

NUCLEAR AND MITOCHONDRIAL SEQUENCE DATA REVEAL AND CONCEAL DIFFERENT DEMOGRAPHIC HISTORIES AND POPULATION GENETIC PROCESSES IN CARIBBEAN REEF FISHES

Ron I. Eytan^{1,2} and Michael E. Hellberg¹

¹Department of Biological Sciences, 107 Life Sciences Building, Louisiana State University, Baton Rouge, Louisiana 70803

²E-mail: ron.eytan@gmail.com

Received March 18, 2010

Accepted June 10, 2010

Mitochondrial and nuclear sequence data should recover historical demographic events at different temporal scales due to differences in their effective population sizes and substitution rates. This expectation was tested for two closely related coral reef fish, the tube blennies *Acanthemblemaria aspera* and *A. spinosa*. These two have similar life histories and dispersal potentials, and co-occur throughout the Caribbean. Sequence data for one mitochondrial and two nuclear markers were collected for 168 individuals across the species' Caribbean ranges. Although both species shared a similar pattern of genetic subdivision, *A. spinosa* had 20–25 times greater nucleotide sequence divergence among populations than *A. aspera* at all three markers. Substitution rates estimated using a relaxed clock approach revealed that mitochondrial *COI* is evolving at 11.2% pairwise sequence divergence per million years. This rapid mitochondrial rate had obscured the signal of old population expansions for both species, which were only recovered using the more slowly evolving nuclear markers. However, the rapid *COI* rate allowed the recovery of a recent expansion in *A. aspera* corresponding to a period of increased habitat availability. Only by combining both nuclear and mitochondrial data were we able to recover the complex demographic history of these fish.

KEY WORDS: *Acanthemblemaria*, Caribbean, reef fish, historical demography, population genetics, substitution rates.

The broad aim of comparative phylogeography is to infer how co-distributed taxa respond to shared evolutionary events (Avice 2000; Hickerson et al. 2009). Each species is treated as a replicate sample of the underlying processes responsible for observed genetic patterns. Several well-documented geological processes have produced concordant patterns of genetic structure and historical demography, including Pleistocene glaciation in Europe (Taberlet et al. 1998; Hewitt 2000) and northwestern North America (Brunsfeld et al. 2001; Carstens et al. 2005), the closure of the Central American seaway (Knowlton et al. 1993; Hickerson et al. 2006; Lessios 2008), and the rise of the Andes (Burney and Brumfield 2009). However, congruent population genetic patterns

may arise due to different processes occurring at different times (i.e., pseudocongruence; Cunningham and Collins 1994) and similar patterns of subdivision may not accurately reflect a shared history of co-occurring taxa.

Just as multiple co-occurring species may afford replication for inferring common historical events acting in a region, multiple genetic markers allow replicate samples of the demographic history of particular species (Brito and Edwards 2009). By combining markers, researchers make the tacit assumption that those markers are behaving in a similar fashion. However, loci, like taxa, may conflict. A major source of this conflict is the inherent stochasticity in the time of lineage sorting for each marker (Hudson and

Turelli 2003). Different estimates of demographic history can also result when markers differ from one another in the mechanisms affecting their evolution, such as mode of transmission, effective population size, or rates of recombination (Graur and Li 2000; Hare 2001; Zhang and Hewitt 2003; Brito and Edwards 2009).

The most common example of this in animal studies is the difference between nuclear and mitochondrial sequence markers. The former is transmitted biparentally and interlocus recombination should mean that most nuclear markers provide replicate estimates of a common demography, whereas the latter is transmitted maternally as a single nonrecombining block. Given the power afforded by the small effective population size of mitochondrial DNA (Moore 1995) and the expense and effort required to survey multiple nuclear markers, the argument has been made that mitochondrial DNA offers more than enough power to address most questions (Barrowclough and Zink 2009). However, as cost concerns recede with the advent of next-generation sequencing (Hudson 2008), concerns about factors that could confound inferences provided by mitochondrial DNA (e.g., non-neutrality, extreme rate variation, recombination) have led some (Galtier et al. 2009) to suggest that mitochondrial DNA “is the worst marker” for population genetics and should not be used at all.

In theory, mitochondrial and nuclear DNA should be able to complement each other in demographic studies. The smaller effective population size of mitochondrial DNA should allow it to capture the signal of demographic events that cannot leave their footprints on the larger effective population size of nuclear markers. The strength of nuclear DNA lies in its ability to provide replicate samples of the underlying demographic history affecting the genome of an organism as well as replicate samples of the coalescent process. For this reason, sampling multiple nuclear markers can substantially reduce the variance of parameter estimates (Felsenstein 2006; Carling and Brumfield 2007; Lee and Edwards 2008; Brito and Edwards 2009; Hey 2010). Discussions of the relative merits of mitochondrial and nuclear markers in phylogeography have centered on two different, but complementary, goals: to identify clades of populations (and thus phylogeographic breaks or cryptic species) using gene trees, and to reconstruct historical demography (Zink and Barrowclough 2008; Barrowclough and Zink 2009; Edwards and Bensch 2009). Although mitochondrial DNA is useful in delimiting geographically restricted clades, its power to estimate demographic parameters on its own is poor. The opposite is true of nuclear DNA. Combining both marker types allows investigators to identify clades and then estimate parameters of interest, such as migration rates (Lee and Edwards 2008; Barrowclough and Zink 2009). However, simple combination does not admit the possibility that mitochondrial and nuclear DNA can reveal different demographic events. In practice, such a dual gene class approach to inferring historical

demography requires robust substitution rate estimates for both types of markers to reconcile them into a single time frame. Incorrect estimates of substitution rates can severely bias parameter estimates such as divergence times, population size changes, and migration, among others. Here, we explore congruence of marker types, phylogeographic patterns, and demographic inferences between co-occurring taxa from a genus of reef fish that contains sister taxa to either side of the Isthmus of Panama, allowing the calibration of taxon-specific substitution rates.

Acanthemblemaria is a genus of blennies occurring on both sides of the Isthmus of Panama and throughout tropical and subtropical waters of the western Atlantic and eastern Pacific. They are members of the Family Chaenopsidae, one of only two coral reef fish families with an exclusively Neotropical distribution (Stephens 1963). The Western Atlantic members of the genus occur throughout the Caribbean basin, the Bahamas, and peninsular Florida (Smith-Vaniz and Palacio 1974). All members in the genus are small (~1.2–3.5 cm SL) and are obligate dwellers of vacated invertebrate holes on shallow (<1 – ~22 m) rocky and coral reefs (Böhlke and Chaplin 1993; Clarke 1994).

Acanthemblemaria aspera and *A. spinosa* are found throughout the Caribbean and the Bahamas and co-occur over large portions of their respective ranges (Fig. 1). The two species are closely related (Hastings 1990 and Eytan et al., unpubl. data), share the same mating system (male resource defense polygyny; Hastings 2002), pelagic larval duration (21–24 days; Johnson and Brothers 1989), and ecologically overlap (Clarke 1994). The two species differ in microhabitat use, though. *Acanthemblemaria spinosa* lives in shelters in live and standing dead coral off the reef surface (high-profile habitat). *Acanthemblemaria aspera* is found at the base of standing dead corals or in coral rubble (low-profile habitat), sometimes at the base of corals housing *A. spinosa* (Clarke 1994).

The differences in microhabitat use and specialization give the two species different propensities to go locally extinct. Clarke (1996) found that *A. spinosa* populations in St. Croix went locally extinct due to habitat degradation, specifically the destruction of standing dead *Acropora palmata* corals. The resulting coral rubble, however, allowed *A. aspera* populations to increase in size. This type of population dynamic may lead to discordant demographic cycles, both between species and between populations. We hypothesize that if the two species do differ in historical demography, despite similar life histories, the contrast in microhabitat requirements for the two blenny species should favor population persistence in *A. aspera* rather than *A. spinosa*, as the latter is more prone to local extinction. However, it is not clear whether the population dynamics observed at ecological time scales will extend to evolutionary ones.

We collected mitochondrial and nuclear sequence data for both species from populations where they co-occur and analyzed

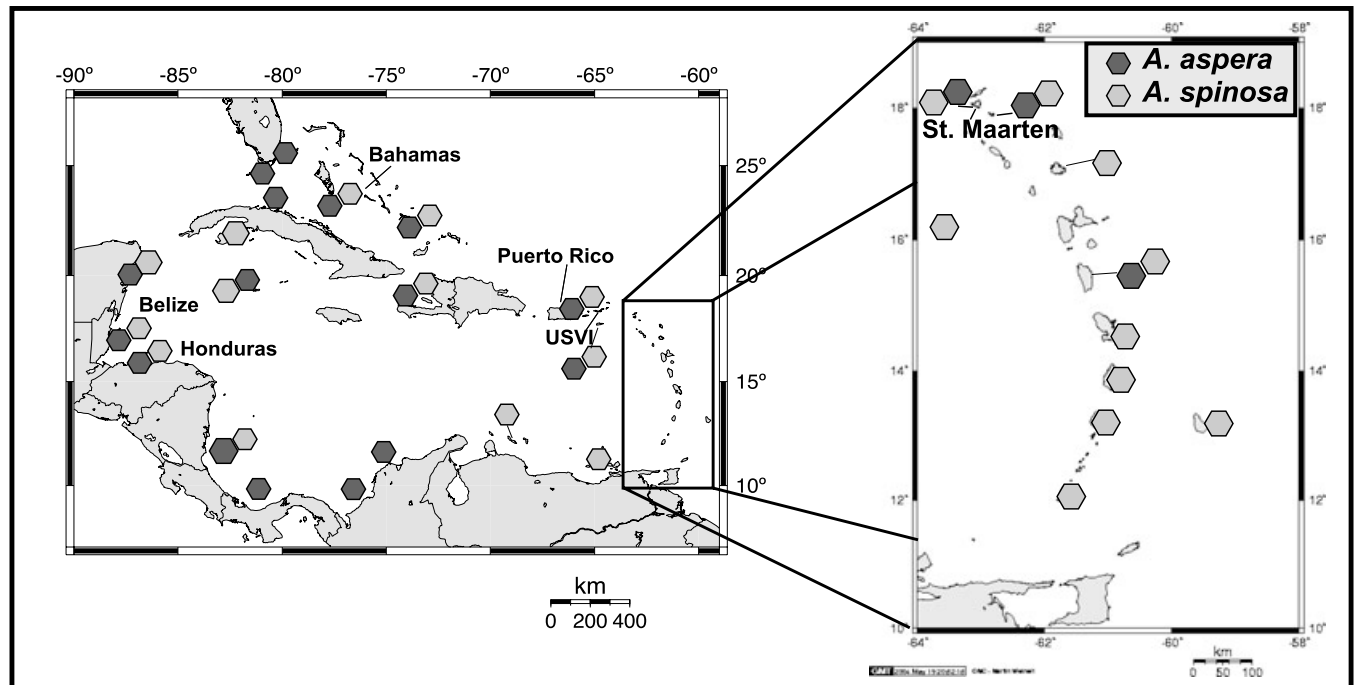


Figure 1. Confirmed locations of *A. aspera* and *A. spinosa* populations in the Caribbean (Smith-Vaniz and Palacio 1974 and RIE, pers. obs.) The localities sampled for this study are listed.

the data with substitution rates calculated using a relaxed molecular clock. We then tested the expectation of phylogeographic and demographic concordance among co-distributed taxa. *Acanthemblemaria aspera* and *A. spinosa* are good candidates for this test because of their similar life histories, close relationship, and nearly identical geographic distributions, but potential differences in historical demography due to different microhabitat use.

Materials and Methods

SAMPLE COLLECTION

Only individuals from populations in which both species co-occur were used in the current study (Fig. 1 and Supporting Information). Samples from species used as outgroups for analysis of substitution rates were collected from Panama and Belize (Supporting Information). Samples were collected on SCUBA. A dilute solution of quinaldine sulfate was squirted into the blenny hole. A small glass vial was then placed over the hole, into which the fish immediately swam. Photo vouchers from freshly collected specimens are available from RIE for a subset of these individuals by request. Whole fish were stored individually in 95% ethanol at -80°C .

DNA EXTRACTION, PCR, AND SEQUENCING

DNA was extracted from 11–16 individuals of each species from each of six populations using the Qiagen (Valencia, CA) QIAMP DNA Minikit. DNA was extracted from a total of 84 individuals

from each species, as well as a single individual of *A. paula*, *A. betinensis*, and *A. exilispinus*.

The polymerase chain reaction (PCR) was performed on a PTC-100 or 200 (MJ Research, Waltham, MA) to amplify three genetic markers: protein-coding genes mitochondrial cytochrome oxidase I (*COI*) and nuclear recombination-activating gene 1 (*rag1*), and intron V from nuclear alpha-tropomyosin (*atrop*). The primers used, primer references, and PCR conditions can be found in the Supporting Information. Amplicons were purified with a Strataprep PCR Purification Kit (Stratagene, La Jolla, CA) or directly sequenced without cleanup in both directions on an ABI 3100 or 3130 XL automated sequencer using 1/8 reactions of BigDye Terminators (Version 3.1, Applied Biosystems) and the amplification primers.

ALIGNMENT AND PHASING

Sequencing reactions for *rag1* and *COI* produced clear reads 539 and 704 bp long, respectively. Reads for *atrop* were 440 bp, 427 bp, 423 bp, and 418 bp long in *A. spinosa*, *A. paula*, *A. betinensis*, and *A. exilispinus*, respectively. The *atrop* sequences for *A. aspera* had length variants (425, 427, and 429 bp) as well as indel-hets, with the latter obscuring reads. Indel heterozygotes containing a single indel were resolved using CHAMPURU (Flot 2007). Sequences containing more than one indel-het were cloned to resolve constituent allelic sequences using the Invitrogen TOPO TA Cloning Kit for Sequencing. The initial direct sequences were always used in determining allelic sequences from cloned DNA

to avoid scoring any changes that resulted from errors introduced by the PCR.

The *COI* and *rag1* sequences contained no gaps and were aligned using MUSCLE (Edgar 2004), implemented in Geneious version 4.5.4 (Drummond et al. 2009), as were the intraspecific *atrop* datasets. The interspecific *atrop* alignment, consisting of one sequence each from *A. spinosa*, *A. paula*, *A. betinensis*, *A. exilispinus*, and *A. aspera*, was more difficult to align due to length polymorphisms. BALi-Phy version 2.0.1 (Suchard and Redelings 2006) was used to align the *atrop* sequences using the HKY substitution model, gamma-distributed rate variation, and the default indel model. BALi-Phy was run four times to ensure concordance among runs. The final output from each run was separately analyzed, with all the samples before convergence were discarded as burnin. The consensus alignment from the run with the highest posterior probability was used for the subsequent analyses.

Alleles from sequences with multiple heterozygous single nucleotide polymorphisms (SNPs) were resolved using PHASE version 2.1.1 (Stephens et al. 2001; Stephens and Donnelly 2003; Stephens and Scheet 2005). Input files were prepared for PHASE using the online software package SeqPHASE (Flot 2010). Alleles determined by cloning heterozygotes or as output from CHAMPURU were used to create a “known” file for PHASE. A default probability threshold of 0.9 was used for all runs.

After initial PHASE runs, all datasets contained some individuals with unresolved SNPs. We cloned a subset of these individuals to directly determine their haplotype phase. The direct haplotype observations were then added to the “known” file and the datasets were reanalyzed. Final datasets for each gene and species contained no more than three individuals for which the phase of a single SNP was not resolved to 0.9 (six total for *rag1*, four for *atrop*). After alignment and phasing of heterozygous SNPs, the final dataset contained 336 nuclear and 84 mitochondrial alleles for each species. The sequences have been submitted to GenBank with the accession numbers HM196865-HM197713.

HAPLOTYPE NETWORKS

Parsimony networks were constructed for *COI* for both species using TCS version 1.21 (Clement et al. 2000). Networks for the *COI* gene in *A. spinosa* failed to connect at the 95% confidence level so the connection limit was fixed at 70 to connect populations. Although this reduces the confidence of the resulting networks, the main purpose is to visualize the number of inferred mutations between populations, rather than infer relationships among populations.

RECOMBINATION

We tested for recombination using the tree-based SBP and GARD methods of Kosakovsky-Pond et al. (Pond et al. 2006) implemented online via the Datamonkey webserver (Pond and Frost

2005) using the model of sequence evolution chosen by Datamonkey using the AIC criterion, with gamma-distributed rate variation.

The GARD and SBP tests failed to find recombination in any of the pooled population datasets for any of the markers of either species. Recombination was detected via GARD and SBP for one population level dataset—the *rag1* alignment for Belize and Honduras in *A. spinosa*. Recombination was detected with a *P* value of 0.1 at position 192. We performed all analyses involving the Belize and Honduras *rag1* alignment twice, both for the whole alignment and for positions 1–192. The results of those analyses did not differ from those using the full alignment (not shown).

POPULATION STRUCTURE

Sequences were collapsed into haplotypes using FaBox, (Villesen 2007) and treated as alleles based solely on identity, not the genetic distance between the haplotypes. Measures of genetic subdivision, as measured by pairwise Φ_{ST} values among populations, were calculated using analysis of molecular variance (AMOVA) (Excoffier et al. 1992; Michalakis and Excoffier 1996), implemented in GenoDive version 2.0b15.1 (Meirmans and Van Tien-deren 2004), which for this purpose are equivalent to Weir and Cockerham's θ (Weir and Cockerham 1984). To allow meaningful comparisons between species and markers and to correct for high levels of variation within populations, which necessarily reduce measures of the proportion of variation partitioned among populations (Hedrick 2005), the Φ_{ST} measures were estimated using a standardizing procedure (Meirmans 2006) implemented by GenoDive. The significance of all comparisons was tested by permutation.

To detect differentiated populations (*k*) without the need to define populations a priori, and to determine if *A. aspera* and *A. spinosa* have congruent patterns of genetic differentiation, a Bayesian clustering analysis was implemented in STRUCTURE version 2.3 (Pritchard et al. 2000) using the admixture model with uncorrelated allele frequencies. A recent extension to the STRUCTURE method (Hubisz et al. 2009), which uses sampling locality as a prior, was employed. The use of this prior does not tend to find population structure when none is present (Hubisz et al. 2009), but has been recommended for situations (like ours) where available data are limited (Pritchard et al. 2009).

We performed 10 replicate runs for *k* values between 1 (no population differentiation) and 6 (the maximum number of populations sampled). Each replicate was run for 10 million iterations following an initial burnin of 100,000 iterations. Best estimates of *k* were inferred using the method of Evanno et al. (2005) as implemented in STRUCTURE Harvester (Earl 2009). The output files for the best estimate of *k* were then processed in CLUMMP (Jakobsson and Rosenberg 2007) using the default parameters.

The Evanno et al. (2005) analysis identified two clusters for both species, but the initial bar plots from these runs showed four clusters (not shown). The Evanno et al. method for determining k is biased toward detecting the highest level of population structure in datasets with hierarchical population structure (Waples and Gaggiotti 2006). This appeared to be the case here, as the initial $k = 2$ corresponds to a split between populations in the eastern and western Caribbean (data not shown). STRUCTURE was run again to determine if additional structure was present in the two identified clusters. The dataset for each species was split to represent membership in the two detected clusters and each was analyzed with the original run conditions, with the exception that $\log \Pr(X | K)$ was estimated for $k = 1-3$.

PAIRWISE SEQUENCE DIVERGENCE

The net average pairwise sequence divergence between each pair of populations for each species and each gene was calculated in MEGA version 4.0 (Tamura et al. 2007) using the model of sequence evolution selected by the AIC in jModelTest (Posada 2008). The models selected by the AIC for each marker and analysis can be found in the Supporting Information. If the chosen model was not available in MEGA, the next less-complex model was used. To aid in direct comparisons, all models of sequence evolution for each marker were the same for both species, using the less-complex model of sequence evolution when these differed.

GENE TREE CONSTRUCTION

Gene trees for each species for each of the three markers were constructed in BEAST version 1.4.8 (Drummond and Rambaut 2007) using a constant size coalescent prior and a strict molecular clock. These were constructed for use in the generalized mixed Yule coalescent (GMYC) analysis (see below), which requires rooted ultrametric gene trees. The BEAST analyses do not require the inclusion of an outgroup to root the tree because a molecular clock is enforced (Huelsenbeck et al. 2002; Drummond and Rambaut 2007).

MCMC analyses were run four times for either 10,000,000 steps (*COI* data) or 25,000,000 steps (nuclear markers), sampling every 1000 steps for all for a total of 10,000 and 25,000 trees, respectively. Convergence onto the posterior distribution was assessed using two methods. The first was by visual inspection of traces in Tracer version 1.4.1 to check for concordance between runs. All parameters had effective sample size (ESS) values > 250 . The second method for assessing convergence was by using AWTY (Nylander et al. 2008). Cumulative posterior probability plots were inferred using the “cumulative” function. Posterior probability estimates for each clade were compared between the four Markov chain Monte Carlo (MCMC) runs by producing scatter plots using the “compare” function. One-fourth of the sampled trees were discarded as burnin for each test. The cumulative poste-

rior probability plots indicated that all runs had stabilized whereas the scatter plots showed that posterior probability for clades were similar for all compared runs. The maximum clade credibility (MCC) tree was calculated for each gene tree using TreeAnnotator version 1.4.8 with a burnin of 6250 for the two nuclear markers and 2500 for *COI*.

SUBSTITUTION RATE AND DIVERGENCE

TIME ESTIMATES

We estimated substitution rates and divergence times using a relaxed clock approach that allows rates to vary among species and divergence times to follow a probabilistic distribution rather than relying on point estimates. The use of a relaxed molecular clock is not appropriate for the analysis of intraspecific data, as the coalescent employs a strict clock (Hein et al. 2005). Any differences in branch lengths in a coalescent tree would only be caused by the variance of the Poisson process describing the number of mutations along a branch and not variation in substitution rates among lineages (Hein et al. 2005). Once all gene copies have coalesced, this restriction is relaxed and rates can vary among lineages.

We determined whether all gene copies had coalesced in individual populations by using the GMYC model (Pons et al. 2006). The GMYC analysis divides a single locus gene tree into a portion in which a Yule speciation process affects the branch lengths and a portion in which there is a shift to a coalescent branching process, using this boundary to define species. Here, it was used to detect the presence of a transition point from a coalescent to Yule process to determine the appropriate use of a relaxed molecular clock. If gene copies were found to belong to distinct clusters then representative alleles from each cluster was used, whereas if no clustering was found then a single representative allele for each species was used. The MCC trees inferred in BEAST were used for the GMYC model, which was implemented in R (R Development Core Team 2009) using the SPLITS package (available at CRAN repository) and the single threshold model.

The substitution rates for the three genes were determined using a relaxed clock analysis in BEAST version 1.4.8. Based on the results of the GMYC analysis, a tree was built including a single representative nuclear gene sequence for *A. aspera*, *A. spinosa*, *A. paula*, *A. betinensis*, and *A. exilispinus* (except for *COI*, which had five *A. spinosa* sequences representing populations delimited by SPLITS as having coalesced). *Acanthemblemia paula* was included because it is hypothesized to be sister to *A. aspera* (Hastings 1990 and Eytan et al., unpubl. data) and could break a possible long branch leading to *A. aspera*. An exponential prior was placed on the time to most recent common ancestor (TMRCA) of the transisthmian geminate pair of *A. betinensis* and *A. exilispinus*. A mean of 7 million years (my) with a zero offset of 3.1 my was used for the exponential prior, which translates into a distribution with lower and upper 95% confidence intervals of

3.28 and 28.92 my, respectively. This prior distribution represents the latest possible divergence between the pair but also allows for the possibility that divergence may have occurred before the final closure of the isthmus, although with a decreasing probability further back in time. The uncorrelated lognormal relaxed clock (Drummond et al. 2006) was used to estimate branch rates. The nucleotide substitution model GTR + Γ was used for *COI* and HKY for the two nuclear markers.

Each dataset was also run using the Kimura 2 parameter (K2P) model so that rate estimates would be directly comparable to those for other transisthmian species pairs in Lessios (2008). Further analyses were also done using both the GTR + Γ or HKY model and the K2P model under fixed substitution rates to determine what inferred transisthmian divergence times would be assuming previously published rates of molecular evolution in geminate coral reef fishes. These rates were taken from Table 3 in Lessios (2008) for species pairs assumed to have begun divergence at the final closure of the isthmus. Upper and lower values, where present, were used. The two rates used for *COI* were 1.03% per million years and 1.77% per million years. The rate used for *rag1* was 0.097% per million years. No *atrop* data were included in Lessios (2008).

Two MCMC searches were conducted for each dataset with a Yule speciation prior on the gene tree for 10,000,000 (*COI*) or 25,000,000 (nuclear genes) generations. The log files from the runs were inspected using TRACER version 1.4.1 to check for convergence in the Markov chain. MCC trees were constructed using TreeAnnotator version 1.4.8. The mean rate for each marker and inferred transisthmian divergence time and their 95% upper and lower highest posterior densities, as provided in TRACER, was recorded.

DEMOGRAPHIC RECONSTRUCTION

To detect departures from a constant population size or neutrality, we used the summary statistics F_s (Fu 1997) and R_2 (Ramos-Onsins and Rozas 2002), which have the greatest power to reveal population growth (Ramos-Onsins and Rozas 2002). Large negative values of F_s and small positive values of R_2 indicate population growth. We also used Tajima's D (Tajima 1989), as it also has good power to detect population growth (Ramos-Onsins and Rozas 2002) but has the added benefit of being a two-tailed test. Significantly negative values of Tajima's D indicate population growth (or a selective sweep), whereas significantly positive values are a signature of genetic subdivision, population contraction, or diversifying selection. The F_s , R_2 , and Tajima's D tests were all implemented in DNAsp version 5.1 (Librado and Rozas 2009) for each gene, population, and all populations combined for each of the two species. The significance of all the tests were determined by 1000 coalescent simulations, also implemented in DNAsp.

We reconstructed the historical demography of each species by using the GMRF skyride plot (Minin et al. 2008), implemented in BEAST version 1.5.2. The GMRF skyride plot is a nonparametric analysis that uses the waiting time between coalescent events in a gene tree to estimate changes in effective population size over time. It differs from the related Bayesian skyline plot (Drummond et al. 2005) by not requiring the specification of a user-defined prior on the number of population size changes in the history of the sample.

GMRF skyride plots using time-aware smoothing were constructed for each population as demarcated by the STRUCTURE analyses, each gene, and all populations combined, for both species. Rates of molecular evolution for each gene were fixed at the values obtained in the substitution rate estimation analysis. Each dataset was run twice for 10 million generations, sampling every 1000, except for the all populations combined datasets, which were run for 100 million generations, sampling every 10,000. Output files were checked in Tracer and all ESS values were greater than 250. Bayesian skyride plots were then visualized in Tracer. Population size changes were deemed significant if the upper and lower 95% confidence intervals at the root of the plot did not overlap with those at the tips.

Results

SIMILAR PATTERNS OF POPULATION SUBDIVISION FOR *A. ASPERA* AND *A. SPINOSA*

Acanthemblemaria aspera and *A. spinosa* had largely congruent patterns of population subdivision. Pairwise Φ_{ST} values showed that both species share few, if any, *COI* alleles among populations, with most populations composed entirely of private alleles (except for Belize–Honduras and Puerto Rico–St. Thomas for *A. aspera* and St. Thomas–Puerto Rico for *A. spinosa*) (Table 1). All pairwise *COI* values were significant, except in the case of Puerto Rico–St. Thomas for *A. aspera*.

Although most *A. spinosa* populations did not share any nuclear alleles, all *A. aspera* populations did. The majority of pairwise comparisons for *A. spinosa* (9 out of 15) had a corrected Φ_{ST} of 1, but none did for *A. aspera*. The significance of pairwise Φ_{ST} values also differed between marker type and between species. Whereas all but one pairwise *COI* comparison was significant for *A. aspera*, four of 15 comparisons using nuclear markers were not significant at $P = 0.05$. This was in contrast to *A. spinosa*, where all nuclear Φ_{ST} values were significant save one, Belize–Honduras, which also lacked significance in *A. aspera*.

The STRUCTURE analyses (Fig. 2) confirmed the presence of hierarchical population structure and recovered the same four clusters for both species, corresponding to the Bahamas, Belize and Honduras, Puerto Rico and St. Thomas, and St. Maarten. For

Table 1. Corrected Φ_{ST} values for *A. aspera* (above diagonal) and *A. spinosa* (below diagonal). All Φ_{ST} values are significant at $P=0.05$ except where marked by an asterisk.

	Bahamas	Belize	Honduras	Puerto Rico	St. Thomas	St. Maarten
<i>COI</i>						
Bahamas	–	1	1	1	1	1
Belize	1	–	0.912	1	1	1
Honduras	1	1	–	1	1	1
Puerto Rico	1	1	1	–	0*	1
St. Thomas	1	1	1	0.682	–	1
St. Maarten	1	1	1	1	1	–
Nuclear						
Bahamas	–	0.44	0.309	0.361	0.617	0.687
Belize	0.776	–	0.072*	0.112*	0.212	0.172
Honduras	0.850	0.061941*	–	0.123*	0.498	0.378
Puerto Rico	1	1	1	–	0.234*	0.297
St. Thomas	1	1	1	0.181	–	0.209
St. Maarten	1	1	1	0.853	0.931	–
All Markers						
Bahamas	–	0.773	0.601	0.788	0.87	0.85
Belize	0.864	–	0.603	0.829	0.845	0.773
Honduras	0.901	0.516	–	0.769	0.862	0.764
Puerto Rico	1	1	1	–	0.089	0.832
St. Thomas	1	1	1	0.448	–	0.81
St. Maarten	1	1	1	0.950	0.972	–

each of the four reduced datasets, the Evanno et al. (2005) method selected $k = 2$. However, in the case of *A. aspera*, there was multimodality in the assignment of individuals from St. Maarten to clusters (see Fig. 2, where there is a probability of 0.6 and 0.4 for assignment of St. Maarten individuals to either the Puerto Rico–St. Thomas cluster or to a separate St. Maarten cluster, respectively). This suggested that k may equal 3 for *A. aspera*. However, in addition to the results from the Evanno et al. test, the values of $L(K)$ for the eastern Caribbean *A. aspera* dataset were at their lowest at $k = 1$ and highest at $k = 2$ (Fig. S1).

The *COI* haplotype networks also showed that geographic structuring of populations is largely congruent between the two species (Fig. 3). For both species, most populations were reciprocally monophyletic, with the exception of Belize and Honduras. There, *A. spinosa* did not share alleles among populations, but *A. aspera* does.

LARGE DIFFERENCES IN POPULATION GENETIC DIVERGENCE BETWEEN SPECIES

Although overall measures of genetic subdivision were similar for the two species, the degree of genetic divergence among populations was not: *A. spinosa* had many more inferred mutations between populations than *A. aspera* (Table 2, Fig. 3). For the two *A. aspera* populations that had the largest *COI* genetic distances

between them, Honduras and Puerto Rico/St. Thomas, pairwise sequence differences based on the number of inferred mutations in the haplotype network were 10 times greater for *A. spinosa COI* than for *A. aspera* (Fig. 3). The true ratio may be even greater than this conservative estimate, because the number of mutations inferred for *A. spinosa* using the TCS analysis did not account for the possibility of multiple mutations at the same site. Model-corrected estimates of net average pairwise sequence divergence between pairs of populations (Table 2) were significantly higher in *A. spinosa* than *A. aspera* for all markers (Mann–Whitney U-test $P \leq 0.05$ for each) (Table 2). The ratio of mean pairwise genetic distance in *A. spinosa* compared to *A. aspera* varied from 19.76 for *rag1* to 25.02 for *atop*, with *COI* at 22.57.

GENE TREES AND BRANCHING PROCESSES ARE DIFFERENT FOR *A. ASPERA* AND *A. SPINOSA*

The gene trees derived from the mitochondrial data recovered a pattern of subdivision similar to that in the STRUCTURE analyses, but the trees constructed from the nuclear markers did not (Fig. 4). The *A. spinosa COI* gene tree recovered five reciprocally monophyletic clades, with all but the Bahamas having good support (BPP > 0.9). Unlike the STRUCTURE results, Belize and Honduras were reciprocally monophyletic. The *A. aspera COI* gene tree supported all clusters recovered in the STRUCTURE

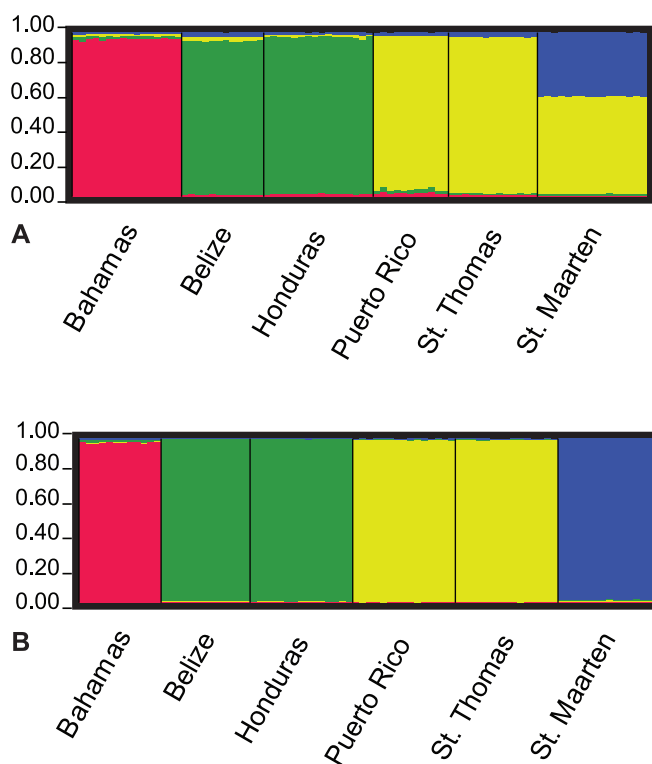


Figure 2. Graphical summary of the results from the STRUCTURE analysis for $k = 4$ for (A) *A. aspera* and (B) *A. spinosa*. Each individual is represented by a vertical line broken into four colored segments to represent the estimated proportions of that individual's genome originating from each of the four inferred clusters.

analysis (with the exception of the Bahamas) as monophyletic with good support.

The *A. spinosa atrop* tree recovered a well-supported western Caribbean clade, but eastern Caribbean populations were paraphyletic (Fig. 4). Neither the *A. spinosa rag1* tree nor either

of the nuclear DNA trees for *A. aspera* had any well-supported nodes that correspond to geography. Here, the eastern and western Caribbean clades are defined as being all populations to the east and west of the Mona Passage, respectively, as in Baums et al. (2005) and Taylor and Hellberg (2006), with the western Caribbean clade including the Bahamas.

The gene trees constructed for each marker for both species were used for the GMYC analyses. The results indicated that gene copies had not coalesced within separate populations in *A. aspera*; there was only one cluster present for all *A. aspera* markers (LRT $P > 0.05$). Based on those results, one representative *A. aspera* allele from each gene was used for the substitution rate estimates.

In *A. spinosa*, the *COI* gene copies coalesced within five of the six populations (Puerto Rico and St. Thomas gene copies coalesced together) (LRT $P = 0.0028$). This indicated that for *COI* there was a transition from a coalescent to a Yule branching process. For that reason, five representative *COI* sequences for *A. spinosa*, corresponding to the five delimited clusters from the GMYC analysis, were used for substitution rate estimates. The GMYC analysis did not indicate any population-level clustering for the *A. spinosa* nuclear genes (LRT $P > 0.05$), so one representative allele for each of the two genes were used for substitution rate estimates.

MITOCHONDRIAL SUBSTITUTION RATES ARE UP TO 37 \times FASTER THAN FOR NUCLEAR DNA

Substitution rate estimates revealed very rapid mitochondrial rates in *Acanthemblemaria*. Mitochondrial *COI* was 37.65 \times and 14.94 \times faster than *rag1* and *atrop*, respectively. The mean and 95% upper and lower highest posterior density (HPD) for substitution rates across all taxa in the analysis in substitutions/site/million years for *COI*, *atrop*, and *rag1* were 5.61×10^{-2} (1.63×10^{-2} ,

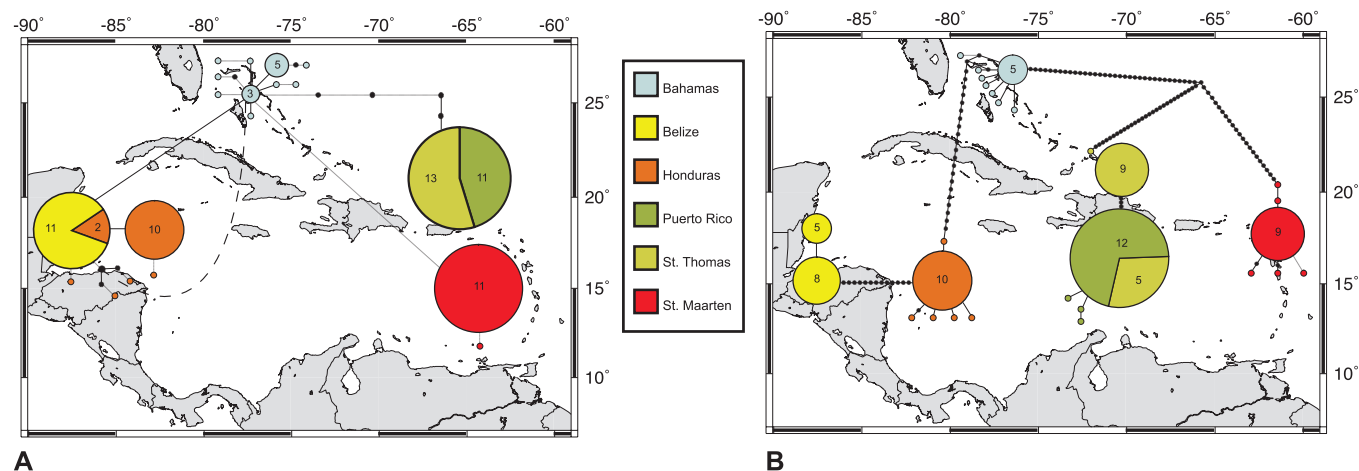


Figure 3. *COI* haplotype networks for (A) *A. aspera* and (B) *A. spinosa*. Circle color indicates population and size is proportional to the number of individuals sharing that haplotype. Haplotypes shared by >1 individual are marked with the sample size. Black dots on branches are inferred mutations.

Table 2. Model-corrected net average pairwise sequence divergence between each pair of populations for *A. aspera* (above diagonal) and *A. spinosa* (below diagonal).

	Bahamas	Belize	Honduras	Puerto Rico	St. Thomas	St. Maarten
<i>COI</i>						
Bahamas	–	0.0023	0.0031	0.0118	0.0118	0.0019
Belize	0.0548	–	0.0006	0.0138	0.0138	0.0034
Honduras	0.0558	0.0288	–	0.0157	0.0157	0.0038
Puerto Rico	0.2458	0.2388	0.259	–	0	0.0124
St. Thomas	0.2578	0.2552	0.2736	0.0029	–	0.0124
St. Maarten	0.2058	0.2486	0.2531	0.1935	0.1915	–
<i>ATROP</i>						
Bahamas	–	0.0012	0.0002	0.0012	0.0024	0.0039
Belize	0.0011	–	0.0004	0.0001	0.0001	0.0006
Honduras	0.0014	0	–	0.0004	0.0015	0.0024
Puerto Rico	0.0375	0.0393	0.04	–	0.0003	0.0009
St. Thomas	0.0386	0.0404	0.0411	0.001	–	0.0001
St. Maarten	0.0444	0.0467	0.0475	0.006	0.0087	–
<i>RAGI</i>						
Bahamas	–	0.0003	0.0002	0.0005	0.0005	0.0005
Belize	0.0091	–	0.0003	0.0004	0.0005	0.0004
Honduras	0.0103	0	–	0.0007	0.0007	0.0007
Puerto Rico	0.0083	0.0116	0.0126	–	0	0.0001
St. Thomas	0.0078	0.0092	0.0102	0.0003	–	0.0001
St. Maarten	0.0046	0.0126	0.0135	0.0033	0.0032	–

9.92×10^{-2} , 4.65×10^{-3} (4.69×10^{-4} , 1.02×10^{-2}), and 1.45×10^{-3} (2.12×10^{-4} , 2.87×10^{-3}), respectively. The posterior distribution for the mean substitution rate for *atrop* was right skewed so the median rate (3.99×10^{-3}) was used instead of the mean for all further analyses. These rates corresponded to values of 11.22%, 0.798%, and 0.298% sequence divergence per million years, respectively. The mean rate we have calculated for *COI* is over six times faster than the highest estimated *COI* substitution rate listed in Lessios (2008) (1.77% per million years) for geminate reef fish taxa assumed to have split at the final closure of the Central America Isthmus. For *ragI*, our mean rate is approximately three times faster than that in Lessios (2008) (0.097% per million years), although the lower bound of the posterior distribution of our estimate overlaps with that of Lessios.

TRANSISTHMIAN DIVERGENCE TIME ESTIMATES ARE CONCORDANT AMONG MARKERS AND ROBUST TO THE MODEL OF SEQUENCE EVOLUTION

Estimates of divergence time for the transisthmian species pair *A. betinensis* and *A. exilispinus* agreed among markers and revealed a split before the final closure of the Isthmus of Panama. The mean and 95% HPDs for dates across the isthmus (TMRCA for *A. betinensis* and *A. exilispinus*) were 4.63 my (3.10, 8.17) for *COI*, 4.62 my (3.10, 8.14) for *atrop*, and 4.67 my (3.10,

8.41) for *ragI*. These estimated values showed that the divergence time estimates did not simply return the calibration prior for the divergence date (7.0 my (3.28, 28.92)) specified in BEAST.

Published estimates for *COI* substitution rates from Lessios (2008) recovered significantly older dates when they were used to estimate divergence across the isthmus, but the use of the published *ragI* rates did not. For the slower *COI* rate from Lessios, 1.03% per million years, the mean and upper and lower 95% HPDs for inferred divergence times using the GTR + Γ model selected by jModelTest were 37.84 (25.94, 50.93) million years ago (mya). When the K2P model was used, dates of 20.31 (16.64, 23.93) mya were recovered. For the faster substitution rate, 1.77% per million years, divergence times when using GTR + Γ are 21.92 (15.08, 29.58) mya. The dates inferred when using K2P were 11.82 (9.72, 14.012) mya. When the *ragI* rate from Lessios (2008) (0.097% per million years) was used, and the K2P model, the inferred mean and upper and lower 95% divergence dates for *A. exilispinus* and *A. betinensis* were 7.25 (1.60, 14.29) mya.

Mean divergence time for *A. exilispinus* and *A. betinensis* did not differ from those inferred using the exponential calibration prior when the K2P model from Lessios (2008) was used, and neither did mean substitution rate estimates. For *COI* the mean, lower, and upper 95% HPDs for divergence time using the K2P

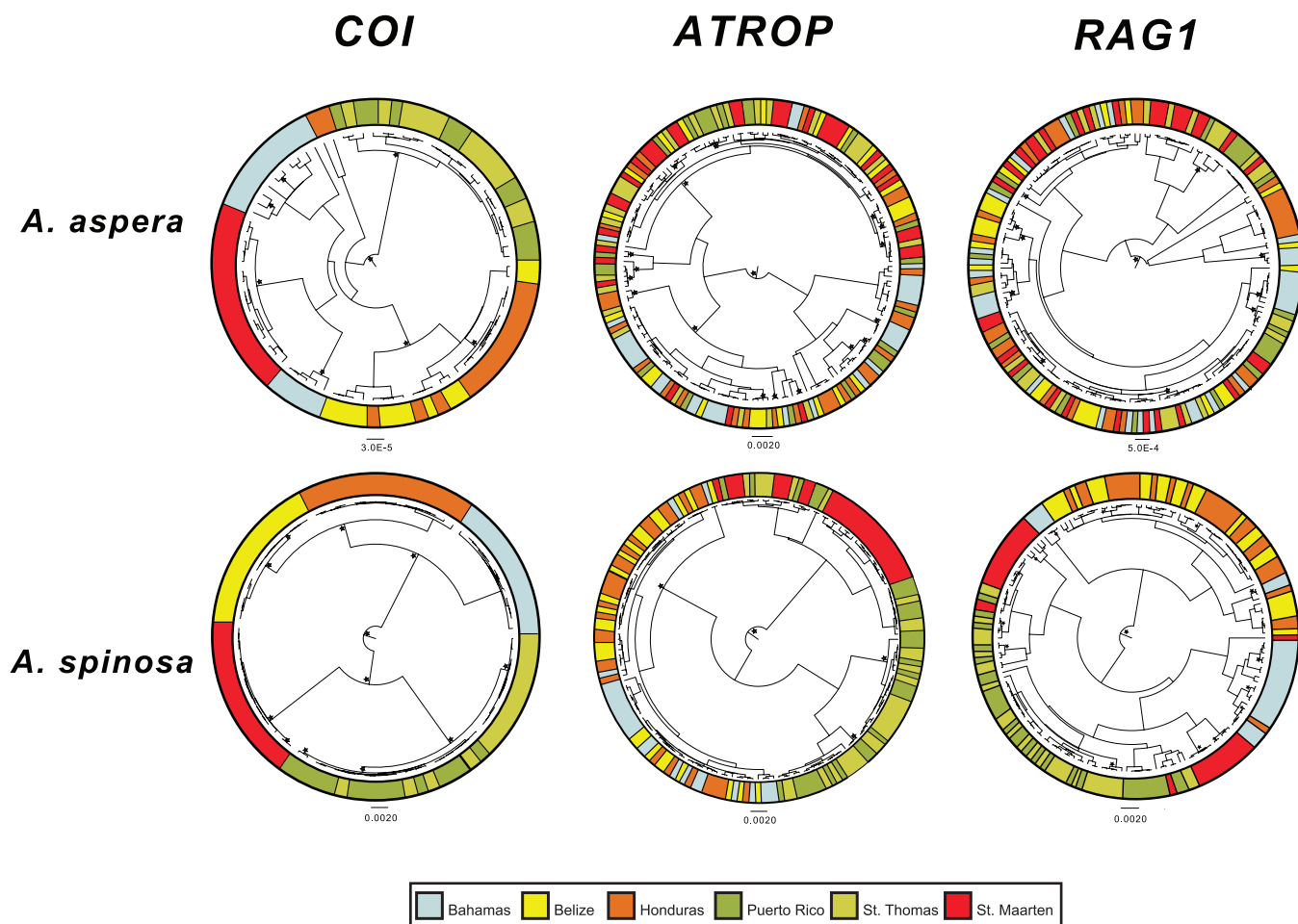


Figure 4. Gene trees for each marker for *A. aspera* and *A. spinosa* inferred in BEAST. Colors at tips of branches indicate population origin for each allele. Stars represent nodes with greater than 0.90 BPP.

model were 4.62 (3.1, 8.19) for *COI*, 4.64 (3.1, 8.21) for *atrop*, and 4.64 (3.1, 8.25) mya for *rag1*. The mean substitution rates using the K2P model for *COI*, *atrop*, and *rag1* were 2.72×10^{-2} (1.0×10^{-2} , 4.32×10^{-2}), 4.05×10^{-3} (5.45×10^{-4} , 1.05×10^{-2}), and 1.47×10^{-3} (2.45×10^{-4} , 2.87×10^{-3}) substitutions/site/million years, respectively.

DEMOGRAPHIC RECONSTRUCTION REVEALS MULTIPLE EXPANSIONS AND LARGE EFFECTIVE POPULATION SIZES

Reconstructions of historical demography revealed a recent population expansion in *A. aspera* and older expansions in both species. Test results for population size change in *A. aspera* based on summary statistics differed between markers (Table 3). For the combined datasets, the two nuclear markers showed a significant signal of population expansion, whereas the *COI* data did not. The results from individual *A. aspera* populations were also mixed. For *COI*, a strong signal of population expansion was only inferred for the Bahamas (although there was some support for an

expansion in Honduras). Likewise, the *atrop* data only recovered a strong signal of expansion in a single population, St. Maarten (Table 3). However, for *rag1*, expansions were detected in five of six individual populations, with the exception of the Bahamas.

In contrast to the results from the summary statistics, the Bayesian skyride tests did not recover significant size changes for individual *A. aspera* populations for any markers surveyed (Supporting Information). However, the skyride plots for the combined datasets recovered signals of significant size change for all three markers (Fig. 5), all in the form of population expansions.

The historical demography of *A. spinosa* inferred using frequency-based tests differed from that of *A. aspera*. Although there was little support from the *A. aspera* *COI* data for expansions in individual populations, four of six *A. spinosa* populations did show a signal of expansion (Table 3). However, like in *A. aspera*, the *atrop* data recovered a strong signal of expansion in only a single population (also in St. Maarten). In addition, two populations had significantly positive Tajima's *D* values (Puerto Rico and St. Thomas *A. spinosa* for *atrop*) and may have undergone

Table 3. Results of demographic tests using summary statistics. Significance was determined by the test itself (only in the case of Tajima's *D*) and/or through coalescent simulations. Significant test results are in bold. Values for *A. aspera* COI are not available for Puerto Rico or St. Thomas because there are no segregating sites in either population.

<i>A. aspera</i> COI	Tajima's <i>D</i>	Fu's <i>F_s</i>	<i>R₂</i>	<i>A. spinosa</i> COI	Tajima's <i>D</i>	Fu's <i>F_s</i>	<i>R₂</i>
Bahamas	-1.5367²	-5.836²	0.0779²	Bahamas	-2.01608¹²	-5.205²	0.0802²
Belize	-1.14053	-0.476	0.2764	Belize	1.30045	1.151	0.2564
Honduras	-1.40459	-1.172	0.0928²	Honduras	-1.98306¹²	-3.537²	0.1018²
Puerto Rico	n.a.	n.a.	n.a.	Puerto Rico	-1.31654	-1.797²	0.1206²
St. Thomas	n.a.	n.a.	n.a.	St. Thomas	1.09767	2.766	0.1969
St. Maarten	-1.16221	-0.700	0.2421	St. Maarten	-1.81096¹²	-2.609²	0.1058²
All populations	-0.89263	-4.872	0.0694	All populations	2.30115¹²	12.718	0.1860
<i>A. aspera</i> ATROP	Tajima's <i>D</i>	Fu's <i>F_s</i>	<i>R₂</i>	<i>A. spinosa</i> ATROP	Tajima's <i>D</i>	Fu's <i>F_s</i>	<i>R₂</i>
Bahamas	-0.89522	-1.672	0.0920	Bahamas	-1.73386²	-4.739²	0.0774²
Belize	-0.34394	-1.358	0.1162	Belize	-0.48367	-3.683	0.1104
Honduras	0.47714	-3.236	0.1388	Honduras	-0.68160	-1.044	0.1071
Puerto Rico	-0.30924	-2.436	0.1185	Puerto Rico	2.53075¹²	0.439	0.2290
St. Thomas	-0.64361	-0.878	0.1012	St. Thomas	1.92666²	-0.334	0.2030
St. Maarten	-1.31859	-2.915²	0.0774²	St. Maarten	0.13775	-1.506	0.1339
All populations	-1.18961	-18.410²	0.0526	All populations	1.07819	-3.683	0.1239
<i>A. aspera</i> RAG1	Tajima's <i>D</i>	Fu's <i>F_s</i>	<i>R₂</i>	<i>A. spinosa</i> RAG1	Tajima's <i>D</i>	Fu's <i>F_s</i>	<i>R₂</i>
Bahamas	-0.37933	-0.869	0.1074	Bahamas	0.74688	-7.156²	0.1625
Belize	-0.82382	-3.133²	0.0936	Belize	0.88035	1.166	0.1700
Honduras	-0.54769	-2.881²	0.1002	Honduras	-0.70511	-0.857	0.0986
Puerto Rico	-1.01239	-2.957²	0.0901	Puerto Rico	0.43585	-1.241	0.1424
St. Thomas	-0.31413	-3.349²	0.1154	St. Thomas	0.50888	-2.998²	0.1455
St. Maarten	-1.11489	-3.854²	0.0377	St. Maarten	0.97316	-5.221²	0.1633
All populations	-1.55085²	-14.866²	0.0377	All populations	0.79666	-25.786²	0.1123

population declines, although that signal was not found for the other two *A. spinosa* markers (Table 3).

Signals of population size changes were found in the combined *A. spinosa* datasets. In the case of the COI data, a significant signal of population decline was found. However, that result may stem from the deep divergence at COI among *A. spinosa* populations, which would cause significant positive values of Tajima's *D*. Of the two nuclear markers, only the *rag1* data showed a signal of expansion in *A. spinosa* (Table 3). A recent meta-analysis by Wares (2010) found that Tajima's *D* values calculated from mitochondrial sequence data in natural populations are biased toward a deviation from neutrality and toward negative values, which could give a false signal of population expansion or a selective sweep. We do not believe that the bias observed by Wares has influenced our results. There are no significantly negative Tajima's *D* values in our results that are not supported by at least one of the other frequency-based tests of neutrality (Table 3). In addition, there were cases in which a significant signal of expansion was detected by the other two tests, but not by using Tajima's *D*.

As in *A. aspera*, the Bayesian skyride tests did not recover significant size changes for any individual populations for any

marker (Supporting Information). The skyride plots for the combined datasets recovered signals of significant size change for two-third of the combined *A. spinosa* datasets, both in the form of a population expansion (Fig. 5). The third plot, COI, showed a signal of decline, which, as with the summary statistics, was likely due to the large genetic divergence among populations.

Mitochondrial and nuclear markers recovered different root heights and times of population size changes both within and between species (Fig. 5). In the case of *A. aspera*, the *atrop* plot showed a signal of population expansion for nearly the entire history of the sample, as evidenced by the median effective population size. The *rag1* dataset showed an older maximum root height than in *atrop*, although the 95% HPD estimates of the root height for the two genes overlapped. The *rag1* dataset did not suggest population expansion until roughly 500,000 years ago—the same time as the inferred expansion using *atrop*, showing concordance between the markers.

The *A. aspera* COI dataset recovered a significantly younger TMRCA and a more recent expansion than the nuclear markers (Fig. 5). The upper root height for the COI dataset was 35,000 years, over 14-fold younger than either of the nuclear

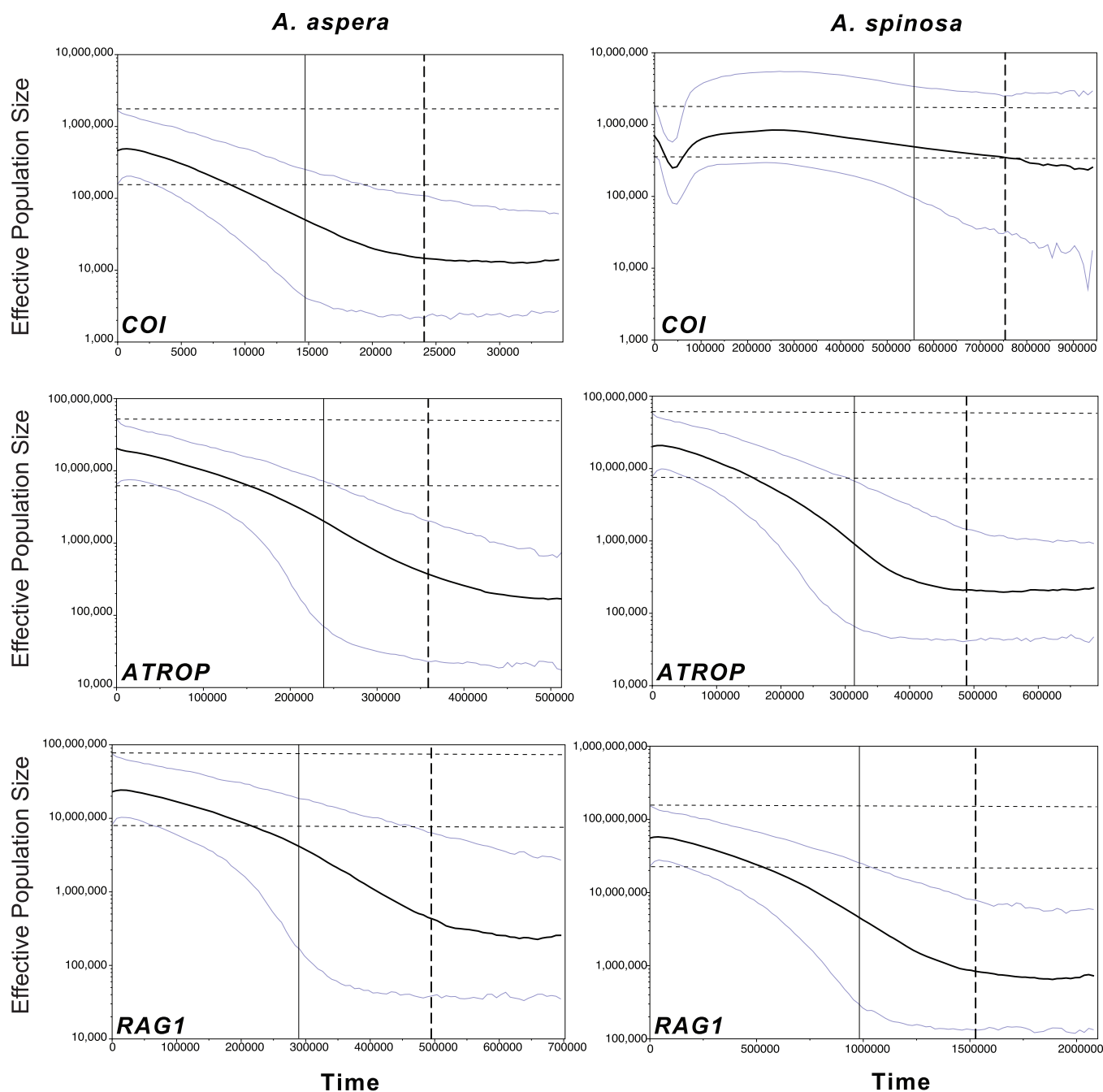


Figure 5. GMRF skyride plots for each marker for *A. aspera* and *A. spinosa* inferred in BEAST. Time, in years, is shown on the x-axis. Effective population size in log number of individuals is shown on the y-axis. The central dark horizontal line in the plot is the median value for effective population size; the light lines are the upper and lower 95% HPD for those estimates. The vertical dashed line represents the median TMRCA. The upper 95% HPD on TMRCA is at the right end of the plot, whereas the lower 95% HPD is the vertical line to the left of the median. Horizontal dashed lines represent the cutoffs used in this study to assess significance of population size changes.

markers, and the inferred expansion began 20,000 years ago. In contrast to the *A. spinosa* mtDNA skyride plot, the *A. aspera* dataset did not show a signature of population decline, even though Φ_{ST} and STRUCTURE analyses indicated similar levels of subdivision among populations.

Both *A. spinosa* nuclear genes recovered signals of expansions, but the inferred timing of the expansions and root heights differed between the markers. The *atrop* dataset showed a population expansion that began ~400,000 years ago and a maximum TMRCA of 690,000 years, neither of which were significantly

different from *A. aspera atrop*. In contrast, the *rag1* data showed an expansion that began earlier, 1.5 mya, and a significantly older maximum root height than *atrop* (and *A. aspera rag1*) of 2.1 my. Individual population-level comparisons (Supporting Information), did not recover significantly different root heights between the species for any of the markers or populations.

The effective population sizes that were estimated from the datasets where all populations were combined did not differ significantly between species for any of the markers (Supporting Information). The results from the nuclear markers showed that both species have large effective population sizes, over 15 million. Individual populations also had large effective sizes for both species (Supporting Information), with median numbers of individuals ranging from 1.2 to 11.7 million, consistent with reported population densities (Clarke 1994) and museum collections (Greenfield 1981; Greenfield and Johnson 1990). The *COI* estimates, representing the effective number of females, were significantly lower than those for the nuclear datasets with median effective population sizes for *A. aspera* and *A. spinosa* of 463,000 and 693,000 individuals, respectively.

Discussion

Recent studies have found that pelagic larval duration can be a poor predictor of differences in Φ_{ST} among marine species (Bowen et al. 2006; Weersing and Toonen 2009). We found the opposite to be true for *A. aspera* and *A. spinosa*: the two species have identical pelagic larval durations (21–24 days) (Johnson and Brothers 1989) and showed near-identical patterns of pairwise Φ_{ST} values (Table 1) and genetic differentiation as reported by STRUCTURE (Fig. 2). These concordant patterns of subdivision recovered from frequency-based analyses were, however, superficial and misleading. *Acanthemblemaria spinosa* has far greater *COI* divergence than *A. aspera* (~20 \times) among populations (Table 2), a difference ignored by the Φ_{ST} and STRUCTURE analyses because they treat all alleles identically, regardless of the number of substitutions separating them. Inferred patterns of historical demography (Fig. 5) differed between mitochondrial and nuclear markers, which may be due to the very rapid rate of mitochondrial evolution in these fish. This rapid rate has obscured signals of old expansions for both species, which were only revealed by nuclear DNA. However, the mitochondrial DNA, in conjunction with the nuclear DNA, allowed us to recover temporally separated population expansions in *A. aspera*.

SUBSTITUTION RATES

The large level of mtDNA sequence divergence among populations that we found for *A. spinosa* (Table 2) is surprising for a marine fish with a 21–24 day pelagic larval duration. This level of sequence divergence may not be unique to *A. spinosa*; a phyloge-

graphic study on *Acanthemblemaria* from the eastern Pacific also found high mitochondrial DNA sequence divergence among populations (Lin et al. 2009). Our results here reveal a mitochondrial substitution rate that is high both in absolute terms and relative to nuclear rates in the same species, and demonstrate the effect that incorrect substitution rate estimates have on the estimation of population genetic parameters.

The inferred *COI* substitution rate of 11.22% per million years is one of the fastest vertebrate mitochondrial rates known (Nabholz et al. 2008, 2009; Welch et al. 2008). One possible reason for the fast mitochondrial rate we estimated would be that the transisthmian calibration we employed was incorrect and that divergence occurred long before the closure of the isthmus, as seen in other taxa (Knowlton et al. 1993; Marko 2002; Lessios 2008). The use of slower rates seen in other teleosts discounts this possibility. When the slowest *COI* rate from Lessios (2008) for transisthmian geminate fish (1.03% per million years) was used, the mean inferred divergence for *A. exilispinus* and *A. betinensis* was 37.8 mya (confidence interval of 50.9 to 25.9 mya). At that time, abyssal water depths connected the eastern Pacific and western Atlantic (Lessios 2008), making an initial divergence between species that live in close association with coral reefs and in 1 m of water unlikely. For the maximum substitution rate listed in Lessios (2008), 1.77% per million years, and using the K2P model, mean divergence time is 11.8 mya with upper and lower bounds of 14.0 and 9.7. These dates would also place the initial divergence of the *A. exilispinus* and *A. betinensis* at a time when appropriate habitat did not exist.

In contrast, the substitution rates inferred for the nuclear genes do not appear to be exceptionally fast. The value we obtained for *rag1*, 0.30% per million years, was faster than that listed in Lessios (2008) for *rag1*. However, when the rate from Lessios (2008) was used in BEAST analyses (0.097% per million years), the confidence interval on the transisthmian divergence overlapped the one obtained here in the analysis using the exponential prior on divergence time. The substitution rate estimated for *atrop*, 0.80% per million years, agreed with published estimates for autosomal introns in birds (0.72% per million years) (Axelsson et al. 2004). Together, these data support an exceptionally fast mitochondrial substitution rate in *Acanthemblemaria*.

These results illustrate the problems that could be introduced by using a published substitution rate that is inappropriate for the taxa under consideration by causing substantial bias in demographic parameters. In addition to the extreme overestimation of divergence times illustrated above, effective population size estimates would be systematically overestimated by using a substitution rate that is too slow. An incorrect substitution rate also leads to biases in coalescent estimation of population size change and migration rates as these values are dependent on substitution rate.

The ratio of mitochondrial to nuclear exon substitution rate, 37.6:1, is also one of the greatest known for animals (Caccone et al. 2004; Willett and Burton 2004; Oliveira et al. 2008; The Nasonia Genome Working Group 2010). This large ratio may have consequences for postzygotic isolation due to epistasis between co-adapted nuclear and mitochondrial genotypes. Proteins encoded in the mitochondrial genome, such as those responsible for oxidative phosphorylation, directly interact with nuclear-encoded proteins. Gene products from each genome must be able to work properly with each other, or organismal breakdown may occur. This has been seen in hybrids with mismatched nuclear and mitochondrial genomes (Rawson and Burton 2002; Burton et al. 2006; The Nasonia Genome Working Group 2010). In *Nasonia* wasps, nuclear genes that interact directly with the mitochondrion have a significantly higher synonymous-to-nonsynonymous substitution ratio (dN/dS) than those that do not (The Nasonia Genome Working Group 2010). Finding similarly high compensatory dN/dS ratios in *Acanthemblemaria* would suggest the possibility of coevolution of nuclear and mitochondrial genomes that could lead to hybrid breakdown through cyto-nuclear disequilibrium in these fish.

DEMOGRAPHIC HISTORIES OF *A. ASPERA* AND *A. SPINOSA*

Our demographic analyses revealed two bouts of population expansion: older expansions of both species and a younger one specific to *A. aspera*. The former was recovered by nuclear DNA, the latter by mitochondrial DNA. Alone, neither marker type, mitochondrial or nuclear, would have provided a complete picture of the historical demography of these fish.

The nuclear data recovered a population expansion dating to 400,000–500,000 years ago for both species, although the *A. spinosa rag1* data indicate a significantly older expansion than *A. spinosa atrop*, beginning ~ 1.5 mya (Fig. 5), as well as a significantly older root age (Supporting Information). This discrepancy between nuclear loci probably arose from coalescent stochasticity, as different markers in the nuclear genome can have different times to their most recent common ancestor.

The *COI* skyride plot for all *A. aspera* populations combined recovered a population expansion beginning $\sim 20,000$ year ago. This coincides with the last glacial maximum, a period of lowered sea levels when there was 89% less available shelf area in the Caribbean basin (Bellwood and Wainwright 2002) than at present. This reduced habitat availability may have caused reduced population sizes in *A. aspera*. Such a population bottleneck would cause the root of the *COI* gene tree to appear to be quite young. As glaciers receded and sea level in the Caribbean basin rose (Lambeck et al. 2002), habitat suitable for coral reef species was restored (Montaggioni 2000) and populations grew. Genetic signatures of population expansion that date to increased habitat

availability following maximum global glaciation have also been found in other coral reef fishes (Fauvelot et al. 2003; Rocha et al. 2005; Thacker et al. 2008).

Within *A. aspera*, then, mitochondrial and nuclear markers recovered different aspects of population history. The older population expansion may reflect the initial spread of both species throughout the Caribbean at the time of speciation and would account for the discrepancy between the inferred gene tree root heights of the mitochondrial and nuclear markers.

A single gene region cannot recover a population history older than the most recent bottleneck (Heled and Drummond 2008) which, due to its smaller effective population size, should be more recent for mitochondrial than nuclear data. The severity and recency of the most recent bottleneck would affect the height of the gene tree, whereas the substitution rate of the mitochondria would determine the amount of signal that would be available to detect the recovery. The pattern of recent expansion revealed by *COI* was nearly hidden by high rates of substitution, which resulted in fixed mitochondrial haplotypes among populations that may have been interpreted as evidence for long-term isolation. This is evidenced by most *A. aspera* populations having private alleles and corrected Φ_{ST} values of 1, which is not the expectation under a scenario of population growth (Excoffier et al. 2009).

Although both species shelter in holes in corals, they differ in their microhabitat use and in their propensity to go locally extinct (Clarke 1994, 1996). *Acanthemblemaria spinosa* is found in shallower water than *A. aspera* (Clarke 1994), although the two overlap at intermediate depths. *Acanthemblemaria spinosa* occurs only in high-profile shelters up off the reef in living or standing dead coral, whereas the less-specialized *A. aspera* can persist in low-profile habitat on the reef surface in coral rubble (Clarke 1994). These differences in microhabitat use give *A. spinosa* a greater propensity to go locally extinct. Thus, the same processes that can cause *A. spinosa* populations to decline (when living and standing dead coral is destroyed and reduced to rubble) can allow *A. aspera* populations to grow (Clarke 1996).

Given what is known about the differences in microhabitat requirements of these fish, the expectation is that *A. aspera* populations should be more stable over time than *A. spinosa* populations, at least at ecological time scales. In addition to local high-frequency demographic cycles, regional changes in habitat availability due to glaciation would also be expected to favor persistence of *A. aspera* populations over *A. spinosa* at evolutionary time scales. Given that *A. spinosa* lives in shallower waters than *A. aspera*, the substantial reduction in shelf area during intervals of low sea level would be expected to have a greater effect on *A. spinosa* than *A. aspera*.

In this study, however, we found demographic patterns that were at odds with those predicted by the ecology and life histories

of these blennies. On the one hand, the similarities in the life histories of the two species did not translate into concordant demographic histories. On the other hand, the differences we did find in the historical demography of the two species were in the opposite direction than expected from their differences in microhabitat use. Although there was a signal of recoveries from bottlenecks in individual *A. spinosa* populations, as suggested by point estimates (Table 3) and TCS haplotype distributions (although Bayesian skyride tests did not), there was no evidence of a range-wide extirpation for the species. However, *A. aspera* appears to have undergone a range-wide bottleneck. Together, our results indicate that, compared to *A. aspera*, *A. spinosa* populations were better able to persist during lower sea levels at the last glacial maximum.

It is not clear then why we recovered a pattern of range-wide population expansion in *A. aspera* and population persistence in *A. spinosa*. In previous comparative studies of historical demography in marine taxa, interspecific ecological differences correlated well with habitat requirements (Marko 2004; Hickerson and Cunningham 2005). However, those studies involved intertidal taxa that were directly affected by glaciation. It may be that less obvious factors than ecological differences are responsible for contrasting demographic patterns in coral reef taxa resulting from sea-level changes. Sampling of additional nuclear markers to test for multiple population size changes using a method such as the extended Bayesian skyline plot (Heled and Drummond 2008) could provide further insight into the timing and degree of demographic changes through multiple glacial cycles.

CONCLUSIONS

Our study illustrates that mitochondrial and nuclear markers can reveal complementary information in historical demographic studies. The smaller effective size and rapid substitution rate of the mitochondrial DNA allowed the inference of a recent population expansion in *A. aspera*, whereas the slower nuclear DNA recovered an older expansion for both species. However, the rapid mitochondrial substitution rate also obscured the recent expansion in *A. aspera*. Analyses of the mitochondrial data using frequency-based metrics alone did not indicate the underlying population expansions in *A. aspera*, neither young, nor old. The results of the frequency-based tests, coupled with the STRUCTURE results, lead to a pattern of subdivision that was very similar to that of *A. spinosa* even though the underlying demography of the two species was quite different.

ACKNOWLEDGMENTS

We thank R. Clarke, B. Holt, and E. Whiteman for field collections and P. Hastings and C. Klepadlo for providing tissue samples from the SIO Marine Vertebrates collection. J. Maley and D. Warren provided computational assistance. We thank D. Warren, A. Caballero, and the staff at Nature Foundation SXM and Dolphin Encounters in Nassau for

providing help in the field and C. Restrepo for helping in the laboratory. We thank B. Carstens, J. Cronin, M. DeBiase, C. Prada, M. Alfaro, and two anonymous reviewers for comments that significantly improved the manuscript. Extensive discussions with R. Clarke and P. Hastings on *Acanthemblemaria* blennies helped greatly. We thank the governments of the Bahamas, St. Maarten, and Belize for providing research permits to collect samples. Support for this research was provided to RIE by LSU BioGrads, the McDaniel Travel Fund, an LSU Graduate School travel award, a Sigma Xi Grant in Aid of Research Award, an ASIH Raney Fund Award, a Lerner-Gray Grant for Marine Research, and the Association of the Marine Labs of the Caribbean, and to MEH by the National Science Foundation (OCE-0550270 to MEH and Iliana Baums). This study was conducted under IACUC protocol No. 05-040 at Louisiana State University. Portions of this research were conducted with high-performance computational resources provided by the Louisiana Optical Network Initiative (<http://www.loni.org>).

LITERATURE CITED

- Avice, J. C. 2000. *Phylogeography: the history and formation of species*. Harvard Univ. Press, Cambridge, MA.
- Axelsson, E., N. Smith, H. Sundstrom, S. Berlin, and H. Ellegren. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. *Mol. Biol. Evol.* 21:1538.
- Barrowclough, G. F., and R. M. Zink. 2009. Funds enough, and time: mtDNA, nuDNA and the discovery of divergence. *Mol. Ecol.* 18:2934–2936.
- Baums, I. B., M. W. Miller, and M. E. Hellberg. 2005. Regionally isolated populations of an imperiled Caribbean coral, *Acropora palmata*. *Mol. Ecol.* 14:1377–1390.
- Bellwood, D. R., and P. C. Wainwright. 2002. The history and biogeography of fishes on coral reefs. Pp. 5–32 in P. F. Sale, ed. *Coral reef fishes*. Elsevier Science, Burlington, MA.
- Böhlke, J. E., and C. G. C. Chaplin. 1993. *Fishes of the Bahamas and adjacent tropical waters*. Univ. of Texas Press, Austin.
- Bowen, B. W., A. L. Bass, A. Muss, J. Carlin, and D. R. Robertson. 2006. Phylogeography of two Atlantic squirrelfishes (Family Holocentridae): exploring links between pelagic larval duration and population connectivity. *Marine Biol.* 149:899–913.
- Brito, P., and S. Edwards. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* 135:439–455.
- Brunsfeld, S. J., J. Sullivan, D. E. Soltis, and P. S. Soltis. 2001. A comparative phylogeography of northwestern North America: a synthesis. Pp. 319–340 in J. Silvertown and J. Antonovics, eds. *Integrating ecology and evolution in a spatial context*. Blackwell Science, Oxford.
- Burney, C. W., and R. T. Brumfield. 2009. Ecology predicts levels of genetic differentiation in neotropical birds. *Am. Nat.* 174:358–368.
- Burton, R. S., C. K. Ellison, and J. S. Harrison. 2006. The sorry state of F-2 hybrids: consequences of rapid mitochondrial DNA evolution in allopatric populations. *Am. Nat.* 168:S14–S24.
- Caccone, A., G. Gentile, C. Burns, E. Sezzi, W. Bergman, M. Ruelle, K. Saltonstall, and J. Powell. 2004. Extreme difference in rate of mitochondrial and nuclear DNA evolution in a large ectotherm, Galapagos tortoises. *Mol. Phylogenet. Evol.* 31:794–798.
- Carling, M. D., and R. T. Brumfield. 2007. Gene sampling strategies for multi-locus population estimates of genetic diversity (θ). *PLoS ONE* 2.
- Carstens, B. C., S. J. Brunsfeld, J. R. Demboski, J. M. Good, and J. Sullivan. 2005. Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: hypothesis testing within a comparative phylogeographic framework. *Evolution* 59:1639–1652.
- Clarke, R. D. 1994. Habitat partitioning by chaenopsid blennies in Belize and the Virgin Islands. *Copeia*:398–405.

- . 1996. Population shifts in two competing fish species on a degrading coral reef. *Mar. Ecol. Prog. Ser.* 137:51–58.
- Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9:1657–1659.
- Cunningham, C. W., and T. M. Collins. 1994. Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. Pp. 405–433 in B. Schierwater, B. Streit, P. Wagner, and R. DeSalle, eds. *Molecular ecology and evolution: approaches and applications*. Birkhauser Verlag, Basel, Switzerland.
- Drummond, A., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.
- Drummond, A., A. Rambaut, B. Shapiro, and O. Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22:1185.
- Drummond, A. J., B. Ashton, M. Cheung, J. Heled, M. Kearse, R. Moir, S. Stones-Havas, T. Thierer, and A. Wilson. 2009. Geneious v4.5.4. Available at: <http://www.geneious.com>.
- Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *Plos Biology* 4:699–710.
- Earl, D. 2009. Structure Harvester v0.3. Available at: http://users.soe.ucsc.edu/~dearl/software/struct_harvest/.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.
- Edwards, S., and S. Bensch. 2009. Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Mol. Ecol.* 18:2930–2933.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14:2611–2620.
- Excoffier, L., R. Petit, and M. Foll. 2009. Genetic consequences of range expansion. *Annu. Rev. Ecol. Evol. Syst.* 40:481–501.
- Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes—application to human mitochondrial-DNA restriction data. *Genetics* 131:479–491.
- Fauvelot, C., G. Bernardi, and S. Planes. 2003. Reductions in the mitochondrial DNA diversity of coral reef fish provide evidence of population bottlenecks resulting from Holocene sea-level change. *Evolution* 57:1571–1583.
- Felsenstein, J. 2006. Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? *Mol. Biol. Evol.* 23:691–700.
- Flot, J. 2007. CHAMPURU 1.0: a computer software for unraveling mixtures of two DNA sequences of unequal lengths. *Mol. Ecol. Notes* 7:974–977.
- Flot, J.-F. 2010. SeqPHASE: a web tool for interconverting phase input/output files and fasta sequence alignments. *Mol. Ecol. Resources* 10:162–166.
- Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.
- Galtier, N., B. Nabholz, S. Glémin, and G. Hurst. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol. Ecol.* 18:4541–4550.
- Graur, D., and W.-H. Li. 2000. *Fundamentals of molecular evolution*. Sinauer Associates, Inc., Sunderland, MA.
- Greenfield, D. W. 1981. The blennioid fishes of Belize and Honduras, Central America, with comments on their systematics, ecology, and distribution (Blenniidae, Chaenopsidae, Labrisomidae, Tripterygiidae). *Fiel-diana Zool.* 8:1–106.
- Greenfield, D. W., and R. K. Johnson. 1990. Community structure of western Caribbean blennioid fishes. *Copeia*:433–448.
- Hare, M. 2001. Prospects for nuclear gene phylogeography. *Trends Ecol. Evol.* 16:700–706.
- Hastings, P. A. 1990. Phylogenetic-relationships of tube blennies of the genus *Acanthemblemaria* (Pisces, Blennioidei). *Bulletin Marine Science* 47:725–738.
- . 2002. Evolution of morphological and behavioral ontogenies in females of a highly dimorphic clade of blennioid fishes. *Evolution* 56:1644–1654.
- Hedrick, P. W. 2005. A standardized genetic differentiation measure. *Evolution* 59:1633–1638.
- Hein, J., M. H. Schierup, and C. Wiuf. 2005. *Gene genealogies, variation and evolution*. Oxford Univ. Press, New York.
- Heled, J., and A. Drummond. 2008. Bayesian inference of population size history from multiple loci. *BMC Evol. Biol.* 8:289.
- Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913.
- Hey, J. 2010. Isolation with migration models for more than two populations. *Mol. Biol. Evol.* 27:905–920.
- Hickerson, M., B. Carstens, J. Cavender-Bares, K. Crandall, C. Graham, J. Johnson, L. Rissler, P. Victoriano, and A. Yoder. 2009. Phylogeography's past, present, and future: 10 years after Avise 2000. *Mol. Phylogenet. Evol.* 54:291–301.
- Hickerson, M., E. Stahl, and H. Lessios. 2006. Test for simultaneous divergence using approximate Bayesian computation. *Evolution* 60:2435–2453.
- Hickerson, M. J., and C. W. Cunningham. 2005. Contrasting quaternary histories in an ecologically divergent sister pair of low-dispersing intertidal fish (*Xiphister*) revealed by multilocus DNA analysis. *Evolution* 59:344–360.
- Hubisz, M., D. Falush, M. Stephens, and J. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Res.* 9:1322–1332.
- Hudson, M. 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Res.* 8:3–17.
- Hudson, R. R., and M. Turelli. 2003. Stochasticity overrules the “three-times rule”: genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution* 57:182–190.
- Huelsenbeck, J. P., J. P. Bollback, and A. M. Levine. 2002. Inferring the root of a phylogenetic tree. *Syst. Biol.* 51:32–43.
- Jakobsson, M., and N. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801.
- Johnson, G. D., and E. B. Brothers. 1989. *Acanthemblemaria paula*, a new diminutive chaenopsid (Pisces, Blennioidei) from Belize, with comments on life history. *Proc. Biol. Soc. Washington* 102:1018–1030.
- Knowlton, N., L. A. Weigt, L. A. Solorzano, D. K. Mills, and E. Bermingham. 1993. Divergence in proteins, mitochondrial-DNA, and reproductive compatibility across the Isthmus of Panama. *Science* 260:1629–1632.
- Lambeck, K., Y. Yokoyama, and T. Purcell. 2002. Into and out of the last glacial maximum: sea-level change during oxygen isotope stages 3 and 2. *Q. Sci. Rev.* 21:343–360.
- Lee, J. Y., and S. V. Edwards. 2008. Divergence across Australia's Carpentarian barrier: statistical phylogeography of the red-backed fairy wren (*Malurus melanocephalus*). *Evolution* 62:3117–3134.
- Lessios, H. 2008. The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annu. Rev. Ecol. Evol. Syst.* 39:63–91.
- Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452.
- Lin, H. C., C. Sánchez-Ortiz, and P. A. Hastings. 2009. Colour variation is incongruent with mitochondrial lineages: cryptic speciation and

- subsequent diversification in a Gulf of California reef fish (Teleostei: Blennioidei). *Mol. Ecol.* 18:2476–2488.
- Marko, P. B. 2002. Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Mol. Biol. Evol.* 19:2005–2021.
- . 2004. ‘What’s larvae got to do with it?’ Disparate patterns of postglacial population structure in two benthic marine gastropods with identical dispersal potential. *Mol. Ecol.* 13:597–611.
- Meirmans, P. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60:2399–2402.
- Meirmans, P. G., and P. H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* 4:792–794.
- Michalakis, Y., and L. Excoffier. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142:1061–1064.
- Minin, V., E. Bloomquist, and M. Suchard. 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* 25:1459.
- Montaggioni, L. 2000. Postglacial reef growth. *Comptes Rendus de l’Academie des Sciences Series IIA Earth and Planetary Science* 331:319–330.
- Moore, W. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49:718–726.
- Nabholz, B., S. Glémin, and N. Galtier. 2008. Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. *Mol. Biol. Evol.* 25:120.
- . 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evol. Biol.* 9:54.
- Nylander, J., J. Wilgenbusch, D. Warren, and D. Swofford. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24:581.
- Oliveira, D., R. Raychoudhury, D. Lavrov, and J. Werren. 2008. Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Mol. Biol. Evol.* 25:2167.
- Pond, S. L. K., and S. D. W. Frost. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531–2533.
- Pond, S. L. K., D. Posada, M. B. Gravenor, C. H. Woelk, and S. D. W. Frost. 2006. GARD: a genetic algorithm for recombination detection. *Bioinformatics* 22:3096–3098.
- Pons, J., T. G. Barraclough, J. Gomez-Zurita, A. Cardoso, D. P. Duran, S. Hazell, S. Kamoun, W. D. Sumlin, and A. P. Vogler. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55:595–609.
- Posada, D. 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25:1253–1256.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Pritchard, J. K., X. Wen, and D. Falush. 2009. Documentation for *structure* software: version 2.3. Available at: <http://pritch.bsd.uchicago.edu>.
- R Development Core Team. 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Ramos-Onsins, S., and J. Rozas. 2002. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19:2092.
- Rawson, P., and R. Burton. 2002. Functional coadaptation between cytochrome c and cytochrome c oxidase within allopatric populations of a marine copepod. *Proc. Natl. Acad. Sci. USA* 99:12955.
- Rocha, L. A., D. R. Robertson, C. R. Rocha, J. L. Van Tassel, M. T. Craig, and B. W. Bowers. 2005. Recent invasion of the tropical Atlantic by an Indo-Pacific coral reef fish. *Mol. Ecol.* 14:3921–3928.
- Smith-Vaniz, W. F., and F. J. Palacio. 1974. Atlantic fishes of the genus *Acanthemblemaria*, with description of three new species and comments on Pacific species (Clinidae: Chaenopsinae). *Proc. Acad. Natl. Sci. Philadelphia* 125:197–224.
- Stephens, J. S. 1963. A revised classification of the blennioid fishes of the American family Chaenopsidae. *Univ. of Calif. Publ. Zool.* 68:1–165.
- Stephens, M., and P. Donnelly. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.* 73:1162–1169.
- Stephens, M., and P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am. J. Hum. Genet.* 76:449–462.
- Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68:978–989.
- Suchard, M. A., and B. D. Redelings. 2006. BAli-Phy: simultaneous Bayesian inference of alignment and phylogeny. *Bioinformatics* 22:2047–2048.
- Taberlet, P., L. Fumagalli, A. Wust-Saucy, and J. Cosson. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* 7:453–464.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585.
- Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24:1596–1599.
- Taylor, M. S., and M. E. Hellberg. 2006. Comparative phylogeography in a genus of coral reef fishes: biogeographic and genetic concordance in the Caribbean. *Mol. Ecol.* 15:695–707.
- Thacker, C., A. Thompson, D. Roje, and E. Shaw. 2008. New expansions in old clades: population genetics and phylogeny of *Gnatholepis* species (Teleostei: Gobioidae) in the Pacific. *Marine Biol.* 153:375–385.
- The *Nasonia* Genome Working Group. 2010. Functional and evolutionary insights from the genomes of three parasitoid *nasonia* species. *Science* 327:343–348.
- Villesen, P. 2007. FaBox: an online toolbox for FASTA sequences. *Mol. Ecol. Notes* 7:965–968.
- Waples, R., and O. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol. Ecol.* 15:1419–1440.
- Wares, J. P. 2010. Natural distributions of mitochondrial sequence diversity support new null hypotheses. *Evolution* 64:1136–1142.
- Weersing, K., and R. Toonen. 2009. Population genetics, larval dispersal, and connectivity in marine systems. *Mar. Ecol. Prog. Ser.* 393:1–12.
- Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370.
- Welch, J., O. Bininda-Emonds, and L. Bromham. 2008. Correlates of substitution rate variation in mammalian protein-coding sequences. *BMC Evol. Biol.* 8:53.
- Willett, C. S., and R. S. Burton. 2004. Evolution of interacting proteins in the mitochondrial electron transport system in a marine copepod. *Mol. Biol. Evol.* 21:443–453.
- Zhang, D. X., and G. M. Hewitt. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects (vol 12, pg 563, 2003). *Mol. Ecol.* 12:1687–1687.
- Zink, R., and G. Barrowclough. 2008. Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* 17:2107–2121.

Supporting Information

The following supporting information is available for this article:

Figure S1. Mean values and standard deviations of $L(K)$ for K 1–6 from 10 STRUCTURE runs of the *A. aspera* eastern Caribbean dataset.

Figure S2. Bayesian skyride plots for the Bahamas.

Figure S3. Bayesian skyride plots for Belize and Honduras.

Figure S4. Bayesian skyride plots for Puerto Rico and St. Thomas.

Figure S5. Bayesian skyride plots for St. Maarten.

Table S1. Collection localities and sample sizes for species used in the present study.

Table S2. PCR conditions, primer sequences, and references.

Table S3. Models of sequence evolution used in the study.

Table S4. Estimates of gene tree root heights and current effective population sizes from skyride analyses.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.