Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 43 (2007) 743–754
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

The Pleistocene history of the sheepshead minnow (*Cyprinodon variegatus*): Non-equilibrium evolutionary dynamics within a diversifying species complex

Robert A. Haney [a,*], Brian R. Silliman [b], Adam J. Fry [c], Craig A. Layman [d], David M. Rand [a]

[a] *Department of Ecology and Evolutionary Biology, Box G-W, Brown University, Providence, RI 02912, USA*
[b] *Department of Zoology, University of Florida, Gainesville, FL 32611, USA*
[c] *Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269, USA*
[d] *Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA*

Received 7 May 2006; revised 24 August 2006; accepted 31 August 2006
Available online 9 September 2006

Abstract

The sheepshead minnow, *Cyprinodon variegatus*, is a widespread fish species that typically inhabits coastal tidal marsh and mangrove swamp environments, ranging from Cape Cod, Massaschusetts to northern Mexico and into the Caribbean. This wide range crosses several biogeographic boundaries which are coincident with genetic structuring within numerous species originating in the Pleistocene. In addition, the more northerly reaches of this species range have been further subject to the evolutionary consequences of Pleistocene glaciation due to local extinction and recolonization of formerly glaciated sites. *C. variegatus* thus provides an excellent vertebrate model system within which to test the extent of genetic differentiation among populations in a dominant coastal ecosystem and examine patterns of historical demography in populations distributed along a latitudinal gradient. Using mitochondrial control region and ND2 sequence data, we discovered monophyletic clades within *C. variegatus* with divergence times within the Pleistocene, and very low gene flow between most sites. Intraspecific genetic breaks appear to correspond broadly to biogeographic or oceanic boundaries. Pleistocene climate change appears to have had dramatic impacts on the size and distribution of populations within and near the glacial margins, but has also affected populations far from formerly glaciated regions.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Gene flow; Demography; Mitochondrial control region; *Cyprinodon*; Phylogeography

1. Introduction

The coastal regions of the eastern United States span a range of climate zones, and the representation of taxa within coastal communities subsequently varies strongly from the northern Atlantic to the Gulf of Mexico. Traditionally, this region has been subdivided into three to four biogeographic zones based on the distribution of marine taxa (Briggs, 1974, 1995), although more recent analysis of estuarine invertebrate community composition suggests the presence of a greater number of biogeographic provinces for these taxa (Engle and Summers, 1999). The discipline of phylogeography originated in this region with the discovery that the boundaries among biogeographic regions were coincident with disruptions of gene flow among populations within numerous species, suggesting a connection between microevolution and macroecological process (Avise et al., 1987a, 2000). Further, the depth of phylogenetic breaks in some taxa indicated that historical factors, in particular Pleistocene glacial cycles, may have initiated differentiation among populations that is maintained to the present by current patterns and ecological gradients (Reeb and Avise, 1990).

* Corresponding author. Fax: +1 401 863 2166.
E-mail address: Robert_Haney@brown.edu (R.A. Haney).

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.08.022

Documented barriers to gene flow occur in a wide range of species along the Atlantic and Gulf coasts (Avise et al., 1987b; Reeb and Avise, 1990; Avise, 2000; Collin, 2001; Young et al., 2002; McMillen-Jackson and Bert, 2003; Weinberg et al., 2003; Caudill and Bucklin, 2004; Gurgel et al., 2004; Kelly et al., 2006). However, data from species residing in the dominant habitat on the east coast, tidal marshes, are limited. Previous studies on tidal marsh teleosts have focused on putative subspecific phylogenetic disjunctions within two species with partially overlapping distributions and similar life history traits. Populations of *Fundulus heteroclitus* north and south of New Jersey are distinct in both the mitochondrial and nuclear genomes, and may constitute separate subspecies (Gonzalez-Villasenor and Powers, 1990; Bernardi et al., 1993; Smith et al., 1998). Similarly, *Cyprinodon variegatus*, a euryhaline species that commonly inhabits environments such as estuarine tidal marshes and mangrove swamps, has also been suggested to consist of two allopatric subspecies, *C.v. ovinus* in the north and *C.v. variegatus* in the south (Hubbs, 1936; Finne, 2001). Within the distributional range of subspecies, levels of gene flow, and in particular the influence of biogeographic boundaries on genetic differentiation, remain unknown.

In addition to the promotion of vicariance, the climatic effects of Pleistocene glaciation may have strongly altered species ranges. This is particularly true of species whose range in whole or in part consisted of regions under glacial ice (Hewitt, 2000). Local extinction and recolonization, or bottlenecks and subsequent population recovery, leave a distinctive signature in the genetic structure of populations. Data for nearshore taxa of the eastern United States are limited. Wares and Cunningham (2001) documented patterns of genetic variation in six species of rocky intertidal invertebrates whose North American habitat was expected to be strongly reduced by past glaciation, due to limited suitable habitat south of the ice sheets. Two primary patterns were found in that study. Four species appeared to have recently colonized the North American coast from Europe after glacial retreat, while two appeared to persist in refugial areas north of Cape Cod. It would be instructive to examine post-glacial recolonization in a coastal species that had extensive suitable habitat to the south of glacial margins on the eastern North American coastline, as is the case for many estuarine species. Studies of this sort are lacking, and *C. variegatus* provides an effective model to investigate this biogeographic scenario.

In this study, we use a combination of phylogeographic and population genetic approaches to elucidate the evolutionary history of *C. variegatus.* We take an hierarchical approach, investigating spatial patterns of genetic variation and testing for the existence of intraspecific genetic breaks through the use of both phylogenetic and population genetic analyses. The application of population genetic techniques within established phylogenetic groupings such as subspecies not only allows insight into divergence events occurring on a shallower timeframe than those identified through phylogenetic analysis, but also allows us to study processes, such as population growth and range expansion, that leave distinct, but fleeting, molecular signatures.

Recent phylogenetic treatments indicate that *C. variegatus* is part of a maritime clade of pupfishes that includes several endemic Caribbean species (Echelle et al., 2005). Within this maritime clade, two divergent mitochondrial lineages occur. Mainland populations of *C. variegatus* belong to both of these lineages, which thus appear to delineate the coastal/mainland northern subspecies *C.v. ovinus* and the more southerly distributed *C.v.variegatus.* The divergence among subspecies appears to date to approximately the Pliocene/Pleistocene boundary (Echelle et al., 2005). Representatives of both lineages have appeared to colonize islands in the Caribbean during the Pleistocene, contributing to the genetic diversity of the insular Caribbean pupfish fauna, and rendering *C. variegatus* paraphyletic with respect to Caribbean species (Echelle et al., 2006). Hence, to address the first goal of this study, that of examining genetic differentiation in *C. variegatus* and establishing its correspondence to biogeographic boundaries, we must establish a broad framework of the evolution of *C. variegatus.* We do this by extensively sampling populations and placing them within the phylogenetic context of the maritime clade of *Cyprinodon.* This will also allow a fuller accounting of distinct lineages in this species and their relation to specific and subspecific boundaries. The geographic distribution of these lineages is then used in conjunction with population genetic estimators of levels of differentiation to infer the relative importance of known gene flow barriers on the Atlantic and Gulf coasts for *C. variegatus.* We test whether large-scale ecological processes that act to define zoogeographic provinces are important for disrupting gene flow, and whether barriers to dispersal also exist on a smaller spatial scale. We examine the timing of clade divergences to establish a chronology of diversification and gain insight into historical factors influencing differentiation.

A second goal of this study is to determine whether climate change related to glacial oscillations may have influenced historical population sizes and distribution, and whether effects varied as a result of proximity to the glacial front. *C. variegatus* is well-suited to this test, with a broad distribution from Cape Cod in the north to Mexico in the south. We expect that populations of the northern subspecies (*C.v. ovinus*) should have been more strongly perturbed by the most recent glacial advance, while southern populations may show long term stable population sizes and little evidence of range shifts. The incorporation of neutrality tests and maximum-likelihood growth models within a phylogeographical framework makes use of the information in haplotype frequency distributions that reflect historical processes such as bottlenecks, population growth and range expansions. Herein we test specific hypotheses related to the demography of *C. variegatus* by combining these tests with more commonly used phylogeographic methodologies. First, are perturbations in population size and

changes in distribution due to the most recent glaciation restricted to populations of the northern subspecies *C.v. ovinus*? The current distribution of *C.v. ovinus* includes the region from northern New Jersey to Cape Cod, MA, which was under ice during the last glacial advance. It follows that residence in this region must be due to recent recolonization from ice-free areas. A clinal decrease in average allozyme heterozygosity in *C. variegatus* occurs from North Carolina to southern Connecticut, consistent with an hypothesis of recolonization from ice-free southern locales (Darling, 1976). No such cline was observed in southern populations of *C. variegatus*. Hence we expect populations occurring within or close to the former ice-sheet margins to show marked reductions in mitochondrial genetic diversity and a signature of recent range expansion and population growth, but that such a signature would be lacking in the south. However, if habitat loss were related most strongly to changes in sea level, the effects may be pronounced for all populations, independent of latitude. To examine these alternatives, we test whether there is evidence for population growth and range expansion that has occurred subsequent to the last glacial maximum at approximately 20,000 years before the present, and determine whether results differ in disjunct clades of this species occupying divergent latitudes.

Table 1
Collection locations and sample sizes for *C. variegatus* along the East and Gulf coasts of North America and in the Caribbean

Site	Abbreviation	Sample Size	Approx. Lat./ Long.
West Falmouth, MA	MA	22	41.60N/70.64W
Barrington, RI	RI	30	41.75N/71.32W
Wequetequock, CT	CT	20	41.36N/71.88W
West Mantoloking, NJ	WMNJ	23	40.04N/74.05W
Tuckerton, NJ	TTNJ	23	39.60N/74.34W
Oyster, VA	VA	23	37.29N/75.92W
Georgetown, SC	SC	21	33.23N/79.17W
Sapelo Island, GA	GA	24	31.39N/81.31W
Windley Key, FL	WK	33	24.92N/80.63W
Summerland Key FL	SK	33	24.67N/81.44W
Port Fourchon, Louisiana	LA	21	29.13N/90.21W
Corpus Christi, Texas	TX	10	27.74N/97.40W
Pain Pond, San Salvador	SS	26	24.05N/74.53W
Crescent Pond, San Salvador	SS	30	24.05N/74.53W
Wild Dilly Pond, San Salvador	SS	7	24.05N/74.53W
Bird Pond, Andros	AN	5	25.06N/78.05W

Sample size refers to number of sequences obtained for control region dataset.

2. Materials and methods

Individuals of *C. variegatus* were collected by dip net, seine or cast net, or in minnow traps from twelve sites along the Atlantic and Gulf coasts of the United States (Table 1; Fig. 1). In addition, individuals were collected from four lakes and ponds on the islands of Andros and San Salvador, Bahamas. Three sympatric forms of *C. variegatus* occur on San Salvador. These forms differ morphologically, ecologically and genetically (Holtmeier, 2001). Only the "normal" forms which are similar to coastal *C. variegatus* were used in this study.

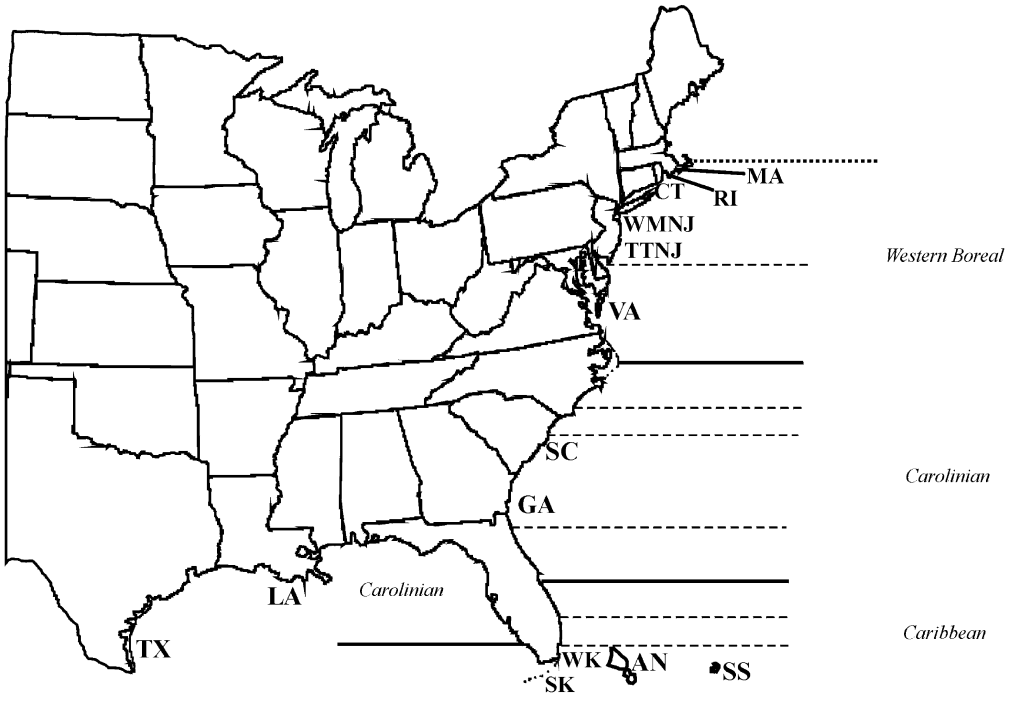

Fig. 1. Sampling locations of *C. variegatus* on the Atlantic and Gulf coasts of North America, and in the Florida Keys. Heavy lines indicate the approximate boundaries between biogeographic provinces as defined by Briggs (1995), with the names of provinces indicated in italics. Dashed lines indicate boundaries of biogeographic provinces based on the distribution of estuarine invertebrates (Engle and Summers, 1999). The uppermost dashed line indicates the northern range boundary of *C. variegatus*.

2.1. Laboratory techniques

Genomic DNA was extracted from muscle tissue by digestion with Proteinase-K, followed by aqueous extraction with phenol/chloroform isoamyl alcohol, and precipitation with 95% ethanol overnight or by using a DNeasy kit (Qiagen, Rockville, MD).

PCR amplification of a portion of the mitochondrial control region was performed with a combination of three primers (K,A,B) described in Lee et al. (1995). The thermal profile for the amplification was as follows: initial denaturation at 94 °C for 4 min, forty cycles of 94 °C for 30 s, 52 °C for 30 s and 72 °C for 45 s, followed by a final extension at 72 °C for 6 min. PCR products were sequenced in both directions on ABI377 or ABI3730 automated sequencers. Amplification of a section of the ND2 gene was accomplished with primers L4633 and H5443 (Miya and Nishida, 1999) using the following reaction conditions: initial denaturation at 94 °C for 2 min, forty to forty-five cycles of 94 °C for 30 s, 52 °C for 30 s and 72 °C for 45 s, followed by a final extension at 72 °C for 6 min. PCR products were sequenced in the forward direction on ABI3730 automated sequencers.

2.2. Sequence alignment, phylogenetic and population genetic analysis

Forward and reverse control region sequences were assembled in Sequencher (Gene Codes Corporation, Ann Arbor, MI) and consensus sequences were aligned in Bioedit (Hall, 1999). For each distinct control region haplotype, we amplified and sequenced ND2 from the same individual. Sequences were deposited in Genbank under Accession No. EF050588-EF050727. Composite control region/ND2 haplotypes were created and then aligned with all composite sequences from earlier studies by Echelle et al. (2005, 2006) obtained from Genbank in preparation for phylogenetic analysis.

Prior to phylogenetic reconstruction of a first dataset comprised of all distinct composite haplotypes, the best-fit model of sequence evolution was determined by hierarchical likelihood-ratio test in Modeltest 3.5 (Posada and Crandall, 1998). Haplotype relationships were ascertained through Bayesian analysis using MrBAYES 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) using the best-fit model of sequence evolution as determined by Modeltest 3.5 with associated parameters estimated from the data. Four Markov chains were run for one million generations with sampling every 100 generations. The log file of the run was examined for the approximate likelihood at which a plateau in improvement was achieved, and all trees prior to this point were discarded. A majority rule consensus of the remaining trees was created in PAUP*v. 4.04b10 (Swofford, 1998). Posterior probability values were used to assess the significance of clades recovered in the analysis.

A second dataset was composed of all ND2 sequences from this and previous studies. A Bayes factor analysis was used to test for evidence in favor of a molecular clock in the mitochondrial ND2 sequence data. Runs with and without a uniform molecular clock were first performed for a reduced dataset of 44 sequences representing all major lineages, and the harmonic mean of the tree likelihoods for all samples taken in the stationary phase of the analysis were used as estimates of the marginal likelihood under each model. The Bayes factor for model comparison was then calculated as twice the difference of the logarithm of the harmonic means and support for the non-clock model was interpreted following Kass and Raftery (1995). We then calculated average net pairwise divergences among clades for ND2 in MEGA 3.1 (Kumar et al., 2001) under a gamma-corrected Tamura–Nei nucleotide substitution model with a shape parameter of 0.5. A clock calibration for ND2 based on the Cyprinodont genus *Aphanius* (Hrbek and Meyer, 2003; Perdices et al., 2001; Echelle et al., 2005) was then used to estimate divergence times.

As a second approach to calculating divergence times among clades, we followed the basic procedure outlined in Echelle et al. (2005) to accommodate unequal substitution rates across lineages. First, a neighbor-joining tree was constructed in PAUP*v. 4.04b10 using Tamura–Nei distances with a gamma correction for rate variation among sites with a shape parameter of 0.5. This tree was imported into Treeedit v. 1.0a10 (Rambaut and Charleston, Department of Zoology, University of Oxford) and submitted to nonparametric rate smoothing (Sanderson, 1997) to moderate unequal rates of substitution across branches and produce an ultrametric tree with branch lengths proportional to time. Branches that were not supported in the Bayesian analysis were collapsed, and adjusted levels of clade divergence read from the tree. These were then converted to times using the available clock calibrations.

A third dataset was used for population genetic analyses and consisted of all control region sequences obtained for this study. Genetic differentiation among sites within clades was assessed by calculating pairwise F_{st} values following Hudson et al. (1992, equation 3) and the S_{nn} statistic of Hudson (2000). S_{nn} is a sequence based statistic that measures how often nearest-neighbors in terms of sequence divergence are found in the same location. Significance of the S_{nn} statistic was determined by permutation test in DNAsp v. 3.99 (Rozas and Rozas, 1999) with 1000 replications and gaps excluded in pairwise comparisons.

Haplotype and nucleotide diversity, along with the average number of pairwise differences within populations, were calculated in DNAsp v.3.99 (Rozas and Rozas, 1999).

We tested whether the pattern of observed polymorphism within populations is consistent with a neutral equilibrium Wright–Fisher model using Fu's F_s (Fu, 1997), computed as $\ln(S'/1-S')$ where S' is the probability of having no fewer than the observed number of alleles conditioned on θ estimated from the average number of nucleotide differences. This statistic takes on a negative value with

an excess of rare haplotypes as may occur under scenarios of background selection, selective sweeps, or population expansions. We chose F_s as it has proven to be the most sensitive statistic with respect to demographic expansion (Fu, 1997; Ramos-Onsins and Rozas, 2002). One thousand replicates of a coalescent simulation were used to determine the empirical distribution of the test statistic and to calculate confidence intervals and the significance level of the observed value in DNAsp v.3.99 (Rozas and Rozas, 1999). We further explored the history of population size using the program Fluctuate v. 1.4 (Kuhner et al., 1998). This program implements a Metropolis–Hastings genealogy sampling approach to obtain joint maximum-likelihood estimates of θ and the population growth parameter g under an exponential growth model. Estimates of time to 1% of current population size were obtained from the equation for exponential growth using a control region mutation rate of 2.5×10^{-8} per site per year. This rate was determined by comparing net levels of interclade divergence for ND2 and control region data. The average divergence level based on the control region data was 95% of that for ND2, hence we used the lower estimate for ND2 of 2.5% per million years (Echelle et al., 2005) as a control region mutation rate. All estimates are averages from three independent runs of 10 short chains of 1000 steps and 2 long chains of 100,000 steps, using Watterson's (1975) estimate of θ as a starting value.

3. Results

3.1. Data characteristics

Partial sequences of the mitochondrial control region were obtained from 365 individual *C. variegatus* for this study. The control region alignment was 479 basepairs and corresponds to the region between tRNA-proline and the first central conserved block identified in other fish species (Lee et al., 1995). Sixty-nine distinct haplotypes were recognized. Partial sequences of the mitochondrial ND2 gene were obtained from 66 individuals, each representing a distinct control region haplotype. These sequences were combined with 52 *Cyprinodon* ND2 sequences from previous studies for a total of 118 sequences of 655 basepairs, which constitutes the region of overlap between the sequences obtained for this study and those from previous studies (Echelle et al., 2005, 2006). The composite control region/ND2 alignment used for phylogenetic analysis included the 66 composite haplotypes obtained for this study plus 60 distinct haplotypes from previous studies (Echelle et al., 2005, 2006). Ten of the 126 sequences in this alignment were drawn from Clade 2B of Echelle et al. (2005) and were designated as the outgroup. Clade 2B contains species from the southern Great Plains and northern Chihuahuan Desert and appears as a sister taxon to the maritime clade of *Cyprinodon* in the earlier phylogenetic analysis. This alignment was 1526 bp, of which 1030 bp overlapped between sequences obtained for previous studies and those from the current study. This alignment had 365 variable sites, of which 266 were parsimony informative.

3.2. Phylogenetic analysis

Analysis of the composite haplotype dataset produced a topology broadly similar to that in Echelle et al. (2006). At the deepest level within the maritime clade of *Cyprinodon* identified by Echelle et al. (2005) there are three clades, each supported by a posterior probability of 1.00, whose relationship to each other is unresolved (Fig. 2).

Clade 1 contains all populations of *C. variegatus* sampled from South Carolina south to the Florida Keys, and west to Texas in the Gulf of Mexico, and from the Caribbean islands of San Salvador and Andros in the Bahamas. It further includes all haplotypes from *C. higuey*, and a subset of the haplotypes sampled from the undescribed species from Lago Enriquillo in the Dominican Republic, *C. dearborni*, and *C. tularosa*, which is the earliest branching lineage within this clade. With broader sampling of northern populations, it is apparent that Clade 2 of Echelle et al. (2006) consists of two distinct monophyletic clades (Fig. 2: Clade 2, Clade 3), one containing populations of *C.v. ovinus* and the second containing all haplotypes sampled from *C. bondi* and *C. nichollsi*, and including a subset of haplotypes from the undescribed species from Lago Enriquillo. Support for the sister relationship of these two clades is weak, as it was in Echelle et al. (2006). Within *C.v. ovinus*, two divergent haplotypes found in South Carolina form a highly supported monophyletic group.

Substantial phylogenetic structure occurs within Clade 1. At the deepest level subsequent to the divergence of *C. tularosa*, there are five clades (1A–E), each supported by a posterior probability of 1.00, whose relationship to each other is unresolved, forming a polytomy. The first (Clade 1A) contains all *C. variegatus* haplotypes from South Carolina, Georgia and coastal Florida including the northern Florida Keys (WK) and two haplotypes from the southern Keys. There is also strong support within this clade for a grouping that ranges from Northern Florida to South Carolina (GA/N.FL *C. variegatus*: posterior probability = 1.00). The second clade (Clade 1B) contains southern Florida Keys (SK) and Andros Island, Bahamas (AN) *C. variegatus*. A monophyletic grouping of the Andros Island haplotypes is also strongly supported. The third clade (Clade 1C) includes all haplotypes from individuals referred to *C. variegatus* from Mississippi to Texas along the Gulf Coast of the United States (LA/MS/TX). Clade 1D is found on San Salvador Island, Bahamas and the Dominican Republic and includes *C. higuey*. Clade 1E contains *C.v. riverendi* from Grand Cayman Island in the Caribbean and a subset of haplotypes from *C. sp.* from Lago Enriquillo, and also has a posterior probability of 1.00.

3.3. Timing of major phylogenetic disjunctions

We compared net average levels of divergence for the ND2 dataset among highly supported clades at the two

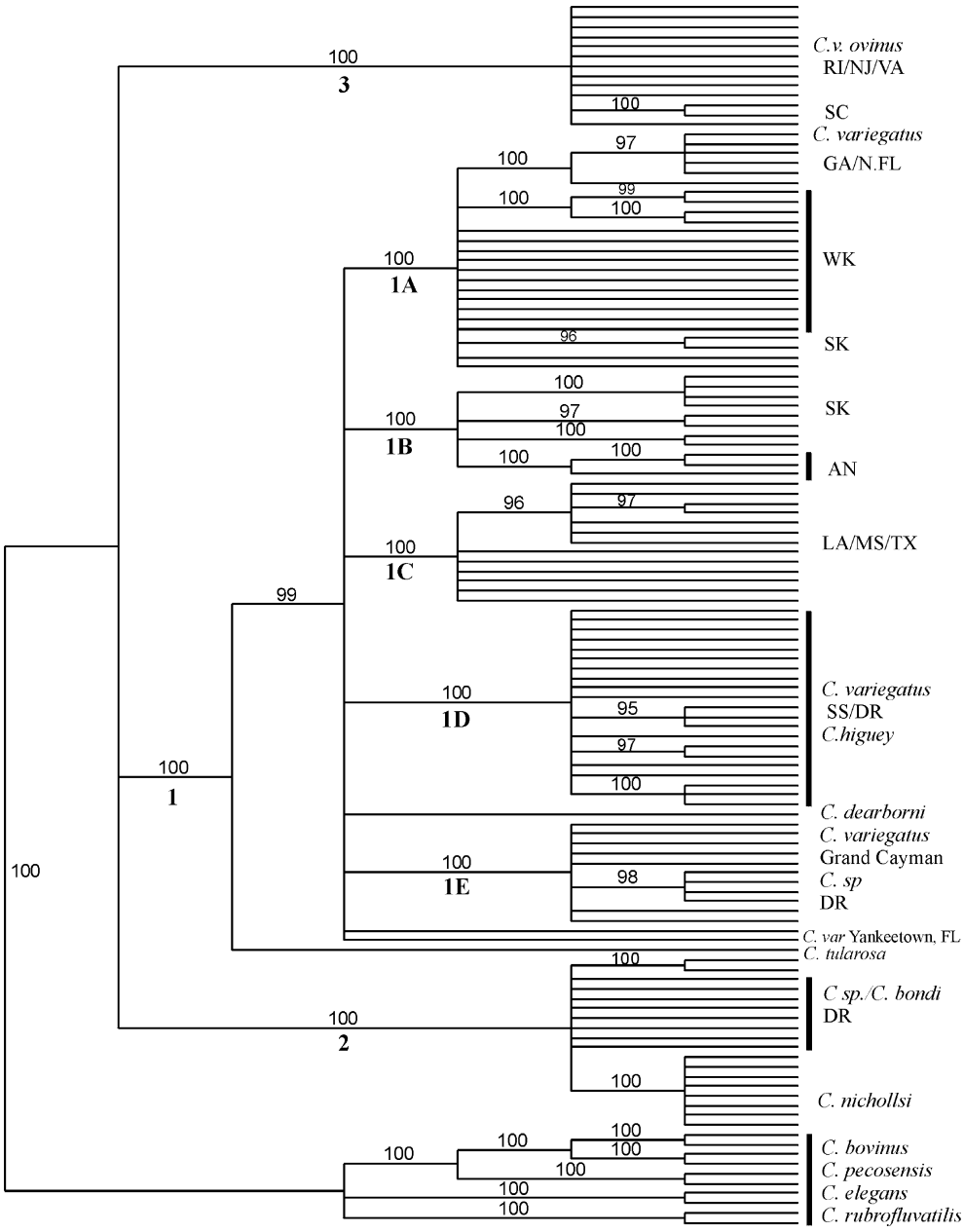

Fig. 2. Majority-rule consensus of 5000 trees sampled after sampling had reached a likelihood plateau during Bayesian analysis. Individual sample labels were omitted for clarity. Species names for a given clade are in italics, some clades contain multiple species. In non-italicized capital letters are site designations for samples taken for this study, full site names are given in Table 1. Numerical values above branches are posterior probability values from Bayesian analysis, only values above 95% are shown. Alphanumeric combinations below branches refer to clades discussed in the text.

deepest nested levels of the recovered tree topology. Average net divergence varies from 1.0% to 5.6% (Table 3). The Bayes factor was 2.0, indicating only very weak support for a non-clock model relative to the molecular clock alternative per Kass and Raftery (1995). Under a clock calibration of 2.5% divergence per million years developed from the closely related genus *Aphanius* (Echelle et al., 2005), levels of divergence observed among clades correspond to divergence times ranging from an approximate minimum of 400,000 to 2.24 million years before present. There are two periods, the first from approximately 2.2–1.7 million years before present which led to the emergence of clades 1, 2 and 3, and the second from 760,000 to 400,000 years before present (clades 1A–E) which seem to contain all of the deeper level clade divergence times within the maritime clade (Table 2). Dates for these two major clade divergence events based on distances from an ultrametric tree obtained by non-parametric rate smoothing are approximately 1.8 and 1.0 million years before present based on a 2.5% clock calibration, or 1.5 million and 800 thousand years before present under a 3.1% per million year clock.

Table 2
Net pairwise population divergence and estimated times of divergence based on 655 bp of ND2 sequence for major clades (see Fig. 2) within the maritime clade of *Cyprinodon*

Clade	1	2	3
1	0	1.8–2.2	1.6–2.0
2	0.056	0	1.3–1.7
3	0.049	0.042	0

Clade	1A	1B	1C	1D	1E
1A	0	0.45–0.56	0.32–0.40	0.35–0.44	0.42–0.52
1B	0.014	0	0.55–0.68	0.61–0.76	0.39–0.48
1C	0.010	0.017	0	0.39–0.48	0.52–0.64
1D	0.011	0.019	0.012	0	0.55–0.68
1E	0.013	0.012	0.016	0.017	0

Below diagonal are Tamura–Nei distances assuming gamma-distributed rate variation among sites with $\alpha = 0.5$. Above diagonal are a range of times for population divergence in millions of years before present assuming a molecular clock calibrated at 2.5% or 3.1% per million years.

3.4. Gene flow and population structure within clades

To more closely examine genetic structuring on a shallower timescale among sites in *C. variegatus*, we employed population genetic estimators of genetic differentiation within highly supported clades for which we had population samples from multiple sites (Table 3).

3.4.1. Clade 1A

As two haplotypes from Summerland Key occur in this clade it was included in gene flow calculations. Significant differentiation is apparent in all comparisons with the exception of that between South Carolina and Georgia (~290 km), with $F_{st} = 0.04$ and S_{nn} not significant ($p = 0.10$), even with the divergent northern clade haplotypes found in South Carolina included in the comparison. Aside from single copies of each of these two haplotypes, all remaining sequences from South Carolina are identical to the most common Georgia haplotype.

Table 3
Estimates of genetic differentiation among sites sampled within clades (see Fig. 2) for the current study

	SC	GA	WK	SK
Clade 1A				
SC	0	0.54	0.98***	1.00***
GA	0.03	0	0.98***	1.00***
WK	0.53	0.62	0	0.93***
SK	0.68	0.77	0.54	0

	MA	RI	CT	WMNJ	TTNJ	VA
Clade 3						
MA	0	0.96***	0.49	0.49	0.51	0.55**
RI	0.95	0	0.89***	0.89***	1.00***	1.00***
CT	0.01	0.84	0	0.58	0.55	0.58*
WMNJ	0.00	0.75	−0.03	0	0.49	0.58**
TTNJ	0.01	0.80	0.05	0.001	0	0.56**
VA	0.06	0.63	0.06	0.05	0.03	0

F_{st} is shown below the diagonal, while S_{nn} values (Hudson, 2000) along with statistical significance assessed by permutation test in DNAsp v. 3.99 (Rozas and Rozas, 1999) are displayed above the diagonal. SC = South Carolina, GA = Georgia, WK = Windley Key, SK = Summerland Key, MA = Massachusetts, RI = Rhode Island, CT = Connecticut, WMNJ = West Mantoloking, New Jersey, TTNJ = Tuckerton, New Jersey, VA = Virginia.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

3.4.2. Clade 1B

Between Summerland Key, Florida and Andros Island, $F_{st} = 0.6054$ and S_{nn} is highly significant ($p = 0.000$). There are no shared haplotypes between these locations, although sorting to reciprocal monophyly has not occurred.

3.4.3. Clade 1C

Between Texas and Louisiana (~720 km), $F_{st} = 0.08$ and S_{nn} is not significant ($p = 0.10$). As in the Georgia/South Carolina comparison, there is a single haplotype that is most common in both regions, but unlike the South Carolina site, the Texas sample also contains three unique haplotypes each differing from the most common by a single mutational difference and is not appreciably less diverse than the Lousiana sample, given it's small sample size.

3.4.4. Clade 3

The sample from Barrington, Rhode Island is genetically distinct. In all comparisons involving this site, F_{st} was large and S_{nn} was highly significant. Although S_{nn} was significant for all comparisons between Virginia and non-Rhode Island sites, it did not remain significant after Bonferroni correction for multiple tests ($p = 0.0033$) for the VA–TTNJ and VA–CT comparisons, but was close to this conservative level of significance for the VA–WMNJ ($p = 0.006$) and VA–MA ($p = 0.009$) comparisons.

3.5. Genetic diversity and historical demography

We used the phylogenetic analysis to identify two clades that include multiple populations of individuals assigned solely to *C. variegatus* with different latitudinal ranges, the first from approximately 24.92N to 33.23N, the second from 33.23N to 41.75N. These two clades were then analyzed using population genetic techniques to take advantage of the information in haplotype frequency distributions in order to investigate recent historical patterns of range shifts and population growth related to the most recent glaciation.

3.5.1. Clade 1 A

Haplotype diversity, nucleotide diversity and the average number of pairwise differences show extensive variability across sites with Windley Key highly variable, while GA and SC are sequentially less so, leading to a cline in genetic diversity along the southeast Atlantic coast (Table 4; Fig. 3B). The values for the South Carolina site are inflated by the presence of rare haplotypes assigned by phylogenetic analysis to the northern clade that is presumed to be a

Table 4
Diversity estimates for *C. variegatus* populations within clades 1A and 3

Population	*N*	*h*	π	*k*
Clade 1A				
WK	33	0.862	0.00767	3.523
GA	24	0.236	0.00051	0.243
SC	31	0.127[a]	0.00369[a]	1.748[a]
Clade 3				
VA	23	0.628	0.0025	1.13
TTNJ	23	0.320	0.00108	0.514
WMNJ	23	0.383	0.00108	0.498
CT	20	0.268	0.00056	0.268
MA	22	0.091	0.00020	0.091
RI	30	0	0	0

h = haplotype diversity, π = nucleotide diversity, k = average number of pairwise differences. Populations are ordered by latitude, from south to north.

[a] values resulting from inclusion of *C. v. ovinus* haplotypes from this site in analysis. In the absence of these two haplotypes, the values in each column become 0.

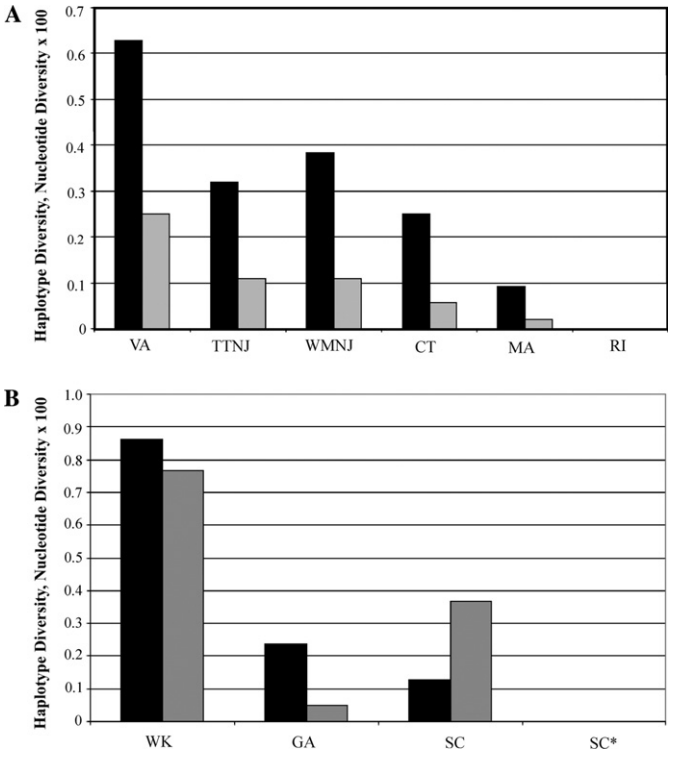

Fig. 3. Patterns of haplotype and nucleotide diversity among sites within clades of *C. variegatus* from the northeast (Graph A) and southeast (Graph B) Atlantic coasts of the Unites States. Populations are arranged left to right in order of latitude. Haplotype diversity is represented by solid bars, nucleotide diversity by striped bars. Nucleotide diversity values are multiplied by 100 to provide a similar scale for comparison of the two sets of values. SC* represents the diversity values for the South Carolina site in the absence of *C. v. ovinus* haplotypes.

separate subspecies, *C.v. ovinus*. If these haplotypes are removed, the South Carolina site has no genetic diversity as it contains only the most common haplotype found in the more diverse Georgia site to the south. We then calculated Fu's F_s to determine whether the observed distribution of polymorphism within populations is consistent with an equilibrium neutral model of molecular evolution. A significant value of a test statistic indicates deviation from this model. This statistic was marginally significant for the Georgia population, and not significant for the Windley Key population (Table 5). Maximum-likelihood estimates of the population growth parameter g show weak positive growth in the Windley key population, but strong positive growth in the Georgia population. This indicates a time to 1% of current population size within approximately the past 20,000 years in Georgia. As the South Carolina site had no diversity after removal of *C.v. ovinus* haplotypes, it could not be analyzed.

3.5.2. Clade 3

Haplotype diversity, nucleotide diversity and the average number of pairwise differences vary strongly among northern sites. In addition, a clinal decrease in genetic diversity is evident in the mitochondrial DNA moving northward from Virginia (Table 4; Fig. 3). The decrease in diversity with latitude is due largely to the different frequencies of two major haplotypes in each sample (Fig. 4). Virginia haplotype 1 (VA1) is the most common haplotype not only at the Virginia site, but also at both New Jersey sites, and in Connecticut and in Massachusetts. It is not found at the Rhode Island site. Several other low frequency haplotypes are found in Virginia and this population is thus relatively diverse. The two New Jersey samples contain low frequency private haplotypes in addition to the common Virginia haplotype, and at the northern New Jersey site the Rhode Island haplotype is present at low frequency. In Connecticut, of 20 haplotypes sampled, 17 were the most common Virginia haplotype, while 3 were the Rhode Island haplotype. In West Falmouth, MA, 21 of the sampled haplotypes were VA1, while one was the Rhode Island haplotype.

When sites are considered separately, Fu's F_s is generally negative with the exception of the Connecticut population,

Table 5
Test statistics for equilibrium-neutral model and maximum-likelihood estimates of growth parameter g (Kuhner et al., 1998) for populations, with estimated time to 1% of current population size calculated from exponential growth equation with mutation rate of 2.5×10^{-8} per site per year and generation time of one year

Population	*N*	F_s	*g*	SD	$t_{1\%}$
Clade 1A					
GA	24	−1.41*	8264.55	1128.01	22 (20–26)
WK	33	−3.91	330.64	88.98	557 (439–762)
Clade 3					
VA	23	−4.07***	3694.99	518.74	50 (44–58)
WMNJ	23	−1.44	5031.60	1449.08	37 (28–51)
TTNJ	23	−1.37	7060.72	415.25	26 (25–28)
CT	20	0.38	4451.33	2868.35	41 (31–116)
MA	22	−0.96	−20.65	2.71	NA

Estimates of g are averages from three independent runs.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

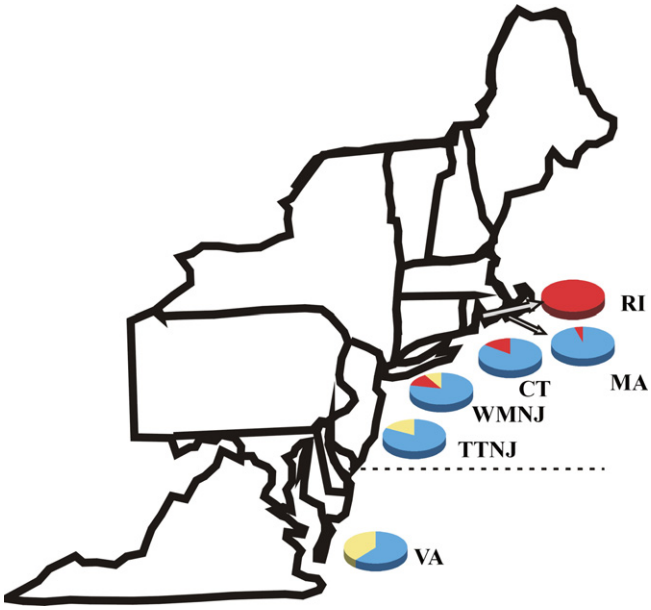

Fig. 4. Haplotype frequencies in populations of the northern subspecies C.v. ovinus. Light blue indicates the proportion of haplotype VA1, the most common haplotype found in the most southerly *C.v. ovinus* population sampled, while red indicates the most common haplotype in the Rhode Island population. Yellow indicates the proportion of other low frequency haplotypes at each site. The dashed line indicates the position of the Cape May biogeographic boundary for estuarine inverebrates (Engle and Summers, 1999).

but is significant only for the Virginia population. Maximum-likelihood estimates of the population growth parameter g indicate strong population growth for all three sites that occupy locations south of the maximum glacial advance, and for Connecticut, although the estimate of g has a large standard deviation (Table 5). West Falmouth, MA has a slightly negative growth rate. Rhode Island could not be analyzed due to lack of diversity. Times to 1% of current population size range from approximately 25-116 thousand years before present.

4. Discussion

The broad sampling of coastal/mainland populations of *C. variegatus* for this study indicates a strong geographic component to the distribution of distinct lineages, and the depth of several of the genetic breaks observed implicates Pleistocene climate change as an initiator of divergence. Although, as is often the case, the limitations of sampling make exact placement of boundaries among clades and barriers to gene flow difficult, it is clear from both the phylogenetic analysis and from population genetic estimators that gene flow across areas of coastline that contain major biogeographic boundaries is limited, as is it is in numerous other species. For example, populations from northern Florida to South Carolina are clearly isolated from those in southern Florida (Fig. 2; Table 3). Between these regions lies a major biogeographic boundary near Cape Canaveral, Florida (Briggs, 1995). There is also strong genetic differentiation between the northern and southern Florida Keys populations of *C. variegatus* (sites WK and SK on map: Windley and Summerland Keys). Although these sites are only approximately 90 km apart, the southern boundary of the Biscayne Bay estuarine biogeographic province of Engle and Summers (1999) lies close to the Windley Key site. A recent genetic study involving the scorched mussel, *Brachidontes exustus*, also found a distinct Key Biscayne mitochondrial lineage that was not found in the southern Florida Keys (Lee and O' Foighil, 2004) suggesting that factors involved in the establishment and maintenance of the southern boundary of the Key Biscayne province may be acting to promote genetic differentiation in coastal organisms. Although both the southeastern Atlantic and Gulf coasts were considered to be part of the Carolinian biogeographic province by Briggs (1995), numerous coastal organisms show a phylogenetic disjunction between these regions (Avise, 2000). This is also true for *C. variegatus*. In the northern subspecies biogeographic boundaries may be playing a role in genetic differentiation although the pattern is less clear. Gene flow would certainly appear to be limited across the region containing the Cape Hatteras boundary. It is possible that this boundary is implicated in the relatively deep phylogenetic divergence between *C.v. ovinus* and *C.v. variegatus*, but that subsequent distributional shifts have obscured this pattern. Further, the two *C.v. ovinus* haplotypes found only in South Carolina are divergent from those found in the next closest northern site (VA), indicating limited gene flow across the region containing the boundary at Cape Hatteras within *C.v. ovinus*. As a possible region of admixture between the two putative coastal subspecies, the North Carolina coast deserves a focused sampling effort in the future. The distribution of genetic variation may also be influenced by the presence of a biogeographic boundary near Cape May, New Jersey uncovered in an analysis of community composition and species range limits in estuarine benthic invertebrates (Engle and Summers, 1999), as there is some indication of limited gene flow between Virginia and sites north of this boundary (Table 3).

As in many other coastal organisms, the extensive population divergence within *C. variegatus* would thus appear to be controlled at least in part by large-scale factors involved in the creation and maintenance of biogeographic regions. Within these regions two comparisons among sites (GA vs. SC, LA vs. TX) showed nonsignificant differentiation and substantial gene flow (although this pattern may be influenced by a lack of equilibrium between migration and drift due to range expansion: see below) over a scale of hundreds of kilometers. This suggests that other factors, such as a low dispersal life-history (Daiber, 1982), and the discrete nature of estuarine habitats, do not necessarily lead to population differentiation in this species in the absence of known major barriers to gene flow, although more detailed sampling within regions is necessary to establish a firmer result.

Interestingly, in *Fundulus heteroclitus*, another eastern United States estuarine fish species with a range and

habitat that substantially overlaps that of *C. variegatus*, the pattern of genetic differentiation in mitochondrial DNA is different, indicating that idiosyncratic aspects of species biology or chance events can play a role in differentiation, even among species with seemingly similar life-histories. In *F. heteroclitus*, there appear to be two largely allopatric clades that correspond to previously described subspecies with a zone of admixture along the New Jersey coast (Smith et al., 1998), suggesting long-term persistence and isolation of northern populations.

4.1. Island colonization

Echelle et al. (2006) identified two major colonization events involving both major lineages found in coastal populations of *C. variegatus*, although support for the sister relationship of *C.v. ovinus* and endemic species of the Dominican Republic was weak, and with broader sampling in our study, this relationship is not well-supported. We have discovered another, more recent island colonization event. Haplotypes found on Andros island are nested within a clade that occurs in the southern Florida Keys (SK), and gene flow estimates show strong genetic differentiation among these sites, although sorting to reciprocal monophyly has not yet occurred.

The results indicate that while factors contributing to the establishment of biogeographic boundaries may have acted to initiate isolation between coastal/mainland populations, colonization of Caribbean islands from the mainland has provided another mechanism for diversification within the maritime clade of *Cyprinodon*. This recurrent phenomenon is also linked to Pleistocene climate change via changes in sea level. Lowered sea level during glacial periods may have facilitated colonization of Caribbean habitats by coastal *C. variegatus* (Echelle et al., 2006), which in turn has led to further isolation and divergence of populations during periods of higher sea-level.

4.2. Taxonomic implications

An early treatment (Hubbs, 1936) suggested the presence of northern and southern subspecies based on morphology, with the northern subspecies designated *C.v. ovinus*. This was later marginally supported by allozyme and mitochondrial sequence data (Darling, 1976; Finne, 2001). This distinction is strongly supported by mitochondrial sequence data in the current study, with broader geographic sampling than in previous studies. However, the southern boundary for the northern clade/subspecies would appear to be in South Carolina rather than North Carolina.

Differentiation occurs within subspecies and several monophyletic clades occur within *C. variegatus*, with *C. variegatus* and other species paraphyletic with respect to each other (Echelle et al., 2006; this study). Although mitochondrial sequence data alone cannot resolve the taxonomic status of these divisions, it is clear that the evolutionary history of *C. variegatus* and related species is not in accord with currently recognized taxonomy.

It is further unclear what role mating incompatibilities have played in the differentiation of these distinct lineages. Turner and Liu (1977) found little evidence of postzygotic isolation in no-choice breeding experiments among *Cyprinodon* species, but knowledge of prezygotic mechanisms is lacking. Within *C. variegatus*, nuclear gene evidence from allozyme data (Darling, 1976), supports allele frequency divergence only between broad scale northern and southern subspecies. Satellite DNA sequences were divergent among all populations sampled except for a shared allele between Cape Cod, MA and Sapelo, GA (Elder and Turner, 1994). Additional sequence data from independent nuclear genes is clearly necessary to better delineate the extent of nuclear divergence among mitochondrial clades and to determine whether species designation might be warranted for any of the groups discovered by analysis of mitochondrial data.

4.3. The role of historical climate change in the distribution of genetic variation

In addition to playing a historical role in initiating genetic divergence in coastal organisms, glaciation has been perhaps the most important extrinsic factor shaping the distribution of genetic variation in high latitude organisms via changes in distribution and local extinction/recolonization events (Hewitt, 2000). In *C. variegatus*, monophyletic clades representing different subspecies occur in regions that were in part formerly glaciated or were far south of the maximum glacial advance. This provides us an opportunity to study how recent glacial advances have differentially affected patterns of genetic variation in closely related clades that differ in their latitudinal position and to gain insight into the mechanisms by which glaciation influences population dynamics. The expectation is that high latitude populations would be more strongly affected by local extinction and recolonization due to loss of habitat in glaciated areas and reduction in population size through unfavorable climatic conditions.

Telltale remnants of such a process are a cline in genetic diversity northwards, the presence in formerly glaciated regions of a subset of southern diversity, and signatures of recent population growth. In fact, evidence for recolonization and growth are found in northern populations that constitute clade 3 of *C. variegatus*. First, a clear decrease in genetic diversity occurs northward from Virginia through Cape Cod (Fig. 3), and northern haplotypes are a limited subset of those found in more diverse populations to the south. The two most common haplotypes in a more diverse population in West Mantoloking, New Jersey are found in similar proportions in Connecticut and Massachusetts. Two other low frequency haplotypes from this population are not found north of the former glacial margin. This pattern is a clear signature of a recent range expansion into formerly glaciated areas from the south (Fig. 4), likely northern New Jersey, as the second haplotype in formerly

glaciated sites is not found in southern New Jersey or Virginia. This second, low-frequency haplotype is the only haplotype found in thirty individuals sampled at a site in Barrington, Rhode Island in Narragansett Bay. This may represent a chance founder event during colonization of *C.v. ovinus* up Narragansett Bay, or may be due to selection by environmental parameters that vary up-estuary, such as salinity. Alternatively, it could be an artifact due to sampling bias, although this site was sampled in the same manner as all other sites, making this scenario unlikely.

The origin of New Jersey populations may also be due to an older recolonization event from further south that occurred prior to the latest post-glacial recolonization. A decline in genetic diversity is also clear from Virginia to New Jersey, and the most common Virginia haplotype is also the most common haplotype in New Jersey. However, some evidence of genetic differentiation occurs between Virginia and New Jersey, and New Jersey contains private alleles. This suggests a two-step recolonization model, with range expansion into New Jersey from the south in the early stages of glacial withdrawal as habitat conditions became favorable for persistence of *C.v. ovinus* populations in areas proximal to the glacial boundary. Subsequent recolonization of more northern habitats recently occurred from northern New Jersey. Population genetic tests of neutrality indicate a significant lack of mutation-drift equilibrium only within Virginia via neutrality test, but maximum likelihood methods consistently indicate population growth with the exception of the West Falmouth, MA sample. While dates vary somewhat among populations, all dates to 1% of current population size indicate an onset of growth prior to the last glacial maximum at approximately 20,000 years before present and ensuing glacial retreat. It may be the case that populations south of the maximum glacial extent were growing throughout the most recent glacial advance, but this seems unlikely for populations within the glacial margin. If polymorphism existed in such populations prior to expansion, this could cause the estimated date to be older than the actual expansion, as may be the case for the Connecticut population, which likely contained two colonizing haplotypes at establishment.

What is also clear from this study is that Pleistocene climate change can affect the population dynamics of coastal organisms far from the glacial margin. A signature of a recent population/range expansion exists in the southeast, south of where we expect that the northern subspecies persisted throughout recent Pleistocene glacial cycles, suggesting a possible difference in environmental optima among subspecies/clades, or nearly complete habitat loss in the area around South Carolina. As in the north, a north/south gradient in genetic diversity exists, and the South Carolina population contains only the most common haplotype from a more diverse population to the south in Georgia. Again, this pattern is complicated somewhat by admixture in South Carolina with clade 3 haplotypes (*C.v. ovinus*) that likely represent a divergent population of this subspecies that is distributed in the unsampled region between Virginia and South Carolina. Population genetic tests suggest strong recent population growth in the coastal southeast (GA population) that occurs after the most recent glacial maximum. This pattern differs from that in allozyme data, where no cline in diversity was evident from northern Florida to South Carolina (Darling, 1976). This suggests the possibility that forces affecting allozyme loci in northern and southern regions may differ. The population in the northern Florida Keys is strongly differentiated from others in the same monophyletic clade. It is highly diverse, does not show deviation from mutation drift equilibrium, and shows weakly positive growth. This population has likely been fairly stable for some time, and indicates that some parts of the range of *C. variegatus* may not have been strongly affected by late Pleistocene glacial oscillations.

The results of this study show that a common vertebrate constituent of tidal marsh communities, *Cyprinodon variegatus*, is strongly genetically differentiated across its range, and that diversification has been a common theme throughout the Pleistocene. More recently, post-glacial range expansion and population growth has left its mark on the genetic variation of this species, in both the northern and southern parts of its range. The spatial pattern of variation from the only other marsh fish for which similar data is available, *F. heteroclitus*, appears to differ, and northern populations do not seem to be the result of recent recolonization (Adams et al., 2006). Further comparative data is needed to reveal underlying commonalties in response to gene flow barriers and climate change in tidal marsh communities, and also the role of chance or species specific traits that act to produce idiosyncratic responses.

Acknowledgments

The authors would like to acknowledge NSF (DEB 0108500 and 0343464 to DMR) and a NOAA-NERRS graduate research fellowship to RAH for funding.

References

Adams, S.M., Lindmeier, J.B., Duvernell, D.D., 2006. Microstaellite analysis of the phylogeography, Pleistocene history and secondary contact hypotheses for the killifish, *Fundulus heteroclitus*. Mol. Ecol. 15, 1109–1123.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., Saunders, N.C., 1987a. Intraspecific phylogeography: the mitochondrial bridge between population genetics and systematics. Ann. Rev. Ecol. Syst. 18, 489–522.

Avise, J.C., Reeb, C.A., Saunders, N.C., 1987b. Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachiodidae). Evolution 41, 991–1002.

Avise, J.C., 2000. Phylogeography. Harvard University Press, Cambridge, MA.

Bernardi, G., Sordino, P., Powers, D.A., 1993. Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish *Fundulus heteroclitus*. Proc. Natl. Acad. Sci. USA 90, 9271–9274.

Briggs, J.C., 1974. Marine Zoogeography. McGraw-Hill, New York.

Briggs, J.C., 1995. Global Biogeography. Elsevier, Amsterdam.

Caudill, C.C., Bucklin, A., 2004. Molecular phylogeography and evolutionary history of the estuarine copepod, *Acartia tonsa*, on the Northwest Atlantic coast. Hydrobiologia 511, 91–102.

Collin, R., 2001. The effects of mode of development on phylogeography and population structure of North Atlantic *Crepidula* (Gastropoda: Calyptraeidae). Mol. Ecol. 10, 2249–2262.

Daiber, F.C., 1982. Animals of the Tidal Marsh. Van Nostrand Reinhold Company, New York, NY.

Darling, J.D.S., 1976. Electrophoretic variation in *Cyprinodon variegatus* and systematics of some fishes of the subfamily Cyprinodontinae. Unpubl. M.S. thesis, Yale University, New Haven, Connecticut.

Echelle, A.A., Carson, E.W., Echelle, A.F., Van Den Bussche, R.A., Dowling, T.E., Meyer, A., 2005. Historical biogeography of the New-World pupfish genus *Cyprinodon* (Teleostei: Cyprinodontidae). Copeia 2005, 320–339.

Echelle, A.A., Fuselier, L., Van Den Bussche, R.A., Rodriguez, C.M.L., Smith, M.L., 2006. Molecular systematics of Hispaniolan pupfishes (Cyprinodontidae: *Cyprinodon*): implications for the biogeography of insular Caribbean fishes. Mol. Phylogenet. Evol. 39, 855–864.

Elder, J.F., Turner, B.J., 1994. Concerted evolution at the population level: pupfish *hind*III satellite DNA sequences. Proc. Natl. Acad. Sci. USA 91, 994–998.

Engle, V.D., Summers, J.K., 1999. Latitudinal gradients in benthic community composition in Western Atlantic estuaries. J. Biogeogr. 26, 1007–1023.

Finne, K.L., 2001. Phylogeographic structure of the Atlantic pupfish, *Cyprinodon variegatus* (Cyprinodontidae), along the eastern coast of North America. Unpubl. M.S. thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 147, 915–925.

Gonzalez-Villasenor, L.I., Powers, D.A., 1990. Mitochondrial-DNA restriction-site polymorphisms in the teleost *Fundulus heteroclitus* support secondary intergradation. Evolution 44, 27–37.

Gurgel, C.F.D., Fredericq, S., Norris, J.N., 2004. Phylogeography of *Gracilaria tikvahiae* (Gracilariaceae, Rhodophyta): a study of genetic discontinuity in a continuously distributed species based on molecular evidence. J. Phycol. 40, 748–758.

Hall, T., 1999. Bioedit sequence alignment editor. Department of Microbiology, North Carolina State University. URL http://www.mbio.ncsu.edu/BioEdit/bioedit.html.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature. 405, 907–913.

Holtmeier, C.L., 2001. Heterochrony, maternal effects and phenotypic variation among sympatric pupfishes. Evolution 55, 330–338.

Hrbek, T., Meyer, A., 2003. Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). J. Evol. Biol. 16, 17–36.

Hubbs, C.L., 1936. Fishes of the Yucatan Peninsula. Carnegie Institution of Washington Publication 457, 157–287.

Hudson, R.R., 2000. A new statistic for detecting genetic differentiation. Genetics. 155, 2011–2014.

Hudson, R.R., Slatkin, M., Maddison, W.P., 1992. Estimation of levels of gene flow from DNA sequence data. Genetics. 132, 583–589.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. Jour. Am. Stat. Assoc. 90, 773–795.

Kelly, D.W., MacIsaac, H.J., Heath, D.D., 2006. Vicariance and dispersal effects on phylogeographic structure and speciation in a widespread estuarine invertebrate. Evolution 60, 257–267.

Kuhner, M.K., Yamato, J., Felsenstein, J., 1998. Maximum likelihood estimation of population growth rates based on the coalescent. Genetics. 149, 429–434.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics 17, 1244–1245.

Lee, W., Conroy, J., Howell, W.H., Kocher, T.D., 1995. Structure and evolution of teleost mitochondrial control region. J. Mol. Evol. 41, 54–66.

Lee, T., O' Foighil, D., 2004. Hidden Floridian biodiversity: mitochondrial and nuclear gene trees reveal four cryptic species within the scorched mussel, *Brachidontes exustus*, species complex. Mol. Ecol. 13, 3527–3542.

McMillen-Jackson, A.L., Bert, T.M., 2003. Disparate patterns of population genetic structure and population history in two sympatric penaeid shrimp species (*Farfantepenaeus aztecus* and *Litopenaeus setiferus*) in the eastern United States. Mol. Ecol. 12, 2895–2905.

Miya, M., Nishida, M., 1999. Organization of the mitochondrial genome of a deep-sea fish *Gonostoma gracile* (Teleostei: Stomiiformes): the first example of transfer RNA gene rearrangements in bony fishes. Mar. Biotechnol. 1, 416–426.

Perdices, A., Carmona, J.A., Fernandez-Delgado, C., Doadrio, I., 2001. Nuclear and mitochondrial data reveal high genetic divergence among Atlantic and Mediterranean populations of the Iberian killifish *Aphanius iberus* (Teleostei: Cyprinodontidae). Heredity. 87, 314–324.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Ramos-Onsins, S.E., Rozas, J., 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19, 2092–2100.

Reeb, C.A., Avise, J.C., 1990. A genetic discontinuity in a continuously distributed species: mitochondrial DNA in the American Oyster, *Crassostrea virginica*. Genetics. 124, 397–406.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rozas, J., Rozas, R., 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15, 174–175.

Sanderson, M.J., 1997. nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14, 1218–1231.

Smith, M.W., Chapman, R.W., Powers, D.A., 1998. Mitochondrial DNA analysis of Atlantic coast, Chesapeake Bay, and Delaware Bay populations of the teleost *Fundulus heteroclitus* indicates temporally unstable distributions over geologic time. Mol. Mar. Biol. Biotechnol. 7, 79–87.

Swofford, D.L., 1998. PAUP* Phylogenetic analysis using parsimony (*and other methods). Version 4.04b10. Sinauer Associates, Sunderland, Massachusetts.

Turner, B.J., Liu, R.K., 1977. Extensive interspecific genetic compatibility in the New World killifish genus *Cyprinodon*. Copeia 1977, 259–269.

Wares, J.P., Cunningham, C.W., 2001. Phylogeography and historical ecology of the North Atlantic intertidal. Evolution 55, 2455–2469.

Watterson, G.A., 1975. On the number of segregating sites in genetical models without recombination. Theor. Pop. Biol. 7, 256–276.

Weinberg, J.R., Dahlgren, T.G., Trowbridge, N., Halanych, K.M., 2003. Genetic differences within and between species of deep-sea crabs (*Chaceon*) from the North Atlantic Ocean. Biol. Bull. 204, 318–326.

Young, A.M., Torres, C., Mack, J.E., Cunningham, C.W., 2002. Morphological and genetic evidence for vicariance and refugium in Atlantic and Gulf of Mexico populations of the hermit crab *Pagurus longicarpus*. Mar. Biol. 140, 1059–1066.