggest that these processes alone rarely drive speciation (Coyne & Orr 2004).

Morphological Evolution in the Fossil Record

Genetic analysis allows an estimate of recent population structure and gene flow, but inferences about persistence and form change come from fossils. Fossils provided the first major insight in evolution and they continue to provide documentation of major features of form change through geological time. Morphological differences found between species have long formed the basis of taxonomy and analysis of phenotypic traits is an essential complement to genetic studies. Analysis of trait variation allows the comparison of modern and fossil populations. The geometric morphometric approach allows inclusion of continuous trait information in analysis rather than relying heavily on one or few, often subjectively chosen, categorical morphological traits. Geometric morphometric analysis has been applied in a wide range of biological studies (Hills et al. 2012; Webster & Sheets 2010; Zelditch et al. 2012) and has the advantage that shape variation can be studied independently of size. In addition, the approach has strong and well-understood statistical and theoretical underpinnings (Bookstein 1996).

By examining fossils of the same lineage sampled at many time points one can study the evolution of morphological traits. Tools have been developed that allow fossil time series to be categorised into different modes based on statistical fit (Hunt 2006; Hunt et al. 2015). The three basic statistical models of morphological evolution

examined are evolutionary stasis (limited fluctuations from a mean state through time); an unbiased random walk (representing stochastic change in a trait through time; and a generalised random walk (representing directional change in a trait with some stochastic variation considered; (Hunt 2006). If the time series (sequence) for the same evolutionary lineage is sampled frequently enough, it is possible to incorporate more than one model to describe the morphological evolution observed (Hunt et al. 2015). Comparing model fit in more than 400 fossil time series has shown that morphological stasis is much more common than directional change (Hopkins & Lidgard 2012; Hunt et al. 2015; Voje 2016).

Snail Evolutionary Examples

Molluscs and snails in particular, provide opportunities to study variation within species, ecotypes and speciation. Previously, studies of evolutionary radiations have been dominated by vertebrates and insects although the fossil record for snails, particularly benthic marine species, is much better (Bookstein 1996). Morphological variability, low vagility, the longevity of several species (particularly island endemics and the large New Zealand species), rarity of reproductive isolation, definitive adult size for morphometric work all make several snail genera ideal model systems for the study of evolutionary biology.

Snail shells also exhibit enormous phenotypic diversity again both within and among species. The diversity of shell morphology is the result of different processes in different lineages. For example in some snails variation in shell form can be a plastic response induced by the interaction of abiotic and biotic stimuli (Appleton & Palmer 1988; Trussell 2000), but in other snail taxa shell shape variation is a fixed inherited trait (Dowle et al. 2015; Johannesson & Johannesson 1996; Stankowski 2013). The high morphological and molecular diversity within snail lineages, combined with relatively high rates of fossilisation provide excellent opportunities to understand the role of natural selection and other evolutionary processes in shaping present-day phenotypic diversity.

Cerion land snails with highly variable shell morphologies among populations are perhaps among the best-studied land snail species. In 1912 Paul Bartsch collected 40,000 live specimens (Woodruff & Gould 1987). More than 600 "species" were subsequently named (Woodruff & Gould 1980) and research has continued with this group until recent times. Genetic and morphometric analysis combined with large sample sizes revealed patterns of morphological variation were simpler than initially thought and the number of species was subsequently revised (Woodruff & Gould 1980). It has been inferred that the high morphological variability resulted from local morphotypes (area effects) arising through three different processes: parapatric interdemic divergence; differentiation in isolation and secondary hybridisation; distant transport of propagules (thought to be relatively common due to the hurricanes that frequent the area). It was also proposed that the genetic environment might be as important as the ecological environment in controlling development of some local morphotypes, narrow zones of allopatric hybridisation and genetic anomalies associated with hybrid zones imply pressure of different co-adapted gene complexes.

In addition to observation of natural variation, Bartsch (Bartsch 1920) transplanted two forms of the Bahamian *Cerion glans* complex to areas in the Florida Keys, then devoid of the native species *C. incanum,* in an attempt to test the Lamarckian transformation hypothesis. Translocated snails remained true to original morphotypes but all populations were eventually replaced by the native *C. incanum*. Woodruff and Gould (1987) studied one of these colonies *C. casablancae* that was considered a biological species despite its ability to hybridise with another taxa when brought together. Approximately 50 years after Bartsch's transplantation of 55 snails a detailed study of the outcomes from hybridisation was undertaken. Genetic hybrids were found to occupy an area larger that that of morphological hybrids (a pattern subsequently found in many other hybrid zones) and suggested that *C. casablancae* had expanded its range since hybridisation began.

Cerion display high morphological variability which Bartsch's translocation experiments show is not the result of a plastic response to environmental factors, much of the resulting variation is due to phylogenetic and physical constraints

(Gould 1989), and local selection. *Cerion* has many populations with local morphotypes (area effects) and it is known that "species" hybridise readily on contact. This combined with an unusually good fossil site within areas with distinct morphotyes, allowed for a historical explanation of immigration and introgression to explain the flat-topped populations of a usually columnar group. Alleles found in populations within the area effect of *C. columna* were not detected within other populations of the species suggesting that these alleles represent introgressed genes produced by interspecific hybridisation. Morphometric analysis of fossil and modern shells provides support for the conclusion that hybridisation occurred within the area of the local morphotype with individuals translocated from the eastern area of Cuba closest to this area; *C. dimidiatum* (Gould & Woodruff 1990). Previous studies of *Cerion* were able to trace the source of anomalous alleles within distinct local morphotypes to known species at nearby locations (Gould & Woodruff 1986; Woodruff & Gould 1987).

Rapid morphological change over several thousand years has been observed in *Mandarina* snails of the Bonin Islands near Japan (Chiba 1996). *Mandarina* snails have an excellent fossil record over 40,000 years to the present. Morphological analysis shows stasis had been interrupted by two periods of rapid change occurring simultaneously in all lineages. The earlier phase of rapid change occurred shortly after the extinction of one of the five species that had previously maintained stasis. The more recent bout of rapid change involved two examples of morphological shift within lineages (pseudo-extinctions) and a branching event in which one species became extinct. Prior to the extinction event all lineages existed in sympatry and are considered different biological species with rapid evolution of reproductive isolation (Chiba 1998).

In this thesis I examine patterns of spatial and temporal structure in New Zealand land snails using genetics and morphological variation from extant and fossil populations. Understanding population differentiation within a single species (or lineage) aids the understanding of evolution because small differences among populations might be the first steps towards new ecotypes, and new ecotypes might be the first steps to new species (Mallet 2008). The processes in both time and space

that produce local change could be the same processes that lead to the formation of new species. If I can determine the extent of evolutionary changes in one lineage then the role of selection and gene flow during speciation can be inferred.

Thesis structure

In chapter 2, fine scale genetic structure of the carnivorous land snail *Powelliphanta* was investigated. Sampling from the known range of *Powelliphanta* Kawatiri complex using short fragment mtDNA (622 bp of COI) and nuclear markers (five microsatellite loci) provide evidence of the structure and cohesion of the complex.

In chapters 3 and 4 tools were developed to investigate population structure of the herbivorous land snail, *Placostylus ambagiosus*. The snail genus *Placostylus* is present in the western Pacific; New Zealand, Vanuatu, Fiji, Papua New Guinea, Solomon Islands, Lord Howe and New Caledonia. In New Caledonia, shell shape differences of *Placostylus* lineages are maintained in sympatry and show ecotype differentiation (Dowle et al. 2015). New Zealand *Placostylus* are important species as large, endangered, endemic invertebrates with cultural significance to local Maori, and have been under active management since the early 1980's (Barker et al. 2016).

In chapter 3 the whole mitochondrial genome of two *Placostylus* species were sequenced and characterised. From the same two species 45S nuclear ribosomal cassette DNA sequences were also aligned and compared. In chapter 4 two-dimensional morphometric methods to characterise shell shape were developed for the same two *Placostylus* species. A set of geometric landmarks and sliding landmarks were optimised for morphometric analysis of shells from modern and fossil populations.

In chapter 5, I examined shell shape and size variation through 5-40 thousand years of *Placostylus ambagiosus* evolution. It was possible to decompose the species trajectory into local (within habitat) components and analyse multiple mathematically independent shape traits in the same lineage in two separate time series. This allowed me to investigate competing explanations for morphological

stasis. I used geometric morphometric analysis of shell shape and size from extant and fossil populations and compared the fit of three models of morphological evolution (random walk, directional, stasis) using a maximum likelihood approach.

In chapter 6, Extant populations of *Placostylus ambagiosus* were examined for both shell shape variation and anonymous nuclear markers (SNPs). Using a modified version of the double digest restriction-site associated DNA sequencing method (ddRAD), I have developed a reduced genome library from multiple populations of *Placostylus ambagiosus* snails. Although I have data for few individuals, the large number of loci (>1000) supported gene flow between populations for neutral markers, despite evidence of population differentiation in shell shape.

I would like to acknowledge here the work of Eddie Dowle in chapter two: development of microsatelllite protocol and chapter six: development of anonymous nuclear markers. Eddie performed the final PCR step prior to cloning, the cloning of microsatellites and intial primer design steps of the microsatellite protocol discussed in chapter two and all of the wetlab work from DNA extraction to sending for sequencing for the anonymous nuclear makers in chapter six.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246.

Appleton, R. D., & Palmer, A. R. (1988). Water-borne stimuli released by predatory crabs and damaged prey induce more predator-resistant shells in a marine gastropod. *Proceedings of the National Academy of Sciences*, *85*(12), 4387–4391.

Barker, G. M., Brodie, G., Bogitini, L., & Pippard, H. (2016). Diversity and current conservation status of Melanesian–New Zealand placostyline land snails (Gastropoda : Bothriembryontidae), with discussion of conservation imperatives, priorities and methodology issues. *Pacific Conservation Biology*, *22*(3), 203.

Bartsch, P. (1920). *Experiments in the breeding of cerions.* (No. 282). Washington, DC: Carnegie Institute.

Berg, M. P., Kiers, E. T., Driessen, G., Van Der HEIJDEN, M., Kooi, B. W., Kuenen, F., … Ellers, J. (2010). Adapt or disperse: understanding species persistence in a changing world. *Global Change Biology*, *16*(2), 587–598.

Bookstein, F. (1996). Applying landmark methods to biological outline data. In *Image fusion and shape variability* (pp. 59–70). Leeds, U.K.: University of Leeds.

Bookstein, F. L. (1996). Biometrics, biomathematics and the morphometric synthesis. *Bulletin of Mathematical Biology*, *58*(2), 313–365.

Bulgarella, M., Trewick, S. A., Minards, N. A., Jacobson, M. J., & Morgan-Richards, M. (2014). Shifting ranges of two tree weta species (*Hemideina spp.*): competitive exclusion and changing climate. *Journal of Biogeography*, *41*(3), 524–535.

Case, T. J., Taper, M. L., & Barton, A. E. N. (2000). Interspecific Competition, Environmental Gradients, Gene Flow, and the Coevolution of Species' Borders. *The American Naturalist*, *155*(5), 583–605.

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195–205.

Chiba, S. (1996). A 40,000-Year Record of Discontinuous Evolution of Island Snails. *Paleobiology*, *22*(2), 177–188.

Chiba, S. (1998). Genetic variation derived from natural gene flow between sympatric species in land snails (*Mandarina*). *Heredity*, *80*(5), 617–623.

Coyne, J., & Orr, H. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.

Darwin, C. (1859). *On the origin of the species by natural selection*.

Dowle, E. J., Morgan-Richards, M., Brescia, F., & Trewick, S. A. (2015). Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. *Molecular Ecology*, *24*(16), 4205–4221.

Ehrlich, P. R., & Raven, P. H. (1969). Differentiation of populations. *Science*, *165*, 1228–1232.

Fiorentino, V., Manganelli, G., Giusti, F., Tiedemann, R., & Ketmaier, V. (2013). A question of time: the land snail *Murella muralis* (Gastropoda: Pulmonata) reveals constraints on past ecological speciation. *Molecular Ecology*, *22*(1), 170–186.

Gould, S. J. (1989). A Developmental Constraint in *Cerion*, with Comments of the Definition and Interpretation of Constraint in Evolution. *Evolution*, *43*(3), 516–539.

Gould, S. J., & Woodruff, D. S. (1986). Evolution and systematics of *Cerion* (Mollusca, Pulmonata) on New Providence Island. Retrieved from

Haldane, J. B. S. (1930). A mathematical theory of natural and artificial selection. (Part VI, Isolation.). *Mathematical Proceedings of the Cambridge Philosophical Society*, *26*(2), 220–230.

Haldane, J. B. S. (1956). The Relation between Density Regulation and Natural Selection. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *145*(920), 306–308.

Hamrick, J. L., Godt, M., Brown, A. (Ed), Clegg, M. (Ed), Kahler, A. (Ed), & Weir, B. (Ed). (1989). Allozymic diversity in plant species. In *Plant population genetics, breeding and gentic resources* (pp. 43–63).

Hills, S. F. K., Crampton, J. S., Trewick, S. A., & Morgan-Richards, M. (2012). DNA and Morphology Unite Two Species and 10 Million Year Old Fossils. *PLOS ONE*, *7*(12), e52083.

Holt, R. D. (1996). Adaptive Evolution in Source-Sink Environments: Direct and Indirect Effects of Density-Dependence on Niche Evolution. *Oikos*, *75*(2), 182–192.

Hopkins, M. J., & Lidgard, S. (2012). Evolutionary mode routinely varies among morphological traits within fossil species lineages. *Proceedings of the National Academy of Sciences*, *109*(50), 20520–20525.

Hunt, G. (2006). Fitting and comparing models of phyletic evolution: random walks and beyond. *Paleobiology*, *32*(4), 578–601.

Hunt, G., Hopkins, M. J., & Lidgard, S. (2015). Simple versus complex models of trait evolution and stasis as a response to environmental change. *Proceedings of the National Academy of Sciences*, *112*(16), 4885–4890.

Hutchinson, G. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, (22), 415–427.

Johannesson, B., & Johannesson, K. (1996). Population differences in behaviour and morphology in the snail *Littorina saxatilis*: phenotypic plasticity or genetic differentiation? *Journal of Zoology*, *240*(3), 475–493.

Kawecki, T. J. (1995). Demography of source—sink populations and the evolution of ecological niches. *Evolutionary Ecology*, *9*(1), 38–44.

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a Species' Range. *The American Naturalist*, *150*(1), 1–23.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183–189. MacArthur, R. H. (1972). *Geographical Ecology: Patterns in the Distribution of Species*. Princeton University Press.

Mallet, J. (2001). The speciation revolution. *Journal of Evolutionary Biology*, *14*(6), 887–888.

Mallet, J. (2008). Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1506), 2971–2986.

Mayr, E. (1963). *Animal species and evolution*. Cambridge, Massachesetts: Belknap Press of Harvard University Press.

Nagylaki, T. (1975). Conditions for the Existence of Clines. *Genetics*, *80*(3), 595–615.

Nei, M. (1987). *Molecular evolutionary genetics*. Columbia university press.

Nosil, P. (2008). Speciation with gene flow could be common. *Molecular Ecology*, *17*(9), 2103–2106.

Nosil, P. (2012). *Ecological Speciation*. Oxford University Press.

Räsänen, K., & Hendry, A. P. (2008). Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, *11*(6), 624–636.

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*(3), 336–352.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*(4803), 787–792.

Stankowski, S. (2013). Ecological speciation in an island snail: evidence for the parallel evolution of a novel ecotype and maintenance by ecologically dependent postzygotic isolation. *Molecular Ecology*, *22*(10), 2726–2741.

Trussell, G. C. (2000). Predator-induced plasticity and morphological trade-offs in latitudinally separated populations of *Littorina obtusata*. *Evolutionary Ecology Research*, *2*(6), 803–822.

Udvardy, M. D. F. (1969). *Dynamic Zoogeography, iwth Special Reference to Land Animals; Illustrated by CS Papp.* New York: van Nostrand Reinhold.

Voje, K. L. (2016). Tempo does not correlate with mode in the fossil record. *Evolution*, *70*(12), 2678–2689.

Webster, M., & Sheets, H. D. (2010). A practical introduction to landmark-based geometric morphometrics. *Quantitative Methods in Paleobiology*, *16*, 168–188.

Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., … Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, *88*(1), 15–30.

Woodruff, D. S., & Gould, S. J. (1980). Geographic differentiation and speciation in *Cerion* - a preilminary discussion of patterns and processes. *Biological Journal of the Linnean Society*, (14), 389–416.

Woodruff, D. S., & Gould, S. J. (1987). Fifty Years of Interspecific Hybridization: Genetics and Morphometrics of a Controlled Experiment on the Land Snail *Cerion* in the Florida Keys. *Evolution*, *41*(5), 1022–1045.

Woodward, F. I., & Kelly, C. K. (2003). Why are species not more widely distributed? Physiological and environmental limits. In *Macroecology: concepts and consequences* (pp. 239–225). Oxford: Blackwell.

Wu, C.-I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, *14*(6), 851–865.

Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric Morphometrics for Biologists: A Primer*. Academic Press.

Chapter 2. Fine scale conservation genetics of the *Powelliphanta* Kawatiri complex land snails.

Introduction

Molluscs have the dubious honour of having the highest number of documented animal extinctions of any major taxonomic group (Lydeard et al. 2004). They suffer the dual disadvantages of extreme sensitivity to habitat modification and pollution, and under-recognition of taxonomic diversity. The grossly disproportionate distribution of taxonomic and conservation effort toward vertebrates and higher plants (Gaston 1992) means that we remain ignorant of the ecology of most mollusc species (Lydeard et al. 2004). Among gastropods, the criteria used to define taxa have traditionally been shell morphology and spatial distribution. However, geographic distance does not provide direct evidence of the evolutionary scale or causes of population partitioning or the extent to which this reflects breakdown in genetic cohesion. Landscapes that appear homogeneous, by virtue of topographic uniformity for example, might encompass cryptic ecological heterogeneity and thus evolutionary (taxonomic) diversity among the inhabitants. Similarly, topographic variation might not be as influential as other factors such as vegetation cover. Habitat fragmentation generates physical fragmentation of populations and poses a serious threat to the genetic diversity of a species. Effects of habitat fragmentation on patterns of genetic diversity and the genetic structure of resulting populations are difficult to predict (Gibbs 2001; Nelson-Tunley et al. 2016; Ricketts 2001; Willi et al. 2007). Consequently, understanding and ameliorating impacts of anthropogenic habitat activity requires an understanding of the underlying "natural" patterns and processes.

Here we examine the link between population fragmentation and landscape, in a group of endemic land snails that occupy mountainous habitat on the West Coast of South Island, New Zealand. *Powelliphanta* O'Connor, 1945 is a genus of carnivorous snails belonging to the family Rhytididae (Mollusca; Gastropoda; Pullmonata) that is distributed in the Southern Hemisphere (Powell 1979; Spencer et al. 2009)**.** Although the systematics and nomenclature of *Powelliphanta* have been questioned (Climo 1978), recent and ongoing molecular and morphological studies (Trewick et al. 2008; Walker 2003; Walker et al. 2008) provide support for much of Powell's classification of 41 taxa (9 species, 34 subspecies, 4 forms; (Walker 2003). In

contrast, diversity within other parts of the Rhytididae is less well resolved by shell morphology (Spencer et al. 2006).

The *Powelliphanta* radiation is endemic to New Zealand, and generally associated with moist, often calcium rich soils, with deep leaf litter. Most species inhabit lowland forested areas, although several alpine taxa are also recognised; mostly from the West Coast of South Island and one in North Island, on Mt Taranaki. *Powelliphanta* species typically have small ranges that render them vulnerable to habitat loss and predation, and all members of this genus are afforded high conservation status (Meads et al. 1984) with at least one recent species extinction (KJW pers comm).

Since the Holocene, but prior to waves of human settlement in the 13th and 19th centuries, the New Zealand landscape was dominated by near contiguous wet forest except in upland areas (Trewick & Morgan-Richards 2009). The increased deforestation that followed European settlement led to the isolation of many snail populations from one another. The greatest threats to *Powelliphanta* continue to be habitat fragmentation and modification associated with forest clearance, open cast mining, acid mine drainage, and predation by introduced mammals (pigs, possums and rats (Walker 2003)).

The *Powelliphanta* Kawatiri complex [1] occurs in fragmented populations over a relatively small area (<11,250km^2) in the West Coast region. Named for the Kawatiri River which flows from Lake Rotoiti entering the Tasman Sea at Westport. Members of the Kawatiri complex are characterised by an inky to bright blue mantle, and blue/purple mucous. Morphological variation of shell size, patterning and colouration initially led to the recognition of eight sub-groups (Walker 2003). Some more recently discovered populations have been assigned to tag-named groups. Tag-names assigned to putative taxa are not subject to International Codes of

[1] Powell (1979) treated these taxa as *Powelliphanta rossiana patrickensis* but Walker

Nomenclature (iczn.org) and therefore represent only observations of spatial distribution and morphological variation. A problem with tag-names is that associated data is minimal so that it is unclear which biological entities they represent (Leschen et al. 2009).

Most populations of the Kawatiri complex exist at high elevation, above the tree line. Field searches over the last 20 years have increased the known ranges and number of populations, but have not revealed snails in the intervening lower elevation habitat. The altitudinal range over which these snail populations occur is in fact quite broad, extending from 550 to 1500 m asl with a single unusually low population at just 50m asl near the Buller River. Elevation is an inexact proxy for habitat distribution that is strongly influenced by local rock type and drainage. For instance, *P. patrickensis* populations at 550–600 m asl on the Stockton and Denniston plateaux are nevertheless ecologically situated above the local tree line because high rainfall and acidic coal measure soils have depleted soil nutrients and supressed forest growth (Overmars et al. 1992).

The fragmented distribution of these *Powelliphanta* Kawatiri complex snail populations may be the result of either of two processes. The Kawatiri complex might be high-elevation specialists whose evolution and distribution originated during the formation of the Southern Alps mountain ranges about 5 mya (Pliocene), and subject to latitudinal and elevation shifts during Pleistocene climate cycling (Trewick et al. 2000). Alternatively, isolation and population genetic structure might be younger due to anthropogenic habitat modification involving deforestation and introduction of mammalian predators that extirpated snails from lowland areas. Sign of introduced predators such as possums is observed less often in high elevation snail habitat than in the surrounding lowlands (Pers Comm K Walker, New Zealand Department of Conservation).

Here I use spatial distance, mtDNA and microsatellite variation over a small geographical area to explore the genetic structure of *Powelliphanta* Kawatiri complex populations. Should the current pattern of population fragmentation result from recent isolation we expect population genetic structure to reflect a pre-existing

pattern of isolation by distance, for all markers (Charlesworth 2009). Alternatively, fragmentation resulting from high-elevation specialisation over a geological timescale would be inferred if populations were differentiated at some loci but retain isolation by distance only at slower evolving loci. Isolated species would be expected to have concordant patterns of nuclear and mitochondrial markers and morphology due to relatively long isolation.

Methods

Material

I explored the genetic structure of a sample of *Powelliphanta* snails representing a monophyletic clade previously revealed by preliminary analysis of mtDNA sequence data (Trewick et al. 2008). Snails in this clade occur in allopatric (fragmented) populations in a contiguous but topographically heterogenous landscape (Figure 2.1). They vary in terms of shell shape, size and patterning that is currently accommodated by eight tag-names and one species; *Powelliphanta patrickensis* (Trewick et al. 2008; Walker 2003).

Snails were located in the field by hand searching. In some cases, whole specimens were collected, killed by freezing and stored at -80°C for genetic analysis. The majority of samples used in the present analysis were collected between 2004 and 2008, using tissue biopsies taken from live snails that were released at their collection location (Trewick et al. 2008). Tissue samples were stored in 95% ethanol at 4°C. More than 600 tissue samples of *Powelliphanta* exist in the collection at present, of which 126 are part of the Kawatiri complex lineage. 101 snails that amplified reliably for COI and microsatellites were used in genetic analysis. GPS coordinates were recorded for each sample with data on mucous and mantle coloration, shell morphology and coloration. The long lifespan and overlapping generations of these snails mean one can treat samples from different years as potentially the same generation for downstream analysis.

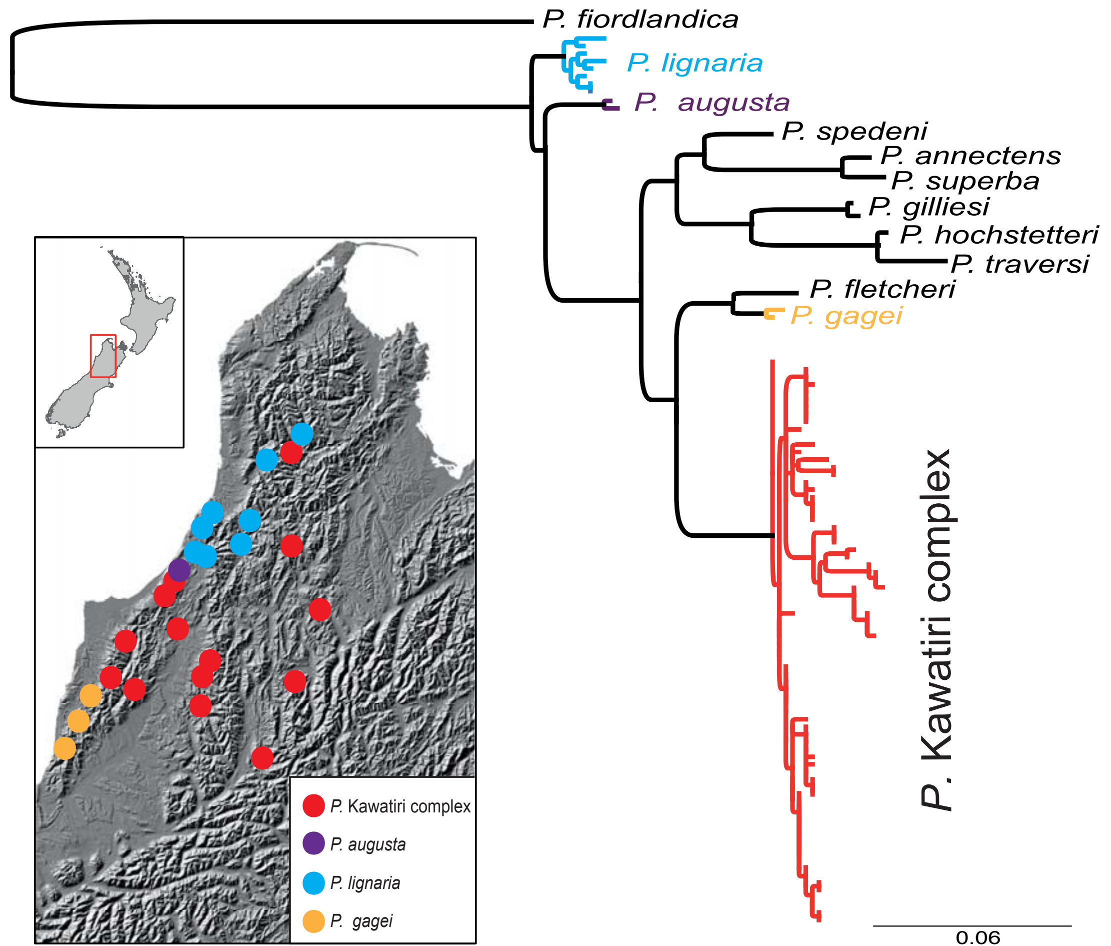


Figure 2.1. Distribution and phylogenetic relationships among representative *Powelliphanta* species in northwest South Island, New Zealand. Members of the Kawatiri complex (red spots) occur in close proximity to three other species. Maximum likelihood phylogeny of mitochondrial COI haplotypes from these species and representative of *Powelliphanta* diversity that are geographically more distant (Trewick et al. 2008).

Laboratory protocols

Genomic DNA was extracted using incubation at 55ºC with Proteinase K and a CTAB buffer (2% Hexadecyltrimethyl ammonium bromide, 100mM Tris-HCl pH8.0, 1.4M NaCl, 20mM EDTA). This was followed by a combined phenol/chloroform/isoamyl alcohol (25:24:1) cleanup based on previously described methods (Stine 1989; Terret, 1992; Thomaz et al. 1996). DNA was precipitated with ethanol and resuspended in a TE buffer (10mM Tris, 0.1mM EDTA) and the quantitiy and qualitiy checked using a Nanodrop® ND-1000 (NanoDrop Technologies) and electrophoresis on 1% agarose gels. DNA was diluted to approximately 10ng/µl to provide templates for amplication of specific fragments.

PCR targeted a fragment of the mitochondiral Cytochrome Oxidase subunit I gene, using a new *Powelliphanta* specific primer (POWCO1r) that replaces H7005 (Hafner et al. 1994) used in previous studies. Together with LCO1490 (Vrijenhoek 1994) POWCO1r targets a fragment of about 700bp at the 5´ end of COI. PCR reactions were performed in 10μL volumes with Red Hot® Taq polymerase using thermocycling conditions 94°C for 2 minutes followed by 36 cycles of 94°C for 15s, 50°C for 30s, and 72°C for 1min 30s with a final extension of 72°C for 8min. PCR products were purified using SAP/EXO (Shrimp Alkaline Phosphatase/exonuclease) enzymatic digest (USB Corporation) and sequenced using BigDye® chemistry (Perkin Elmer) following the manufacturer's protocols on an ABI prism DNA capillary sequencer (Applied Biosystems Inc. California).

Sequence reads were checked against ABI trace files using Sequencher v4.6 (Applied Biosystems, Inc., Foster City, California) and aligned by eye in SeAl v2.0a11 (Se-Al 2007). The nucleotide sequences for COI were translated to check for the presence of stop codons and frame shifts that might be indicative of nuclear copies (Bensasson et al. 2001; Blanchard & Lynch 2000) representing outgroup taxa were obtained from Genbank .

mtDNA analysis

The *Powelliphanta* Kawatiri complex ingroup clade was identified using phylogenetic analysis of a 622 bp alignment of COI sequence. A maximum likelihood tree was generated by PhyML in GENEIOUS v6.0.5 using the general time reversible (GTR) model, with outgroup sampling as previously reported (Trewick et al. 2008). As phylogenetic trees are not always ideal to describe the relationship between haplotypes within closely related taxa, where ancestral sequence variants might persist in extant populations, we used POPART (Leigh & Bryant 2015) to construct a median joining network among the mtDNA COI ingroup haplotypes. Networks inferred using other algorithms available in POPART (Minimum Spanning, Ancestral MP, Integer NJ, Tight Span Walker, TCS) were also examined.

To assess mtDNA sequence variation present within each population sample, nucleotide (π) and haplotype (h) diversities and other statistics were calculated in

ARLEQUIN v3.5.1.2 (Excoffier & Lischer 2010). Population pairwise φ_{ST} were estimated and significant deviations from 0 determined. A standard AMOVA was used to test for significant genetic differences based on the estimate of genetic partitioning among groups (F_{CT}). For this, population samples were alternatively grouped according to tag names that reflect shell traits, geographic clusters or other characteristics.

Evidence of isolation by distance (IBD) was sought using a Mantel test of the correlation of pairwise linear geographic distances and pairwise φ_{ST} from the COI haplotype data using 1000 permutations implemented in IBDWS v3.23 (Jensen et al. 2005). MtDNA haplotypes were imported to IBDWS as a text alignment. Geographic distances were log transformed as recommended for an expanded stepping stone model (Slatkin 1993).

Microsatellite development

A genomic library was developed to isolate potential microsatellite markers for *Powelliphanta* Kawatiri complex using a modified enriched microsatellite library protocol (Vaughan Symonds protocol Pers. Comm. 2008). Whole genomic DNA was digested with the restriction enzyme BfuCI (4 U/µL), and fragments ligated to Sau3AI (BfuCI) linker (25 µL). Successful ligation was ensured by amplification with Sau L-A oligo (20 µL). Replicate PCRs using 12, 24 or 36 cycles were performed in 20µL volumes with Red Hot® Taq polymerase company and cycling conditions; 95°C for 3 minutes followed by 95°C for 45s, 56°C for 45s, and 72°C for 1min 15s with a final extension of 72°C for 5min. 12 cycles of standard PCR protocol were used for hybridisation. A linker created using 10X T4 Ligase buffer, Sau L-B oligo and polynucleotide kinase, was annealed suing 10X buffer, NaCl (600 µL) and Sau L-A oligo (100 µL). Hybridisation used to 5´ biotin-labelled (CA_{12} or GA_{12}) oligonucleotides (100 µM), the linker primer Sau L-B (50µM), HYB-A (12 x SCC/0.2% SDS) solution (heated to 65°C to re-suspend), 2.5µL probe (100µM), 5µL primer and 7.5µl water. Thermocycling conditions denatured PCR roduct at 95°C for 3min and then used a step down protocol: 70.5°C for 1min reducing 0.5°C every cycle to 55.5°C. Cycle length was then changed to 2 minutes and dropped 0.5°C every cycle from 55°C–53°C. The final cycle held at 52.5°C for 30s. Enriched DNA was then

selectively captured using streptividin magnetic beads (Promega). The biotinylated beads were prepared by washing 3 times in HYB-B (6 x SCC/0.1% SDS) and then resuspended in 150µL of HYB-B. 50µL of the hybridisation mix was added to beads and gently agitated for 30min. The following washes were then performed with all buffers pre-heated accordingly, capturing the magnetic beads with a magnetic stand between each step: 300µL HYB-C (3 x SCC/0.1% SDS) at room temperature twice; 300µL HYB-D (1 x Scc/0.1% SDS) at 42°C for 15 minutes, repeated at 60°C for 15 minutes and repeated at 52°C for 15 minutes. A final wash stage at 100µL TE at 95°C for 5 minutes with the supernatant removed but without bead capture was repeated once. DNA was precipitated from each supernatant using 22µL of 3M NaOAc and EtOH dried and resuspended in 25µL TE. Enriched DNA was subjected to PCR amplification to generate double stranded DNA.

The resulting DNA fragments were ligated in a plasmid vector and transformed into competent cells using the TOPO® Cloning Kit (Invitrogen). The *E. coli* cells plated and cultured at 37 °C overnight before a total of 672 recombinant colonies were picked and suspended in 20µL of broth (tryptone, yeast extract and NaCl) and incubated overnight. These were denatured at 95°C for 10 minutes, before PCR amplification of inserts with M13 primers. The resulting 196 PCR positive products were sequenced using Perkin Elmer BigDye® v3.1 chemistry following the manufacturer's protocols, and read on an ABI 3730 capillary sequencer (Applied Biosystems, Inc, Foster City, California). Primer pairs were designed to target 78 putative microsatellite loci using the programme Primer3 (Koressaar & Remm 2007; Untergasser et al. 2012). The forward primer of each pair was synthesized with an M13 tail to allow for incorporation of a fluorescent label.

Putative microsatellite loci were tested by PCR amplification across seven individuals comprising five *Powelliphanta* Kawatiri complex from separate geographical locations and two *P. lignaria johnstoni*. Each forward primer was labelled with either of the fluorescent dyes, 6-Fam or HEX (Invitrogen), during PCR using M13 tails on the forward primers. PCR reactions were performed in 10µL volumes with Red Hot® Taq polymerase using the cycling conditions 94°C for 2 minutes followed by 36 cycles of 94°C for 15s, 50°C for 30s, and 72°C for 1min 30s

with a final extension of 72ºC for 8min. Loci were pooled using combinations of dye colour and allele size range and genotyped on an ABI 3730 with an internal LIZ size standard and scored using GENEMAPPER® v4.7 (Applied Biosystems).

Microsatelite analysis

Five reliable polymorphic loci were used to genotype population samples. Loci were assessed for large allele dropout and stuttering using 1000 randomisations in the programme MICROCHECKER v2.2.3 (Van Oosterhout et al. 2004). Evidence of linkage disequilibruim among loci was sought using the largest sample of 31 snails from the Stockton population and the software GENEPOP ON THE WEB (Raymond & Rousset 1995; Rousset 2008). We compared observed genotype distributions in each population sample with those expected under Hardy–Weinberg equilibrium using χ^2 tests. Population pairwise F_{ST} estimates were examined to see which were significantly greater than zero as expected if gene flow was limited (ARLEQUIN v3.5.1.2).

Evidence for Isolation by Distance (IBD) was sought using a Mantel test of the correlation of pairwise geographic distances and pairwise F_{ST} with 1000 permutations implemented in IBDWS v3.23 (Jensen et al. 2005). Diploid genotypes were imported into IBDWS as raw data. Geographic distances were calculated and log transformed as described above. Loci were examined separately and as a single group.

We searched for evidence of population structuring among our multi-locus genotype data using Bayesian assignment in STRUCTURE v2.3.2 (Pritchard et al. 2000). This approach infers genotype similarity and assigns individuals to clusters (populations), as well as determine isolation-by-barrier. We used Bayesian MCMC of 100,000 steps following a burnin 10,000 steps. The number of clusters (K) was set at 1–15 in separate runs with 11 iterations for each. This analysis was performed using the no-admixture model, appropriate for studying fully discrete populations and suitable for detecting subtle structure. Analysis was repeated with both correlated and independent allele frequency models and then repeated using the admixture model. We also examined the effect of including sampling location information as

recommended for analyses with relatively small sample sizes (Hubisz et al. 2009). Including sampling location categories or phenotype-based groups allows the models implemented by STRUCTURE v2.3.2 to modify the prior distribution for each individual's population assignment. Results were processed in STRUCTURE HARVESTER v0.6.93 (Earl & vonHoldt 2011) to assess and visualise likelihood values across multiple values of K and hundreds of iterations to determine the number of genetic groups that best fit the data (Earl & vonHoldt 2011). In addition, STRUCTURE HARVESTER v0.6.93 was used to reformat data for analysis in CLUMPP v1.1.2 (Jakobsson & Rosenberg 2007). This cluster matching and permutation program deals with label switching and multimodality in analysis of population structure. DISTRUCT v1.1 (Rosenberg 2003) was used to graphically display results.

We used OBSTRUCT (Gayevskiy et al. 2014) to determine which populations lent significance to the structure of the group. OBSTRUCT objectively analyses the nature of structure revealed in Bayesian ancestry profiles using established statistical methods to evaluate the extent of structural similarity between sampled and inferred populations. It tests the significance of population differentiation, provides information on the contribution of sampled and inferred populations to the observed structure and, crucially, determines whether the predetermined factor of interest correlates with inferred population structure. Comma Seperated Value (CSV) files were created from STRUCTURE output files manually for K= 2, 5, 6, 7 and processed with OBSTRUCT. The overall R^2 value for each hypothesized number of populations (K) was compared with each other and to R^2 values calculated without predefined and inferred populations respectively. Results were graphically visualised using R v3.2.1 (R Core Team 2013).

Heirarchical analyses of haplotype and microsatellite data were performed separately across hypothesised groupings of populations in ARLEQUIN 3.5.1.2 (Excoffier & Lischer 2010) Population samples were grouped according to closest proximity to the eight recognised tag names (8 groups) and the groups identified in the mtDNA analysis (7 groups) and microsatellite data (6 groups), and into groups based on geographic proximity (6 groups) and by which axial mountain range they occurred on (3 groups) (see Table 4).

Results

Mitochondrial DNA

Phylogenetic analysis of the mtDNA COI sequences yielded a well resolved monophyletic *Powelliphanta* Kawatiri complex clade, sister to a clade comprising *P. fletcheri* and *P. gagei* (Figure 2.1). Within the *Powelliphanta* Kawatiri complex clade were 101 snails collected from 16 locations (Table 2.1) and elevations ranging from 550m asl to 1500m asl.

Among these data were 36 haplotypes that differed by up to 7.2%, (ML distance) and nucleotide diversity (π) of 0.009. No deletions, inversions or insertions were detected. Nucleotide diversity was low within population samples from the Kawatiri complex clade, with six population samples being monomorphic. The highest sequence divergence within a population sample was between the two snails from the Mounts (0.013; Table 2.1). Haplotype sharing among location samples of Kawatiri complex was limited with most population samples having private haplotypes. Even the most densely sampled locations (Stockton), which had eleven haplotypes, shared only one (Hap12) and this with the nearest site (Denniston). In most cases haplotypes from the same location grouped closely on the network. The exception being two haplotypes from Stockton plateau that differed by 16/622 bp (2.6%) from other Stockton haplotypes (Figure 2.2). Most genetic variation in our sample was therefore partitioned among sample locations with all pairwise φ_{ST} estimates significantly greater than zero. There was strong evidence of genetic partitioning between the populations with an overall φ_{ST} of 0.7369. Although the haplotype network shows some indication of geographic partitioning, Overall mtDNA variation was consistent with a model of isolation by distance (mantel test: r = 0.453; $P < 0.001$).

Table 2.1. mtDNA haplotype and diversity table of New Zealand *Powelliphanta* Kawatiri complex snails listed by sampling location.

		Haplotype																																				Diversity	
Locations	**n**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	***π***	***h***
Garabaldi	3																	1	2																			0.00	0.67
Matiri	3				1	2																																0.00	0.67
Braeburn	2						2																															0.00	0.00
Matakitaki	2						2																															0.00	0.00
Springs Junction	4	4																																				0.00	0.00
Brunner Nth	2			2																																		0.00	0.00
Brunner Mid	2															2																						0.00	0.00
Kirwans	5																4										1											0.00	0.40
Mounts	2																								1	1												0.01	1.00
3 Sisters	4		4																																			0.00	0.00
Bucklands Peak	3																						1	1				1										0.01	1.00
Buller	7							6														1																0.00	0.29
Denniston Sth	7										2		4																					1				0.00	0.67
Denniston Nth	6												2																	1	1	1	1					0.00	0.93
Stockton	47								3	5		5	12	4	13					1	1								1						1	1		0.00	0.84
Baton	2																																				2	0.00	0.00
Total	101	4	4	2	1	2	4	6	3	5	2	5	18	4	13	2	4	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	0.009	

Microsatellites

All 101 snails were genotyped for five microsatellite loci. No evidence of long allele drop out, or null alleles was detected for any of these loci. However, there was evidence of departure from the expectations of the Hardy-Weinberg equilibrium for four of the five loci within the population sample at Stockton; likely caused by the presence of null alleles. Pairwise linkage disequilibrium tested using the genotypes of the 31 Stockton snails indicated that the alleles at different loci were randomly associated. The number of alleles per locus ranged from 4 to 20. Per locus allelic richness from all genotyped population samples (n = 15) reached 8 (Appendix 2.1). Average gene diversity (Table 2.2) was highest in the population sample from Brunner Mid (0.733). Braeburn and Matakitaki samples had no allelic variation, but each consisted of only two snails. Allele sharing among population samples was common but half the population pairwise F_{ST} values were significantly greater than 0 (46/91; Table 2.3). Stockton differed significantly from all other population samples (pairwise F_{ST} 0.247 – 0.504). Denniston South and Dennsiton North differed significantly from one another although the geographic distance between them was only ~5.4 km. In general F_{ST} estimates were large but small sample sizes prevented detection of significant differentiation between half the populations. Each population was significantly differentiated from at least three other populations, with the exception of Brunner North (n = 2) which differed significantly only from the Stockton sample (Table 2.3). Microsatellite differentiation followed a model of isolation-by-distance (mantel test: r = 0.3207; P = 0.0190).

Table 2.2. Genetic diversity at 5 microsatellite loci among population samples of the landsnail *Powelliphanta* Kawatiri complex. Monomorphic samples are indicated with *. Number of individuals (n), average allelic richness (Ar), observed heterozygosity (Ho), expected heterozygosity (He), and average gene diversity (G).

Location	n	Ar	Ho	He	G
Garabaldi	3	2.00	0.11	0.47	0.28
Matiri	3	4.00	0.33	0.87	0.17
Braeburn	2	1.00	*	*	0.00
Matakitaki	2	1.00	*	*	0.00
Springs Junction	4	3.00	0.00	0.71	0.14
Brunner Nth	1	2.00	1.00	1.00	0.40
Brunner Mid	2	2.60	0.60	0.73	0.73
Kirwans (Sth Brunner)	4	3.50	0.31	0.62	0.49
Mounts	2	2.50	0.63	0.67	0.53
Bucklands Peak	3	2.75	0.17	0.68	0.55
Buller	7	2.75	0.61	0.52	0.42
Sth Denniston	7	2.67	0.52	0.54	0.32
Nth Denniston	6	3.40	0.33	0.56	0.56
Stockton	31	6.20	0.25	0.60	0.60

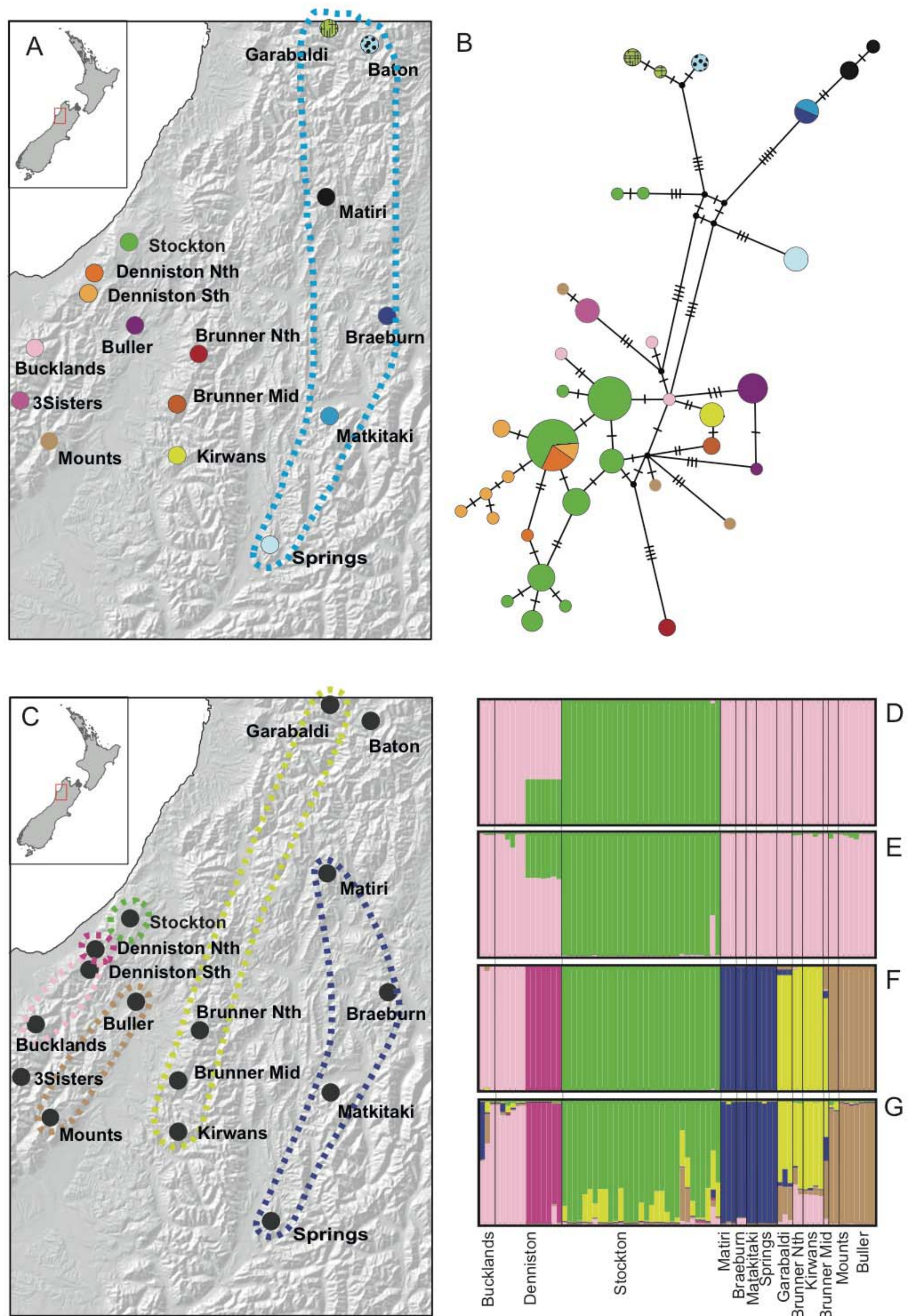


Figure 2.2 Population genetic structure among land snail populations in the New Zealand *Powelliphanta* Kawatiri complex. Median joining network of mtDNA COI (622bp) haplotypes: B. Colours correspond to location markers in map: A. Bayesian genotype assignment probability graphs for five microsatellite loci: (D) no-admixture model, K=2; (E) admixture model, K=2; (F) non admixture model, K=6; (G) admixture model, K=6. Map C shows population samples sharing genotypic clusers (dashed lines). Colours represent alternative genotype assignments, at each K.

Population structure

Bayesian assignment of microsatellite genotypes using STRUCTURE v2.3.2 yielded similar results using correlated and independent allele frequency models. The maximum value of Delta K determined using the independent allele frequency model, which is reported as less likely to overestimate K, was for K = 2 with the next highest value for K = 6. Visual examination of output from DISTRUCT (Figure 2.2) shows that K = 6 captured more information about subdivision of population with similar clustering of population samples returned with and without admixture. There was little difference in the R^2 values for K = 2 (0.84), K = 5(0.84) or K = 6(0.83), but these were all higher than the R^2 value for K = 7 (0.69) returned by OBSTRUCT. We infer from this steep decline in R^2 that K = 6 captures the majority of population structuring. For K = 6 the Stockton population had the greatest average contribution to structure within the data, followed by Buller and Springs Junction respectively. Other population samples had little effect on the structure of the data. Results were similar for K = 5 and K = 2 where it was again the Stockton and Buller populations that contributed most to the structure of the group. Overall clustering is evident and several clusters contained individuals with mixed genotypes.

We further explored support for alternative clustering hypotheses using hierarchical analysis with seven alternative arrangements of population samples that reflect taxonomic and geographical hypotheses (Table 2.3). Unsurprisingly the highest φ_{CT} were obtained for mtDNA data based on the structure within the haplotype network, K = 7 (0.60605; $p < 0.001$). The highest φ_{CT} for microsatellite data was obtained with two groups (K = 2) based on the clusters identified by Bayesian assignment (0.22413; $p < 0.005$). Mitochondrial DNA and genotype data were not concordant. The lowest estimate of φ_{CT} (which did not differ significantly from zero) was obtained for mtDNA and microsatellite data using a grouping hypothesis based on the six populations inferred using STRUCTURE (Table 2.3). The tag-name hypothesis (K = 8) provided reasonable support for mtDNA (0.49809; $p< 0.001$), and grouping by mountain ranges (K = 3) provided the second highest φ_{CT} for microsatellite data (0.18708; $p < 0.001$).

Table 2.3. Hypothetical groupings of population samples used for hierarchical analysis of data from mtDNA haplotypes and five microsatellite loci. FCT scores for each grouping are found below table. The groupings indicated by STRUCTURE analysis (Structure), a geographical east/west split with the populations found in the middle in the second group rather than the first (Other), grouped by mountain ranges (Mount), grouped by geographic proximity (Geo), microsatellite STRUCTURE analysis K=6 (Micro), groupings by mitochondrial haplotype data (Micro) and by closest proximity to the recognised tag names (Tagnames).

Populations	Structure K=2	Other K=2	Mount K=3	Geo K=6	Micro K=6	Haplo K=7	Tagnames K=8
Denniston Nth	1	1	1	1	1	1	1
Dennsiton Sth	1	1	1	1	2	1	1
Stockton	1	1	1	1	3	1	1
Buller	1	1	1	4	5	5	6
3 Sisters	1	1	1	5	1	6	7
Bucklands Peak	1	1	1	5	1	6	7
Mounts	1	1	1	5	5	6	7
Brunner Nth	1	2	2	6	6	7	8
Brunner Mid	1	2	2	6	6	7	8
Kirwans (Sth Brunner)	1	2	2	6	6	7	8
Braeburn	2	2	3	2	4	2	2
Matakitaki	2	2	3	2	4	2	2
Springs Junction	2	2	3	2	4	3	2
Matiri	2	2	3	3	4	2	3
Garabaldi	2	2	3	3	6	4	4
Baton	2	2	3	3	6	4	5
FCT mt DNA	0.44	0.21	0.36	0.40	*	0.61	0.50
FCT micros	0.22	0.14	0.19	0.15	*	0.13	0.11

Discussion

The *Powelliphanta* Kawatiri complex mtDNA clade comprises snails from high elevations (or ecologically associated) habitat in northwest South Island. The lowland species *P. lignaria* is found in the intervening lowland areas toward the northern end of the Kawatiri complex distribution, within just 3km of Garabaldi and Baton populations and 15–20km from the Stockton population. Analysis of mtDNA sequence data revealed that the Kawatiri complex includes snails in the Northern Paparoa Range at high elevations that had previously been assumed to be a population of *Powelliphanta gagei*. *Powelliphanta gagei* is a species that inhabits lowland areas of the Paparoa Range especially areas south of Punakaiki and Reefton (Walker 2003) (Figure 2.1). Thus a clear pattern of high elevation Kawatiri complex snail populations isolated on mountain ranges has emerged, with other *Powelliphanta* species occupying low altitude forest between the Kawatiri complex habitat at its northern and southern reaches. The only exception to this pattern is the Buller River population that is clearly within the Kawatiri complex mtDNA clade, and found to group with the nearby Mounts population in microsatellite analysis. Unlike all other known populations of *Powelliphanta* Kawatiri complex the population at Buller (Kawatiri) River is at low elevation (50m asl).

Genetic diversity is partitioned among the fragmented populations, the lack of concordance among markers is consistent with a single genotypic cluster (Mallet 1995) or single species. The exception to this may be the eastern group which cluster separately from other populations in both the haplotype network and to a certain extent within the structure analysis (Figure 2.2). Genetic diversity within our sampling of the Kawatiri complex (mtDNA and nuclear markers) follows a model of moderate isolation by distance suggesting some population connectivity in the recent past. Microsatellites showed evidence of isolation by distance, however there was also clear evidence of uneven gene flow among populations. Bayesian assignment indicated genotype clusters that each included several distant populations rather than grouping populations that are geographically closest. The sampling regime may have influenced results but there are limitations on the scale of sampling that is possible with these protected and morphologically cryptic snails.

It has been shown that where organisms are spatially clustered a sampling scheme of only 1–4 individuals can be sufficient for detecting isolation by distance or isolation by barrier.

The current data cannot discriminate between gene flow during the last glacial maxima (about 15,000 years ago) and the first wave of human settlement ~800y bp or further colonisation by Europeans ~ 250 y bp. It is thought that settlement by Polynesians ~800 y bp would have had little effect on the snails in this area as this first wave of settlers brought only one predator of snails, kiore (Polynesian rat). Although habitat fragmentation from forest burning is evident in other parts of New Zealand there is little evidence for this on the West Coast of the South Island, and finally although other species of snail are known to have been a food source, there is no evidence in middens of *Powelliphanta* being used this way. The fact that lowland populations occur at the northern and southern reaches of the Kawatiri complex range adds to evidence that genetic structuring is older than anything caused by human settlement of New Zealand.

During the last glacial maxima it has been inferred that the tree line would have been lower and thus habitat suitable for *Powelliphanta* Kawatiri complex could have been at lower elevation, potentially connecting populations that are now separated (McGlone et al. 2010; Trewick & Bland 2012). Human arrival in New Zealand has resulted in forest fragmentation, and the introduction of predatory land mammals is implicated in the extinction of many endemic species (Duncan & Blackburn 2004; Towns & Daugherty 1994).

The existence of other *Powelliphanta* species in intervening lowland elevations suggests that mammal predation does not explain the current fragmented distribution of Kawatiri complex. The relatively long life span (12–14 years and up to ~20 years), and low reproductive rate (K J Walker Pers Comm) of *Powelliphanta*, combined with the genetic structuring, provides support for a scenario of allopatry and population genetic structure in response to climate changes that restricted distribution to high elevation sites since the end of the Last Glacial Maximum 15,000 years ago. It is unlikely that the level of differentiation described here could have

accumulated over the ~250 years, since extensive anthropogenic habitat modification, in a snail such as *Powelliphanta*. In fact the genetic structure and differentiation described among Kawatiri complex populations suggests that there has been restricted gene flow for many generations; sufficient to leave a genetic signature. Therefore we suggest that these populations were last separated during climate warming in the Holocene (<15k ya), although habitat fragmentation and isolation of populations is likely to have occurred prior to this too. Since the Holocene clusters have been separated by potentially unsuitable lowland forest habitat impeding further gene flow.

It is not possible to determine the exact timing of isolation between populations within the Kawatiri complex lineage but we can exclude the hypothesis that population differentiation evident today was due to high elevation specialisation associated with initial phases of Pliocene mountain building. Although emergence of the Southern Alps provided the topography in which climate linked habitat adaptation could occur, this process will have been influence by the repeated changes in conditions through the Pleistocene (Hewitt 2000). Episodes of climate change are likely to have had a major influence on population fragmentation and thus population size, and we note that the implied north-south trend in gene flow is consistent with the intersection of altitudinal and latitudinal range shifting (Bulgarella et al. 2014). This contrasts with the more pronounced spatial pattern seen among other invertebrates in the Southern Alps (e.g. alpine scree weta: Trewick et al. 2000; cidadas: Hill et al. 2009; stoneflies: McCulloch et al. 2010)where high altitude populations appear to have persisted through repeated climate cycles.

Analysis of *Powelliphanta lignaria* and *P. augusta* clades (Buckley et al. 2014) confirmed the specific status of *P. augusta* and found population genetic structure in *P. liganaria* similar to the present study. Genotype clusters were not concordant with mtDNA or subspecies based on morphology alone. In the Hawaiian tree snail *Achatinella mustelina* genetic distances were independent of geographical distance (Holland & Hadfield 2002) but correlated with deep, largely deforested valleys and steep mountain peaks. The factors that isolate snail populations within *Powelliphanta* on the West Coast of the South Island appear to be more subtle and

similar to Bonnin Island *Mandarina* snails where genetically close populations are sometimes separated by genetically distant populations in a dynamic environment (Davison & Chiba 2006). The exact nature of the timing and patterns observed during the Last Glacial Maxima (LGM) remains unknown because the paleoclimate of the South Islands West Coast is not well understood (Alloway et al. 2007; Marske et al. 2009).

We saw a tendency for north-south connectivity (east/west division) sometimes over quite large distances, rather than more regional groups which might have been expected. This is concordant with a landscape where ridges and mountain ranges have a north/south orientation. Thus when climate was cooler we expect there were larger patches of habitat suitable for Kawatiri complex snails, along the coastal ranges and the inland ranges. The land surface features inferred for the LGM (Alloway et al. 2007) show glaciation or permanent snow and fellfeild in the surrounding area between the eastern and western groups, lending support to this idea. It is interesting to note that even large rivers do not appear to provide a barrier to gene flow in this snail lineage.

There are two major groupings seen within the genetic diversity of the Kawatiri complex lineage, one comprises the populations found on the western edge of their range (including Brunner Range), and the other that includes the more inland (eastern) populations and six shallower clusters. We advocate management that retains the evolutionary potential of these and recognises the variation encompassed by the six clusters described in this study. Conservation managers and reserachers should explicitly separate taxonomic diversity from recovery planning, and recognise that different units are appropriate to each purpose (Isaac et al. 1994). Ideally decisions about how conservation effort is distributed should be based on genetic, morphological and ecological evidence, in the context of ecological and genetic exchangeability, although there may also be other non-evolutionary (geological, economic, aesthetic) reasons for ascribing conservation value to particular populations (Moritz 1994). Currently two clusters within the Kawatiri complex (Stockton and Denniston) are threatened by destruction of habitat due to economic reasons. Mining of coal on the Stockton/Denniston plateaux is currently

considered to be of higher importance than retention of biodiversity or ecosystem integrity (Trewick et al. 2008; Walker et al. 2008). Until recently the distribution in high altitude and remote habitats, considered economically unviable for development, has yielded some protection for these snail populations and other taxa.

This is a group with some evidence of genetic partitioning in habitat that is highly vulnerable to destruction. We suggest a careful review of taxonomy to encompass the genetic diversity and structure identified outlined here as well as morphological and ecological information. Regardless of taxonomic classification there are several groups of populations that deserve recognition as discrete entities.

References

Alloway, B. V., Lowe, D. J., Barrell, D. J. A., Newnham, R. M., Almond, P. C., Augustinus, P. C., … Williams, P. W. (2007). Towards a climate event stratigraphy for New Zealand over the past 30 000 years (NZ-INTIMATE project). *Journal of Quaternary Science*, *22*(1), 9–35.

Bensasson, D., Zhang, D.-X., Hartl, D. L., & Hewitt, G. M. (2001). Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology & Evolution*, *16*(6), 314–321.

Blanchard, J. L., & Lynch, M. (2000). Organellar genes: why do they end up in the nucleus? *Trends in Genetics*, *16*(7), 315–320.

Buckley, T. R., White, D. J., Howitt, R., Winstanley, T., Ramón-Laca, A., & Gleeson, D. (2014). Nuclear and Mitochondrial Dna Variation Within Threatened Species and Subspecies of the Giant New Zealand Land-Snail Genus Powelliphanta: Implications for Classification and Conservation. *Journal of Molluscan Studies*, eyu014.

Bulgarella, M., Trewick, S. A., Minards, N. A., Jacobson, M. J., & Morgan-Richards, M. (2014). Shifting ranges of two tree weta species (Hemideina spp.): competitive exclusion and changing climate. *Journal of Biogeography*, *41*(3), 524–535.

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195–205.

Climo, F. M. (1978). The *Powelliphanta gilliesi — traversi — hochstetteri — rossiana — lignaria — superba* ring species (Mollusca: Pulmonata). *New Zealand Journal of Zoology*, *5*(2), 289–294.

Davison, A., & Chiba, S. (2006). The recent history and population structure of five Mandarina snail species from subtropical Ogasawara (Bonin Islands, Japan). *Molecular Ecology*, *15*(10), 2905–2919.

Duncan, R. P., & Blackburn, T. M. (2004). Extinction and endemism in the New Zealand avifauna. *Global Ecology and Biogeography*, *13*(6), 509–517.

Earl, D. A., & vonHoldt, B. M. (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567.

Gaston, K. J. (1992). Taxonomy of taxonomists. *Nature*, *356*(6367), 281–281.

Gayevskiy, V., Klaere, S., Knight, S., & Goddard, M. R. (2014). ObStruct: A Method to Objectively Analyse Factors Driving Population Structure Using Bayesian Ancestry Profiles. *PLOS ONE*, *9*(1).

Gibbs, J. P. (2001). Demography versus habitat fragmentation as determinants of genetic variation in wild populations. *Biological Conservation*, *100*(1), 15–20.

Hafner, M. S., Sudman, P. D., Villabanca, F. X., Spradling, T. A., Demastes, J. W., & Nadler, S. A. (1994). "Disparate Rates of Molecular Evolution in Cospeciating Hosts and Parasites. *Science*, *265*(5175), 1087–1090.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), 907–913.

Hill, K. B. R., Simon, C., Marshall, D. C., & Chambers, G. K. (2009). Surviving glacial ages within the Biotic Gap: phylogeography of the New Zealand cicada *Maoricicada campbelli*. *Journal of Biogeography*, *36*(4), 675–692.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, *9*(5), 1322–1332.

Isaac, M., Herzer, R., Brook, F., & Hayward, B. (1994). *Cretaceous and Cenozoic sedimentary basins of Northland, New Zealand.* (Vol. monograph 8). Institute of Geological & Nuclear Sciences.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806.

Jensen, J. L., Bohonak, A. J., & Kelley, S. T. (2005). Isolation by distance, web service. *BMC Genetics*, *6*, 13.

Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, *23*(10), 1289–1291.

Leigh, J. W., & Bryant, D. (2015). popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116.

Leschen, R. A. B., Buckley, T. R., & Hoare, R. (2009). The Use of Tag-Names and New Zealand Taxonomy. *New Zealand Entomologist*, *32*(1), 85–87.

Lydeard, C., Cowie, R. H., Ponder, W. F., Bogan, A. E., Bouchet, P., Clark, S. A., … Thompson, F. G. (2004). The Global Decline of Nonmarine Mollusks. *BioScience*, *54*(4), 321–330.

Mallet, J. (1995). A species definition for the modern synthesis. *Trends in Ecology & Evolution*, *10*(7), 294–299.

Marske, K. A., Leschen, R. a. B., Barker, G. M., & Buckley, T. R. (2009). Phylogeography and ecological niche modelling implicate coastal refugia and trans-alpine dispersal of a New Zealand fungus beetle. *Molecular Ecology*, *18*(24), 5126–5142.

McCulloch, G. A., Wallis, G. P., & Waters, J. M. (2010). Onset of Glaciation Drove Simultaneous Vicariant Isolation of Alpine Insects in New Zealand. *Evolution*, *64*(7), 2033–2043.

McGlone, M. S., Newnham, R. M., & Moar, N. T. (2010). The vegetation cover of New Zealand during the Last Glacial Maximum: do pollen records under-represent woody vegetation. *Terra Australis*, *32*, 49–68.

Meads, M. J., Walker, K. J., & Elliott, G. P. (1984). Status, conservation, and management of the land snails of the genus *Powelliphanta* (Mollusca: Pulmonata). *New Zealand Journal of Zoology*, *11*(3), 277–306.

Moritz, C. (1994). Defining 'Evolutionary Significant Units' for conservation. *Trends in Ecology & Evolution*, *9*(10), 373–375.

Nelson-Tunley, M., Morgan-Richards, M., & Trewick, S. A. (2016). Genetic diversity and gene flow in a rare New Zealand skink despite fragmented habitat in a volcanic landscape. *Biological Journal of the Linnean Society*, *119*(1), 37–51.

Overmars, F. B., Kilvington, M. J., Gibson, R. S., Newell, C. L., & Rhodes, T. J. (1992). Ngakawau Ecological District survey preport for the Protected Natural Areas Programme. New Zealand Protected Natrual Areas Programme No. 11. Department of Conservation, Hokitika.

Powell, A. (1979). *New Zealand Mollusca: marine, land, and freshwater shells*. Collins.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Raymond, M., & Rousset, F. (1995). GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity*, *86*(3), 248–249.

Ricketts, T. H. (2001). The Matrix Matters: Effective Isolation in Fragmented Landscapes. *The American Naturalist*, *158*(1), 87–99.

Rosenberg, N. A. (2003). distruct: a program for the graphical display of population structure: PROGRAM NOTE. *Molecular Ecology Notes*, *4*(1), 137–138.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106.

Slatkin, M. (1993). Isolation by Distance in Equilibrium and Non-Equilibrium Populations. *Evolution*, *47*(1), 264–279.

Spencer, H. G., Brook, F. J., & Kennedy, M. (2006). Phylogeography of Kauri Snails and their allies from Northland, New Zealand (Mollusca: Gastropoda: Rhytididae: Paryphantinae). *Molecular Phylogenetics and Evolution*, *38*(3), 835–842.

Spencer, H. G., Marshall, B. A., Maxwell, P. A., Grant-Mackie, J. A., Stilwell, J. D., Willan, R. C., … Pojeta, J. (2009). Phylum Mollusca: chitons, clams, tusk shells, snails, squids, and kin. In *ew Zealand inventory of biodiversity. Vol. 1, Kingdom Animalia: Radiata, Lophotrochozoa, Deuterostomia.* (Vol. 1, pp. 161–254). Christchurch: Canterbury University Press.

Stine, O. C. (1989). Cepaea nemoralis from Lexington, Virginia: the isolation and charactereization of their mitochondrial DNA, the implications for their origin and climatic selection. *Malacologia*, *30*(1–2), 305–315.

Terret, J. A. (1992). *The mitochondrial genome of Cepea nemoralis* (PhD Thesis). University of Nottingham, Nottingham, UK.

Thomaz, D., Guiller, A., & Clarke, B. (1996). Extreme Divergence of Mitochondrial DNA within Species of Pulmonate Land Snails. *Proceedings of the Royal Society of London B: Biological Sciences*, *263*(1368), 363–368.

Towns, D. R., & Daugherty, C. H. (1994). Patterns of range contractions and extinctions in the New Zealand herpetofauna following human colonisation. *New Zealand Journal of Zoology*, *21*(4), 325–339.

Trewick, S. A., & Bland, K. J. (2012). Fire and slice: palaeogeography for biogeography at New Zealand's North Island/South Island juncture. *Journal of the Royal Society of New Zealand*, *42*(3), 153–183.

Trewick, S. A., & Morgan-Richards, M. (2009). New Zealand Biology. In *Encyclopedia of Islands (Eds R.G. Gillespie and D.A. Clague).* (pp. 665–673). Berkeley: University of California.

Trewick, S. A., Walker, K. J., & Jordan, C. J. (2008). Taxonomic and conservation status of a newly discovered giant landsnail from Mount Augustus, New Zealand. *Conservation Genetics*, *9*(6), 1563–1575.

Trewick, S. A., Wallis, G. P., & Morgan-Richards, M. (2000). Phylogeographical pattern correlates with Pliocene mountain building in the alpine scree weta (Orthoptera, Anostostomatidae). *Molecular Ecology*, *9*(6), 657–666.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Research*, *40*(15), e115–e115.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535–538.

Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*(5), 294–299.

Walker, K. J. (2003). Recovery plans for *Powelliphanta* land snails. Department of Conservation, Wellington.

Walker, K. J., Trewick, S. A., & Barker, G. M. (2008). *Powelliphanta augusta*, a new species of land snail, with a description of its former habitat, Stockton coal plateau, New Zealand. *Journal of the Royal Society of New Zealand*, *38*(3), 163–186.

Willi, Y., Van Buskirk, J., Schmid, B., & Fischer, M. (2007). Genetic isolation of fragmented populations is exacerbated by drift and selection. *Journal of Evolutionary Biology*, *20*(2), 534–542.

Appendix

Appendix 2.1. Microsatellite loci characteristics for each of the *Powelliphanta patrickensis* sampling locations * indicates a monomorphic locus, n= number of individuals Na = total number of aleles (n) = number of private alleles Ar = alleleic richnesss; Ho = observed heterozygosity; He = expected heterozygosity

Location	n		Loc1	Loc2	Loc3	Loc4	Loc5
Garabaldi	3	Na	6	6	6	6	6
		Ar	2	1	2	2	1
		Ho	0	*	0.333	0	*
		He	0.533	*	0.333	0.533	*
Matiri	3	Na	6	6	6	6	6
		Ar	4	1	1	1	1
		Ho	0.333	*	*	*	*
		He	0.867	*	*	*	*
Braeburn	2	Na	4	4	4	4	4
		Ar	1	1	1	1	1
		Ho	*	*	*	*	*
		He	*	*	*	*	*
Matakitaki	2	Na	4	4	4	4	4
		Ar	1	1	1	1	1
		Ho	*	*	*	*	*
		He	*	*	*	*	*
Springs Junction	4	Na	8	8	8	8	8
		Ar	3	1	1	1	1
		Ho	0	*	*	*	*
		He	0.714	*	*	*	*
Brunner Nth	1	Na	2	2	2	2	2
		Ar	1	1	2	1	2
		Ho	*	*	*	*	*
		He	*	*	*	*	*
Brunner Mid	2	Na	4	4	4	4	4
		Ar	2	4	2	3	2
		Ho	0	1	1	0.500	0.500
		He	0.667	1	0.667	0.833	0.500
Kirwans (Sth Brunner)	4	Na	8	8	8	8	8
		Ar	5	5	2	2	1
		Ho	0.250	0.250	0.250	0.250	*
		He	0.893	0.893	0.429	0.250	*
Mounts	2	Na	4	4	4	4	4
		Ar	2	3	3	2	1
		Ho	0.500	0.500	1	0.500	*
		He	0.500	0.833	0.833	0.500	*
Bucklands Peak	3	Na	6	6	6	6	6
		Ar	1	4	2	2	3

		Ho		0.667	0	0	0
		He		0.867	0.533	0.533	0.800
Buller	7	Na	14	14	14	14	14
		Ar	3	2	3	3	1
		Ho	0.714	0.429	0.857	0.429	*
		He	0.692	0.363	0.659	0.385	*
Denniston Sth	7	Na	14	14	14	14	14
		Ar	2	1	4	1	2
		Ho	0.714	*	0.714		0.143
		He	0.495	*	0.758		0.363
Denniston Nth	6	Na	12	12	12	12	12
		Ar	4	4	3	4	2
		Ho	0.167	0.333	0.333	0.167	0.667
		He	0.561	0.712	0.318	0.742	0.485
Stockton	31	Na	62	62	62	62	62
		Ar	4	8	8	7	4
		Ho	0.032	0.355	0.323	0.161	0.355
		He	0.211	0.788	0.745	0.793	0.471
Total	77	Na	154	154	154	154	154
		Ar	13	14	18	20	4
		Ho					
		He	0.728	0.858	0.888	0.912	0.533

Appendix 2.2. Details of *Powelliphanta patrickensis* snail morphology by "Tagname" and location sample. Elevations and shell morphology from Walker (2003), Mucous and soft tissue coloration noted at time of specimen collection.

Powell's	Tagnames	Elevations	Diameter	Height	d/h	Populations	Mucous Coloration	Soft Tissue Coloration
	P. "Garabaldi"	1200-1400	38	18	2.11	Garabaldi	4 faint blue	4 blue
	P. "Baton"	1100-1200	39	19	2.05	Baton	3 faint blue	3 blue
	P. "Martiri"	1200-1500 (uncommonly to 600)	42.5	21	2.02	Matiri	1 blue 2 clear 2 not noted	1 blue 2 grey 2 not noted
	P. "Matakitaki"	1200-1700 (at one site to 1000)	44.5	21.5	2.07	Braeburn	not noted	not noted
						Matakitaki	not noted	not noted
						Springs Junction	4 faint blue	4 blue-green
P. rossiana gagei	P. "Kirwans"	1250-1280	40	18	2.22	Nth brunner	1blue 1 not noted	1blue 1 not noted
						Mid Brunner	not noted	not noted
						Kirwans (Sth Brunner)	4 clear	4 blue
	P. gagei	800-1450	42.5	24	1.77	Gagei *	*	*
						Mts	1 purple 1 clear 1 not noted	1 purple 2 not noted
						Three sisters (no micros)	2 purple 2 clear	not noted
						Bucklands Peak	2 purple 1 dark blue 3 not noted	not noted
	P. "Buller River"	50-100	40	20	2	Buller	5 blue 2 clear 3 not noted	5 grey 5 not noted
P. rossiana patrickensis	P. patrickensis	550-850	35	21	1.67	Sth Denniston	5 blue 2 faint blue	not noted
						Nth Denniston	6 blue	not noted
						Stockton	25 blue 4 faint blue 5 clear 24 not noted	3 blue 4 black 5 grey 49 not noted

Chapter 3. Mitochondrial genome and 45S RNA cassette sequence of New Zealand land snails, *Placostylus ambagiosus* and *Placostylus hongii* (Mollusca: Gastropoda: Orthalicoidea: Bothriembryontidae: Placostylinae)

Introduction

Mollusca is the second most diverse animal phylum in terms of described species spanning huge phylogenetic and morphological diversity (Lydeard et al. 2004). Form and neurological complexity ranges from cuttlefish and squid to slugs and snails. These Gastropoda (snails, limpets and slugs) include phylogenetically distinct lineages and assemblages on land and in water. Despite their diversity the Gastropoda are an overlooked animal clade with unresolved taxonomy and high species extinction (Lydeard et al. 2004). DNA sequence information is fundamental to our understanding of genome structure, function and evolution. Many genetic markers are conserved across eukaryotes and thus provide homologous traits for phylogenetic inference. The mitochondrial genome and the nuclear ribosomal RNA cassette (45S) are relatively accessible due to the high copy number of both within cells. This combined with improvements in sequencing technologies provide access to a wealth of data to test hypotheses of evolution, divergence and gene flow in non-model organisms.

Despite the huge described and undescribed diversity the scale of genomic information is evident from the relatively small number of molluscan mitochondrial genomes that have been sequenced. On public databases (GenBank) only 706 mollusc complete mitochondrial DNA sequences are reported (May 2016). In comparison approximate numbers of publically available mitochondrial genome sequences for Mammalia total 36,696 and Sauropsida (the class containing the Aves) total 1,913 results (more than twice that of Mollusca). Of the results for Mollusca only 262 are Gastropod snails and under half (114) of those Heterobranchia (Table 3.1). The superfamily Orthalicoidea that includes *Placostylus* snails contains only 2 entries, (Table 3.1). There are currently no published rDNA cassettes of gastropods on GenBank, although partial gene sequences have been generated with targeted PCR for a number of taxa.

Table 3.1 Approximate numbers of whole mtDNA genome sequences available (May 2016) for Mollusca on the public GenBank database. These are ordered by taxonomic rank for the classification leading to the family containing *Placostylus.*

Taxonomic rank	mtDNA genomes
Mollusca	706
Gastropoda	262
Heterobranchia	114
Euthyneura	110
Panpulmonata	87
Eupulmonata	54
Stylomatophora	34
Sigmurethra	34
Orthalicoidea	2

Mitochondrial genomes

Reconstruction of deeper phylogenetic relationships among metazoans using complete mitochondrial genomes has gained increasing popularity (e.g. Dabney et al. 2013; Guschanski et al. 2013; Carr et al. 2015; Mitchell et al. 2014). The approach is much more informative than using sequences of individual genes as the complete genome also provides sets of genome-level characteristics, such as relative arrangements of genes, which can be especially powerful in resolving ancient relationships (e.g. Gibb et al. 2007; Zhou et al. 2014; Boore 2006; Shen et al. 2009). Gene (and genome) duplications play an important role in the diversification of eukaryotic functions and the formation of new species (Ohno 1970). These common phenomena contribute to sequence heterogeneity at homologous loci in all eukaryote genomes.

The mitochondrial genome, present in nearly all eukaryotic cells, contains genetic information that has greatly facilitated systematic and population genetic research since the 1980's (Wilson et al. 1985; Avise et al. 1987). Characteristics include comparatively small size (13–36kb), typically circular, and containing 36–37 genes, including 12–13 protein-coding genes, 2 ribosomal RNA (rRNA) genes and 22 transfer RNA (tRNA) genes (Boore 1999). The mitochondrial genome is maternally inherited (with a few exceptions) in a single unit without recombination, and each

cell contains numerous copies. There are some bivalves that have Doubly Uniparental Inheritance of two distinct mtDNA genomes (Passamonti et al. 2011; Plazzi et al. 2015), but even these genomes seem to be inherited without recombination. Although the mitochondrion has a fairly conserved genomic organization, in animals it has a rapid substitution rate compared with the nuclear genome. These attributes have resulted in its use in studies of systematics, phylogenetic analysis, phylogeography, and population structure of diverse taxonomic groups (e.g. birds: Morgan-Richards et al. 2008; Gibb et al. 2015; Gibb et al. 2007; insects: Pratt et al. 2009; Trewick 2000 molluscs: Trewick, et al. 2008; Hills et al. 2012; mammals: Matisoo-Smith and Daugherty 2012; Dalebout et al. 2005; land snails: Hirano et al. 2014; Morii et al. 2015).

Ribosomal DNA (45S)

The nuclear ribosomal DNA cassette (45S) is inherited from both parents. In Eukaryotes the Large Subunit (LSU) or 45S ribosomal cassette is generally composed of a high number of transcription units (TU) per genome (>500 in plants) arranged in tandem repeats on one or several chromosomes. Each Transcription Unit contains three coding regions (18S, 5.8S and 28S) separated by two non-coding Internal Transcribed Spacers (ITS-1 and ITS-2). The 18S and 28S coding regions are flanked by two external transcribed spacers or ETS (Schaal and Learn 1988; Poczai and Hyvönen 2009). The 5´-external transcribed spacer is subdivided into three regions: the ETS region I which contains the TATA box corresponding to the transcription initiation site (TIS); the ETS region II which includes several and highly variable subrepeats; and the ETS region III which is adjacent to the 18S region (Volkov et al. 1995). Each transcription unit (TU) is separated from the other by an intergenic spacer (IGS) that contains the two ETS bordering the non-transcribed spacer region (NTS) composed of serial repeats of 80bp. (Figure 3.1).

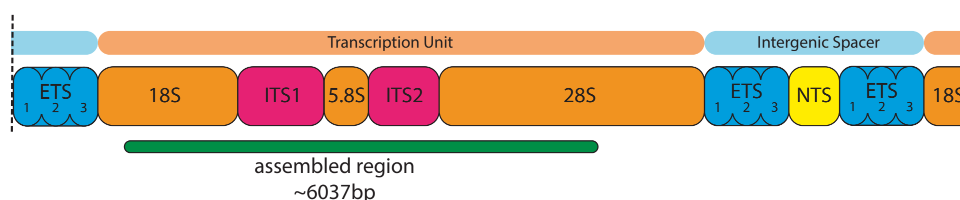


Figure 1: Representation of the 45S rDNA ribosomal cassette, showing gene order in the Transcription Unit. Green region indicates the section assembled, aligned in *Placostylus*.

Within 45S, 18S rDNA is the eukaryotic nuclear homologue of the 16S ribosomal DNA in Prokaryotes and mitochondria. Sequence data from these genes is widely used in molecular analysis to reconstruct the evolutionary history of organisms (esp. vertebrates) as its slow evolutionary rate makes it suitable to reconstruct ancient divergences. Internal Transcribed Spacer 1 refers to the spacer DNA situated between the small subunit ribosomal DNA (rDNA) and large–subunit rDNA genes in the chromosome or corresponding transcribed gene region in the polycistronic rRNA precursor transcript. 5.8S ribosomal DNA (5.8S rDNA) is a non-coding RNA component of the large subunit of the eukaryotic ribosome and so it plays an important role in protein translation. 5.8S is transcribed by DNA polymerase-I as part of the 45S precursors that also contains 18S and 28S rDNA. Its function is though to be in ribosome translocation. Internal Transcribed Spacer 2 is not present in Prokaryotes and likely to have originated as an insertion that interrupted the ancestral 23S rDNA gene. 28S ribosomal DNA is the structural DNA for the 45S or large subunit (LSU) of eukaryotic cytoplasmic ribosomes, and thus one of the basic components of cells. It is the eukaryotic nuclear homologue of the prokaryotic 23S ribosomal DNA (Rogers 2011; Lafontaine and Tollervey 2001).

The ribosomal cassette is highly replicated within the genome, being represented by tandem repeats and copies on different chromosomes (Hillis and Dixon 1991). This high duplication is because 45S ribosomal RNA has to be made in great quantity as they perform critical functions in the ribosome that allow protein synthesis to occur. The translation of a single messenger RNA requires attachment of several ribosomes and a single cell is likely to be producing many copies of hundreds of different kinds of messenger RNA. As a result as much as 85% of the total RNA extracted from ordinary somatic cells is ribosomal RNA (Ohno 1970). The number of repeats of ribosomal DNA and the length of each vary considerably among species (Poczai and Hyvönen 2009). The different repeats are known to undergo a process of homogenization by gene conversion leading to the observation of “concerted evolution” of these ribosomal genes (Nei and Rooney 2005). This homogenization limits the number and divergence of paralogous copies in this gene family, which has promoted its use in molecular phylogenies.

The high copy number of rDNA genes means that regions of the 45S cassette are relatively easy to amplify even from small quantities of genomic DNA. ITS typically varies in sequence even between closely related species. Sequence comparison of Internal Transcribed Spacer (ITS) (a part of the ribosomal cassette), using traditional Polymerase Chain Reaction (PCR) amplification and Sanger sequencing techniques with primers located in the flanking genes has been widely used in molecular systematics of plants and animals (e.g. Morgan-Richards and Trewick 2005; Hörandl et al. 2005; Schoch et al. 2012). The 18S, 5.8S, and 28S regions of the 45S cassette are widely used for working out evolutionary relationships among organisms as they are of ancient origin and are found in all known forms of life (Wu et al 2015; Hedges et al. 1990; Xiaet al. 2003; Torczynski et al. 1983). The differential rates of variation along the transcriptional unit (the coding regions are more conserved than the spacers) allow phylogenetic inferences at different levels of the taxonomic hierarchy (Álvarez and Wendel 2003). Next Generation Sequencing tools allow the entire cassette to be assembled, studied and the variation within an individual (due to its repeated nature in the genome) examined and compared with other individuals.

Generating large-scale sequence data for non-model taxa is increasingly affordable. Here I report assembly and annotation for mitochondrial and nuclear ribosomal DNA sequences from two species of *Placostylus* using next generation sequencing strategy. It is hoped this will provide a resource for future studies of gastropod phylogenetics and conservation genetics.

Methods

Tissue collection

Foot muscle tissue samples for two species *Placostylus ambagiousus* (PS185/GS4) and *Placostylus hongii* (PS257/WG865) originated from a frozen tissue collection originally developed from whole body samples for allozyme analysis (Triggs and Sherley 1993).

DNA extraction

50 mg sections of foot muscle tissue were cut from preserved specimens using a sterile scalpel blade. These were pressed and dried to remove ethanol and were then diced into a dozen pieces. Tissue was transferred to a clean 2 ml Eppendorf microtube. Genomic DNA was extracted using incubation at 55°C with Proteinase K and a CTAB buffer (2% Hexadecyltrimethyl ammonium bromide, 100mM Tris-HCl pH8.0, 1.4M NaCl, 20mM EDTA). DNA was isolated using high-salt precipitation, following a CTAB tissue digestion protocol using chloroform (24:1 chloroform-isoamyl alcohol), sodium acetate (3 M NaOAc), and -20 °C chilled 70% ethanol, which is a modification of previous molluscan DNA extraction methods (Thomaz et al. 1996; S. Trewick et al*.* 2009). This extraction method has been found to be the most successful for attaining high molecular weight DNA while avoiding the potential problem of mucopolysaccharide contamination interfering with enzymatic reactions using Neogastropod tissue (Winnepenninckx et al. 1993). Samples were re-suspended in 50 µl of TE buffer (10 mM Tris, 0.1 mM EDTA), or 100 µl for larger yields of DNA. DNA was quantified using the Qubit Fluorometric Quantitation kit (Life Technologies, Thermo Fisher Scientific Inc.). Total DNA extracts from two individuals of two putative species were processed through massive parallel, high-throughput sequencing using the ThruPLEX® DNA-seq Kit (Rubicon Genomics). Fragmented genomic DNA was paired-end sequenced on an Illumina HiSeq 2500. Reads for each of the 2 individuals were de-multiplexed using standard indexes incorporated in the library preparation kit. Resulting Illumina short-sequence reads that passed standard quality filters had adapter sequences removed using cutadapt 1.11 (Martin 2011) to remove sample specific DNA barcodes. Illumina paired end reads were assembled in Geneious v8 (Kearse et al. 2012).

Mitochondrial Genome assembly and annotation

Complete mitochondrial genomes were characterized separately for individual snails. A partial sequence of *Placostylus* mtDNA Cytochrome Oxidase subunit I was used as a reference for mapping paired end reads from the whole genome using Medium-Low Sensitivity with 25 iterations to the reference sequences using the Geneious Map to reference tool (Kearse et al. 2012). This process iteratively extended the aligned sequence by sequentially mapping additional reads to the existing read alignment. The resulting alignment was used to generate a putative partial mtDNA consensus sequence for the target species. This was subsequently used as the reference for mapping raw reads, and the process repeatedly iteratively until gaps were filled by extension with new sequence data. Final consensus sequences were then checked for mapping idiosyncrasies, which might indicate an error in the generated consensus sequence. Final sequences were then checked for overlap, trimmed and circularised.

Whole mitochondrial genome DNA sequences of *Placostylus ambgiosus* and *Placostylus hongii* were aligned by eye in Geneious. A well annotated genome from a species within the same gastropod order was download from GenBank (*Naesiotus nux*; accession number KT821554; Hunter et al. 2016). Genes were extracted from *Naesiotus nux* and mapped back to the two *Placostylus* genomes. Protein coding genes were translated into amino acid sequences using the invertebrate mitochondrial genetic code, and aligned using the translation frames and start/stop codons in Geneious. tRNAs were aligned by identifying the anticodon portion in the stem-loop anticodon arm.

Ribosomal cassette

Short DNA reads were iteratively mapped to reference ribosomal sequences in Geneious v.8 using the method described above for mtDNA genomes. Final consensus sequences were then checked for mapping idiosyncrasies, which might indicate an error in the generated consensus sequence. Genes were annotated by aligning consensus sequences to previously annotated RNA cassettes compiled within our lab group (or from genbank). The two annotated consensus sequences (*P. ambagiosus* and *P. hongii*) were then aligned and trimmed to the same length. To

confirm unbiased representation of rDNA cassettes in each target genome raw reads were mapped to each consensus sequences using custom settings (10 iterations, 5% maximum mismatch). Single Nucleotide Polymorphisms (SNPs) were identified within each individual using the find SNPs function in Geneious v8 set with a 0.1 minimum variant frequency. SNP annotations were then copied to the consensus and the consensus sequences exported and aligned by eye for the two individuals. This alignment was used to identify SNPs between individuals using the method and settings described above.

Results and Discussion

In total after initial quality control *Placostylus ambagiosus* returned 21,795,718 paired 125bp length reads and *Placostylus hongii* returned 24,869,360 paired 125bp length reads for assembly.

Mitochondrial Genome

The mitochondrial genomes of *Placostylus ambagiosus* and *P. hongii* are 14,654bp and 14,711bp in length, respectively (Table 3.2). A final mapping to the completed mitochondrial genome mapped 16,999 reads to *P. ambagiosus* with a minimum read depth of 16 (mean read depth: 134.2, max read depth: 203). The final mapping to the completed mitochondrial genome of *Placostylus hongii* mapped 14,136 reads with a minimum read depth of 9 (mean read depth: 98.2, max read depth 176). Both mitochondrial genomes contain the expected genes of metazoan mitochondria: 13 protein coding, 2 rRNA and 22 tRNA genes. One strand encodes 25 genes, the other encodes 13 genes (Table 3.3). As in other invertebrate mitochondrial genomes (Shioiri and Takahata 2001) nucleotide composition is skewed with a high proportion of Adenine and Thymine (76.3% *P. ambagiosus* and 76.2% *P. hongii*) (Table 3.2) that is within the range recorded in Sigmurethra snails for which data are available (n = 16, range = 60–77%, mean = 69%). Non-coding regions form a small proportion of the mitochondrial genome accounting for less than 3% (2.7% *P. ambagiosus*, 2.8% *P. hongii*).

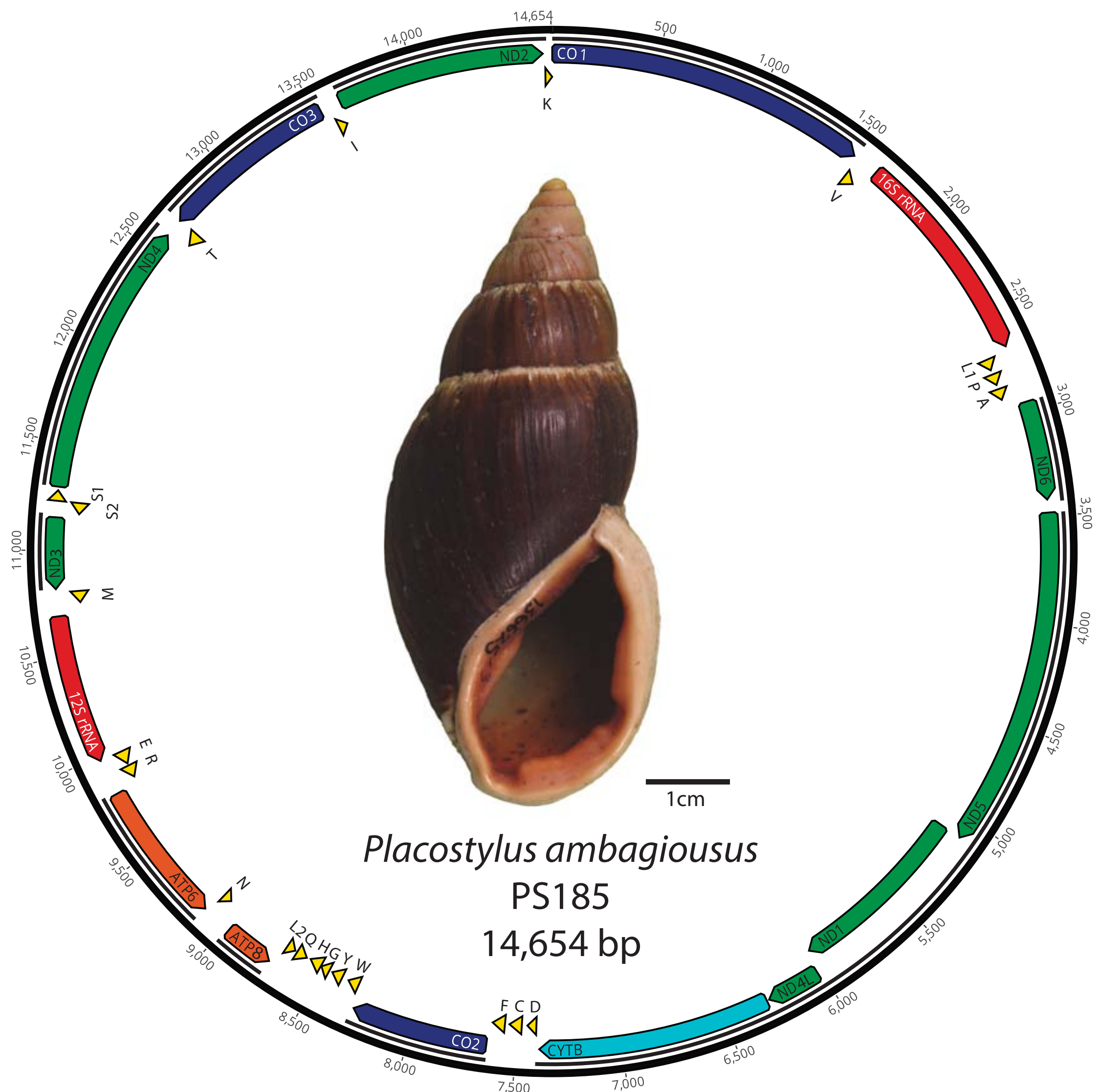


Figure 3.2. The mitochondrial genome of the New Zealand flax snail *Placostylus ambagiosus*: Bothriembryontidae. The tRNA genes are labelled based on the IUPAC IUB single letter amino acid codes. Direction of gene translation is indicated by arrow tips on annotation bars.

Gene content and arrangement

Thirteen protein-coding genes are commonly found in metazoan mitochondrial genomes. The order and transcriptional orientation of the two New Zealand *Placostylus* are identical. Both *Placostylus* mitochondrial genomes have 24 genes in the same strand, including 9 protein coding genes, 1 rRNA gene and 14 tRNA genes (Figure 3.2; Table 3.3). The content and order of these component genes varies only slightly compared to their nearest available relative in the family Orthalicoidae (*Naesiotus nux*) in that *Placostylus* have a reciprocal translocation of tRNA W and

tRNA Y and they do not share the gene duplication of ND4L. The gene duplication within *Naesiotus nux* is the only annotated duplication so far observed within the Stylommatophora Sigmurethra mitochondrial diversity (Figure 3.3). The duplicated ND4L gene is inserted within the tRNAs found between 16S and the ND/CytB group. Comparing assembled mitochondrial genomes from within Sigmurethra (Figure 3.3) one can see that the gene order within *Placostylus* is unique in having tRNA W immediately adjacent to COII. There are few non-coding regions in the *Placostylus* mitochondrial genomes and the largest non-coding region falls between the PCG–ATP8 and the tRNA–Asparagine totalling only 83 bp. When aligned *Placostylus ambagiosus* contains 175 indels and *Placostylus hongii* contains 118 indels.

Table 3.2. Comparison of genetic information of the Mitochondrial genome of two New Zealand *Placostylus spp* used in this study.

	P. ambagiousus	*P. hongii*
Size (bp)	14654	14711
A + T proportion	0.76	0.76
Non Coding Region (bp) and (%)	401 (2.7%)	412 (2.8%)
Genes	37	37
Gene Region (bp)*	14189	14224
tRNA	22	22
Identical sites	13279	
Pairwise % Identity	89.50%	

* gene region calculated as the sum of all gene lengths

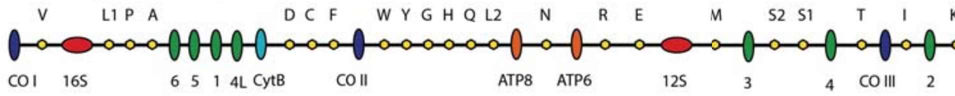
It is not surprising that these two newly sequenced and assembled mitochondrial genomes should not be identical in order to other published snail mitochondrial genomes. Mitochondrial gene order varies considerably among the classes of Mollusca (Kurabayashi and Ueshima 2000; Boore et al. 2004; Grande et al. 2008). Within the Gastropoda, the mitochondrial gene order recorded in Panpulmonata, a clade established by (Jörger et al. 2010), is regarded as standard (White et al. 2011; Knudsen et al. 2006). However, within the Stylommatophora/Sigmurethra only minor gene rearrangements are observed, involving mostly tRNA genes but occasionally single protein coding genes (Figure 3.3). For example in the Helicidae family (*Cepea, Cylindrus, Helix*) ND6 is separated from ND5 by a tRNA and in a

member of the Bradybaenidae family (*Aegista*) ND3 has moved to the other side of ATP8, ATP6, 12S and associated tRNAs.

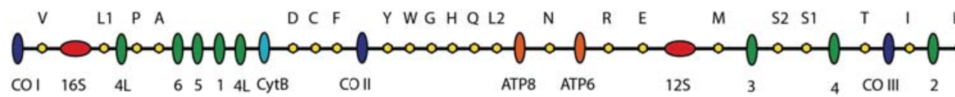
Protein coding genes

Three common start (initiation) codons occur in invertebrate mitochondrial protein coding genes including ATG (methionine), ATA (methionine), and ATT (isoleucina) (Shen et al. 2012; Shen et al. 2009). The two *Placostylus* sequenced use ATG and ATA and the previously unreported in molluscs TTG (methionine). Identical initiation codons are shared in 12 of the 13 protein coding genes of *Placostylus ambagiosus* and *Placostylus hongii*. The exception being ND3 where *Placostylus ambagiosus* has the initiation codon TTG (methionine) and *Placostylus hongii* has ATG (methionine). There are three types of stop codons in the *Placostylus* mitochondrial genome. Different termination codons are found in 3 of protein coding genes (ND6, ND3, ND2). *Placostylus ambagiosus* has TAA for 9 protein coding genes (COI, ND1, CYTB, COII, ATP8, ATP6, ND3, ND4, COIII), TAG for three protein coding genes (ND6, ND5, ND2) and the single codon T for ND4L. *Placostylus hongii* has TAA for 10 protein-coding genes (COI, ND1, fND4, CYTB, COII, ATP8, ATP6, ND4, COIII, ND2), TAG for ND5 and the single codon T for ND4L and ND3. The protein coding gene with the least variability between the two species was COI showing similarity at 93.5% of nucleotide sites and 99.2% amino acid sites. COI is the region most commonly amplified for population genetic studies of invertebrates (Folmer et al. 1994; Dawnay et al. 2007) and interspecific comparisons of *Placostylus spp* in New Zealand (Buckley et al. 2011; Ponder et al. 2003; Trewick et al. 2009). It is evident that other gene regions would carry more sequence variation and therefore provide more information. The most variable protein coding gene was ATP8 with 82.4% similarity between nucleotide sequences and 73.7% differences between amino acid sites (Table 3.3). The highly conserved gene order within this group (Figure 3.3) suggests that a primers targeting the ATP 8 to ATP 6 region (a region including the tRNA asparagine, of ~957bp) may provide significantly more information than the markers traditionally used within CO1, 12S or 16S. Dowle and co-workers used the ND2 gene (830 bp) of the mitochondria of New Caledonia *Placostylus* to resolve haplotype variation among populations of two species (Dowle et al. 2015).

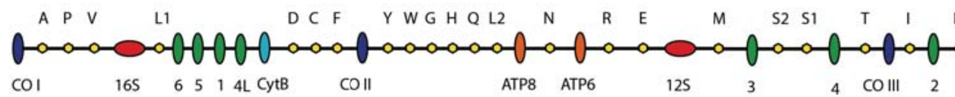
Placostylus ambagiosus and hongii *Orthalicoidea; Bothriembryontidae*



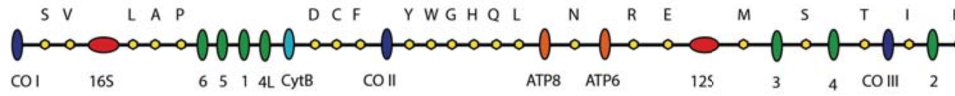
Naesiotus nux *Orthalicoidea; Orthalicoidae (NC_028553)*



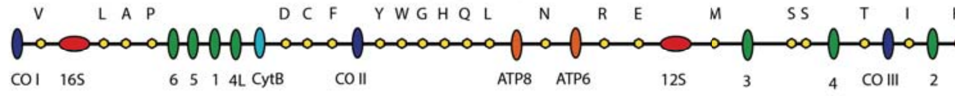
Cerion incanum *Urocoptoidea; Cerionidae (NC_025645)*



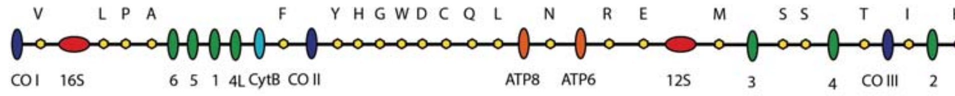
Achatina fulica *Achatinoidea; Achatinoidae (KM114610)*



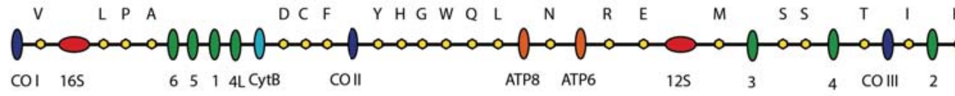
Achatina fulica *Achatinoidea; Achatinoidae (NC_024601)*



Pupilla muscorum *Pupilloidea; Pupullidae(NC_0206044)*



Gastrocopta cristata *Pupiloidea; Pupullidae (NC_026243)*



Vertigo pusilla *Pupilloidea; Vertiginidae (NC_026045)*

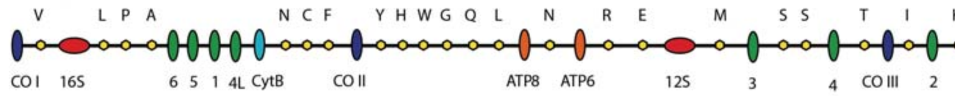


Figure 3.3. Mitochondrial gene order of snails in the molluscan Sigmurethra suborder drawn in linear format starting at Cytochrome Oxidase I (COI). Gene order rearrangements within the group mostly involve position of tRNAs.

Table 3.3: Comparison of gene organisation in mitochondrial genomes of the New Zealand land snails *Placostylus ambagiosus* (PS185) and *Placostylus hongi* (PS257). Amino acid = AA.

Name	Type	Direction	Anti-codon	Identical % bp	Identical % AA	Length bp		Length AA		Start codon		Stop Codon		Start		End	
						PS185	PS257	PS185	PS257	PS185	PS257	PS185	PS257	PS185	PS257	PS185	PS257
CO1	CDS	F		93.5	99.2	1,533	1,533			TGG	TGG	TAA	TAA	1	1	1,533	1,533
Valine (V)	tRNA	F	TAC	86.4		62	66							1,547	1,547	1,608	1,618
16S	rRNA	F		90.5		1,036	1,036							1,649	1,652	2,684	2,687
Leucine (L1 UUR)	tRNA	F	TAG	92.5		67	65							2,696	2,699	2,762	2,763
Proline (P)	tRNA	F	TGG	98.5		66	66							2,774	2,775	2,839	2,840
Alanine (A)	tRNA	F	TGC	88.7		68	68							2,851	2,860	2,918	2,927
ND6	CDS	F		89.1	88.8	480	486	159	161	ATG	ATG	TAG	TAA	2,953	2,965	3,432	3,450
ND5	CDS	F		91.1	90.2	1,623	1,623	540	540	ATA	ATA	TAG	TAG	3,482	3,497	5,104	5,119
ND1	CDS	F		90.8	96.1	918	918	305	305	ATG	ATG	TAA	TAA	5,091	5,106	6,008	6,023
ND4L	CDS	F		93.0	89.0	301	301	100	100	ATG	ATG	T	T	5,989	6,004	6,286	6,301
CYTB	CDS	F		91.6	95.1	1,113	1,113	370	370	ATG	ATG	TAA	TAA	6,290	6,305	7,402	7,417
Aspartic (D)	tRNA	F	GTC	84.5		54	58							7,409	7,422	7,462	7,479
Cysteine (C)	tRNA	F	GCA	84.0		73	76							7,479	7,495	7,551	7,570
Phenylalanine (F)	tRNA	F	GAA	89.3		74	74							7,560	7,585	7,633	7,658
CO2	CDS	F		90.8	94.5	660	666	219	221	ATG	ATG	TAA	TAA	7,637	7,662	8,296	8,327
Tryptophane (W)	tRNA	F	TCA	95.7		69	69							8,307	8,346	8,375	8,414
Tyrosine (Y)	tRNA	F	GTA	86.1		67	71							8,396	8,459	8,462	8,529
Glycine (G)	tRNA	F	TCC	96.4		55	53							8,469	8,550	8,523	8,602
Histidine (H)	tRNA	F	GTG	84.4		70	65							8,525	8,604	8,594	8,668
Glutamine(Q)	tRNA	R	TTG	85.7		66	65							8,596	8,670	8,661	8,734
Leucine (L2 CUN)	tRNA	R	TAA	94.2		51	51							8,666	8,739	8,716	8,789

ATP8	CDS	R		82.4	73.7	246	252	81	83	TTG	TTG	TAA	TAA	8,742	8,802	8,987	9,053
Asparagine (N)	tRNA	R	GTT	96.6		57	58							9,063	9,130	9,119	9,187
ATP6	CDS	R		90.8	88.9	654	654	229	217	ATG	ATG	TAA	TAA	9,128	9,198	9,781	9,851
Arginine ¨	tRNA	R	TCG	81.4		67	69							9,845	9,914	9,911	9,982
Glutamic (E)	tRNA	R	TTC	89.0		80	78							9,912	9,983	9,991	10,060
12S	rRNA	R		90.8		720	734							9,982	10,051	10,743	10,817
Methionine (M)	tRNA	R	CAT	97.0		66	67							10,745	10,833	10,810	10,899
ND3	CDS	R		90.8	89.6	348	350	115	116	TTG	ATG	TAA	T	10,818	10,896	11,165	11,245
Serine (S2 UCN)	tRNA	R	TGA	88.1		64	67							11,178	11,259	11,241	11,325
Serine (S1 AGN)	tRNA	F	GCT	86.2		63	63							11,232	11,316	11,294	11,378
ND4	CDS	F		90.8	91.3	1,311	1,311	436	436	ATG	ATG	TAA	TAA	11,298	11,382	12,608	12,692
Threonine (T)	tRNA	R	TGT	82.4		78	72							12,627	12,711	12,704	12,782
CO3	CDS	R		91.9	95.0	846	849	258	264	ATG	ATG	TAA	TAA	12,687	12,747	13,532	13,595
Isoleucina (I)	tRNA	F	GAT	94.0		58	57							13,571	13,634	13,628	13,690
ND2	CDS	F		89.6	88.7	989	984	329	327	ATA	ATA	TAG	TAA	13,646	13,708	14,623	14,673
Lysine (K)	tRNA	F	TTT	91.7		36	36							14,626	14,683	7	7

Transfer and ribosomal RNA genes

Twenty-two single copy tRNA genes are found in *Placostylus ambagiosus* and *Placostylus hongii*. The exact length and structure of the tRNAs was not determined as the secondary folding structure has not yet been examined in detail, however identity was confirmed by comparison to suitable annotated examples.

The 12S and 16S ribosomal genes in *Placostylus ambagiosus* and *Placostylus hongii* are located almost directly opposite each other in the circular genome. 12S is 720 bp (plus 16 indels) and 734 bp (plus 2 indels) respectively and 16S is the same length for both species (1036bp) and contains the same number of indels for both specimens (13 indels). 12S shows 90.8% similarity between the two species and 16S shows 90.5% similarity (Table 3.3).

Non-coding regions and gene overlap

There are 23 and 24 non-coding regions of 5 bp or more in the *Placostylus ambagiosus* and *Placostylus hongii* mitochondrial genomes. There are five regions of gene overlap of 5 bp or more. The site of DNA replication (control region) has not yet been identified in most gastropod mitochondrial genomes. A 56–58 bp region with the same secondary structure (a strong stem–loop element and a second small stem–loop element) has been identified in seven published neogastropods (McComish et al. 2010). The longest intergenic spacer in the marine snail *Amalda northlandica* is located between the genes trnF and COIII, and contains 11 bp of sequence that is identical (not complementary) on the light and heavy strand. This extremely unusual structural element may represent a highly reduced control region (McComish et al. 2010). In general, gene order within the gastropod group is very stable but lineage-specific arrangements of structure-bearing intergenic spacers are likely to be phylogenetically informative. DNA sequence variation within the commonly sequenced gene COI is relatively low compared to other mitochondrial genes.

Ribosomal Cassette –partial 45S

The consensus partial 45S sequence for *Placostylus ambagiosus* was 6,031bp and when remapped a total of 15,062 paired reads produced a minimum coverage depth of 79 reads (max: 3982 reads, mean: 258.7 reads). The same section of the 45S cassette for *Placsotylus hongii* was 6,029bp long and remapping included 12,056 paired reads with a minimum depth of 50 (max: 3199 reads, mean: 237.0 reads). One exception with a minimum depth of a single read related only to the last nucleotide at the downstream end of the partial 45S cassette. This read overlapped by 100bp with 50 other paired reads and the nucleotide position in question was the same nucleic acid in both species.

Table 3.4. Comparison of the genetic information of the rDNA cassette of two New Zealand *Placostylus* species.

	Length		**Identical sites**
	***P. ambagiosus* PS185**	***P. hongii PS257***	**%**
18S	2311	2312	99.9
ITS1	590	589	98.0
5.8S	159	159	99.4
ITS2	365	366	92.6
28S	2598	2597	99.9
Total	**6023**	**6023**	**99.2**

The aligned ribosomal cassette of 6,037bp, *Placostylus ambagiosus* and *P. hongii* were 6,023bp (+14 INDELS) and 6,022bp (+15 INDELS) in length, respectively (Table 3.4). Both contain the expected components of ribosomal cassettes: 18S, ITS1, 5.8S, ITS2 and 28S. The AT content was 44.2% for both *P. ambagiosus* and *P. hongii* (Table 3.5). ITS2 shows the most variation between the two individuals with only 92.6% identical sites (Table 3.4). In comparison, the highly constrained 28S and 18S are 99.9% identical in sequence. ITS1 is also more variable than other regions with 98% identical sites. Due to the non-coding nature of these regions this higher level of variance is not surprising (Table 3.4).

Table 3.5. Single Nucleotide Polymorphism (SNP) data within (and between) each individual, shown for each region of the available 45S and the comparison for the two New Zealand *Placostylus sp.* Y(C/T), W(A/T), S(C/G), R(A/G), M(A/G), K(A/T).

P. ambagiosus (PS185)								
		SNPs						
Region	Total	Y	W	S	R	M	K	INDELS
18S	7	2	2	-	1	1	1	-
ITS1	0	-	-	-	-	-	-	-
5.8S	1	-	-	-	1	-	-	-
ITS2	0	-	-	-	-	-	-	-
28S	1	-	-	-	-	-	-	1
Total	9	2	2	0	2	1	1	1
P. hongii (PS257)								
		SNPs						
Region	Total	Y	W	S	R	M	K	INDELS
18S	5	2	1	-	1	-	1	-
ITS1	0	-	-	-	-	-	-	-
5.8S	0	-	-	-	-	-	-	-
ITS2	1	-	-	-	1	-	-	-
28S	1	-	-	-	-	-	1	-
Total	7	2	1	0	2	0	2	0
Interspecific								
		SNPs						
Region	Total	Y	W	S	R	M	K	INDELS
18S	3	1	1	-	-	1	-	-
ITS1	13	1	1	-	2	1	-	7
5.8S	1	-	-	-	1	-	-	-
ITS2	28	3	-	1	3	-	-	21
28S	3	-	1	-	-	-	1	1
Total	48	5	3	1	6	2	1	29

Single Nucleotide Polymorphism (SNP)

There were a total of eight SNPs within *Placostyus ambagiosus* and seven within *P. hongii.* Interspecific SNPs in the aligned dataset totalled 48. The majority of these (29) were indels. ITS regions contained most of the interspecific variation with ITS1 containing 13 SNPs, seven of which were indels. ITS2 contained 28 interspecific SNPs 21 of which were indels (Table 3.5). In all instances any given intraspecific SNP from one specimen was represented in the other by either an identical SNP or by one of the two nucleotides present in the SNP.

Conclusion

Traditional sequence based markers using COI in the mitochondrial genome or the ITS region in the 45S cassette seem unlikely to provide information suitable for population genetic studies. Other genes within the mitochondrial genome that may prove more suitable are ATP8 –ATP6. It seems unlikely that the 45S cassette will provide enough information for within species population genetic studies. To examine population structure and dynamics such as gene flow microsatellites or new tools such as double digest RAD sequencing may be more useful.

References

Álvarez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, *29*(3), 417–434.

Avise, J., Arnold, J., Ball, R., Bermingham, E., Lamb, T., Neigel, J., … Saunders, N. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, *18*(1), 489–522.

Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research*, *27*(8), 1767–1780.

Boore, J. L. (2006). The use of genome-level characters for phylogenetic reconstruction. *Trends in Ecology & Evolution*, *21*(8), 439–446.

Boore, J. L., Medina, M., & Rosenberg, L. A. (2004). Complete Sequences of the Highly Rearranged Molluscan Mitochondrial Genomes of the Scaphopod *Graptacme eborea* and the Bivalve *Mytilus edulis*. *Molecular Biology and Evolution*, *21*(8), 1492–1503.

Buckley, T., Stringer, I., Gleeson, D., Howitt, R., Attanayake, D., Parrish, R., … Rohan, M. (2011). A revision of the New Zealand *Placostylus* land snails using mitochondrial DNA and shell morphometric analyses, with implications for conservation. *New Zealand Journal of Zoology*, *38*(1), 55–81.

Carr, L. M., (Trish) McLenachan, P. A., Waddell, P. J., Gemmell, N. J., & Penny, D. (2015). Analyses of the mitochondrial genome of *Leiopelma hochstetteri* argues against the full drowning of New Zealand. *Journal of Biogeography*, *42*(6), 1066–1076.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., … Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences*, *110*(39), 15758–15763.

Dalebout, M. L., Robertson, K. M., Frantzis, A., Engelhaupt, D., Mignucci-Giannoni, A. A., Rosario-Delestre, R. J., & Baker, C. S. (2005). Worldwide structure of mtDNA diversity among Cuvier's beaked whales (*Ziphius cavirostris*): implications for threatened populations. *Molecular Ecology*, *14*(11), 3353–3371.

Dawnay, N., Ogden, R., McEwing, R., Carvalho, G. R., & Thorpe, R. S. (2007). Validation of the barcoding gene COI for use in forensic genetic species identification. *Forensic Science International*, *173*(1), 1–6.

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*(5), 294–299.

Gibb, G. C., England, R., Hartig, G., McLenachan, P. A. (Trish), Smith, B. L. T., McComish, B. J., … Penny, D. (2015). New Zealand Passerines Help Clarify the Diversification of Major Songbird Lineages During the Oligocene. *Genome Biology and Evolution*, evv196.

Gibb, G. C., Kardailsky, O., Kimball, R. T., Braun, E. L., & Penny, D. (2007). Mitochondrial Genomes and Avian Phylogeny: Complex Characters and Resolvability without Explosive Radiations. *Molecular Biology and Evolution*, *24*(1), 269–280.

Grande, C., Templado, J., & Zardoya, R. (2008). Evolution of gastropod mitochondrial genome arrangements. *BMC Evolutionary Biology*, *8*, 61.

Guschanski, K., Krause, J., Sawyer, S., Valente, L. M., Bailey, S., Finstermeier, K., … Savolainen, V. (2013). Next-Generation Museomics Disentangles One of the Largest Primate Radiations. *Systematic Biology*, *62*(4), 539–554.

Hedges, S. B., Moberg, K. D., & Maxson, L. R. (1990). Tetrapod phylogeny inferred from 18S and 28S ribosomal RNA sequences and a review of the evidence for amniote relationships. *Molecular Biology and Evolution*, *7*(6), 607–633.

Hillis, D. M., & Dixon, M. T. (1991). Ribosomal DNA: Molecular Evolution and Phylogenetic Inference. *The Quarterly Review of Biology*, *66*(4), 411–453.

Hills, S. F. K., Crampton, J. S., Trewick, S. A., & Morgan-Richards, M. (2012). DNA and morphology unite two species and 10 million year old fossils. *PloS One*, *7*(12), e52083.

Hörandl, E., Paun, O., Johansson, J. T., Lehnebach, C., Armstrong, T., Chen, L., & Lockhart, P. (2005). Phylogenetic relationships and evolutionary traits in Ranunculus s.l. (Ranunculaceae) inferred from ITS sequence analysis. *Molecular Phylogenetics and Evolution*, *36*(2), 305–327.

Hunter, S. S., Settles, M. L., New, D. D., Parent, C. E., & Gerritsen, A. T. (2016). Mitochondrial Genome Sequence of the Galápagos Endemic Land Snail *Naesiotus nux*. *Genome Announcements*, *4*(1), e01362-15.

Jörger, K. M., Stöger, I., Kano, Y., Fukuda, H., Knebelsberger, T., & Schrödl, M. (2010). On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. *BMC Evolutionary Biology*, *10*, 323.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649.

Knudsen, B., Kohn, A. B., Nahir, B., McFadden, C. S., & Moroz, L. L. (2006). Complete DNA sequence of the mitochondrial genome of the sea-slug, *Aplysia californica*: Conservation of the gene order in Euthyneura. *Molecular Phylogenetics and Evolution*, *38*(2), 459–469.

Kurabayashi, A., & Ueshima, R. (2000). Complete Sequence of the Mitochondrial DNA of the Primitive Opisthobranch Gastropod *Pupa strigosa*: Systematic Implication of the Genome Organization. *Molecular Biology and Evolution*, *17*(2), 266–277.

Lydeard, C., Cowie, R. H., Ponder, W. F., Bogan, A. E., Bouchet, P., Clark, S. A., … Thompson, F. G. (2004). The Global Decline of Nonmarine Mollusks. *BioScience*, *54*(4), 321–330.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*(1), 10–12.

Matisoo-Smith, E., & Daugherty, C. (2012). Africa to Aotearoa: the longest migration. *Journal of the Royal Society of New Zealand*, *42*(2), 87–92.

Mitchell, K. J., Llamas, B., Soubrier, J., Rawlence, N. J., Worthy, T. H., Wood, J., … Cooper, A. (2014). Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution. *Science*, *344*(6186), 898–900.

Morgan-Richards, M., & Trewick, S. A. (2005). Hybrid origin of a parthenogenetic genus? *Molecular Ecology*, *14*(7), 2133–2142.

Morgan-Richards, M., Trewick, S. A., Bartosch-Härlid, A., Kardailsky, O., Phillips, M. J., McLenachan, P. A., & Penny, D. (2008). Bird evolution: testing the Metaves clade with six new mitochondrial genomes. *BMC Evolutionary Biology*, *8*, 20.

Nei, M., & Rooney, A. P. (2005). Concerted and Birth-and-Death Evolution of Multigene Families. *Annual Review of Genetics*, *39*, 121–152.

Ohno, S. (1970). *Evolution by Gene Duplication*. Springer Science & Business Media.

Poczai, P., & Hyvönen, J. (2009). Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. *Molecular Biology Reports*, *37*(4), 1897–1912.

Ponder, W. F., Colgan, D. J., Gleeson, D. M., & Sherley, G. H. (2003). Relationships of Placostylus from Lord Howe Island: an investigation using the mitochondrial cytochrome c oxidase 1 gene. *Molluscan Research*, *23*(2), 159–178.

Pratt, R. C., Gibb, G. C., Morgan-Richards, M., Phillips, M. J., Hendy, M. D., & Penny, D. (2009). Toward Resolving Deep Neoaves Phylogeny: Data, Signal Enhancement, and Priors. *Molecular Biology and Evolution*, *26*(2), 313–326.

Schaal, B. A., & Learn, G. H. (1988). Ribosomal DNA Variation within and among Plant Populations. *Annals of the Missouri Botanical Garden*, *75*(4), 1207–1216.

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., … Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, *109*(16), 6241–6246.

Shen, X., Tian, M., Meng, X., Liu, H., Cheng, H., Zhu, C., & Zhao, F. (2012). Complete mitochondrial genome of *Membranipora grandicella* (Bryozoa: Cheilostomatida) determined with next-generation sequencing: The first representative of the suborder Malacostegina. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *7*(3), 248–253.

Shen, X., Wang, H., Ren, J., Tian, M., & Wang, M. (2009). The mitochondrial genome of *Euphausia superba* (Prydz Bay) (Crustacea: Malacostraca: Euphausiacea) reveals a novel gene arrangement and potential molecular markers. *Molecular Biology Reports*, *37*(2), 771–784.

Shioiri, C., & Takahata, N. (2001). Skew of Mononucleotide Frequencies, Relative Abundance of Dinucleotides, and DNA Strand Asymmetry. *Journal of Molecular Evolution*, *53*(4–5), 364–376.

Torczynski, R., Bollon, A. P., & Fuke, M. (1983). The complete nucleotide sequence of the rat 18S ribosomal RNA gene and comparison with the respective yeast and frog genes. *Nucleic Acids Research*, *11*(14), 4879–4890.

Trewick, S. A. (2000). Molecular evidence for dispersal rather than vicariance as the origin of flightless insect species on the Chatham Islands, New Zealand. *Journal of Biogeography*, *27*(5), 1189–1200.

Trewick, S. A., Walker, K. J., & Jordan, C. J. (2008). Taxonomic and conservation status of a newly discovered giant landsnail from Mount Augustus, New Zealand. *Conservation Genetics*, *9*(6), 1563–1575.

Trewick, S., Brescia, F., & Jordan, C. (2009). Diversity and phylogeny of New Caledonian *Placostylus* land snails; evidence from mitochondrial DNA. In *Grandcolas P ed. Zoologia Neocaledonica 7* (pp. 421–436).

Triggs, S. J., & Sherley, G. H. (1993). Allozyme genetic diversity in *Placostylus* land snails and implications for conservation. *New Zealand Journal of Zoology*, *20*(1), 19–33.

Volkov, R., Kostishin, S., Ehrendorfer, F., & Schweizer, D. (1995). Molecular organization and evolution of the external transcribed rDNA spacer region in two diploid relatives of *Nicotiana tabacum*. *Plant Systematics and Evolution*, *201*(1–4), 117–129.

White, T. R., Conrad, M. M., Tseng, R., Balayan, S., Golding, R., de Frias Martins, A. M., & Dayrat, B. A. (2011). Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. *BMC Evolutionary Biology*, *11*, 295.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helm-Bychowski, K. M., … Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society*, *26*(4), 375–400.

Wu, S., Xiong, J., & Yu, Y. (2015). Taxonomic Resolutions Based on 18S rRNA Genes: A Case Study of Subclass Copepoda. *PLOS ONE*, *10*(6), e0131498.

Xia, X., Xie, Z., & Kjer, K. M. (2003). 18S Ribosomal RNA and Tetrapod Phylogeny. *Systematic Biology*, *52*(3), 283–295.

Zhou, X., Lin, Q., Fang, W., & Chen, X. (2014). The complete mitochondrial genomes of sixteen ardeid birds revealing the evolutionary process of the gene rearrangements. *BMC Genomics*, *15*, 573.

Chapter 4. Optimising landmark selection for geometric morphometric analysis of the New Zealand land snail *Placostylus ambagiosus*.

Introduction

The analysis of shape is fundamental to much biological research. The systematic taxonomic classification of organisms, and interpretation of the diversity of biological life, were established on the descriptions of morphological forms; the visible and preservable phenotype (Hennig 1966; Hennig 1950; Linnaeus 1758; Zimmerman 1834, 1930, 1931, 1943). Morphometrics is the quantitative study of biological shape variation (Bookstein 1996). Although a variety of processes influence the shape variation observed among individuals including disease, injury, ontogenic development, it is the adaptation to local geographic features and long term evolutionary sorting that is the focus of this work. Geometric morphometrics allows empirical comparison of complex shapes using a landmark configuration. “A landmark configuration is a discrete sample from a homology mapping across specimens” (Bookstein 1996), a labeled series of points that are each homologous over an entire sample of individuals (Bookstein 1996). The use of biologically meaningful landmarks allows the shape of organisms (or parts of organisms) to be compared in a way that removes the confounding effects of scale (size), rotation and location (using procrustes superimposition). The distributions of populations shapes can be compared and deformations that relate to different shapes visualised. The approach has strong and well-understood statistical and shape-theoretical underpinnings (Bookstein 1996). When a variety of methods of studying shape are compared (dimensionless shape descriptors, elliptic Fourier analysis, Fast Fourier Transform of 2D outline, geometric morphometrics), geometric morphometrics is frequently found to be efficient in distinguishing species differences (Dieleman et al. 2015; Ponton 2006; Van Bocxlaer & Schultheiß 2010).

Landmarks provide the only data in any analysis within the geometric framework, and therefore shape variation in the space between landmarks cannot be detected (Webster and Sheets 2010). Difficulties in using Landmark analysis to document morphological variation in taxa with a relatively low morphological complexity (molluscs relative to vertebrates) “remain a potential obstacle to examining underlying paleobiological and evolutionary signals” (Van Bocxlaer & Schultheiß 2010). If curves or perimeters of structures between landmarks appear critical to

the research question, then the use of semi-landmarks is appropriate. The inclusion of semi-landmarks is a relatively new refinement to geometric morphometric methods, and the concept refers to points that are placed arbitrarily but consistently along a curve of interest: the curve is biologically homologous, but the placement of semi-landmarks on that curve is not (Bookstein 1997, 1996; Green 1996; Sheets et al. 2004). The incorporation of semi-landmarks into a data set permits analysis of components of shape that lack explicitly defined landmarks. During processing, semi-landmarks are adjusted to 'slide' along the curve (anchored at each end by a permanent landmark) in order to minimize variation in shape that is due simply to arbitrary placement of the points.

Logically more comprehensive placement of landmarks (and semi-landmarks) results in an increase in insight into the structure of shape variation. Mostly it is preferable to attempt to achieve the most complete coverage of the structure possible, given the constraints of time, sample size and patience (Webster and Sheets 2010). However, when specimens within a sample are highly variable for a given structure and fine scale variation of that structure does not add information for the discrimination of groups, it is worth removing. A problem then arises not only to balance complete coverage with the constraints of time, but also with the variability of specimens. A trade off results between including many specimens of the population sample with few landmarks or few specimens with complete landmarks. Surprisingly few studies explicitly compare results from varying numbers of semi-landmarks to optimize dataset size. Here I examine the optimal number of landmarks required to give sufficient information in the least amount of time so that a larger number of individuals can be included in subsequent analyses. I investigate how the choice of semi-landmarks affects discrimination of sample groups.

Shape analysis is fundamental to palaeontology, where the phenotype of most taxa are represented in the fossil record only by hard-parts. The shells of snails (Gastropoda) are well represented in the fossil record, and have therefore contributed extensively to the study of past biodiversity (Crampton 2010) and analysis of occupancy curves (Foote et al. 2007). But shells produced by snails are

highly sensitive to local environmental conditions and have provided material for studying local adaptation (Giokas et al. 2013) and ecotypes. Although shell morphology in some species of gastropods has been shown to be plastic with respect to the environment and predators (Appleton & Palmer 1988; Goodfriend 1986; Hollander & Butlin 2010; Trussell 2000b, 2000a), shell characteristics can also reflect fixed genetic attributes (e.g. Goodfriend 1986; Johannesson & Johannesson 1996; Stankowski 2013).

Geometric morphometric techniques, not previously used for New Zealand *Placostylus* snails are applied to the New Zealand taxa to explore variation in shell shape and size in time and space. Recently, a combination of landmark and sliding semi-landmark (analysis of curves) has been used successfully to identify morphologically distinct sympatric species in New Caledonian *Placostylus* (Dowle et al. 2015). These techniques have also been applied to fossil and modern taxa (Aguirre et al. 2006). Such analyses are challenging because mollusc shells, and in particular gastropod shells that lack hinges and visible muscle attachment features, have few homologous points on which to place landmarks. Fortunately, what they lack in landmark points they make up for in curves associated with body whorl and aperture. Although the three species of New Zealand *Placostylus* are now rare and have threatened conservation status there is abundant shell material of both modern and Holocene fossil populations available for study. Thus, ensuring good representation of the available individuals while maximising the information gained per sample is possible and appropriate.

The configuration of landmarks and/or semi-landmarks taken from a two dimensional digital image will always be associated with some degree of measurement error. Inconsistent tilting (pitch/roll) of specimens relative to the plane of digitisation, non-coplanarity of landmarks and or difficulty in pinpointing the landmark locus due to biological vagueness, or differences in illumination or focus, and variation between operator interpretation can all be influential. Image distortion can also introduce non–biological signal. Measurement error can be minimized by careful mounting, photography and landmark selection, but can never be totally eliminated (Collins & Gazley 2016).

Here I determine the degree of variation that can be attributed to orientation of specimens for photographing and placement of landmarks on *Placostylus* shell photographs. Working with a dataset with nine landmarks and sixty-three semi-landmarks, the configuration (Figure 5.1) and minimum number of semi-landmarks required to retain sufficient information to identify population samples of *Placostylus* shells was determined.

Methods

Material

Material used in this analysis came from the collections of the Museum of New Zealand Te Papa Tongarewa (NMNZ), and originated from the far north of New Zealand. Modern samples represent extant populations of *Placostylus ambagiosus* at Maungapika, Waikuku, Unuwhao, Whareana, Surville Cliffs and Cape Maria van Diemen, (n = 197), *P. hongii* from Tawhiti Rahi Island (Poor Knights Island Group) (n = 5), and *P. bollonsi* from Manawatawhi (Three Kings Island) (n = 5). Holocene fossil shells of *Placostylus ambagiosus* represent four temporal samples from sites near Tom Bowling Bay a region not currently occupied by *Placostylus* snails. Fossil populations were coded by date estimated from radiometric Carbon dating (see supplementary material for collection details).

Photography and digitisation

Intact adult snail shells from each sample population were bedded in sand of contrasting colour. A digital image of the ventral surface of each shell with a scale marker was captured using a high-resolution digital single-lens reflex Cannon EOS 600d camera with EF100mm f2.8 USM macro lens. Photographic equipment was mounted on a high-precision Kaiser stand to allow reproducible positioning and orientation.

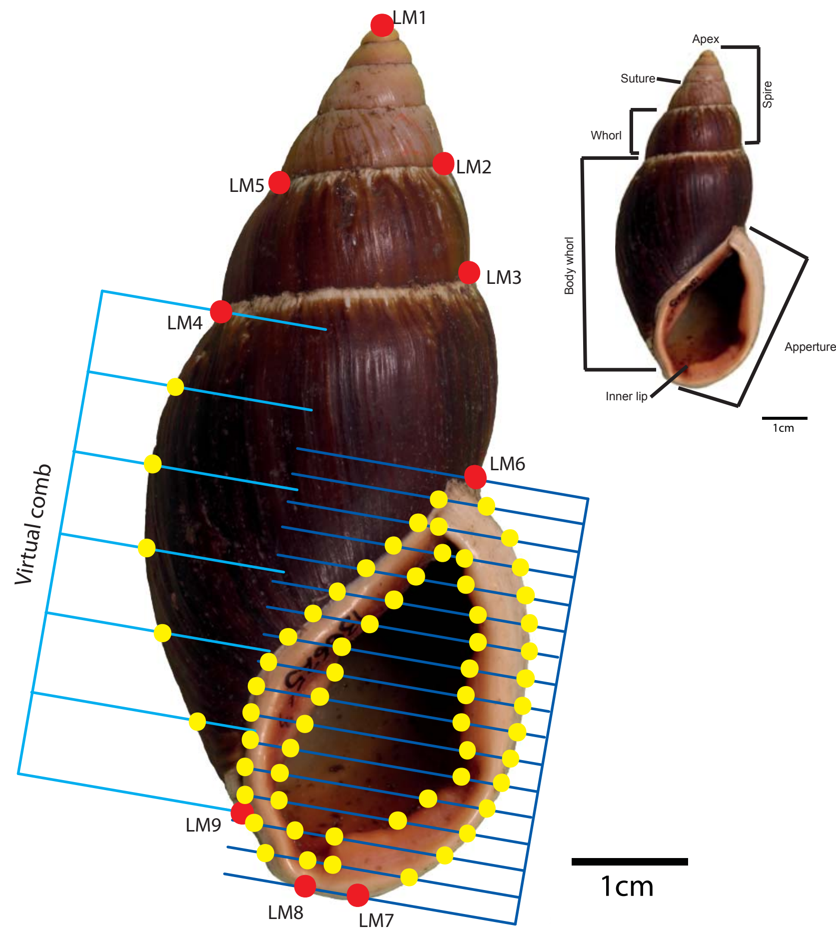


Figure 4.1. *Placostylus ambagiosus* showing orientation and placement of 9 landmarks (red) and 63 semi-landmarks (yellow) for 2D geometric morphometric analysis. Inset shows parts of the shell.

Nine permanent and 63 semi-landmarks were placed around the aperture and shell outline when placed in the ventral view (Figure 4.1). This configuration was made up of five open curves of equally spaced semi-landmarks (5, 15, 15, 15, 15 respectively) that were anchored by starting and ending landmarks. Two virtual combs used to locate semi-landmarks were first placed on images and aligned manually in Adobe Photoshop CS2 v 9.0.2. Combs were aligned to a centreline running from the apex of the shell to the intercept of the aperture and body whorl (see Figure 3). Digitizing was under-taken in TPSDIG v 2.17 (Rohlf 2013) on a Wacom Cintiq 22HD digitizing tablet. Digitized semi-landmarks were slid using SEMILAND, part of the IMP714 package (Sheets 2012; Zelditch et al. 2012), implementing the Procrustes distance method. Landmark X/Y coordinates were then imported into MORPHOJ 1.05F (Klingenberg 2011). Sliding eliminates variation

due to differences in translation, orientation and size. The Procrustes distance method is preferred when morphological variation is relatively small (Perez et al. 2006). Inclusion of semi-landmarks in a geometric morphometric analysis reduces the degrees of freedom in ways that are difficult to quantify, so statistical analyses based on such data employed nonparametric bootstrap resampling approaches (Hills et al. 2012; Zelditch et al. 2012).

The information content of data form a range of landmark and semi landmark permutations (Figure 2; A–N) were assessed. All permutations included landmarks 1, 3, 4, 6, 7, 8, and 9. Semi-landmark combinations were either reduced in frequency along curves (G, I, and M in Figure 2) and/or excluded for a given curve (C, D, E, F, H, J, K, L and N in Figure 2).

Sensitivity analysis

Measurement error was separately assessed for one specimen of *Placostylus ambagiosus* from Cape Maria van Diemen. This specimen was mounted photographed and dismounted a total of five times. A landmark configuration was digitized from each of the independent replicate images providing a measure of the error associated with mounting specimens. One of these images was then digitised five separate times providing a measure of the error associated with the digitization process. Down stream analysis was as for all other datasets and a Principal Component analysis was performed to visualise the error associated with photography and digitisation. A disparity analysis was performed, to calculate the proportion of variance contributed by each source of error (photographic error and digitization error), in R (R Core Team 2013).

Shape variation

To display the major features of shape variation between individuals and as an ordination method, principal component analysis (PCA) was performed using the covariance matrix generated from the dataset in the software MORPHOJ. Resulting Eigen values were subject to broken-stick analyses calculating the number of components (PCs) that exceed the theoretical scree plot of random, uncorrelated noise in R-studio (RStudio Team 2015) using "Vegan" (Oksanen et al. 2016). The

resulting PCs were exported from MORPHOJ and saved in comma-separated value format. To find the shape features that best distinguish among multiple groups of specimens (rather than among individuals as does PCA), the allometry-free variables were used as input for a canonical variate analysis (CVA). CVA determines whether the predefined populations can be statistically distinguished based on the relative warp matrix and was used to display the major features of shape variation between populations.

The success of the CVA in assigning specimens to populations was determined using a leave-one-out cross-validation procedure (jack-knife test) in which each specimen is sequentially omitted from the initial calculation of the CV axes and used as a test set (Webster & Sheets 2010). The omitted specimen is then treated as an unknown and assigned to the population whose mean is closest using the CV axes performed in R-studio using "MASS". Projected means were calculated from the dataset and linear discriminant analysis was used to test assignment of individuals to the projected mean. This CVA procedure was repeated for each of the landmark configurations.

To visualise the shape features that best distinguish among multiple groups and to calculate Mahalanobis distances a CVA was also performed in MORPHOJ using the default settings of 10,000 iterations. Graphs and figures presented are from this analysis. The Mahalanobis distances are given as *D* values, (not *D*-squared values) which are interpretable, approximately, as distance measures scaled by the within-group standard deviation in the respective direction. The Mahalanobis distances for all pairs are computed using the pooled within-group covariance matrix for all the groups jointly. This usage reflects the assumption of CVA that the within-group covariance matrices are identical (Klingenberg 2011).

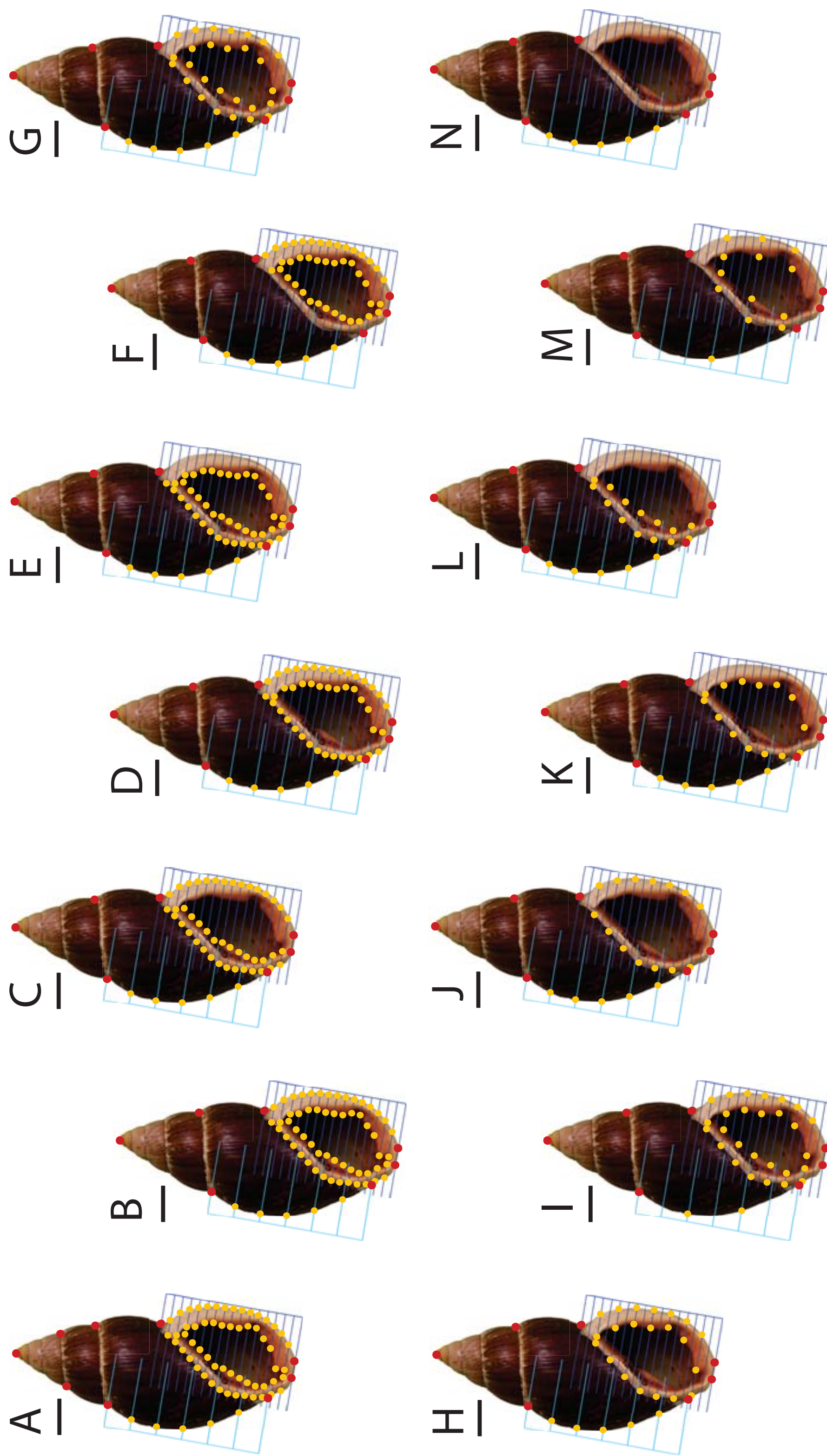


Figure 4.2. Permanent landmarks (red) and semi-landmarks (yellow) digitised in the geometric dataset for New Zealand *Placostylus*. shells. Images A–N represent the different configurations of landmarks to test optimised shape variation around the aperture and body whorl, which may vary between populations and current taxonomic species. Cross validation scores for each configuration are presented in Table 1. Scale bar = 1cm.

Mahalanobis distance matrices were saved as nexus files and imported into Splitstree v4.13.1 (Huson & Bryant 2006). The NeighbourNet function with EqualAngle method, which computes a planar split network for a circular (sub-) set of splits was used to visualise phenotype similarity. Other settings were explored however this one provided the most information-rich networks. Inferred phenotype relationships of each network were then compared by eye.

The criterion used to establish the optimal set of semi-landmarks was that it should be the minimum number that returned a cross-validation score for the population sample of *Placostylus bollonsi* of 100%. *Placostylus bollonsi* is morphologically and genetically distinct from the other New Zealand species of *Placostylus* (Buckley et al. 2011; Trewick et al. 2009). Extant population samples of *Placostylus ambagiosus* were required to return a cross-validation score above 70%. Inferred phenotype relationships from the selected semi-landmarks were required to be consistent with those of the full dataset.

Results

Sensitivity analysis

Sensitivity analysis was undertaken to quantify operator error associated with orientation of specimens for photography, placement of the comb, and digitisation of landmarks. The disparity analysis calculated the proportion of variance attributed to photographic and digitisation error combined as 3.5% of the variance within the Cape Maria van Diemen population (from which the sample used for error measurement was taken). Overall variation from photographic and digitisation accounts for <2.3% of the variation within the dataset as a whole. This indicates that operator error was negligible compared to intraspecific shape variation in *Placostylus* and was unlikely to influence biological interpretation of the results presented here. Placement of landmarks (digitisation error) introduced less variation (1.8%) than the repositioning of the shell for photography (3.2%) (for the Cape Maria van Diemen population). Comparison of two digitisation sets for the same specimens showed good repeatability for x, y coordinates. Principal

component analysis shows considerably less variation between replicates of photography and digitisation than within population samples (Figure 3).

Population samples of New Zealand *Placostylus* occupy overlapping regions of morphospace, regardless of landmark configuration, as shown by Principal Components Analyses (Figure 4). Means of many population samples were, however, significantly different in shape based on a procrustes ANOVA (partial Procrustes distance between means, F =17.84, p =<0.0001 for all extant populations). This test showed no significant variation for centroid size among populations. Using just PC1 and PC2 the *Placostylus bollonsi* snail shells form a distinct cluster of individuals (Figures 4 & 5).

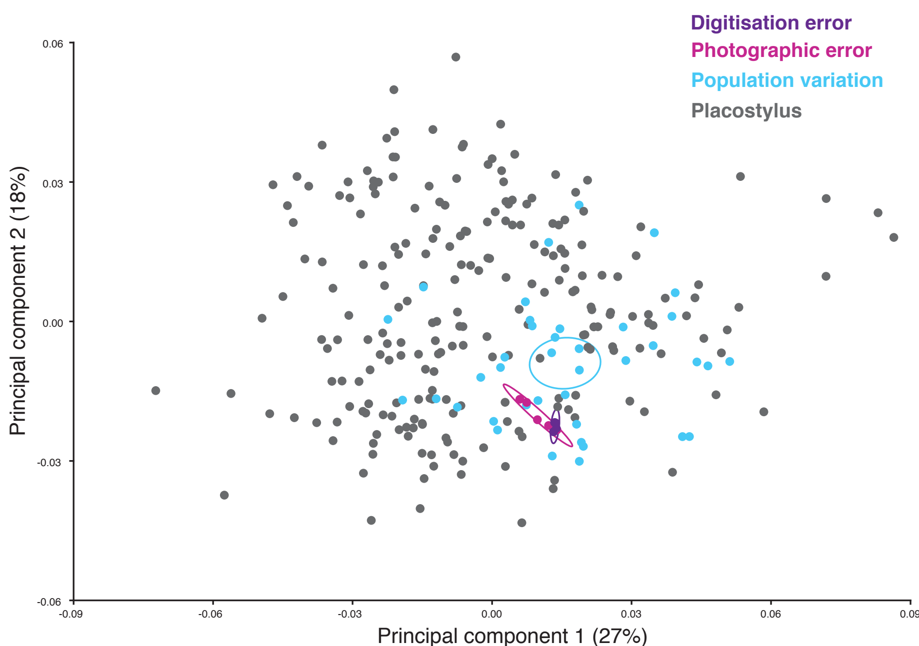


Figure 3: Principal components 1 and 2 of shell shape variation within New Zealand *Placostylus* to illustrate measurement error. The population dataset is used as a background dataset **(blue)** Replicate (five times) digitisation of one image of *Placostylus*; **(purple)** Replicate (five times) photography and subsequent digitisation of one specimen of *Placostylus ambagiosus* from Cape Maria van Diemen; **(pink)** Other populations of *Placostylus* including extant and fossil groups; **(grey)**. Elipses indicate 95% confidence intervals of the mean.

Nine combinations of landmarks and semi-landmarks were analyses independently for the same set of 275 snail shells. Depending on the landmark configuration tested between four and nine principal components exceeded the theoretical scree plot of random, uncorrelated noise when tested against a broken-stick model (Table 1, shown in grey italics). The initial landmark configuration (A) with all data showed the first nine PCs accounted for a total of 84.7% of the variation, whereas landmark configuration G showed the first seven PCs accounting for a total of 82.5% of the variation (Table 2). Landmark configurations J, M and N had the fewest significant principal components (5, 5, and 4 respectively; Table 2).

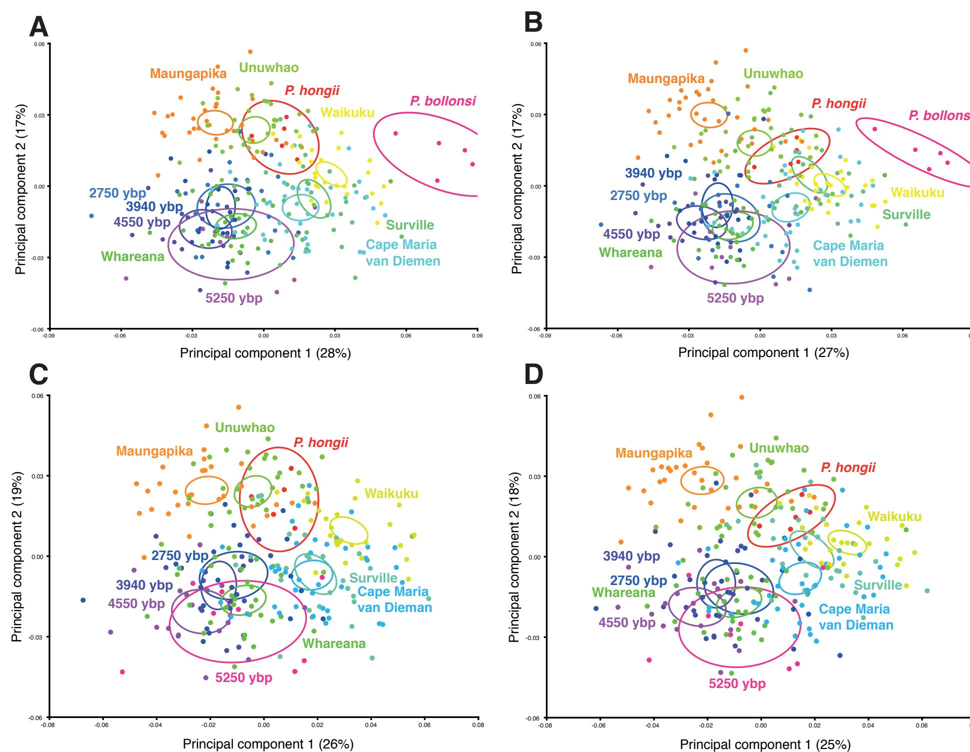


Figure 4: Principal component analysis of New Zealand *Placostylus* land snails test dataset. A: All 72 landmarks (A in figure 2), B: Optimal set of semi-landmarks (G in figure 2). C: All 72 landmarks without the *P. bollonsi* population. D: Optimal set of semi-landmarks without the *P. bollonsi* population. Ellipses show 95% confidence intervals of the mean of each population sample. Overlap in morphospace of population sample data is not increased by the exclusion of a subset of semi-landmarks.

Placostylus bollonsi (n=5) returned cross validation scores of either 80% or 100% depending on the landmarks used in the analysis (Table 1). *Placostylus hongii* (n=5) returned a cross validation scores between 40% and 100% (once each) and returned scores of 80% and 60% (4 times and 8 times respectively).

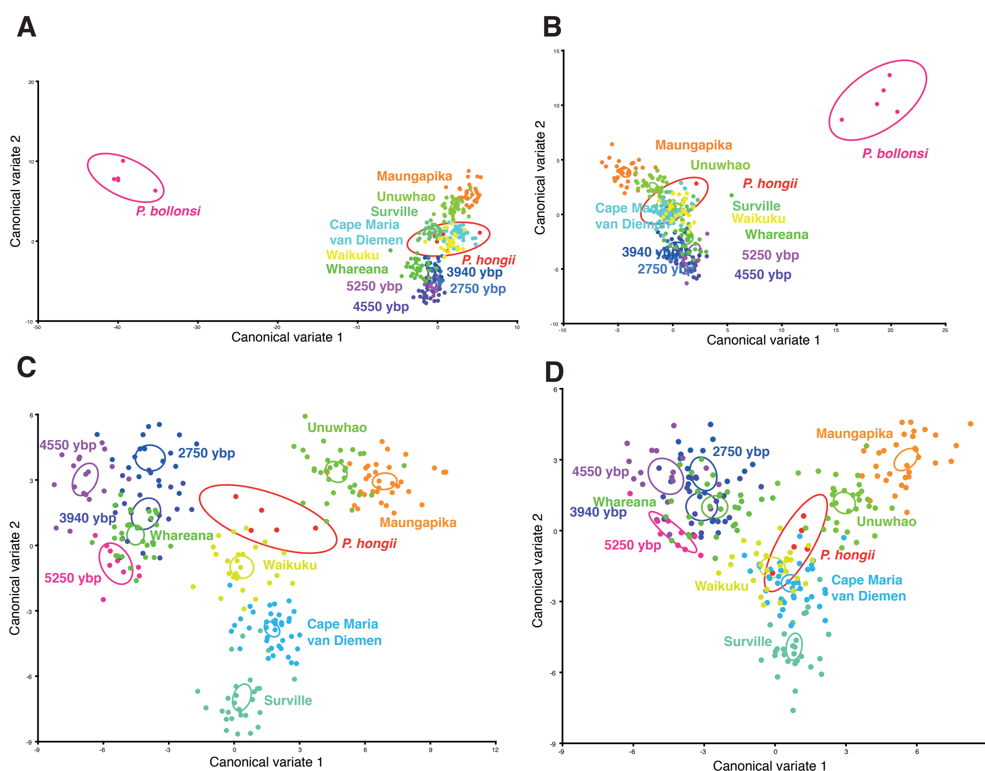


Figure 4.5. Canonical variance analysis plots of New Zealand *Placostylus* land snail test dataset. A. All 72 landmarks; B. Optimal configuration (G in Figure 2); C. All 72 landmarks without the *P. bollonsi* population; D. Optimal configuration without the *P. bollonsi* population. Ellipses show 95% confidence intervals of the means.

In contrast, many of the *Placostylus ambagiosus* extant population samples were successfully assigned to their local population using the cross-validation score. For example the snail shells from the extant population of Maungapika were correctly assigned to the *a priori* population more than 90% of the time in nine of the fourteen landmark combinations (Table 1). However, even with the full complement of landmarks and semi-landmarks (Figure 1 and 2A) several populations returned low assignment scores. This was most apparent in fossil population samples with the oldest population (5250 ybp) returning a cross-validation score of 0% five times out

of fourteen. Fossil misallocation tended to be caused by allocation into another fossil population or an extant but geographically close population (Table 4).

Overall, changing the number or configuration of semi-landmarks appeared to have little effect on the correct assignment to *a priori* groups of individual *Placostylus* specimens. Population samples with robust assignment scores continued to score highly and groups with weak scores continued to score low in this cross-validation test even when the semi-landmark set was reduced to half its original size. Removing entire curves from the landmark configuration did not improve the subsequent data, suggesting that even if the curve structure does not add information it is not so variable between individuals that it needs to be removed. In fact in most cases removal of an entire curve diminished the ability to determine correct assignment to a priori groups.

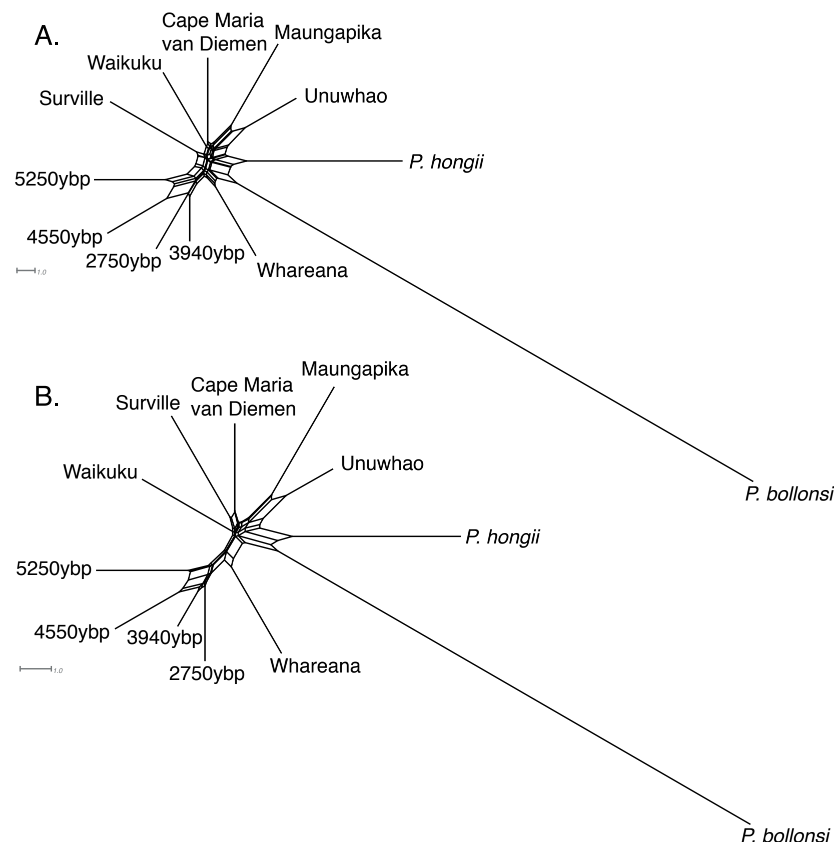


Figure 4.5. Phenotypic relationships among populations inferred from Mahalanobis distances using Neighbournet Equal Angle networks in Splitstree. for New Zealand *Placostylus* land snails. A. uses the landmark configuration shown in Figure 2A. B. uses the optimal landmark configuration shown in Figure 2G.

Table 4.1. Population samples of 275 *Placostylus* showing the results of the cross validation test (Jacknife) analysis by % correct in cross validation performed in R-studio using the "MASS" library. Labels A-O represent different configuration of landmarks shown in figure 1. Numbers in *italics* below labels A-N are the number of Principal Components exceeding the theoretical scree plot of random, uncorrelated noise and therefore used in downstream analysis. Columns in **bold** highlight configurations resulting in 100% correct allocation of *P. bollonsi*. The column in green is the configuration chosen for future landmark analysis.

				A	B	C	D	E	F	G	H	I	J	K	L	M	N
	no. of semi landmarks			63	63	49	49	47	47	33	33	30	19	19	19	14	5
	Population name	NMNZ	n	*9*	*9*	*8*	*8*	*9*	*8*	*7*	*6*	*7*	*5*	*6*	*6*	*5*	*4*
P. ambagiousus	Maungapika	102324	32	93.75	93.75	90.63	90.63	93.75	78.13	93.75	90.63	93.75	93.75	68.75	68.75	87.50	68.75
	Waikuku	102328	29	88.46	88.46	92.31	96.15	88.46	88.46	92.31	80.77	84.62	84.62	88.46	88.46	76.92	84.62
	Unuwhao	102503	36	87.50	93.75	75.00	78.13	87.50	81.25	75.00	68.75	71.88	56.25	43.75	43.75	71.88	56.25
	Whareana	136614	31	80.00	70.00	56.67	60.00	70.00	73.33	63.33	60.00	66.67	36.67	40.00	40.00	70.00	30.00
	Surville	136621	30	86.21	86.21	89.66	86.21	82.76	79.31	89.66	86.21	89.66	89.66	86.21	86.21	79.31	82.76
	CMvD	136625	39	92.31	89.74	69.23	82.05	89.74	87.18	82.05	69.23	82.05	58.97	64.10	64.10	69.23	61.54
	2750ybp	180400	20	35.00	45.00	30.00	15.00	40.00	20.00	15.00	25.00	20.00	15.00	15.00	15.00	30.00	15.00
	3940ybp	180421	20	40.00	35.00	25.00	20.00	40.00	20.00	20.00	25.00	20.00	30.00	20.00	20.00	20.00	25.00
	4550ybp	180437	18	61.11	55.56	66.67	55.56	44.44	55.56	44.44	50.00	44.44	61.11	44.44	44.44	38.89	0.00
	5250ybp	180465	10	44.44	33.33	0.00	33.33	22.22	22.22	22.22	22.22	22.22	0.00	0.00	0.00	33.33	0.00
	P. hongii	5786	5	80.00	80.00	60.00	100.00	80.00	80.00	60.00	40.00	60.00	60.00	60.00	60.00	60.00	60.00
	P. bollonsi	183761	5	100.00	100.00	80.00	100.00	100.00	80.00	100.00	100.00	100.00	80.00	100.00	100.00	100.00	0.00

Table 4.2. The percentage of variance accounted for by each principal component and the cumulative variance for two landmark configurations: all landmarks and chosen configuration (G in Figure 2) for *Placostylus* snails. Only principal components that exceed the theoretical scree plot of random, uncorrelated noise and therefore used in downstream analysis are shown.

	PC	1	2	3	4	5	6	7	8	9
all landmarks	% Variance	27.92	17.44	15.02	7.40	5.03	3.76	3.14	2.59	2.41
	Cumulative %	27.92	45.36	60.39	67.79	72.82	76.58	79.71	82.31	84.71
chosen configuration	% Variance	26.76	17.01	15.55	9.12	5.77	4.65	3.69	-	-
	Cumulative %	26.76	43.76	59.32	68.43	74.20	78.86	82.54	-	-

An optimal set of landmarks and semi-landmarks (Figure 2G) was selected for future work. Positions of populations relative to one another in morphospace do not vary greatly between the two configurations (Figure 2A: all landmarks; and 2G) (See Figures 4 & 5). CVA classification for extant populations did not vary greatly between the configurations A and G (Table 3). Fossil populations proved difficult to assign under all configurations although A performed better than the G. As all fossil populations in this study come from the same site (but multiple time periods) we might expect to find little difference between them. The phenotypic relationships inferred from Mahalanobis distances do not vary greatly between the optimal landmark configuration (Figure 2 G) and the landmark configuration containing all landmarks (Figure 2 A) (Figure 5). The cross validation test assigned extant individuals with relative certainty to their *a priori* groups (Table 1). The optimal dataset has seven landmarks & 33 semi-landmarks and includes all four aperture curves and the curve from the body whorl. This significantly decreases the time spent digitising shells, but maintains biologically important information allowing discrimination of extant populations.

Table 4.3. CVA classification for New Zealand Placostylus using cross validation scores from geometric morphometric analysis. (Upper) nine landmarks and 63 semi-land (A in Figure 2). (Lower) seven landmarks and 33 semi-landmarks (G in Figure 2).

Population name	NMNZ code	Maungapika	Waikuku	Unuwhao	Whareana	Surville	Cape Maria	2750ybp	3940ybp	4550ybp	5250ybp	*P. hongii*	*P. bollonsi*
Maungapika	102324	**30**	0	2	0	0	0	0	0	0	0	0	0
Waikuku	102328	0	**23**	1	0	0	2	0	0	0	0	0	0
Unuwhao	102503	1	1	**28**	0	0	2	0	0	0	0	0	0
Whareana	136614	0	1	0	**24**	0	0	3	2	0	0	0	0
Surville	136621	0	2	1	0	**25**	1	0	0	0	0	0	0
CMvD	136625	0	0	1	0	2	**36**	0	0	0	0	0	0
2750ybp	180400	0	1	0	9	0	0	**7**	1	1	1	0	0
3940ybp	180421	0	0	0	6	1	0	4	**8**	1	0	0	0
4550ybp	180437	0	0	0	1	0	0	1	4	**11**	1	0	0
5250ybp	180465	0	0	0	0	0	0	1	2	2	**4**	0	0
P. hongii	5786	0	0	1	0	0	0	0	0	0	0	**4**	0
P. bollonsi	183761	0	0	0	0	0	0	0	0	0	0	0	**5**
Total		31	28	34	40	28	41	16	17	15	6	4	5
Maungapika	102324	**30**	0	2	0	0	0	0	0	0	0	0	0
Waikuku	102328	0	**24**	1	0	0	1	0	0	0	0	0	0
Unuwhao	102503	3	1	**24**	0	0	2	0	0	2	0	0	0
Whareana	136614	0	1	0	**19**	1	3	3	2	0	1	0	0
Surville	136621	0	1	1	0	**26**	1	0	0	0	0	0	0
CMvD	136625	0	0	1	1	5	**32**	0	0	0	0	0	0
2750ybp	180400	1	0	2	5	0	3	**3**	1	2	3	0	0
3940ybp	180421	0	0	1	5	1	1	3	**4**	4	1	0	0
4550ybp	180437	0	0	1	0	0	1	0	5	**8**	3	0	0
5250ybp	180465	0	1	0	1	0	0	1	3	1	**2**	0	0
P. hongii	5786	0	0	2	0	0	0	0	0	0	0	**3**	0
P. bollonsi	183761	0	0	0	0	0	0	0	0	0	0	0	**5**
Total		34	28	35	31	33	44	10	15	17	10	3	5

Discussion

Sensitivity analysis of photographing, positioning and digitising was successfully used to establish that measurement error was small compared to biological variation among populations, regardless of the landmark configuration. Sensitivity analysis should be calculated for future studies as different species and the resulting differences in shell morphology and resultant landmark configurations are likely to produce variations in error. Photographic error is further reduced in New Zealand *Placostylus* as size variation between shells is minimal and therefore the camera lens and distance from lens can be kept constant. If the same operator orientates and positions the shells for photographing the resulting error is much less than the naturally occurring variation among individuals of the same populations. Because placement of landmarks incorporates even less error in this study, it is possible that different operators could perform this task within a study although it would be prudent to calculate early on in the study the digitisation error this introduces.

There are few examples of comparisons between the density and placement of landmarks and semi-landmarks in the published literature. Although more comprehensive placement of landmarks (and semi-landmarks) might result in increased resolution of shape variation, time constraints demand a trade off between inclusion of many specimens with fewer landmarks or fewer specimens with more landmarks. Semi-landmark analysis has been found to capture overall variability in a comparable manner to Elliptic Fourier Analysis (EFA) (Dieleman et al. 2015; Hills et al. 2012; Van Bocxlaer & Schultheiß 2010) and has the advantage of being much more time efficient.

The optimal configuration of landmarks found for New Zealand *Placostylus* uses 7 permanent landmarks and 33 semi-landmarks, it includes all four curves of the aperture and the curve of the body whorl. The use of a smaller number of semi-landmarks greatly decreases the time required to process specimens (by ~five minutes per sample), while maintaining biologically informative shape variation. These landmarks are used in the subsequent chapters to explore shape variation of New Zealand *Placostylus* shells in time and space.

References

Aguirre, M. L., Perez, S. I., & Sirch, Y. N. (2006). Morphological variability of *Brachidontes swainson* (Bivalvia, Mytilidae) in the marine Quaternary of Argentina (SW Atlantic). *Palaeogeography, Palaeoclimatology, Palaeoecology*, *239*(1–2), 100–125.

Appleton, R. D., & Palmer, A. R. (1988). Water-borne stimuli released by predatory crabs and damaged prey induce more predator-resistant shells in a marine gastropod. *Proceedings of the National Academy of Sciences*, *85*(12), 4387–4391.

Bookstein, F. L. (1996). Biometrics, biomathematics and the morphometric synthesis. *Bulletin of Mathematical Biology*, *58*(2), 313–365.

Bookstein, F. L. (1997). Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis*, *1*(3), 225–243.

Bookstien, F. (1996). Applying landmark methods to biological outline data. In *Image fusion and shape variability* (pp. 59–70). Leeds, U.K.: University of Leeds.

Buckley, T., Stringer, I., Gleeson, D., Howitt, R., Attanayake, D., Parrish, R., … Rohan, M. (2011). A revision of the New Zealand *Placostylus* land snails using mitochondrial DNA and shell morphometric analyses, with implications for conservation. *New Zealand Journal of Zoology*, *38*(1), 55–81.

Collins, K. S., & Gazley, M. F. (2016). Does my posterior look big in this? - The effect of photographic distortion on morphometric analyses. *In Prep.*

Crampton, J. S., Cooper, R., Beu, A. G., Foote, M., & Marshall, B. A. (2010). Biotic influences on species duration: interactions between traits in marine molluscs. *Paleobiology*, *36*(2), 204–223.

Dieleman, J., Bocxlaer, B. V., Manntschke, C., Nyingi, D. W., Adriaens, D., & Verschuren, D. (2015). Tracing functional adaptation in African cichlid fishes through morphometric analysis of fossil teeth: exploring the methods. *Hydrobiologia*, *755*(1), 73–88.

Dowle, E. J., Morgan-Richards, M., Brescia, F., & Trewick, S. A. (2015). Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. *Molecular Ecology*, *24*(16), 4205–4221.

Foote, M., Crampton, J. S., Beu, A. G., Marshall, B. A., Cooper, R. A., Maxwell, P. A., & Matcham, I. (2007). Rise and Fall of Species Occupancy in Cenozoic Fossil Mollusks. *Science*, *318*(5853), 1131–1134.

Giokas, S., Páll-Gergely, B., & Mettouris, O. (2013). Nonrandom variation of morphological traits across environmental gradients in a land snail. *Evolutionary Ecology*, *28*(2), 323–340.

Goodfriend, G. A. (1986). Variation in Land-snail Shell form and Size and its Causes: a Review. *Systematic Biology*, *35*(2), 204–223.

Green, W. D. K. (1996). The thin-plate spline and images with curving features. In *Image fusion and shape variability* (pp. 79–87). Leeds, U.K.: University of Leeds.

Hennig, Q. (1966). *Phylogenetic Systematics*. Urbana: University of Illinois Press.

Hennig, W. (1950). *Grundzuge einer Theorie der phylogenetischen Systematik*. Berlin: Deutscher Zentraverlag.

Hills, S. F. K., Crampton, J. S., Trewick, S. A., & Morgan-Richards, M. (2012). DNA and Morphology Unite Two Species and 10 Million Year Old Fossils. *PLOS ONE*, *7*(12), e52083.

Hollander, J., & Butlin, R. K. (2010). The adaptive value of phenotypic plasticity in two ecotypes of a marine gastropod. *BMC Evolutionary Biology*, *10*, 333.

Huson, D. H., & Bryant, D. (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, *23*(2), 254–267.

Johannesson, B., & Johannesson, K. (1996). Population differences in behaviour and morphology in the snail Littorina saxatilis: phenotypic plasticity or genetic differentiation? *Journal of Zoology*, *240*(3), 475–493.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*(2), 353–357.

Linnaeus, C. (1758). *Systema Naturae* (10th ed., Vol. 1). Stockholm.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2016). vegan: Community Ecology Package (Version 2.4-1). Retrieved from https://cran.r-project.org/web/packages/vegan/index.html

Perez, S. I., Bernal, V., & Gonzalez, P. N. (2006). Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy*, *208*(6), 769–784.

Ponton, D. (2006). Is geometric morphometrics efficient for comparing otolith shape of different fish species? *Journal of Morphology*, *267*(6), 750–757.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Rohlf, F. (2013). TpsDig2 (Version 2.17).

RStudio Team. (2015). *RStudio: Integrated Development for R.* Boston, MA: RStudio, Inc. Retrieved from https://www.rstudio.com/

Sheets, H. (2012). *IMP714 - Integrated Morphometrics Package.* Retrieved from http://www3canisius.edu/~sheets/imp714.htm

Sheets, H. D., Keonho, K., & Mitchell, C. E. (2004). A combined landmark and outline-based approach to ontogentic shape change in the Ordovician Trilobite Trianthrus becki. In *Applications of morphometrics in paleontology and biology* (pp. 67–81). New York: Springer.

Stankowski, S. (2013). Ecological speciation in an island snail: evidence for the parallel evolution of a novel ecotype and maintenance by ecologically dependent postzygotic isolation. *Molecular Ecology*, *22*(10), 2726–2741.

Trewick, S., Brescia, F., & Jordan, C. (2009). Diversity and phylogeny of New Caledonian Placostylus land snails; evidence from mitochondrial DNA. In *Grandcolas P ed. Zoologia Neocaledonica 7* (pp. 421–436).

Trussell, G. C. (2000a). Phenotypic Clines, Plasticity, and Morphological Trade-Offs in an Intertidal Snail. *Evolution*, *54*(1), 151–166.

Trussell, G. C. (2000b). Predator-induced plasticity and morphological trade-offs in latitudinally separated populations of Littorina obtusata. *Evolutionary Ecology Research*, *2*(6), 803–822.

Van Bocxlaer, B., & Schultheiß, R. (2010). Comparison of morphometric techniques for shapes with few homologous landmarks based on machine-learning approaches to biological discrimination. *Paleobiology*, *36*(3), 497–515.

Webster, M., & Sheets, H. D. (2010). A practical introduction to landmark-based geometric morphometrics. *Quantitative Methods in Paleobiology*, *16*, 168–188.

Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric Morphometrics for Biologists: A Primer*. Academic Press.

Zimmerman, W. (1834). Research on phylogeny of species and of single characters. *The American Naturalist*, *68*(717), 381–384.

Zimmerman, W. (1930). *Die Phylogenie der Pflanzen*.

Zimmerman, W. (1931). Abeitsweise der botanischen Phylogenetik und anderer Gruppierungswissenscheften. In *Handbuch der biolgischen Arbeitsmethoden* (pp. 941–1053). Berlin: Urban and Schwarsenberg.

Zimmerman, W. (1943). Die Methoden der Phylogenetik. In *Die Evolution der Organismen* (pp. 20–56). Germany: G Fischer, Jena.

Chapter 5. Understanding morphological stasis in a heterogenous environment.

"Understanding the potential of gene flow and population structure to promote stasis requires…. realistic temporal and spatial variation [while] tracking phenotypic evolution at the level of populations and species. … Paleontological examples that decompose species trajectories into local or within-habitat components would be ***invaluable****"*

(Hunt & Rabosky, 2014)

Introduction

Fossils provided the first major insights into evolution and they continue to provide empirical evidence of major features of form change through geological time. Fossils have also been seen as intimately linked to stratification and when William Smith published the first geological map of Britain in 1815, much of the information he used came from index fossils; species recognisable as belonging to a particular geological horizon. This observation that pre-empted accepted evolutionary theory signalled a pattern involving prolonged periods of morphological stasis. Reconciling observations of variation among living taxa with that observed in the fossil record, time needs to account for differing scales and causes of variation. Modelling phenotypic variation within lineages through geological time has become a standard tool in the paleontological literature, but identification of the processes that result in these patterns of variation is less common. Phenotypic change over ecological timescales within single populations, where evolution takes place, is rarely associated with paleontological data, but see recent work on molluscs (Bocxlaer & Hunt 2013; Chiba 1996; Ubukata et al. 2009) and three-spine sticklebacks (Hunt et al. 2008).

An empirical approach to document form-change seeks to assess the degree of variation among individuals through time. Since Darwin, evolutionary biologists have questioned whether change occurs gradually or in a stepwise manner. Underlying this question is the fact that individuals within populations are not identical and survival and reproduction may not be randomly distributed across that variation. Even stochastic processes can result in a change in mean variation from one generation to the next but this might not result in net change over many generations.

Morphological stasis is one of three modes of phenotypic change considered to be represented in the fossil record, and contrasts with random walks and directional evolution (Hunt 2006). Stasis is modelled (Figure 5.1 A) as uncorrelated variation (white noise) around a steady mean resulting in a pattern of little net change in mean population morphology through time. Trait increments in the Random walk model are independent and equally likely to be increases or decreases. "This model is characterised by a parameter, the step variance, that determines the size of the steps and hence the evolutionary rate" (Hunt et al. 2015) (Figure 5.1 B) resulting in patterns of varying population means but in a way that actually varies little over evolutionary timescales. "Directional change (Figure 5.1 C) is a generalisation of the random walk model allowing for evolutionary steps that are biased toward increase or decreases" (Hunt et al. 2015). This results in a pattern of population mean morphology changing in a stepwise and directional manner through time. Identifying the mode of evolution, provides the starting point for the more challenging task of identifying specific causal mechanisms (Hunt & Rabosky 2014).

A single lineage may vary in its mode of evolution through time (stasis – directional change – stasis) and these complex models are favoured if fossil sequences are preserved in a long time-series (Hunt et al. 2015). When multiple traits are examined per linages it is possible to find differing modes of evolution for different traits. So it is surprising how rarely analyses of fossil lineages try to avoid the non-independence of phenotypic traits by using principal component analysis (Hopkins & Lidgard 2012).

Well documented fossil lineages typically display some variation within and between spatial and temporal samples, but the distribution of this variation often models stasis and morphological stasis is cited as a common feature in the fossil record (Hunt 2007; Hunt et al. 2015). Empirical evidence for directional evolution appears rare with 5% – 13% of studies finding support for directional evolution (see Figure 5.1C) (Hopkins & Lidgard 2012; Gene Hunt et al. 2015).

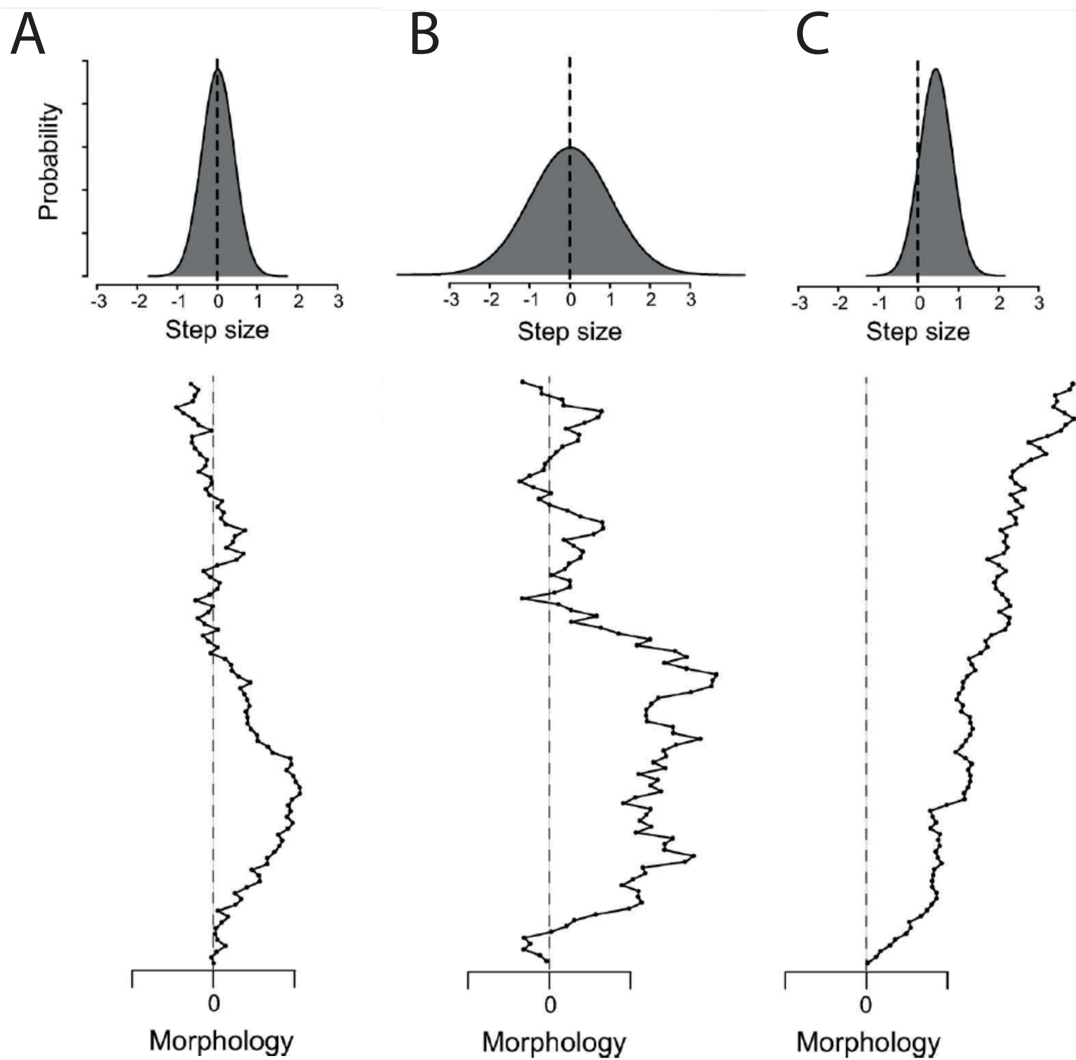


Figure 5.1. Three example distributions (top) used to generate corresponding evolutionary sequences of 100 steps (bottom). When the mean of the step distribution is zero, increases and decreases are equally likely and the overall dynamics are non-directional (A: stasis, B: random walk). Step distribution C has a positive mean, and therefore will tend to produce positively trended evolutionary sequences (directional change). With increasing step variance, evolutionary sequences are more volatile, with larger positive and negative excursions (compare A with B). Stasis (A) is characterised by variance around an optimum but no accumulation of net morphological difference over time. Figure modified from Hunt (2006).

The general inference that stasis is common in the fossil record does not sit perfectly with what is known from micro-evolutionary studies. In populations of living taxa genetic variation is abundant for the majority of traits (Hansen & Houle 2004), quantitative traits typically harbour substantial genetic variation (Hansen et al. 2011; Houle 1992; Houle 1998), intense natural selection is commonplace (Hereford et al. 2004), and evolutionary change is observable over generations (Hendry & Kinnison 1999; Kinnison & Hendry 2001; but see Merilä et al. 2001). The contrasting

observations of evolutionary change across long- and short-time scales are known as the "paradox of stasis" (Charlesworth et al. 1982; Ehrlich & Raven 1969; Eldredge et al. 2005; Estes & Arnold 2007).

Constraining selection, niche tracking and species cohesion have been proposed as possible explanation of stasis (Charlesworth et al. 1982; Ehrlich & Raven 1969; Eldredge et al. 2005; Estes & Arnold 2007). It is also quite possible that in at least some studies, fossils collected within a horizon came from populations that experienced different local conditions. This could result in sample pooling that conceals the spatial partitioning of trait variation.

Constraining selection

Studies of living species reveal that individuals are generally well adapted to their environment and decreased fitness is associated with increasingly extreme trait values suggesting that constraining selection is widely influential. This pattern is seen in both natural populations and in artificial selection experiments. Selection studies find stabilizing selection to be fundamental and suggest that external forces such as predation and competition, and internal constraints that are a consequence of functional interactions are influential in trait evolution (Estes & Arnold 2007; Kingsolver et al. 2001). Therefore, constraining selection should result in a fossil sequence that fits a stasis model of evolution.

Niche tracking

Stasis within species may result from their ability to track a favoured niche as it moves through space and time. Most paleontological data (especially that drawn from macroscopic animals) is gathered from samples spanning many millions of years, incorporating change over relatively long biological timescales. Evidence for habitat tracking comes from the apparent absence of morphological response to climatic change, implying instead a change in geographic distribution (Brett et al. 2007). Range shifting rather than adaptation has been interpreted as selection favouring populations that are evolutionarily conservative at the species level (Eldredge et al. 2005).

Species cohesion in maintaining stasis

Species cohesion is an extension of J. B. S. Haldane's proposition that asymmetrical gene flow prevents local adaptation (Bridle & Vines 2007; Kirkpatrick & Barton 1997; Lenormand 2002). Haldane suggested that gene flow from the bulk of the population would prevent local adaptation at the periphery. A novel allele (or trait) appearing at the periphery of a species range might be swamped by gene flow from higher density regions even if the allele was well suited to conditions experienced at the edge (Case & Taper 2000). Thus over time, little morphological change might result of gene flow within species preventing local adaptation from being preserved (Hansen & Houle 2004; Lieberman & Dudgeon 1996; Eldredge et al. 2005). Increasingly detailed paleontological studies corroborate the potential importance of species cohesion/gene flow in maintaining stasis (Eldredge et al. 2005).

Difficulties in sampling through time and space

The geological time scales of many millions of years over which stasis has been usually inferred are unlikely to represent periods of unchanging environmental conditions experienced by populations (Hansen & Houle 2004). Sampling over geological timescales (through millennia) may mean changes occurring at ecological timescales (from changes in environmental conditions) are imperceptible. However, it has also been suggested that evolutionary response to environmental fluctuations might fit a model of stasis if environmental changes are fluctuating and not accumulating (Hunt et al. 2015). Therefore the scale of sampling in space and time needs to match the scale of habitat variation and environmental perturbations that is relevant to the organism being studied.

If spatially separate populations respond differently to local conditions through time this would be consistent with observations made of living populations (i.e. population genetic structuring). Sampling across such populations in one time horizon, as would readily happen when collecting some fossils of a given "species", could yield an estimate of variance that has no net difference from similar arbitrary sampling at another horizon. Palaeontology frequently struggles to explore separate populations but tends to aggregate fossil from what may have been multiple populations, in several discrete environments, into a single group. Sampling from

multiple populations in a single time horizon as would readily happen when collecting some fossils of a given “species” could give an estimate of variance that has no net difference from the sampling at another given horizon even if some populations had gone extinct or local conditions had changed and populations had responded with changes to morphology. The artefact of sampling might also be attributed to the way in which fossil deposits are formed, so that a single horizons may contain individuals from ecologically variable habitats/times, this combined with the difficulty in dating horizons at very fine scales may result in significant time series averaging within lineages.

In this study I focus on two lineages within a single terrestrial snail taxon; *Placostylus ambagiosus* in New Zealand. Rather than assuming I am sampling a uniform species, I analyses two time series (two locations) independently, thus reducing the pooling effect that might lead to erroneous inferences of stasis. Geometric morphometric analysis of shell shape is used to identify and model the evolution of independent traits in each lineage over ~4000 years. Sampling each lineage from a narrow geographic region (<3km^2) is within the range of extant populations of this species and reduces the possibility of niche tracking (Figure 5.2). This contrasts with the spatial phenotypic variation observed at seven extant populations. The time series of between 12 and 15 points, spans the scale of environmental perturbations likely to be biologically relevant to this large, herbivorous land snail.

If stasis is observed within *Placostylus ambagiosus* populations, then different morphological features that might be subject to different pressures (constraining selection and species cohesion). Alternatively these mechanisms might operate at different times or geographic scales. If evolution within a lineage is the result of constraining selection, rather than species cohesion, then stasis might not be found in all traits. Traits could respond independently within and among lineages and therefore fit different models of evolution and ultimately little to no differentiation between populations. Finding that evolutionary modes differ among morphological traits for the same set of samples would be inconsistent with the idea that gene flow is the sole factor preventing local divergence, and patterns of multiple models of

evolution have frequently been identified per species' lineage (Chiba 1996; Hopkins & Lidgard 2012; Hunt 2007).

Methods

Rationale

Morphometric analysis and covariation typically involve the application of multivariate methods to linear measurements. Geometric morphometric methods, however, allow the partitioning of shape and size components, preserving the main geometric properties of the specimens while generating a visual representation and determining shape variables that can be analysed statistically. Geometric approaches therefore allow much more information to be obtained and provide more biologically meaningful information. Using combinations of landmark and sliding semi-landmark (analysis of curves) data, geometric techniques have been successfully applied to the shell shape of extant *Placostylus* from New Caledonia (Dowle et al. 2015) and other molluscs (Crampton & Gale 2005; Haase & Misof 2009; Hills et al. 2012; Stankowski 2011) including comparisons of fossil and modern taxa (Aguirre et al. 2006). For each specimen, centroid size (estimating the specimen size) were obtained. Centroid size is the square root of the sum of squared distances of landmarks to their centroid (the average x and y coordinate points) of the landmark configuration (Bookstein 1991).

Material –shell specimens

The large land snails *Placostylus* are present in the western Pacific, New Zealand, Vanuatu, Fiji, Papua New Guinea, Solomon Islands, Lord Howe and New Caledonia. New Zealand has three extant species of which two (*P. ambagiousus, P. hongii*) form a monophyletic pair, sister to *Placostylus* from New Caledonia (Trewick et al. 2009). *Placostylus ambagiosus* and *P. hongii* are restricted to the northern third of North Island and near shore islands (Figure 5.2 map). Within this range extant populations are fragmented and restricted, but deposits of fossil shells suggest that past populations had larger distributions and were connected (Buckley et al. 2011; Sherley 1996). Differences in shell shape associated with tubercles around the

operculum and patchy distributions of these snails led to the description of many sub-species, including some known only from fossils (Powell 1979).

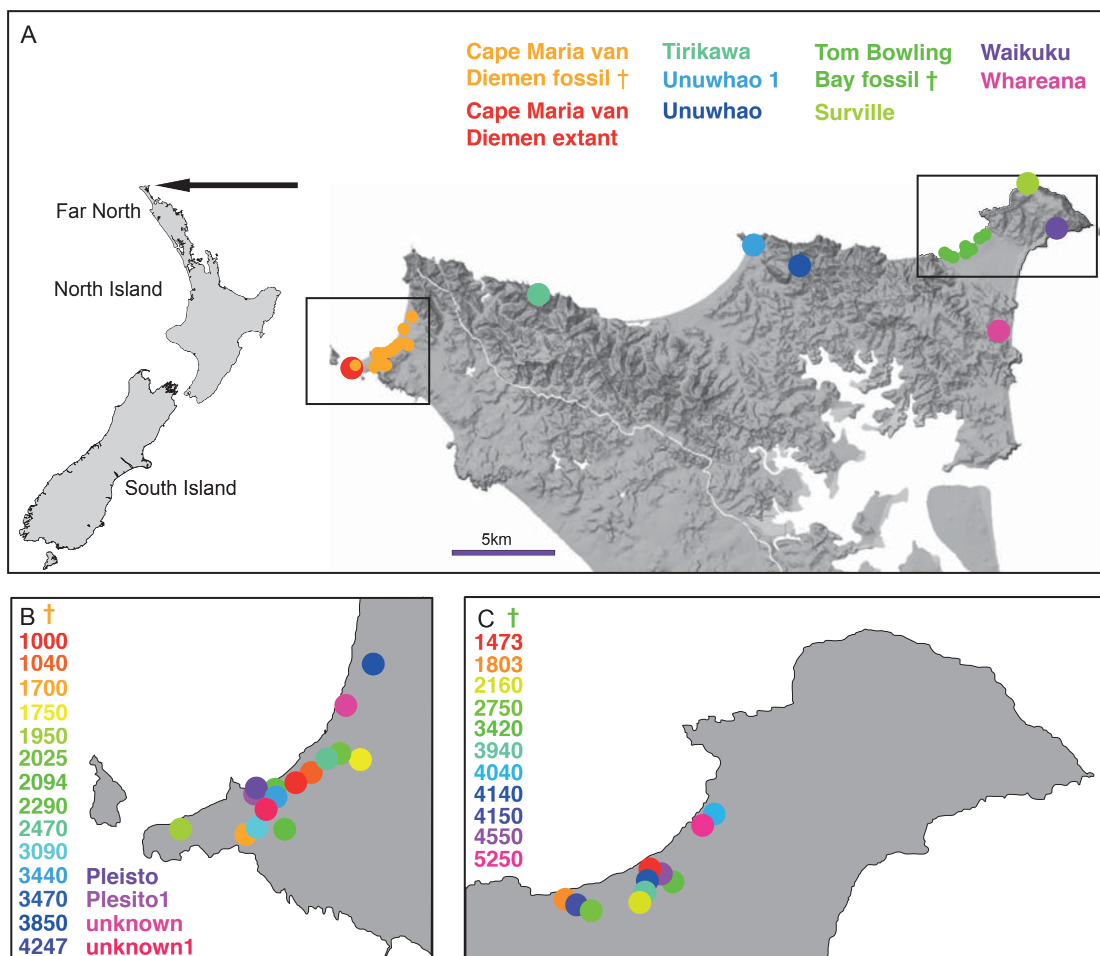


Figure 5.2. Approximate locations of extant and fossil population of the land snail *Placostylus ambagiosus* in the Te Paki district in Northland New Zealand. A. Extant and fossil populations; B. Fossil populations at Cape Maria van Diemen; C. Fossil populations in the vicinity of Tom Bowling Bay. Sites are colour coded with respect to their radiocarbon age in years bp (B &C).

In the present study 823 shells of *Placostylus ambagiosus* representing extant and extinct populations were examined. The material lodged in the National Museum of New Zealand Te Papa Tongarewa and the Auckland War Memorial Museum (Table 5.1). Radiometric dates have previously been obtained for most of the population samples using radio carbon isotope ratio analysis (Brook 1999, 2000; Millener 1981).

Table 5.1. *Placostylus ambagiosus* specimens used in this study. Radiocarbon ages as previously reported (Brook, 1999, 2000; Millener, 1981).

Location	Type	Museum	Accession	n	Conventional radiocarbon age (BP 1950)	Latitude	Longitude
Cape Maria van Diemen	extant	Te Papa	M.136625	39	-		
Pandora	extant	Te Papa	M.136627	21	-	-34.4445106	172.7480075
Surville Cliffs	extant	Te Papa	M.136621	29	-	-34.39618086	172.0165032
Unuwhao (Maungapiko)	extant	Te Papa	M.102324	32	-		
Unuwhao1 (Spirits Bay)	extant	Te Papa	M.102503	32	-	-34.43184351	172.8885088
Waikuku	extant	Te Papa	M.102328	26	-	-34.41517633	173.0265097
Whareana	extant	Te Papa	M.136614	30	-	-34.46117682	172.9951763
Cape Maria van Diemen	fossil	AUCKLAND	MA84863	20	1000±45YBP	-34.4695666	172.6679515
Cape Maria van Diemen	fossil	AUCKLAND	MA92194	20	1040+50YBP	-34.46845547	172.6701737
Cape Maria van Diemen	fossil	AUCKLAND	MA84853	20	1700+90YBP	-34.47651114	172.6593404
Cape Maria van Diemen	fossil	AUCKLAND	MA84849	10	1750±100YBP	-34.46678876	172.6776737
Cape Maria van Diemen	fossil	AUCKLAND	MA84865	20	1950±89YBP	-34.47567784	172.6493403
Cape Maria van Diemen	fossil	AUCKLAND	MA84862	20	2025±44YBP	-34.46595543	172.6746182
Cape Maria van Diemen	fossil	AUCKLAND	MA84850	20	2094±43YBP	-34.47567778	172.6657293
Cape Maria van Diemen	fossil	AUCKLAND	MA84851	20	2290±90YBP	-34.47039995	172.6646181
Cape Maria van Diemen	fossil	AUCKLAND	MA84854	20	2470±190YBP	-34.46678878	172.6723959
Cape Maria van Diemen	fossil	AUCKLAND	MA84858	20	3090±60YBP	-34.47567779	172.6612848
Cape Maria van Diemen	fossil	AUCKLAND	MA131756	20	3440±50YBP	-34.47123329	172.6646181
Cape Maria van Diemen	fossil	AUCKLAND	MA131758	20	3470±120YBP	-34.47567779	172.6612848
Cape Maria van Diemen	fossil	AUCKLAND	MA100013	20	3850±60YBP	-34.45401086	172.6798959
Cape Maria van Diemen	fossil	AUCKLAND	MA84852	20	4247±71YBP	-34.47567778	172.6657293
Cape Maria van Diemen	fossil	AUCKLAND	MA84860	18	>40,000YPB	-34.47123331	172.6612848
Cape Maria van Diemen	fossil	AUCKLAND	MA89524	24	Pleistocene	-34.47039997	172.6612848
Cape Maria van Diemen	fossil	AUCKLAND	MA131757	19	unknown	-34.45956648	172.6754515
Cape Maria van Diemen	fossil	Te Papa	M.183771	21	unknown	-34.47151108	172.6643404
Tom Bowling Bay	fossil	Te Papa	M.158307	40	1473±50YBP	-34.42350987	172.977676
Tom Bowling Bay	fossil	Te Papa	M.180610	25	1803±62YBP	-34.42617659	172.9665093
Tom Bowling Bay	fossil	Te Papa	M.318186	20	2160±50YBP		
Tom Bowling Bay	fossil	Te Papa	M.180400	22	2750±60YBP		
Tom Bowling Bay	fossil	Te Papa	M.180449	25	3420±67YBP	-34.4243432	172.9810094
Tom Bowling Bay	fossil	Te Papa	M.180421	21	3940±70YBP	-34.42517655	172.9776761
Tom Bowling Bay	fossil	Te Papa	M.180476	25	4040±70YBP	-34.41784312	172.9885094
Tom Bowling Bay	fossil	Te Papa	M.158309	40	4140±54YBP	-34.4243432	172.9788427
Tom Bowling Bay	fossil	Te Papa	M.180389	24	4150±60YBP	-34.42700993	172.967676
Tom Bowling Bay	fossil	Te Papa	M.180437	20	4550±70YBP	-34.4243432	172.9798427
Tom Bowling Bay	fossil	Te Papa	M.180465	20	5250±70YBP	-34.41984315	172.9851761

Paleoclimate of northland area, New Zealand

Despite global changes in temperature and aridity resulting in vegetation shifts throughout the Pleistocene, Northland New Zealand remained mostly forested (Alloway et al. 2007). Deforestation began with the arrival of people (Elliot et al. 1998; Newnham 1999; Newnham 1992). The climate of Northland in the Holocene (about eight thousand years ago) was warmer and moist until five thousand years ago when there was a change to a cooler and drier climate. At 3.4 thousand years ago signs of climate changes are inferred, with some evidence of forest disturbance from cyclones typical of warm moist and windy conditions with dry clear winters. The warmer period 3.4 thousand years ago correlates with a rise of 1.2m in sea level. There is evidence for increased cyclone activity at around 1.8 thousand years ago and climatic variability, with cool wet winters and warm dry summers (Alloway et al. 2007; Newnham 1999; Newnham 1992).

Shape Analysis

Intact adult snail shells from each sample population were bedded in sand of contrasting colour. A digital image of the ventral surface of each shell with a scale marker was captured using a high-resolution digital single-lens reflex Cannon EOS 600d camera with EF100mm f2.8 USM macro lens. Photographic equipment was mounted on a high-precision Kaiser stand to allow reproducible positioning and orientation. Up to 40 adult shells were photographed for each time period/location where available, with samples comprising of fewer than 10 shells discarded from analysis.

Digitising was under-taken in TPSDIG v2.17 (Rohlf 2013) on a Wacom Cintiq 22HD digitizing tablet. Two virtual combs used to locate semi-landmarks were first placed on images and aligned manually in Adobe Photoshop CS2 v9.0.2. Combs were aligned to a centreline running from the apex of the shell to the intercept of the aperture and body whorl (Figure 5.3). Six permanent and 33 semi-landmarks were placed around the aperture and shell outline when place in the ventral view (Figure 5.3). Digitised semi-landmarks were slid using SEMILAND, part of the IMP714 package (Sheets 2012; Zelditch et al. 2012) implementing the Procrustes distance method. Landmark X/Y coordinates were then imported into MORPHOJ 1.05F

(Klingenberg 2011). Sliding eliminates variation due to differences in translation, orientation and size. The Procrustes distance method is preferred when morphological variation is relatively small (Perez et al. 2006). Inclusion of semi-landmarks in a geometric morphometric analysis reduces the degrees of freedom in ways that are difficult to quantify, so statistical analyses based on such data employed nonparametric bootstrap resampling approaches (Hills et al. 2012; Zelditch et al. 2012).

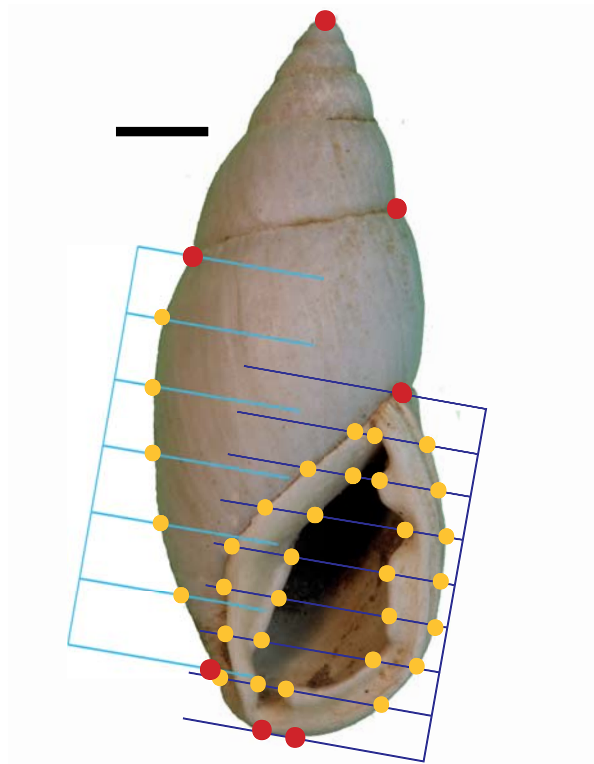


Figure 3: Fossil *Placostylus ambagiosus* shell showing orientation and placement of 7 landmarks (red) and 33 semi-landmarks (yellow) for 2D geometric morphometric analysis. Additional details in Chapter 4.

Measurement error was assessed using a specimen of *Placostylus ambagiosus* from Cape Maria van Diemen. This specimen was mounted, photographed and dismounted a total of five times. A landmark configuration was digitized from each of the independent replicate images providing a measure of the error associated with mounting specimens. One of these images was then digitised five separate times providing a measure of the error associated with the digitization process. Downstream analysis was as for all other datasets and disparity analysis was performed to calculate the error associated with photography and digitisation.

To display the major features of shape variation and as an ordination method a Principal Component Analysis (PCA) was performed using the covariance matrix generated from the dataset in the software MORPHOJ (Klingenberg 2011). Fossil populations from each of the two locations Cape Maria van Diemen and Tom Bowling Bay (Figure 5.2), were treated separately as previous analysis of extant populations indicate they harbour separate micro-lineages (Buckley et al. 2011). Principal component analysis of the three combined datasets was also performed. Resulting eigen values were used to perform broken-stick analyses calculating the number of principal omponents (PCs) that exceed the theoretical scree plot of random, uncorrelated noise using the package VEGAN (Dixon 2003) in R-studio (RStudio Team 2015). Resulting principal component scores were exported from MORPHOJ in comma separated value format. To identify the shape features that best distinguish among multiple groups of specimens (rather than among individuals as does PCA), the allometry-free variables were used as input for a canonical variance analysis (CVA) was performed in MORPHOJ. CVA determines whether the predefined populations can be statistically distinguished based on the relative warp matrix and was used to display the major features of shape variation between populations. Variation of key principal components was visualised using MORPHOJ. Variation within and between modern populations was compared to that of fossil populations using box and whisker plots created in R. Plots were produced for each of the first three principal component scores and centroid size.

Bayesian Clustering

Statistically discrete clusters of morphological variation were identified in each of three data sets (extant, Cape Maria van Diemen fossil and Tom Bowling Bay fossil) using the package MCLUST v4.0 (Fraley et al. 2012; Fraley & Raftery 2003). This treats the total data set as a mixture of multivariate normal data sets, with a selection of covariance structures and vectors of expectation (Fraley & Raftery 1999). The MCLUST analysis allows for a total of 90 models to be examined (10 different models with various combinations of parameterization and 1 to 9 clusters/components). MCLUST uses Bayesian Information Criterion (BIC) for parameter estimation via the EM algorithm for normal mixture modelling to cluster samples. BIC is a model selection tool that uses negative log-likelihood to maximise

statistical goodness of fit and parsimony by penalising models with more parameters, allowing it to filter out overly complicated models. Assignment is done without the specification of priors and so generates unbiased estimates of observation clusters. The optimal model was then selected based on BIC, which is the value of the maximized negative log likelihood, with a penalty for the number of parameters in the model. The model with the BIC score closest to zero was selected as the most likely and individual shells assigned to each cluster. Analysis of Principal Component scores that exceeded the theoretical scree plot of random, uncorrelated noise were used to assign clusters, without priors.

Analysis of fossil populations

Three models of morphological evolution were fitted to morphometric geometric data obtained from spatial and temporal population samples. The models describe three plausible modes of change through time: directional change (Generalised Random Walk), non-directional change (Unbiased Random Walk), and stasis. Datasets consisted of principal component scores fitted using maximum likelihood in the package PaleoTS (Hunt 2011) in the statistical programming environment R (R Core Team 2013). Carbon dates provided confidence intervals for age estimates and for several samples these overlap (no stratigraphic information is available for these samples to inform on the relative ages). Therefore data for shells from each of the fossil sampling areas where carbon dates had overlapping confidence intervals were reordered. The fit of the three models were assessed for each reordering of the data.

PaleoTS (Hunt 2011) assesses model support using the bias-corrected Akaike Information Criterion (AICc). AICc balances goodness of fit (log-liklihood) with model complexity (the number of model parameters). The model with the lowest AICc value is best supported, with the degree of support quantified by Akaike weights. Akaike weights are a simple transformation of AICc values that sum to one and are interpreted as approximate probabilities that each model is the best of those considered. “An important advantage of this model-based approach is that no evolutionary mode is granted privileged null status; all models are compared on an equal footing based on their empirical support” (Hunt 2007). Data for shells from

each of the fossil sampling areas (Cape Maria can Diemen and Tom Bowling Bay) where carbon dates had overlapping confidence intervals reordered within the time series for those samples because no stratigraphic information was available to inform on relative ages. The three models were fitted for each reordering of the data.

Results

Eight hundred and twenty three shells representing seven modern populations, eighteen fossil population from Cape Maria van Diemen spanning ~ 40,000 years and 11 fossil populations from Tom Bowling Bay spanning >5000 years were analysed. Photographing and digitising error was found to be negligible and well within the population based variation (Chapter 4).

Principal Component Analysis

The first six principal components were found to exceed the theoretical scree plot of random, uncorrelated noise calculated using brokenstick analysis (Figure 5.4).

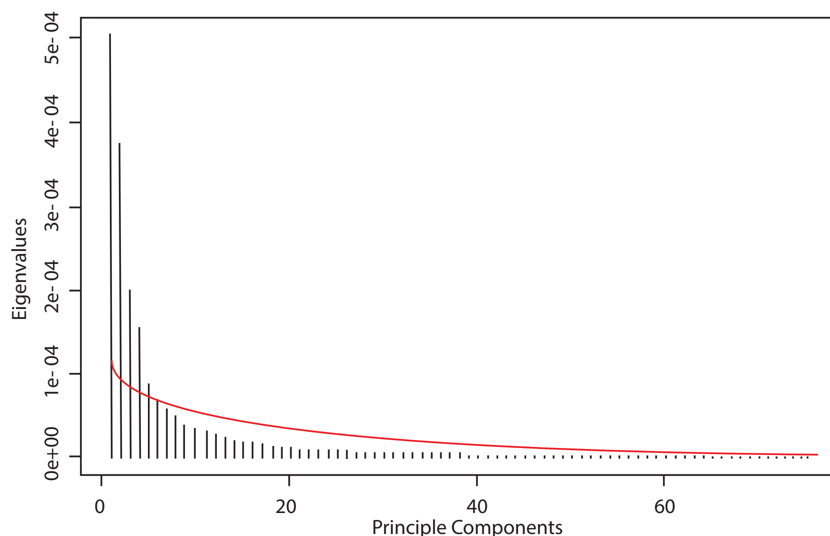


Figure 5.4. Eigenvalues from principal component analysis of the dataset including all *Placostylus ambagiousus* samples from Cape Maria van Diemen and Tom Bowling Bay (black lines) and a Brokenstick model for 76 components. The black bars that exceeded red curve line represent the principal components that exceed the theoretical scree plot of random, uncorrelated noise calculated.

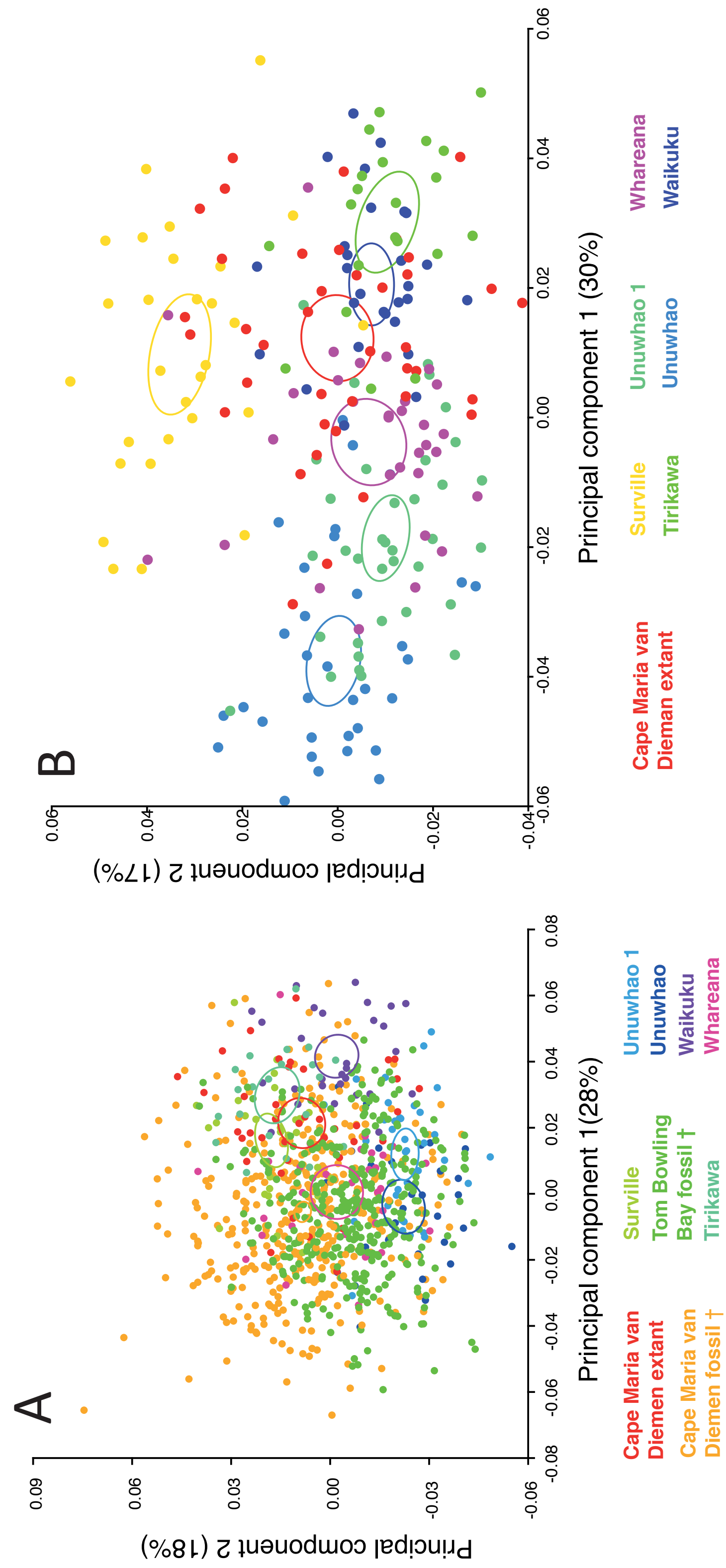
Principal components one to six accounted for 76.4% of the variation seen among individuals across all populations (Table 5.2). Pairwise scatterplots of principal components one and two revealed some clustering of population samples but as expected of a single species, population samples overlapped in morphospace (Figure 5.5). Extant and fossil populations cannot be separated using the first two principal components (Figure 5.5 A). When only extant snail population samples were examined for PC 1 and PC 2, three of the seven location samples had distributions with little overlap (Figure 5.5 B). In contrast, at each of the two fossil sites, population samples separated by time had similar shell shapes (Figure 5.5 C & D). When individuals were assigned to their *a priori* populations, canonical variates one and two separated the majority of specimens to either of the two fossil lineages (Figure 5.6) and each of the modern sample localities fell into an adjacent but distinct cluster.

Table 5.2. *Placostylus ambagiosus* from extant and all populations. Eigenvalues, % variance and cumulative variance for the principal components that exceed the theoretical scree plot of random, uncorrelated noise calculated from Principal Component Analysis of the dataset including all samples from Cape Maria van Diemen and Tom Bowling Bay.

Principle Component	Eigenvalue	% variance	Cumluative variance
1	0.00050343	27.80	27.80
2	0.00037373	20.64	48.44
3	0.00019836	10.96	59.40
4	0.00015436	8.53	67.93
5	0.00008712	4.81	72.74
6	0.00006693	3.70	76.44

Analysis of modern populations

Modern populations showed less variability for all traits than the pooled fossil populations (Figure 5.7). Many of the modern population samples had distinctive shell shapes exhibited by all individuals from a location (Figure 5.6). Shells from Surville Cliffs had the smallest average centroid size, although the individual shell with the smallest centroid comes from the Cape Maria van Diemen fossil population. Modern Cape Maria van Diemen shells also had a relatively small mean centroid size. The population from Unuwhao 1 comprised shells with the largest mean centroid size encountered. Shell centroids from Whareana and Waikuku are almost indistinguishable.


A
Principal component 2 (18%)
Principal component 1(28%)
Cape Maria van Diemen extant
Cape Maria van Diemen fossil †
Surville
Tom Bowling Bay fossil †
Tirikawa
Unuwhao 1
Unuwhao
Waikuku
Whareana
B
Principal component 2 (17%)
Principal component 1 (30%)
Cape Maria van Dieman extant
Surville
Tirikawa
Unuwhao 1
Unuwhao
Whareana
Waikuku

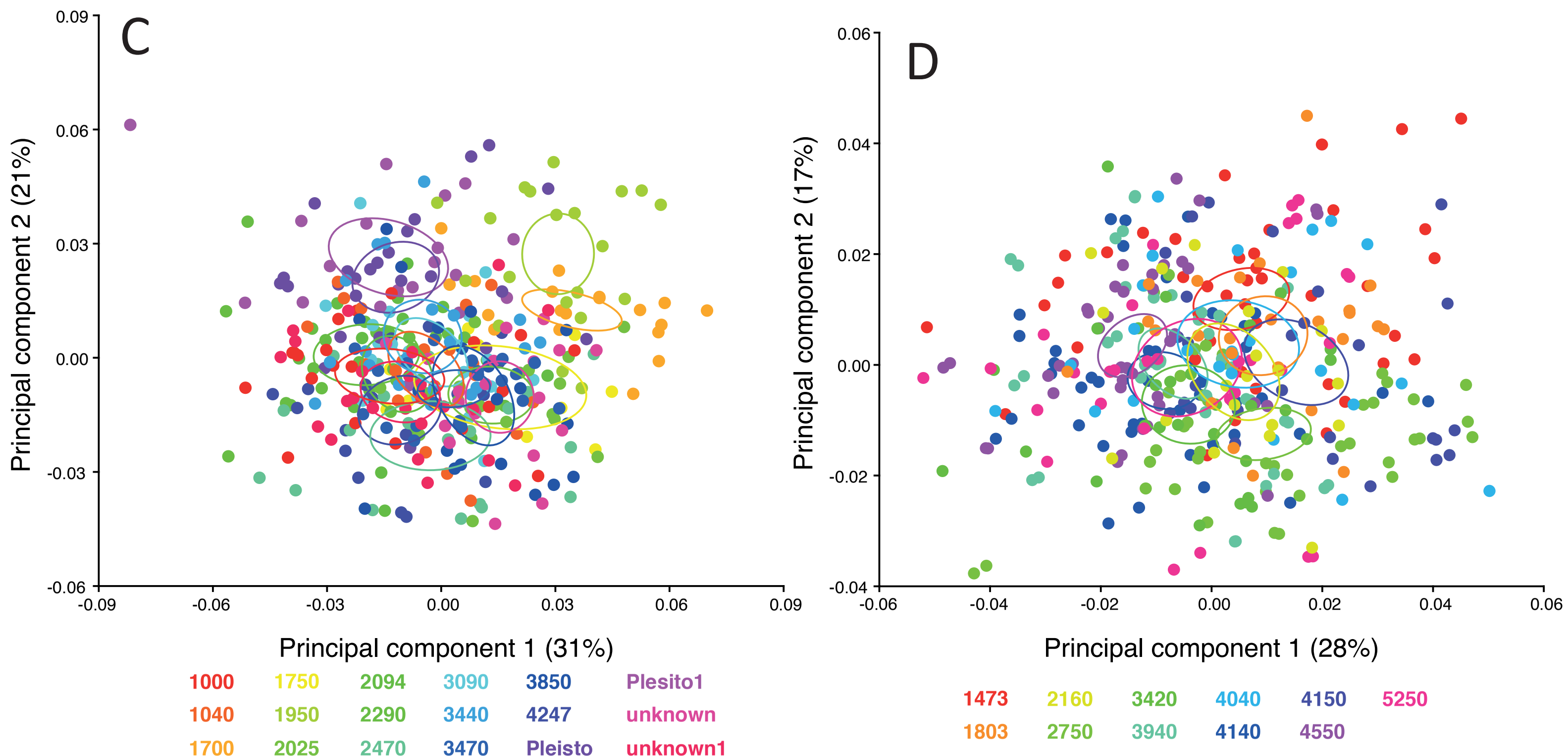
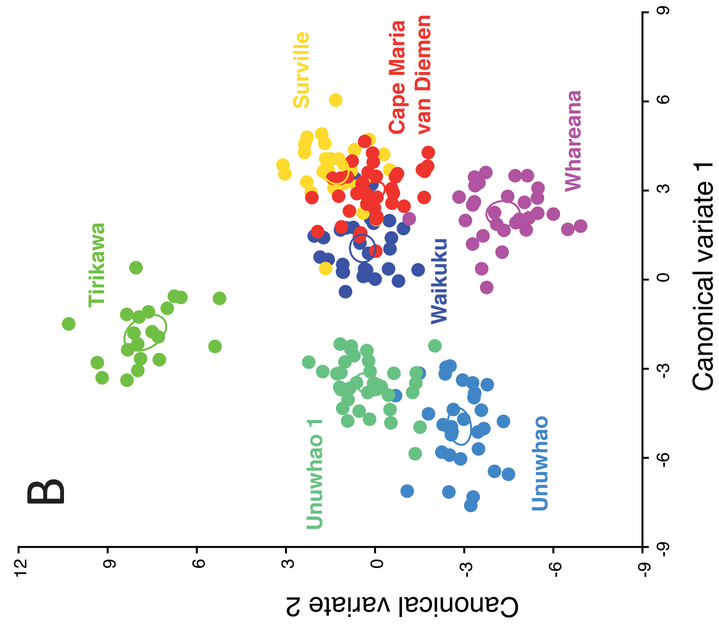
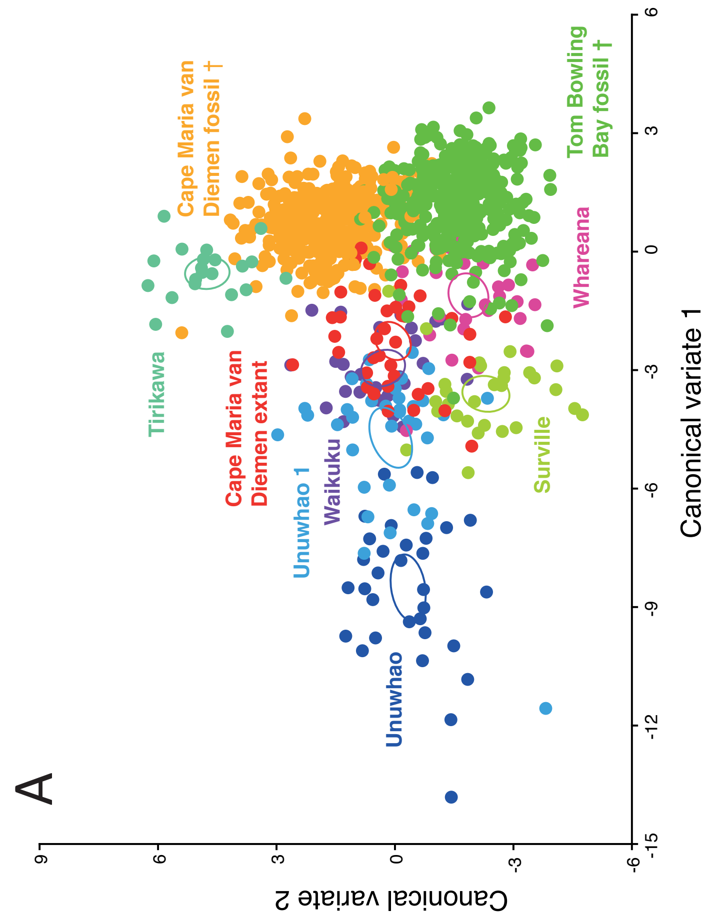


Figure 5.5. *Placostylus ambagiosus* shell shape variation from Principal Component Analyses using all permanent and semi-landmarks. A) All samples. B) Extant populations only. C) Fossil populations from Cape Maria van Diemen. D) Fossil populations from Tom Bowling Bay. Ellipses show 95% confidence intervals of the mean. Graphs are colour coded with respect to location or radiocarbon age in years bp. Colours match those in Figure 2 Map.


B
Tirikawa
Surville
Cape Maria
van Diemen
Whareana
Waikuku
Unuwhao 1
Unuwhao
Canonical variate 2
Canonical variate 1


A
Tirikawa
Cape Maria van
Diemen fossil †
Cape Maria van
Diemen extant
Unuwhao 1
Waikuku
Unuwhao
Surville
Whareana
Tom Bowling
Bay fossil †
Canonical variate 2
Canonical variate 1

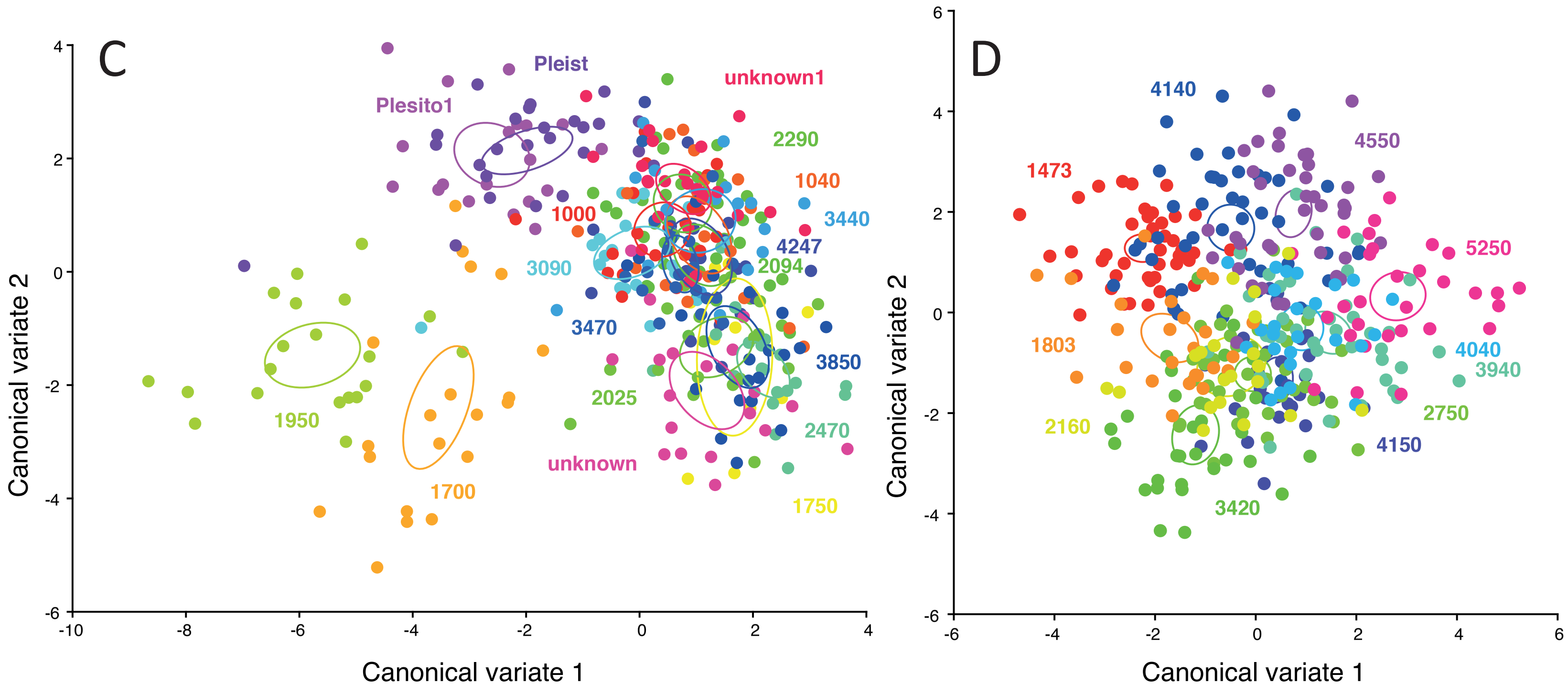


Figure 5.6. Canonical Variance Analyses of *Placostylus ambagiosus* shell shape variation using data that includes all permanent and semi-landmarks. A) All sampling. B) Extant populations only. C) Fossil populations from Cape Maria van Diemen. D) Fossil populations from Tom Bowling Bay. Ellipses indicate 95% confidence intervals of the mean. Graphs are colour coded with respect to location or radiocarbon age in years bp. Colours match those in Figure 2 Map.

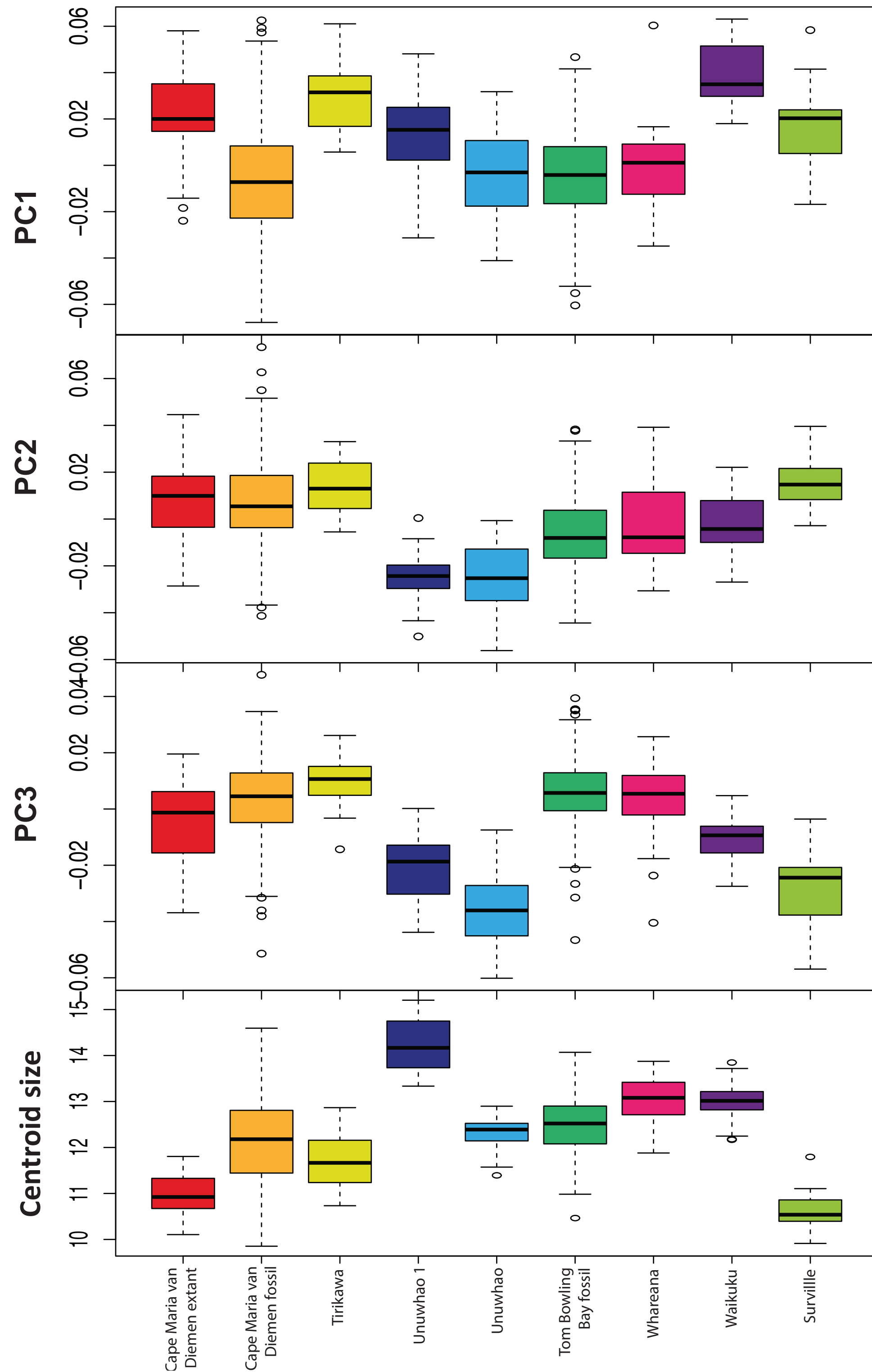


Figure 5.7. Shape and size variation among *Placostylus ambagiousus* shells from nine sampling locations. Boxplots for the first three Principal component scores and centroid size for all modern and fossil populations. Populations are ordered west to east.

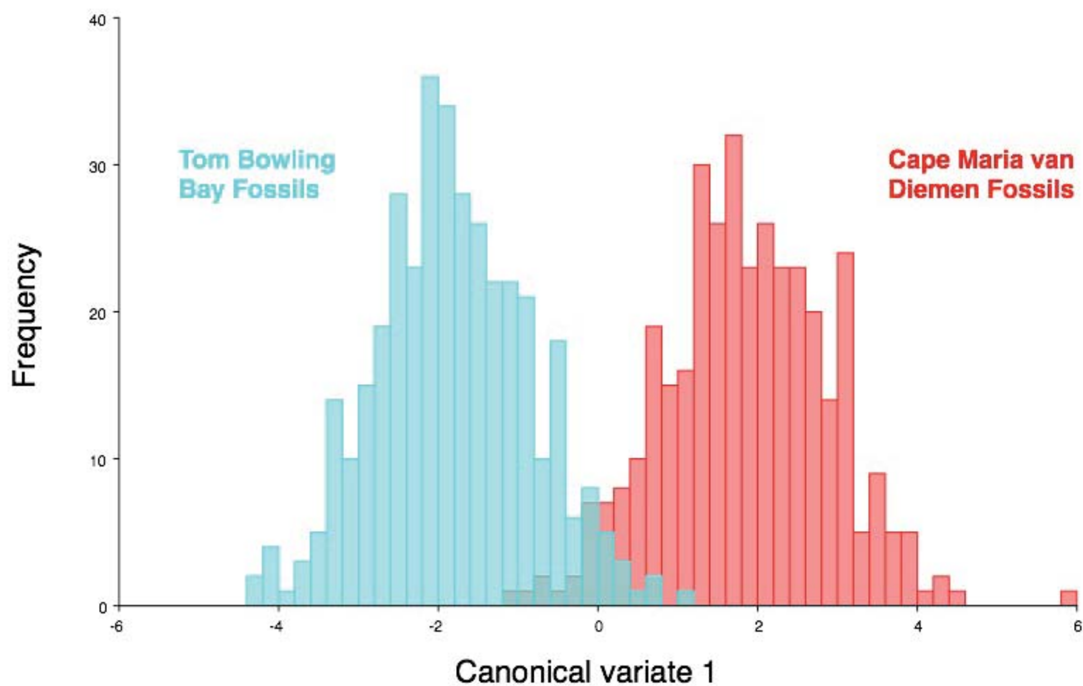


Figure 5.8. Canonical variate analysis of fossil shell populations of *Placostylus ambagiosus* shows a binomial distribution of sample site morphology for all temporal populations.

Bayesian Clustering

When geometric morphometric data of all extant population samples (Appendix 5.1) was analysed without priors in a Bayesian model-based clustering analysis (MCLUST) the optimal model identified three distinct population clusters (with PCs 1–3): [Surville] [CMvD, Tirikawa, Whareana] and [Unuwhao, Unuwhao1, Waikuku]. Increasing the number of principal components used in the analysis tended to increase the number of clusters identified. PC 1–4 identified four clusters with Unuwhao forming a separate cluster (others remained the same), PC 1–5 identified five clusters with individuals from Unuwhao and Tirikawa each forming separate clusters. PC 1–6 identified the same clusters as PC 1–5 but yielded slightly better assignment probabilities.

When the same approach was applied to data for the time-series fossil from Cape Maria van Diemen (Appendix 1), a single cluster was identified with PCs 1–3 and 1–4, however two clusters were identified using PCs 1–5 and 1–6. The second cluster for PCs 1–5 consisted of individuals from eight out of eighteen (44%) population

samples and between one and nine (10% to 45%) of the individuals within those populations being assigned to that cluster. For PCs 1–6 only three of the eighteen populations (17%) had any individuals assigned to the second cluster. Of these populations between three and six individuals (15% to 30%) were assigned to that cluster. Tom Bowling Bay fossil populations samples showed a single morphological cluster was identified by all iterations of varying the total PCs included.

Models of morphological evolution

There was overwhelming support for stasis in the shape of *Placostylus ambagiosus* shells from Cape Maria van Diemen for all Principal Components of shape and also for centroid size (Figure 5.9 C & D); Appendix 5.1). Support for stasis was not affected by reordering of temporal samples that have overlapping radio carbon age estimates. Running the model test again without Pleistocene samples and/or without the extant population from Cape Maria van Diemen also had little effect on the preferred model: stasis. In contrast, although stasis was the preferred model of evolution of the shell shape of in the Tom Bowling Bay snails (Figure 5.9 A; Supplementary material Table 5.2) for six Principal Components of shape, centroid size fitted the Unbiased Random Walk (URW) (Figure 5.9 B; Supplementary material Table 5.2) model better than other models. The ordering of the Tom Bowling Bay samples did influence the preferred evolutionary model (Appendix 5.2).

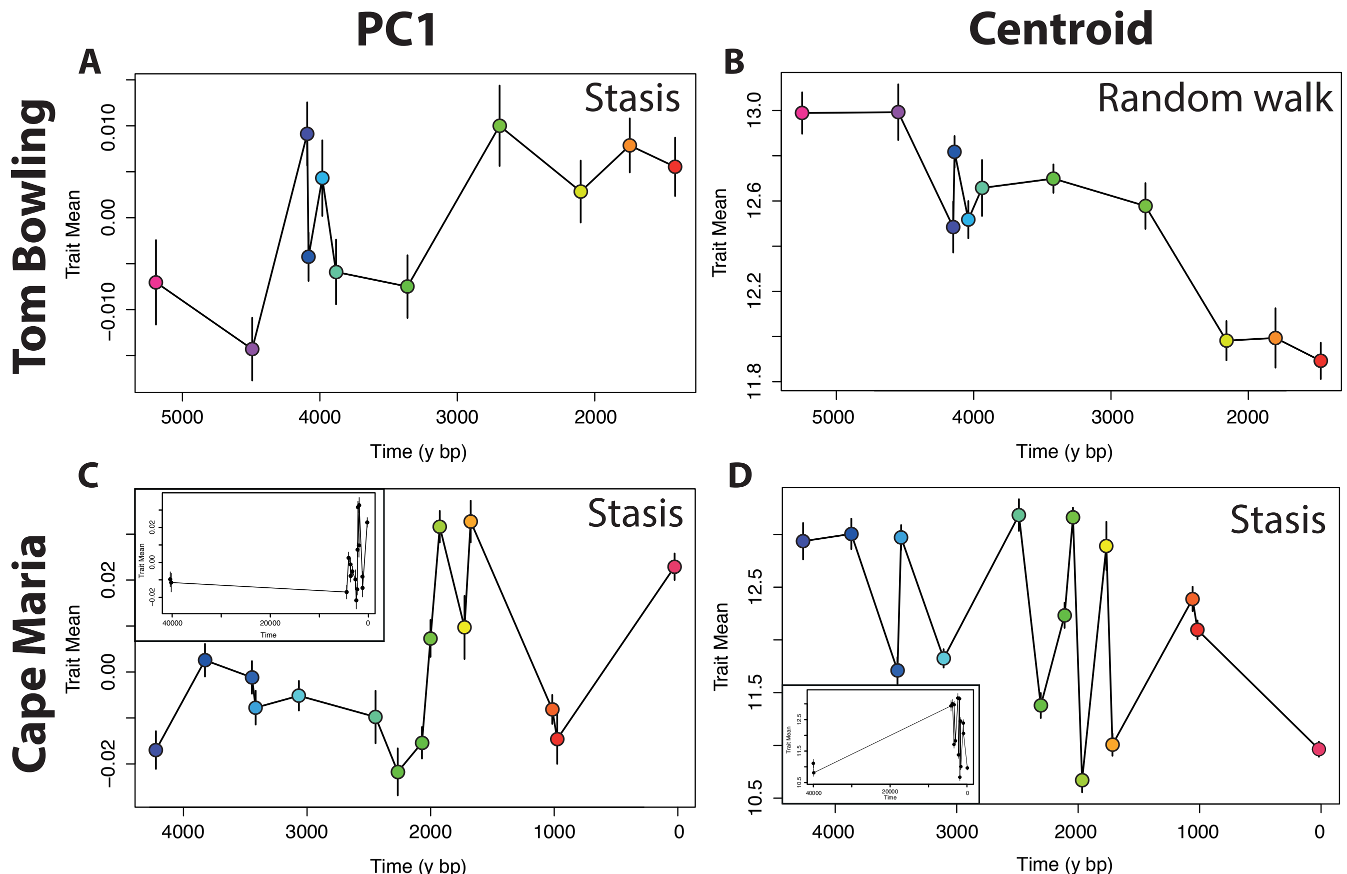


Figure 5.9. Trait means for fossil populations of *Placostylus ambagiosus* from Tom Bowling Bay and Cape Maria van Diemen. Principal Component 1 shows strong support for Stasis at Tom Bowling Bay (A) and Cape Maria van Diemen (C). Centroid size finds support for the Random Walk model for the fossil populations at Tom Bowling Bay (B) and Stasis for fossil populations at Cape Maria van Diemen (D). Insets of C & D show patterns when Pleistocene samples are included in analysis. Data points are colour coded to match Figure 2, 5 and 6.

Discussion

The potential for more than one population/lineage of the same species to be traced through time with sampling that allows estimates of population phenotypic variance. This unusual situation allows us to assess the patterns and processes within and between lineages to help understand the role of gene flow in uniting populations (and influencing model of morphological evolution).

Modern populations sampled geographically throughout the range of *Placostylus ambagiosus* show shell shape differences, but analysis of extinct and extant populations, together reduces discrete clusters. One needs to study datasets constrained by either space or time in order to perceive the significant differentiation that exists. The geometric morphometric shape variation for extant populations sampled within 100 km^2 of each other was best described by a model with five distinct clusters, correlating with distinct geographic regions. Thus six principal components of shell shape variation could distinguish some (but not all) of the differences used by the taxonomist A.W.B. Powell 70 years ago to describe 16 subspecies (Powell 1947, 1951). The degree of morphological differentiation within a single species suggests that local environments differ in some influential manner, within the very narrow range of this species. Although vegetation type and moisture level within local sites has previously been studied in relation to *Placostylus amabgious* snails, the information used was insufficient to discriminate shape differences (and potential environmental effects) (Buckley et al. 2011; Penniket 1981). Higher moisture levels and/or more sheltered sites were found to be occupied by larger shelled snails compared to dry or coastal sites (Buckley et al. 2011; Penniket 1981 Powell 1947, 1951). In this study no relationship between distance from coast and shell centroid was observed (results not shown). Shell shape variation described here using geometric morphometrics is independent of shell size, and there was little support for influence from moisture levels or vegetation type as currently understood (Penniket 1981). It appears likely that shell shape is influenced by a number of factors yet to be identified. However local differentiation strongly suggest local selection has resulted in populations adapting to local environments.

Just as one can distinguish differences in shell shape between populations that are geographically isolated today there are also differences in morphology between Cape Maria van Diemen and Tom Bowling Bay material. These differences in morphology are relatively unchanged through time with different traits in the same lineage exhibiting the same mode of evolution. This was especially evident at Cape Maria van Diemen, where we find overwhelming evidence for stasis among fossils dating back 40 k years.

The snails in the modern population at Cape Maria van Diemen were the most frequently incorrectly assigned to clusters, often falling almost equally into two clusters. In fact two subspecies of *Placostylus ambagiosus* were described from the Cape Maria van Diemen peninsula (*P. a. ambagiosus*: Powell 1947, p.182 and *P. a. paraspiritus*: Powell 1951, p.137 with a further subspecies on the detached headland Motuopao Island *P. a. consobrinus*: Powell 1938, p. 148 that are all now synonymised (Buckley et al. 2011)) and this may reflect collection of two morphs (type localities separated by 0.5-1.2km) from a single or nearby location(s). Interestingly the fossil *Placostylus ambagiosus* at Cape Maria van Diemen were not considered synonymous with any of the extant snail populations (Powell 1951b). There has also been speculation that Māori may have established new *Placostylus* populations (Buckley et al. 2011). As extant populations of *Placostylus ambagiosus* at Cape Maria van Diemen are confined to old Pā sites (Māori settlements protected by fortifications) (Triggs & Sherley 1993) it is possible that human mediated movement explains the lack of continuity between fossil and extant populations and the variation within the extant populations.

In contrast to the snail population shape variation seen among extant shells morphometric variation within locations sampled temporally was distributed around a constant optima (steady mean). All six independent shape traits conformed to a model of evolutionary stasis through 40 thousand years at Cape Maria van Diemen, and through 5 thousand years at Tom Bowling Bay. The only trait not following a model of stasis was centroid size in the Tom Bowling Bay lineage where snail shell centroid size decreased over time. Simulated evolutionary

sequences under a simple temperature tracking model have been found to provide a surprisingly close match to the empirical distribution of stasis (Hunt et al. 2015). Although this could be considered inconsistent with neo-biologists interpretations of stasis that emphasise stabilising selection and niche tracking, it appears stasis can be an expected outcome even in the face of dynamic environmental change. Hunt et al. (2015) suggests two explanations to make sense of this result. First: stasis implies constancy of form with changes to morphology modest and fluctuating rather than accumulating, therefore morphological response to changing environment could be consistent with stasis so long as the environment is not changing in an accumulating manner (i.e. it is tracked non-directional). Second: Climate often changes rapidly over short periods but maintains a mean state over the longer duration of a species. The recent plaeoenvironments experience dramatic climatic swings but strongly bounded by glacial-interglacial limits. Accordingly most traits that track temperature (and associated variables) will also show the bounded, fluctuating pattern that we recognise as stasis i.e. variation within limits. However, at Tom Bowling Bay, modelling with centroid size showed support for non-directional change (Unbiased Random Walk). Size of shells over the same time period, at Cape Maria van Diemen evolved via stasis. The fact that independent shell traits are following different models from one another in the same geographic location, and that some of these patterns differ from those inferred for a population of the same species 30 km away provides compelling evidence that gene flow is not limiting all local adaptation.

Extant populations are not uniform so local adaptation is possible in the timeframe we can observe as extant. However local populations although different, show stasis in their shell shape. Thus one can infer that constraining selection is likely to be the mechanism resulting in morphological stasis for some traits over this time scale. However, species cohesion could explain longer term inference of stasis in other taxa. My findings emphasise the importance of biological systems where time and location can be isolated for analysis.

References

Aguirre, M. L., Perez, S. I., & Sirch, Y. N. (2006). Morphological variability of Brachidontes Swainson (Bivalvia, Mytilidae) in the marine Quaternary of Argentina (SW Atlantic). *Palaeogeography, Palaeoclimatology, Palaeoecology*, *239*(1–2), 100–125.

Alloway, B. V., Lowe, D. J., Barrell, D. J. A., Newnham, R. M., Almond, P. C., Augustinus, P. C., … Williams, P. W. (2007). Towards a climate event stratigraphy for New Zealand over the past 30 000 years (NZ-INTIMATE project). *Journal of Quaternary Science*, *22*(1), 9–35.

Bocxlaer, B. V., & Hunt, G. (2013). Morphological stasis in an ongoing gastropod radiation from Lake Malawi. *Proceedings of the National Academy of Sciences*, *110*(34), 13892–13897.

Brett, C. E., Hendy, A. J. W., Bartholomew, A. J., Bonelli, J. R., & Mclaughlin, P. I. (2007). Response of Shallow Marine Biotas to Sea-Level Fluctuations: A Review of Faunal Replacement and the Process of Habitat Tracking. *PALAIOS*, *22*(3), 228–244.

Bridle, J. R., & Vines, T. H. (2007). Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution*, *22*(3), 140–147.

Brook, F. J. (1999). Stratigraphy, landsnail faunas, and paleoenvironmental history of coastal dunefields at Te Werahi, northernmost New Zealand. *Journal of the Royal Society of New Zealand*, *29*(4), 361–393.

Brook, F. J. (2000). Prehistoric predation of the landsnail *Placostylus ambagiosus* Suter (Stylommatophora: Bulimulidae), and evidence for the timing of establishment of rats in northernmost New Zealand. *Journal of the Royal Society of New Zealand*, *30*(3), 227–241.

Buckley, T., Stringer, I., Gleeson, D., Howitt, R., Attanayake, D., Parrish, R., … Rohan, M. (2011). A revision of the New *Placostylus* land snails using mitochondrial DNA and shell morphometric analyses, with implications for conservation. *New Zealand Journal of Zoology*, *38*(1), 55–81.

Charlesworth, B., Lande, R., & Slatkin, M. (1982). A Neo-Darwinian Commentary on Macroevolution. *Evolution*, *36*(3), 474–498.

Chiba, S. (1996). A 40,000-Year Record of Discontinuous Evolution of Island Snails. *Paleobiology*, *22*(2), 177–188.

Crampton, J. S., & Gale, A. S. (2005). A plastic boomerang: speciation and intraspecific evolution in the Cretaceous bivalve Actinoceramus. *Paleobiology*, *31*(4), 559–577.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927–930.

Dowle, E. J., Morgan-Richards, M., Brescia, F., & Trewick, S. A. (2015). Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. *Molecular Ecology*, *24*(16), 4205–4221.

Ehrlich, P. R., & Raven, P. H. (1969). Differentiation of populations. *Science*, *165*, 1228–1232.

Eldredge, N., Thompson, J. N., Brakefield, P. M., Gavrilets, S., Jablonski, D., Jackson, J. B. C., … Miller, W. (2005). The dynamics of evolutionary stasis. *Paleobiology*, *31*(sp5), 133–145.

Elliot, M. B., Flenley, J. R., & Sutton, D. G. (1998). A late Holocene pollen record of deforestation and environmental change from the Lake Tauanui catchment, Northland, New Zealand. *Journal of Paleolimnology*, *19*(1), 23–32.

Estes, S., & Arnold, S. J. (2007). Resolving the Paradox of Stasis: Models with Stabilizing Selection Explain Evolutionary Divergence on All Timescales. *The American Naturalist*, *169*(2), 227–244.

Fraley, C., Raferty, A., Murphy, T. B., & Scrucca, L. (2012). *mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation* (Technical Report No. 597). Department of Statistics: University of Washington.

Fraley, C., & Raftery, A. E. (1999). MCLUST: Software for Model-Based Cluster Analysis. *Journal of Classification*, *16*(2), 297–306.

Fraley, C., & Raftery, A. E. (2003). Enhanced Model-Based Clustering, Density Estimation, and Discriminant Analysis Software: MCLUST. *Journal of Classification*, *20*(2), 263–286.

Haase, M., & Misof, B. (2009). Dynamic gastropods: stable shell polymorphism despite gene flow in the land snail *Arianta arbustorum*. *Journal of Zoological Systematics and Evolutionary Research*, *47*(2), 105–114.

Hansen, T. F., & Houle, D. (2004). Evolvability, Stabilizing Selection, and the Problem of Stasis. In M. P. P. of E. and Evolution & K. P. D. of B. S. S. University (Eds.), *Phenotypic Integration : Studying the Ecology and Evolution of Complex Phenotypes: Studying the Ecology and Evolution of Complex Phenotypes*. Oxford University Press, USA.

Hansen, T. F., Pélabon, C., & Houle, D. (2011). Heritability is not Evolvability. *Evolutionary Biology*, *38*(3), 258.

Hendry, A. P., & Kinnison, M. T. (1999). Perspective: The Pace of Modern Life: Measuring Rates of Contemporary Microevolution. *Evolution*, *53*(6), 1637–1653.

Hereford, J., Hansen, T. F., & Houle, D. (2004). Comparing Strengths of Directional Selection: How Strong Is Strong? *Evolution*, *58*(10), 2133–2143.

Hills, S. F. K., Crampton, J. S., Trewick, S. A., & Morgan-Richards, M. (2012). DNA and Morphology Unite Two Species and 10 Million Year Old Fossils. *PLOS ONE*, *7*(12), e52083.

Hopkins, M. J., & Lidgard, S. (2012). Evolutionary mode routinely varies among morphological traits within fossil species lineages. *Proceedings of the National Academy of Sciences*, *109*(50), 20520–20525.

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, *130*(1), 195–204.

Houle, D. (1998). How should we explain variation in the genetic variance of traits? *Genetica*, *102–103*(0), 241.

Hunt, G. (2006). Fitting and comparing models of phyletic evolution: random walks and beyond. *Paleobiology*, *32*(4), 578–601. https://doi.org/10.1666/05070.1

Hunt, G. (2007). The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. *Proceedings of the National Academy of Sciences*, *104*(47), 18404–18408.

Hunt, G. (2011). *paleoTS: analyze paleontological time-series. R package version 0.4–1.* Retrieved from http://CRAN.R-project.org/package=paleoTS

Hunt, G., Bell, M. A., & Travis, M. P. (2008). Evolution Toward a New Adaptive Optimum: Phenotypic Evolution in a Fossil Stickleback Lineage. *Evolution*, *62*(3), 700–710.

Hunt, G., Hopkins, M. J., & Lidgard, S. (2015). Simple versus complex models of trait evolution and stasis as a response to environmental change. *Proceedings of the National Academy of Sciences*, *112*(16), 4885–4890.

Hunt, G., & Rabosky, D. L. (2014). Phenotypic Evolution in Fossil Species: Pattern and Process. *Annual Review of Earth and Planetary Sciences*, *42*(1), 421–441.

Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hill, C. E., … Beerli, P. (2001). The Strength of Phenotypic Selection in Natural Populations. *The American Naturalist*, *157*(3), 245–261.

Kinnison, M. T., & Hendry, A. P. (2001). The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica*, *112–113*(1), 145–164.

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a Species' Range. *The American Naturalist*, *150*(1), 1–23.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*(2), 353–357.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183–189.

Merilä, J., Sheldon, B. C., & Kruuk, L. E. B. (2001). Explaining stasis: microevolutionary studies in natural populations. *Genetica*, *112–113*(1), 199–222.

Millener, P. R. (1981). *The Quaternary avifauna of the North Island, New Zealand.* (Doctoral dissertation, Geology). University of Auckland, New Zealand.

Newnham, R. (1999). Environmental change in Northland, New Zealand during the last glacial and Holocene. *Quaternary International*, *57–58*, 61–70.

Newnham, R. M. (1992). A 30,000 Year Pollen, Vegetation and Climate Record from Otakairangi (Hikurangi), Northland, New Zealand. *Journal of Biogeography*, *19*(5), 541–554.

Penniket, A. S. W. (1981). *Population studies of land snails of the genus* Placostylus *in the north of New Zealand.* (Unpublished MSc thesis). Auckland University, Auckland.

Perez, S. I., Bernal, V., & Gonzalez, P. N. (2006). Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy*, *208*(6), 769–784.

Powell, A. (1979). *New Zealand Mollusca: marine, land, and freshwater shells*. Collins.

Powell, A. W. B. (1947). Distribution of *Placostylus* Land Snails in Northernmost New Zealand on JSTOR. *Records of the Auckland Institute and Museum*, *3*(3), 173–188.

Powell, A. W. B. (1951). On Further Colonies of *Placostylus* Land Snails from Northernmost New Zealand. *Records of the Auckalnd Institute and Museum*, *4*(2), 134–140.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Rohlf, F. (2013). TpsDig2 (Version 2.17).

RStudio Team. (2015). *RStudio: Integrated Development for R.* Boston, MA: RStudio, Inc. Retrieved from https://www.rstudio.com/

Sheets, H. (2012). *IMP714 - Integrated Morphometrics Package.* Retrieved from http://www3canisius.edu/~sheets/imp714.htm

Sherley, G. (1996). Morphological variation in the shells of Placostylus species (Gastropoda: Bulimulidae) in New Zealand and implications for their conservation. *New Zealand Journal of Zoology*, *23*(1), 73–82.

Stankowski, S. (2011). Extreme, continuous variation in an island snail: local diversification and association of shell form with the current environment: EXTREME VARAITION IN AN ISLAND SNAIL. *Biological Journal of the Linnean Society*, *104*(4), 756–769.

Trewick, S., Brescia, F., & Jordan, C. (2009). Diversity and phylogeny of New Caledonian Placostylus land snails; evidence from mitochondrial DNA. In *Grandcolas P ed. Zoologia Neocaledonica 7* (pp. 421–436).

Triggs, S. J., & Sherley, G. H. (1993). Allozyme genetic diversity in Placostylus land snails and implications for conservation. *New Zealand Journal of Zoology*, *20*(1), 19–33.

Ubukata, T., Kitamura, A., Hiramoto, M., & Kase, T. (2009). A 5000-Year Fossil Record of Larval Shell Morphology of Submarine Cave Microshells. *Evolution*, *63*(1), 295–300. https://doi.org/10.1111/j.1558-5646.2008.00546.x

Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric Morphometrics for Biologists: A Primer*. Academic Press.

Appendix

Appendix 5.1 Summary table for PaleoTS results for all samples from Cape Maria van Diemen. The model with the highest Akaike Weight is the preferred model. Where the error on radiocarbon dates overlapped the model test was run iteratively for different scenarios (orders were reversed and averaged.

Cape Maria van Diemen Fossils and one extant population			with "Plesitocene" samples				without "Pleistoncene" samples				averaged (all overapppig times) with Plist			
		Model	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight
	PC1	GRW	37.79887	2	-70.67466	0.002	34.29437	2	-63.49783	0.025	24.05995	2	-42.6199	0.001
		URW	37.79535	1	-73.30499	0.007	34.23495	1	-66.13656	0.094	24.05289	1	-45.66135	0.002
		Stasis	**44.00693**	**2**	**-83.09078**	**0.991**	**37.85576**	**2**	**-70.62062**	**0.881**	**31.65201**	**2**	**-57.80402**	**0.997**
	PC2	GRW	40.15263	2	-75.38218	0.000	36.36576	2	-67.64061	0.000	**41.92169**	**2**	**-78.34337**	**0.556**
		URW	40.14265	1	-77.99958	0.001	36.36262	1	-70.39191	0.000	40.16964	1	-77.89484	0.444
		Stasis	**48.08794**	**2**	**-91.2528**	**0.998**	**46.54704**	**2**	**-88.00317**	**1.000**	33.79719	2	-62.09439	0.000
	PC3	GRW	42.87037	2	-80.81765	0.000	38.68457	2	-72.27823	0.000	36.3527	2	-67.20541	0.072
		URW	42.86744	1	-83.44916	0.000	38.68218	1	-75.03102	0.001	36.33476	1	-70.22507	0.326
		Stasis	**52.75677**	**2**	**-100.5905**	**1.000**	**47.17417**	**2**	**-89.25743**	**0.999**	**38.47709**	**2**	**-71.45419**	**0.602**
	PC4	GRW	56.26481	2	-107.6065	0.000	50.80442	2	-96.51794	0.000	39.669	2	-73.838	0.000
		URW	56.26454	1	-110.2434	0.000	50.80317	1	-99.27302	0.002	39.65965	1	-76.87486	0.001
		Stasis	**67.10112**	**2**	**-129.2792**	**1.000**	**58.63774**	**2**	**-112.1846**	**0.998**	**47.95906**	**2**	**-90.41813**	**0.999**
	PC5	GRW	48.87118	2	-92.81929	0.000	43.94989	2	-82.80888	0.000	47.78468	2	-90.06937	0.027
		URW	48.87014	1	-95.45457	0.000	43.94772	1	-85.56211	0.000	47.04743	1	-91.65042	0.060
		Stasis	**64.96146**	**2**	**-124.9999**	**1.000**	**56.66008**	**2**	**-108.2293**	**1.000**	**51.30076**	**2**	**-97.10152**	**0.913**
	PC6	GRW	54.21162	2	-103.5002	0.000	48.55029	2	-92.00967	0.000	44.51321	2	-83.52642	0.006
		URW	54.16804	1	-106.0504	0.000	48.50937	1	-94.68541	0.000	44.49395	1	-86.54346	0.025
		Stasis	**67.14418**	**2**	**-129.3653**	**1.000**	**59.18834**	**2**	**-113.2858**	**1.000**	**49.66337**	**2**	**-93.82674**	**0.969**
	Centroid Size	GRW	-32.28	2	69.55157	0.000	-26.93247	2	59.06494	0.000	-17.64109	2	40.78217	0.005
		URW	-32.27579	1	66.85928	0.000	-26.95726	1	56.27816	0.000	-17.64118	1	37.7268	0.022
		Stasis	**-19.19865**	**2**	**43.3973**	**1.000**	**-15.76329**	**2**	**36.72657**	**1.000**	**-12.33224**	**2**	**30.16448**	**0.973**

Appendix 5.1 cont.

Cape Maria van Diemen Fossils and one extant population		averaged all -no Pleist				reversed dates				reversed dates no pleist			
	Model	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight
PC1	GRW	20.36525	2	-34.7305	0.007	33.37852	2	-61.83397	0.000	30.51973	2	-55.94856	0.001
	URW	20.26678	1	-37.96212	0.036	33.37462	1	-64.46352	0.000	30.46967	1	-58.60601	0.002
	Stasis	**25.26885**	**2**	**-44.53771**	**0.957**	**44.00693**	**2**	**-83.09078**	**1.000**	**37.85576**	**2**	**-70.62062**	**0.997**
PC2	GRW	35.03905	2	-64.07809	0.158	39.42651	2	-73.92995	0.000	35.80602	2	-66.52114	0.000
	URW	**34.65498**	**1**	**-66.73853**	**0.596**	39.42609	1	-76.56646	0.001	35.74936	1	-69.16538	0.000
	Stasis	35.48306	2	-64.96611	0.246	**48.08794**	**2**	**-91.2528**	**0.999**	**46.54704**	**2**	**-88.00317**	**1.000**
PC3	GRW	30.52292	2	-55.04584	0.041	41.40257	2	-77.88205	0.000	37.4454	2	-69.79989	0.000
	URW	30.52039	1	-58.46936	0.229	41.40251	1	-80.51931	0.000	37.44218	1	-72.55103	0.000
	Stasis	**33.39271**	**2**	**-60.78541**	**0.729**	**52.75677**	**2**	**-100.5905**	**1.000**	**47.17417**	**2**	**-89.25743**	**1.000**
PC4	GRW	34.15968	2	-62.31937	0.005	57.5505	2	-110.1779	0.000	52.35335	2	-99.61579	0.002
	URW	34.03173	1	-65.49203	0.023	57.52254	1	-112.7594	0.000	51.96199	1	-101.5907	0.005
	Stasis	**39.50131**	**2**	**-73.00262**	**0.973**	**67.10112**	**2**	**-129.2792**	**1.000**	**58.63774**	**2**	**-112.1846**	**0.993**
PC5	GRW	39.24734	2	-72.49468	0.015	49.26474	2	-93.6064	0.000	44.27187	2	-83.45283	0.000
	URW	39.24236	1	-75.91329	0.081	49.26453	1	-96.24334	0.000	44.26697	1	-86.2006	0.000
	Stasis	**43.37572**	**2**	**-80.75144**	**0.905**	**64.96146**	**2**	**-124.9999**	**1.000**	**56.66008**	**2**	**-108.2293**	**1.000**
PC6	GRW	37.60822	2	-69.21644	0.011	56.85091	2	-108.7787	0.000	50.74557	2	-96.40024	0.000
	URW	37.12781	1	-71.68418	0.038	56.84422	1	-111.4027	0.000	50.65248	1	-98.97163	0.001
	Stasis	**42.07389**	**2**	**-78.14778**	**0.951**	**67.14418**	**2**	**-129.3653**	**1.000**	**59.18834**	**2**	**-113.2858**	**0.999**
Centroid Size	GRW	-13.78501	2	33.57001	0.003	-30.32177	2	65.56661	0.000	-25.1134	2	55.31771	0.000
	URW	-13.91768	1	30.40679	0.013	-30.32178	1	62.92927	0.000	-25.15441	1	52.64215	0.001
	Stasis	**-7.902134**	**2**	**21.80427**	**0.984**	**-20.08516**	**2**	**45.09339**	**1.000**	**-16.5173**	**2**	**38.12551**	**0.999**

Appendix 5.2. Summary table for PaleoTS results for all samples from Tom Bowling Bay. The model with the highest Akaike Weight is the preferred model. Where the error on radiocarbon dates overlapped the model test was run iteratively for different scenarios (orders were reversed and averaged.

			carbon dated order (youngest -abc -oldest)				carbon dated order (acb)				carbon dated order (bca)				carbon dated order (bac)				carbon dated order (3940 and 4040swapped)			
		Model	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight
Tom Bowling Bay Fossils	PC1	GRW	28.259	2	-50.804	0.030	30.687	2	-55.660	0.025	31.992	2	-58.270	0.068	32.457	2	-59.200	0.092	28.481	2	-51.247	0.003
		URW	28.039	1	-53.579	0.100	30.626	1	-58.752	0.119	31.909	1	-61.318	0.311	32.268	1	-62.037	0.379	28.339	1	-54.178	0.014
		Stasis	**34.210**	**2**	**-62.705**	**0.987**	**34.210**	**2**	**-62.705**	**0.856**	**34.210**	**2**	**-62.705**	**0.622**	**34.210**	**2**	**-62.705**	**0.529**	**34.210**	**2**	**-62.705**	**0.983**
	PC2	GRW	37.735	2	-69.756	0.055	37.508	2	-69.302	0.042	37.645	2	-69.576	0.048	38.923	2	-72.132	0.120	37.676	2	-69.638	0.052
		URW	37.025	1	-71.551	0.134	37.344	1	-72.188	0.177	37.352	1	-72.205	0.177	38.352	1	-74.204	0.337	37.017	1	-71.534	0.133
		Stasis	**40.435**	**2**	**-75.156**	**0.812**	**40.435**	**2**	**-75.156**	**0.781**	**40.435**	**2**	**-75.156**	**0.775**	**40.435**	**2**	**-75.156**	**0.543**	**40.435**	**2**	**-75.156**	**0.815**
	PC3	GRW	32.718	2	-59.721	0.003	33.597	2	-61.480	0.006	35.801	2	-65.888	0.043	37.229	2	-68.745	0.101	34.663	2	-63.612	0.017
		URW	32.642	1	-62.784	0.012	33.569	1	-64.639	0.030	35.800	1	-69.100	0.216	**37.190**	**1**	**-71.881**	**0.485**	34.600	1	-66.700	0.079
		Stasis	**38.639**	**2**	**-71.563**	**0.985**	**38.639**	**2**	**-71.563**	**0.964**	**38.639**	**2**	**-71.563**	**0.741**	38.639	2	-71.563	0.414	**38.639**	**2**	**-71.563**	**0.904**
	PC4	GRW	34.232	2	-62.750	0.003	34.694	2	-63.675	0.005	36.698	2	-67.681	0.034	35.158	2	-64.601	0.008	35.631	2	-65.548	0.013
		URW	34.229	1	-65.958	0.017	34.694	1	-66.889	0.027	36.621	1	-70.741	0.155	35.143	1	-67.787	0.041	35.600	1	-68.699	0.063
		Stasis	**39.885**	**2**	**-74.055**	**0.979**	**39.885**	**2**	**-74.055**	**0.968**	**39.885**	**2**	**-74.055**	**0.812**	**39.885**	**2**	**-74.055**	**0.950**	**39.885**	**2**	**-74.055**	**0.923**
	PC5	GRW	47.685	2	-89.655	0.721	43.356	2	-80.998	0.092	42.726	2	-79.737	0.055	**47.098**	**2**	**-88.481**	**0.691**	**47.245**	**2**	**-88.776**	**0.651**
		URW	44.821	1	-87.142	0.205	42.532	1	-82.564	0.201	42.091	1	-81.682	0.146	44.165	1	-85.829	0.183	44.667	1	-86.834	0.247
		Stasis	**45.396**	**2**	**-85.078**	**0.073**	**45.396**	**2**	**-85.078**	**0.707**	**45.396**	**2**	**-85.078**	**0.799**	45.396	2	-85.078	0.126	45.396	2	-85.078	0.102
	PC6	GRW	41.951	2	-78.187	0.002	44.157	2	-82.599	0.016	48.428	2	-91.141	0.163	42.751	2	-79.788	0.004	45.340	2	-84.966	0.043
		URW	41.933	1	-81.366	0.010	44.120	1	-85.740	0.078	**48.288**	**1**	**-94.075**	**0.709**	42.738	1	-82.977	0.021	45.309	1	-88.117	0.210
		Stasis	**48.183**	**2**	**-90.653**	**0.989**	**48.183**	**2**	**-90.653**	**0.906**	48.183	2	-90.653	0.128	**48.183**	**2**	**-90.653**	**0.975**	**48.183**	**2**	**-90.653**	**0.746**
	Centroid Size	GRW	-2.441	2	10.596	0.354	0.811	2	4.093	0.411	0.045	2	5.625	0.333	0.256	2	5.203	0.352	-1.637	2	8.989	0.377
		URW	**-3.723**	**1**	**9.946**	**0.489**	**-0.450**	**1**	**3.400**	**0.582**	**-0.887**	**1**	**4.274**	**0.654**	**-0.757**	**1**	**4.014**	**0.638**	**-2.872**	**1**	**8.244**	**0.548**
		Stasis	-3.252	2	12.218	0.157	-3.252	2	12.218	0.007	-3.252	2	12.218	0.012	-3.252	2	12.218	0.011	-3.252	2	12.218	0.075

Appendix 5.2 cont.

			carbon date order (3940 & 4040 swapped and 4140 & 4150				carbon dates averaged (4040, 4140, 4150)				carbon dates averaged (4140, 4150)				carbon dates averaged (4140, 4150 and 3940, 4040)			
		Model	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight	logL	K	AICc	Akaike weight
Tom Bowling Bay Fossils	PC1	GRW	27.895	2	-50.076	0.002	26.029	2	-45.659	0.093	29.740	2	-53.481	0.102	26.680	2	-46.959	0.103
		URW	27.892	1	-53.284	0.009	**25.875**	**1**	**-49.083**	**0.517**	**29.532**	**1**	**-56.492**	**0.458**	**26.570**	**1**	**-50.473**	**0.595**
		Stasis	**34.210**	**2**	**-62.705**	**0.989**	27.460	2	-48.520	0.390	31.204	2	-56.409	0.440	27.762	2	-49.123	0.303
	PC2	GRW	37.488	2	-69.262	0.041	31.574	2	-56.748	0.120	36.065	2	-66.130	0.132	31.466	2	-56.531	0.102
		URW	37.337	1	-72.174	0.176	**31.160**	**1**	**-59.653**	**0.512**	**35.670**	**1**	**-68.768**	**0.493**	**31.196**	**1**	**-59.725**	**0.505**
		Stasis	**40.435**	**2**	**-75.156**	**0.783**	32.697	2	-58.993	0.368	37.112	2	-68.223	0.375	32.811	2	-59.221	0.393
	PC3	GRW	33.324	2	-60.933	0.005	30.054	2	-53.708	0.100	34.490	2	-62.980	0.116	31.405	2	-56.410	0.119
		URW	33.260	1	-64.020	0.022	**30.052**	**1**	**-57.438**	**0.643**	**34.483**	**1**	**-66.396**	**0.642**	**31.395**	**1**	**-60.124**	**0.761**
		Stasis	**38.639**	**2**	**-71.563**	**0.973**	31.001	2	-55.603	0.257	35.222	2	-64.443	0.242	31.421	2	-56.441	0.121
	PC4	GRW	35.428	2	-65.142	0.011	29.505	2	-52.611	0.043	29.637	2	-53.274	0.003	31.574	2	-56.749	0.081
		URW	35.384	1	-68.269	0.052	29.505	1	-56.344	0.281	29.635	1	-56.698	0.016	**31.545**	**1**	**-60.422**	**0.506**
		Stasis	**39.885**	**2**	**-74.055**	**0.937**	**32.250**	**2**	**-58.100**	**0.676**	**35.487**	**2**	**-64.974**	**0.982**	33.210	2	-60.019	0.414
	PC5	GRW	42.858	2	-80.001	0.062	36.466	2	-66.532	0.243	42.237	2	-78.474	0.404	36.419	2	-66.438	0.226
		URW	42.132	1	-81.764	0.150	**35.438**	**1**	**-68.210**	**0.563**	**40.694**	**1**	**-78.816**	**0.479**	**35.458**	**1**	**-68.250**	**0.558**
		Stasis	**45.396**	**2**	**-85.078**	**0.788**	36.239	2	-66.077	0.194	40.996	2	-75.992	0.117	36.376	2	-66.353	0.216
	PC6	GRW	44.519	2	-83.323	0.023	40.194	2	-73.989	0.111	38.200	2	-70.401	0.006	36.560	2	-66.720	0.038
		URW	44.362	1	-86.223	0.096	**40.172**	**1**	**-77.677**	**0.705**	38.200	1	-73.829	0.032	36.545	1	-70.423	0.243
		Stasis	**48.183**	**2**	**-90.653**	**0.881**	40.693	2	-74.987	0.184	**43.320**	**2**	**-80.640**	**0.962**	**39.495**	**2**	**-72.591**	**0.719**
	Centroid Size	GRW	-0.424	2	6.563	0.307	1.968	2	2.464	0.448	2.161	2	1.679	0.473	1.790	2	2.820	0.389
		URW	**-1.245**	**1**	**4.989**	**0.675**	**0.305**	**1**	**2.057**	**0.549**	**0.552**	**1**	**1.468**	**0.525**	**0.372**	**1**	**1.924**	**0.609**
		Stasis	-3.252	2	12.218	0.018	-3.238	2	12.875	0.002	-3.166	2	12.332	0.002	-3.227	2	12.854	0.003

Appendix 5.3. Results of Bayesian clustering for all shell populations of *Placostylus ambagaiosus*.

		using Principle Componnets 1-3										using Principle Componnets 1-4									
		clusters										clusters									
		1	**2**	**3**	**4**	**5**	**6**	**7**	***n***	no. of individuals assigned incorrectly	correctly assigned	**1**	**2**	**3**	**4**	**5**	**6**	**7**	***n***	no. of individuals assigned incorrectly	correctly assigned
extant	Surville	**29**	0	-	-	-	-	-	*29*	0	**1.00**	**29**	0	-	-	-	-	-	*29*	0	**1.00**
	CMvD	**19**	**20**	-	-	-	-	-	*39*	19	0.51	**20**	**19**	-	-	-	-	-	*39*	19	0.51
	Tirikawa	**10**	**11**	-	-	-	-	-	*21*	10	0.52	**14**	7	-	-	-	-	-	*21*	7	0.67
	Whareana	4	**26**	-	-	-	-	-	*30*	4	0.87	4	**26**	-	-	-	-	-	*30*	4	0.87
	Unuwhao 1	**27**	5	-	-	-	-	-	*32*	5	0.84	**31**	1	-	-	-	-	-	*32*	1	**0.97**
	Unuwhao	**31**	1	-	-	-	-	-	*32*	1	**0.97**	**32**	0	-	-	-	-	-	*32*	0	**1.00**
	Waikuku	**26**	0	-	-	-	-	-	*26*	0	**1.00**	**26**	0	-	-	-	-	-	*26*	0	**1.00**
Tom Bowling Bay	1473±50YBP	4	**36**	-	-	-	-	-	*40*	4	0.90	9	**31**	-	-	-	-	-	*40*	9	0.78
	1803±62YBP	1	**24**	-	-	-	-	-	*25*	1	**0.96**	3	**22**	-	-	-	-	-	*25*	3	0.88
	2160±50YBP	0	**20**	-	-	-	-	-	*20*	0	**1.00**	0	**20**	-	-	-	-	-	*20*	0	1.00
	2750±60YBP	3	**19**	-	-	-	-	-	*22*	3	0.86	3	**19**	-	-	-	-	-	*22*	3	0.86
	3420±67YBP	0	**25**	-	-	-	-	-	*25*	0	**1.00**	0	**25**	-	-	-	-	-	*25*	0	1.00
	3940±70YBP	4	**17**	-	-	-	-	-	*21*	4	0.81	5	**16**	-	-	-	-	-	*21*	5	0.76
	4040±70YBP	1	**24**	-	-	-	-	-	*25*	1	**0.96**	2	**23**	-	-	-	-	-	*25*	2	0.92
	4140±54YBP	2	**38**	-	-	-	-	-	*40*	2	**0.95**	4	**36**	-	-	-	-	-	*40*	4	0.90
	4150±60YBP	3	**21**	-	-	-	-	-	*24*	3	0.88	4	**20**	-	-	-	-	-	*24*	4	0.83
	4550±70YBP	1	**19**	-	-	-	-	-	*20*	1	**0.95**	1	**19**	-	-	-	-	-	*20*	1	**0.95**
	5250±70YBP	2	**18**	-	-	-	-	-	*20*	2	0.90	3	**17**	-	-	-	-	-	*20*	3	0.85
Cape Maria van Diemen	1000±45YBP	0	**20**	-	-	-	-	-	*20*	0	**1.00**	0	**20**	-	-	-	-	-	*20*	0	**1.00**
	1040+50YBP	0	**20**	-	-	-	-	-	*20*	0	**1.00**	0	**20**	-	-	-	-	-	*20*	0	**1.00**
	1700+90YBP	**12**	8	-	-	-	-	-	*20*	8	0.60	**17**	3	-	-	-	-	-	*20*	3	0.85
	1750±100YBF	1	**9**	-	-	-	-	-	*10*	1	0.90	3	**7**	-	-	-	-	-	*10*	3	0.70
	1950±89YBP	**15**	5	-	-	-	-	-	*20*	5	0.75	**16**	4	-	-	-	-	-	*20*	4	0.80
	2025±44YBP	2	**18**	-	-	-	-	-	*20*	2	0.90	3	**17**	-	-	-	-	-	*20*	3	0.85
	2094±43YBP	2	**18**	-	-	-	-	-	*20*	2	0.90	2	**18**	-	-	-	-	-	*20*	2	0.90
	2290±90YBP	0	**20**	-	-	-	-	-	*20*	0	**1.00**	0	**20**	-	-	-	-	-	*20*	0	**1.00**
	2470±190YBF	1	**19**	-	-	-	-	-	*20*	1	**0.95**	2	**18**	-	-	-	-	-	*20*	2	0.90
	3090±60YBP	2	**18**	-	-	-	-	-	*20*	2	0.90	2	**18**	-	-	-	-	-	*20*	2	0.90
	3440±50YBP	5	**15**	-	-	-	-	-	*20*	5	0.75	6	**14**	-	-	-	-	-	*20*	6	0.70
	3470±120YBF	0	**20**	-	-	-	-	-	*20*	0	**1.00**	0	**20**	-	-	-	-	-	*20*	0	**1.00**
	3850±60YBP	0	**20**	-	-	-	-	-	*20*	0	**1.00**	2	**18**	-	-	-	-	-	*20*	2	0.90
	4247±71YBP	0	**20**	-	-	-	-	-	*20*	0	**1.00**	0	**20**	-	-	-	-	-	*20*	0	**1.00**
	>40,000YPB	**12**	7	-	-	-	-	-	*19*	7	0.63	**13**	6	-	-	-	-	-	*19*	6	0.68
	Pleistocene	7	**17**	-	-	-	-	-	*24*	7	0.71	10	**14**	-	-	-	-	-	*24*	10	0.58
	unknown	0	**21**	-	-	-	-	-	*21*	0	**1.00**	2	**19**	-	-	-	-	-	*21*	2	0.90
	unknown	3	**16**	-	-	-	-	-	*19*	3	0.84	4	**15**	-	-	-	-	-	*19*	4	0.79

Chapter 6. Investigating population structure and gene flow within the New Zealand land snail *Placostylus ambagiosus* using anonymous nuclear markers.

Introduction

The value of combining evidence from fossil samples with genetic data from extant populations of the same evolutionary lineage can not be underestimated (Michaux 1989; Hunt 2010). For example, understanding the link between lineage splitting (speciation) and changes in morphology (Eldredge & Gould 2014) and understanding the forces that produce stasis (Hunt & Rabosky 2014) should come via an integration of evidence from both palaeontology and population genetics. This is only possible where fossil lineages extend into the present; a fairly rare situation. Organisms where both fossil and living representative have been studied provide support for processes that cannot be inferred directly from the fossil record. For example the detection of morphological stasis within biological species (Michaux 1989), biological invasions (Van Bocxlaer et al. 2008), ecotypes recognised within single species (Hills et al. 2012), and rapid morphological divergence (Chiba 1996).

Evolution of shell morphology of the New Zealand land snail *Placostylus ambagiosus* over 5–40 thousand years was characterised by morphological stasis for many independent traits (Chapter 5). Evolutionary lineages could be identified because the samples through time accumulated sequentially at the same location (within a radius of 3km) over thousands of years, and the shell shape variation for each location was encompassed within a single morphological cluster. When shells from the two locations were combined, two overlapping morphological clusters were observed. Bayesian assignment without priors I identified distinct clusters of extant population samples from geometric morphometric data. The morphological variation within extant *Placostylus ambagiosus* might be the result of drift in isolation, or selection. Here I test for concordance of shell morphology and neutral genetic markers using two-dimensional geometric morphometrics and numerous anonymous single nucleotide polymorphic nuclear loci. The neutral genetic markers also allow estimation of past gene flow among populations.

New Zealand has three species of the large *Placostylus* land snails. All have limited distributions and are protected as endangered species. The three species were described using morphological characteristics of the shell including apical angles,

the ratio of height to maximum diameter, number of tubercles in the aperture, and colours of the periostracum and internal aperture. Suter (1908) added some details of the radula and reproductive organs of the species from the Three Kings Islands (*Placostylus bollonsi*). The three New Zealand *Placostylus* species are naturally geographically restricted, and each has been severely affected by human land-use and invasive species over the past few centuries. New Zealand *Placostylus* are protected under local legislation (*Wildlife Act* 1953), are and have been under active management since the early 1980s (Barker et al. 2016) within the context of a Recovery Plan (Parrish et al. 1995).

New Zealand *Placostylus* are an important species ecologically as a prominent endemic invertebrate and have cultural significance to local Maori (Barker et al. 2016). *Placostylus hongii*, occurs along the east coast of Northland Province, North Island, and on some of the nearby small island groups (Brook & McArdle 1999; Buckley et al. 2011; Powell 1979). Anecdotal evidence suggests part of this range may be due to translocation by Maori (Buckley et al. 2011, and references therein) either in leaf mould whilst translocating karaka trees (Climo 1971) or as a live food source (Hayward & Brook 1981).

Placostylus ambagiosus occurs naturally in the northern end of Aupouri Peninsula at the northern extremity of North Island, New Zealand (Figure 6.1). An area of <40000ha (<40 km^2) along a narrow coastal distribution meaning the distance between the two furthest populations is less than 45 km. The original forest habitat of *P. ambagiosus* on northern Aupouri Peninsula has been severely depleted and fragmented by Maori and European occupation, such that native coastal forest now covers only ~3% of the area and is largely replaced by native shrublands or exotic grasslands (Lux et al. 2009). All *P. ambagiosus* populations are considered to be remnants of what would have been a continuous range across the northern end of the peninsular prior to human settlement about 800 years ago (Barker et al. 2016). This idea is supported by the extensive distribution of Holocene fossil shells and subsequent habitat fragmentation. Eleven subspecies of *P. ambagiosus* were recognised by Powell (Powell 1938, 1947, 1951) on the basis of differences in shell morphology. Since the work of Powell, 15 additional small populations of *P.*

ambagiosus have been located (Brook 2002) but their taxonomic status within the framework of Powell's subspecies remains unknown (Barker et al. 2016).

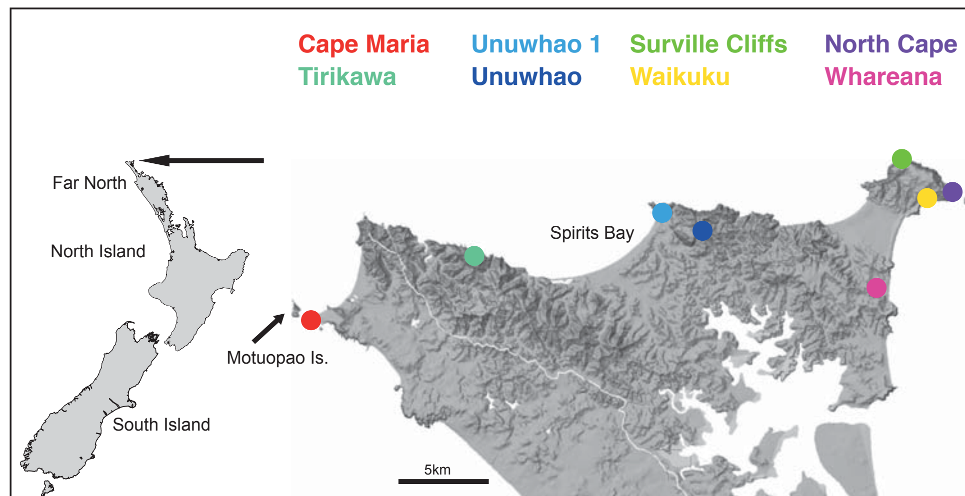


Figure 6.1. The Te Paki district in Northland, New Zealand and sites of *Placostylus ambagiosus* sampled for shell morphology and/or anonymous nuclear data.

Genetic structuring among populations of *P. ambagiosus* has previously been examined with five polymorphic isozyme loci using (Triggs & Sherley 1993) and limited mtDNA analyses (Buckley et al. 2011; Ponder et al. 2003). This structuring of genetic variation is as expected of a single species but quite high given their restricted distribution. The nuclear data did not exhibit diagnostic alleles for any of the named subspecies within *P. ambagiosus.* Two major clusters were distinguishable, separating the east and west of Spirits Bay (genetic distance D = 0.08 – 0.18). Mitochondrial DNA analysis showed support for the same east/west cluster although two haplotypes were shared between east and west groups (Figure 6.2).

That *P. ambagiosus* sub-species share mtDNA haplotypes is indicative of either on-going gene flow, or very recent partitioning not yet subject to complete lineage sorting. Buckley et al. (2011) recognised five distinct groups of *P. ambagiosus* on the basis of mtDNA variation: populations from Cape Maria van Diemen and Motuopao Island; populations in the Tirikawa/Pandora area; populations run along the coast from the eastern end of Spirits Bay to Te Huka; populations from North Cape; and the isolated population at Whareana.

Shell morphology (size and shape) of the New Zealand *Placostylus* has been examined using traditional linear measurments (Buckley et al. 2011; Penniket 1981; Sherley 1996). Buckley et al. (2011) suggested that at least some of the variation in shell height and shape can be explained by local environmental factors, as observed earlier (Penniket 1981). Buckley et al (2011) thus formally synonymised the 11 described subspecies with *P. ambagiosus*.

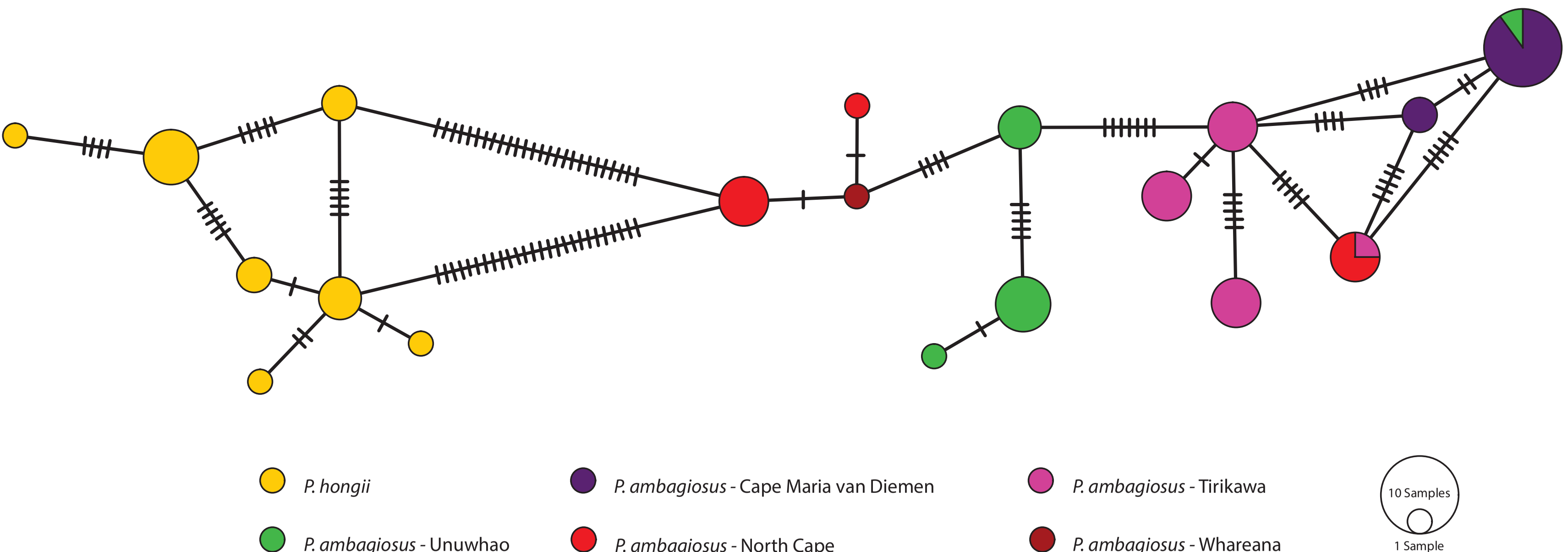


Figure 6.2. mtDNA (COI 610 bp) Haplotype network of *Placostylus ambagiosus and Placostylus hongii* (in yellow) using published sequences (Buckley et al. 2011) and one unpublished sequence of *P. hongii* (from the mt genome in chapter 5).

Geometric morphometric techniques, not previously applied to New Zealand *Placostylus* but successfully used to identify morphologically distinct sympatric species in New Caledonian *Placostylus* (Dowle et al. 2015), are applied to the New Zealand taxa to analyse shell shape variation. To estimate gene flow more polymorphic loci are needed as it has been shown that maximising the number of markers is likely to increase the power of landscape genetic inferences (Landguth et al. 2012) and as this is a non-model organism it is necessary to apply techniques that require little or no prior knowledge of the organism or the system in which it lives. Thus, genetic diversity will be examined with a single nucleotide polymorphism (SNP) marker data set derived from RAD sequence (Senn et al. 2013; Wagner et al. 2013) where next-generation sequencing of restriction fragments is used to create anonymous nuclear loci for analysis (Peterson et al. 2012). High throughput DNA sequencing now enables large datasets to be gathered quickly and is fast becoming a common tool in population genetic studies from model and non-model organisms (Dowle et al. 2015).

Comprehensive morphological techniques (discussed in detail in previous chapters) and large numbers of nuclear loci are combined with Bayesian model-based clustering to examine concordance of shell shape and neutral genetic variation. Concordance might result from fragmentation and divergence due to genetic drift. However, if shell shape is under strong selection it might override gene flow, and neutral genetic clusters would then reflect prior connectivity of populations (and one might detect isolation by distance). Selection on shell shape might result in local adaptation and morphological clusters reflecting environmental convergence.

Methods

This study focuses on variation within *Placostylus ambagiosus.* Evolutionary relationships inferred from the mtDNA (COI) haplotype network (Figure 6.2) shows differentiation between Placostylus *hongii* and *Placostylus ambagiosus*, as have previous studies (Buckley et al. 2011), therefore in this study *P. hongii* has been excluded. Differences in shell shape associated with tubercles around the operculum and restricted distributions of these snails resulted in many sub-species being described, including some known only from fossils (Powell 1979). At least five of the ten named sub-species were included in this study although I relied on material collected for earlier genetic studies and museum collections.

Double Digest Sequencing Restriction site Associated DNA

Material

No new sampling of the endangered species *Placostylus ambagiosus* was required for this study as material from five extant populations was available from previous collections. Tissue obtained for allozyme analysis (Triggs & Sherley 1993) and later used for mtDNA COI studies (Buckley et al. 2011; Trewick et al. 2009) were resampled from the National Arthropod Collection held by Landcare Research in Auckland (Table 6.5).

RAD-Seq

Anonymous single nucleotide polymorphic nuclear markers were generated using high-throughput sequencing, with DNA fragments tagged, so I could assign genotypes to individual snails. The double-digest RAD-Seq protocol was used with minor modifications (Peterson et al. 2012). The restriction endonucleases PstI and BamHI were used to digest whole-genomic DNA from 54 individual snails. The DNA fragments were ligated (Invitrogen T4 ligase) with short DNA barcode sequences to enable identification post sequencing, before the DNA was pooled and indexed. Data generated on an Illumina Hi-Seq (New Zealand Genomics Ltd) were sorted using the STACKS 1.01 pipeline (Catchen et al. 2013; Catchen et al. 2011).

Table 6.1. Role of software (Stacks) core parameters in the assembly of loci and potential sources of genotyping error (from Mastretta-Yanes et al., 2015).

Parameter	How it affects assemble and genotyping error*
Minimum number of identical raw reads required to create a stack (-m) default 3	Reads with convergent sequencing errors can be erroneously labelled in stacks if –m is too low. True alleles will not be recorded and will drop out if –m is too high. –m can decrease genotyping error by distinguishing real loci from PCR and sequencing error, but it can increase error by calling a heterozygous locus as a homozygous when minimum coverage is set too high and one of the alleles is therefore excluded.
Number of mismatches allowed between alleles when processing a single individual (-M) default 2	If –M is too low, some real loci will not be formed, and their alleles will be treated as different loci (under merging). If –M is too large, repetitive sequences and paralogs will form large nonsensical loci (over merging)
Number of mismatches allowed between alleles within a locus when building the catalogue (-n) default 0	For n=0, there would be loci represented independently across individuals that are actually alleles of the same locus. If n>0, the consensus sequence from each locus is used to attempt to merge loci. This is important for population studies where monomorphic or fixed loci may exist in different individuals. Merging fixed alleles as a single locus can increase the probability of assembling real loci and therefore decrease the allele error rate. However, erroneous loci will be created if –n is too high. The setting of n should consider the genetic similarity expected among samples within each study.

*Parameters explanation as in Catchen et al. (2013) and Stacks documentation, effect on genotyping error as from Mastretta-Yanes et al (2015).

STACKS pipeline

Selection of nuclear markers was undertaken so analyses would be performed on loci likely to be single copy and for which the maximum number of individuals could be genotyped (Harvey et al. 2015). In STACKS, a range of parameter settings relating to read coverage, individual number and population coverage were implemented. Initial exploration of the data suggested coverage varied for the 54 individuals. Recommended read coverage settings vary in the literature (Buerkle & Gompert 2013; Peterson et al. 2012), so we experimented with parameter optimisation. Table 1 describes the core parameters discussed here and how varying them affects assembly and genotyping error (Mastretta-Yanes et al. 2015).

Optimising coverage (-m)

Variability in coverage was a problem in this dataset, probably due to variance in the quality of whole genomic DNA, a character I had no control over as this study was using samples collected for earlier studies. Removing poorer quality samples increased the number of loci retained per population, when using a population proportion cut-off. Therefore individuals with low coverage were removed from the dataset. Low coverage combined with a high error rate has the potential to reduce the number of true loci by a substantial amount (up to 51%; Catchen et al. 2011), however at 20x coverage the majority of loci should be identified and correctly assembled (Catchen et al. 2011).

Short-read sequences were sorted into exactly matching stacks of sequences and stacks were then compared and parameters (set by user) determined a set of loci and detected SNPs at each locus using a maximum likelihood framework in USTACKS (Catchen et al. 2013; Catchen et al. 2011). Poor locus coverage for an individual was determined by examining the average stack depth and number of stacks returned for values of (-m:3—15). This was later confirmed by running the data through the denovo_map pipeline and generating a dataset using POPULATIONS: part of the STACKS pipeline allowing computation of population-level summary statistics, the output of site level single nucleotide polymorphism (SNP) calls for subsequent analysis with the software STRUCTURE. For this analysis parameters were set with low stringency (r: 0.5 p: 1) for a dataset containing all individuals. -p is the number of populations an locus must be present in to be included in downstream analysis and –r is the number of individuals, as a proportion, within the population that must contain that locus for it to be included in downstream analysis. The SNP table was then exported and the total number of SNPs present for each individual was calculated. The number of SNPs for each individual was then calculated as a proportion of all possible SNPs under these parameters.

Parameter Optimisation (-M; -n)

I explored the effect of different de novo parameters within the stacks assembly pipeline to maximise the number of loci recovered, using a subset of the data to speed up processing times. The following key parameters were tested with the

values specified in parentheses; the maximum number of mismatches allowed between stacks when processing an individual (-M: 2, 4, 6, 8, 10), the maximum number of mismatches between loci when building the catalogue (-n: 0-5). Only one parameter at a time was varied while keeping other parameters fixed to m=3, M=2, n=0, The value of max_locus_stacks=3 (the maximum number of stacks at a single de novo locus) and the value of –N (the number of mismatches allowed when aligning secondary reads to primary stacks) was always defined as M+2.

SNP Identification

To create the dataset used to examine genetic structuring we used the optimised parameters from above. Here, we report analysis using a minimum of three reads per individual (-m) as providing a reliable set of markers for downstream analysis and excluding all stacks with a lower coverage. Potentially spurious highly repetitive stacks were removed. Within an individual, we allowed a maximum of two mismatches between alleles (-M) and four mismatches between primary and secondary reads (-N= M+2) within ustacks. As above, USTACKS aligns short-read sequences into exactly matching stacks and stacks were then compared and parameters (set by user) determined a set of loci and detected SNPs at each locus using a Maximum Likelihood framework (Catchen et al. 2013; Catchen et al. 2011). In CSTACKS, three mismatches were allowed between alleles in different individuals (-n) when generating the SNP set. CSTACKS forms a catalogue from any set of samples processed by USTACKS generating a set of consensus loci among individuals (Catchen et al. 2013; Catchen et al. 2011). Thus parameter settings were as follows: -m 3, –N 4, –M 2, –n 4. Analysis was restricted to a single SNP site per putative locus (always the first) avoiding potential problems of data nonindependence. Data sets with different parameter settings were analysed, and patterns observed were consistent among settings.

Sample representation per population

A range of values relating to the number of populations a SNP marker was required to be present in before being recorded was tested using POPULATIONS which is part of the STACKS pipeline allowing computation of population-level summary statistics, the output of site level SNP calls for subsequent analysis in STRUCTURE. The *P.*

ambagiosus data set required that each putative locus included was genotyped in individuals from three or more of the five populations (-p 3) and genotyped in ≥90% of individuals within those populations (-r 0.9). The minor allele frequency was set at 0.1 removing rare alleles from the data set. This data set was used to estimate pairwise F_{ST} values as it maximized the data available for each pairwise comparison and each comparison was independent of others. Two smaller, more stringent data sets were also examined. The first required a locus to be detected in all populations for inclusion (-p 5), regardless of the proportion of individuals genotyped (-r 0.0). The other required that ≥90% of individuals be genotyped (-r 0.9) at each SNP locus in at least one populations before being recorded (-p 1). However, given the agreement we found between analyses of each data set, we report only results from the first set of SNP loci here.

Detection of genetic structuring

Isolation by Distance

Analysis of isolation by distance (IBD) was conducted for SNP loci (below), using a Mantel test of the correlation of pairwise linear geographic distances and pairwise F_{ST} with 1000 permutations implemented in IBDWS v3.23 (Jensen et al. 2005). Geographic distances were log transformed as recommended for an expanded stepping stone model (Slatkin 1993).

Structure

Population genetic structure based on nuclear genotypes of 23 *Placostylus ambagiosus* snails was inferred using STRUCTURE 2.3.4 (Pritchard et al. 2000). Initially SNP loci were examined with 10 replications of an admixture model with correlated allele frequencies using a burn-in of 30 000 followed by 100 000 generations, with the number of groups (K) set from 1 to 7. Once concordance across the runs was confirmed, a longer run on the full data set was implemented using an admixture model with correlated allele frequencies, a burn-in of 100 000 followed by 100 000 generations (10 replications), for numbers of groups (K) ranging from 1 to 7. Results from structure were examined using STRUCTURE HARVESTER (Earl & vonHoldt 2011), before being averaged across the 10 replications using CLUMPP

1.1.2 (Jakobsson & Rosenberg 2007). Output genotype structure graphs were then regenerated in DISTRUCT 1.1 (Rosenberg 2003).

Morphometric Analysis

Two hundred and nine shells of extant *Placostylus ambagiosus* representing seven putative populations were examined (material lodged in the Museum of New Zealand (Te Papa Tongawera) (Table 6.2). Only adult shells, determined by the thickening of the apertural lip, were used for analysis.

Table 6.2. *Placostylus ambagiosus* specimens representing extant populations used for morphometric analysis. No shells were analysed from the North Cape population.

Location	Type	Museum	Accession	n	Latitude	Longitude
Cape Maria van Diemen	extant	Te Papa	M.136625	39	-34.2862	172.3887
Tirikawa	extant	Te Papa	M.136627	21	-34.4445	172.7480
Surville Cliffs	extant	Te Papa	M.136621	29	-34.3962	172.0165
Unuwhao (Maungapiko)	extant	Te Papa	M.102324	32	-34.2537	172.5180
Unuwhao1 (Spirits Bay)	extant	Te Papa	M.102503	32	-34.4318	172.8885
Waikuku	extant	Te Papa	M.102328	26	-34.4152	173.0265
Whareana	extant	Te Papa	M.136614	30	-34.4612	172.9952

Photography and digitisation methods were followed as for previous morphology chapters.

To display the major features of shape variation and as an ordination method a principal component analysis was performed using the covariance matrix generated from the dataset in the software MORPHOJ (Klingenberg 2011). Resulting Eigen values were used to perform broken-stick analyses calculating the number of principal components (PCs) that exceed the theoretical scree plot of random, uncorrelated noise using the package VEGAN (Dixon 2003) in R-studio (RStudio Team 2015). Resulting principal component scores were then exported from MORPHOJ in comma separated value format. To identify and measure associations between groups using collection site and population as priors, a canonical variance analysis was performed in MORPHOJ.

Bayesian Clustering

To confirm that morphometric data were consistent with the hypotheses of distinct micro-lineages (Buckley et al. 2011) a model-based clustering method was

implemented using the package MCLUST v4.0 (Fraley & Raftery 1999) in the R Project statistical environment. Analysis of principal component scores that exceeded the theoretical scree plot of random, uncorrelated noise were used to determine how many morphological clusters could be distinguished without *a priori* identification of specimens. The total dataset is considered as a mixture of multivariate normal datasets, with a selection of covariance structures and vectors of expectation. The MCLUST analysis allows for a total of 90 models to be examined (10 different models with various combinations of parameterization and 1 to 9 clusters/components). MCLUST uses Bayesian Information Criterion (BIC) for parameter estimation via the EM algorithm for normal mixture modelling to cluster samples. BIC is a model selection tool that uses negative log-likelihood to maximise statistical goodness of fit and parsimony by penalising models with more parameters, allowing it to filter out overly complicated models. It does not require the specification of priors and so generates unbiased estimates of observation clusters. The model with the BIC score closest to zero was selected as the most likely and individual shells assigned to each cluster.

Results

Double Digest Sequencing Restriction site Associated DNA

RAD -seq

The available analytical software for the paired-end Illumina sequencing protocol was not compatible with the sequence tags on the DNA fragments. Base composition bias in the first four base pairs (SNP site) in the second read was outside of the limits set by the software and thus were excluded from analysis (Table 3). Results are therefore presented from the first read only.

Table 3. Radtags summary: Summary of raw ddRAD New Zealand *Placostylus* data after demultiplexing using radtags in STACKS

Read number	Lane	Retained Reads	Low Quality	Ambiguous Barcodes	Ambiguous RAD-Tag	Total
R1	C	7080683	560372	79587	25616	7746258
	D	6996421	553752	172704	23381	7746258
	Total	*14077104*	*1114124*	*252291*	*48997*	*15492516*
R2	C	775	753	7738341	6389	7746258
	D	761	724	7738911	5862	7746258
	Total	*1536*	*1477*	*15477252*	*12251*	*15492516*

STACKS pipeline

Optimising Coverage

In total 45 *Placostylus ambagiosus* and nine *Placostylus hongii* were sequenced. Average stack depth and number of stacks were compared to selected high quality samples (Appendix 6.1). In order for specimens to be appropriate for use downstream they needed to have average stack depth of 20 or more and over 200 stacks (Appendix 6.1). Individuals that did not have both of these for any given value of –m were not used in downstream analysis and are shown in grey in Table 6.4. Poor coverage resulted in five *P. hongii* and 21 *P. ambagiosus* being removed from the dataset. This left 24 *P. ambagiosus* and four *P. hongii* for downstream analysis. To confirm that the individuals removed from downstream analysis did lack reads for most loci, the full dataset was processed through the STACKS pipeline for the parameters (-m:3, -M:2, -n:4). Loci were tallied for each individual and divided by the total of loci present, displayed as a ratio (Table 6.4). Individuals with a ratio of ≥0.90 were included for downstream analysis.

Table 6.4. *Placostylus ambagiosus* and *P. hongii* individuals showing the proportion of loci returned for each given individual within a group (max number of loci observed for this setting divided by the number of those loci present in a given individual). The specimens shown in grey were removed from further analyses due to low coverage.

Species	ID	Population	all species number	all species ratio
P. ambagiosus	PS186	Te Werahi/ Cape Maria van Diemen	842	1
	PS183		841	1
	PS185		839	1
	PS182		839	1
	PS252		835	0.99
	PS253		812	0.96
	PS184		799	0.95
	PS250		205	0.24
	PS251		7	0.01
	PS189	Tirikawa/ Pandora	841	1
	PS204		839	1
	PS248		834	0.99
	PS243		821	0.98
	PS202		806	0.96
	PS190		760	0.9
	PS244		760	0.9
	PS201		545	0.65
	PS246		293	0.35
	PS203		87	0.1
	PS247		38	0.05
	PS194		16	0.02
	PS249		16	0.02
	PS242		5	0.01
	PS245		0	0
	PS235	Unuwaho/ Matirarau	842	1
	PS197		835	0.99
	PS198		812	0.96
	PS199		819	0.97
	PS205		810	0.96
	PS200		770	0.91
	PS232		155	0.18
	PS196		7	0.01
	PS231		1	0
	PS233		0	0
	PS234		0	0
	PS193	TitirangiPoint/ NorthCape	834	0.99
	PS187		821	0.98
	PS191		293	0.35
	PS195		205	0.24
	PS192		38	0.05
	PS238	Hikurua/ SurvilleCliffs	355	0.42
	PS240		117	0.14
	PS239		3	0
	PS241		1	0
	PS188	Whareana	839	1
P. hongii	PS226	Aorangi Island	819	0.97
	PS230		810	0.96
	PS227		770	0.91
	PS228		545	0.65
	PS229		87	0.1
	PS255	Whangaruru	155	0.18
	PS254		1	0
	PS256		0	0
	PS257		0	0

Intuitively, as the minimum stack depth (–m) increased so does average stack depth but at a cost to the number of stacks formed (Figure 6.3).

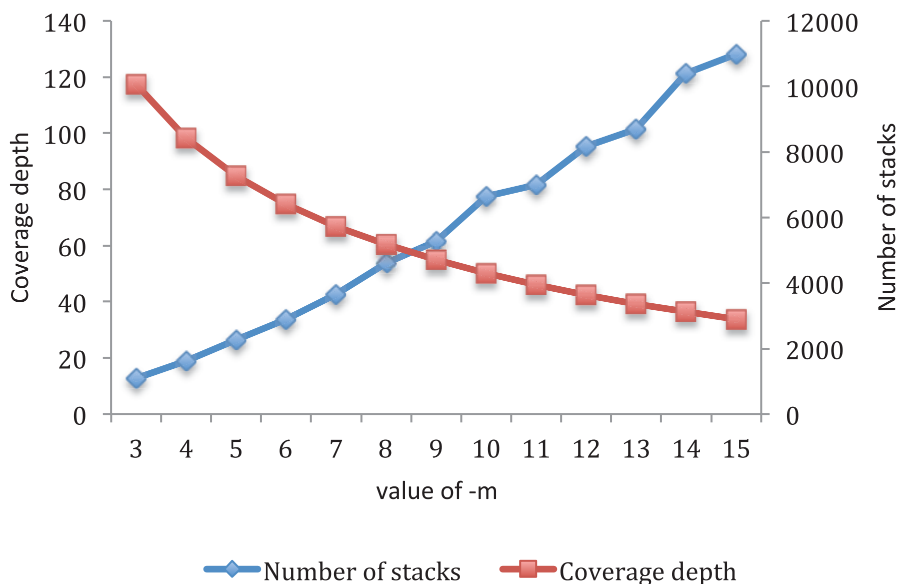


Figure 6.3. Average number of stacks (blue), and coverage depth (red) for all individuals for varying values of -m (number of reads required to form a stack), from RADseq data of the New Zealand snail *Placostylus ambagiosus*.

Parameter Optimisation (-M; -n)

Varying the number of mismatches allowed between alleles when processing a single individual (M) has little effect on coverage depth (Table 6.5). As M increases the number of loci remaining after merging decreases (although the proportion of loci with SNPs might increase) and the number of stacks removed increases in most cases.

Table 6.5. Variation in the parameter settings of the STACKS pipeline is shown compared to the number of loci common for x individuals for a RADseq data from the New Zealand snail *Placostylus ambagiosus*.

Settings	**number of loci common for x individuals**					
default (m3 M2 n0)	**x = 1**	**x = 10**	**x = 20**	**x = 30**	**x = 40**	**x = 50**
m2	*6026*	*3324*	*1745*	*502*	*19*	*0*
m3	5152	2780	1392	200	9	0
m15	1444	685	90	9	0	0
M3	*9338*	*7509*	*4523*	*562*	*26*	*12*
M4	4308	2621	1413	196	8	1
M5	4098	2563	1408	187	2	0
n1	9170	6923	3999	492	24	10
n2	9766	7819	4680	634	56	30
n3	9870	8212	5060	773	119	67
n4	9826	8344	5280	868	171	97

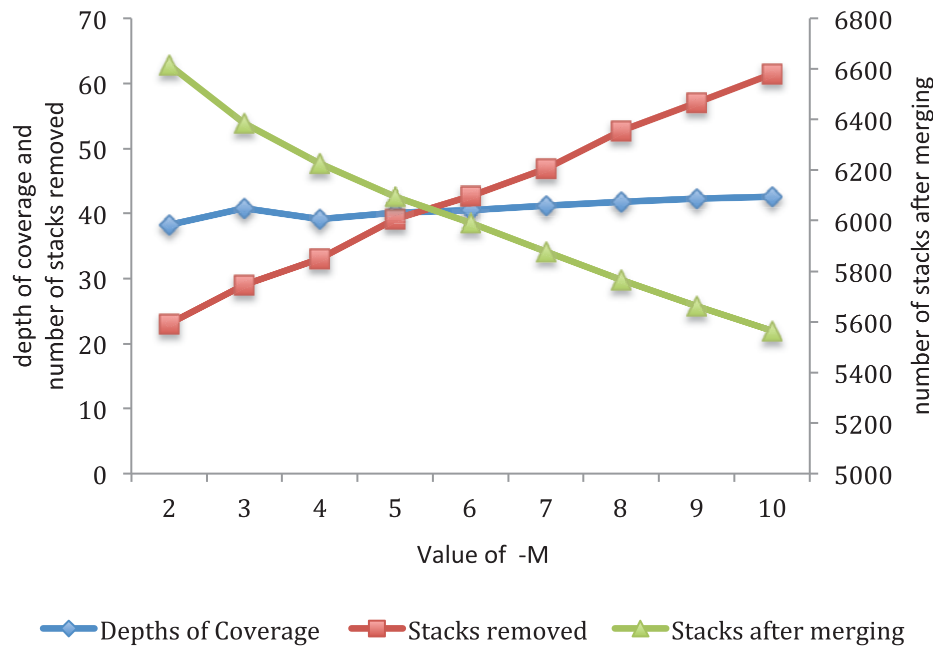


Figure 6.4. Similarity thresholds used in DNA sequence assembly from short reads can reduce the number of loci for downstream analyses. Average depths of coverage (blue), number of stacks removed (red), and stacks remaining after merging (green) for a subsample of 30 individuals for varying values of -M (number of mismatches allowed between loci when processing a single individual) from RADseq data of the New Zealand snail *Placostylus ambagiosus*.

Detection of Genetic Structure

Using the optimal settings described above 2268 single instance SNPs were returned across our 23 individuals. The dataset was run with and without the samples of P. *hongii* included. For all downstream analyses the dataset without *P. hongii* was used in order to focus on variation within *P. ambagiosus*.

Strong evidence of genetic partitioning between the population samples with an overall F_{ST} of 0.7369 was found. Some genetic variation in the sample was partitioned among sample locations with all pairwise F_{ST} estimates significantly greater than zero (Table 6.6). However, SNP variation did not follow a model of isolation by distance (Mantel test: $r = -0.515$; $P < 0.075$).

Table 6.6. Geographic distances (kilometres, top right) and population pairwise F_{ST} (bottom left) from SNP data for *Placostylus ambagiosus* land snails.

Population	**Cape Maria**	**North Cape**	**Whareana**	**Tirikawa**	**Unuwaho**
Cape Maria	*	37.10	31.93	9.91	21.85
North Cape	0.21	*	7.10	27.44	15.65
Whareana	0.03	0.39	*	22.75	11.78
Tirikawa	0.08	0.10	0.06	*	11.76
Unuwhao	0.12	0.11	0.15	0.15	*

Bayesian clustering of genotypes (using STRUCTURE) found the optimal number of genetic clusters within my sample was one or two. Selection of the optimal number of groups relies on the change in likelihood score and therefore resolving K = 1 using deltaK is not possible. The log likelihood probability of the data fitting the model is highest for K = 1. Examination of the assignment probabilities (STRUCTURE plots) suggest a single group of recently separated populations which until recently had on-going gene flow (Figure 6.5). When the K = 2 (with strongest support using deltaK) plot is examined it is clear that several populations have genotypes from each of the two clusters and individuals with mixed genotypes are common.. Assignment probability plots are also shown for K = 3 (the number of clusters resolved by morphology) and K = 5 the number of clusters inferred by Buckley et al (2011) and the number of sampling locations in this study. Results were consistent between SNP datasets created by different parameters. Note that the population sample from Unuwhao contains genotypes that fall into four clusters when K = 5.

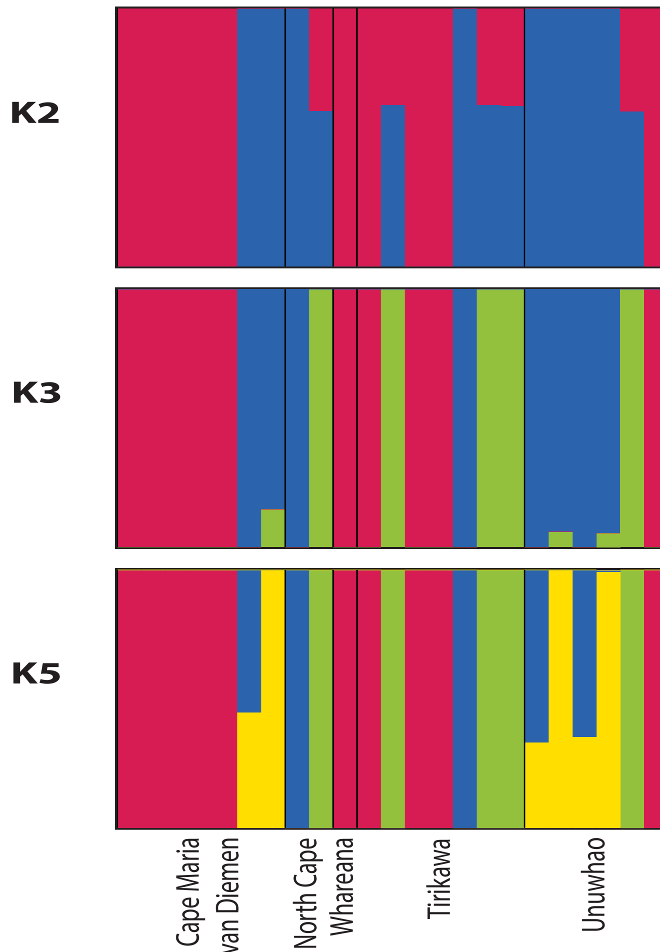


Figure 6.5. Genetic structure is not concordant with sampling location as shown by assignment probabilities of genotypes of New Zealand snails *Placostylus ambagiosus* from five populations. Co-ancestry coefficients from STRUCTURE using 2268 SNP loci for three models are shown but best fitting model (K=1) is not illustrated. K=5 appears to only have 4 colours present as four clusters appears to be the maximum found for this dataset. K=4 found the same pattern of clustering as K=5 and so is not shown.

Morphometric analysis

Shape variation of *Placostylus ambagiosus* shells from extant populations resolved six statistically informative principal components (81% of the variation). As expected of a single species populations overlap in morphospace (Figure 6.6 A and B)). When individuals were assigned to their *a priori* populations canonical variates 1, 2 and 3 separated the majority of specimens successfully, with the exception of specimens from Cape Maria van Diemen and Surville (Figure 6.6 C and D).

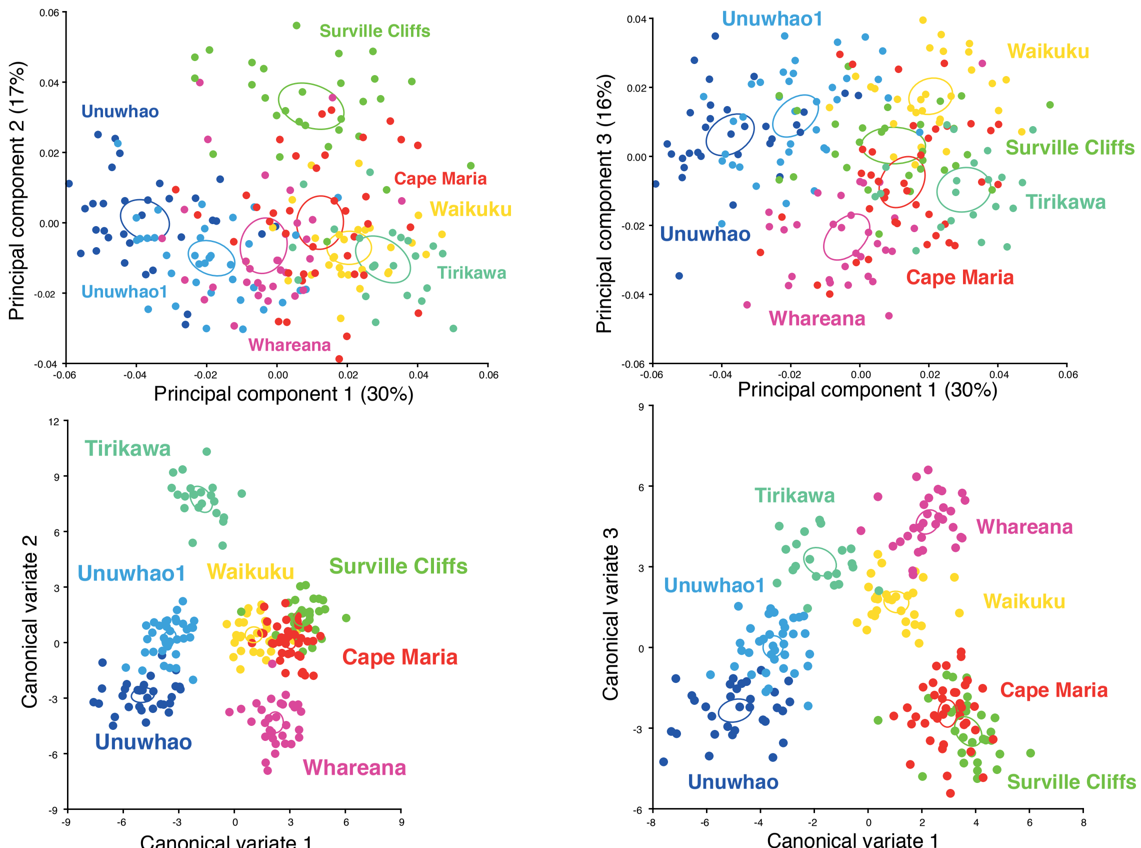


Figure 6.6. *Placostylus ambagiosus* shell shape variation from principal component analysis. A) principal components one and two B) principal components one and three. C) canonical variate one and two D) canonical variate one and three. Colours match those on the map Firgure 1.

Bayesian clustering

Variation of shell shape was greater between population samples than within. *A priori* clustering of between 3 and 6 principal components (PCs) resolved more than one cluster (optimal models from mclust are summarised in Table 6.7). Variation in the first three Principal Components (62.5% of the variation) fitted a model of three morphological clusters (Figure 6.7). The majority of specimens fell into a cluster with specimens from the same location. The three clusters consisted broadly of: Surville Cliffs (plus 11 specimens from Cape Maria van Diemen); Cape Maria van Diemen, Whareana and Tirikawa; both Unuwhao populations and Waikuku. For six of the seven population samples most individuals were assigned to the same cluster (0.97 – 0.85)("correct" assignments Table 6.6). The exception to this was the population sample from Cape Maria van Diemen in which the majority of individuals fell into two distinct groups. When more shell shape variation was included in the analyses (PC4, PC5 & PC6), the best clustering models resolved up to five morphological clusters. Using 5 and 6 significant principal components the optimal model had five clusters (Table 6.6). For five of the seven population samples most individuals still assigned to the same cluster (0.75 -1.00). The exception to this was again Cape Maria van Diemen and the Unuwhao1 populations that grouped a number of samples (8 and 7) with the other population from Unuwhao (Table 6.7).

Table 6.7. Bayesian clustering of morphometric variation of *Placostylus ambagiosus* shells from extant populations: the clustering of each population sample using variations of Principal Components. Correct assignment to the same population cluster is given as specimen number and proportion of the sample.

		clusters					*n*	no. of individuals assigned incorrectly	correct assignment
		1	**2**	**3**	**4**	**5**			
using Principle Componnets 1-3	Surville	**27**	1	1	-	-	*29*	2	0.93
	CMvD	11	**24**	4	-	-	*39*	15	0.62
	Tirikawa	1	**20**	0	-	-	*21*	1	0.95
	Whareana	2	**27**	1	-	-	*30*	3	0.9
	Unuwhao 1	0	1	**31**	-	-	*32*	1	0.97
	Unuwhao	0	1	**31**	-	-	*32*	1	0.97
	Waikuku	2	2	**22**	-	-	*26*	4	0.85
using Principle Componnets 1-4	Surville	**28**	1	0	0	-	*29*	1	0.97
	CMvD	9	**23**	7	0	-	*39*	16	0.59
	Tirikawa	1	**18**	2	0	-	*21*	3	0.86
	Whareana	2	**27**	1	0	-	*30*	3	0.9
	Unuwhao 1	0	3	**23**	6	-	*32*	9	0.72
	Unuwhao	0	2	3	**27**	-	*32*	5	0.84
	Waikuku	2	0	**24**	0	-	*26*	2	0.92
using Principle Componnets 1-5	Surville	**27**	0	0	1	1	*29*	2	0.93
	CMvD	7	**17**	6	1	8	*39*	22	0.44
	Tirikawa	0	0	0	0	**21**	*21*	0	1
	Whareana	2	**25**	1	1	1	*30*	5	0.83
	Unuwhao 1	0	2	**22**	8	0	*32*	10	0.69
	Unuwhao	0	3	2	**27**	0	*32*	5	0.84
	Waikuku	2	0	**23**	0	1	*26*	3	0.88
using Principle Componnets 1-6	Surville	**28**	0	0	0	1	*29*	1	0.97
	CMvD	11	**16**	3	0	9	*39*	23	0.41
	Tirikawa	0	1	0	0	**20**	*21*	1	0.95
	Whareana	2	**24**	2	1	1	*30*	6	0.8
	Unuwhao 1	0	1	**24**	7	0	*32*	8	0.75
	Unuwhao	0	1	2	**29**	0	*32*	3	0.91
	Waikuku	2	0	**24**	0	0	*26*	2	0.92

Discussion

Double Digest Restriction site Associated DNA

Although new tools using high throughput DNA sequences (such as ddRAD) provide greatly improved opportunities to study non-model organisms (Andrews & Luikart 2014) processing of the DNA reads is sensitive to a range of study-specific parameters (Catchen et al. 2013; Harvey et al. 2015; Huang & Knowles 2014). In addition, variation in quality of data was amplified in this study because the endangered status of the snail prevented the ideal sampling and extraction strategy. Post hoc approaches were explored to both remove low quality samples and optimise parameter settings used in this study for compiling an optimal set of variable loci.

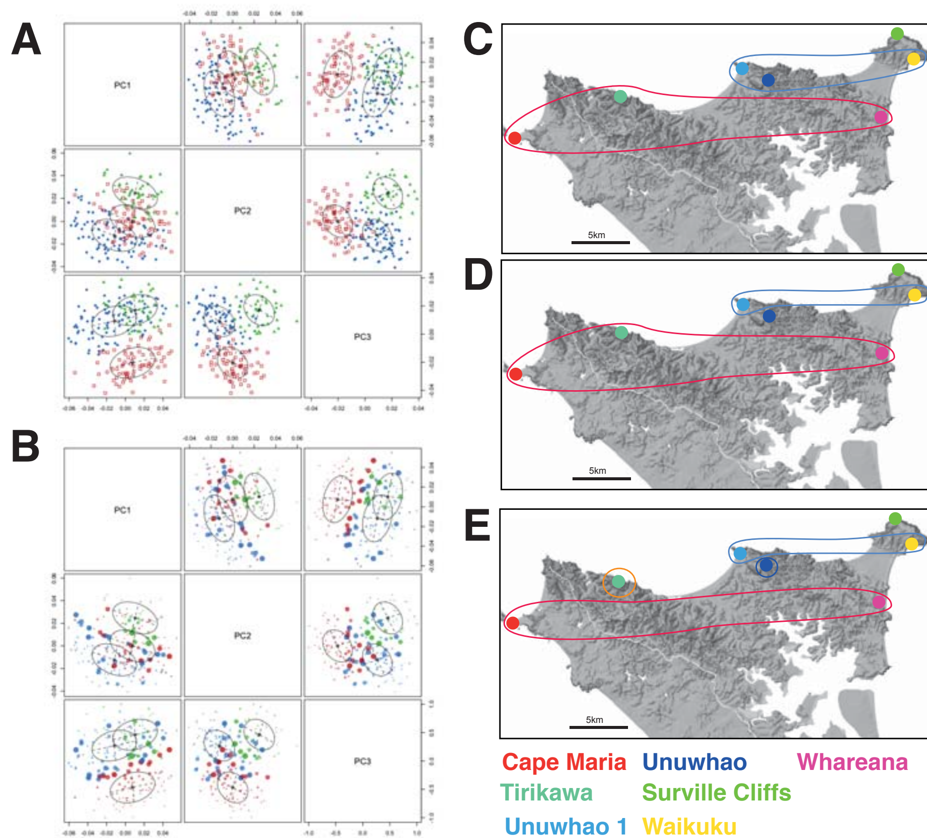


Figure 6.7. Shell shape variation of *Placostylus ambagiosus* is not concordant with geographic proximity. (A) Using PC1, PC2, PC3 three clusters are resolved. Specimens marked as green triangles are mostly composed of samples collected from Surville Cliffs (with some Cape Maria); red squares were collected from Tirikawa, Cape Maria and Whareana; blue circles were collected from Unuwhao and Waikuku. (B) In combination PC1, PC2, PC3 fit a model of three clusters but assignment probabilities vary, the uncertainty is illustrated by spot size, the larger the circle the more uncertain of cluster assignment. Maps show distribution of Bayesian clustering for PC 1-3 (C); PCs 1-4 (D); PCs 1-5 (E). Clustering using all six PCs show the same pattern as for (E).

Polymorphic anonymous nuclear markers were successfully generated. It was found that reducing the number of reads required to form a stack increased the number of loci and number of individuals available for analysis, however, a low read number runs the risk of introducing genotyping error caused by sequencing error.

Optimising Coverage (-m)

Two methods were used to identify suitable specimens and gave similar results. Identifying individuals with low coverage in USTACKS allows for them to be removed from the dataset early on. Using this method I would recommend initially removing

any individual with an average coverage for less than 20x at the selected value of *–m,* as suggested by Catchen et al (2013). Poor coverage does not appear to be something that can be improved greatly by optimisation of parameters. I suggest datasets with poor or varying coverage be run through the stack pipeline and downstream analysis for two values of *–m*: one low, allowing for a greater number of loci and another higher stringency setting allowing more reliable identification of loci but fewer loci. Comparing results from downstream analysis will inform the user of the best parameter to optimise: number of total loci or number of loci common to a percentage of the population. Although sample sizes for snail specimens were small (n=23) a large number of polymorphic nuclear loci (2268) were generated for downstream analysis. Amplifying more (and more variable) loci is likely to increase the power of landscape genetic inferences more than increasing number of individuals (Landguth et al. 2012).

Parameter Optimisation (-M –n)

The average depth of coverage varies little between different values of *-M* (mean for all samples: 38-43) although there is a tendency for increases as values of *-M* increase. This is to be expected as the main parameter that controls average depth of coverage is *-m* (number of raw reads required to form a stack) however it follows that as you allow more differences within loci more stacks will merge to form loci resulting in an increase in average coverage. This in turn results in an average decrease in the number of stacks after merging (mean for all samples: 6617 – 5565) though is this is variable between individuals. If the number of mismatches allowed between alleles when processing a single individual (*–M*) is too low, some real loci will not be formed, and their alleles will be treated as different loci (under merging). However, if *–M* is too large, repetitive sequences and paralogs will form large nonsensical loci (over-merging) (Catchen et al. 2011; Mastretta-Yanes et al. 2015).

Sample representation per population

Varying the stringency of settings in the POPULATIONS module of the STACKS pipeline has an effect on the number of loci generated, however little difference was observed in the results of downstream analysis.

Detection of genetic structure

Genetic structure within the narrow range of extant *Placostylus ambagiosus* did not resolve the three clear clusters identified using morphological variation. No evidence of isolation by distance was detected as expected if geographic distance was a barrier to gene flow. Assignment probabilities of individual snail genotype to >1 cluster were often low, as expected of genetic mixing. The SNP data reflects a single species without strong genetic structure, in contrast to the observed morphological clusters and mitochondrial haplotype data of this species.

Analysis of the SNP data suggests that populations of *Placostylus ambagiosus* have been subject to admixing throughout their range. The lack of strong structure in the SNP data suggests that any fragmentation and subsequent isolation between populations is recent and likely to be human induced. The original forest habitat of *P. ambagiosus* on northern Aupouri Peninsula has been severely depleted and fragmented by Maori and European occupation, such that native coastal forest now covers only ~3% of the area and is largely replaced by native shrublands or exotic grasslands (Lux et al. 2009). All populations are highly restricted in their distributions and considered small remnant populations (Barker et al. 2016) of a once much larger population, an idea supported by the extensive distribution of Holocene fossil shells and subsequent habitat fragmentation.

The fragmented populations show some evidence of divergence via genetic drift if one examines pairwise F_{ST} values. However, samples are small and genetic assignment of individual genotypes do not suggest snails from the same locations cluster together (population structure) and thus contradict an inference of population isolation.

Morphological variation

Shell shape variation of living populations of *Placostylus ambagiosus* resolved three clusters, which although clearly concordant with location sampling, did not group populations geographically adjacent to one-another. For example Tirikawa is not neighbour to Whareana, however they form a cluster with samples from Cape Maria van Diemen. However, this geometric morphometric analyses is clearly identifying

statistically significant and biologically interesting shell variation in contrast to the tools employed earlier that failed to resolve population differentiation (Buckley et al. 2011). The majority of morphological variation was explained by three clusters: Cape Maria van Diemen, Whareana and Tirikawa; Surville Cliffs; both Unuwhao populations and Waikuku. When size of the snail shells is also considered the population sample from Surville Cliffs contains significantly smaller individuals, and the sample from Unuwhao1 are on average larger than snails from all other locations (including Unuwhao, only 2km distant; see Figure 5.7). The clustering into three groups is slightly different to that observed by Sherley (1996) (using traditional morphological methods) which in turn was different from Penniket (1981). None of these studies have matched the 15 named subspecies of Powell (1938, 1947, 1951). The strength of the geometric morphometric technique used here is that shape is studied without the confounding influence of location, scale (size), and rotation using partial Procrustes superimposition, so that pure 'shape' can be studied in isolation. It is perhaps for this reason that my study has resolved considerably more significant variation among populations.

Ecotypes arise through local adaptation to local environments; a process not necessarily requiring many generations (Hendry et al. 2007; Räsänen & Hendry 2008; Rundle & Nosil 2005). There are many examples of ecotypes within snail species, including the rocky shore intertidal snail *Littorina saxatilis* (Galindo & Grahame 2014) and the *Mandarina* land snails of the Bonin Island (Chiba 2004; Davison & Chiba 2006). Land snails are very sensitive to environmental conditions and desiccation is a major environmental pressure, particularly in those taxa lacking an operculum. A link between morphology and environment has been demonstrated in New Caledonian *Placostylus* snails (Dowle et al. 2015; Penniket 1981). The morphological clusters observed within *Placostylus ambagiosus* could represent plastic responses to similar environments or fixed morphological differences between locally adapted populations. Because I only looked at neutral genetic variation there may be structure from genetic markers under selection that are concordant with shell shape. In *Placostylus ambagiosus* I found no evidence for genetic structuring supporting the idea that these populations are ecotypes. Shell shape traits can evolve independently (Chapter 5) so there may be a combination of

evolutionary drivers (environmental plasticity and fixed morphological differences). The well resolved morphological differentiation that groups the sample of seven populations into three clusters could well reflect environmental convergence across the top of northland with oceanic weather conditions likely creating differing habitats and microclimates. Some of the morphological variation documented here might be a plastic response to the local conditions but as every principal component is mathematically independent, each can now be studied independently in relation to environmental variables (e.g. Dowle et al 2015).

The pattern of morphological clustering is not reflected in the mitochondrial haplotype network. For example the mtDNA haplotype sharing seen between different populations involves populations from distinct morphological clusters: Cape Maria populations and Unuwhao share mtDNA haplotypes, but are each part of distinct morphological clusters.

With *Placostylus ambagiosus* I found evidence for what might be environmental convergence in shell shape, and gene flow between what are now highly fragmented isolated populations. If morphology were solely a plastic response to environmental conditions then we might expect little genetic divergence between geographically close populations, but we might expect to see support of isolation by distance. Although there is a lack of strong genetic structure suggesting a model of gene exchange between all populations I see no evidence of isolation by distance in *Placostylus ambagiosus*. And morphological clusters are not geographically adjacent to each other. Some traits may be highly plastic in response to environment whilst others a response to different environmental pressures likely to drive the formation of morphologically distinct snails. Rapid morphological adaptation has been documented in other land snails (Chiba 1996; Dowle et al. 2015).

Over just 40km three clusters of morphologically distinct *P. ambagiosus* populations occur. In contrast, study of fossil *P. ambagiosus* shells revealed morphological stasis for all shape traits, over thousands of years (Chapter 5). If the environment at each of the two locations was uniform over this time then constraining selection could explain the lack of morphological change. We have detected evidence of gene flow

among extant *P. ambagiosus* populations so it is possible for gene flow to be responsible for morphological stasis (Chapter 5), however, gene flow has not prevented local differentiation suggesting both stasis and morphological adaption is possible in the face of gene flow if selection is operating on these traits.

Gene flow between diverging lineages is an important process of evolution that is relatively common and does not necessarily constrain further divergence (Abbott et al. 2013; Mallet 2008; Nosil 2008). Gene flow is commonly thought to homogenise variation within species and it has traditionally been considered that speciation (or population differentiation) is unlikely to occur in the presence of gene flow (Coyne & Orr 2004; Mayr 1963). Many models of speciation have emphasised partitioning of populations by some extrinsic process (e.g. allopatry) as a prerequisite for evolution (Coyne & Orr 2004; Mayr 1963). However the effects of gene flow can have multiple consequences; it can allow the passing of adaptive traits, reinforce reproductive barriers, cause hybrid speciation, limit adaptation or result in the loss of a parental species (Abbott et al. 2013; Morgan-Richards et al. 2009) and if selection on particular loci is intense enough the effects of gene flow could be mitigated. How these processes impact the diversity of *Placostylus ambagiosus* populations is unclear. But the evidence of morphological clusters, although there is a lack of both isolation by distance and clear genetic structure, suggest that the homogenisation of all populations was not straightforward.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246.

Alex Buerkle, C., & Gompert, Z. (2013). Population genomics based on low coverage sequencing: how low should we go? *Molecular Ecology*, *22*(11), 3028–3035.

Andrews, K. R., & Luikart, G. (2014). Recent novel approaches for population genomics data analysis. *Molecular Ecology*, *23*(7), 1661–1667.

Barker, G. M., Brodie, G., Bogitini, L., & Pippard, H. (2016). Diversity and current conservation status of Melanesian–New Zealand placostyline land snails (Gastropoda : Bothriembryontidae), with discussion of conservation imperatives, priorities and methodology issues. *Pacific Conservation Biology*, *22*(3), 203.

Brook, F. J. (2002). Uncommon and rare landsnails in the Northland region of New Zealand, and an assessment of conservation management priorities. Departmnet of Conservation, Northland Conservancy.

Brook, F. J., & McArdle, B. H. (1999). Morphological variation, biogeography and local extinction of the northern New Zealand landsnail *Placostylus hongii* (Gastropoda: Bulimulidae). *Journal of the Royal Society of New Zealand*, *29*(4), 407–434.

Buckley, T., Stringer, I., Gleeson, D., Howitt, R., Attanayake, D., Parrish, R., … Rohan, M. (2011). A revision of the New Zealand *Placostylus* land snails using mitochondrial DNA and shell morphometric analyses, with implications for conservation. *New Zealand Journal of Zoology*, *38*(1), 55–81.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3: Genes, Genomes, Genetics*, *1*(3), 171–182.

Chiba, S. (1996). A 40,000-Year Record of Discontinuous Evolution of Island Snails. *Paleobiology*, *22*(2), 177–188.

Chiba, S. (2004). Ecological and morphological patterns in communities of land snails of the genus *Mandarina* from the Bonin Islands. *Journal of Evolutionary Biology*, *17*(1), 131–143.

Climo, F. M. (1971). Additions to the land snail fauna of the Poor Knights Islands, New Zealand. *Journal of the Royal Society of New Zealand*, *1*(1), 65–69.

Coyne, J., & Orr, H. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.

Davison, A., & Chiba, S. (2006). The recent history and population structure of five *Mandarina* snail species from subtropical Ogasawara (Bonin Islands, Japan). *Molecular Ecology*, *15*(10), 2905–2919.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927–930.

Dowle, E. J., Morgan-Richards, M., Brescia, F., & Trewick, S. A. (2015). Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. *Molecular Ecology*, *24*(16), 4205–4221.

Earl, D. A., & vonHoldt, B. M. (2011). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361.
Eldredge, N., & Gould, S. J. (2014). Puntated equilibira: An alteranative to phyletic gradualism. In *Essential Readings in Evolutionary Biology* (pp. 239–273). Baltimore, Maryland: JHU Press.
Fraley, C., & Raftery, A. E. (1999). MCLUST: Software for Model-Based Cluster Analysis. *Journal of Classification*, *16*(2), 297–306.
Galindo, J., & Grahame, J. W. (2014). Ecological Speciation and the Intertidal Snail Littorina saxatilis. *Advances in Ecology*, *2014*, e239251.
Harvey, M. G., Judy, C. D., Seeholzer, G. F., Maley, J. M., Graves, G. R., & Brumfield, R. T. (2015). Similarity thresholds used in DNA sequence assembly from short reads can reduce the comparability of population histories across species. *PeerJ*, *3*, e895.
Hayward, B., & Brook, F. J. (1981). Exploitation and redistribution of flax snail (*Placostylus*) by the prehistoric maori. *New Zealand Journal of Zoology*, *4*, 33–36.
Hendry, A. P., Nosil, P., & Rieseberg, L. H. (2007). The speed of ecological speciation. *Functional Ecology*, *21*(3), 455–464.
Hills, S. F. K., Crampton, J. S., Trewick, S. A., & Morgan-Richards, M. (2012). DNA and Morphology Unite Two Species and 10 Million Year Old Fossils. *PLOS ONE*, *7*(12), e52083. h
Huang, H., & Knowles, L. L. (2014). Unforeseen Consequences of Excluding Missing Data from Next-Generation Sequences: Simulation Study of RAD Sequences. *Systematic Biology*, syu046.
Hunt, G. (2010). Evolution in Fossil Lineages: Paleontology and The Origin of Species. *The American Naturalist*, *176*(S1), S61–S76.
Hunt, G., & Rabosky, D. L. (2014). Phenotypic Evolution in Fossil Species: Pattern and Process. *Annual Review of Earth and Planetary Sciences*, *42*(1), 421–441.
Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801–1806.
Jensen, J. L., Bohonak, A. J., & Kelley, S. T. (2005). Isolation by distance, web service. *BMC Genetics*, *6*, 13.
Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*(2), 353–357.
Landguth, E. L., Fedy, B. C., OYLER-McCANCE, S. J., Garey, A. L., Emel, S. L., Mumma, M., … Cushman, S. A. (2012). Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources*, *12*(2), 276–284.
Lux, J., Holland, W., Rate, S., & Beadel, S. (2009). Natural areas of Te Paki Ecological District. Whangarei, Department of Conservation.
Mallet, J. (2008). Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1506), 2971–2986.
Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*(1), 28–41.

Mayr, E. (1963). *Animal species and evolution*. Cambridge, Massachesetts: Belknap Press of Harvard University Press.

Michaux, B. (1989). Morphological variation of species through time. *Biological Journal of the Linnean Society*, *38*(3), 239–255.

Morgan ‐ Richards, M., Smissen, R. D., Shepherd, L. D., Wallis, G. P., Hayward, J. J., Chan, C., ⋯ Chapman, H. M. (2009). A review of genetic analyses of hybridisation in New Zealand. *Journal of the Royal Society of New Zealand*, *39*(1), 15–34.

Nosil, P. (2008). Speciation with gene flow could be common. *Molecular Ecology*, *17*(9), 2103–2106.

Parrish, R., Sherley, G., & Aviss, M. (1995). Giant Land Snail Recovery Plan, *Placostylus* Spp., *Paryphanta* Sp. Threatened Species Unit, Department of Conservation.

Penniket, A. S. W. (1981). *Population studies of land snails of the genus Placostylus in the north of New Zealand.* (Unpublished MSc thesis). Auckland University, Auckland.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLOS ONE*, *7*(5), e37135.

Ponder, W. F., Colgan, D. J., Gleeson, D. M., & Sherley, G. H. (2003). Relationships of *Placostylus* from Lord Howe Island: an investigation using the mitochondrial cytochrome c oxidase 1 gene. *Molluscan Research*, *23*(2), 159–178.

Powell, A. (1979). *New Zealand Mollusca: marine, land, and freshwater shells*. Collins.

Powell, A. W. B. (1938). The Paryphantidae of New Zealand No. IV. and The Genus *Placostylus* in New Zealand. *Records of the Auckland Institute and Museum*, *2*(3), 133–150.

Powell, A. W. B. (1947). Distribution of Placostylus Land Snails in Northernmost New Zealand on JSTOR. *Records of the Auckland Institute and Museum*, *3*(3), 173–188.

Powell, A. W. B. (1951). On Further Colonies of *Placostylus* Land Snails from Northernmost New Zealand. *Records of the Auckalnd Institute and Museum*, *4*(2), 134–140.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Räsänen, K., & Hendry, A. P. (2008). Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, *11*(6), 624–636.

Rosenberg, N. A. (2003). distruct: a program for the graphical display of population structure: PROGRAM NOTE. *Molecular Ecology Notes*, *4*(1), 137–138.

RStudio Team. (2015). *RStudio: Integrated Development for R.* Boston, MA: RStudio, Inc. Retrieved from https://www.rstudio.com/

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*(3), 336–352.

Senn, H., Ogden, R., Cezard, T., Gharbi, K., Iqbal, Z., Johnson, E., … McEwing, R. (2013). Reference-free SNP discovery for the Eurasian beaver from restriction site–associated DNA paired-end data. *Molecular Ecology*, *22*(11), 3141–3150.

Sherley, G. (1996). Morphological variation in the shells of Placostylus species (Gastropoda: Bulimulidae) in New Zealand and implications for their conservation. *New Zealand Journal of Zoology*, *23*(1), 73–82.

Slatkin, M. (1993). Isolation by Distance in Equilibrium and Non-Equilibrium Populations. *Evolution*, *47*(1), 264–279.

Suter, H. (1908). A new Placostylus from New Zealand. *Transactions and Proceedings of the New Zealand Institute*, *40*, 340–343.

Trewick, S., Brescia, F., & Jordan, C. (2009). Diversity and phylogeny of New Caledonian *Placostylus* land snails; evidence from mitochondrial DNA. In *Grandcolas P ed. Zoologia Neocaledonica 7* (pp. 421–436).

Triggs, S. J., & Sherley, G. H. (1993). Allozyme genetic diversity in *Placostylus* land snails and implications for conservation. *New Zealand Journal of Zoology*, *20*(1), 19–33.

Van Bocxlaer, B., Damme, D. V., & Feibel, C. S. (2008). Gradual Versus Punctuated Equilibrium Evolution in the Turkana Basin Molluscs: Evolutionary Events or Biological Invasions? *Evolution*, *62*(3), 511–520.

Wagner, C. E., Keller, I., Wittwer, S., Selz, O. M., Mwaiko, S., Greuter, L., … Seehausen, O. (2013). Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, *22*(3), 787–798.

Appendix

Appendix 6.1. Specimens grouped by population showing the average coverage per stack and the number of stacks assembled for each specimen at different values of –m (minimum read depth to form a stack). Values in grey were lower than required value of 20 for coverage depth and 200 for number of stacks. Specimens highlighted in grey were subsequently removed from downstream analysis.

				Minimum read depth (-m)												
Species	**ID**	**Population**		**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
P. ambagiosus	PS182	Cape Maria van Diemen	Coverage depth	10	11	12	14	15	16	17	18	19	20	22	23	25
			Number of stacks	19811	16946	14625	12673	11065	9604	8249	7007	5899	4936	4065	3332	2736
	PS183		Coverage depth	22	23	25	26	27	28	29	30	31	32	33	33	34
			Number of stacks	25998	24024	22331	21031	19811	18879	18006	17156	16391	15666	14971	14307	13677
	PS184		Coverage depth	6	7	9	10	12	15	18	20	24	29	36	43	51
			Number of stacks	11914	8250	5615	3781	2481	1627	1071	712	484	334	231	167	120
	PS185		Coverage depth	20	22	23	25	26	27	28	28	29	30	31	32	33
			Number of stacks	25143	23164	21533	20119	18978	18007	17103	16315	15549	14855	14179	13461	12763
	PS186		Coverage depth	23	25	26	28	29	30	31	32	33	34	35	36	36
			Number of stacks	26078	24232	22619	21262	20071	19099	18214	17422	16679	15958	15320	14691	14087
	PS252		Coverage depth	16	18	19	20	21	22	23	24	25	26	27	28	29
			Number of stacks	24270	21867	20068	18576	17277	16076	14989	14007	12955	12019	11113	10209	9417
	PS253		Coverage depth	9	10	11	12	13	14	15	17	18	20	22	24	27
			Number of stacks	17166	14169	11813	9808	7975	6474	5174	4073	3176	2443	1865	1437	1108
	PS250		Coverage depth	4	7	13	22	38	102	118	132	139	148	148	161	161
			Number of stacks	1483	461	165	68	34	22	19	17	16	14	14	13	13
	PS251		Coverage depth	10	22	39	74	92	104	104	120	120	140	140	140	226
			Number of stacks	62	22	12	10	8	7	7	6	6	5	5	5	5
	PS189	Tirikawa	Coverage depth	22	23	25	26	27	28	29	30	31	32	33	34	35
			Number of stacks	26036	24055	22376	21075	19858	18936	18045	17201	16456	15728	15010	14368	13751
	PS190		Coverage depth	6	7	9	10	12	14	17	21	26	31	37	44	53
			Number of stacks	11709	8083	5575	3761	2505	1625	1051	650	435	313	227	178	130
	PS201		Coverage depth	5	6	8	12	16	23	30	40	57	73	86	99	136
			Number of stacks	4948	2530	1228	625	325	190	113	81	65	48	39	33	30
	PS202		Coverage depth	6	7	9	10	12	15	17	21	23	28	37	42	48
			Number of stacks	12315	8710	5993	4078	2727	1804	1213	798	551	377	266	187	139
	PS204		Coverage depth	20	22	23	25	26	27	28	28	29	30	31	32	33
			Number of stacks	25148	23168	21535	20120	18983	18009	17111	16323	15553	14860	14181	13471	12770
	PS243		Coverage depth	12	13	15	16	17	18	19	20	21	22	23	24	25
			Number of stacks	22658	19828	17577	15818	14223	12768	11403	10151	8955	7846	6921	6035	5258
	PS244		Coverage depth	6	7	9	10	12	14	17	21	26	31	37	44	53
			Number of stacks	11709	8083	5575	3761	2505	1625	1051	650	435	313	227	178	130
	PS248		Coverage depth	21	23	24	26	27	28	29	30	31	32	33	33	34
			Number of stacks	25741	23587	21839	20449	19258	18211	17360	16543	15781	15045	14345	13720	13041
	PS194		Coverage depth	8	18	37	48	53	95	106	119	136	136	136	156	156
			Number of stacks	119	44	19	14	11	10	9	8	7	7	7	6	6
	PS203		Coverage depth	5	9	18	26	35	60	71	80	91	102	123	133	133
			Number of stacks	824	246	95	56	35	24	20	17	14	12	10	9	9
	PS242		Coverage depth	11	23	28	37	47	54	54	179	179	179	229	229	229
			Number of stacks	33	14	11	8	6	5	5	5	5	5	4	4	4
	PS245		Coverage depth	36	37	43	104	112	159	324	324	324	324	324	324	324
			Number of stacks	6	6	5	4	3	2	1	1	1	1	1	1	1
	PS246		Coverage depth	5	7	9	13	17	20	37	47	51	63	71	75	87
			Number of stacks	2252	1048	523	261	142	96	61	44	38	29	24	22	18
	PS247		Coverage depth	6	14	33	42	57	62	68	70	77	77	80	92	92
			Number of stacks	299	86	36	28	20	18	16	14	12	12	11	9	9
	PS249		Coverage depth	4	7	13	22	38	102	118	132	139	148	148	161	161
			Number of stacks	1483	461	165	68	34	22	19	17	16	14	14	13	13

Appendix 6.1 cont.

Species	ID	Population		Minimum read depth (-m)												
				3	4	5	6	7	8	9	10	11	12	13	14	15
	PS197	Unuwaho	Coverage depth	16	18	19	20	21	22	23	24	25	26	27	28	29
			Number of stacks	24270	21867	20068	18576	17277	16076	14989	14007	12955	12019	11113	10209	9417
	PS198		Coverage depth	9	10	11	12	13	14	15	17	18	20	22	24	27
			Number of stacks	17166	14169	11813	9808	7975	6474	5174	4073	3176	2443	1865	1437	1108
	PS199		Coverage depth	12	14	15	16	17	18	19	20	21	22	23	24	26
			Number of stacks	22421	19631	17489	15738	14200	12688	11322	10075	8945	7944	7006	6155	5331
	PS200		Coverage depth	6	7	8	10	11	13	14	17	20	26	34	44	56
			Number of stacks	11195	7949	5507	3756	2555	1745	1169	764	503	339	225	155	110
	PS235		Coverage depth	23	25	26	28	29	30	31	32	33	34	35	36	36
			Number of stacks	26078	24232	22619	21262	20071	19099	18214	17422	16679	15958	15320	14691	14087
	PS205		Coverage depth	6	7	8	10	11	13	15	17	21	26	34	40	46
			Number of stacks	13370	9767	7016	4886	3327	2238	1480	966	611	413	274	203	150
	PS196		Coverage depth	10	22	39	74	92	104	104	120	120	140	140	140	226
			Number of stacks	62	22	12	10	8	7	7	6	6	5	5	5	5
	PS231		Coverage depth	5	15	45	84	88	129	133	133	133	133	108	501	501
			Number of stacks	94	24	7	4	4	3	3	3	3	3	2	2	2
	PS232		Coverage depth	4	7	12	24	32	59	83	108	121	140	194	213	213
			Number of stacks	1151	328	117	50	33	24	17	16	14	12	9	8	8
	PS233		Coverage depth	20	88	110	110	205	205	205	205	205	351	351	351	351
			Number of stacks	6	5	4	4	2	2	2	2	2	1	1	1	1
	PS234		Coverage depth	60	60	87	87	87	87	87	87	87	87	106	106	106
			Number of stacks	3	3	2	2	2	2	2	2	2	2	1	1	1
	PS187	North Cape	Coverage depth	12	13	15	16	17	18	19	20	21	22	23	24	25
			Number of stacks	22655	19820	17572	15814	14216	12765	11398	10140	8951	7843	6914	6034	5253
	PS193		Coverage depth	21	23	24	26	27	28	29	30	31	32	33	33	34
			Number of stacks	25741	23587	21839	20449	19258	18211	17360	16543	15781	15045	14345	13720	13041
	PS191		Coverage depth	5	7	9	13	17	20	37	47	51	63	71	75	87
			Number of stacks	2252	1048	523	261	142	96	61	44	38	29	24	22	18
	PS192		Coverage depth	6	14	33	42	57	62	68	70	77	77	80	92	92
			Number of stacks	299	86	36	28	20	18	16	14	12	12	11	9	9
	PS195		Coverage depth	4	7	13	22	38	102	118	132	139	148	148	161	161
			Number of stacks	1483	461	165	68	34	22	19	17	16	14	14	13	13
	PS188	Whareana	Coverage depth	10	11	12	14	15	16	17	18	19	20	22	23	25
			Number of stacks	19811	16946	14625	12673	11065	9604	8249	7007	5899	4936	4065	3332	2736
	PS238	Surville Cliffs	Coverage depth	4	6	8	11	19	28	42	80	95	109	125	153	163
			Number of stacks	3537	1536	704	304	127	66	46	32	27	23	20	16	15
	PS239		Coverage depth	11	39	59	70	107	107	107	107	107	107	107	107	107
			Number of stacks	28	11	7	6	4	4	4	4	4	4	4	4	4
	PS240		Coverage depth	4	7	19	33	49	86	87	112	124	124	124	124	124
			Number of stacks	812	237	79	38	25	14	14	11	10	10	10	10	9
	PS241		Coverage depth	16	41	63	63	63	63	63	498	498	694	694	694	694
			Number of stacks	16	6	4	4	4	4	4	3	3	2	2	2	2

Appendix 6.1 cont.

Species	ID	Population		Minimum read depth (-m)												
				3	4	5	6	7	8	9	10	11	12	13	14	15
P. hongii	PS226	Aorangi Island	Coverage depth	12	14	15	16	17	18	19	20	21	22	23	24	26
			Number of stacks	22440	19646	17508	15754	14220	12717	11345	10105	8963	7963	7026	6181	5352
	PS227		Coverage depth	6	7	8	10	11	13	14	17	20	26	34	44	56
			Number of stacks	11195	7949	5507	3756	2555	1745	1169	764	503	339	225	155	110
	PS230		Coverage depth	6	7	8	10	11	13	15	17	21	26	34	40	46
			Number of stacks	13370	9767	7016	4886	3327	2238	1480	966	611	413	274	203	150
	PS228		Coverage depth	5	6	8	12	16	23	30	40	57	73	86	99	136
			Number of stacks	4948	2530	1228	625	325	190	113	81	65	48	39	33	30
	PS229		Coverage depth	5	9	18	26	35	60	71	80	91	102	123	133	133
			Number of stacks	824	246	95	56	35	24	20	17	14	12	10	9	9
	PS254	Whanganuru	Coverage depth	4	15	46	87	91	133	137	137	137	137	115	505	505
			Number of stacks	94	24	7	4	4	3	3	3	3	3	2	2	2
	PS255		Coverage depth	4	7	12	24	32	59	83	108	121	140	194	213	213
			Number of stacks	1151	328	117	50	33	24	17	16	14	12	9	8	8
	PS256		Coverage depth	20	88	110	110	205	205	205	205	205	351	351	351	351
			Number of stacks	6	5	4	4	2	2	2	2	2	1	1	1	1
	PS257		Coverage depth	60	60	87	87	87	87	87	87	87	87	106	106	106
			Number of stacks	3	3	2	2	2	2	2	2	2	2	1	1	1
		Average	*Coverage depth*	*13*	*19*	*26*	*34*	*43*	*54*	*61*	*78*	*82*	*95*	*102*	*121*	*128*
			Number of stacks	*10068*	*8432*	*7278*	*6414*	*5726*	*5171*	*4704*	*4302*	*3950*	*3642*	*3368*	*3120*	*2893*

Appendix 6.2. Variations of the parameter M (differences allowed between stacks within individuals) in the STACKS pipeline and the effect on coverage depth, number of stacks after merging and number of stacks removed.

ID	m12 Mx n0	M2	M3	M4	M5	M6	M7	M8	M9	M10
PS182	Depths of Coverage	21	22	21	21	21	21	22	21	21
	Stacks after merging	4954	4838	4759	4679	4628	4566	4490	4418	4346
	Stacks removed	16	19	19	30	32	34	38	41	45
PS183	Depths of Coverage	32	33	32	32	32	32	32	32	32
	Stacks after merging	15872	15287	14873	14543	14279	13995	13719	13451	13211
	Stacks removed	53	66	77	90	99	111	123	132	142
PS184	Depths of Coverage	36	55	28	30	31	37	31	33	35
	Stacks after merging	338	328	319	315	311	310	309	309	306
	Stacks removed	7	8	8	7	8	8	8	8	8
PS185	Depths of Coverage	30	31	31	31	30	30	31	31	31
	Stacks after merging	15024	14441	14041	13717	13477	13206	12940	12702	12485
	Stacks removed	48	59	71	85	93	101	116	122	131
PS186	Depths of Coverage	34	34	34	34	34	34	34	34	34
	Stacks after merging	16143	15491	15030	14675	14404	14102	13804	13553	13314
	Stacks removed	52	65	77	90	98	106	123	129	141
PS187	Depths of Coverage	23	24	23	23	23	22	23	22	23
	Stacks after merging	7913	7698	7551	7424	7318	7189	7067	6943	6828
	Stacks removed	24	32	36	42	47	49	54	65	69
PS188	Depths of Coverage	21	22	21	21	21	21	22	21	21
	Stacks after merging	4954	4838	4759	4679	4628	4566	4490	4418	4346
	Stacks removed	16	19	19	30	32	34	38	41	45
PS189	Depths of Coverage	32	33	32	32	32	32	32	33	33
	Stacks after merging	15935	15343	14928	14598	14332	14047	13769	13499	13259
	Stacks removed	53	66	77	90	99	111	123	132	142
PS190	Depths of Coverage	32	32	32	39	32	32	35	38	38
	Stacks after merging	317	303	291	287	282	281	281	280	276
	Stacks removed	7	9	10	9	9	9	9	9	9
PS193	Depths of Coverage	32	33	32	32	32	32	32	32	32
	Stacks after merging	15194	14562	14166	13851	13585	13327	13064	12825	12608
	Stacks removed	55	71	78	90	99	112	120	134	145
PS197	Depths of Coverage	26	26	26	26	26	26	26	26	26
	Stacks after merging	12149	11794	11521	11281	11075	10873	10665	10474	10278
	Stacks removed	32	46	55	67	73	81	92	101	112

Appendix 6.2 cont.

ID	m12 Mx n0	M2	M3	M4	M5	M6	M7	M8	M9	M10
PS198	Depths of Coverage	21	23	23	21	23	21	22	21	21
	Stacks after merging	2456	2408	2367	2335	2317	2278	2240	2208	2174
	Stacks removed	12	12	10	11	13	19	23	24	24
PS199	Depths of Coverage	23	24	23	23	23	23	23	23	23
	Stacks after merging	8038	7824	7659	7509	7400	7276	7135	7002	6893
	Stacks removed	21	27	34	48	50	55	66	71	76
PS200	Depths of Coverage	26	25	24	24	25	28	27	28	30
	Stacks after merging	340	321	312	308	305	301	298	295	294
	Stacks removed	8	9	9	8	8	8	8	8	8
PS202	Depths of Coverage	36	53	27	29	30	35	30	32	33
	Stacks after merging	382	371	361	357	352	351	349	349	346
	Stacks removed	7	8	8	7	8	8	8	8	8
PS204	Depths of Coverage	30	31	31	31	30	30	31	31	31
	Stacks after merging	15029	14445	14045	13721	13481	13210	12944	12705	12487
	Stacks removed	48	59	71	85	93	101	116	122	131
PS205	Depths of Coverage	27	29	28	28	33	29	32	31	33
	Stacks after merging	417	406	396	390	383	381	379	374	373
	Stacks removed	7	8	8	7	7	7	7	8	8
PS226	Depths of Coverage	23	24	23	23	23	23	23	23	23
	Stacks after merging	8058	7843	7676	7527	7416	7291	7151	7016	6909
	Stacks removed	21	27	34	48	50	55	66	72	77
PS227	Depths of Coverage	26	25	24	24	25	28	27	28	30
	Stacks after merging	340	321	312	308	305	301	298	295	294
	Stacks removed	8	9	9	8	8	8	8	8	8
PS230	Depths of Coverage	27	29	28	28	33	29	32	31	33
	Stacks after merging	417	406	396	390	383	381	379	374	373
	Stacks removed	7	8	8	7	7	7	7	8	8
PS232	Depths of Coverage	140	150	155	159	164	169	172	176	178
	Stacks after merging	13	12	12	12	12	12	12	12	12
	Stacks removed	2	2	2	2	2	2	2	2	2
PS235	Depths of Coverage	34	34	34	34	34	34	34	34	34
	Stacks after merging	16143	15491	15030	14675	14404	14102	13804	13553	13314
	Stacks removed	52	65	77	90	98	106	123	129	141

Appendix 6.2 cont.

ID	m12 Mx n0	M2	M3	M4	M5	M6	M7	M8	M9	M10
PS243	Depths of Coverage	23	24	23	23	23	22	23	22	23
	Stacks after merging	7916	7701	7554	7427	7321	7192	7070	6946	6831
	Stacks removed	24	32	36	42	47	49	54	65	69
PS244	Depths of Coverage	32	32	32	39	32	32	35	38	38
	Stacks after merging	317	303	291	287	282	281	281	280	276
	Stacks removed	7	9	10	9	9	9	9	9	9
PS246	Depths of Coverage	63	65	66	73	76	80	82	84	70
	Stacks after merging	30	29	29	27	27	27	27	27	26
	Stacks removed	2	2	2	2	2	2	2	2	3
PS247	Depths of Coverage	77	80	82	84	86	90	95	99	102
	Stacks after merging	13	13	13	13	13	13	13	13	13
	Stacks removed	2	2	2	2	2	2	2	2	2
PS248	Depths of Coverage	32	33	32	32	32	32	32	32	32
	Stacks after merging	15194	14562	14166	13851	13585	13327	13064	12825	12608
	Stacks removed	55	71	78	90	99	112	120	134	145
PS249	Depths of Coverage	136	146	153	158	162	166	167	168	168
	Stacks after merging	8	8	8	8	8	8	8	8	8
	Stacks removed	1	1	1	1	1	1	1	1	1
PS252	Depths of Coverage	26	26	26	26	26	26	26	26	26
	Stacks after merging	12149	11794	11521	11281	11075	10873	10665	10474	10278
	Stacks removed	32	46	55	67	73	81	92	101	112
PS253	Depths of Coverage	21	23	23	21	23	21	22	21	21
	Stacks after merging	2456	2408	2367	2335	2317	2278	2240	2208	2174
	Stacks removed	12	12	10	11	13	19	23	24	24
Mean	*Depths of Coverage*	*38*	*41*	*39*	*40*	*41*	*41*	*42*	*42*	*43*
	Stacks after merging	*6617*	*6386*	*6225*	*6094*	*5990*	*5878*	*5765*	*5661*	*5565*
	Stacks removed	*23*	*29*	*33*	*39*	*43*	*47*	*53*	*57*	*62*

Chapter 7. Thesis Discussion

Species concepts and conservation

Biologists commonly apply biological species concept to define taxonomic units (Mayr 1942). This is the idea that species are discrete genetic entities that do not exchange alleles, and so are reproductively isolated. However, population geneticists who investigate hybridisation tend to regard the biological species concept as deficient, instead accepting that successful mating does take place between different species (Abbott et al. 2013). The value in delimitating the boundaries of evolutionary lineages is acknowledged, while recognising that it is not a simple problem (Mallet 1995; Vaux et al. 2016a, 2016b).

The criteria used to define taxonomic units within gastropods have traditionally been shell morphology and spatial distribution (Clarke & Murray 1969), but geographical distance does not provide direct evidence of the extent to which this reflects breakdown in genetic cohesion or the causative factors of partitioning. Landscapes that appear homogenous, for example by virtue of topology, might encompass ecological heterogeneity and thus evolutionary diversity among the inhabitants.

Different species and groups of species have arisen under different evolutionary scenarios and so have different species-associated properties (de Queiroz 2007). Therefore, species delimitation is not a one-size-fits-all analysis and the number of species delimited depends on the criteria being applied. For example in New Zealand Powell described 10 species and 27 subspecies of *Powelliphanta* land snails in a classification based largely on detailed examination of shell size, colour and pattern (Powell, 1930, 1930, 1932, 1936, 1946, 1949). In contrast, for the same populations Climo (1978) proposed an alternative classification comprising a ring species with only one or two species and five subspecies. The systematics of *Powelliphanta* is currently under review (K Walker unpublished). My examination of fine grain population structure of one species complex within this genus using a combination of mtDNA haplotype data and nuclear markers confirmed that ecological specialisation has lead to fragmented distributions of some lineages. Improved

species delimitation tools will very likely increase the recognised biodiversity of this genus.

Similarly, New Zealand *Placostylus* subspecies described by Powell were subsumed after traditional morphometric analysis and examination of haplotype diversity (Buckley et al. 2011). Improved tools for studying shell shape used here support the morphological diversity recognised by “naturalists [of] sound judgement and wide experience” (Darwin 1859).

Boundaries among some taxa may be fairly obvious but most groups require intense study using a combination of information (morphology, behaviour, distribution, mitochondrial and nuclear genetic data) in order to formulate and test specific hypothesis. In snails, and other molluscs, it is apparent that morphology or genetics (especially single markers) alone may be insufficient to reliably identify taxonomic units and explore population histories; numerous examples exist of variation in shell colour, pattern and shape within a single species (e.g. *Littorina:* Carvajal-Rodríguez et al. 2005; *Partula* Clarke & Murray 1969; *Cerion:* Woodruff & Gould 1980) and examples of snail species differentiated by similar levels of phenotypic variation (e.g. *Mandarina*: Chiba 1998; *Euhadra*: Davison et al. 2005; *Placostylus:* Dowle et al. 2015).

Recent years have seen an explosion of genetic and ecological studies of forms generally classified as infraspecific, but which show some characteristics of species, such as distinct morphological, ecological and/or genetic differences in sympatry, and a degree of reproductive isolation. Some such biotypes may be created by phenotypic plasticity or single-locus polymorphisms (Mallet 2008).

Sympatric populations that can be inferred to show ‘dumbell’ or ‘bimodal’ distributions of genotypes at multiple loci, but which nonetheless are considered members of the same species (Dres & Mallet 2002; Jiggens & Mallet 2000; Kondrashov & Mina 1986). Such forms might be called ecological or genetic races if they form ‘genetically differentiated, sympatric populations between which there is appreciable gene flow’(Mallet 2008). A combination of morphological, ecological

and genetic information is essential if we are to disentangle the evolutionary histories of land snails and provide a sound appraisal of their biodiversity (Lydeard et al. 2004).

To determine the population structure of New Zealand land snails *Placostylus ambagiosus* and *Powelliphanta* Kawatiri complex I used a combination of information. Many New Zealand invertebrate taxa have relied heavily on short sequence mitochondrial markers. A mitochondrial phylogeny tends to be the starting point of many phylogenetic investigations (Avise et al. 1987). mtDNA has been useful for resolving phylogenetic relationships at many taxonomic levels (Simon et al. 1994). The relatively high rate of molecular evolution of the mtDNA genome and the fact that it exists as many copies per cell make this genome especially useful for analysis of species diversity. Mitochondrial haplotype data and microsatellite (nuclear) markers were used to estimate population diversity within the *Powelliphanta* Kawatiri complex. Genetic diversity was found to be partitioned among the fragmented populations and the lack of concordance among markers supported a single genotypic cluster or species complex. Lack of gene flow at nuclear loci (microsatellites) has allowed population differentiation. These results are to be combined with allozyme, morphological and environmental data to inform on the taxonomy of this group and subsequent conservation management plans (K Walker Pers Comm).

Of equal importance to using a combination of molecular markers, is to consider the many aspects of an organism's biology that contribute to current patterns. Morphology has traditionally played a pivotal role in studying animal phylogeny since the first evolutionary biologists began to decipher the tree of life. In recent times, however, morphological characters have lost prominence in inferring deep relationships among animals due to fundamental issues with homology and the choice of higher taxa as terminals, but most importantly, due to declining costs of generating molecular data (Giribet 2015). It is "thus ironic that fossils have become formidable allies of molecules wherever inference of divergence times is required, while molecules have relegated morphological characters of living organisms to an almost marginal role in phylogenetics" (Giribet 2015). Morphological traits are the

outcome of multi-locus variation, and thus constitute a more thorough reflection of variation among individuals than a particular, pre-determined section of the DNA sequence (Will et al. 2005). Furthermore if divergent selection is the main driver of speciation, then adaptive (morphological) traits may provide best insights onto species' limits (Rundle & Nosil 2005; Schluter 2000). Only with an understanding of the context within which genetic patterns arise can evolutionary processes be fully understood. Morphological quantification was essential to allowing us to link extant variation with fossil data. Modern populations of *Placostylus ambagiosus* show morphological differentiation between populations however clustering of morphologically distinct populations is not with geographically adjacent populations suggesting that local environments differ in some influential, but currently unknown, manner within the narrow range of this species. There are also differences in morphology between Cape Maria van Diemen and Tom Bowling Bay fossil populations. These differences in morphology are relatively unchanged through time (five thousand years) and fit a model of stasis for many traits. The fact that the independent traits follow different models from one another in the same lineage and that some of these patterns differ from those inferred for a population of the same species 30 km away provides compelling evidence that gene flow is not limiting all local adaptation nor preventing the formation of ecotypes is. Determining the way in which local environments differ and how this correlates to shell shape change is the next step. Although many traits showed evidence of stasis, more complex models of change in evolutionary mode through time should be examined. It is well known that the time period in which these lineages span includes periods of climate change and increased seasonality. Although the climate record for Northland New Zealand is not well resolved, finding evidence of periods of directional evolution or random walks bounded by stasis within lineages could be compared to not only climate but also local environment inferred from fossil pollen found within shells allowing a reconstruction of paleo-habitat.

Stasis, gene flow and the maintenance of ecotypes

Gastropods are a morphologically diverse group displaying huge differences between species in shell shape variation and examples of species with multiple morphotypes described, however this diversity is not reflected in the fossil record. Several reasons for this general lack of interspecific evolutionary diversity in the fossil record have been proposed (see Chapter 5). Also ecotypes are sometimes classified as different species and not included in lineage studies (Aze et al. 2013; Hills et al. 2012). My research has shown that although it is possible to observe ecotype diversity within a species, when sampling across the range of a species this morphological differentiation is averaged out. However when I sampled several discrete lineages within a species each from a highly restricted geographic area, I observe the same morphological ecotypes in each region through time (stasis).

The demographic structure of a species over evolutionary time periods influences the creative or constraining forces of gene flow (Slatkin 1987). Gene flow is likely to play a conservative role constraining evolution in situations where local populations persist for long time periods over relatively stable geographic distribution because here gene flow is likely to prevent the genetic differentiation of local populations, inhibiting speciation. However where gene flow is infrequent there is the potential for each population to evolve independently. Of course the absence of gene flow does not "trigger" evolution, it simply allows for mutations to accumulate in the population (Slatkin 1987).

Local extinctions and recolonisations of habitats can be an important source of gene flow. Even if there is no exchange of individuals between established populations, there will be little differentiation of local populations due to genetic drift if the average time that populations persists in one area is less than the time is takes for genetic drift to fix neutral alleles. Species in metapopulations are expected to have active and dynamic patterns of local extinction and colonization among habitat patches (Hanski 1998; Levins 1969). For putatively fragmented populations, analysis of genetic structure can reveal the degree of isolation (Hedrick 2005) and the rate of successful migration between populations (Slatkin 1987). Gene flow

between populations of *Placostylus ambagiosus* was common with SNP data showing a single admixed cluster. Lack of Isolation by Distance (IBD) in extant populations suggests that recent gene flow although present at neutral markers is perhaps not as straightforward as initial results imply.

RAD-seq is a relatively new technology and allows for sampling of a fraction of the genome so that many individuals of non-model organisms can be included in population genetic studies. While the excitement surrounding data types like RAD-seq is warranted, the field has transitioned so rapidly that basic assessments of pipeline robustness and consistency have not been fully pursued (Shafer et al. 2016). Within this study I have been able to add to the growing list of examples illustrating the importance of analysing RADseq data with multiple parameter settings.

The dataset created by ddRAD could also be used to quantify directionality and amount of gene flow between populations using the sequence data rather than SNPs and a programme like IMa (Hey & Nielsen 2004, 2007). IMa uses Bayesian inferences to fit a model of Isolation with Migration to haplotype data drawn from a pair (or groups) of closely related populations or species. This would be well suited to our dataset as it requires few individuals per population, but high numbers of loci.

I detected little genetic differentiation among extant *P. ambagiosus* populations so it is possible for gene flow to be responsible for morphological stasis (see chapter 5), however, gene flow has not prevented local differentiation suggesting both stasis and morphological adaptation is possible in the face of gene flow if selection is operating on these traits. Gene flow between populations has not been continuous or straightforward and a history of population extinction, and occasional gene flow between populations is likely. The results of chapter 5 where populations remain in stasis through time is inconsistent with an interpretation of phenotypic plasticity, although the possibility cannot be excluded entirely. Instead, the observed morphological divergence but lack of genetic structure is consistent with strong character specific selection in the presence of high levels of gene flow. With high

levels of gene flow, neutral markers may not detect population structuring (Thibert-Plante & Hendry 2010).

We usually do not class these ecotypes or races as species, but they clearly have many of the same characteristics, including ecological and genetic differences. These are therefore forerunners of species, but high rates of gene flow and likely sensitivity to ecological conditions probably means that the fate of many snail ecological races is extinction.

Gene flow between populations is no longer considered an impediment to speciation but part of the process of speciation (Abbott et al. 2013). Earlier theory predicted that gene flow inhibited divergence by preventing natural selection and genetic drift from maintaining local differentiation (Mayr 1963). However gene flow is neither uncommon nor necessarily an impediment to divergence and natural selection can cause populations to diverge and adapt regardless of species distributions (Mallet 2001; Nosil 2008; Wu 2001). Strong selection on particular traits can produce adaptations to local conditions in the presence of substantial gene flow (Ehrlich & Raven 1969)

This work illustrates not only that ecological races can form despite gene flow but that morphological stasis is not (in this instance) a product of gene flow but rather maintained within populations by constraining selection in the face of it.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246.

Avise, J., Arnold, J., Ball, R., Bermingham, E., Lamb, T., Neigel, J., … Saunders, N. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, *18*(1), 489–522.

Aze, T., Ezard, T. H. G., Purvis, A., Coxall, H. K., Stewart, D. R. M., Wade, B. S., & Pearson, P. N. (2013). Identifying anagenesis and cladogenesis in the fossil record. *Proceedings of the National Academy of Sciences*, *110*(32), E2946–E2946.

Buckley, T., Stringer, I., Gleeson, D., Howitt, R., Attanayake, D., Parrish, R., … Rohan, M. (2011). A revision of the New Zealand *Placostylus* land snails using mitochondrial DNA and shell morphometric analyses, with implications for conservation. *New Zealand Journal of Zoology*, *38*(1), 55–81.

Carvajal-Rodríguez, A., Conde-Padín, P., & Rolán-Alvarez, E. (2005). Decomposing Shell Form into Size and Shape by Geometric Morphometric Methods in Two Sympatric Ecotypes of *Littorina saxatilis*. *Journal of Molluscan Studies*, *71*(4), 313–318.

Chiba, S. (1998). Genetic variation derived from natural gene flow between sympatric species in land snails (*Mandarina*). *Heredity*, *80*(5), 617–623.

Clarke, B., & Murray, J. (1969). Ecological genetics and speciation in land snails of the genus *Partula*. *Biological Journal of the Linnean Society*, *1*(1–2), 31–42.

Climo, F. M. (1978). The *Powelliphanta gilliesi — traversi — hochstetteri — rossiana — lignaria — superba* ring species (Mollusca: Pulmonata). *New Zealand Journal of Zoology*, *5*(2), 289–294.

Darwin, C. (1859). *On the origin of the species by natural selection*.

Davison, A., Chiba, S., Barton, N. H., & Clarke, B. (2005). Speciation and Gene Flow between Snails of Opposite Chirality. *PLOS Biology*, *3*(9), e282.

de Queiroz, K. (2007). Species Concepts and Species Delimitation. *Systematic Biology*, *56*(6), 879–886.

Dowle, E. J., Morgan-Richards, M., Brescia, F., & Trewick, S. A. (2015). Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. *Molecular Ecology*, *24*(16), 4205–4221.

Dres, M., & Mallet, J. (2002). Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *357*, 471–492.

Ehrlich, P. R., & Raven, P. H. (1969). Differentiation of populations. *Science*, *165*, 1228–1232.

Giribet, G. (2015). Morphology should not be forgotten in the era of genomics–a phylogenetic perspective. *Zoologischer Anzeiger - A Journal of Comparative Zoology*, *256*, 96–103.

Hanski, I. (1998). Metapopulation dynamics. *Nature*, *396*, 41–49.

Hedrick, P. W. (2005). A Standardized Genetic Differentiation Measure. *Evolution*, *59*(8), 1633–1638.

Hey, J., & Nielsen, R. (2004). Multilocus Methods for Estimating Population Sizes, Migration Rates and Divergence Time, With Applications to the Divergence of Drosophila pseudoobscura and D. persimilis. *Genetics*, *167*(2), 747–760.

Hey, J., & Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(8), 2785–2790.

Hills, S. F. K., Crampton, J. S., Trewick, S. A., & Morgan-Richards, M. (2012). DNA and Morphology Unite Two Species and 10 Million Year Old Fossils. *PLOS ONE*, *7*(12), e52083.

Jiggens, C., & Mallet, J. (2000). Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution*, *15*(6), 250–255.

Kondrashov, A. S., & Mina, M. V. (1986). Sympatric speciation: when is it possible? *Biological Journal of the Linnean Society*, *27*(3), 201–223.

Levins, R. (1969). Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control. *Bulletin of the Entomological Society of America*, *15*(3), 237–240.

Lydeard, C., Cowie, R. H., Ponder, W. F., Bogan, A. E., Bouchet, P., Clark, S. A., … Thompson, F. G. (2004). The Global Decline of Nonmarine Mollusks. *BioScience*, *54*(4), 321–330.

Mallet, J. (1995). A species definition for the modern synthesis. *Trends in Ecology & Evolution*, *10*(7), 294–299.

Mallet, J. (2001). The speciation revolution. *Journal of Evolutionary Biology*, *14*(6), 887–888.

Mallet, J. (2008). Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1506), 2971–2986.

Mayr, E. (1942). *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. Harvard University Press.

Mayr, E. (1963). *Animal species and evolution*. Cambridge, Massachesetts: Belknap Press of Harvard University Press.

Nosil, P. (2008). Speciation with gene flow could be common. *Molecular Ecology*, *17*(9), 2103–2106.

Powell, A. (1930). The Paraphantidae of New Zealand; their hypothetical ancestry, with descriptions of new species and a new genus. *Records of the Auckland Institute and Museum*, *1*, 17–55.

Powell, A. (1932). The Parypahantidae of New Zealand. Descriptions of further new species. *Records of the Auckalnd Institute and Museum*, *1*(155–162).

Powell, A. (1936). The Paryphantidae of New Zealand, no 3. Furhter new spceis of *Paryphanta* and *Wainuia*. *Records of the Auckalnd Institute and Museum*, *2*, 29–41.

Powell, A. (1946). The Paryphantidae of New Zealand, no 5. Further new species of *Paryphanta*, *Wainuia* and *Rhytida*. *Records of the Auckalnd Institute and Museum*, *3*(99–136).

Powell, A. (1949). The Parayphantidae of New Zealand, no 6. Distribution, hybrids and new species of Paryphantidae, *Rhytida* and *Schizoglossa*. *Records of the Auckalnd Institute and Museum*, *3*(347–372).

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*(3), 336–

Schluter, D. (2000). Ecological Character Displacement in Adaptive Radiation. *The American Naturalist*, *156*(S4), S4–S16.

Shafer, A. B. A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. W. (2016). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, n/a-n/a.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., & Flook, P. (1994). Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers. *Annals of the Entomological Society of America*, *87*(6), 651–701.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*(4803), 787–792.

Thibert-Plante, X., & Hendry, A. P. (2010). When can ecological speciation be detected with neutral loci? *Molecular Ecology*, *19*(11), 2301–2314.

Vaux, F., Trewick, S. A., & Morgan-Richards, M. (2016a). Lineages, splits and divergence challenge whether the terms anagenesis and cladogenesis are necessary. *Biological Journal of the Linnean Society*, *117*(2), 165–176.

Vaux, F., Trewick, S. A., & Morgan-Richards, M. (2016b). Speciation through the looking-glass. *Biological Journal of the Linnean Society*, n/a-n/a.

Will, K. W., Mishler, B. D., & Wheeler, Q. D. (2005). The Perils of DNA Barcoding and the Need for Integrative Taxonomy. *Systematic Biology*, *54*(5), 844–851.

Woodruff, D. S., & Gould, S. J. (1980). Geographic differentiation and speciation in Cerion - a preilminary discussion of patterns and processes. *Biological Journal of the Linnean Society*, (14), 389–416.

Wu, C.-I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, *14*(6), 851–865.