Biological Journal of the Linnean Society, 2018, , 1–17. With 4 figures.

MICHAEL R. GEMMELL[1], STEVEN A. TREWICK[1], JAMES S. CRAMPTON[2,3], FELIX VAUX[1], SIMON F. K. HILLS[1], ELIZABETH E. DALY[1], BRUCE A. MARSHALL[4], ALAN G. BEU[2] and MARY MORGAN-RICHARDS[1*]

[1]*Ecology Group, College of Sciences, Massey University, Palmerston North, New Zealand*
[2]*GNS Science, PO Box 30-368, Lower Hutt, New Zealand*
[3]*School of Geography, Environment & Earth Sciences, Victoria University, Victoria University of Wellington, PO Box 600, Wellington, New Zealand*
[4]*Museum of New Zealand Te Papa Tongarewa, Wellington, New Zealand*

*Corresponding author. E-mail: m.morgan-richards@massey.ac.nz

a wide variety of genetic and environmental influences. In marine environments, factors such as water depth, temperature, currents, wave energy and predation have been shown to affect plastic shell traits (Palmer, 1990; Baker *et al.*, 2004; Olabarria & Thurston, 2004; Shelmerdine *et al.*, 2007; Hollander & Butlin, 2010; Hills *et al.*, 2012; Avaca *et al.*, 2013; Butlin *et al.*, 2014; Haig *et al.*, 2015; Irie & Morimoto, 2016; reviewed by Bourdeau *et al.*, 2015). Different ecological conditions across the geographical range of a species can generate polymorphisms that can lead to the erroneous taxonomic

To determine whether the local environment can explain the particular shell phenotype recognized as *B. colensoi*, we also examined an independent lineage from the same environment. The rocky shore snail *Cominella maculosa* (Martyn, 1784), although superficially similar in appearance, is not closely related to *Buccinulum* (Donald *et al.*, 2015; Vaux *et al.*, 2017). Both species lay eggs that hatch as small snails (Ponder, 1971; Morley, 2013; Donald *et al.*, 2015), with no pelagic larval stage, and so are expected to have similar dispersal rates. The direct larval development of a marine species can facilitate the maintenance of biogeographical patterns (Lee & Boulding, 2009) when water currents are considered (Schiel, 2004; Sponaugle *et al.*, 2005). On the eastern coast of North Island New Zealand north of East Cape, the coastal ocean current flows south until reaching

was estimated using cross-validation scores (number of individuals correctly assigned to each group with 1000 permutations).

For each genus, morphometric data were examined using a model-based clustering approach to explore the distribution of variation and to test for natural clusters without a priori classification. We used the MCLUST v.5.0.2 package (Fraley *et al.*, 2012) in R that implements Gaussian modelling where the total dataset is considered as a mixture of multivariate normal datasets, with a selection of covariance structures and vectors of expectation (Nanova, 2014; Scrucca *et al.*, 2017). Unlike discriminant analysis (CVA), MCLUST analysis does not require prior information about specimen identity to classify sample data (Fraley & Raftery, 1999, 2002, 2003). The optimal model and number of clusters in the

T4 ligase to enable identification of individuals after pooling of samples. Fragments between 250 and 350 bp long were selected for sequencing. Single-end high-throughput sequencing was performed using an Illumina Hi-seq (NZGL), and resulting data were processed using the STACKS 1.0.1 pipeline (Catchen *et al.*, 2013).

Selection of nuclear markers was undertaken so that 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 were implemented relating to read coverage, individual number and population coverage. Initial exploration of the data suggested that depth coverage varied for the 73 individuals. Recommended read depth coverage settings vary in the literature (Peterson *et al.*, 2012; Buerkle & Gompert, 2013); therefore, we experimented with parameter optimization. After initial tests using the STACKS pipeline to assess information content, samples with file sizes of < 2 megabytes were excluded from further analysis. These samples had low DNA read numbers, and their inclusion reduced analytical power downstream. Low coverage combined with high error rate has the potential to reduce the number of true loci by a substantial amount, up to 51% (Catchen *et al.*, 2011).

The number of mismatches allowed between alleles when processing a single individual ($-M$) was explored. If $-M$ is too low, some real loci are not formed, and subsequent alleles will be treated as different loci (undermerging). If $-M$ is too high, repetitive sequences and paralogues will form large nonsensical loci (overmerging). We tested a range of values ($-M$ = 5–25) and found that catalogue construction failed at high values (e.g. $-M$ = 25). When alleles were allowed to differ by a maximum of 15 substitutions ($-M$ = 5–15), some variation in *omorph*

with a burn-in of 200 000 and 1 000 000 generations, were used. Potential populations (K) was set from one to six (the number of population samples). To determine the optimal number of clusters, we examined estimates of the posterior probability of the data for a given K (Pritchard *et al.*, 2000) and ΔK, the rate of change in logarithmic probability of the data (Evanno *et al.*, 2005) implemented in STRUCTURE

visualize shape variation, there was some clustering of specimens, but overlapping shell shape distributions (Fig. 2). *Buccinulum* specimens from the eastern region (the range of *B. colensoi*; Fig. 1) were the most distinct for PC1. Clusters of shell shape for other regions were also apparent, with non-overlapping confidence regions on group means, but without clear separation of the four a priori groups. Not all individuals were allocated to their original sample groupings in cross-validation tests (*P*-values for difference between means ranging between 0.37 and 0.97; Table 2). A single morphological cluster of specimens was inferred as optimal for the *Buccinulum* data using the model-based clustering approach with principal components 1–5 (diagonal multivariate normal model).

Given that the majority of shell shape variation was explained by PC1 (51.7%; Fig. 2), we explored the data further using a model-based clustering of specimens performed with only PC1. The Bayesian information criterion selected a model with two morphological clusters as being the optimal clustering strategy for PC1 (Fig. 3). One cluster contained 15 specimens collected from the eastern region (geographical range of *B. colensoi*) and two specimens from the northern region (Oakura). These shells were characterized by relatively short spires and wide apertures (Fig. 2). The other cluster comprised seven specimens from the eastern region (Mahia, Akitio and Castlepoint) with samples from the northern, southern and Chatham samples (assignment probabilities for each specimen are shown in Fig. 3).

Shell shape variation within the sympatric whelk *C. maculosa* was examined using 46 shells from six sampling locations. The first three principal components provided statistically significant information (broken-stick test) and explained 73.74% of shell shape variation (see Supporting Information, Table S2). When the first two principal components (65.26% of variation) were used to visualize shape variation, little clustering of regional samples was evident. There was overlap of regional samples in morphospace, although eastern region shells had a 90% mean confidence ellipse that did not overlap the ellipses for the other two regions (Fig. 2). Snails collected from the eastern region had shells that were, on average, tall and narrow (Fig. 2). Nevertheless, a single morphological cluster was identified by model-based clustering of PC1 and from the first three principal components using MCLUST.

D A D AD A B

A fragment of *cox1* sequence (443 bp) was used for phylogenetic inference of *B. v. vittatum*, *B. v. littorinoides*, *B. v. bicinctum* and *B. colensoi*, with the outgroup *B. powelli*. Within the *B. vittatum* complex, three mtDNA clades were resolved (see Supporting information, Fig. S2), one representing the northern subspecies *B. v. vittatum*, which was sister to a clade of Chatham Island specimens (*B. v. bicinctum*), and a third clade with the southern subspecies *B. v. littorinoides*

Supporting informatTJ/T1 1 TS (n).6 (d).T1 1 Sp sus

A set of 849 anonymous nuclear loci (SNPs) were produced for 28 individuals from the *B. vittatum* complex, from six locations (Table 1). The genetic clustering model with the best fit to the data (optimal *K* value = 2; see Supporting Information, Table S3 and Fig. S5) divided the *Buccinulum* specimens into two clusters (STRUCTURE HARVESTER; Fig. 3). Ten specimens (from Oakura, Waiheke Island and Hicks Bay) grouped with assignment probabilities between

B. v. littorinoides than between either of these two groups and *B. v. vittatum* (Table 4).

have slightly different shell shapes compared with conspecifics, they did not show the relatively short spire and wide aperture that characterizes *B. colensoi* (Fig. 2). In addition, *C. maculosa* from this region do not exhibit the differing shell surface sculpturing observed in *Buccinulum*.

In contrast to the single morphometric cluster identified with five principal shape components, the mitochondrial DNA sequence (*cox1* gene) resolved distinct partitioning within the *B. vittatum* complex. Individuals from the Chatham Islands (identified as *B. v. bicinctum* from their location) formed a distinct mitochondrial clade. Individuals from within the range of *B. v. vittatum* (north) formed a distinct genetic cluster, and individuals from within the range of *B. v. littorinoides* (south) and *B. colensoi* (east) together formed a cluster of haplotypes. Analysis of 849 anonymous nuclear loci (without the Chatham Island population sample) showed this same pattern, resolving two clusters [northern *B. v. vittatum* distinct from samples from eastern and southern locations (*B. colensoi* and *B. v. littorinoides*)]. Thus, the genetic partitioning of both mitochondrial and nuclear markers was in conflict with the recognized taxonomy. In contrast to the mtDNA data, the nuclear markers suggest that gene flow between populations from within the range of *B. v. vittatum* (north) and *B. colensoi* (east) is ongoing. This finding is in agreement with observations by Ponder (1971) of rare 'intermediate specimens' on the north-east side of East Cape, where the two taxa meet. The three snails in our study from Hicks Bay, East Cape, which grouped with *B. v. vittatum* in mtDNA analysis, shared nuclear (SNP) alleles with both *B. vittatum* complex genetic clusters.

There is a morphotype of *Buccinulum* currently identified as *B. colensoi* present on the east coast of the North Island, but the shell traits that are distinctive to this taxon were not found to be diagnostic of a separate lineage. Variation within population samples suggest that the *B. colensoi* form might be better considered an ecotype. There is no clear indication of environmental conditions that might be influencing phenotype. In contrast to the limited shell shape variation, we found strong genetic clustering in *Buccinulum*, revealing the western and northern samples of *B. vittatum* as

convergent evolution and cannot determine whether there is an environmental factor that selects for phenotypic differences between *B. colensoi* and *B. vittatum*. One might infer from the signature of population differentiation of neutral nuclear genetic markers that shell shape and texture variation could simply be the result of genetic drift. However, the ecotype *B. colensoi* shares mtDNA haplotypes with *B. v. littorinoides*, suggesting that selection is preventing the homogenizing effect of gene flow.

Local adaptation among animal populations has been inferred from phenotypic traits, but determining the role of natural selection in shaping geographically partitioned variation is not simple (Merilä & Hendry, 2014). Shell shape variation of snails can result from a combination of genetic drift in isolation or selection on functional difference (Conde-Padín *et al.*, 2009) and/or represent ecophenotypic plasticity (Iguchi *et al.*, 2005) and historical phylogeographical effects. If variation results from divergence owing to drift, then phenotype and neutral genetic markers are expected to show similar patterns. This is not observed in *Buccinulum*. Phenotypic variation within a single species maintained by selection or ecophenotypic plasticity would explain the pattern we observed. The factors that distinguish the niche of these two similar whelks might be key to understanding why *B. littorinoides* has a distinct East Coast ecotype, but *C. maculosa* does not.

ACKNOWLEDGEMENTS

We thank members of the Phoenix laboratory (evolves.massey.ac.nz) for discussion over earlier versions of the manuscript. We thank Rose Gregersen and Logan Penniket for DNA extractions of specimens. Two anonymous reviewers provided constructive suggestions that improved this work. Financial support for this project was provided by the New Zealand Royal Society Marsden fund (12-MAU-008) and a Te Tipu Pūtaiao Postdoctoral Fellowship (CONT-22922-TTP-MAU).

REFERENCES

. 2013. Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution* **4** 393–399.

. 2011. What, if anything, can we learn from the fossil record about speciation in marine gastropods? Biological and geological considerations. *American Malacological Bulletin* **2** 247–276.

. 2014. The geography and ecology of plant speciation: range overlap and niche divergence in sister species. *Proceedings of the Royal Society B: Biological Sciences* **2 1** 20132980.

2013. Shell shape variation in the Nassariid *Buccinanops globulosus*. *Helgoland Marine Research* **6** 567–577.

. 2004. Mitochondrial DNA phylogenetic structuring suggests similarity between two morphologically plastic genera of Australian freshwater mussels (Unionoida: Hyriidae). *Molecular Phylogenetics and Evolution* **32** 902–912.

. 1 . Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16** 37–48.

. 1 1. *Morphometric tools for landmark data: geometry and biology*. Cambridge: Cambridge University Press.

. 2005. Classification and nomenclator of gastropod families. *Malacologia* **4** 1–368.

. 201 . Revised classification, nomenclator and typification of gastropod and monoplacophoran famalies. *Malacologia* **61** 1–526.

. 2015. What c .

gastropods: the effect of mode of larval development on variation at one mitochondrial and two nuclear DNA markers. *Molecular Ecology* **1** 2165–2184.

. 2012. Geneious Basic: an integrated and extendable desktop

and Stewart Island. The Chatham clade consists of samples solely from the Chatham Islands. The outgroup is *Buccinulum pallidum powelli*. Bayesian posterior probabilities are indicated at nodes.

3. A positive relationship between geographical and genetic distance between populations the New Zealand rocky shore whelk *Buccinulum vittatum* complex suggests isolation by distance. Pairwise Φ_{ST} estimates from mitochondiral DNA sequence (*cox1*; orange spots) and pairwise F_{ST} estimates (blue) inferred from 849