Population subdivision in marine environments:
the contributions of biogeography, geographical distance
and discontinuous habitat to genetic differentiation
in a blennioid fish, Axoclinus nigricaudus

C. RIGINOS* and M. W. NACHMAN
Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

Abstract
The relative importance of factors that may promote genetic differentiation in marine organisms is largely unknown. Here, contributions to population structure from a biogeographic boundary, geographical distance and the distribution of suitable habitat were investigated in Axoclinus nigricaudus, a small subtidal rock-reef fish, throughout its range in the Gulf of California. A 408-bp fragment of the mitochondrial control region was sequenced from 105 individuals. Variation was significantly partitioned between 28 of 36 possible combinations of population pairs. Phylogenetic analyses, hierarchical analyses of variance and a modified Mantel test substantiated a major break between two putative biogeographic regions. This genetic discontinuity coincides with an abrupt change in ecological characteristics, including temperature and salinity, but does not coincide with known oceanographic circulation patterns or any known historic barriers. There was an overall relationship of increasing genetic distance with increasing geographical distance between population pairs, in a manner consistent with isolation-by-distance. A significant habitat-by-geographical-distance interaction term indicated that, for a given geographical distance, populations separated by discontinuous habitat (sand) are more distinct genetically than are populations separated by continuous habitat (rock). In addition, populations separated by deep open waters were more genetically distinct than populations separated by continuous habitat (rock). These results indicate that levels of genetic differentiation among populations of A. nigricaudus cannot be explained by a single factor, but are due to the combined influences of biogeography, geographical distance and availability of suitable habitat.

Keywords: gene flow, isolation-by-distance, mitochondrial DNA, phylogeny, phylogeography, population structure

Received 14 September 2000; revision received 17 January 2001; accepted 17 January 2001

Introduction
A continuing challenge in evolutionary biology is to understand the processes by which populations become genetically distinct. In general, genetic drift and local adaptation are counteracted by the unifying effects of gene flow. In many marine organisms, gene flow is high due to a planktonic life stage that can result in movement over large distances (reviewed in Palumbi 1994; Shulman 1998). The combination of high dispersal and few barriers to larval movement presents a challenge for understanding how divergence occurs in marine environments. Although the predominant mechanisms leading to population differentiation are not always clear (Palumbi 1994), several factors may be important either singly or in combination, including limited dispersal ability (Waples 1987; Duffy 1993; Hunt 1993; Doberiy et al. 1995), local adaptation (Koehn et al. 1980; Powers et al. 1991; Schmidt & Rand 1999), oceanographic currents (Shulman & Bermingham 1995; Benzie & Williams 1997; Palumbi et al. 1997; Rocha-Olivares & Vetter 1999; Stepien 1999), habitat discontinuities (Winans 1980; Burton & Feldman 1981; Bell et al. 1982; Stepien & Rosenblatt...
Biogeographic regions are often described based on the overlapping ranges of many species, and boundaries between these regions may derive from historical discontinuities or from present-day environmental differences, such as differences in temperature or salinity. Although the underlying causes of such interspecific boundaries are not always well understood, these boundaries represent natural places to look for genetic discontinuities within species as well. For example, deep intraspecific divisions and the presence of many sister species pairs on either side of Cape Canaveral, Florida (separating Atlantic and Gulf of Mexico regions) point to historic vicariance in this region (reviewed in Avise 1992). Similarly, several Indo-Pacific taxa show increased divergence between Pacific and Indian ocean populations (Lavery et al. 1995, 1996; McMillan & Palumbi 1995; Chenoweth et al. 1998b; Duda & Palumbi 1999). However, not all biogeographic divisions correspond to intraspecific division, for example, intraspecific divergences are negligible across Point Conception, California where Californian and Oregonian regions meet (reviewed by Burton 1998).

Despite the generally high levels of gene flow in marine organisms, isolation-by-distance (Wright 1943) has been reported for a number of fishes and invertebrates (e.g. Johnson & Black 1995, 1998a,b; Pogson et al. 1995; Chenoweth et al. 1998a,b; Gold & Richardson 1998; Benzie 1999; Mamuris et al. 1998a,b; Duda & Palumbi 1999). How- ever, not all biogeographic divisions correspond to intraspecific division, for example, intraspecific divergences are negligible across Point Conception, California where Californian and Oregonian regions meet (reviewed by Burton 1998).

Despite the generally high levels of gene flow in marine organisms, isolation-by-distance (Wright 1943) has been reported for a number of fishes and invertebrates (e.g. Johnson & Black 1995, 1998a,b; Pogson et al. 1995; Chenoweth et al. 1998a,b; Gold & Richardson 1998; Benzie 1999; Mamuris et al. 1998a,b; Duda & Palumbi 1999). There is some evidence that relationships between geographical and genetic distances depend on sampling scale. For example, high levels of genetic divergence are observed in Pacific urchins (Palumbi et al. 1997) and coconut crabs (Lavery et al. 1995, 1996) but only over large geographical scales (> 5000 km). In surgeonfish, in contrast, genetic and geographical distance are correlated within, but not among, archipelagos (Planes et al. 1996). Comparisons between low and high dispersal solitary corals (Hellberg 1996) and mud snails (Wilke & Davis 2000) reveal isolation-by-distance only for the low dispersal species but not for the high dispersal species.

Finally, discontinuity in suitable habitat may reduce gene flow among populations of marine organisms. For example, several studies have found that island populations isolated by open water are genetically divergent from mainland populations (Winans 1980; Bell et al. 1982; Stepien & Rosenblatt 1991; Doherty et al. 1995; Johnson & Black 1995), although divergence is minimal in a coastal urchin between western and eastern Pacific populations spanning 5400 km of open water (Lessios et al. 1998). Along coastlines, populations of tidepool copepods from rocky outcrops separated by sandy beach are more divergent than populations from the same outcrop (Burton & Feldman 1983), and estuarine populations of an atherinid fish (Johnson et al. 1994), a catfish (Ayvazian et al. 1994) and a littorinid snail (Johnson & Black 1998b) are more divergent than shoreline (continuous habitat) populations. Among pelagic organisms, deep oceanic channels are associated with genetic divergence in both cod (Bentzen et al. 1996; Rozzante et al. 1998) and squid (Shaw et al. 1999). Clearly, habitat requirements differ among species but, where discontinuities in habitat have been identified, they often appear to contribute to genetic divergence.

Although individual studies point to the importance of biogeographic boundaries, geographical distance and habitat discontinuities for genetic differentiation, relatively few studies have considered these factors simultaneously (Johnson & Black 1995, 1998a,b; Lavery et al. 1995, 1996; Benzie 1999). Here, we test the hypothesis that these three factors in combination are important for generating population differentiation in the Cortez triplefin, Axoclinus nigricaudus (Allen & Robertson 1991).

Axoclinus nigricaudus is a small (25–40 mm standard length) subtidal fish with demersal (benthic) eggs and planktonic larvae. A. nigricaudus is found only on rocky shores in the Gulf of California, Mexico. On the western coast of the Gulf (Baja) rocky shores are almost continuous (Fig. 1). In contrast, rocky regions on the eastern coastline of the Gulf (Sonora) are few and are surrounded by extensive sandy shores (beaches and estuaries). Thus, populations are separated by either continuous suitable habitat (rocky coast) or by unsuitable habitat (sandy shores or open water). The Gulf of California contains at least two distinct biogeographic regions (northern and central, Fig. 1), based on range distributions, community composition of fishes and differences in salinity, temperature and tidal regimes (Walker 1960; Thomson & Gilligan 1983). There are no reported examples of sister taxa between these two regions. Thus, A. nigricaudus presents an opportunity to simultaneously evaluate the importance of a biogeographic boundary, geographical distance and discontinuities in suitable habitat for population differentiation. We sampled A. nigricaudus throughout its range and found that all three factors in combination contribute to patterns of mitochondrial DNA (mtDNA) differentiation among populations.

Materials and methods

**Populations sampled**

In total 105 Axoclinus nigricaudus individuals were collected from nine geographical locations in the Gulf of California (Fig. 1, Table 1). The study sites were selected to include...
representative sites in the northern and central Gulf biogeographic regions and on both coastlines. In addition, a single congener, *A. carminalis*, was also collected to provide an outgroup for genealogical analyses. Fish were frozen whole in liquid nitrogen. Subsequently, a piece of muscle tissue was removed for DNA extraction. Each individual was deposited as a voucher specimen in the University of Arizona Fish Collection (UAZ).
mtDNA amplification and sequencing

Genomic DNA was prepared from muscle tissue following Sambrook et al. (1989). Universal fish primers A and E (Lee et al. 1995), which target a portion of transfer RNA (tRNA)-pro and the central conserved region of the mitochondrial control region, were used in a polymerase chain reaction (PCR) to amplify the first hypervariable region with 0.5 units of Taq DNA polymerase (Amersham) in 10-µl reactions. The final concentrations were 100 nM forward and reverse primers, 200 µM dNTPs, 10 mM Tris pH 8.3, 50 mM KCl and 2.5 mM MgCl₂. Reactions were optimized at 40 cycles of 95°C for 30 s, 52°C for 30 s and 72°C for 45 s. PCR products were sequenced using a chain termination protocol in which each of the four deoxyxynucleotide terminators (ddNTPs) was labelled with 32P following the manufacturer’s directions (USB Thermosequenase radiolabelled terminator cycle sequencing kit). Both PCR primers and internal sequencing primers (AN.F3: 5’-AGCGATACACCAACTAACAAT-3’, AN.R4: 5’-TGGTCCGTTCTTACTACATTA-3’) were used to generate overlapping contigs such that > 90% of each reported fragment was sequenced in both forward and reverse directions. Sequencing products were electrophoresed on an 8% acrylimide gel and exposed to film overnight.

Descriptive statistics

All A. nigricaudus sequences were aligned manually. The number of unique haplotypes and number of transversions and transversions were counted. The average number of pairwise differences, θ (Nei & Li 1979), diversity based on the number of segregating sites, θ̂ (Watterson 1975), and their standard errors were calculated using ARLEQUIN (Schneider et al. 2000). Both θ and θ̂ estimate the neutral parameter θ = 2Nμ for mtDNA, where N is the effective population size for females and μ is the mitochondrial neutral mutation rate. To check for deviations from neutral expectations for the frequency spectrum of polymorphisms, significance values were calculated for Tajima’s D (Tajima 1989), Fu and Li’s D (Fu & Li 1993), and Fu’s F (Fu 1996) using a program made available by Fu (http://hgc.sph.uth.mcm.edu/fu/). These statistics compare values for θ and θ̂, the distribution of mutations on internal branches to external branches on a gene tree, and the number of haplotypes observed to a given value of θ̂, respectively.

Genealogical estimations

To estimate genealogical relationships among haplotypes and among populations, trees were constructed using maximum parsimony and neighbour-joining algorithms. For rooted parsimony analyses, A. nigricaudus haplotypes were aligned with a congener, A. armnulis, usingPILEUP in GCG. Parsimony analyses were conducted in PAUP 4.0b2a (Swoford 2000), with transitions and transversions weighted equally and also with transversions given a weight six times that of transitions (observed number of transitions to transversions in aligned sequences was 69 : 11). Heuristic searches were performed using stepwise addition for initial trees and the tree-bisection-reconnection method of branch swapping. For each search, the maximum limit of trees was set at 1000 and a strict consensus tree was generated. These conditions were repeated 10 times under both transition : transversion weighting schemes. Haploetype trees were also constructed in PAUP 4.0b2a using the neighbour-joining (Saitou & Nei 1987) with distances between sequences estimated under a Kimura (1980) two-parameter model and no among-site rate variation. A neighbour-joining tree based on genetic divergence, KST (Hudson et al. 1992a), between populations was estimated, in which each of the nine populations sampled was an operational taxonomic unit. This tree was bootstrapped 1000 times by resampling haplotypes with replacement within each population, recalculating KST, and generating a neighbour-joining tree for each bootstrap replicate. Randomization and KST calculations were performed using a program kindly provided by P. J. Palsboll and neighbour-joining trees from each bootstrap were constructed using PHYLIP (Felsenstein 1993).

Population subdivision

Two analyses of molecular variance (AMOVA: Excoffier et al. 1992) were used to test for significant population structure. Both AMOVAS were performed in ARLEQUIN (Schneider et al. 2000) using uncorrected pairwise differences between haplotypes and the haplotypes were permuted 1000 times. In the first analysis, a one-factor AMOVA was employed to assess the degree of population structure over all populations. Here, variation within populations was compared with variation between populations. In the second test, sequences from all nine geographical sites were included in a two-factor AMOVA in which variation was partitioned among biogeographic regions, populations and individuals. Populations were assigned to biogeographic regions following Thomson & Gilligan’s (1983) biogeographic regions (northern Gulf = Gonzaga + Los Angeles + Lobos + Libertad, central Gulf = Chivato + La Paz + Muertos + Kino + Venecia, Fig. 1).

Genetic distance and FST (Wright 1951) are often used to measure variation among populations. Because FST is inflated by anything that reduces within-population variation (Charlesworth 1998), an absolute measure of genetic distance, Dm (Nei 1973), was used to measure divergence among population pairs in this study. The minimum population distance, Dm (Nei 1973), is equivalent to the between-population distance corrected for a within-population
component of diversity (Nei & Li 1979). Average numbers of pairwise nucleotide differences between and within populations were calculated in ARLEQUIN (Schneider et al. 2000). Genetic distances ($D_{st}$) were estimated from between and within population diversity for the 36 possible population pairs using a weighting scheme described by Charlesworth (1998) to correct for unequal sample sizes. In order to test whether the genetic distances were significantly greater than random, all individual pairs of populations were tested for geographical subdivision using Hudson et al.’s (1992a) statistic $K$s, which is a weighted measure of average nucleotide difference between haplotypes from different populations. This statistic was calculated for each population pair and the significance was then assessed by randomly assigning haplotypes to each of the two populations (holding relative population sizes constant) and recalculating the statistic 10,000 times. The overall significance for population pairs was adjusted for multiple tests using a sequential Bonferroni correction. Because species-ance for population pairs was adjusted for multiple tests (holding relative population sizes constant) and randomly assigning haplotypes to each of the two populations (holding relative population sizes constant) and recalculating the statistic 10,000 times. The overall significance for population pairs was adjusted for multiple tests using a sequential Bonferroni correction.

Genetic distances between $A. nigricaudus$ populations may result primarily from one factor. Alternatively, several factors in combination might be responsible for the observed patterns of genetic differentiation. Whereas multiple regression analyses allow the contributions of individual parameters to be estimated, associated tests of significance rely on the assumption that all data points are independent. In our study, there are 36 nonindependent pairwise comparisons among nine populations. In order to avoid statistical tests that assume independence of data points, we first explored the data structure using multidimensional scaling. Second, we explicitly tested for contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation by comparing observed results with null distributions generated by permutations of the original data. Multidimensional scaling (Kruskal & Wish 1978) takes the distances between objects and tries to approximate those distances in a reduced number of dimensions through an iterative fitting procedure. A two-dimensional plot fitted to genetic distances ($K$s) between populations was created after 100 iterations in VISTA version 5.6 (http://forrest.psych.unc.edu/). The goodness of fit between the fitted and observed distances was measured by a stress test (Kruskal & Wish 1978). In order to assess whether eastern Gulf populations (separated by unsuitable sandy habitat) are, on average, more genetically distant from each other than western Gulf populations (connected by nearly continuous rocky habitat), two polygons were drawn connecting each set of populations and the area of each polygon was measured.

Whereas multidimensional scaling is a useful method for describing data, permutation-based tests can be used to test for statistical significance without assuming independence of data points. For example, a simple Mantel procedure (Mantel 1967) assesses the significance of a regression between two distance matrices by calculating the sum of element-by-element products and then compares that statistic with a null distribution of values created from permutations of elements within one of the matrices. This approach has been extended for three or more matrices using a number of approaches (Dow & Cheverud 1985; Hubert 1985; Manly 1986; Smouse et al. 1986). Here, we follow the procedures of Manly (1986, 1997b) in which a standard multiple regression is conducted, but significance is assessed by comparing observed statistics for the regression against those generated by permuting the dependent variables. This particular method has been used in a number of recent studies that attempt to identify factors associated with genetic differentiation (e.g. Brown et al. 1991; Lugon-Moulin et al. 1999; Pestano & Brown 1999).

To explore the effects of biogeography, geographical distance and discontinuous habitat on genetic distance in $A. nigricaudus$, we first defined six prediction vectors reflecting distances between population pairs based on: (i) biogeography (0 = same biogeographic region, 1 = different); (ii) geographical distance (km separating the two populations); (iii) discontinuous habitat due to sandy shores (0 = population pairs separated by rocky shore or open water, 1 = population pairs separated by sandy shore); (iv) discontinuous habitat due to open water (0 = population pairs separated by sandy shore, 1 = population pairs separated by rocky and sandy shore); (v) geographical distance * discontinuous habitat due to sandy shores; and (vi) geographical distance * discontinuous habitat due to open water. (The abbreviations BIOG, DIST, SAND, WATER, DIST * SAND and DIST * WATER are used throughout this study to refer to these six vectors.) There are three types of habitats separating the population pairs: open water, sandy shoreline and rocky shoreline. However, assigning three presence/absence binary codes, one for each habitat type (for example, for the habitat type of rock: 0 = sand or open water, 1 = rock), resulted in a design matrix that was not invertible and thus could not be solved for a least-squares solution. Therefore, the effects of the three habitat types could not be estimated independently but could be estimated only relative to each other. Thus, in order to test whether discontinuous habitat (water or sand) increases genetic distance relative to continuous habitat (rock), we created two variables, WATER and SAND, which represent the effect of open water minus the effect of rocky shore and the effect of sandy shore minus the effect of rocky shore, respectively. In both cases, a positive nonzero
correlation coefficient would indicate that the populations separated by discontinuous habitat (water or sand) are more genetically distant than populations separated by continuous habitat (rock). The interaction terms (DIST * SAND and DIST * WATER) were included because if discontinuous habitat reduces the opportunities for genetic interchange, the magnitude of the effect may be contingent upon the geographical distance. The vector of genetic distances (estimated by \( D_e \)) was used as the response variable. Thus, the full multiple regression model was:

\[
Y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 s + \beta_6 w + \beta_7 (s, w) + \epsilon
\]

where \( x_1, x_2, x_3, x_4, s, w \) are the prediction vectors BIOG, DIST, SAND and WATER, respectively, \( \beta_1, ..., \beta_7 \) are the estimated partial regression coefficients, \( \epsilon \) is error and \( Y \) is genetic distance. Following Manly (1986), each vector was normalized (zero mean, unit variance). Sums of squares and associated statistics for the whole model were calculated following standard regression procedures. Significances for the whole model and partial regression coefficients were estimated by permuting the order of values in the vector of genetic distances and recalculating the regression 10,000 times using the program version 2.1 (Manly 1997a). This Monte Carlo procedure was used to create null distributions for all statistics reported for this model. Because each factor was expected to have a positive effect, one-sided t-tests were used to determine whether the observed t-statistic was greater than the permuted distribution of t-statistics. Significances of the linear regression and six partial regression coefficients were adjusted using a sequential Bonferroni correction.

In order to visualize the interplay of habitat type and geographical distance while controlling for biogeography, residual genetic distances were calculated as follows:

\[
\text{Residuals} = Y - \hat{\beta} x
\]

where \( \hat{\beta} \) is the estimated partial regression coefficient for BIOG, and \( Y \) and \( x \) are the normalized vectors of genetic distance and BIOG. These residual genetic distances were plotted against geographical distance for each habitat type.

Results

Descriptive statistics

A 408-bp fragment was sequenced from 105 Auxilinus migrandus individuals. The first base in our sequence corresponds to position 16619 in the sequence of Zardoya et al. (1995). All sequences have been deposited in GenBank, with Accession nos AF333610–AF333695. Diversity within A. migrandus mtDNA was high but was similar to diversity among other fishes for the control region (Bowen & Grant 1997; Chenoweth et al. 1998a; McMillan et al. 1999; Rocha-Olivares & Vetter 1999). Among the 105 total sequences, there were 86 haplotypes (Fig. 2). Within the 408-bp sequence, there was one indel and 72 polymorphic sites, 27 of which were singletons. There were 69 observed transitions and 11 transversions in the aligned sequences. The species-level estimate of 2N\( e \) from average pairwise differences was \( \theta = 2.71\% \) (SE = 1.38%) and from the number of segregating sites was \( \theta = 3.37\% \) (SE = 0.89%) (both expressed per site). Tajima’s \( D \) (Tajima 1989) did not reject neutral expectations for the total data set (\( D = -0.827, P > 0.1 \)), however, both Fu and Li’s \( D \) (Fu & Li 1993), and Fu’s \( F \) (Fu 1996), which are more sensitive to the presence of singletons in a sample, were significantly different from neutral expectations (\( D = -2.843, P < 0.02; F = -19.904, P < 0.005 \)). These deviations indicate an excess of low frequency polymorphisms, as previously reported for mtDNA in other organisms (Excoffier 1990; Nachman et al. 1996) and may be consistent with the presence of mildly deleterious mutations, a recent selective sweep or nonequilibrium demographic effects (Tajima 1989). These deviations from a standard neutral model may have relatively little impact on the attainment of isolation-by-distance. Isolation-by-distance depends on a migration–drift equilibrium that may be achieved more quickly than mutation–drift equilibrium, provided that the migration rate is much higher than the mutation rate (Crow & Aoki 1984).

Genalogical estimations

All estimates of genealogical relationships showed substantial concordance with biogeography. Parsimony analyses of the 86 haplotypes identified several thousand equally parsimonious trees regardless of weighting scheme. When transversions and transitions were weighted equally, the tree lengths were 189 and the consistency indices (CI) equalled 0.42. A strict-consensus tree contained two clades, one consisting mostly of individuals from the northern Gulf and the other consisting mostly of individuals from the central Gulf. The central Gulf clade contained two smaller clades, corresponding mostly to haplotypes from western (Baja) Gulf and eastern (Sonora) Gulf populations, respectively. A haplotype network (Fig. 3A) was constructed based on a
Fig. 3. (A) Haplotype network constructed from a single unweighted parsimony tree. Individual haplotypes are indicated by circles that are proportional in area to the number of individuals observed to have that haplotype. Hash marks indicate the number of mutational steps separating each haplotype. The number in each circle refers to the haplotype and the letter(s) refers to the geographical location from which it was sampled (G = Gonzaga, LA = Los Angeles, C = Chivato, P = La Paz, M = Muertos, Ls = Lobos, Ld = Libertad, K = Kino, V = Venecia). Three major clades are labelled and consist largely of individuals from the indicated geographical regions. Branches present in the strict consensus of all 1000 equally parsimonious trees are indicated in bold. (B) Neighbour-joining tree based on distances between populations estimated by $K_{ST}$. Support for individual branches is given by bootstrap percentages, where haplotypes were drawn randomly with replacement within each of the nine populations and $K_{ST}$ was recalculated for each of the 1000 bootstrap replicates. Three major clades are labelled by geographical region.
randomly chosen unweighted parsimony tree. Trees based on a 6:1 weighting of transversions to transitions had tree lengths equal to 258 and consistency indices equal to 0.52. The consensus trees obtained with a 6:1 weighting contained the central western Gulf and central eastern Gulf clades, but the large northern Gulf clade collapsed into a basal paraphyletic assemblage. The neighbour-joining tree based on haplotypes contained three clades corresponding to populations from northern, western-central and eastern-central Gulf locations. Bootstrap support was high (95%) for the branch separating northern, western-central and eastern-central Gulf locations. Different genealogical analyses are concordant in suggesting separate evolutionary histories among groups, and within populations, revealed significant partitioning both among regions (biogeographic regions), among populations (P < 0.001). The two-factor AMOVA (Fig. 3A) also contained three clades corresponding to populations from northern, western-central and eastern-central Gulf locations. Bootstrap support was high (95%) for the branch separating northern and central Gulf clades. These different genealogical analyses are concordant in suggesting separate evolutionary histories for northern and central Gulf populations and also an east–west division within the central Gulf populations.

Population subdivision

Haplotypic diversity was significantly partitioned among populations and biogeographic regions in A. nigricaudus. In the one-factor AMOVA of all populations, much variation was explained by differences among populations (ΦST = 0.485, P < 0.001). The two-factor AMOVA that partitioned variation among groups (biogeographic regions), among populations in groups, and within populations, revealed significant partitioning both among regions (ΦST = 0.307, P < 0.001) and among populations (ϕCT = 0.536, P < 0.001). Haplotypes were also nonrandomly partitioned between many pairs of populations as shown by Hudson et al.’s (1992a) permutation test; in 28 of 36 possible populations pairs, the statistic Ks* was higher than expected by chance (Table 2). In many cases there was significant partitioning between populations, including those separated by < 200 km (Table 2). However, there was no significant genetic differentiation between populations separated by < 100 km, and some population pairs up to 483 km apart were also not significantly differentiated. For A. nigricaudus, species-level Fs0 was equal to 0.536.

Biogeography, geographical distance and discontinuous habitat

A multidimensional scaling of genetic distances generated a good fit to the observed data (stress = 0.037 after 100 iterations, where a zero stress value indicates a perfect fit). The multidimensional scaling plot revealed the same clusters of populations similar to those observed from genealogical analyses (Fig. 4A): northern Gulf, western-central Gulf and eastern-central Gulf. The polygon enclosing eastern Gulf populations that are connected by almost continuous rock habitat (eastern Gulf polygon area = 1.17, western Gulf polygon area = 0.89), despite the fact that the populations separated by discontinuous habitat cover a much smaller geographical area (Fig. 1). The multiple regression of six vectors (BIOG, DIST, SAND, WATER, DIST * SAND and DIST * WATER) on genetic distance was highly significant and explained 71% of the variance in Dst between populations (R2 = 0.705, F = 10.32, P < 0.05). The partial regression coefficients of BIOG, DIST, WATER and DIST * SAND were significantly greater than random (P < 0.05; Table 3), indicating that these factors contribute to genetic distances in A. nigricaudus, where partial regression coefficients represent the slope of a given variable when all other variables are held constant. A scatterplot of geographical distance vs. residual genetic distance illustrates an overall pattern of increasing genetic distance with increased geographical distance (Fig. 4B).

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Los Angeles</th>
<th>Muertos</th>
<th>Lobos</th>
<th>Libertad</th>
<th>Kino</th>
<th>Venice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gonzaga</td>
<td>10 0.666</td>
<td>121 267</td>
<td>388</td>
<td>731 612</td>
<td>791 672</td>
<td>156</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>10 0.690**</td>
<td>1.562</td>
<td>2.273</td>
<td>345 404</td>
<td>365 329</td>
<td>192 154</td>
</tr>
<tr>
<td>Chivato</td>
<td>9 1.428**</td>
<td>1.093**</td>
<td></td>
<td>2.097 2.084</td>
<td>64 704</td>
<td>670</td>
</tr>
<tr>
<td>La Paz</td>
<td>6 1.828**</td>
<td>1.632**</td>
<td>0.000</td>
<td>0.062 1.877</td>
<td>759 730</td>
<td>577 483</td>
</tr>
<tr>
<td>Muertos</td>
<td>10 1.826**</td>
<td>1.782**</td>
<td>0.160</td>
<td>0.062 1.877</td>
<td>759 730</td>
<td>577 483</td>
</tr>
<tr>
<td>Lobos</td>
<td>10 1.924**</td>
<td>0.317</td>
<td>1.457**</td>
<td>2.043 2.254**</td>
<td>1.343 28 183</td>
<td></td>
</tr>
<tr>
<td>Libertad</td>
<td>10 1.621**</td>
<td>0.208</td>
<td>0.032</td>
<td>1.159 1.743**</td>
<td>−0.029 1.802 154 230</td>
<td></td>
</tr>
<tr>
<td>Kino</td>
<td>10 1.810**</td>
<td>1.372**</td>
<td>1.131**</td>
<td>0.998 1.271**</td>
<td>1.788 1.503**</td>
<td>1.365 100</td>
</tr>
<tr>
<td>Venice</td>
<td>30 2.133**</td>
<td>2.065**</td>
<td>1.778*</td>
<td>1.512 1.576</td>
<td>2.453 2.244**</td>
<td>0.115 1.13</td>
</tr>
</tbody>
</table>

*Within population nucleotide diversity, †is on the diagonal. Genetic distances between populations (Dst) are below the diagonal. Geographical distances (km) separating populations are above the diagonal. Significance of genetic partitioning was evaluated using the statistic Ks* with a sequential Bonferroni correction for multiple tests and is indicated by asterisks: * < 0.05, ** < 0.001. See text for further details.

© 2001 Blackwell Science Ltd, Molecular Ecology, 10, 1439–1453
Fig. 4 (A) Multidimensional scaling plot of populations based on genetic distances ($D_m$) between pairs of populations. Populations from the western Gulf that are connected by rocky habitat are shown by triangles and populations from the eastern Gulf that are separated by sandy habitat are shown by squares. The area of the polygon defined by the western Gulf populations equals 0.89 and the area of the polygon defined by the eastern Gulf populations equals 1.17. Greater genetic distances among eastern Gulf populations relative to western Gulf populations are consistent with discontinuous habitat (sand) reducing genetic interchange. (B) Scatterplot of residual genetic distances (controlling for biogeography) on geographical distance for pairs of populations separated by discontinuous habitat (sand and open water) and continuous habitat (rock). A modified Mantel permutation test shows that population pairs separated by open water are more genetically distant overall than populations pairs connected by rocky coastline (Table 3: $t = 2.94$, $P < 0.05$). In addition, as geographical distance increases, populations separated by sandy coastline become more genetically distant than those population pairs connected by rocky habitat (Table 3: $t = 3.01$, $P < 0.05$). See text for further details.
Although population pairs from different biogeographic regions are generally more distant than pairs from the same region, the significant positive coefficients for both BIOG and DIST demonstrate that each factor independently contributes to genetic distance. Because WATER represents the difference between population pairs separated by open water (discontinuous habitat) and rocky shore (continuous habitat), the positive partial regression coefficient for WATER indicates that population pairs separated by open water are on average more genetically distant than those connected by continuous habitat (Fig. 4B). The positive coefficient for DIST * SAND demonstrates that, as geographical distance increases, the difference in genetic distances between populations separated by sand (discontinuous habitat) and rocky shore increases (Fig. 4B).

Discussion

In Axochinus nigricaudus, we find that there is substantial partitioning of genetic variation over relatively short geographical distances (< 200 km, in some cases), and that the patterns of genetic differentiation are best explained by the combined effects of biogeography, geographical distance and discontinuous habitat.

Biogeography, geographical distance and habitat discontinuities

Phylogenetic relationships among mtDNA haplotypes revealed a deep division between northern and central Gulf populations. Unweighted parsimony and neighbour-joining analyses of haplotypes, as well as a neighbour-joining tree of populations, produced trees that contained two large clades corresponding mostly to northern and central Gulf haplotypes and populations. Among all parsimony trees, northern and central Gulf clades were separated by nine mutational steps (Fig. 3A), suggesting separate evolutionary histories for these two groups. The two-factor AMOVA, based on haplotype distances, also reflected the strong partitioning of haplotypes between biogeographic regions (\(F_{CT} = 0.307, P < 0.007\)). Overall, the concordance between A. nigricaudus genealogy and previously identified biogeographic regions is striking.

Delineations of biogeographic regions in the Gulf of California have been largely based on changes in fish community composition and do not necessarily point to an underlying cause of the north–central division. Ecological differences, particularly cold winter temperatures, probably prevent range expansion of many tropically derived species into the northern Gulf (Thomson & Lehner 1976). Although temperature may contribute to changes in community composition, it seems unlikely that selection due to environmental factors (such as temperature) would cause a deep divergence among A. nigricaudus mtDNA haplotypes. The observed partitioning of A. nigricaudus mtDNA between biogeographic regions is also puzzling because there are no obvious geological events or oceanographic circulation patterns (Badan-Dangon et al. 1985; Bray 1988; Paden et al. 1991) that point to a northern–central Gulf division. In addition, there are no known sister taxa on opposite sides of this division, and northern–central Gulf partitioning has not been observed in Malacocentrus hubbsi (C. Riginos unpublished) or Girdula nigricans (Terry et al. 2000), the only other species surveyed for genetic variation across both regions. Nonetheless, the genetic partitioning of A. nigricaudus populations between northern and central Gulf regions suggests that genetic interchange across this boundary may be more restricted than previously recognized. If genetic surveys of additional taxa reveal other examples of northern–central Gulf divergence, this pattern would point to an historical factor affecting many species.

Although 31% of genetic variation in A. nigricaudus was partitioned between biogeographic regions, discontinuous habitat and geographical distance also contributed to

| Variable‡ | Partial regression coefficient (β|β|) | t | P‡ |
|-----------|--------------------------------|---|----|
| BIOG      | 0.437                          | 3.53 | 0.0009* |
| DIST      | 0.472                          | 2.39 | 0.0164* |
| SAND      | -0.235                         | -0.93 | 0.8209 |
| WATER     | 0.709                          | 2.94 | 0.0036* |
| DIST * SAND | 0.669                          | 3.01 | 0.0019* |
| DIST * WATER | -0.392                        | -1.56 | 0.8331 |

‡Multiple regression: \(R^2 = 0.705, F = 10.32, P = 0.0001. \) BIOG, biogeography; DIST, geographical distance; SAND, sandy shore; WATER, open waters. See Materials and Methods for a full explanation. §Equivalent to β values in full regression model. See Materials and Methods. ¶Probability was assessed by following a Monte Carlo procedure where observed t and F-values were compared against a null distribution of t and F-values based on 10 000 randomizations of response values (genetic distances). Asterisks indicate values that remain significant (P < 0.05) following a sequential Bonferroni correction.

© 2001 Blackwell Science Ltd, Molecular Ecology, 10, 1439–1453
divergence among populations. Overall, genetic distance increased with geographical distance (Fig. 4B), consistent with isolation-by-distance. Although BIOG and DIST were correlated (R² = 0.181), the significant partial regression coefficient for DIST (t = 2.93, P < 0.05) showed that there was an effect of geographical distance beyond that explained by biogeography. Studies of coconut crabs (Lavery et al. 1995, 1996) also demonstrate additive effects from geographical distance and biogeographic regions, where genetic distances between Pacific populations and a single Indian Ocean population are greater than genetic distances among Pacific populations.

Discontinuous habitat also accounted for substantial genetic divergence. The partial regression coefficient of WATER was significantly greater than zero (t = 2.94, P < 0.05), indicating that population pairs separated by open water were more genetically distant, on average, than population pairs connected by rocky shoreline. However, there was much scatter among the open water data points, with large genetic distances found both over small and large geographical distances (Fig. 4B). This scatter was largely due to differing patterns of genetic connectivity across open water within each biogeographic region. Distinct central-west Gulf and central-eastern Gulf groups were present in all genealogical estimations and in a multidi- mensional scaling plot, and population distances were high between these regions (D⁰ ≥ 0.998, Table 2). This west–east division suggests that the open waters of the central Gulf are a strong barrier to genetic interchange. In contrast, among northern Gulf fish, several haplotypes from Los Angeles (west) were most closely related to eastern haplotypes from Lobos and Libertad (Figs 2 and 3A), resulting in low D⁰ values (Table 2: D⁰ < 0.317). It is likely that stepping-stone populations on the islands Isla de la Guardia and Isla Tiburon facilitate limited genetic interchange between Los Angeles and north-eastern populations (Fig. 1). Overall, genetic distances were higher across open water than across rocky shoreline, and this result is consistent with other studies that have demonstrated a reduction in gene flow across open water in coastal marine organisms (e.g. Bell et al. 1982; Stepien & Rosenblatt 1991; Doherty et al. 1995; Johnson & Black 1995).

Multidimensional scaling (Fig. 4A) showed that population pairs connected by sandy shoreline were more genetically distant than population pairs connected by rocky coast, even though rocky coast populations span a greater geographical distance than sandy shore populations (791 and 281 km, respectively, Fig. 1). In the multiple regression model, in contrast, the partial regression of SAND was not significant (P > 0.05, Table 3). The interaction DIST * SAND, however, was significant (t = 3.01, P < 0.05). This interaction is best understood by examining Fig. 4B: the partial residuals of genetic distance (with the effect of biogeography removed) for population pairs separated by both sand and rock habitats showed a generally linear relationship of increasing genetic distance with increased geographical distance. However, the partial regression slope for population pairs separated by sand was greater than the slope for populations separated by rock; this difference in slope is expressed by the statistical significance of the DIST * SAND interaction term. Over increased distance, sandy shoreline reduces genetic interchange among A. nigricaudus populations. Similar differences in isolation-by-distance slopes are observed between estuarine and shoreline populations of an atherinid fish (Johnson et al. 1994) and a littorinid snail (Johnson & Black 1998).

Extreme genetic partitioning in A. nigricaudus

In A. nigricaudus, a very high level of genetic partitioning was often found over relatively small distances (< 200 km; Table 2). Differences among populations explained 48.5% of the variation (one-factor AMOVA, P < 0.001), and Hudson et al.’s (1994b) permutation test provided strong statistical support for differences between many pairs of populations (Table 2). In fact, the degree of genetic partitioning found in A. nigricaudus is higher than that observed for nearly all marine fishes. In A. nigricaudus, FST = 0.536, where FST (Hudson et al. 1992b) for mtDNA estimates the parameter 1/(2Nmt + 1), N is the female effective population size and m is the female migration rate. Our estimate of 0.536 corresponds to a value of 0.224 for autosomal loci, assuming a sex ratio of one and equal migration of males and females. This estimate of partitioning is generally higher than reported values for reef fishes based on allozymes (FST ranges from 0.002 to 0.032, Waples 1987; average FST = 0.062, Ward et al. 1994; FST ranges from 0.0023 to 0.1124, Shulman 1998) or mtDNA (e.g. FST ranges from −0.012 to 0.172, Shulman & Bermingham 1999). Both Acanthochromis polyacanthus (FST = 0.7919, Doherty et al. 1995) and Embiotocidae jacksoni (FST = 0.444, Waples 1987) populations have greater population structure than A. nigricaudus, but both also lack a planktonic larval stage. A. nigricaudus has a planktonic larval stage, but its duration is unknown. Among fishes that do have a larval stage, the single example of an FST greater than that observed for

© 2001 Blackwell Science Ltd, Molecular Ecology, 10, 1439–1453


A. nigricaudus in Abrolhos Islands Ctenocephalus caproli (FST = 0.437 over 35 km), although surprisingly FST was considerably less when C. caproli were sampled over a greater geographical range (FST = 0.073 over 850 km, Johnson et al. 1994).

Two aspects of A. nigricaudus life history may contribute to the high degree of genetic partitioning. First, A. nigricaudus larvae probably remain close to shore after hatching (Brogan 1994), which might reduce migration. Second, the species range of A. nigricaudus is geographically complex, spanning two parallel shorelines, a number of islands and two ecologically distinct regions (Fig. 1). Here, we decomposed some of that geographical complexity into three factors (biogeography, geographical distance and discontinuous habitat) and found that each factor contributed to population divergence. These results suggest that where multiple factors can limit dispersal, a planktonic larval stage does not preclude substantial genetic partitioning over short geographical distances.

Acknowledgements

We thank J. S. Shomer, J. McConnell and J. Francis for assistance in the field, and V. E. Riginos for assistance with computer programs. We also thank P. A. Hastings, G. Bernardi, S. R. Palumbi, P. J. Palsbøll, and two anonymous reviewers for comments on the manuscript. Specimens were collected under SEMARNAP permit no. 210597-213-03. This work was funded by Sigma Xi, the American Museum of Natural History, the Flinn Foundation, the Department of Ecology and Evolutionary Biology at the University of Arizona, and NSF.

References


Benzie JA, Williams ST (1997) Genetic structure of giant clam (Tridacna maxima) populations in the west Pacific is not consistent with dispersal by present-day ocean currents. Evolution, 51, 768–783.


Johnson MS, Black R (1996b) Increased genetic divergence and reduced genetic variation in populations of the snail Rembicium vitatum in isolatedtidal pools. Heredity, 80, 163–172.


This research is part of a larger project that looks at how extrinsic and intrinsic factors shape genetic variation in Gulf of California blennioid fishes. This work represents part of Cynthia Riginos’s PhD dissertation. Michael W. Nachman’s research interests include molecular population genetics and the genetics of speciation.