



Conservation of pūpū whakarongotaua - the snail that listens for the war party

Elizabeth E. Daly^{1,2}, Steven A. Trewick¹, Eddy J. Dowle^{1,3}, James S. Crampton⁴ and Mary Morgan-Richards^{1*}

ABSTRACT

Snail ecotypes are an indicator of local forms that can have cultural value and also represent evolutionary potential. Conservation of both regional diversity and evolutionary potential can be improved by recognition and documentation of genetically determined phenotypic variation. The large threatened terrestrial snail pūpū whakarongotau (*Placostylus ambagiosus*) is a range restricted species that is valued by the indigenous people of the northern most region of New Zealand, valued as both a security alarm and food source. Some populations of this rare species are restricted to sites of human occupation and oral history suggests these distinct populations represent prehistoric translocations. We used 2-D geometric morphometrics and nuclear markers to document the distribution of variation. We aimed to determine whether shell shape variation was entirely an ecophenotypic plastic response, and whether or not prehistoric human movement of snails and establishment of food populations on pā sites (fortified settlements) had resulted in the loss of spatial genetic structure within this species. We could discriminate specimens due to shape variation among population samples without reference to their geographic origin. Using samples from prehistoric sites we infer shell shape in this species is not purely a plastic response to the local environment but reveals evolutionary history and potential. Genotypes with 1738 nuclear loci for 19 individuals from ten locations revealed some evidence of population mixing, but genetic variation was mostly partitioned among locations, with strong spatial structure revealed. Thus, we advocate conservation measures that will preserve local forms that represent evolutionary potential within this species.

Keywords: Prehistoric translocations; *Placostylus ambagiosus*; shell shape.

1 Wildlife & Ecology Group, Massey University, Palmerston North, New Zealand

2 Horizons Regional Council, Palmerston North, New Zealand

3 Department of Anatomy, University of Otago, Dunedin, New Zealand

4 School of Geography, Environment & Earth Sciences, Victoria University, Wellington, New Zealand

* Corresponding author. ✉E-mail address: EED (Elizabeth.Daly@horizons.govt.nz), SAT (s.trewick@massey.ac.nz), EJD (eddy.dowle@otago.ac.nz), JSC (james.crampton@vuw.ac.nz), MMR (m.morgan-richards@massey.ac.nz)

Significance Statement

In northern New Zealand the indigenous people (Ngāti Kurī) treasure giant land snails (pūpū whakarongotau; *Plactostylus ambagiosus*) because, in the past, these molluscs were both security alarms and a food source. We studied this endangered species using old material stored in museums and national collections and found shell shape of *P. ambagiosus* can be used to identify which local population a snail came from. By using a population produced from pre-historic cultivation and human-aided dispersal of snails, we determined that shell shape differences are maintained even when the snails of mixed provenance occupy the same environment. These differences record pre-cultivation spatial genetic structure in this species. Shell shape variation thus reveals evolutionary potential and should be conserved by protecting all living populations. We recommend spending money building fences to conserve all the shape forms (subspecies) rather than on more DNA sequences.

INTRODUCTION

Ecotypes are distinct forms of a species occupying particular habitats that are usually considered part of the continuum between populations, varieties and species (Darwin 1859; Mallet 2008). Ecotypes represent evolutionary potential within a lineage (Segura *et al.* 2006; Eilaguirre and Baltazar-Soares 2014; Nichols *et al.* 2016) and biological conservation should aim to preserve ecotypic and phenotypic variation where it is partly or wholly underpinned by allelic variation. Phenotypic variation can also be a valuable resource in species propagated for food or traditional tools and local forms of a species can have cultural and spiritual value (Wehi 2009).

Ecotypic variation is recognized widely among land snails (e.g. Jones *et al.* 1977; Goodfriend 1986; Chiba 2004; Parent and Crespi 2009; Cameron and Cook 2012; Galindo and Grahame 2014), a group that is also characterized by a high incidence of threatened species. Diversity decline of terrestrial molluscs is a growing conservation concern (Lydeard *et al.* 2004; Parkyn and Newell 2013; 2016 IUCN Red list of Threatened Species). Land snails are sensitive to human modification of the environment by forest clearing and

introduction of novel predators. For example, it has been estimated that 65-90% of endemic snail species on Pacific islands have gone extinct since human colonization (Cowie *et al.* 2017). Although variation within snail lineages provides a valuable resource for the study of ecological speciation (e.g. Parent and Crespi 2009; Stankowski 2013), the rapid loss of biodiversity limits our chances of fully understanding many systems. In addition, human-mediated dispersal (movement of food species, conservation translocations, invasive populations) might result in genetic mixing and the loss of natural geographical structure.

Pūpū whakarongotau (*Plactostylus ambagiosus*) is a snail highly valued by tangata whenua Māori (the indigenous people) of northern Aotearoa New Zealand (Te Aupōuri me Ngāti Kurī), because in the past: (1) It was an important food source, (2) Empty shells are used as musical instruments, (3) It provided alarm calls at night warning of approaching invaders. In Māori tradition all flora and fauna are part of the extended family and genealogies (whakapapa) reveal how all people and life forms descend from a common ancestor, so for the local people these snails are whānau (family). For the Ngāti Kurī people of the Far

North of New Zealand the 'song' of *P. ambagiosus* has made it a treasured guardian. The sounds these pūpū (snails) make as they hastily retreat into their shells when disturbed at night once alerted the people to approaching invaders and so saved their lives, and consequently the snails are known as pūpū whakarongotaua (the snail that listens for the war party). Māori ate *Placostylus* species and oral histories recount that snails were moved to establish new *P. ambagiosus* populations (Powell 1947 1979). On a former island, now the north-western part of Cape Maria van Diemen, the extant snails differ from the fossil subspecies at the same site (Powell 1979; Triggs and Sherley 1993), and this lack of continuity between fossil and extant populations is best explained by translocation (Powell 1947). Where snail populations occupy the sites of former pā (fortified settlements on hills and headlands) they are likely to be descendants of translocated snails (Powell 1947 1951).

Snail shell shape represents genetic and environmental influence and can be challenging to quantify (Van Bocxlaer and Schultheiß 2010). However, two-dimensional landmark based geometric morphometric techniques, which use the spatial distribution of biologically meaningful, morphological landmarks to describe the shape of organisms, are superior to traditional linear measurements (Mitteroecker and Gunz 2009; Webster and Sheets 2010; Zelditch *et al.* 2010). Geometric morphometrics methods allow comparison of shape whilst controlling for variation in the size and orientation of objects, and coupled with naïve Gaussian Mixture Models provide an unbiased tool to cluster specimens without a priori assumptions or hypotheses. The simultaneous description of shape variation and identification of the most efficient model

to group specimens strengthens studies of natural phenotypic variation, identification of ecotypes and characterization of evolutionary potential.

Geometric morphometric techniques have been used successfully to identify distinct sympatric species of New Caledonian *Placostylus snails* (Dowle *et al.* 2015; Quenu *et al.* 2019), and are here applied to New Zealand *Placostylus* populations to analyze the spatial distribution of shell shape variation. *Placostylus ambagiosus* is now endangered with 19 small and isolated populations representing a once common forest species (Stringer *et al.* 2014; 2017; Parrish *et al.* 2014). A holistic treatment of shell size and shape differences led to the description of 15 subspecies (Powell 1979), but a single taxon has been proposed largely on the basis of linear measurements and low mtDNA divergence (Buckley *et al.* 2011; NZ Department of Conservation species management report 2015). Shell size of adult snails, which had been used in part to identify subspecies, has now been shown to be plastic (Parrish *et al.* 2014; Stringer *et al.* 2014). Variation in shell shape of *P. ambagiosus* might also represent plastic responses to local and temporal environmental conditions, in which case we would expect low phenotypic variation within a cohort exposed to the same conditions at the same site. Genetically controlled shell shape variation will be detected when shape variation is maintained across different genotypes exposed to the same local conditions, such as would be the case if populations established by translocation combined snails from multiple sources. Localised adaptation to the local environment might result in genetic structure concordant with phenotype clusters, but as shell shape is likely to be under selection and most genetic markers are neutral,

correlations might be weak, and selection can overcome gene flow leading to divergent forms (Slatkin 1987; Stankowski 2013). In addition, human-mediated movement of snails might disrupt population structure resulting in genetically mixed populations that lack spatial structuring.

Here we use model-based clustering to examine population structure in terms of shell shape variation and genotypic variation using a large number of polymorphic nuclear loci. Due to the current rarity of this snail species we have not collected fresh specimens but have made use of samples held in museums and the New Zealand National Arthropod Collection. This approach has the disadvantage that we cannot obtain both morphological and genetic information from the same individuals, but it has allowed us to: 1) Determine whether phenotypic variation in this species is a purely plastic response to local conditions, and 2) Determine whether human mediated dispersal has resulted in loss of spatially structured genetic variation. Both these elements will help identify evolutionary potential and thus inform conservation priorities for this endangered species as well as enhancing cultural understanding.

MATERIAL AND METHODS

Study system

The large herbivorous land snails of the family Bothriembryontidae, subfamily Placostylinae are present on islands in the western Pacific from Papua New Guinea to Fiji and New Zealand (Barker *et al.* 2016). New Zealand has three extant species of genus *Placostylus*, restricted to habitat fragments in the northern tip of North Island and nearby Three King islands. Deposits of fossil shells suggest more widespread

distributions in the past (Brook 1999; Powell 1979; Sherley 1996). *Placostylus ambagiosus* is afforded high conservation status (Nationally Critical) because numbers have declined to fewer than 2000 individuals in total, distributed over 19 remnant populations (Stringer *et al.* 2017). These snails are large as adults (up to 10 cm long), slow growing, with a lifespan of 10-22 years and they show strong site fidelity (Stringer *et al.* 2017). Powell (1979) described 15 subspecies of *P. ambagiosus* among patchily distributed extant populations and fossils noting differences in aperture shape as well as overall shell shape and size. On the basis of shell measurements (that revealed variation but not diagnostic traits) and low mtDNA divergence synonymy of the subspecies has been proposed (Buckley *et al.* 2011); however the names remain valuable when referencing to particular populations (Brook 2002) and as putative ecotypes we use them in this study.

Placostylus ambagiosus material for shell morphometrics was available from seven population samples and for genetic work from ten extant populations (Table 1). Four populations provided both phenotypic and genotypic data. At least three extant populations of *P. ambagiosus* are the result of past translocations by local Māori who moved and maintained a range of edible species so they were readily accessible to their settlements (Powell 1947; 1951; Leach and Stowe 2005; Costall *et al.* 2006; Wehi 2009). Snail populations south-east of Cape Maria van Diemen and at Maungapiko are confined to the sites of former pā (Māori settlements protected by fortifications), where they are found associated with plant species propagated by Māori – harakeke (*Phormium tenax*) and karaka (*Corynocarpus laevigatus*). The population formally recognised as subspecies *P. a.*

annectens occupies an area considerably larger than a pā site (Powell 1947), but its range includes Unuwahao Pā. Our shell shape analysis was confined to adult individuals to avoid shape variation due to allometry (Outomuro and Johansson 2017), but tissue for genetic work was collected from juvenile snails to reduce impact on endangered populations (Triggs and Sherley 1993).

Shell shape

Shell specimens used in this analysis came from the collections of the Museum of New Zealand Te Papa Tongarewa (NMNZ). Shells were collected between 1949 and 1997 from seven sites within 40km of each other in Far North, New Zealand (Figure 1). With one exception these represent extant populations that have formerly been considered different subspecies - here we

designate populations sampled by locality but give subspecies names *sensu* Powell (1947) in parentheses as they may represent unique ecotypes. Two populations sampled for shell analysis are here considered the result of human-mediated dispersal: Maungapiko Pā (*P. a. keenorum*), and south-east Cape Maria van Diemen Pā (*P. a. paraspiritus*), and five population samples are assumed natural: Tirikawa (*P. a. 'nouvelle'*); Unuwahao (*P. a. annectens*); Waikuku (*P. a. watti*); Whareana (*P. a. whareana*), and Surville Cliffs (*P. a. michiei*; Table 1). See Appendix A for further details.

Two-dimensional landmark-based geometric morphometrics was used to characterize shell shape and to provide input shape variables for Gaussian Mixture Model (GMM) clustering analysis. A total of 209 *P. ambagiosus* shells were sampled from seven sites across the species' geographic range (Figure 1A., Table 1). Only intact adult

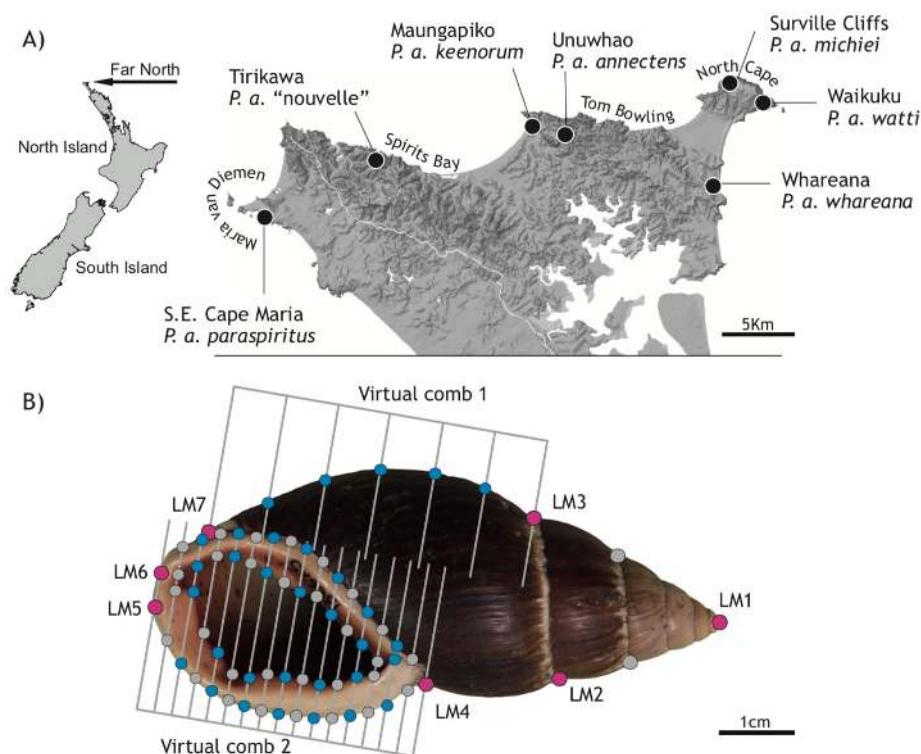


Figure 1. A) Sampling locations in the now fragmented native range of the New Zealand snail *Placostylus ambagiosus*. B) Orientation and placement of landmarks for 2D geometric morphometric analysis of shells: landmarks LM1–7 (magenta), 33 semi-landmarks (blue), and landmarks removed during optimisation phase (grey).

Table 1. Sampling of extant *Placostylus ambagiosus* snails from New Zealand used for morphometric and genetic analyses (n= sample size). Three populations are closely associated with pā sites (not natural). Nine of the populations used by Powell (1947, 1951) to describe subspecies were sampled (type location = yes). Collectors: Greg Shirley (GS, 1990-1997), Richard Prouse (RP, 1949-1950), unknown (WG). Te Papa = National Museum Te Papa Tongarewa registration numbers provided for all material used for geometric morphometrics. Haplotype codes follow Buckley et al. (2011) with the addition of haplotype Z. Nuclear loci (single nucleotide polymorphisms) generated using RADseq. NAC = National Arthropod Collection.

Region	Location	Natural?	Powell subspecies	Subsp. type location?	Collector	Morph. (n)	Te Papa code	mtDNA (n)	SNP loci (n)	NAC code
WEST	Cape Maria van Diemen	Pā site	<i>consobrinus</i>	yes	GS	0	n/a	Y(4), Z	5	GS1-5
	South-east of Cape Maria van Diemen	Pā site	<i>paraspiritus</i>	yes	GS	38	M.136625	X(4)	2	WG860,861
NORTH-WEST	Motuopao Island	yes	<i>ambagiosus*</i>	yes	GS	0	n/a	U(4)	0	n/a
	Pandora (cliff face amongst boulders)	yes	<i>pandora</i>	yes	GS	0	n/a	W, V	1	GS10
	NNW of Pandora	yes	n/a	no	WG	0	n/a	V(2)	1	WG855
	Te Paki	yes	<i>lesleyae</i>	no	GS	0	n/a	T(4)	0	n/a
	Tirikawa (cliff face)	yes	<i>ambagiosus**</i>	no	WG	0	n/a	U(4)	0	n/a
NORTH-EAST	between Tirikawa and Pandora Bay	yes	"nouvelle"	no	GS	21	M.136627	V	2	WG857
	Darkies Ridge	yes	n/a	no	GS	0	n/a	V	0	n/a
	Maungapiko	Pā site	<i>keenorum</i>	yes	RP	32	M.102324	R(4), Y	0	n/a
	Unuwahao	plus Pā	<i>annectens</i>	yes	RP	32	M.102503		0	n/a
	not recorded	n/a	<i>annectens</i>	no	GS	0	n/a	P, T	2	GS17,25
EAST	Matirarau	yes	<i>annectens</i>	no	GS	0	n/a	R, S	2	GS18,20
	Te Huka	yes	<i>annectens</i>	no	GS	0	n/a	P(2)	1	GS19
	Survile Cliffs	yes	<i>michiei</i>	yes	GS	30	M.136621	O(4)	0	n/a
	Waikuku (cliff face)	yes	<i>watti</i>	yes	RP	26	M.102328	N, W(3)	2	GS6,13
	Whareana (in steep valley)	yes	<i>whareana</i>	yes	GS	30	M.136614	M	1	GS8

* = *ambagiosus sensu stricto*, ** = *ambagiosus sensu lato*

snail shells, inferred from the thickening of the aperture lip (Stringer *et al.* 2014), were included. Shells were photographed and digitised as previously described for gastropods (Dowle *et al.* 2015; Vaux *et al.* 2017; Vaux *et al.* 2018) following established procedure (Webster and Sheets 2010; Collins and Gazley 2017). All positioning, photography and subsequent digitisation was conducted by one person to minimise experimental error (Schilthuizen and Haase 2010), which was confirmed to be negligible (Appendix Figure A.1). Shells were photographed with the aperture facing upward and a total of six biologically homologous and 33 sliding semi-landmarks were digitised using tpsUtil, tpsDig (Rohlf 2013) and CoordGen8 (Sheets 2014). Landmarks were selected following a preliminary optimisation process (Figure 1B, Appendix Tables A.1, A.2). Between 21 and 38 adult shells were photographed for each of the seven population samples (Table 1).

Principal components analysis (PCA) implemented in MorphoJ 1.06c (Klingenberg 2011) was used to analyse shell shape. The principal components (PCs) generated by PCA reflect mathematically independent variation in the shape of objects and analysis included all PCAs that exceeded the theoretical scree plot of random, uncorrelated noise using the broken-stick test on Eigenvalues. We sought groupings of shell PCA scores without reference to taxonomy or geographical sampling information using Gaussian Mixture Models implemented by MCLUST v4.0 (Fraley and Raftery 2012) in R (R Core Team 2016). MCLUST uses a model-based approach to infer structure in datasets that do not have a priori labels (Fraley and Raftery 2002; Hastie *et al.* 2009). Shell shape variation is treated as a mixture of population samples with different Gaussian distributions, and the

modelling process divides the data into different clusters, each corresponding to a single Gaussian distribution (Fraley and Raftery 2003) by iterative expectation–maximisation (Fraley and Raftery 2012). MCLUST allows for a total of 90 models (10 different models with various combinations of parameterization and 1 to 9 clusters/groups) to be evaluated with the Bayesian Information Criterion (BIC), which balances improvements to model fit that result from increasing model complexity against the risk of over-parameterizing the model. The optimal model was used to generate assignment plots (probability of shells belonging to each possible cluster) for each individual specimen.

Genetic markers

Tissue obtained for allozyme analysis (Triggs and Sherley 1993) and later used for mtDNA COI studies (Buckley *et al.* 2011), were resampled from the National Arthropod Collection held by Landcare Research in Auckland for double digest restriction associated DNA (RADseq) population genetic analysis (Table 1). Available *P. ambagiosus* cytochrome oxidase I (COI) mitochondrial partial gene sequence (Genbank) were supplemented using polymerase chain reaction with standard primers and conditions (Buckley *et al.* 2011). MtDNA COI sequences were aligned by eye in GENEIOUS v6.0.5 and POPART (Leigh and Bryant 2015) was used to infer a median-joining network (Bandelt *et al.* 1999).

Anonymous single nucleotide polymorphic (SNP) nuclear markers were generated using high-throughput sequencing. The double-digest RADseq protocol was used with minor modifications (Peterson *et al.* 2012; Dowle *et al.* 2015). Restriction endonucleases PstI and BamHI

were used to digest whole-genomic DNA from 54 individual snails, with DNA fragments ligated (Invitrogen T4 ligase) with short DNA barcode sequences to enable identification after sequencing. DNA was then pooled, indexed and data were generated on an Illumina Hi-Seq platform (New Zealand Genomics Ltd).

Although new tools using high-throughput DNA sequences (such as RADseq) provide greatly improved opportunities to study non-model organisms (Andrews and Luikart 2014) processing DNA reads is sensitive to several study-specific parameters (Catchen *et al.* 2013; Huang and Knowles 2016; Harvey *et al.* 2015). In the present study variation in data quality was amplified by the endangered status of the snail constraining the sampling strategy. Post hoc approaches were used to optimise parameter settings and sample inclusion using the STACKS 1.01 pipeline (Catchen *et al.* 2013; Catchen *et al.* 2011). Selection of nuclear markers targeted likely single copy loci for which the maximum number of individuals could be genotyped (Harvey *et al.* 2015; Languth *et al.* 2012). Parameter optimisation (described in Appendix B) resulted in a dataset comprising 19 individuals (Table 1).

Here, we report analysis using a minimum of eight reads per individual (-m) as providing a reliable set of markers for downstream analysis. Within an individual, we allowed a maximum of four mismatches between alleles (-M) and six mismatches between primary and secondary reads (-N=M+2) within USTACKS. In CSTACKS, 10 mismatches were allowed between alleles in different individuals (-n) when generating the SNP set. 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. Our settings required that each putative locus included was genotyped in individuals from three or more of the five populations (-p 3) and genotyped in $\geq 90\%$ of individuals within those populations (-r 0.9). The minor allele frequency was set at 0.1 removing rare alleles that occurred only in a single individual (final N=19) 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. Two smaller, more stringent data sets were also examined. One required that a locus be detected in all populations for inclusion (-p 5), regardless of the proportion of individuals genotyped (-r 0.0). The other required that $\geq 90\%$ of individuals be genotyped (-r 0.9) at each SNP locus in at least one population sample before being recorded (-p 1).

Detection of genetic structuring

Evidence of isolation by distance (IBD) was sought using a mantel test of the correlation of pairwise linear geographic distances and pairwise F_{ST} (SNP data) 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).

Population genetic structure based on nuclear genotypes of 19 *P. 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 30000 followed by 100000 generations, with the number of groups (K) set from one to seven. 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 100000 followed by 100000 generations (10 replications), for numbers of groups (K) ranging from one to seven. Results from STRUCTURE were examined using STRUCTURE HARVESTER (Earl and vonHoldt 2011), before being averaged across the 10 replications using CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007). Output genotype structure graphs were then regenerated in DISTRUCT 1.1 (Rosenberg 2003).

Nuclear genetic variation was used to generate PCA plots based on the SNP dataset. This is a model-free method that uses eigen decomposition to visualise underlying population structure. To avoid high levels of missing data, which can have the effect of drawing samples towards the middle (mean) of a PC plot, we included loci only where we knew genotypes for a minimum of 50%, 70%, 80% or 90% of individuals. Here we present the results from a minimum coverage of 90% of individuals. Many loci were by chance not represented in the sequence sets from individual snails, thus the total number of loci with only 0–10% missing was 54. The PCA analysis was implemented via the adegenet package in R v. 3.1.2 (R Core Team, 2016; Jombart 2008) and PCA plots were drawn using the ggplot2 package (Wickham 2016). The grouping of individuals that were from the same location or adjacent locations suggest there is spatial genetic structure.

RESULTS AND DISCUSSION

Shell shape

Two hundred and nine shells representing seven extant population samples of *Placostylus ambagiosus* were analysed. Photographing and digitising error was found to be negligible and well within the

population-based variation (<2.3% of total variance; Appendix A). Shape variation of *P. ambagiosus* shells from extant populations resolved six statistically informative principal components (combined, 81% of the variation explained). As expected of a single species, population samples overlapped in morphospace, although their means were distinct (Figure 2).

Linearly uncorrelated shape variables (PCs) were used in Gaussian Mixture Model assignment analyses to cluster specimens without a priori labels. When all six significant PCs of shape variation were included for the seven populations, the data were best described by a model with five distinct clusters, corresponding to sampling locations (Table 2; Figure 3). Five population samples were largely confined to their own distinct cluster with 83–95% of individuals from a sample given the same group assignment: Surville, Tirikawa, Whareana, Maungapiko, and Waikuku (Table 2). In contrast, 75% of shells from Unuwaho were grouped with the population sample from Waikuku, and seven Unuwaho shells grouped with the population from Maungapiko. The sample from Maungapiko Pā, separated by only ~3km from Unuwaho, is morphologically distinct (91% of shells from Maungapiko Pā correctly assigned to their own unique cluster, *P. a. keenorum*). Lastly, the population sample from the pā at SE Cape Maria van Diemen was dispersed amongst four different clusters (Table 2).

When only the three largest components of shape variation (PC1, 2, 3; 63%) were used to cluster shells, the variation was best described by a model with three distinct clusters (Table 2; Appendix Figure A.2). Again, the SE Cape Maria van Diemen Pā sample had individuals assigned to all three clusters, contrasting with the six other population samples where most individuals

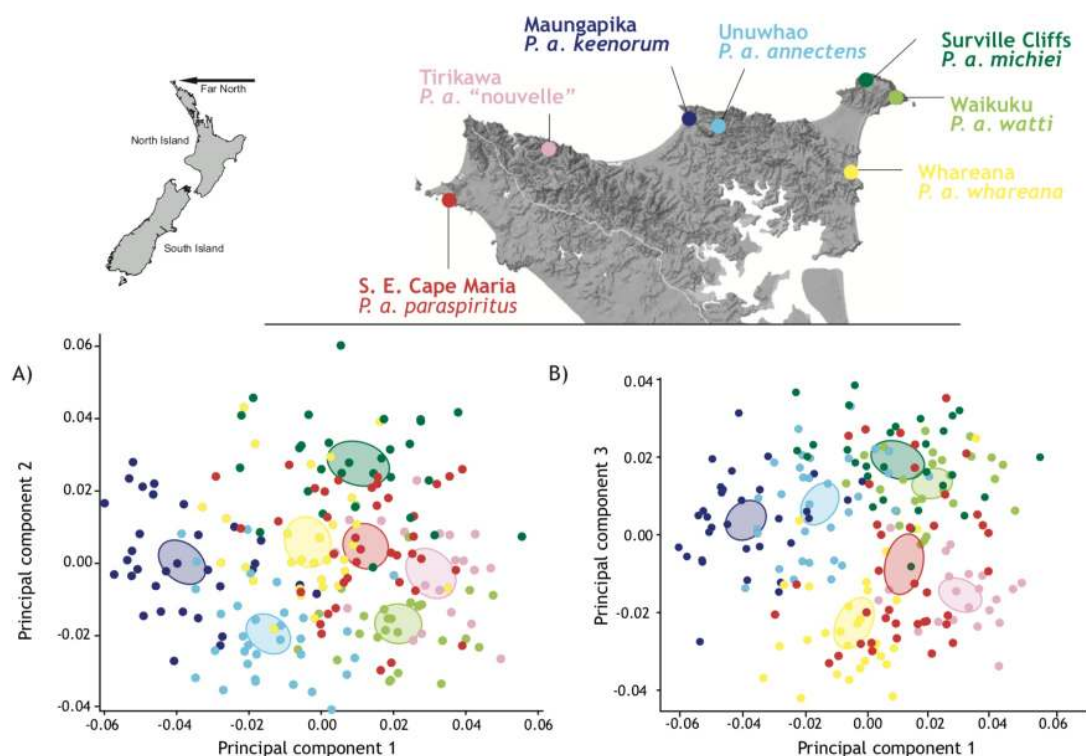


Figure 2. Sampling of shell shape variation of the endangered snail *Placostylus ambagiosus* in Far North, New Zealand. Population sample locations and specimens are colour coded in the analysis using principal components of variation from geometric morphometric data. A) Principal components one and two (47% of shape variation). B) Principal components one and three (46% of shape variation) with 95% confidence ellipses around sample means.

from the same location clustered together (84%–97% of the sample; Table 2). The grouping of population samples into three clusters was not concordant with geographical proximity (see appendix Figure A.2).

Genetic markers

We resolved 13 unique mtDNA COI haplotypes among *P. ambagiosus* individuals. Although most haplotypes were observed in more than one population sample, six (M, N, O, S, X, Z) were found in only one population sample each (Table 1), and evidence of geographic partitioning was revealed in the haplotype network (Figure 3). Haplotypes resolve into two main clusters associated predominantly with the eastern

and western parts of the snail range (Fig. 3A). Three haplotypes (T, W, Y) were found in individuals from non-adjacent locations with haplotypes found in western population samples also found in two eastern population samples, including Maungapiko Pā and Waikuku population samples (Table 1).

Although DNA extractions from 54 *P. ambagiosus* museum specimens were initially processed, down-stream analysis of RADseq SNPs was restricted to 19 individuals with high quality DNA that provided good coverage of loci. Using the optimal settings inferred from extensive parameter trials (see Appendix Table B.2) 1738 single instance SNPs were suitable for genotype analysis. Strong evidence of genetic partitioning among the population

samples was found (overall $F_{ST} = 0.7369$). Some genetic variation was partitioned among sample locations with all pairwise F_{ST} estimates significantly greater than zero. However, SNP variation did not follow a model of isolation by distance (Mantel test: $r = -0.5152$; $P = 0.0747$).

Clustering and assignment of genotypes (using STRUCTURE with 1738 loci) found the optimal number of genetic clusters within the sample was two ($K = 2$ had strongest support using deltaK; see appendix Figure A3). These two genetic clusters correlate with samples from eastern and western parts of their distribution. Most individuals (15/19) had high assignment probabilities for a single cluster (>0.98) in the optimal model, but four individuals from the eastern part of the range (North Cape and Tom Bowling Bay) provided evidence of genetic mixing as their assignment probabilities were < 0.86 . In

general, the nuclear data is concordant with the mtDNA haplotypes, dividing the sample into eastern and western parts of the species range. One *P. a. annectens* snail (PS205=GS25) with haplotype T characteristic of snails from the western part of the distribution (Spirits Bay; Figure 1), unlike adjacent snails, also clustered with western snails based on nuclear alleles. The nuclear SNP dataset was reduced in size to examine the genetic signal when only loci represented in 90% or more of the specimens were considered. With this criterion fewer independent nuclear loci ($n=54$) were scored for 19 individuals, but these loci provided evidence of the same genetic structure (Figure 3C). The first PC of genetic variation explains 57.28% of variation and separates samples from the eastern part of the range from the western samples. Principal component 2 (22.36% of

Table 2. Shape variation in seven population samples of *Placostylus ambagiosus* snail shells from New Zealand allowed specimens to be assigned to groups using model-based clustering without a priori labels; the body of the table indicates the numbers of individuals from each sample that are assigned to each cluster. Optimal Gaussian mixture model unsupervised analysis resolved five distinct phenotypes using six principal components of variation ($K = 5$; PCs 1–6) and three phenotype clusters with three principal components of variation ($K = 3$; PCs 1–3).

Population sample	<i>P. ambagiosus</i> subspecies	PCs		1–6			Clusters (K=5)	n	# incorrect	correct assignment
		1	2	3	4	5				
Surville Cliffs	<i>michiei</i>	28	1				1	30	2	0.933
S.E. Cape Maria	<i>paraspiritus</i>	11	16	3			8	38	na	na
Tirikawa	"nouvelle"		1				20	21	1	0.952
Whareana	<i>whareana</i>	2	25	2			1	30	5	0.833
Unuwahao	<i>annectens</i>		1	24	7			32	8	0.75
Maungapiko	<i>keenorum</i>		1	2	29			32	3	0.906
Waikuku	<i>watti</i>	2		24				26	2	0.92

Population sample	<i>P. ambagiosus</i> subspecies	PCs		1–3		Clusters (K=3)	n	# incorrect	correct assignment
		A	B	C					
Surville Cliffs	<i>michiei</i>	27	2	1			30	3	0.871
S.E. Cape Maria	<i>paraspiritus</i>	11	23	4			38	na	na
Tirikawa	"nouvelle"	1	20				21	1	0.952
Whareana	<i>whareana</i>	2	27	1			30	3	0.9
Unuwahao	<i>annectens</i>		1	31			32	1	0.967
Maungapiko	<i>keenorum</i>		1	31			32	1	0.969
Waikuku	<i>watti</i>	2	2	22			26	4	0.84

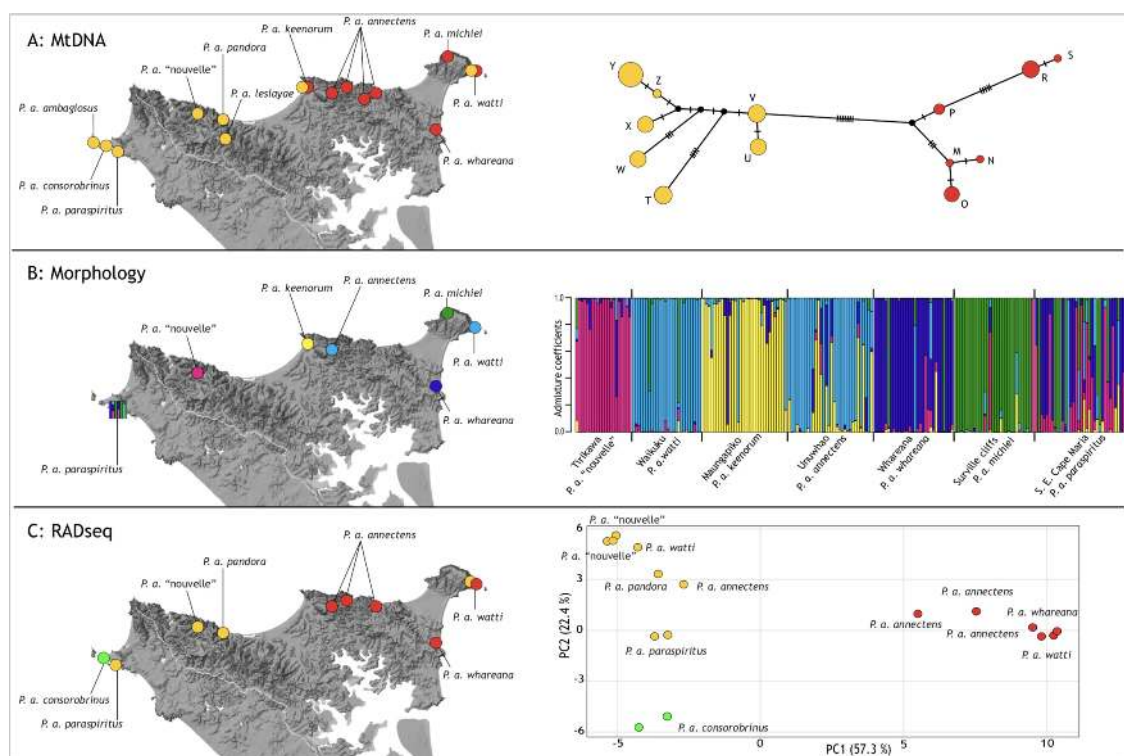


Figure 3. Spatial structure of variation within the endangered snail *Placostylus ambagiosus* in Far North New Zealand. A. mtDNA COI haplotype network with location of samples coloured by haplotype cluster. B. Assignment probabilities of adult shell shape variation using model-based clustering without a priori labels, the optimal Gaussian mixture model resolved five distinct phenotypes using six principal components of variation (K = 5; PCs 1–6), collecting locations coloured by predominate shape cluster. C. PCA of genotypes using 54 nuclear SNP loci.

variation) separates the five Cape Maria van Diemen *P. a. consobrinus* samples from all others.

Detecting population differences

Our geometric morphometric analyses clearly identify statistically significant and biologically interesting shell shape variation, in contrast to the tools previously employed that failed to resolve population differentiation (Buckley *et al.* 2011). The strength of the technique used here is that shape is studied without the confounding influence of size and the inclusion of semi-landmarks allows analysis of the curves that are integral to snail shell shape (Van Bocxlaer and Schultheiß 2010; Vaux *et al.*

2017; Quenu *et al.* 2019). Use of a naïve clustering approach that avoids imposing a priori assumptions about groups has revealed that meaningful shell shape variation does exist among these snails. Strikingly, shell shape variation of seven population samples of *P. ambagiosus* resolved five clusters broadly concordant with Powell’s taxonomy that recognized six subspecies within our sampling (Powell, 1951 1979). Importantly, these clusters were resolved without reference to collecting location. We have not sampled all previously recognised subspecies of *P. ambagiosus* but our shell shape results presented here suggest that using the taxon names of Powell (1947, 1951) provides a valuable way to ensure evolutionary potential is conserved

when setting management priorities (Parrish *et al.* 1995).

Our genetic data relied on tissue samples from juvenile specimens and therefore are not associated with shell morphometric data. Two geographic clusters of snails are seen in the mtDNA data; west and east (Buckley *et al.* 2011), but superimposed on this pattern is sharing of haplotypes between Spirits Bay and North Cape (Waikuku *P. a. watti*) and between two samples collected from old pā sites (S. E. Cape Maria van Diemen and Maungapiko). Nuclear genotypic variation also resolved east/west partitioning of diversity. The nuclear SNP data suggest genetic structure existed before humans moved snails. Although populations of *P. ambagiosus* have been subject to admixing, this has not resulted in complete homogenising of spatial genetic structure.

Translocation of snails

Preserving variation that partly or entirely reflects expressed genetic differences can ensure conservation of ecotypes and evolutionary potential. Therefore, it is important to differentiate developmental plasticity in response to local conditions in shell shape from differences arising from genetic differentiation. Translocation of snails can have the effect of combining populations that were otherwise spatially isolated, and so provide a convenient common garden test of phenotypic plasticity. Where morphology such as shell size reflects developmental plasticity, variation that differentiated ancestral populations would be lost in the first generation. However, if phenotypic difference is genetically controlled, significant morphological variation within a mixed population sample from the same

environment would be expected. If the individuals interbreed, then over a number of generations the discrete forms would be expected to be lost.

Māori propagated food species close to their settlements (Joseph Banks as cited by Beaglehole 1962; Wehi 2009) including land snails (Powell 1979). Prehistoric snail translocation probably explains morphological mismatch between fossils spanning the last few thousand years and extant populations at some locations. Three *P. ambagiosus* population samples that provided geometric morphometric data were likely the result of prehistoric translocation to established pā sites between 1200 and 1800AD (Powell 1949; Triggs and Sherley 1993). Snails sampled from the population at the pā site south-east of Cape Maria van Diemen were assigned to four different morphological clusters (Figure 3). Our population sample (Museum code: M. 136625) was collected in 1990 by a biologist who could readily distinguish the subspecies *P. a. paraspiritus* from at least four other subspecies (Sherley 1996) following Powell's (1951, 1979) descriptions. Therefore, we are confident that these shells were collected from the same site and represent snails exposed to the same environment. Two (or more) distinct shell phenotypes within the same environment is unlikely if shell shape variation is entirely plastic, and so this indicates that Māori brought together two or more genetically distinct forms. Thus, we infer that much of the shell shape variation within *P. ambagiosus* must have a genetic component as it cannot be explained simply from developmental plasticity responding to environmental stimuli.

Ecotypes arise through local adaptation to local environments; a process not necessarily requiring many generations (Rundle and Nosil 2005; Hendry *et al.* 2007;

Räsänen and Hendry 2008). A link between shell shape variation and environment has been demonstrated in genetically distinct New Caledonian *Placostylus* lineages (Dowle *et al.* 2015), suggesting ecotypes in this genus readily evolve in response to the local environment. Higher moisture levels and/or more sheltered sites in the Far North of New Zealand are occupied by larger snails compared to dry or coastal sites (Powell 1947 1951; Penniket 1981; Buckley *et al.* 2011) but the size of adults has been shown to be a plastic trait (Parrish *et al.* 2014; Stringer *et al.* 2014). In contrast, shell shape variation described here is independent of shell size and is resolved in specimens that have been translocated to the same environment. Estimates of dispersal for this species suggest that individuals seldom move more than a few metres from where they hatch, are long-lived (10–22 years), and show strong site fidelity (Parrish *et al.* 2014; Stringer *et al.* 2017). These traits, in conjunction with habitat fragmentation, could lead to isolated populations and random shape differentiation. Our observations, therefore, could be explained either by local selection and adaptation to local environments, or by genetic drift.

CONCLUSIONS

The original forest habitat of *P. ambagiosus* on northern Aupouri Peninsula has been severely depleted and fragmented by Māori and European occupation, such that native coastal forest now covers only ~3% of the area (Lux *et al.* 2009). All pūpū whakarongotau populations are highly restricted in their distributions and considered remnants of what were once much larger populations (Barker *et al.* 2016). The combination of loss of habitat,

introduced predators and competitors threaten this species with extinction. Conservation requires identification of management units and although snail shells can often be a source of complex and intriguing diversity, we have demonstrated that documenting the variation is valuable. We encourage conservation of all shape variation within *P. ambagiosus* because some elements of shell shape variation are not simply developmental ecoplasticity and human movement of snails has not resulted in a complete loss of spatial structure of genetic and shape variation. Therefore, shell shape variations (as recorded in the subspecies described by Powell) are suitable proxies for the evolutionary potential contained within this species. This is evolutionary potential that may be beneficial to this endangered species in the face of global warming, invasive species and human modification of local habitats. Importantly, local forms of a species can have cultural and spiritual significance and provide evidence to support oral traditions of propagation and trading. Rather than spending resources on further genetic studies and their publication we advocate investing in predator control, fence building and habitat restoration to extend kaitiakitanga to all extant *P. ambagiosus* shape and genetic diversity, the land they occupy and the mana whenua.

ACKNOWLEDGEMENTS

Snail shells were collected by Greg Sherley and Richard Prouse. Bruce Marshall provided access to specimens from the National Museum Te Papa Tongarewa. Tissue samples were collected by Sue Triggs and her co-workers. The NZ National Arthropod collection (at Landcare Research) stored tissue samples. Mathieu Quenu ran

some clustering models for us. Snail expertise and advice was provided by Sheridan Waitai (Ngāti Kurī) and Alan Beu, Kath Walker, Kerry Walton, Bruce Marshall and Ian Stringer. The manuscript was improved by anonymous review.

Funding: This work was supported by a grant from the Royal Society of New Zealand Te Apārangi Marsden Fund 12-MAU-008, and a grant from Massey University.

DATA AVAILABILITY

The geometric morphometric and genetic data used to support the findings of this study are available at http://evolves.massey.ac.nz/DNA_Toolkit.htm.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

CONTRIBUTION STATEMENT

Conceived of the research idea: EED, MMR, JSC, SAT.

Collected the data: EED, EJD, MMR.

Carried out the data analyses: EED, EJD, SAT.

Wrote the first draft of the manuscript: EED.

Review and editing of the manuscript: MMR, SAT, JSC, EED, EJD.

Obtained funding: MMR, JSC, SAT.

REFERENCES

Andrews KR, Luikart G (2014) **Recent novel approaches for population genomics data analysis.** *Molecular Ecology* 23:1661–1667. <https://doi.org/10.1111/mec.12686>

Bandelt HJ, Forster P, Rohlf A (1999) **Median-joining networks for inferring intraspecific phylogenies.** *Molecular Biology and Evolution* 16:37–48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036>

Barker GM, 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: 203. <https://doi.org/10.1071/PC14929>

Beaglehole JC (1962) **The Endeavour journal of Joseph Banks: 1768-1771. Vol. 1.** The Trustees of the Public Library of New South Wales in association with Angus and Robertson, Sydney, Australia.

Brook FJ (1999) **Stratigraphy, landsnail faunas, and paleoenvironmental history of coastal dunefields at Te Werahi, northernmost New Zealand.** *Journal of the Royal Society of New Zealand*, 29:361-393. DOI: 10.1080/03014223.1999.9517603

Brook FJ (2002) **Uncommon and rare landsnails in the Northland region of New Zealand, and an assessment of conservation management priorities.** Whangarei, New Zealand, Department of Conservation.

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* 38:55–81. <https://doi.org/10.1080/03014223.2010.527997>

Cameron RAD, Cook LM (2012) **Habitat and the shell polymorphism of *Cepaea nemoralis* (L.): interrogating the Evolution Megalab database.** *Journal of Molluscan Studies* 78:179–184. <https://doi.org/10.1093/mollus/eyr052>

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait, JH (2011) **Stacks: building and genotyping loci de novo from short-read sequences.** *G3: Genes, Genomes, Genetics* 1:171–182. <https://doi.org/10.1534/g3.111.000240>

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) **Stacks: an analysis tool set for population genomics.** *Molecular Ecology*, 22:3124–3140. <https://doi.org/10.1111/mec.12354>

- Chiba S (2004) **Ecological and morphological patterns in communities of land snails of the genus *Mandarina* from the Bonin Islands.** *Journal of Evolutionary Biology* 17:131–143. <https://doi.org/10.1046/j.1420-9101.2004.00639.x>
- Collins KS, Gazley MF (2017) **Does my posterior look big in this? The effect of photographic distortion on morphometric analyses.** *Paleobiology* 43:508-520. DOI: 10.1017/pab.2016.48
- Costall JA, Carter RJ, Shimada Y, Anthony D, Rapson GL (2006) **The endemic tree *Corynocarpus laevigatus* (karaka) as a weedy invader in forest remnants of southern North Island, New Zealand.** *New Zealand Journal of Botany* 44: 5–22.
- Cowie RH, Régnier C, Fontaine B, Bouchet P (2017) **Measuring the Sixth Extinction: what do mollusks tell us?** *The Nautilus* 113:3-41.
- Darwin C (1859) **On the origin of the species by natural selection.** John Murray, London, UK
- Dowle EJ, Morgan-Richards M, Brescia F, Trewick SA (2015) **Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails.** *Molecular Ecology* 24:4205–4221. <https://doi.org/10.1111/mec.13302>
- Earl DA, vonHoldt BM (2011) **STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.** *Conservation Genetics Resources* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Eilaguire C, Baltazar-Soares M (2014) **Evolutionary conservation-evaluating the adaptive potential of species.** *Evolutionary Application* 7:963-967. <https://doi.org/10.1111/eva.12227>
- Fraley C, Raftery AE (2002) **Model-based clustering, discriminant analysis and density estimation.** *Journal of the American Statistical Association* 97:611-631.
- Fraley C, Raftery AE (2003) **Enhanced model-based clustering, density estimation, and discriminant analysis software: MCLUST.** *Journal of Classification* 20:263–286.
- Fraley C, Raftery AE (2012) **Mclust version 4 for R: normal mixture modelling for model-based clustering, classification, and density estimation.** Technical Report 597, University of Washington, USA
- Galindo J, Grahame JW (2014) **Ecological Speciation and the Intertidal Snail *Littorina saxatilis*.** *Advances in Ecology* e239251. <https://doi.org/10.1155/2014/239251>
- Goodfriend GA (1986) **Variation in land-snail shell form and size and its causes: a review.** *Systematic Zoology* 35:204-223.
- Harvey MG, Judy CD, Seeholzer GF, Maley JM, Graves GR, Brumfield RT (2015) **Similarity thresholds used in DNA sequence assembly from short reads can reduce the comparability of population histories across species.** *PeerJ*, 3: e895. <https://doi.org/10.7717/peerj.895>
- Hastie T, Tibshirani R, Friedman J (2009) **The Elements of Statistical Learning: Data Mining, Inference, and Prediction.** Springer Series in Statistics. Springer-Verlag, New York, USA
- Hendry AP, Nosil P, Rieseberg LH (2007) **The speed of ecological speciation.** *Functional Ecology* 21:455–464. <https://doi.org/10.1111/j.1365-2435.2007.01240.x>
- Huang H, Knowles LL (2016) **Unforeseen Consequences of Excluding Missing Data from Next-Generation Sequences: Simulation Study of RAD Sequences.** *Systematic Biology* 65:357-365 <https://doi.org/10.1093/sysbio/syu046>
- Jakobsson M, Rosenberg NA (2007) **CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure.** *Bioinformatics* 23:1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Jombart T (2008) **Adegenet: a R package for the multivariate analysis of genetic markers.** *Bioinformatics* 24:1403–1405.
- Jones JS, Leith BH, Rawlings P (1977) **Polymorphism in *Cepaea* – a problem with too many solutions?** *Annual Review of Ecology and Systematics* 8:109–143.
- Jensen JL, Bohonak AJ, Kelley ST (2005) **Isolation by distance, web service.** *BMC Genetics* 6:13. <https://doi.org/10.1186/1471-2156-6-13>
- Klingenberg CP (2011) **MorphoJ: an integrated software package for geometric morphometrics.** *Molecular Ecology Resources* 11:353–357. <https://doi.org/10.1111/j.1755-0998.2010.02924.x>

- Landguth EL, Fedy BC, Oyler-McCance SJ, Garey AL, Emel SL, Mumma M, Cushman SA (2012) **Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern.** *Molecular Ecology Resources* 12:276–284. <https://doi.org/10.1111/j.1755-0998.2011.03077.x>
- Leach H, Stowe C (2005) **Oceanic arboriculture at the margins – the case of the karaka (*Corynocarpus laevigatus*) in Aotearoa.** *The Journal of the Polynesian Society* 114:7-27.
- Leigh JW, Bryant D (2015) **Popart: full-feature software for haplotype network construction.** *Methods in Ecology and Evolution* 6:1110–1116. <https://doi.org/10.1111/2041-210X.12410>
- Lux J, Holland W, Rate S, Beadel S (2009) **Natural areas of Te Paki Ecological District.** Department of Conservation, Whangarei NZ.
- Lydeard C, Cowie RH, Ponder WF, Bogan AE, Bouchet P, Clark SA, Cummings KS, Frest TJ, Tompson FG (2004) **The global decline of nonmarine mollusks.** *BioScience* 54:321-330.
- 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* 363:2971–2986. <https://doi.org/10.1098/rstb.2008.0081>
- Mitteroecker P, Gunz P (2009) **Advances in Geometric morphometrics.** *Evolutionary Biology* 36:235–247.
- Nichols KM, Kozfkay CC, Narum SR (2016) **Genomic signatures among *Oncorhynchus nerka* ecotypes to inform conservation and management of endangered Sockeye Salmon.** *Evolutionary Applications* 9:1285-1300.
- Outomuro D, Johansson F (2017) **A potential pitfall in studies of biological shape: does size matter?** *Journal of Animal Ecology* 86:1447–1457.
- Parent CE, Crespi BJ (2009) **Ecological opportunity in adaptive radiation of Galápagos endemic land snails.** *The American Naturalist* 174:898-905
- Parkyn J, Newell DA (2013) **Australian land snails: a review of ecological research and conservation approaches.** *Molluscan Research* 33:116-129 doi.org/10.1080/13235818.2013.782793.
- Parrish R, Sherley G, Aviss M (1995) **Giant land snail recovery plan (*Placostylus* spp., *Paryphanta* sp.).** Threatened Species Recovery Plan 13. 40 p. Department of Conservation, Wellington, New Zealand.
- Parrish GR, Stringer IAN, Sherley GH (2014) **The biology of *Placostylus ambagiosus* (Pulmonata: Bulimulidae) in New Zealand: Part 1. Behaviour, habitat use, abundance, site fidelity, homing and the dimensions of eggs and snails.** *Molluscan Research* 34:139-154.
- Penniket ASW (1981) **Population studies of land snails of the genus *Placostylus* in the north of New Zealand.** Unpublished MSc thesis. Auckland University, Auckland, NZ.
- Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) **Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species.** *PLOS ONE* 7:e37135. <https://doi.org/10.1371/journal.pone.0037135>
- Powell AWB (1947) **Distribution of *Placostylus* Land Snails in Northern most New Zealand.** *Records of the Auckland Institute and Museum* 3:173–188.
- Powell AWB (1951) **On Further Colonies of *Placostylus* Land Snails from Northernmost New Zealand.** *Records of the Auckland Institute and Museum* 4: 134–140.
- Powell AWB (1979) **New Zealand Mollusca: marine, land, and freshwater shells.** Collins, Hong Kong.
- Pritchard JK, Stephens M, Donnelly P (2000) **Inference of population structure using multilocus genotype data.** *Genetics* 155:945–959.
- Quenu M, Trewick SA, Brescia F, Morgan-Richards M (2020) **Geometric morphometrics and machine learning challenge the division of *Placostylus* snails from the Isle of Pines, New Caledonia into just two species.** *Journal of Molluscan Studies*. In Press.
- R Core Team (2016) **R: A Language and Environment for Statistical Computing.** Version: 3.2.5. R Foundation for Statistical Computing, Vienna, Austria.
- Räsänen K, Hendry AP (2008) **Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification.** *Ecology Letters* 11:624–636. <https://doi.org/10.1111/j.1461-0248.2008.01176.x>

- Rohlf F (2013) tpsUtil 1.58 and tpsDig 2.17. (Version 2.17) **Department of Ecology and Evolution, State University of New York.** Retrieved from <http://life.bio.sunysb.edu/morph/>
- Rosenberg NA (2003) **distruct: a program for the graphical display of population structure.** *Molecular Ecology Notes* 4:137–138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Rundle HD, Nosil P (2005) **Ecological speciation.** *Ecology Letters* 8:336–352. <https://doi.org/10.1111/j.1461-0248.2004.00715.x>
- Schilthuizen M, Haase M (2010) **Disentangling true shape differences and experimenter bias: are dextral and sinistral snail shells exact mirror images?** *Journal of Zoology (London)* 282:191–200. <https://doi.org/10.1111/j.1469-7998.2010.00729.x>
- Segura I, Rocha-Olivares A, Flores-Ramírez S, Rojas-Bracho L (2006) **Conservation implications of the genetic and ecological distinction of *Tursiops truncatus* ecotypes in the Gulf of California.** *Biological Conservation* 133:336–346.
- Sheets HD (2014) **Integrated Morphometrics Package (IMP)8.** Retrieved from <http://www3.canisius.edu/~sheets/morphsoft.html>
- 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* 23:73–82. <https://doi.org/10.1080/03014223.1996.9518067>
- Slatkin M (1987) **Gene flow and the geographic structure of natural populations.** *Science*, 236:787–792.
- Slatkin M (1993) **Isolation by Distance in Equilibrium and Non-Equilibrium Populations.** *Evolution* 47:264–279. <https://doi.org/10.2307/2410134>.
- 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* 22:2726–2741. <https://doi.org/10.1111/mec.12287>
- Stringer IAN, Parrish GR, Sherley GH, Mackenzie DI (2014) **The biology of *Placostylus ambagiosus* (Pulmonata: Bulimulidae) in New Zealand: Part 2. Population changes, growth, mortality and life expectancy.** *Molluscan Research* 34:155–175.
- Stringer IAN, Parrish GR, Sherley GH (2017) **Homing, dispersal and mortality after translocation of long-lived land snails *Placostylus ambagiosus* and *P. hongii* (Gastropoda: Bothriembryontidae) in New Zealand.** *Molluscan Research* 38. <https://doi.org/10.1080/13235818.2017.1323368>
- Triggs SJ, Sherley GH (1993) **Allozyme genetic diversity in *Placostylus* land snails and implications for conservation.** *New Zealand Journal of Zoology* 20:19–33. <https://doi.org/10.1080/03014223.1993.10423239>
- 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* 36:497–515.
- Vaux F, Crampton JS, Marshall BA, Trewick SA, Morgan-Richards M (2017) **Geometric morphometric analysis reveals that the shells of male and female siphon whelks *Penion chathamensis* are the same size and shape.** *Molluscan Research* 37:194–201. <https://doi.org/10.1080/13235818.2017.1279474>
- Vaux F, Trewick SA, Crampton JS, Marshall BA, Beu AG, Hills SFK, Morgan-Richards M (2018) **Evolutionary lineages of marine snails identified using molecular phylogenetics and geometric morphometric analysis of shells.** *Molecular Phylogenetics and Evolution* 127:626–637. <https://doi.org/10.1016/j.ympev.2018.06.009>
- Webster M, Sheets HD (2010) **A practical introduction to landmark-based geometric morphometrics.** *Quantitative Methods in Paleobiology* 16:168–188.
- Wehi PM (2009) **Indigenous ancestral sayings contribute to modern conservation partnerships: examples using *Phormium tenax*.** *Ecological Application* 19:267–275. <https://doi.org/10.1890/07-1693.1>
- Wickham H (2016). **ggplot2: Elegant Graphics for Data Analysis.** Springer-Verlag New York, USA. <https://ggplot2.tidyverse.org>.
- Zelditch ML, Swiderski DL, Sheets HD, Fink WL (2010) **Geometric Morphometrics for Biologists: A Primer.** Elsevier Academic Press, London, UK.

Received: 05 January 2020

Accepted: 27 April 2020

Published: 13 May 2020

Appendix A – Shell geometric morphometrics

Collection locations of shell specimens

Three population samples of *Placostylus ambagiosus* were collected by Richard Prouse 1949-1950 (*P. a. watti*; *P. a. annectens*; *P. a. keenorum*). The location descriptions for these samples match Powell's type location descriptions for these subspecies (1947, 1951). Powell said *P. a. keenorum*'s distribution was limited to the east and west side of Maungapika Hill (pā) and might be the result of human movement (Powell 1947), but *P. a. watti*'s distribution at the base of a cliff was natural, and *P. a. annectens* had a much wider distribution than the Māori pā site at Unuwahao (Powell 1947). *Placostylus a. annectens* collected by Prouse was from "Spirits Bay, Unuwahao" where there was a pā, and is probably not exactly the same as either genetic sample locations called "Matirarau Bay" or "Te Huka Bay". However, we do have two tissue samples of *P. a. annectens* that have no location information, just subspecies name (samples GS17 =PS197; GS25 =PS205).

Four population samples of *Placostylus ambagiosus* were collected between 1990 and 1997 by Greg Sherley, a biologist who has written extensively on the shell size and shape variation of this snail species (Sherley 1996) and their genetics and biology (Triggs and Sherley 1993; Stringer *et al.* 2017). The *P. a. paraspiritus* sample he collected was from an old pā site south-east of Cape Maria van Diemen. *P. a. michiei* was collected from its type location at Surville cliffs (although on a headland, this is not a known pā site, and is not associated with other propagated species such as karaka). *P. a. wheneana* was collected from its type location in a steep valley (not pā site). The sample collected "SE of Cape Reinga, between Tirikawa and Pandora" is referred to as *P. a. "nouvelle"*, and came from a stretch of coast that has known pā sites on headlands. Details of this collection site are lacking but three populations referred to as *P. a. "nouvelle"* each consist of fewer than 50 individuals (Brooks, 2002).

Genetic samples came from type locations of ten subspecies of *Placostylus ambagiosus* (Triggs and Sherley 1993) plus four additional sites. These tissue samples included four individuals of *P. a. "nouvelle"* - two of which were in the National Arthropod frozen collection (and included in our SNP dataset).

Determining experimental error in geometric morphometrics

2D-geometric morphometric analysis using landmarks is used to compare shapes, and the inclusion of semi-landmarks permits analysis of components of shape, such as the curves of a snail shell profile, that lack explicitly defined landmarks. During processing, semi-landmarks are adjusted to 'slide' along the curve (anchored at each end by a biologically explicit landmark) in order to minimize variation in shape that is due simply to arbitrary placement of the points along the curve (Zelditch *et al.* 2004). Two virtual "combs" used to locate semi-landmarks were placed on images and aligned manually in Adobe Photoshop CS@ v9.0.2. Combs were aligned to a centre-line running from the apex of the shell to the intercept of the aperture and last whorl (Figure 1B). Digital images were organised into thin plate spine (TPS) files using tpsUtil (Rohlf, 2013) with the order of

specimens randomised to reduce potential experimental bias. Landmarks were then identified on each image using a Wacom Cintiq 22HD Pen Display tablet and the scale calibrated.

Sensitivity analysis was undertaken to quantify operator error associated with orientation of specimens for photography, placement of the comb, and digitisation of landmarks. A specimen of *Placostylus ambagiosus* from S. E. Cape Maria van Diemen 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 then followed the procedure for all other datasets and disparity analysis was performed to calculate the error associated with photography and digitisation. The disparity analysis calculated the proportion of variance attributed to photographic and digitisation error combined as 3.5% of the variance within the population sample. Overall variation from photographic and digitisation accounts for <2.3% of the variation within the dataset as a whole, as illustrated by the scatter of replicate points on a principal component plot (PCA, Fig. A1). 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%; Figure A.1).

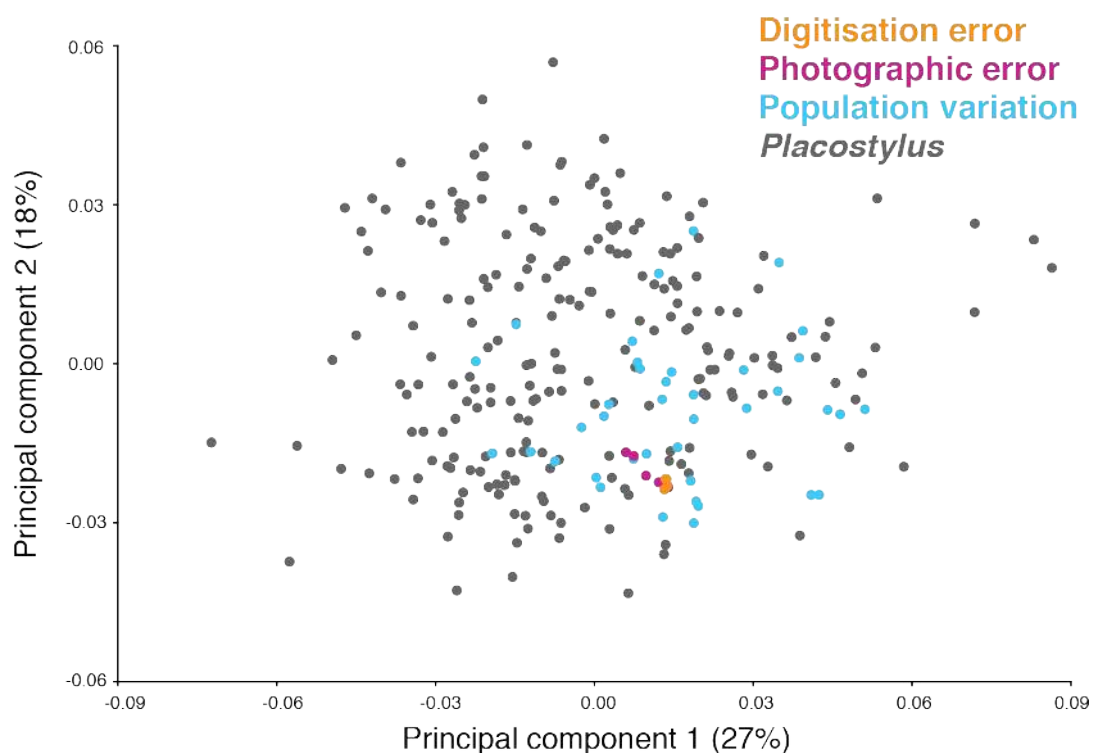


Figure A.1: Principal components 1 and 2 of shell shape variation within New Zealand **Placostylus** to illustrate measurement error. The population dataset is used as a background dataset (blue). Replicate (five times) digitisation of one image of **Placostylus**; (purple). Replicate (five times) photography and subsequent digitisation of one specimen of *Placostylus ambagiosus* from S. E. Cape Maria van Diemen; (pink). Other populations of *Placostylus*; (grey).

Landmark optimisation

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). If one is including a long, smooth, evenly curving profile (e.g., outer lip), then there is little point in using hundreds of landmarks (but this might only be determined after both digitizing and analysing multiple landmarks). In addition, the reduction of redundant landmarks improves statistical power because it raises the degrees of freedom relative to the dimensionality of the data. A problem then arises not only to balance complete coverage with the constraints of time, but also with the variability of specimens, and statistical power. Surprisingly few studies explicitly compare results from varying numbers of semi-landmarks to optimize dataset size (but see Vaux *et al.* 2017 supplementary information). Here we determine 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. We investigated how the choice of semi-landmarks affected discrimination of sample groups.

Nine combinations of landmarks and semi-landmarks were analysed 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 A1). The initial landmark configuration (A) with all data showed the first 9 PCs accounted for 84.7% of the variation, whereas landmark configuration G showed the first 7 PCs accounting for a total of 82.5% of the variation (Table A2). Landmark configurations J, M and N had the fewest significant principal components (5, 5, and 4 respectively; Table A2).

Canonical variate analysis (CVA) was used to test statistically the ability to differentiate *a priori* sample population groups, with discrimination success determined using cross-validation scores. The cross-validation used the number of individuals correctly assigned to each *a priori* group based on 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 and Sheets, 2010). The omitted specimen is then treated as an unknown and assigned to the population based on Mahalanobis distance. This was

Table A.1. 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 are therefore used in down-stream 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	-	-

Table A.2. Population samples of 275 *Placostylus* specimens showing the results of the cross-validation test (Jackknife) analysis by percentage correctly assigned to a priori groups; analysis performed using the “MASS” library in R. Labels A - O represent different configurations of landmarks. 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 further landmark analysis.

no. of semi landmarks	Population name	NMNZ	n	A	B	C	D	E	F	G	H	I	J	K	L	M	N
				63 9	63 9	49 8	49 8	47 9	47 8	33 7	33 6	30 7	19 5	19 6	19 6	14 5	5 4
	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
	Unuwahao	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
<i>P. ambagiosus</i>	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
	<i>P. hongii</i>	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
<i>P. bollonsi</i>	183761	5	100.00	100.00	80.00	100.00	100.00	80.00	100.00	100.00	100.00	100.00	80.00	100.00	100.00	100.00	0.00

conducted within both MorphoJ 1.06c (Klingenberg, 2011), analysing the original X-Y landmark coordinates, and using the R package MASS 7.3-26 (R Core Team, 2013; Venables & Ripley, 2002) and the PC scores generated from PCA in MorphoJ. *Placostylus bollonsi* (n=5) returned cross-validation scores of either 80% or 100% depending on the landmarks used in the analysis (Table A2). *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; Table A.2). 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 at Maungapika were correctly assigned to their *a priori* population more than 90% of the time in nine of the fourteen landmark combinations (Table A2). However, even with the full complement of landmarks and semi-landmarks several populations returned low assignment scores. This was most apparent in fossil population samples with the oldest population (5250 y bp) 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.

Overall, changing the number or configuration of semi-landmarks appeared to have little effect on the correct assignment to a *a priori* groups of individual *Placostylus* specimens. Removing entire curves from the landmark configuration (e.g. configurations I, J, K, N; Table A.2 and Fig. 1) did not improve the subsequent data, in fact, in most cases removal of an entire curve diminished the ability to determine correct assignment to a *a priori* groups.

Canonical variate analysis classification for extant populations did not vary greatly between the configurations A and G (Table A2). An optimal set of landmarks and semi-landmarks was selected for future work (Figure 1) which had 7 landmarks and 33 semi-landmarks and included all four aperture curves and the curve from the body whorl. This reduction from 63 to 33 semi-landmarks significantly decreased the time spent digitising

shells, but maintained biologically important information, allowing discrimination of extant populations.

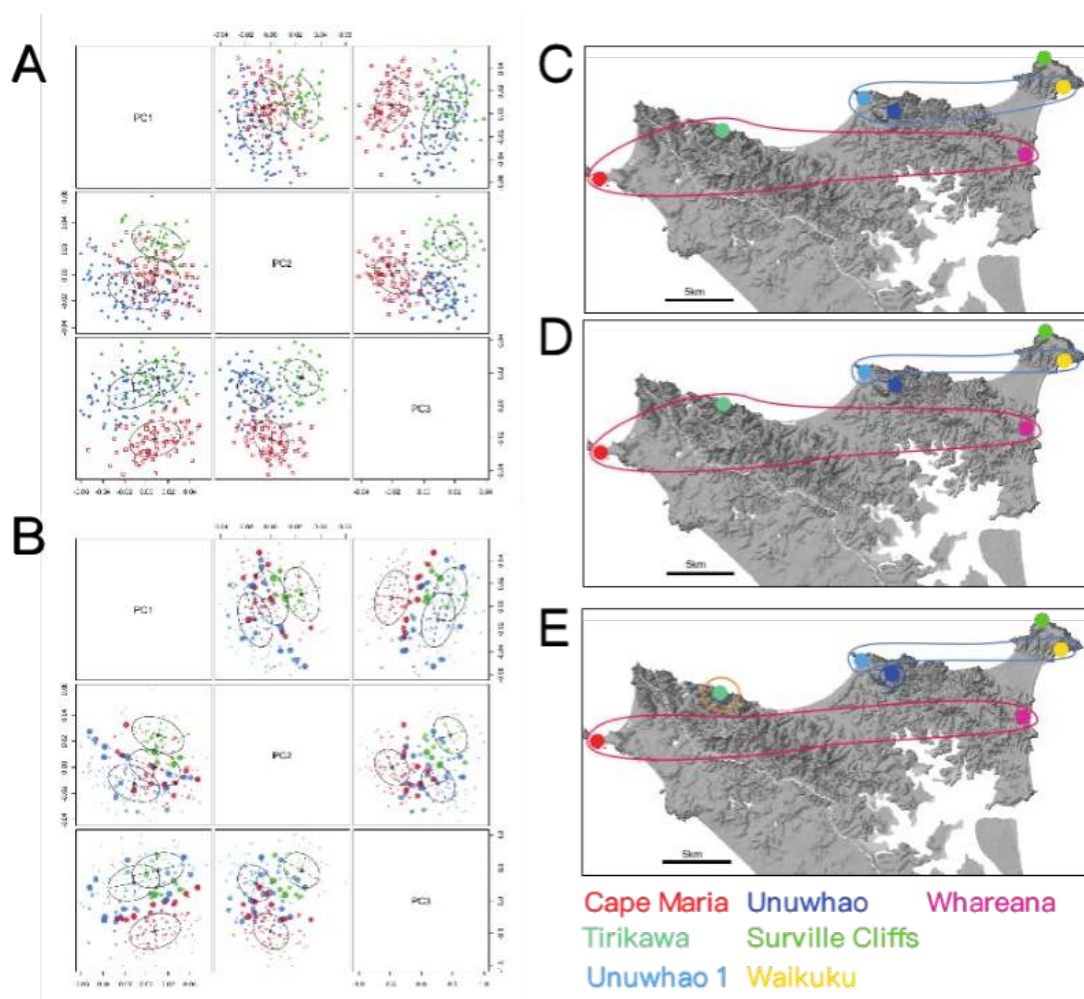


Figure A.2. Shell shape variation of *Placostylus ambagiosus* is not concordant with geographic proximity. (A) Using PC1, PC2, PC3, three clusters are resolved using Gaussian mixture models in MCLUST. Specimens marked as green triangles were mostly collected from Surville Cliffs (with some S. E. Cape Maria); red squares were collected from Tirikawa, S. E. Cape Maria and Whareana; blue circles were collected from Unuwahao and Waikuku. (B) The same Gaussian mixture model as illustrated in (A) with variation of assignment probabilities specimen shown, the uncertainty is illustrated by spot size, the larger the circle the more uncertain of cluster assignment. Maps show distribution of 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).

Appendix B. Single Nucleotide Polymorphism Genotyping

Removing samples with missing data

DNA sequences were obtained using a double digest restriction-site associated DNA (ddRADseq) protocol described by Dowle *et al.* 2015 (Table B1). Short-read DNA sequences were sorted into exactly matching stacks of sequences and stacks were then compared and parameters (set by the 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, and the output of site-level 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 a 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.

Table B.1. 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
	<i>Total</i>	<i>14077104</i>	<i>1114124</i>	<i>252291</i>	<i>48997</i>	<i>15492516</i>
R2	C	775	753	7738341	6389	7746258
	D	761	724	7738911	5862	7746258
	<i>Total</i>	<i>1536</i>	<i>1477</i>	<i>15477252</i>	<i>12251</i>	<i>15492516</i>

In total 45 *Placostylus ambagiosus* and 9 *Placostylus hongii* were sequenced. Average stack depth and number of stacks were compared to selected high-quality samples. 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. Individuals that did not have both of these for any given value of -m were not used in downstream analysis. Poor coverage resulted in 9 *P. hongii* and 26 *P. ambagiosus* being removed from the dataset, retaining 19 *P. ambagiosus* for downstream analysis.

Parameter Optimisation (-M; -n)

We 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 $\text{max_locus_stacks}=3$ (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 $M+2$. Varying the number of mismatches allowed between alleles when processing a single individual (M) has little effect on coverage depth (Table B2). 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 B.2. Variation in the parameter settings of the STACKS pipeline is shown compared to the number of loci common 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

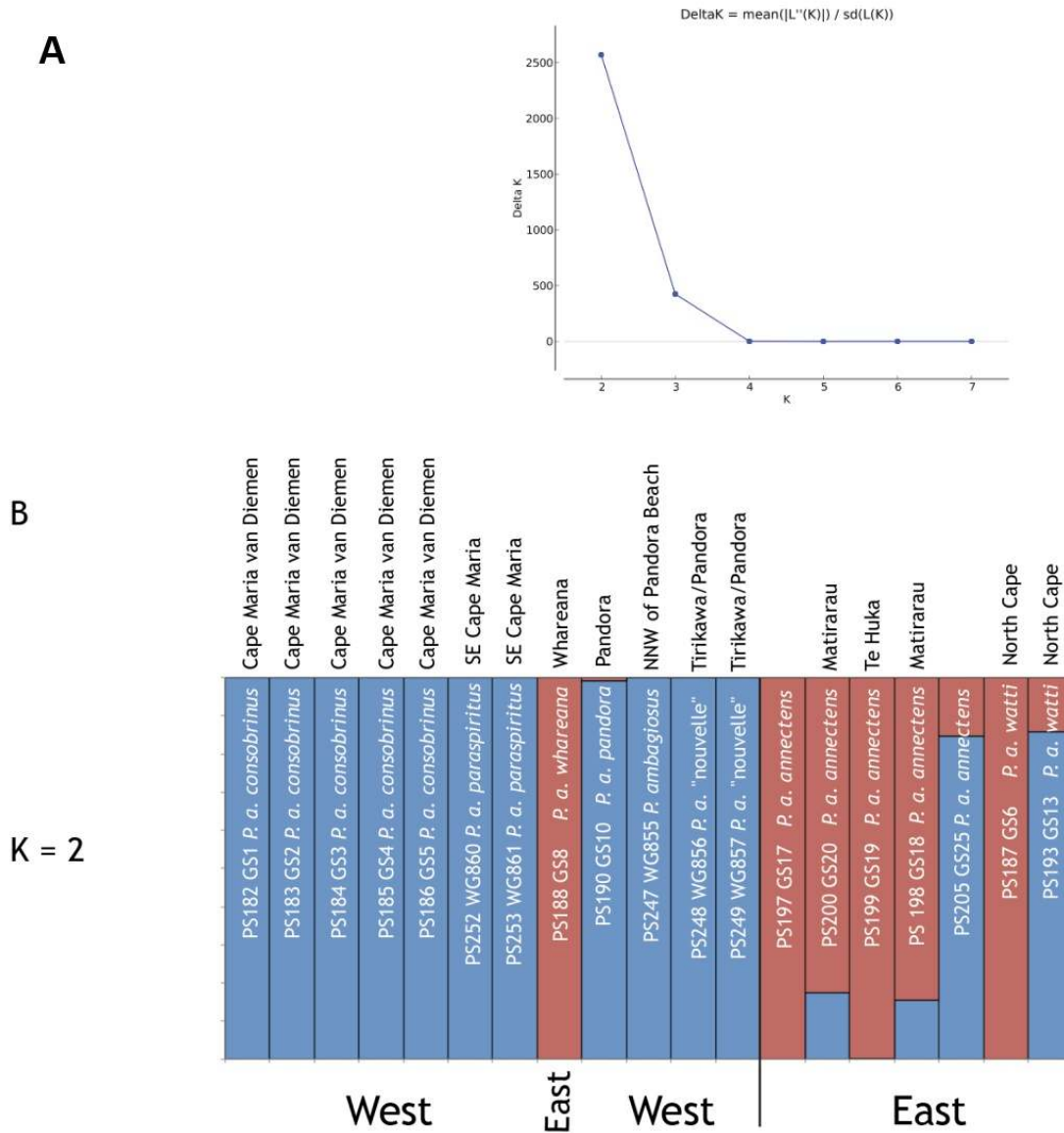


Figure A.3. Genetic STRUCTURE analysis of genotypes using 1738 nuclear SNP loci (RADseq) for the endangered land snail *Placostylus ambagiosus*. (A) To identify optimal K based on the posterior probability of the data for a given K, and the maximum value of Delta K was determined using the independent allele frequency model. (B) Genotype assignment probability (K=2) for 19 *Placostylus ambagiosus* individuals (1738 nuclear SNP loci).

REFERENCES

- Brook FJ (2002) **Uncommon and rare landsnails in the Northland region of New Zealand, and an assessment of conservation management priorities.** New Zealand Department of Conservation. ISBN 0-478-22343-9
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) **Stacks: building and genotyping loci de novo from short-read sequences.** *G3: Genes, Genomes, Genetics*, 1(3), 171–182. <https://doi.org/10.1534/g3.111.000240>
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) **Stacks: an analysis tool set for population genomics.** *Molecular Ecology*, 22(11), 3124–3140. <https://doi.org/10.1111/mec.12354>
- Dowle EJ, Morgan-Richards M, Brescia F, Trewick SA (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. <https://doi.org/10.1111/mec.13302>
- Klingenberg CP (2011) **MorphoJ: an integrated software package for geometric morphometrics.** *Molecular Ecology Resources* 11, 353–357. doi:10.1111/j.1755-0998.2010.02924.x
- Powell AWB (1947) **Distribution of *Placostylus* Land Snails in Northern most New Zealand on JSTOR.** *Records of the Auckland Institute and Museum*, 3(3), 173–188.
- 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. <https://doi.org/10.1080/03014223.1996.9518067>
- Stringer IAN, Parrish GR, Sherley GH (2017) Homing, dispersal and mortality after translocation of long-lived land snails *Placostylus ambagiosus* and *P. hongii* (Gastropoda: Bothriembryontidae) in New Zealand. *Molluscan Research* 38. <https://doi.org/10.1080/13235818.2017.1323368>
- Triggs, SJ, Sherley, GH (1993) **Allozyme genetic diversity in *Placostylus* land snails and implications for conservation.** *New Zealand Journal of Zoology*, 20(1), 19–33. <https://doi.org/10.1080/03014223.1993.10423239>
- Vaux F, Crampton JS, Marshall BA, Trewick SA, Morgan-Richards M (2017) **Geometric morphometric analysis reveals that the shells of male and female siphon whelks *Penion chathamensis* are the same size and shape.** *Molluscan Research* 37, 194–201.
- Venables WN, Ripley BD (2002) **Modern applied statistics with S.** Fourth Edition. Springer, New York, USA.
- Webster M, Sheets HD (2010) **A practical introduction to landmark-based geometric morphometrics.** *Quantitative Methods in Paleobiology* 16, 163–188.
- Zelditch ML, Swiderski DL, Sheets HD, Fink WL (2004) **Geometric morphometrics for biologists: a primer.** Elsevier Academic Press, London, UK.