Tracing geographical patterns of population differentiation in a widespread mangrove gastropod: genetic and geometric morphometrics surveys along the eastern African coast

CAROLINA MADEIRA1,2*, MARIA JUDITE ALVES2,3, NATACHA MESQUITA2,3, SARA EMA SILVA2 and JOSÉ PAULA1

1Centro de Oceanografia, Laboratório Marítimo da Guia, Faculdade de Ciências da Universidade de Lisboa, Avenida Nossa Senhora do Cabo 939, 2750-374 Cascais, Portugal
2Museu Nacional de História Natural e da Ciência, Universidade de Lisboa, Rua da Escola Politécnica 56/58, 1250-102 Lisboa, Portugal
3Centro de Biologia Ambiental, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal

Received 1 April 2012; revised 10 May 2012; accepted for publication 11 May 2012

In the present study, we assessed the inter- and intrapopulation genetic and morphological variation of Cerithidea decollata along the eastern coast of Africa. The population structure of C. decollata along the latitudinal gradient was examined by sequencing 420 bp of the mitochondrial cytochrome c oxidase I (COI) gene in 172 snails from 29 sites, in a combined analysis with geometric morphometrics in 1799 snails from 32 sites. Analysis of molecular variance and spatial analysis of molecular variance showed a moderate spatial population differentiation from Kenya to the Republic of South Africa, suggesting genetic divergence between the northern, central, and southern regions. This structure appears to be the result of life-history traits combined with oceanographic features. Haplotype network and mismatch analysis suggest a recent population expansion during the Holsteinian inter-glacial period in the northern region and several colonization events in the central and southern regions. The morphometric approach suggests that morphological variation in shell shape is somewhat independent of the genetic divergence, revealing an overlap of shape across the latitudinal gradient but significant differences among-population at a local level. This may indicate that similar ecological pressures are acting along the coast, leading to the occurrence of similar morphological characters. © 2012 The Linnean Society of London, Biological Journal of the Linnean Society, 2012, 107, 647–663.


INTRODUCTION

In the marine environment, the population genetic structure of a species is mainly influenced by species-specific ecological requirements and life-history traits (Wilke & Davis, 2000; Riginos & Nachman, 2001). Genetic differentiation in marine organisms (invertebrates in particular) is therefore highly influenced by their dispersal capacity (Levin, 2006) and, consequently, by their mode of reproduction (Palumbi, 1995). Most marine invertebrates have limited mobility when adults (gastropods: Thorpe, Solé-Cava & Watts, 2000). In this way, the presence or absence of planktonic larval stages has shown to be an important factor in the determination of populations’ spatial structure (Avise, 2004).

A number of population genetics comparative studies support the hypothesis that marine invertebrate species with direct development have less
potential for gene flow, showing higher levels of spatial structure compared to species with planktonic larvae (gastropods: Kyle & Boulding, 2000; Collin, 2001; bryozoa: Watts & Thorpe, 2006; holothurians: Arndt & Smith, 1998; sea stars: Hunt, 1993; solitary corals: Hellberg, 1996). In species with little or no dispersal during their development (i.e. species with nondispersive larval stages), substantial spatial structure is expected even at geographical microscales (Sokolova & Boulding, 2004). By contrast, larvae dispersal through oceanic currents connects populations that are geographically separated, resulting in a partial homogeneity of their genetic diversity (Bohonak, 1999), where most of the genetic variance is a result of within population variation (Martel et al., 2004). In addition, for planktonic larvae of dispersing organisms, coastal oceanographic dynamics appears to play an important role in a population’s connectivity and differentiation (Silva et al., 2009).

Several studies suggest that habitat size and continuity are factors possibly shaping the evolutionary trajectories of marine organisms (Colborn et al., 2001; Silva, Mesquita & Paula, 2010a, b). When a species has a continuous distribution and there are no restrictions to dispersal along its natural geographical range, we might consider either one of the following: (1) panmixia (Cook, Bunn & Hughes, 2002) or (2) isolation-by-distance (IBD) (Martel et al., 2004), where the exchange of genes and individuals preferentially occurs locally (Cook et al., 2002), decreasing with increasing geographical distance. According to Selkoe & Toonen (2011), in a study based on new datasets obtained from recent marine genetic research, there is moderate fit between the IBD slope and pelagic larval duration proxies of dispersal, typically reflecting scales of dispersal when the sampling design is robust. When species have a patchy discontinuous distribution, a high potential for dispersal will lead to regular exchange of individuals between local subpopulations, connected in a metapopulation complex (Armonies, 2001; Hixon, Pacala & Sandin, 2002), whereas a low potential for dispersal will lead to a spatial network of genetically differentiated populations (Dahdouh-Guebas et al., 2002).

It has been shown that the level of genetic structuring might be reflected in phenotypic differences (Palumbi, 1995; Parsons, 1998). Several studies also suggest that the occurrence of local morphological variation can be explained by ecophenotypic plasticity (Miner et al., 2005; Silva & Paula, 2008), although the relative contribution of genetic and environmental factors in phenotypic expression has not been solved (Smith & Ruiz, 2004). The latter include physical and chemical water properties, desiccation, food availability and quality, predatory pressures (Trowbridge, 1994), and parasitism (Levri et al., 2007), all acting on ecophenotypic plasticity.

Gastropods have shown to be particularly adequate in population differentiation studies for two reasons: (1) they inhabit heterogeneous environments and exhibit conspicuous variation in morphology, life-history, and behaviour (Rolan et al., 2004) and (2) both types of development, direct and indirect (variable dispersion capacity), occur in gastropods, making them good models for comparison between species with different life strategies. Cerithidea decollata (Linnaeus, 1758) snails inhabit shallow tropical and warm temperate seas along the Indian African coast (Reid et al., 2008). Parsimonious reconstruction of ancestral habitats suggests that the living snails in Potamidae family (in which the genus Cerithidea is included) result from adaptive radiation closely associated with mangroves and that the specialized tree-climbing group Cerithidea is derived from mud-dwelling ancestors (Reid et al., 2008). However very little is known about the biology of most of these tropical gastropods and even their systematic classification is still not settled, despite recent studies clarifying phylogenetic relationships (Kojima et al., 2006; Miura, Torchin & Bermingham, 2010). According to D. G. Reid (pers. comm.), the life cycle of C. decollata should be similar to that of Cerithidea rhizophorarum, for which Kojima et al. (2006) report a planktonic life of 12–20 days. In this way, a high dispersal ability can be expected for this species.

The eastern coast of Africa offers a set of unique characteristics, especially for studying the relative importance of oceanographic processes in the genetic structure of marine organisms (Silva et al., 2010b). There are three main current systems along this coast: (1) the Agulhas’ warm current; (2) the Mozambique Channel eddy system; and (3) the South Equatorial current (Lutjeharms, 2006). The patterns of these currents influence the temperature regimes, coastal topography, composition, and distribution of species (Neethling et al., 2008). In addition, the anticyclonic circulation facilitates an almost random dispersal of planktonic larvae (Silva et al., 2010b), contributing to the homogenization of populations along the coast. The mesoscale phenomena occurring in the continental platform, such as eddies and counter currents (Lutjeharms & da Silva, 1988), may also contribute to larval dispersal or retention (Paula, Dray & Queiroga, 2001). Other events, such as upwelling, tidal regime, and estuarine flow, also influence dispersal at a smaller spatial scale (Abelson & Denny, 1997; Quinteiro, Rodriguez-Castro & Reyn-Méndez, 2007). Most of these oceanographic factors can act in a double way: dispersing planktonic larvae and acting as corridors to gene flow or alternatively, acting as physical barriers (Quinteiro et al., 2007;
The combination of data from genetic and morphological surveys obtained when studying population dynamics allows an integrated approach of different levels of an organism's complexity. This is important when interpreting variability patterns and recognizing the factors involved in population differentiation (Silva et al., 2010a). In the present study, we aimed to assess the level of populations' morphological and genetic diversity and spatial structure of Cerithidea decollata (Mollusca, Gastropoda) along the eastern coast of Africa. The specific questions addressed were: (1) do populations of C. decollata show evidence of genetic and morphological subdivision within the geographical area under study? (2) do oceanographic factors contribute to the patterns observed? (3) how do morphological characteristics vary along the latitudinal gradient? and (4) do genetic and morphological patterns of variation vary similarly?

**MATERIAL AND METHODS**

**SAMPLING DESIGN AND PROCESSING**

Samples of C. decollata were collected from mangrove forests from Kenya to South Africa, ranging a geographical north–south gradient of approximately 5000 km.

A spatially-nested sampling design was adopted (Fig. 1), consisting of four regions, 12 areas, and 33 sites, following a hierarchical geographical scale: macrogeographical scale for the regions (> 500 km); mesogeographical scale for the areas (> 50 km < 100 km); and microgeographical scale for the sites (> 1 km < 50 km). The design consisted of four major regions, each one comprising different sampling areas (adding up a total of 12): Kenya/Tanzania north (areas A and B); Tanzania south/Mozambique north (area C to F); Mozambique centre (area G); Mozambique south/Republic of South Africa (area H to L). Except for areas G, J, and L, which were represented by single point locations, the other sampling areas included several independent sampling sites (replicates; adding up a total of 33). The geographical distances between sampling sites were measured by the shortest sea distance (km), using the software GOOGLE EARTH, version 6.1 (Google Inc).

Upon collection by hand at each location, 60 captured specimens were immediately fixed in ethanol 70% for morphometric analysis, whereas another 15 specimens were preserved in absolute ethanol (shells were broken to allow adequate tissue penetration) for genetic analyses.

**GENETIC ANALYSIS**

A total of 172 individuals from 29 sites were used for the genetic analysis. Total DNA was isolated from a portion of mantle tissue, using EZNA Mollusc DNA Isolation Kit (Omega Biotek) in accordance with the manufacturer's instructions.

A 1000–1300-bp fragment of the mitochondrial (mt)DNA cytochrome c oxidase I (COI) gene was amplified by polymerase chain reaction (PCR) using, as primer pairs, COI-bf (5'-GGGGGCTCTTGTAGT AGCTTTTCC-3', Miura et al., 2006b) and COI-6 (5'-GGRTARTCNSWRTANCGNCGNGGYAT-3', Shimayama et al., 1990) or LCO1490 (5'-GGTCAACAA ATCATAAAGATATTGG-3', Folmer et al., 1994) and COI-6. The 25-μl PCR reaction mixture included: 0.1–0.7 μg ml⁻¹ total DNA, 1 × Taq buffer (Fermentas), 2.5 μM MgCl₂, 0.4 μg ml⁻¹ bovine serum albumin, 0.1 μM dNTPs, 1 μM of forward and reverse primers, and 0.02 μl⁻¹ Taq DNA polymerase (Fermentas). The reaction mixtures were subjected to the temperature conditions: denaturation at 94 °C for 60 s; five cycles of denaturation at 94 °C for 30 s, annealing at 45 °C for 60 s, and extension at 72 °C for 60 s; followed by 30 cycles of denaturation at 94 °C for 60 s, annealing at 50 °C for 90 s, extension at 72 °C for 60 s; with a final extension step at 72 °C for 5 min.

PCR products of the samples were then purified by Exo-SAP clean-up protocol (Werle et al., 1994) and sent to Stab-Vida (http://www.stabvida.com/) for sequencing in the forward or reverse directions using a 3700 ABI DNA Sequencer (Applied Biosystems). Sequences analyzed in this study were deposited in GenBank (accession numbers JX026281–JX026452).

Chromatograms were analyzed and sequences were verified, edited and aligned using SEQUENCER, version 4.8 (Gene Codes Corporation). A segment of 420 bp of the partial gene for COI was recovered for all samples. A diversity analysis of the genetic data was performed using DNASP, version 5.0 (Librado & Rozas, 2009): the number of polymorphic sites, segregating sites, mutations, the number of mtDNA haplotypes, and the mean number of nucleotide differences were calculated. Haplotype diversity (h) and nucleotide diversity (π) were calculated for each location, as well as for the entire pooled populations using ARLEQUIN, version 3.5 (Excoffier & Lischer, 2010). The amounts of sequence divergence (p-distances) were calculated using PAUP* 4.0 (Swofford, 2002).

Phylogeographical relationships among haplotypes were explored by haplotype network analysis using the median-joining method, as implemented in NETWORK, version 4.6.0.0 (Shareware Phylogenetic Network Software Website).

© 2012 The Linnean Society of London, Biological Journal of the Linnean Society, 2012, 107, 647–663
Genetic differentiation among samples was assessed through pairwise $F_{ST}$ as implemented in ARLEQUIN, version 3.5 (Excoffier & Lischer, 2010). Significant deviations from the null hypothesis of no differentiation were assessed with 10 000 permutations of individuals among populations after sequential Bonferroni correction.

To assess correlations between genetic and geographical distances and to test whether genetic differentiation among samples could be explained by IBD, values of $F_{ST}/(1 - F_{ST})$ were plotted against geographical distances (ln transformed) for each pair of areas. Furthermore, the significance of the correlation achieved was tested by Mantel $Z$-test with 10 000 iterations, using software IBD, version 1.52 (Bohonak, 2002).

Population genetic structure was explored through a spatial analysis of molecular variance (SAMOVA) approach, using the software SAMOVA, version 1.0 (Dupanloup, Schneider & Excoffier, 2002). Partitioning of mtDNA variation and correlation of haplotypes at different levels of hierarchical subdivision was investigated by analysis of molecular variance (AMOVA), using ARLEQUIN, version 3.5. In a first approach, overall patterns of genetic differentiation among samples could be explained by IBD, values of $F_{ST}/(1 - F_{ST})$ were plotted against geographical distances (ln transformed) for each pair of areas. Furthermore, the significance of the correlation achieved was tested by Mantel $Z$-test with 10 000 iterations, using software IBD, version 1.52 (Bohonak, 2002).

Population genetic structure was explored through a spatial analysis of molecular variance (SAMOVA) approach, using the software SAMOVA, version 1.0 (Dupanloup, Schneider & Excoffier, 2002). Partitioning of mtDNA variation and correlation of haplotypes at different levels of hierarchical subdivision was investigated by analysis of molecular variance (AMOVA), using ARLEQUIN, version 3.5. In a first approach, overall patterns of genetic differentiation among samples could be explained by IBD, values of $F_{ST}/(1 - F_{ST})$ were plotted against geographical distances (ln transformed) for each pair of areas. Furthermore, the significance of the correlation achieved was tested by Mantel $Z$-test with 10 000 iterations, using software IBD, version 1.52 (Bohonak, 2002).

Mismatch distributions were analyzed with ARLEQUIN, version 3.5 to explore the demographic history of populations. Goodness of fit between the observed and expected distribution was tested with 1000 permutation replicates under a sudden expansion model. To further examine the population history, neutrality tests of Tajima’s $D$ (Tajima, 1989) and Fu’s $F$ statistics (Fu, 1997) were carried out using the same software. We calculated the time since expansion ($t$) by converting $\tau$ in the equation $\tau = 2ut$, to time in years where $u$ is the mutation rate/nucleotide year$^{-1} \times$ sequence length. We used a mutation rate of 2.81%/Myr estimated for Cerithidea snails using K2P distances in Miura et al., (2010).

**Morphometric analysis**

The morphology of the shell was characterized in two dimensions (using geometric morphometric methods) for 1799 specimens from 32 sites. These methods yield detailed information about variation in the shape of objects at the same time as retaining a visual representation of them throughout the analysis (Mitteroecker & Gunz, 2009). Landmarks, which are coordinates of points, were recorded from high-resolution digital images taken with a Nikon D70 digital camera with a 55-mm micro lens and using consistent capture conditions for all specimens. The 15 homologous points chosen (Fig. 2) were digitized on the shells using TPSDIG, version 2.16 (Rohlf, 2010a). The criteria used to choose the landmarks were their relative ease in identification across samples and the ability of the suite of landmarks to capture the general shape of the shell.

Gastropods are known to have allometric growth and so it was necessary to evaluate the effect of size in shape variation, using the TPSREGR, version 1.37 (Rohlf, 2009). This was achieved by regressing each shape variable (relative warps; RW) against a...
measure of body size and estimating the residual shape variation, which was then used for further morphometric analyses. Landmarks of each specimen were optimally aligned using generalized Procrustes analysis (Rohlf & Slice, 1990), in which the configurations of the specimens are superimposed via translation, scaling, and rotation, using the minimal bending energy method (Mitteroecker & Gunz, 2009). Centroid size, which is used as scaling factor during the superimposition process, was subsequently used as a measure of size for each specimen. The term ‘shape’ used in the present study is thus defined as a geometric representation of an object after removing all nonshape variation as a result of measurement-associated errors (Claverie, Chan & Patek, 2011).

From the aligned specimens, shape variables were generated by performing a RW analysis (Bookstein, 1991), which is the analogue of a principal components analysis (Zelditch et al., 2004). This analysis reduces the dimensionality of multivariate data by transforming a set of many correlated variables into a small number of significant uncorrelated variables (Claverie et al., 2011) called RWs. These new sets of shape variables are then used for statistical comparisons of shape variation within and among groups. The present study also used the thin-plate spline approach, using the TPSRELW, version 1.49 (Rohlf, 2010b), which allowed the visualization of shape change as deformation grids.

STATISTICA, version 9.0 (StatSoft Inc) was used to investigate shell shape variation significance. A nested multivariate analysis of variance (MANOVA) was performed on the RW scores. In this way, we determined the degree and significance of variation between regions, study areas within regions, and populations within areas. To test which groups differed within each hierarchical level, we performed post-hoc Tukey's honestly significant difference tests for unequal N, which perform pairwise comparisons. The best structure obtained in the genetic analysis was also used to test whether differences at morphological level were also significant, using the same approach as before. A stepwise discriminant analysis was also performed on the shape variables. Classification methods were then used to evaluate the power of the functions in discriminating groups.

RESULTS
GENETIC ANALYSIS
The sequencing of 420 bp from 172 specimens revealed a total of 60 distinct haplotypes defined by 49 polymorphic sites, of which 26 were parsimony informative. Forty-nine (81.67%) haplotypes were unique, accounting for 28.49% of the overall specimens. The most common haplotype (H1) was present in 60 of the 172 individuals (34.88%). H1 was found in every sampling locality from Kenya to south Mozambique (except for Shirazi and Saco), although it did not appear in the populations of the Republic of South Africa. The second (H10) and third (H6) most common haplotypes accounted for 16.27% and 6.40% of the total individuals, respectively. H10 was found across the entire geographical range of this study. Haplotype 6 was private to Mozambique. Almost all haplotypes (seven out of ten) found in Mozambique central region were private to this region. Within a given location, the highest haplotype richness was found in Rio Savane (Mozambique centre), with ten haplotypes, and the lowest haplotype richness was found in Pemba (Mozambique north) samples with only one haplotype.

Considering the total study area, the eastern coast of Africa, the overall values of haplotype diversity and nucleotide diversity (Table 1) for the entire mtDNA data set were 0.8478 ± 0.023 and 0.00912 ± 0.00052, respectively. The levels of within population genetic diversity (Table 1) were quite variable among localities. The regions with the highest overall level of h were Mozambique centre (h = 0.955 ± 0.057), Mozambique south/Republic of South Africa (h = 0.916 ± 0.026), and Kenya/Tanzania north (h = 0.909 ± 0.039), and the lowest level was found in Tanzania south/Mozambique north (h = 0.643 ± 0.056). Considering nucleotide diversity, the highest overall value of π was found in Mozambique centre (π = 0.01205 ± 0.00138) and Mozambique south/Republic of South Africa (π = 0.01105 ± 0.00091), and the lowest value was found in Tanzania south/north Mozambique (π = 0.00504 ± 0.00069). The calculations of sequence divergence among haplotypes resulted in a mean genetic distance of 1.32%, and a minimum and maximum of 0.24% and 3.81%, respectively.

The mtDNA haplotype network for C. decollata (Fig. 3) was not fully resolved, presenting various ambiguities (loops). The three most common haplotypes (H1, H10, and H6), presented central positions within the network, connected by a few mutational steps. H1 was in the centre of a clear star-like topology. Several exclusive alleles were interspersed in the network, connected by a few mutational steps, without a clear assignment to a cluster.

Values of pairwise Φst ranged from −0.0296 to 0.4367 and allowed us to identify significantly different areas. Pairwise Φst values were higher and most of them were significant when comparing northern areas with central and southern areas of the latitudinal gradient. The full COI data set revealed a positive and significant correlation between genetic
divergence and geographical distance, as indicated by the results of the Mantel test \((Z = 68.543, r = 0.289, P < 0.05; \text{Fig. 4})\), fitting the IBD model.

The SAMOVA algorithm was used to investigate the hypothesis of finding population groups using both genetic information and geographical location of the localities sampled. This analysis was performed several times by varying the user-defined number of clusters, from \(K = 2\) to \(K = 6\). The cluster formed by the northern localities did not collapse with the increase in the number of groups, while the two other clusters were broken progressively. In most cases, samples were grouped not breaking the defined sampling areas (A to L; \text{Fig. 1}).

When considering the hierarchical partitioning of mtDNA variation, for AMOVA I, in which individuals from the same locality were treated as individual populations, the large majority of the genetic variance was a result of differences within populations (78.76%) and 21.24% to differences among populations. A significant \(F_{ST}\) value was achieved \((F_{ST} = 0.21245, P < 0.001)\), indicating significant overall genetic differentiation among populations (Table 2). In AMOVA II (Table 2), the groups were defined according to our a priori expectations: four regions as described in the nested sampling design: (1) Kenya/Tanzania north; (2) Tanzania south/Mozambique north; (3) Mozambique centre; and (4) Mozambique south/Republic of South Africa. The results show significant differentiation among groups, representing 21.24% of the total variance \((F_{CT} = 0.21244, P < 0.001)\), although the greatest source of variation revealed to be among populations (73.21%, \(F_{ST} = 0.26787, P < 0.001)\). When performing

Table 1. Summary statistics: number of specimens, haplotypes and haplotype and nucleotide diversity for \(C. decollata\) populations

<table>
<thead>
<tr>
<th>Locality</th>
<th>Number of specimens</th>
<th>Number of haplotypes</th>
<th>Number of private haplotypes</th>
<th>(h) (mean ± SD)</th>
<th>(\pi) (meaqn ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mi</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0.67 ± 0.31</td>
<td>0.003 ± 0.003</td>
</tr>
<tr>
<td>G</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0.83 ± 0.22</td>
<td>0.008 ± 0.006</td>
</tr>
<tr>
<td>Sh</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.00 ± 0.00</td>
<td>0.000 ± 0.000</td>
</tr>
<tr>
<td>K</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>1.00 ± 0.13</td>
<td>0.008 ± 0.005</td>
</tr>
<tr>
<td>MK</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0.90 ± 0.16</td>
<td>0.008 ± 0.005</td>
</tr>
<tr>
<td>RD</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>0.93 ± 0.12</td>
<td>0.011 ± 0.007</td>
</tr>
<tr>
<td>Mn</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>0.80 ± 0.17</td>
<td>0.005 ± 0.004</td>
</tr>
<tr>
<td>RE</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>0.83 ± 0.22</td>
<td>0.002 ± 0.002</td>
</tr>
<tr>
<td>OL</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0.60 ± 0.22</td>
<td>0.005 ± 0.004</td>
</tr>
<tr>
<td>MP</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>0.29 ± 0.20</td>
<td>0.005 ± 0.003</td>
</tr>
<tr>
<td>U</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>0.33 ± 0.22</td>
<td>0.001 ± 0.001</td>
</tr>
<tr>
<td>Lu</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>0.42 ± 0.19</td>
<td>0.003 ± 0.002</td>
</tr>
<tr>
<td>IK</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>0.87 ± 0.13</td>
<td>0.006 ± 0.004</td>
</tr>
<tr>
<td>T</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>0.71 ± 0.18</td>
<td>0.003 ± 0.003</td>
</tr>
<tr>
<td>O</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>1.00 ± 0.18</td>
<td>0.012 ± 0.009</td>
</tr>
<tr>
<td>P</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>0.00 ± 0.00</td>
<td>0.000 ± 0.000</td>
</tr>
<tr>
<td>Me</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>0.80 ± 0.17</td>
<td>0.005 ± 0.004</td>
</tr>
<tr>
<td>NV</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>0.81 ± 0.13</td>
<td>0.010 ± 0.006</td>
</tr>
<tr>
<td>N</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>0.80 ± 0.17</td>
<td>0.008 ± 0.006</td>
</tr>
<tr>
<td>CG</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0.73 ± 0.16</td>
<td>0.006 ± 0.004</td>
</tr>
<tr>
<td>RS</td>
<td>12</td>
<td>10</td>
<td>7</td>
<td>0.95 ± 0.06</td>
<td>0.012 ± 0.007</td>
</tr>
<tr>
<td>CS</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>0.86 ± 0.09</td>
<td>0.011 ± 0.007</td>
</tr>
<tr>
<td>MF</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>0.90 ± 0.10</td>
<td>0.010 ± 0.006</td>
</tr>
<tr>
<td>S</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>0.90 ± 0.16</td>
<td>0.012 ± 0.008</td>
</tr>
<tr>
<td>PR</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1.00 ± 0.17</td>
<td>0.012 ± 0.008</td>
</tr>
<tr>
<td>KB</td>
<td>11</td>
<td>6</td>
<td>4</td>
<td>0.83 ± 0.30</td>
<td>0.009 ± 0.008</td>
</tr>
<tr>
<td>SL</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0.80 ± 0.16</td>
<td>0.006 ± 0.004</td>
</tr>
<tr>
<td>HB</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0.83 ± 0.22</td>
<td>0.009 ± 0.007</td>
</tr>
<tr>
<td>Mlz</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>1.00 ± 0.12</td>
<td>0.014 ± 0.010</td>
</tr>
<tr>
<td>Total</td>
<td>172</td>
<td>60</td>
<td>53</td>
<td>0.85 ± 0.02</td>
<td>0.009 ± 0.001</td>
</tr>
</tbody>
</table>

For the complete name of the sampling sites see Fig. 1. \(h\), haplotype diversity; \(\pi\), nucleotide diversity.
AMOVAs to the genetic structures defined by SAMOVA for the various user-defined number of clusters (from $K = 2$ to $K = 6$), the results (not shown) actually indicate a tendency of $\Phi_{CT}$ to increase and $\Phi_{SC}$ to decrease with increasing $K$, reaching a plateau at $K = 3$ (AMOVA III; Table 2). The genetic structure tested in AMOVA III consisted of three groups: (1) from Mikindani to Cabaceira Grande (area A to F) – northern; (2) from Rio Savane to Ponta Rasa (area G to I) – central; and (3) from Kosi bay to Mlalazi (area J to L) – southern. For this population genetic structure, the partition of total variance reached its highest value for genetic differentiation among groups (30.96%) and its lowest value for genetic differentiation among populations within groups (near 0.00%), with the genetic differentiation among populations within groups being nonsignificant ($P > 0.05$).

Gene flow estimates, $M (= N_m$ for haploid populations), were performed pairwise between groups (for groups in AMOVA III): (1) $M = 1.24$ migrants/generation among northern-central groups; (2) $M = 1.87$ migrants/generation among central–
southern groups; and (3) $M = 0.78$ migrants/generation among northern–southern groups.

For the mismatch analysis and neutrality tests (Table 3), populations were grouped according to the three-group structure generated by SAMOVA that maximized the variance among groups. Both $P_{\text{SSD}}$ and $P_{\text{Ragged}}$ showed that the observed mismatch distribution patterns did not significantly differ from the expected distribution under a sudden expansion model (except the $P_{\text{SSD}}$ value for the central group). Tajima’s $D$ and Fu’s $F_s$ tests for the northern group (areas A to F) were negative with significant $P$-values, and therefore the hypothesis of sudden population expansion could not be rejected. The calculated time of expansion for northern group was approximately 227 155 years (between 201 746 and 252 640 years; $\tau$: 95% confidence interval $= 4.756–5.956$). For central (areas G to I) and southern (areas J to L) groups, the results of the several tests are not all consistent with the hypothesis of sudden demographic expansion.

**Table 3.** Mitochondrial DNA haplotype mismatch distribution analysis for *Cerithidea decollata*; Tajima’s $D$ and Fu’s $F_s$ tests and $P$ values are also presented

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Goodness-of-fit tests</th>
<th>Tajima’s $D$ test</th>
<th>Fu’s $F_s$ test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_0$, $\theta_1$, $\tau$</td>
<td>$SSD$</td>
<td>$P_{\text{SSD}}$</td>
<td>$P_{\text{Ragged}}$</td>
</tr>
<tr>
<td>North</td>
<td>0.002</td>
<td>2.799</td>
<td>5.355</td>
</tr>
<tr>
<td>Centre</td>
<td>0.000</td>
<td>14.707</td>
<td>6.398</td>
</tr>
<tr>
<td>South</td>
<td>0.000</td>
<td>7.998</td>
<td>5.625</td>
</tr>
</tbody>
</table>

$\theta_0$ and $\theta_1$, pre- and post-expansion population size; $\tau$, time in number of generations elapsed since the episode of sudden expansion; $SSD$, sum of squared deviations; $P_{\text{SSD}}$ and $P_{\text{Ragged}}$, probability that the expected mismatch distributions have significantly larger frequencies than the observed mismatch distributions. *Significant values at $P < 0.05$.

revealed differences at the shell aperture level, mainly in the distances between landmarks 7, 8, 9, and 10, characterizing the progressive narrowing/opening of the aperture. In Figure 5, the thin plane spline grids allow the visualization and interpretation of the deviation values (both positive and negative) in geometric terms, for RW1 and RW2. The results graphically support the visual differences in shell morphology that can be observed.

Results from the nested MANOVA following the sampling design showed that there were no significant differences among regions and among areas within the regions, revealing an overlap of shape across the latitudinal gradient. When testing for differences among populations (Fig. 6), however, the results showed significant differences among them (Wilks' $\lambda = 0.055; F = 10.676; P < 0.0001$), which indicate that differences in shell shape are mainly local. The post-hoc Tukey's honestly significant difference tests performed on the populations to test which ones were different confirmed this possibility showing that, within all areas, almost all of the populations were significantly different from each other. Results from the MANOVA testing morphological differences for the best genetic structure obtained (three groups from the SAMOVA analysis: northern, central, southern)
were identical to the previous ones, with no significant differences among groups, corroborating the possibility of an overlap of shape across the gradient.

The degree of separation between regions, areas, and populations was quantified using a stepwise discriminant analysis applied to the shape variables. Significant differences were observed for all cases for the discriminant function: regions \((R = 0.571; \ P < 0.0001)\), study areas \((R = 0.625; \ P < 0.0001)\), and populations \((R = 0.795; \ P < 0.0001)\), and the most efficient variable for discrimination was RW1. A cross validation statistical analysis also verified the efficiency of the discrimination.

Correct classification of specimens according to shape similarities/differences was 44.7% for Kenya/Tanzania north; 77.4% for Tanzania south/Mozambique North; 90.2% for Mozambique Centre; and, finally, 67.4% for Mozambique south/Republic of South Africa. The overall rate of specimens correctly classified into regions was 58.3%. When analyzing each area and each locality separately, 47.4% and 58.4% of the overall specimens, respectively, were correctly classified into groups. Considering localities, correct classification ranged from 41% (Pemba, Cabaceira Grande) to 84% (Rio Savane). Most misclassified individuals were assigned to geographically adjacent localities or areas, although there were cases of misclassification that showed no obvious pattern of assignment.

**DISCUSSION**

The assessment of population differentiation along the eastern African coast for the mangrove gastropod *C. decollata* through the combined analysis of mtDNA and shell shape variation revealed different patterns of divergence for genetic and morphological markers. Considering the genetic differentiation pattern, the major source of variation was found to be at within-localities level. However, differences at a macrogeographical scale were found to be significant and correlated with geographical distance, suggesting the possibility of the existence of discrete groups (northern–central–southern) along the coast with diminished gene flow among them, as demonstrated by the low levels of gene flow estimated \((1 < N_m < 2\) for estimates between the central and the two other groups, and \(N_m < 1\) for estimate between north–south groups). Concerning morphometric geometrics analyses, the major source of morphological variation occurs within areas at among-localities level, where significant differences were obtained, although not at a macro spatial scale, showing an overlap of shape across the major regions of the latitudinal gradient.

In summary: (1) Genetic approach: significant differentiation among regions, with a three-group structure.
(macrospatial scale), and significant genetic differentiation within localities (microspatial scale) along the coast but nonsignificant differentiation among localities within groups (mesospatial scale), and (2) Morphological approach: nonsignificant differentiation among regions (macrospatial scale), nonsignificant differentiation within localities (microspatial scale) but significant differentiation among localities within areas (mesospatial scale).

*Cerithidea decollata* proved to be somewhat genetically variable, as shown by the *p*-distances. Considering nucleotide diversity values, similar ranges of values (0.0000–0.0143) were found compared to other *Cerithidea* species (Kojima et al., 2006; *Cerithidea cingulata*: 0.0004–0.0119, *Cerithidea djajariensis*: 0.0000–0.0088, *Cerithidea largillieri*: 0.0059–0.0068, *C. rhizophorarum*: 0.0007–0.0054). Regarding haplotype diversity, it was generally high within most sampling sites. Furthermore, the presence of many rare haplotypes suggests that the female effective population size is very large (Lewontin, 1974). These results are concordant with the expectation that gastropod species with pelagic larvae have high levels of genetic diversity as a result of high effective population sizes (*N_e*) and so the effects of genetic drift might be greatly diminished. In this way, rare mitochondrial haplotypes will be more likely to persist in the population (Slatkin, 1985). Our results further indicate that there is no apparent haplotype saturation level with increasing sample size, suggesting that many unidentified haplotypes may still exist in the populations sampled, as also observed by Cassone & Boulding (2006) and Silva et al. (2010a) for crabs in the Pacific and Indian oceans, respectively. Similar to our results, high levels of genetic variation within populations have been reported for several gastropod species with pelagic larvae such as *Littorina scutulata* along the west coast of North America (Kyle & Boulding, 2000), *Nerita atramentosa* in south-eastern Australia (Waters et al., 2005), and *Littoraria scabra* in eastern Africa (S. E. Silva, unpubl. data). The low levels of haplotype and nucleotide diversities observed in Tanzania south/Mozambique north may be the result of a population bottleneck followed by demographic expansion, as suggested by the mismatch analysis (discussed below).

Our analysis of DNA sequences suggests that *C. decollata* displays a moderate population structure along the 5000 km of the eastern African coast. This was reflected by: (1) a reasonable and significant overall value of $\Phi_{ST}$ (0.21245); (2) high pairwise $\Phi_{CT}$ values among areas from the northern versus central versus southern regions of the latitudinal gradient; (3) the existence of private haplotypes in most populations; and (4) significant $\Phi_{CT}$ values in the AMOVAs performed. These results are consistent with a number of other intertidal and nearshore species with planktonic larvae that, despite the high potential for dispersal and gene flow, display a significant population structure (*Batillaria zonalis*, Kojima et al., 2005; *Carcinus maenas*, Roman & Palumbi, 2004; Pascoal et al., 2009; *Tectarius striatus*, Van den Broeck et al., 2008; *Conus ebraeus*, Duda & Lessios, 2009).

A similar macrogeographical zonation has been obtained for the larvae dispersal crab *Perisesarma guttatum* in the same geographical area (Silva et al., 2010a). The different genetic differentiation patterns attained at different geographical scales in the present study appear to indicate that the presence and duration of the planktonic larval phase are key determinants of the magnitude of spatial population structure in marine invertebrate species, as observed by Lee & Boulding (2009).

We should note that the moderate genetic differentiation obtained at a macrogeographical scale (three-group structure obtained by SAMOVA) might have been underestimated. This is because fixation indices derived from the AMOVA are not completely accurate when used for the analysis of haplotypic data and, hence, should be regarded as indicators rather than absolute values (Schneider, Roessli & Excoffier, 2000). This is important because samples from large populations, as in the case of *C. decollata*, along the sampled geographical gradient tend to underestimate population differentiation, causing the results to be conservative.

The attained macrogeographical structure pattern is consistent with increasing evidence that a variety of physical oceanographic factors, including temperature gradients, wind patterns, ocean mesoscale currents and eddies can restrict larval dispersal (McCartney, Keller & Lessios, 2000). As a result of its planktonic larval stage of 2–3 weeks, we predicted a high dispersal ability for *C. decollata*. Therefore, the observed differentiation may be a consequence of physical mechanisms, where a correlation between coastal hydrology and population connectivity exists. In particular, the complex hydrographical dynamics of Maputo Bay and the adjacent neritic waters, which include eddy systems and inshore countercurrents derived from Mozambique Channel, are the most likely physical mechanisms that could act as barriers reducing dispersal. These phenomena may contribute to the retention of larvae and restrict offshore dispersal as suggested by Paula et al. (2001), which would lead to the differentiation of the central group (because larvae would be kept in this area), as well as prevent gene flow between northern and southern regions, thus explaining the observed differences. Also, Larson & Julian (1999) maintain that oceanographic currents and the entrainment of larvae lead to genetic patchiness, which may counteract the effect.
of fecundity and dispersal time. Considering the non-
significant values of $F_{SC}$ for the three-group structure
obtained by SAMOVA, the East African coastal
current and the anti-cyclonic circulation pattern in
north Mozambique coast appear to be the most
important features contributing to the homogenization of
populations along the coast in the northern region
because they allow an almost random dispersal of the
larval planktonic phases (Silva et al., 2010a). Similarly,
the warm southward Agulhas Current that
forms between south Mozambique and the Republic of
South Africa (Lutjeharms, 2006) appears to play an
identical role, and might be responsible for the south-
ward dispersion of larvae from areas J, K and L,
leading to the homogenization of populations within
the southern region.

In IBD analysis, genetic differentiation was found
to be correlated with coastline distance among popu-
lations, suggesting that dispersal occurs across
limited distances and through successive events (i.e.
following a step-by-step process along the coastline),
fitting an IBD model. This is concordant with other
studies in which isolation by distance has been sug-
gested for rocky and shallow intertidal organisms in
south Mozambique and South Africa, irrespective of
their pelagical larval duration (Evans et al., 2004;
Teske et al., 2007). A review of IBD slopes for several
marine invertebrates inferred that planktonically
developing gastropods dispersed over 20–140 km on
average (Kimlan & Gaines, 2003). In addition, it
appears that, within each group defined in the three-
group genetic structure, the larval export strategy
(Paula et al., 2004) might apply, once the degree of
differentiation among populations within groups was
really low and nonsignificant (AMOVA III).

Genetic differences observed at microgeographical
cscale may result from the zonation of plant commu-
nities in intertidal habitats to which $C. decollata$ is
associated. Mangrove zonation quite often occurs in a
mosaic style that varies with the complex of physical,
chemical, and biological interactions occurring in a
particular area (Neukermans et al., 2008; Rey &
Rutledge, 2009). In this way, $C. decollata$ should be
distributed in clumps near Avicennia marina tree
patches, which serve as shelter during the high tide
to escape from predators. Such a correlation between
mangrove vegetation and fauna distribution has been
suggested for several intertidal species, in particular
for grapsid crabs (Dahdouh-Guebas et al., 2002). This
type of distribution would lead to a network of spatial
genetic variation explaining the significant overall
subdivision observed in the study area.

The network topology, the mismatch distribution
and Tajima's $D$ and Fu's $F$, tests for $C. decollata$ for
the northern region are indicative of a population
that has recently expanded in size from one or a small
number of founders following a population bottlenecks
(Slatkin & Hudson, 1991), showing an excess of
recent mutations. Thus, the limited degree of gene-
differentiation among populations within this area may also be a reflection of recent expansion.

The estimated time of expansion [approximately
227 155 years (95% confidence interval = 201 746–
252 640 years)] coincides with the beginning of a
warm high sea level period in the Holsteinian inter-
glacial period. Similar patterns appear to be shared
by many other western and central Pacific marine
taxa (Duda & Lessios, 2009). For the central and
southern groups, the results of the tests described
above were not all concordant with the hypothesis of
sudden expansion, which appears to indicate the
existence of equilibrium between mutation and drift
and constancy in population size in these regions. In
line with these findings, Benzie et al. (2002) and
Kochzius & Nuryanto (2008) suggest that the fall in
sea level during the Pleistocene ice ages is likely to
have removed shelf habitat in southern Africa.
Indeed, the mangrove habitat is highly sensitive to
sea-level change, being greatly reduced in global area
during low sea-level stands (Sun et al., 2000). Conse-
sequently, the distributions of marine species must have
been profoundly disrupted along the broad continen-
tal shelves in the central Indo-West-Pacific, when
these were exposed during Pleistocene sea-level fluc-
tuations (Voris, 2000). These habitats would then
have been reinvaded by the surviving populations,
when the sea level rose, at the end of the glacial
periods (Forbes et al., 1999), enabling the recoloniza-
tion and growth of the reduced populations. This
scenario also appears to be applicable for many
marine organisms from other disparate geographical
regions (Gopurendo & Hughes, 2002; Couceiro et al.,
2007; Espinosa et al., 2010).

Considering the geometric morphometrics analysis,
differences in morphological characters were not so
obvious when analyzing the thin-plate deformations,
and the MANOVAs showed morphological similarity
along the gradient at a macrogeographical scale as a
result of the overlap of morphotypes. Nevertheless,
we observed the existence of significant among-
locality within areas divergence of shape. The main
morphometric differences concerned shell length and
aperture size. The greatest deviations to the mean
configuration were found in the central region (area
G) versus the northern (areas A to F) and southern
regions (areas H to L), when displaced in a mor-
phological space. Interestingly, morphotypes of different
regions appear to be alike. Therefore, our data do not
show any particular geographical assignment of
shapes among the regions defined, which was con-
firmed by the low levels of correct classification
obtained by classification methods. This may indicate
that similar ecological pressures are acting along the coast, leading to the similar development of morphological characters. However, when acting in the opposite direction, site-specific ecological pressures may cause changes in the expression of a particular morphological trait, which may explain the significant differences observed among sampling sites. A similar phenomenon was also observed by Silva et al. (2010a) for a crab species, where the same environmental factors that possibly justified the lack of morphological differentiation across several hundreds of kilometers of coast were the same factors contributing to local shape variation, when acting in opposite directions. These ecological pressures can have diverse origins, such as foraging, defence, habitat availability, mating, food acquisition, and even human caused disturbances (Smith, 2004; Brian et al., 2006).

In C. decollata, the combination of a high dispersal potential with the environmental features, namely oceanographic currents, appears to have led to a moderate genetic divergence, suggesting the existence of three groups (northern, central and southern) with low gene flow among them. Yet, in view of the fact that we only looked at one single gene fragment, and given that there was no haplotype saturation in our samples, it is possible that there may be far more differentiation than expected, although this may be difficult to detect because a more intensive sampling effort would be needed. In addition, genetic analyses of marine population structure often find slight geographical differentiation in species with a high dispersal potential but interpreting this genetic signal has been difficult. Many more studies of widespread species, namely those combining more than one genetic marker, will be needed before the factors that contribute to modern-day patterns can be elucidated. This should help solve questions such as the effective population sizes of marine populations, how these change through time, and the movements among marine populations that shape their evolutionary history and sustain their dynamics today. For the geometric morphometrics survey, the patterns of shape variation along the coast were identical, suggesting the possibility of similar environmental pressures acting along the coast. Understanding the cues inducing morphological changes and their immediate adaptive value is being improved (Schlichting & Pigliucci, 1998), although there is still much left to comprehend about how this phenomenon contributes to broader spatial (and temporal) patterns of variation, especially in complex ecosystems such as mangroves. The combined application of different approaches is therefore very valuable and should help resolve important questions about different aspects of the past and present differentiation patterns of populations over large geographical scales.

ACKNOWLEDGEMENTS

This study is part of the research project TRANSMAP (INCO-CT2004-510862) funded by the European Commission, 6th Framework. N. Mesquita acknowledges a post-doc grant by Fundação para a Ciência e Tecnologia (Portugal), reference SFRH/BPD/14557/2003. The authors extend their gratitude to the following individuals who assisted in field work: I. Silva, D. Abreu, A. Alfiado, F. Amade, D. Glassom, C. Hassane, N. Jiddawi, A. Macia, J. Maggs, P. Muianga, E. Nhaca, and G. Penha-Lopes. We thank M. Drago, M. Lourenço, and C. Mateus for all the help given during the laboratory work, as well as R. N. Mendes for map elaboration. Finally, we would like to thank the referees for their helpful comments.

REFERENCES


Sokolova IM, Boulding EG. 2004. A neutral DNA marker suggests that parallel physiological adaptations to open shore and salt marsh habitats have evolved more than once within two different species of gastropods. *Marine Biology* **145**: 133–147.


