ISSN 0962-1083

VOLUME 22
NUMBER 10
MAY
2013

# MOLECULAR ECOLOGY

Published by
Wiley Blackwell

MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 2726–2741 doi: 10.1111/mec.12287

# Ecological speciation in an island snail: evidence for the parallel evolution of a novel ecotype and maintenance by ecologically dependent postzygotic isolation

SEAN STANKOWSKI

*School of Animal Biology (M092), University of Western Australia, Crawley, WA 6009, Australia*

## Abstract

**Speciation is the process by which reproductive isolation evolves between populations. Two general models of speciation have been proposed: ecological speciation, where reproductive barriers evolve due to ecologically based divergent selection, and mutation-order speciation, where populations fix different mutations as they adapt to similar selection pressures. I evaluate these alternative models and determine the progress of speciation in a diverse group of land snails, genus *Rhagada*, inhabiting Rosemary Island. A recently derived keeled-flat morphotype occupies two isolated rocky hills, while globose-shelled snails inhabit the surrounding plains. The study of one hill reveals that they are separated by a narrow hybrid zone. As predicted by ecological speciation theory, there are local and landscape level associations between shell shape and habitat, and the morphological transition coincides with a narrow ecotone between the two distinct environments. Microsatellite DNA revealed a cline of hybrid index scores much wider than the morphological cline, further supporting the ecological maintenance of the morphotypes. The hybrid zone does not run through an area of low population density, as is expected for mutation-order hybrid zones, and there is a unimodal distribution of phenotypes at the centre, suggesting that there is little or no prezygotic isolation. Instead, these data suggest that the ecotypes are maintained by ecologically dependent postzygotic isolation (i.e. ecological selection against hybrids). Mitochondrial and Microsatellite DNA indicate that the keeled-flat form evolved recently, and without major historical disruptions to gene flow. The data also suggest that the two keeled-flat populations, inhabiting similar rocky hills, have evolved in parallel. These snails provide a complex example of ecological speciation in its early stages.**

*Keywords*: hybrid zone patchwork, land snail, mosaic landscape, parapatric divergence, shell shape, unimodal hybrid zone

*Received 4 November 2012; revision received 11 January 2013; accepted 16 January 2013*

## Introduction

Speciation is the process by which reproductive isolation evolves between populations (Coyne & Orr 2004). Two models of speciation have been proposed (Schluter 2009). The first, ecological speciation, describes the evolution of reproductive barriers as a result of divergent ecological selection. Selection may arise from any ecological source, including environmental differences, sexual selection or ecological interactions, and may cause reproductive isolation in a variety of ways, both pre- and postzygotic (Rundle & Nosil 2005; Schluter & Conte 2009). While recognized as a legitimate isolating mechanism in its own right (Nosil *et al.* 2005; Rundle & Nosil 2005), ecological selection can drive the evolution of other intrinsic isolating mechanisms and, ultimately, complete reproductive isolation (Barton 2001; Schluter 2009). The alternative model is mutation-order speciation, where reproductive isolation evolves as a by-product of the fixation of

Correspondence: Sean Stankowski, Fax: 6488 1029;
E-mail: stanks01@student.uwa.edu.au

different advantageous mutations in separate populations experiencing similar selection pressures (Schluter 2009). Although selection does not favour divergence *per se*, populations diverge because they acquire different mutations, or because the same mutations are fixed in a different order (Schluter 2009). Among population, genomic incompatibilities reduce the fitness of recombinant genotypes, generating a partial or complete intrinsic reproductive barrier between the divergent forms. Reinforcement can strengthen a weak barrier, resulting in the evolution of complete reproductive isolation (Coyne & Orr 2004).

While both models are theoretically plausible, and are supported by clear empirical examples, there is considerable debate surrounding their relative importance in nature (Coyne & Orr 2004; Rundle & Nosil 2005; Schluter 2009; Schluter & Conte 2009). Some authors have argued that stochastic divergence and genomic incompatibility are the primary drivers of speciation (Barton & Hewitt 1985, 1989; Hewitt 1988, 2011); others place primary importance on ecological differences (Jiggins & Mallet 2000; Mallet 2008). Part of the problem is that initial barriers, evolving by ecological selection or mutation-order divergence, favour the evolution of additional reproductive barriers that are common to both modes of speciation; so, for most species pairs, it is difficult to determine which reproductive barrier evolved first (Schluter 2009).

Closely related geographic races and their hybrid zones provide excellent material to study this and other related problems (Mallet 2008). They are predicted to evolve under both models of speciation (Kruuk *et al.* 1999). In the ecological model, geographic races form when habitat-specific selection drives populations in contrasting environments towards different adaptive peaks (Kruuk *et al.* 1999). Where the transition between habitats is an abrupt step or ecotone, ecological speciation theory predicts the formation of narrow hybrid zones coincident with the environmental interface. In the early stages of their evolution, ecological hybrid zones are expected to be maintained by ecologically dependent postzygotic isolation (Kruuk *et al.* 1999; Rundle & Nosil 2005). Specifically, hybrid offspring, which are genetically intermediate of the alternative ecotypes, are poorly adapted to both parental environments. The resulting selection against hybrid immigrants acts to maintain the genetic integrity of the parental ecomorphs. In the mutation-order model, populations diverge stochastically in physical isolation (Hewitt 1988, 2011; Coyne & Orr 2004). Upon recontact, hybrid unfitness restricts the area of recombination to a narrow hybrid zone, so that an intrinsic postzygotic barrier to gene flow separates the parental races (Barton & Hewitt 1985; Rundle & Nosil 2005).

This study focuses on the mechanisms driving the formation of geographic races on Rosemary Island, a small (11 $km^2$) continental island in the Dampier Archipelago off the coast of northwestern Australia. *Rhagada*, an arid-zone genus of land snail, can be found over most of the island, but vary dramatically in their form among locations (Fig. 1). The most striking variation is in the shape of their shells, which range from globose to keeled-flat (Stankowski 2011). Not only does this variation span the entire range of variation in the genus *Rhagada* (Johnson *et al.* 2012), but it exceeds levels of variation observed in most genera of land snails. There is also evidence that the recently derived keeled-flat form has evolved on Rosemary Island, as detailed surveys of the Dampier Archipelago and mainland have failed to recover the same phenotype (Stankowski in prep.) and the *Rhagada* of Rosemary Island form a well-supported monophyletic group (Stankowski 2011).

The distributions of the extreme shell forms appear to be related to the local habitat (Stankowski 2011). Keeled-flat populations are found on two hilltops which have a very rocky substrate, while globose-shelled populations inhabit the surrounding spinifex plains, which are low-lying and have a sandy or stony substrate. Narrow clines of shell shape, almost certainly hybrid zones, approximately coincide with the transition between habitats (Stankowski 2011). A preliminary analysis of mtDNA revealed high haplotypic diversity, but low levels of sequence divergence and no association between shell shape and the pattern of mitochondrial evolution (Stankowski 2011). Based on the patterns of phenotypic continuity, the association between shell shape and habitat and the results of the phylogenetic analysis, it was proposed that the keeled-flat phenotype form has evolved due to divergent ecological selection and in the presence of gene flow (Stankowski 2011).

This study integrates morphological, molecular and ecological data in an attempt to understand the mode and progress of speciation in this diverse group of land snails. The investigation was conducted at two geographic scales: (i) at the landscape scale, including both of the rocky hills inhabited by flat-shelled *Rhagada*; and (ii) across transects, which capture the transition between the globose- and flat-shelled phenotypes. At the broad scale, I provide a detailed picture of the distributions of the different shell forms and the pattern of population subdivision obtained from microsatellite genotypes and measurements of the local habitat. A population level analysis of mtDNA sequences is conducted to examine the role that history has played in the morphological divergence. At the fine scale, I estimate parameters that describe the clines, and with a time-scaled estimate of dispersal obtained from microsatellite DNA under an isolation-by-distance model, estimate the strength of selection required to maintain them. Patterns of environmental variation, population

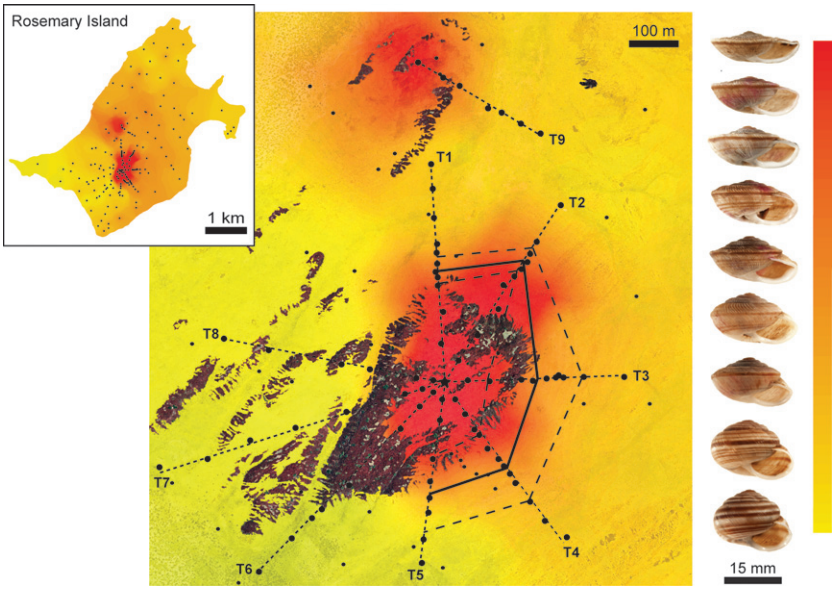


**Fig. 1** Geographic variation of shell shape on Rosemary Island. The prediction surface is based on mean estimates of shape (si score) from snails collected at 183 locations (black circles). The aerial photograph shows the main study area, including both rock outcrops inhabited by flat-shelled Rhagada. Areas not included in the prediction surface lacked adequate vegetation, and were not occupied by snails. The fine dashed lines connect samples collected along nine transects, eight of which radiate from a single sample site at the centre of the southern rock outcrop. The heavy solid and dashed lines connect neighbouring cline centres and widths, respectively. The area between the upper and lower estimates of width is the '2D' cline.

density and microsatellite differentiation are also examined across the transition zone to test the expectations of ecological and mutation-order models of speciation. Together, the results reveal a striking hybrid zone patchwork, the evolution of which is explained most parsimoniously by the current ecological setting.

## Materials and methods

### *Samples*

Snails were collected from 183 locations in two consecutive years. In 2008, 103 sites were sampled across the island to establish the distributions of the different forms (Stankowski 2011). The keeled-flat form has been described as *Rhagada dampierana*, and other variants on the island have been described as four separate species, although all have globose shells (Solem 1997). As a detailed study of these samples provided no support for the taxonomic divisions (Stankowski 2011), I refer to populations based only on shell shape. In 2009, additional samples were collected within a 1.5 × 1.7 km area. This landscape, which is the focal area for the current study, includes both rock outcrops inhabited by keeled-flat *Rhagada* (Fig. 1). Seventy-nine sample sites were arranged among nine transects. Eight transects radiated from a single sample site at the centre of the southern rock outcrop, each capturing a different aspect of the transition between the flat and globose forms. The ninth transect was positioned onto the smaller northern rock outcrop. At each site, a 14 $m^2$ (2 × 7 m) quadrant was thoroughly searched to ensure that all live snails and empty adult shells were collected. If a reasonable sample could not be obtained, the search area was widened in 7 $m^2$ increments. The final size of the each site was recorded, so that population density could be estimated. The maximum search area (28 $m^2$) was smaller than the neighbourhood estimated for a congeneric species that lives in a similar environment and has a similar life history (Johnson & Black 1991).

### *Measurement of shell shape*

In a previous study of these snails, Relative Warp Analysis was used to quantify variation in shell shape (Stankowski 2011). The first relative warp, which explained over 67% of the total variation in shape, was almost perfectly correlated with the spire index, which is the ratio of the shell height to shell width ($r^2$ = 0.98). The spire index is a well studied aspect of land snail shell morphology and is highly heritable in species that have been characterized genetically (Cain 1977; Welter-Schultes 2000; Okajima & Chiba 2009). Here, I use the spire index (*si*), which was estimated for up to 10 adult shells collected at sites visited in 2008 and 20 adult shells from sites visited in 2009. In some locations, a full sample of live adults could not be obtained, so I

included empty shells with 'live' colour and texture. Height and width of 2269 shells were measured according to Cain (1977), using digital calipers accurate to 0.01 mm.

### *DNA extraction and molecular markers*

Genomic DNA was extracted from 675 snails from 82 sites (mean 8.2 ± SD 6.8 snails per site; range 1–20), according to Ivanova *et al.* (2006). Each sample was genotyped at 8 polymorphic, unlinked microsatellite loci designed for the congener *Rhagada convicta*. Marker development methods and tests of Hardy–Weinberg equilibrium and linkage disequilibrium are provided in the Appendix S1, Tables S1 and S2 (Supporting information). Each locus was amplified in 10 μl reaction volumes containing 50 mM KCl, 20 mM Tris HCl (pH 8.4), 0.2 mM of each DNTP, 0.3 mM of each primer, 0.6 μl of *Taq* DNA polymerase, between 0.75 and 1.25 mM $MgCl_2$ (Table S1, Supporting information) and 20 ng of DNA. Cycling conditions were as follows: 94 °C for 3 m, followed by 35 cycles of 94 °C for 40 s, 40 s at 52, 54 or 56 °C (Table S1, Supporting information), 72 °C for 30 s and a final extension period of 30 s at 72 °C. Labelled products were separated on an ABI 377 96-well capillary sequencer and scored using GENEMARKER version 1.7.

A 666-bp fragment of the mitochondrial gene cytochrome oxidase I (COI) was obtained for 267 individuals from 78 sites (mean 3.37 ± SD 1.68, range 1–5) using the forward (L1490-Alb) and reverse (H2198-Alb) primers of Gittenberger *et al.* (2004), according to Stankowski (2011). Cleanup and sequencing reactions were conducted by the Australian Genome Research Facility using Big Dye chemistry (v.3.1.). Sequences were edited and aligned with SEQUENCHER v.4.6.

### *Estimation of morphological cline centre and width*

For each transect, patterns of clinal variation in *si* scores were described using a four-parameter sigmoid function, $y = y_0 + a/(1 + \exp(-(x-x_0)/b))$, where $x_0$ is the value of $x$ at the inflection point, which provides an estimate of the geographic position of the cline centre ($c$) (Barton & Gale 1993). The sigmoid fits were conducted using the program SIGMAPLOT V. 11 (Systat Inc.), which uses the Marquardt–Levenberg algorithm (Marquardt 1963) to find the parameter values that give the best fit between the equation and the data. The process is iterative (each analysis consisted of 10 000 fits, each with a maximum of 1000 iterations) and seeks the values of the parameters that minimize the sum of the squared differences between the observed and predicted values of the dependent variable. Cline width ($w$) was obtained according to Barton & Gale (1993) as the ratio between $\Delta z$, the difference in the mean trait value on either end of the cline and the gradient at the cline centre ($\Delta z/(\partial z/\partial x)$).

With the cline centre for each transect established, sites were scaled to a common centre and collapsed into a one-dimensional transect. Some additional sample sites were included by measuring their minimum distance to the fitted two-dimensional cline centre, which was estimated by connecting neighbouring cline centres with linear segments (Fig. 1). This increased the total number of sample localities from 50 to 67. Cline parameters were estimated as described above.

### *Analysis of Microsatellite data*

Two methods were used to examine population subdivision. First STRUCTURE v. 2.3.2 (Pritchard *et al.* 2000) was used to infer population genetic structure based on the 675 microsatellite genotypes. The STRUCTURE algorithm probabilistically assigns individuals to $K$ clusters, so that departures from Hardy–Weinberg equilibrium and linkage disequilibrium are minimized within each cluster. Ten replicate runs were conducted for each $K$ between 1 and 12, assuming admixture and uncorrelated allele frequencies. Each run consisted of 50 000 iterations of burn-in, followed by 150 000 iterations. The optimal $K$ was determined using the $\Delta K$ method of Evanno *et al.* (2005), as implemented in STRUCTURE HARVESTER (Earl & van Holdt 2011). $Q$ scores (probability of membership of each individual to each $K$) were obtained by averaging values from 40 independent runs using the Greedy algorithm in CLUMPP (Jackobsson & Rosenberg 2007).

Second, estimates of differentiation were calculated between the divergent shell forms. Individuals were first allocated to one of four groups. The 'Intermediate' group contained all individuals from sites located within the bounds of the 2D cline (Fig. 1). The second and third groups contained 'Globose' and 'South Keeled-Flat' samples from outside the cline. The fourth group, 'North Keeled-Flat' contained flat-spired samples from the northern rock outcrop. Pairwise estimates of $F_{ST}/(1-F_{ST})$ were calculated between each pair of sample sites, excluding those with fewer than 10 individuals, according to Weir & Cockerham (1984) in Genepop v. 4.0.10 (Raymond & Rousette 2009). These estimates were plotted against pairwise estimates geographic distance, made taking the undulations of the landscape into account, so the level of divergence between shell groups could be examined relative to the pattern of spatial divergence within groups. Mantel tests were used to determine the significance of the relationship for selected sets of comparisons.

If a cline in a quantitative trait is maintained by hybrid unfitness rather than divergent ecological selection, then its distribution should mirror the spatial pattern of recombination between two alternative stable equilibria. To test this, STRUCTURE $Q$ scores were also examined along the collapsed transect. Mean $Q$ scores were obtained for each site. Sites with fewer than 5 genotyped individuals were not included in the analysis, reducing the number of localities from 67 to 49. The mean $Q$ scores were plotted against their distance from the morphological cline centre ($x_0$), and estimates of the centre and width were obtained as described for the morphological data.

### *Geostatistical analysis*

Patterns of geographic variation in mean *si* and mean STRUCTURE $Q$ scores were illustrated by the construction of prediction surfaces in the geostatistical analyst extension of ARCGIS 9 (ESRI). Ordinary Kriging, an interpolation technique, was used to produce both surfaces. For both data sets, the J-Bessel model best reflected the semivariogram cloud. Anisotropy was incorporated into the model, because there was evidence for directional autocorrelation. A suitable lag size was determined as recommended in the manual, and the optimal values of the partial sill and nugget were calculated by the software. The accuracy of the final surfaces was confirmed by comparing the empirical values at each sample site with the values predicted from the semivariogram model.

### *Patterns of variation at the cline centre*

Stable hybrid zones can be identified by the presence of increased variation at the cline centre, which is generated when alternative stable equilibria recombine (Endler 1977). To test for elevated quantitative and molecular variation at the cline centre, mean *si* and $Q$ scores were plotted against estimates of within-population variation. Gaussian (normal) curves were then fitted to these data using the iterative least-squares method in SIGMAPLOT, as described above for the sigmoid fits. To accommodate different shaped Gaussian distributions, the fit of a three-parameter, $y = a*\exp(-0.5*((x-x_0)/b)^2)$, and a four-parameter function, $y = y_0 + a*\exp(-0.5*((x-x_0)/b)^2)$, were compared.

The distribution of phenotypes at the centre of a cline can provide information about the forces that are maintaining it (Jiggins & Mallet 2000; Gay *et al.* 2008). Where a hybrid zone is maintained by weak to moderate postzygotic selection rather than prezygotic barriers, a unimodal distribution of phenotypes is predicted (Jiggins & Mallet 2000). Where postzygotic selection is strong, or prezygotic barriers are present, the distribution of phenotypes is expected to become flatter and, ultimately, bimodal (Jiggins & Mallet 2000). Here, the distribution of phenotypes at the centre of the cline was examined by producing a frequency histogram of spire index scores. Care must be taken when pooling samples from different locations within a hybrid zone because the mean trait value can vary over a fine spatial scale, which would always have the effect of making a zone appear 'flatter', or more bimodal. To reduce this noise, I included the three sites closest to the centre of each transect included in the one-dimensional collapse, totalling 15 sites. A three-parameter Gaussian function (as described in the above section) was used to test the fit of a unimodal distribution.

### *Analysis of mitochondrial data*

Relationships between mtDNA haplotypes were inferred via a minimum-spanning network, calculated in ARLEQUIN v.3.11 (Excoffier *et al.* 2005) and drawn in HAPSTAR (Teacher & Griffiths 2011). To determine the pattern of differentiation between the divergent shell forms, the pairwise estimates of $\Phi_{ST}$ were calculated between the South Keeled-Flat, Globose and North Keeled-Flat groups (as established for the analysis of microsatellite data) based on the number of pairwise differences, and their significance determined via 10 000 permutations of the data.

If a cline has evolved recently in the presence of gene flow, there should be little or no association between morphological variation and the pattern of mitochondrial evolution (Avise 2000). To test this hypothesis, Analysis of Molecular Variance (AMOVA) was used to partition COI haplotype variation (i) among the a priori groups, (ii) among sample sites within groups and (iii) within sample sites across the collapsed transect. Two AMOVAS were conducted in ARLEQUIN. The first analysis included the Globose, Intermediate and South Keeled-Flat groups, totalling 59 sites. In the second analysis, the Intermediate group was excluded, which reduced the number of sites to 48. Sites with a single genotyped individual were excluded from AMOVA analyses. The significance of variation at each level was estimated by 10 000 permutations.

### *Collection and analysis of habitat data*

If a cline is maintained by divergent ecological selection favouring different genotypes in different environments, its position should coincide with a sharp transition in the local habitat. To test this hypothesis, environmental data were collected at each site included in the collapsed transect, with the exception of the eight sites sampled in 2008. Measurements were also made at two sites at the northern area where

flat-shelled *Rhagada* were found, taking the total number of sites with habitat data to 60.

The rocky hilltop and spinifex plain environments differ in several ways, including the rockiness of the substrate, elevation, soil type and vegetation, and would almost certainly differ in ways that are not immediately discernable, for example, the microclimate and faunal composition. Here, two conspicuous environmental variables were selected as surrogates to determine positions and scale of the transition between the contrasting environments. The first was the 'rockiness' of each site, which was quantified by estimating the rugosity. Rugosity, $f_r$, is a measure of fine-scale variation in the amplitude of a surface and is calculated as $f_r = A_r/G_r$, where $A_r$ is the actual surface area and $G_r$ is the geometric surface area. At each site, a 3-m-metal chain, with links similar to the diameter of a *Rhagada* shell (15 mm), was laid out in a straight line, but allowed to conform to the vertical profile of the substrate. $G_r$ was then obtained by measuring the 'new' geometric length of the chain. The process was repeated five times within each site, and the mean estimates of rugosity were arcsine square root transformed. The second measured variable was the elevation of each site, estimated to the nearest metre using a hand-held GPS.

Preliminary analysis revealed a strong correlation between the rugosity and elevation scores, so that the sites at higher elevations tended to be rockier. To reduce the number of variables in further analyses, the variables were summarized into a single principal component (PC1) that explained over 80% of the total variance. Simple linear regression was used to test the relationship between habitat PC1 score and mean *si* scores. Spatial variation in habitat was also examined across the collapsed transect by plotting habitat PC1 scores against the distance of each site from the fitted cline centre. A four-parameter sigmoid curve was fitted as described for the morphological data, and estimates of the cline centre and width were obtained as described above for the morphological data.

### *Analysis of population density data*

If a cline is maintained primarily by intrinsic hybrid unfitness rather than divergent environmental selection, its position is expected to coincide with an area of low population density, either because it is pushed into an area of unsuitable habitat by a net flux of genes (Barton & Hewitt 1985) or because selection against hybrids is strong relative to the rate of recombination (hybrid sink effect; Barton & Hewitt 1985). To determine whether this cline runs through an area of low density, estimates of population density were obtained at each of the sites along the collapsed transect, with the exception of the eight sites sampled in 2008 ($n = 54$). One-way ANOVA was used to compare mean estimates of density inside and on either side of the phenotypic cline. Two separate analyses were conducted. The first, performed on the density of live snails per square metre, including snails from all age classes, is an appropriate test if a density trough is driven by an area of unsuitable habitat, because all age classes should be affected, independent of their genetic background. The second, conducted on estimates of juvenile density per square metre, is more appropriate if the reduction in density is driven by a hybrid sink effect, because adult density is likely to be buffered by dispersal into the cline (Nichols & Hewitt 1988).

### *Estimate of the rate of dispersal*

A time-scaled estimate of dispersal is required to estimate the strength of selection required to maintain a cline with a given width. Dispersal rates are difficult to measure and can vary in space and time (Rousset 1997, 2003). Direct estimates, such as those obtained from mark–recapture, tend to underestimate dispersal, because occasional long-distance migrants are easily missed (Rousset 2003). Estimates based on genetic markers are more likely to reflect true patterns (Rousset 2003).

Here, the dispersal rate, $\sigma$, defined as the mean axial parent–offspring distance per generation$^{-1/2}$ (Rousset 1997), was estimated from the microsatellite data under an isolation-by-distance model. In a two-dimensional space, the slope of the regression between $F_{ST}/(1-F_{ST})$ and the natural logarithm (ln) of geographic distance provides an estimate of $1/(4D\pi\sigma^2)$, where $D$ is the local population density per generation and $\sigma^2$ is the average squared mean axial parent–offspring distance (Rousset 1997). As the model assumes that population differentiation results only from migration and drift, the analysis was restricted to a subset of 12 'pure' populations on one side of the cline. As it is preferable to confine samples to an area approximately $10\sigma \times 10\sigma$, and not more than $20\sigma$ (Rousset 2003), an estimate of $\sigma = 23$ m [the mean dispersal rate of land snails in Endler (1977)] was used as a guide during site selection. Pairwise estimates of $F_{ST}$ were calculated as described by Weir & Cockerham (1984) using GENEPOP v. 3.3 (Raymond & Rousset 1995). The significance of the regression between pairwise estimates of $F_{ST}/(1-F_{ST})$ and ln geographic distance was tested using a Mantel test (999 permutations) in GENALEX6 (Peakall & Smouse 2006).

With the slope of the regression obtained, an estimate of population density per generation, $D$, is required to derive an estimate of $\sigma^2$. First, the mean observed

density of live adult snails per square metre was obtained at each site. Standardizing this multigenerational estimate per generation requires some knowledge of the generation time and lifespan of the organism. Based on a mark–recapture study of the congener *Rhagada capensis* in a similar climate and habitat, Johnson & Black (1991) estimated that it takes five years to reach adulthood and that adults live for an average of five years, with a single cohort produced each year. Assuming that these values are accurate for these populations (i.e. that the observed density of live adults is the sum of the adults across five consecutive generations), $D$ was estimated as 0.2 of the mean observed density of adult snails per square metre.

## Results

### *Variation in shell shape*

The individual transects revealed considerable variation in the transition between the globose and flat-spired forms (Table 1; Fig. 2). Five transects, all from the southern transition zone, showed continuous variation in shell shape, while sharp discontinuities were observed along the other four. These contrasting patterns were associated with the position and continuity of suitable habitat across each transect. All four transects without intermediate phenotypes were interrupted by a barrier of exposed rocky substrate, absent vegetation. Thorough searches of these areas failed to find snails.

The five transects with continuous variation had steep clines in shell shape (Fig. 2; Table 1). The four-parameter sigmoid function provided an excellent summary of each transition, with $r^2$ values ranging from 0.777 to 0.894. There were, however, some striking differences among clines, including the range of variation captured by each transect ($\Delta_{si}$ = 0.225–0.360) and the estimate of the *si* score at each cline centre ($c_{si}$ = 0.490–0.504). The most striking variation among transect was in cline width, which varied 4.5-fold, from 76.0 m to 343.5 m (mean $w$ = 202.4 m ± SD 114.64).

**Table 1** Estimates of cline parameters obtained from the analysis of *si* scores along each cline. $r^2$, the adjusted $r^2$ value obtained from the best sigmoid fit; $c$, position of the cline centre; $c_{si}$, *si* score at the cline centre; $\Delta si$, change in the *si* score across the cline; $w$, cline width

| Transect | $r^2$ | $c$ ± SE (m) | $c_{si}$ | $\Delta_{si}$ | $w$ (m) |
|---|---|---|---|---|---|
| 1 | 0.817 | 362.4 ± 5.4 | 0.490 | 0.246 | 92.8 |
| 2 | 0.777 | 469.3 ± 3.6 | 0.500 | 0.225 | 76.0 |
| 3 | 0.822 | 341.6 ± 18.2 | 0.495 | 0.236 | 343.5 |
| 4 | 0.805 | 320.9 ± 9.8 | 0.504 | 0.245 | 259.5 |
| 5 | 0.894 | 360.4 ± 6.9 | 0.548 | 0.360 | 240.0 |
| 1D collapse | 0.890 | −3.9 ± 8.8 | 0.502 | 0.246 | 169.7 |

Despite the variation in characteristics of the individual clines, the sigmoid function explained a large proportion of the site-level variation along the collapsed transect ($r^2$ = 0.890, $P \leq 0.0001$; Fig. 3A). The estimated position of the cline centre was −3.9 m from the fitted cline centre, and the cline width ($w$ = 169.7 m) was similar to the mean width calculated from the individual clines. Likewise, the estimate of shell shape at the centre ($c_{si}$ = 0.502) fell within range of estimates obtained from the individual transects.

Phenotypic variation increased at the cline centre (Fig. 3B). Within-site variation was similarly low at both ends of the cline, with the CV falling consistently between 3 and 9% of the mean *si* score. Variation increased with distance towards the cline centre, reaching a maximum of 18.4%, with most values above 9.5%. A four-parameter Gaussian function explained over 42% of the variation between the mean *si* and CV scores ($r^2$ = 0.427, $P$ = 0.0001).

There was no evidence for a bimodal distribution of spire index scores at the cline centre (Fig. 4). Rather, there was a single unimodal distribution. The three-parameter Gaussian function explained over 87% of the variation ($r^2$ = 0.876, $P < 0.0001$); the peak of function coinciding with the intermediate range of spire index scores.

### *Microsatellite differentiation*

The STRUCTURE analysis revealed clear subdivision, indicating that $K$ = 2 was most probably ($\Delta K$ = 243.93; Table S3, Supporting information). Although most individuals were assigned with a high probability to one cluster (70% of individuals had $Q$ scores >0.85), there was a complete continuum of $Q$ scores rather than two distinct groups (Fig. 5A). Among sites, there was a strong linear relationship between the mean $Q_{K2}$ scores and the mean spire index ($r^2$ = 0.691, $P < 0.0001$), as the probability of membership to *K*2 increased with decreasing shell height. However, the flat-spired populations from the northern rockpile departed strongly from this relationship, both being assigned to *K*1 with a very high level of probability (mean $Q_{K1}$ = 0.94 and 0.93). Additional STRUCTURE analyses of $K$ = 3 and $K$ = 4 failed to distinguish the northern flat and globose samples. Comparison of the prediction surfaces generated from the mean $Q$ (Fig. 5B) and *si* scores (Fig. 1) confirmed the close spatial relationship between the geographic distribution of *K*2 and the southern flat-shelled *Rhagada*, and the local incongruence between the mean *si* and $Q_{K2}$ scores, as the flat-shelled snails from the

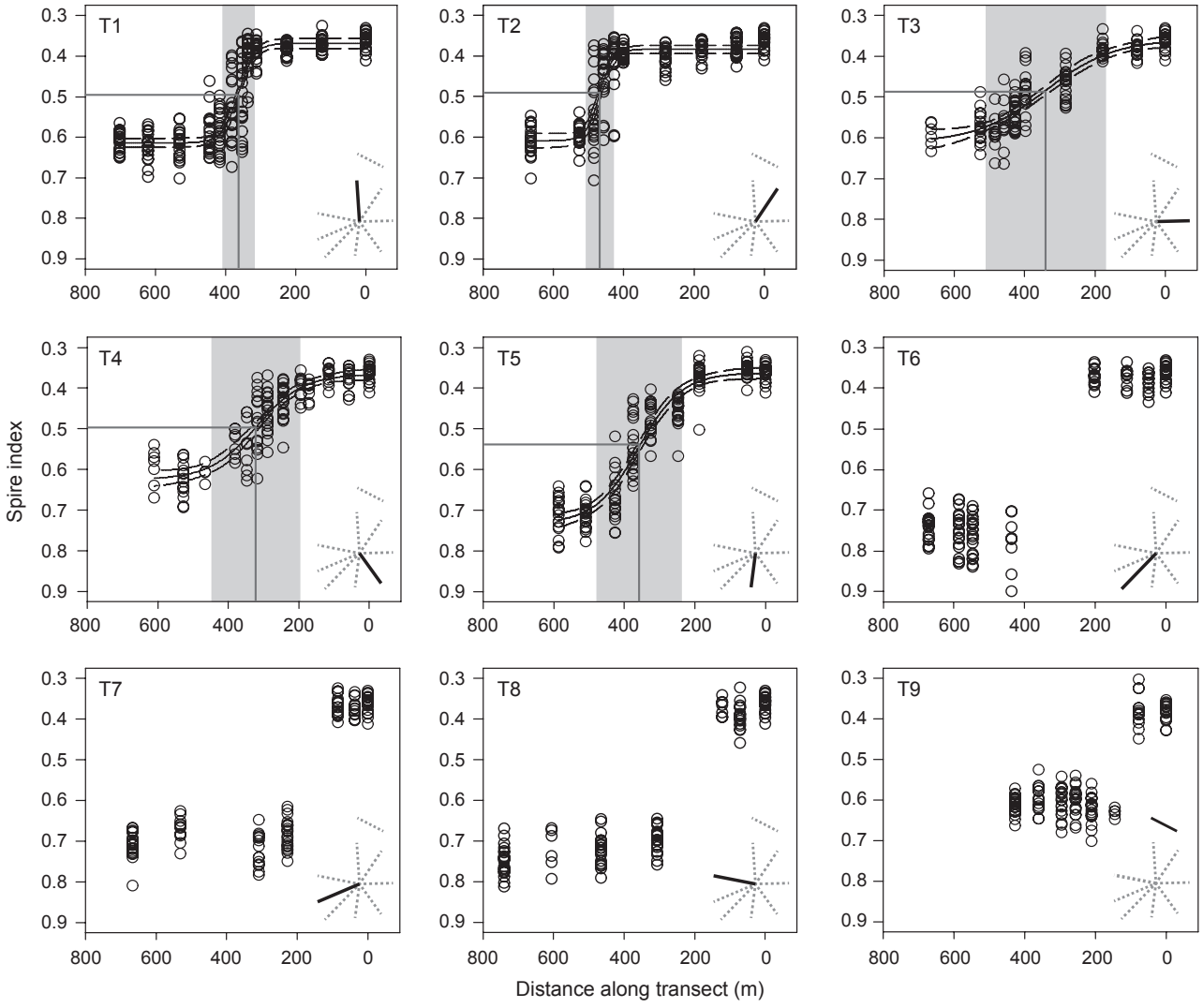


**Fig. 2** Variation in shell shape (*si* score) along each transect. The diagram in the bottom right corner of the plot shows the geographic position of each transect (as in Fig. 3). The regression lines are a fourparameter sigmoid function (with 95% intervals) fitted using an iterative least-squares method. The solid vertical and horizontal lines represent the geographic cline centre and *si* score at the cline centre, respectively. The shaded gray areas represent the estimated cline width. Each point represents a single shell.

northern rockpile were indistinguishable from the surrounding globose-shelled populations.

The pairwise estimates of $F_{ST}$ agreed with the results of the STRUCTURE analysis (Fig. 6). There was a positive linear relationship between $F_{ST}/(1\text{-}F_{ST})$ between Globose samples ($r^2 = 0.323$; Mantel $P \leq 0.0001$; Fig. 5A), with the divergence increasing at an average rate of 3.93E−05/m. Almost all of the comparisons between South Keeled-Flat samples fell within that band of estimates. The same pattern was observed for the comparisons between the Globose and North Keeled-Flat samples, indicating that they were not more divergent than expected once the effects of space were taken into consideration. In contrast, the comparisons between the Globose and South Keeled-Flat samples were more differentiated than expected based on the pattern of isolation by distance, differing by a mean estimate of 0.095 (Fig. 6B). The most striking result was that the strong divergence between the South Keeled-Flat and North Keeled-Flat groups, which were more divergent from one another (0.149) than either was to the globose-shelled group.

STRUCTURE scores varied clinally along the collapsed transect (Fig. 7A). The four-parameter sigmoid function explained 69% of the variation between the distance from the fitted morphological cline centre and mean $Q_{K2}$ score. The cline was approximately 2.8 times as wide as the cline in *si* scores (Table 2), although the estimates of the cline centre were not significantly different ($P > 0.05$). There was increased within-site varia-

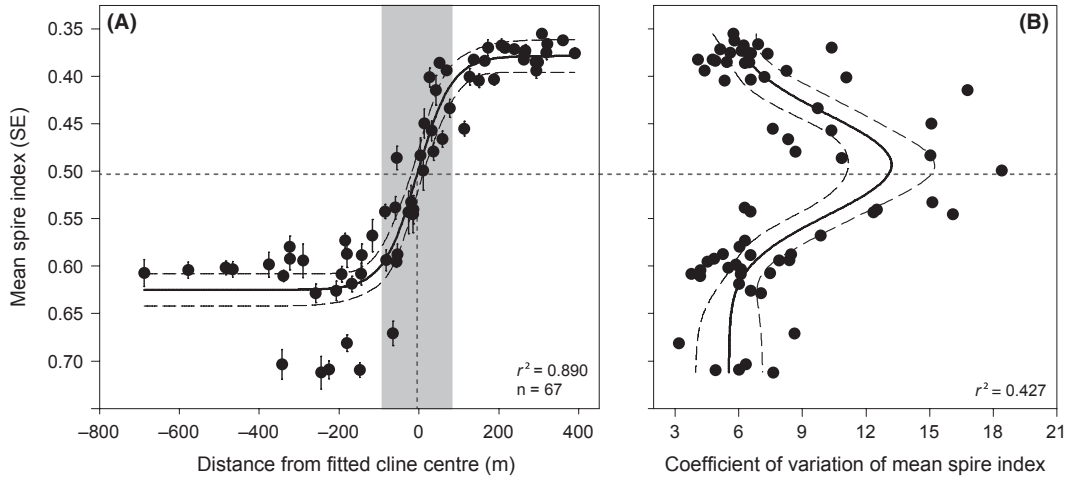


**Fig. 3** Variation in (A) shell shape (mean *si* score ± SE) and (B) within-population variation (CV) along the collapsed transect. The regression line in plot A is a four-parameter sigmoid function (with 95% confidence intervals) fitted using an iterative least-squares method. The vertical and horizontal dashed lines represent the geographic cline centre and *si* score at the cline centre, respectively. The shaded gray area represents the estimated cline width. The regression line in plot B is a four-parameter Gaussian function (with 95% confidence intervals) fitted using an iterative least-squares method.

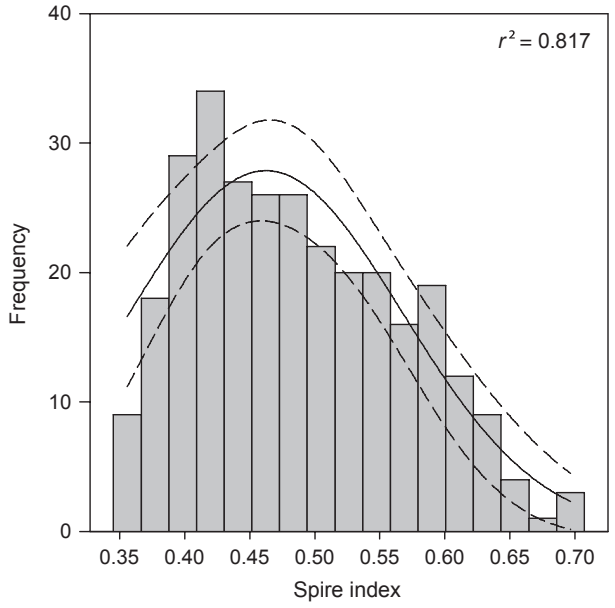


**Fig. 4** Frequency histogram of spire index scores at the cline centre. The regression line is a threeparameter Gaussian function (with 95% confidence intervals) fitted using an iterative least-squares method.

tion in samples with intermediate *Q* scores (Fig. 6B). A three-parameter Gaussian function explained 77% of the variation between the mean $Q_{K2}$ score and SD.

### *Mitochondrial variation*

Mitochondrial diversity was high, with 117 COI haplotypes recovered from 269 snails (Fig. 8). The minimum-spanning network revealed complex but close relationships between most haplotypes and many alternative connections. Half the connections (49%) were single substitution (0.15% divergence), and 83% consisted of three or fewer substitutions (0.45% divergence). Moreover, there was no strong association between molecular and phenotypic variation along the collapsed transect, as would be expected if the divergent shell forms had evolved in isolation. An AMOVA revealed that the Globose, Intermediate and South Keeled-Flat groups explained less than 4.1% of total molecular variance, with over 86.7% distributed within sample sites. The remaining 9% was distributed among sample sites within groups. A second AMOVA, excluding the Intermediate group, saw the among-group variation rise to 8.0%, while majority (84.07%) remained distributed within sample sites (Table 3).

Population estimates of mitochondrial differentiation between the shell groups supported the conclusion of the STRUCTURE and $F_{ST}$ analyses. Although the South Keeled-Flat and North Keeled-Flat populations were indistinguishable based on their *si* scores, estimates of $\Phi_{ST}$ indicated that they were more different from one another than they were from the globose-shelled group (Fig. 8).

### *Variation in local habitat*

There was a strong relationship between shell shape and habitat, with the habitat PC1 scores explaining 57% of the variation in the mean spire index ($r^2$ = 0.573,

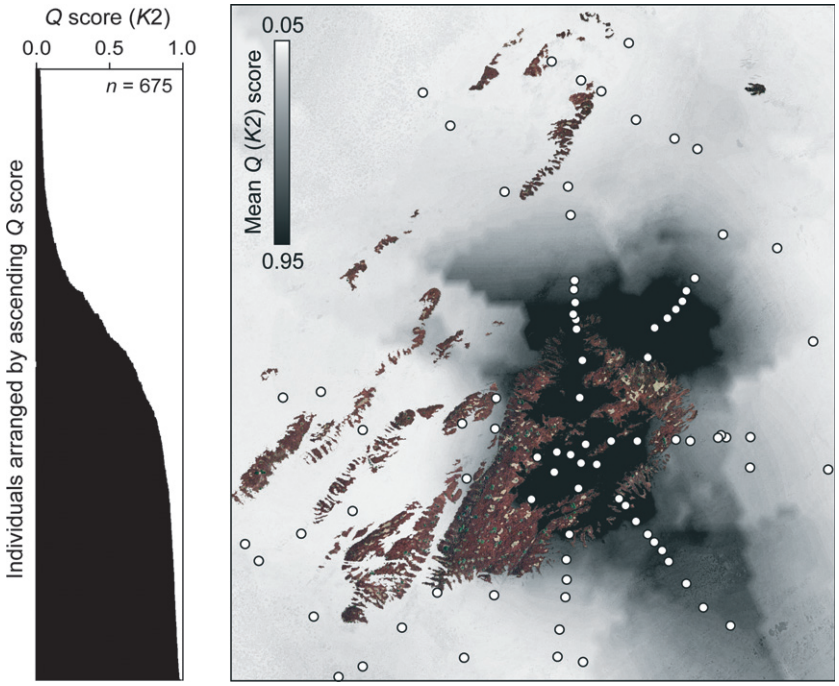


**Fig. 5** Population genetic structure within the study area, inferred from multilocus microsatellite genotypes using the program STRUCTURE ($n$ = 675), in which $K$ = 2 was most likely. (B) Prediction surface illustrating the geographic distribution of the clusters based on mean $Q_{K2}$ scores at each sample locality. Areas not included in the prediction surface lacked adequate vegetation, and were not occupied by snails.

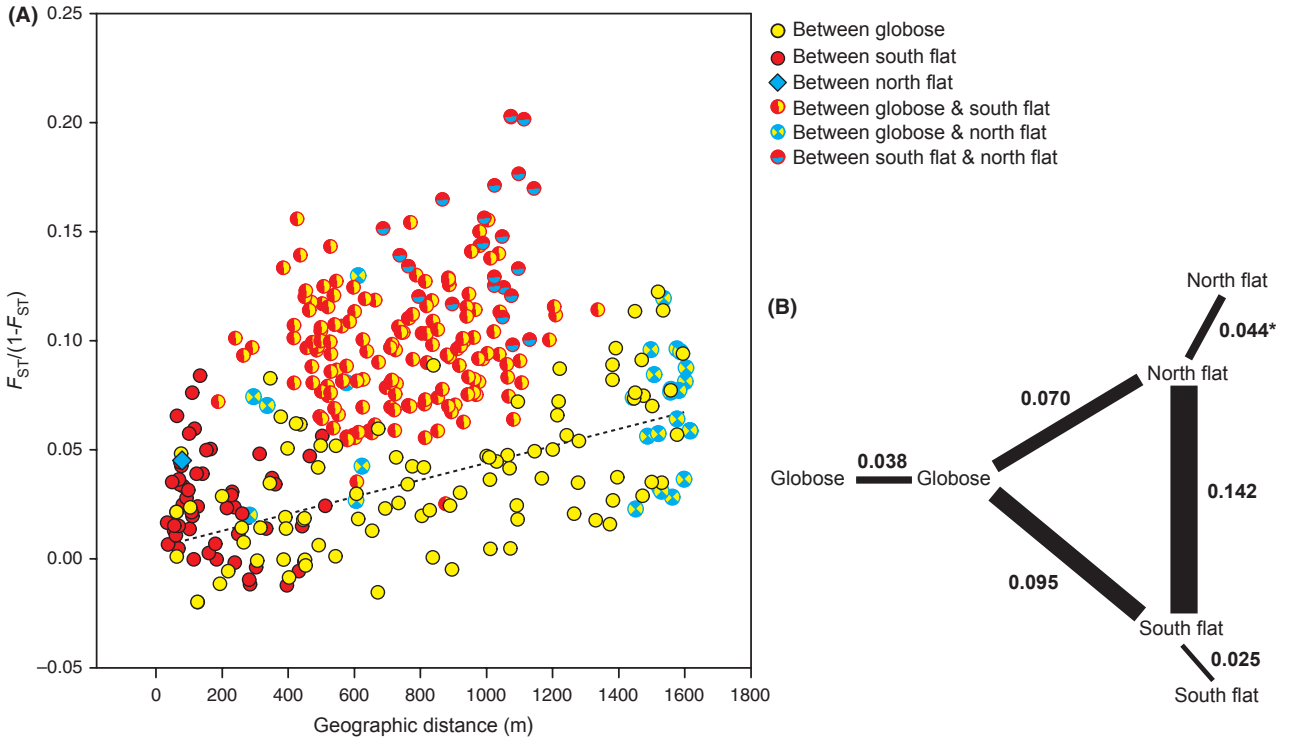


**Fig. 6** Pairwise estimates of divergence based on microsatellites genotypes. (A) Pairwise estimates of $F_{ST}(/1-F_{ST})$ averaged over loci versus geographic distance between sample sites. The regression line (y = 3.93E−05x + 0.0046; $r^2$ = 0.323) is through the comparisons between Globose samples, and is significant based on a Mantel test ($P \leq 0.0001$). (B) mean pairwise divergence within and between the South Keeled-flat, Globose and North Keeled-flat groups. The length and weight of the lines are proportional the level of divergence.

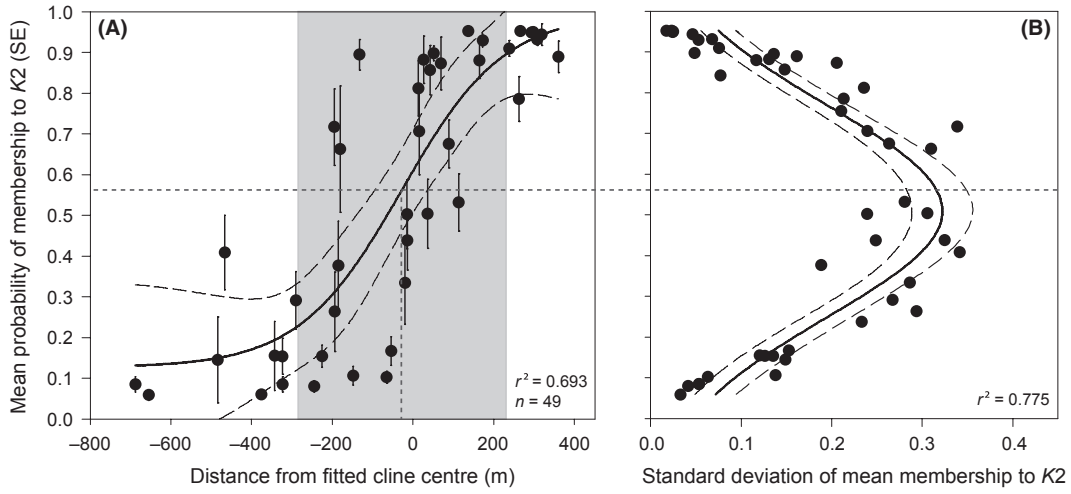


**Fig. 7** Variation in (A) the mean $Q_{K2}$ score (±SE) and (B) within-population variation (SD) along the collapsed transect. The regression line in plot A is a four-parameter sigmoid function (with 95% confidence intervals) fitted using an iterative least-squares method. The vertical and horizontal dashed lines represent the geographic cline centre and $Q_{K2}$ score at the cline centre, respectively. The shaded gray area represents the estimated cline width. The regression line in plot B is a three-parameter Gaussian function (with 95% confidence intervals) fitted using an iterative least-squares method

**Table 2** Estimates of the geographic position of the cline centre and cline width obtained for habitat PC1 score, mean *si* scores and mean $Q_{K2}$ scores across the collapsed transect

| | $c \pm$ SE (m) | w (m) |
|---|---|---|
| Habitat PC1 | 38.2 ± 10.3 | 44.0 |
| *si* | −3.9 ± 8.8 | 169.7 |
| $Q_{K2}$ | −28.3 ± 57.8 | 480.7 |

$P \leq 0.0001$; Fig. 9A). Snails from rockier, more elevated sites had flatter shells. Populations with extreme *si* scores clustered tightly on the basis of habitat, while those of intermediate shape were associated with a wider range of PC1 scores.

There was a sharp transition in habitat PC1 scores along the collapsed transect. The four-parameter sigmoid function explained 49% of the variation between the habitat PC1 scores and distance from the fitted cline centre ($r^2 = 0.494$, $P \leq 0.0001$; Fig. 9B). The transition was abrupt, with an estimated width of 44 m, which is approximately 25% of the width of the phenotypic cline and 9% of the cline in the $Q_{K2}$ scores. There was a bimodal distribution of PC1 scores within the transition zone, rather than intermediate scores, indicating that it is a mosaic of the alternative habitat types, rather than an area of intermediate habitat. The habitat transition occurred within the width of the phenotypic cline, although the centres, positioned 38 m apart (95% limits 22.89–60.51 m), were significantly different (Fig. 10).

### *Variation in population density*

Although the density of snails ranged from 0.07 to over 6.7/m$^2$ among the 58 sites, mean density was relatively uniform across the collapsed transect (Fig. 10). There was no significant difference in total density (outside globose = 0.66 ± SE 0.38/m$^2$; inside 0.85 ± SE 0.20/m$^{2;}$ outside keeled-flat = 1.35 ± SE 0.33/m$^2$; $F_{2,52} = 1.340$, $P = 0.271$) or juvenile density (outside globose = 0.35 ± SE 0.14/m$^2$; inside 0.22 ± SE 0.10/m$^{2;}$ outside keeled-flat = 0.35 ± SE 0.09/m$^2$; $F_{2,52} = 0.607$, $P = 0.548$) inside and outside the cline.

### *Estimate of dispersal*

There was a significant, positive linear relationship between $F_{ST}/(1-F_{ST})$ and the logarithm of geographic distance for the 12 'pure' keeled-flat populations ($r^2 = 0.119$, Mantel test $P = 0.046$; Fig. S1, Supporting information). Based on the slope of the relationship ($a = 0.0093$, 95% limits = 0.0061–0.0125) and the time-scaled estimate of density (0.0957 per m$^2$/generation), $\sigma$ was estimated at 9.45 (95% limits 7.84–12.52) m/generation$^{-1/2}$.

## Discussion

### *An ecological hybrid zone patchwork*

The globose and keeled-flat phenotypes are separated by a hybrid zone that is almost certainly maintained by a balance between dispersal and divergent selection. This inference is based on the shape of the phenotypic transition, its narrow width relative to the rate of

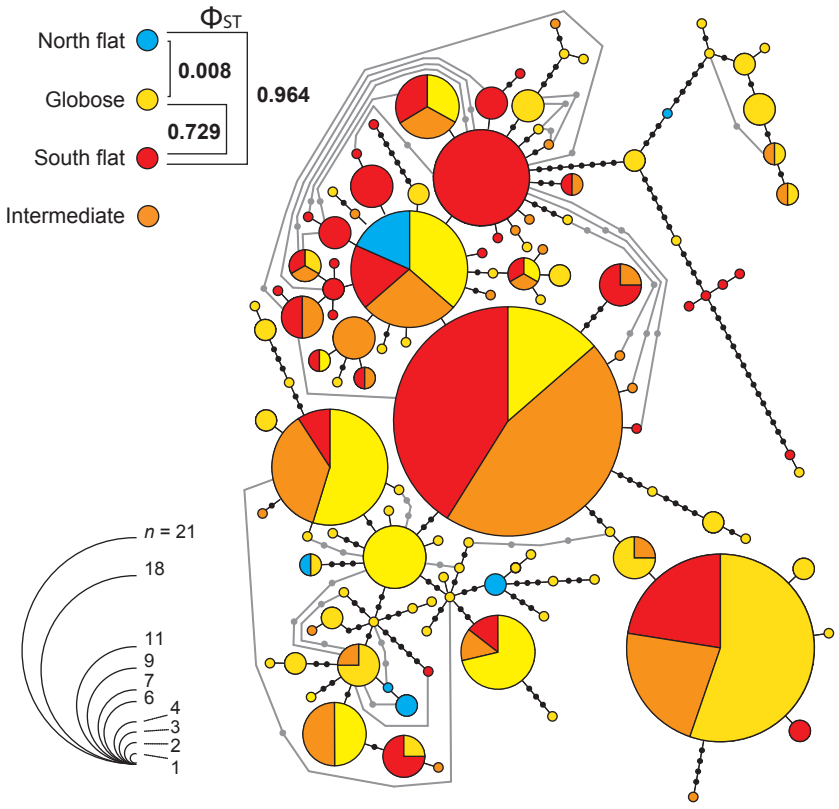


**Fig. 8** Minimum spanning network of Cytochrome Oxidase I (COI) haplotypes (666 bp) obtained from 269 snails. Each haplotype is represented by a circle, the size of which is proportional to its frequency. Each circle is divided into red, blue, orange and yellow segments, which show the frequency of that haplotype in South Keeled-flat, North Keeled-flat, Intermediate and Globose-shelled populations, respectively. Haplotypes connected by a line differ by a single mutational step. Black dots represent missing' haplotypes that were not sampled. Pairwise estimates of $\Phi_{ST}$, calculated between the South Keeled-flat, Globose and North Keeled-flat groups, are provided in the figure legend. The depth of the bracket is proportional to the degree of differentiation.

**Table 3** Results of analyses of molecular variances (AMOVAS) of mitochondrial sequence data across the collapsed transect

| | Level of comparison | df | ss | Variance component | % mol. variance | *P* |
|---|---|---|---|---|---|---|
| With intermediates | Among phenotypic groups | 2 | 29.64 | 0.16 | 4.13 | <0.0001 |
| | Among sites, within groups | 53 | 206.95 | 0.36 | 9.16 | <0.0001 |
| | Within sites | 144 | 485.53 | 3.36 | 86.71 | <0.0001 |
| | Total | 189 | 720.11 | 3.87 | | |
| Without intermediates | Among phenotypic groups | 1 | 27.69 | 0.32 | 8.03 | <0.0001 |
| | Among sites, within groups | 33 | 152.96 | 0.31 | 7.90 | <0.0001 |
| | Within sites | 110 | 367.71 | 3.34 | 84.07 | <0.0001 |
| | Total | 144 | 548.33 | 3.98 | | |

dispersal ($w = 17.9\sigma$), and the elevated variance at the centre (Barton & Hewitt 1985). Understanding the exact nature of the selection maintaining the morphological variation is difficult, because ecological and mutation-order models of hybrid zone maintenance can result in clines with similar shapes (Kruuk *et al.* 1999).

Taken together, these results suggest that the ecological setting plays a primary role in the maintenance of the cline in shell shape. Ecological models of cline formation and maintenance predict that a hybrid zone will settle on a narrow boundary between two contrasting environments (Endler 1977; Kruuk *et al.* 1999). Consistent with this expectation, the cline in shell shape coincides with a narrow ecotone, $4.6\sigma$ wide, which separates the contrasting spinifex plain and rocky hilltop environments. The pattern of microsatellite differentiation

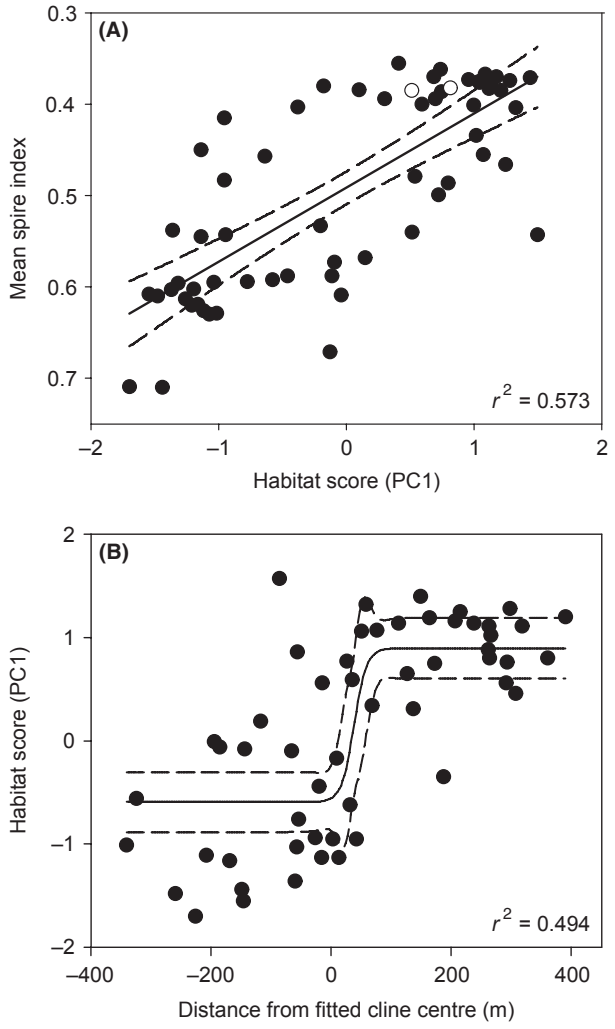


**Fig. 9** Association between (A) shell shape (mean si score) and habitat PC1 score and (B) habitat PC1 score position along the collapsed transect. The open symbols in plot A are from the northern rock outcrop. The regression line in plot B is a four-parameter sigmoid function (with 95% confidence intervals) fitted using an iterative least-squares method.

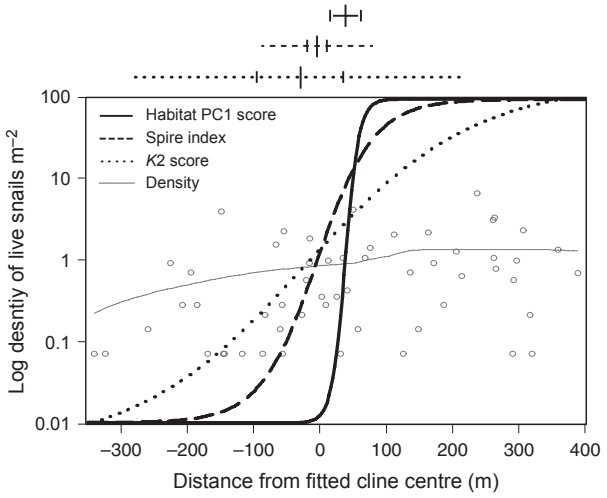


**Fig. 10** Variation in shell shape (*si* scores), neutral genetic variation ($Q_{K2}$ scores), local habitat (PC1 scores) and population density (log snails per square metre) across the collapsed transect. The horizontal lines above the plot represent the cline widths, while the central and outer vertical marks represent the cline centres and 95% confidence intervals around those estimates, respectively. The line through the estimates of density is a Loess spline.

across the zone also fits the expectations of ecological speciation theory. If the morphological cline was maintained by hybrid unfitness, as expected for mutation-order hybrid zones, one would expect the cline in molecular hybrid index scores to have a similar width (Barton & Hewitt 1985). Contrary to that expectation, the cline in shell shape is far narrower, indicating that its width is not governed by the pattern of recombination.

While the evidence for the ecological maintenance of the hybrid zone is compelling, there are alternative explanations for the association between shell shape and habitat that must also be considered. First, mutation-order hybrid zones, maintained by hybrid unfitness, tend to be pushed towards and trapped in areas of low populations density that often coincide with zones of environmental transition (Barton & Hewitt 1985; Hewitt 1988). This cline, however, is not positioned in an area of low density. Another possible explanation of the association of the cline with the ecotone is that it is a random coincidence. This also seems unlikely, given that the association between shell shape and habitat occurs at two separate locations on Rosemary Island and has been observed in other species of land snail with keeled-flat populations (Teshima *et al.* 2003). Keeled-flat shells are thought to be an adaptive response to dry, open habitats (Goodfriend 1986; Elejalde *et al.* 2008; Fiorentino *et al.* 2012). On Rosemary Island, the environment is arid, and snails are highly susceptible to desiccation if they are exposed during dry periods (Solem 1997). Keeled-flat *Rhagada* are often found aestivating in narrow rock crevices, while globose-shelled *Rhagada* typically burrow into the soil (personal observation). Thus, flat-spired shells may reflect the evolution of an alternative aestivation strategy in this genus.

### *Local divergence in the face of gene flow*

Unlike mutation-order speciation, which occurs when populations are physically isolated, ecological speciation can occur in the presence of gene flow, although proceeds most easily when gene flow is absent (Rundle & Nosil 2005; Schluter 2009). The results of this study suggest that the divergence on Rosemary Island is recent and has occurred without major historical disruptions

to gene flow. A previous analysis of mtDNA variation in these snails provided no evidence that the derived flat-shelled phenotype evolved during a period of physical isolation, but did reveal patterns of differentiation consistent with current and historical marine barriers, suggesting that the COI sequences provide a resolution of at least 8 ky (Stankowski 2011). The pattern of mitochondrial evolution revealed here, based on longer COI sequences and within-population sampling, provides strong support for the original conclusion. While impossible to rule out the possibility that the morphological divergence occurred in physical isolation, the mitochondrial data indicate that any period of isolation must have been relatively short.

The pattern of microsatellite DNA divergence provides strong evidence that the morphological divergence occurred in the presence of gene flow. There is weak divergence between the South Keeled-Flat and Globose-shelled populations, and the North Keeled-Flat and Globose populations have indistinguishable neutral genetic backgrounds. More striking is that the two keeled-flat populations, isolated on separate rock outcrops (~63$\sigma$ apart), are more genetically divergent from one another that they are from the globose form. A similar pattern of divergence has been observed in the threespine stickleback, where differentially adapted ecotypes occupy marine and freshwater habitats (Schluter & Conte 2009). Where habitats merge into one another, hybrids are present. The association between morph and habitat is repeated in many locations and at a great range of geographic scales. At the broad scale, the pattern of neutral divergence reflects geographic location rather than ecotype, while, at the local scale, there is evidence for fine-scale molecular divergence between ecotypes. As with the threespine stickleback, the pattern of differentiation in *Rhagada* on Rosemary Island is explained most simply by the repeated evolution of the keeled-flat ecotype. Based on the pattern of microsatellite and mtDNA variation, it is most parsimonious to conclude that the flat-spired form evolved on the southern hill first, because there is significant molecular genetic divergence between the ecotypes.

The subsequent evolution of the same form on the northern hill would not necessarily require the *de novo* evolution of the genetic variation itself. Rather, it may have evolved from standing genetic variation, as has been suggested to explain the repeated evolution of the freshwater stickleback ecotype (Schluter 2009). In this scenario, referred to as the 'transporter' hypothesis (Schluter 2009), the hybrid zone breaks down ecotypes and quantitative traits into their building blocks. Alleles from one ecotype can recombine onto the alternative genetic background where they may pass through areas of low fitness. Selection and recombination may drive low frequency alleles to fixation in their favoured ecological setting, leading to the reassembly of the ecotype in a new location.

### *Progress of speciation*

Recent commentary on the nature of species and speciation has seen a shift from the idea that species are 'real' distinct biological entities (Mallet 2008). Instead, greater emphasis is placed on how different mechanisms contribute to the formation of reproductive barriers (Mallet 2008; Schluter 2009). In ecological speciation, ecologically based divergent selection is the primary reproductive barrier. It acts as a legitimate isolating mechanism, reducing the rate of recombination between divergent ecotypes (Nosil *et al.* 2005), and can drive the evolution of additional intrinsic isolating mechanisms through reinforcement (Rundle & Nosil 2005).

These data suggest that the *Rhagada* of Rosemary Island are in the early stages of ecological speciation. There is no evidence to suggest that the divergent ecotypes are separated by a significant intrinsic postzygotic barrier, and there is a unimodal distribution of phenotypes at the centre of the hybrid zone, suggesting that prezygotic have not evolved, or that they have little effect in the natural setting (Jiggins & Mallet 2000). Thus, it is most parsimonious to assume that divergent ecomorphs are maintained primarily by ecologically dependent postzygotic isolation (Rundle & Nosil 2005), as is the case for most ecological hybrid zones (Schluter 2009). Ultimately though, experimental studies are required to understand how different isolating mechanisms contribute to the reproductive barrier.

Assuming that this conclusion is correct, it is possible to estimate the rate of selection against hybrids that is required to generate the partial reproductive barrier. The effective selection pressure at the centre of a cline, $s^*$, required to maintain a cline with a width of $w$ can be estimated from the equation $s^* = 8\sigma^2/w^2$, where $\sigma$ is the mean axial parent–offspring distance per generation$^{-1/2}$ (Bazykin 1969; Szymura & Barton 1986). While developed from a single-locus model, the equation provides meaningful estimates in multilocus systems providing that selection is weak to moderate and linkage is loose (Kruuk *et al.* 1999). Based on the estimate of $\sigma$ obtained here (9.45 m/generation$^{-1/2}$, 95% limits = 7.84–12.52), and the width of the collapsed cline (169.7 m), $s^*$ = 2.5% per gen. (1.8–4.4%). This relatively weak estimate of selection is consistent with the abundance of later-generation hybrid phenotypes and genotypes and uniform population density across the zone. It is, however, worth noting that this estimate is also accurate if the cline is maintained by intrinsic postzygotic isolation, because the relationships between selection, dispersal and cline width are mathematically very similar for zones

maintained by hybrid unfitness or divergent ecological selection (Kruuk *et al*. 1999)

What effect does the ecological-dependent postzygotic barrier have on loci that are not under selection? In ecological hybrid zones, selection is acting on genes of ecological importance, but neutral loci experience indirect selection if they are linked to functional loci or are spatially associated with differentially adapted alleles. For rapidly evolving markers such as microsatellites, where newly evolved alleles may spend considerable time in their native ecological setting before they cross the hybrid zone, ecological selection can drive neutral divergence. Assuming that the neutral marker cline between the South Keeled-Flat and Globose ecotypes evolved under this scenario, which is reasonable given that there is no such differentiation between the North Keeled-flat and Globose populations, it is possible to describe the strength of the barrier to the diffusion of neutral alleles. Based on the slope of the relationship between pairwise genetic and geographic distance observed between globose-shelled populations, the mean divergence across the hybrid zone ($F_{ST}$ = 0.095; 95% limits 0.092–0.098) is expected between samples separated by 2290 linear m (2226–2370 m). Thus, while the zone has a physical width of only $w$ = 170 m, it can be thought of as a barrier to neutral gene flow with the dimensions of distance = 13.5$w$, which, in these snails, is equivalent to 242 generations of dispersal on a linear axis.

## Conclusion

The *Rhagada* of Rosemary Island provide a complex example of ecological speciation in its early stages. Progress towards complete speciation is weak, with the reproductive barrier generated primarily by divergent ecological selection against hybrid immigrants. As closely related diverse group, they provide excellent material to study the genetics of ecotype formation and to test several aspects of ecological speciation theory.

## Acknowledgements

I thank my supervisor, Michael S. Johnson, for advice and stimulating discussion, Jeremy Feichtinger and Robert MacDonald for assistance with the fieldwork, and Simon Smith, Robert Black and Danielle Stephens for advice and discussion. Viviana Fiorentino, Menno Schilthuizen, an anonymous reviewer and the subject editor, Angus Davison, provided comments that greatly improved the manuscript. The Department of Environment and Conservation, Western Australia (Karratha) provided transport to Rosemary Island. Financial support was provided by Woodside Energy Ltd, Rio Tinto, The University of Western Australia School of Animal Biology, an Australian Postgraduate Award, and the Department of Environment and Conservation, Western Australia.

## References

Avise JC (2000) *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge.

Barton NH (2001) The role of hybridization in evolution. *Molecular Ecology*., **10**, 551–568.

Barton NH, Gale KS (1993) Genetic analysis of hybrid zones. In: *Hybrid Zones and the Evolutionary Process* (ed Harrison R. G.), pp. 13–45. Oxford University Press, Oxford, UK.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology, Evolution and Systematics*, **16**, 113–148.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–502.

Bazykin AD (1969) Hypothetical mechanism of speciation. *Evolution*, **23**, 685–687.

Cain AJ (1977) Variation in the spire index of some coiled gastropod shells, and its evolutionary significance. *Philosophical Transactions of the Royal Society of London B*, **277**, 377–428.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.

Earl DA, van Holdt BM (2011) Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Elejalde MA, Madeira MJ, Munoz B, Arrébola JR, Gómez-Moliner BJ (2008) Mitochondrial DNA diversity and taxa delineation in the land snails of the *Iberus gualtieranus* (Pulmonata, Helicidae) complex. *Zoological Journal of the Linnean Society*, **154**, 722–737.

Endler JA (1977) *Geographic Variation, Speciation and Clines*. Princeton University Press, Princeton, New Jersey.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the program STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, **1**, 47–50.

Fiorentino V, Manganelli G, Giusti F, Tiedmann R, Ketmaier V (2012) A question of time: the land snail Murella muralis (Gastropoda: Pulmonata) reveals constraints on past ecological speciation. *Molecular Ecology*, **1**, 170–186.

Gay L, Crochet PA, Bell DB, Lenormand T (2008) Comparing clines on molecular and phenotypic traits in hybrid zones: a window on tension zone models. *Evolution*, **62**, 2789–2806.

Gittenberger E, Piel WH, Groenenberg D (2004) The Pleistocene glaciations and the evolutionary history of the polytypic snail species *Arianta arbustorum* (Gastropoda, Helicidae). *Molecular Phylogenetics and Evolution*, **30**, 64–73.

Goodfriend GA (1986) Variation in land-snail shell form and size and its causes: a review. *Systematic Zoology*, **35**, 204–223.

Hewitt GM (1988) Hybrid zones - natural laboratories for evolutionary studies. *Trends in Ecology and Evolution*, **3**, 158–167.

Hewitt GM (2011) Quaternary phylogeography: the roots of hybrid zones. *Genetica*, **139**, 617–638.

Ivanova N, deWaard JR, Hebert PND (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes*, **6**, 998–1002.

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

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution*, **15**, 250–255.

Johnson MS, Black R (1991) Growth, survivorship and population size in the land snail *Rhagada convicta* Cox, 1870 (Pulmonata: Camaenidae) from a semiarid environment in Western Australia. *Journal of Molluscan Studies*, **57**, 367–374.

Johnson MS, Hamilton ZR, Teale R, Kendrick PG (2012) Endemic evolutionary radiation of *Rhagada* land snails (Pulmonata: Camaenidae) in a continental archipelago in northern Western Australia. *Biological Journal of the Linnean Society*, **106**, 316–327.

Kruuk LEB, Baird SJE, Gale KS, Barton NH (1999) A comparison of multilocus clines maintained by environmental adaptation or selection against hybrids. *Genetics*, **153**, 1959–1971.

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

Marquardt DW (1963) An algorithm for least squares estimation of parameters. *Journal of the Society for Industrial and Applied Mathematics*, **11**, 431–441.

Nichols RA, Hewitt GM (1988) Genetical and ecological differentiation across a hybrid zone. *Ecological Entomology*, **13**, 39–49.

Nosil P, Vines TH, Funk DJ (2005) Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution*, **59**, 705–719.

Okajima R, Chiba S (2009) Cause of bimodal distribution in the shape of a terrestrial gastropod. *Evolution*, **63**, 2877–2887.

Peakall R, Smouse PE (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

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

Raymond M, Rousset F (1995) Genepop: population genetics software for exact tests and eucumenicism. *Journal of Heredity*, **86**, 248–249.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rousset F (2003) Inferences from spatial population genetics. In: *Handbook of Statistical Genetics*, 2nd edn (eds Balding D. J., Bishop M. & Cannings C), pp. 681–712. John Wiley and Sons, Ltd, Chichester, UK.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336–352.

Schluter D (2009) Evidence for ecological speciation and its alternative. *Science*, **323**, 737–741.

Schluter D, Conte GL (2009) Genetics and ecological speciation. *Proceedings of the National Academy of Sciences*, **106**, 9955–9962.

Solem A (1997) Camaenid land snails from Western and Central Australia (Mollusca: Pulmonata: Camaenidae): VII taxa from Dampierland through the Nullarbor. *Records of the Western Australian Museum*, **50**, 1461–1906.

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

Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegata*, near Cracow in southern Poland. *Evolution*, **40**, 1141–1159.

Teacher AGF, Griffiths DJ (2011) Hapstar: automated haplotype network layout and visualization. *Molecular Ecology Resources*, **11**, 151–153.

Teshima H, Davison A, Kuwahara Y *et al*. (2003) The evolution of extreme shell shape variation in the land snail *Ainohelix editha*: a phylogeny and hybrid zone analysis. *Molecular Ecology*, **121**, 869–1878.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the estimation of population structure. *Evolution*, **38**, 1358–1370.

Welter-Schultes FW (2000) The pattern of geographical and altitudinal variation in the land snails *Albinaria idaea* from Crete (Gastropoda: Clausiliidae). *Biological Journal of the Linnean Society*., **71**, 237–250.

---

This work forms part of S.S. PhD thesis, which focused on the evolution of Rhagada land snails in the Dampier Archipelago, Western Australia. S.S. is interested the roles that history and selection play in the evolution of differentiated gene pools.

---

## Data accessibility

Lats, longs and mean spire (shell shape) index for each site; Individual shell measurements at each transect site; Microsatellite genotypes; Estimates of habitat variables at each site; Estimates of population density across the collapsed transect; Rosemary Rhagada COI sequences. All available from DRYAD doi: 10.5061/dryad.7fn18.

## Supporting information

Additional supporting information may be found in the online version of this article.

**Table S1** Characteristics of eight microsatellite markers, designed for *Rhagada convicta* (cox), in a single population of *Rhagada* from Rosemary Island ($n$ = 46 individuals).

**Table S2** Tests for Linkage disequilibria between pairs of loci in a single population of Rhagada from Rosemary Islands.

**Table S3** Results of the Evanno method for $K1…12$, based on the analysis of 675 8-locus microsatellite genotypes in the program STRUCTURE.

**Fig. S1** Differentiation among the 12 locations selected for the estimate of dispersal.

**Appendix S1** Details of microsatellite marker development.