Journal of Heredity 2013:104(4):521–531
doi:10.1093/jhered/est024
Advance Access publication April 24, 2013

© The American Genetic Association. 2013. All rights reserved.
For permissions, please email: journals.permissions@oup.com.

American Genetic Association
The Journal of Heredity
www.jhered.oxfordjournals.org

Spatiotemporal Genetic Structure in a Protected Marine Fish, the California Grunion (*Leuresthes tenuis*), and Relatedness in the Genus *Leuresthes*

Rosemary J. Byrne, Giacomo Bernardi, and John C. Avise

From the Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697 (Byrne and Avise) and the Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060 (Bernardi).

Address correspondence to Rosemary J. Byrne at the address above, or e-mail: rbyrne@uci.edu.

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

Abstract

The genus *Leuresthes* displays reproductive behavior unique among marine fish in which mature adults synchronously emerge completely out of the water to spawn on beach land. A limited number of sandy beaches, which are suitable for these spawning events, are present in discontinuous locations along the geographic range of the species, potentially limiting gene flow and the degree of genetic homogeneity between intraspecific populations. Here, we tested for molecular genetic differentiation between 363 individuals, representing 3 populations of California grunion, *Leuresthes tenuis*, by employing 2 mitochondrial and 4 nuclear DNA markers. We include temporally diverse sampling to evaluate contemporary and temporal divergence, and we also analyze 28 individuals from one population of Gulf grunion (restricted to the Gulf of California), *Leuresthes sardina*, at the same markers to evaluate the molecular evidence for their separate species distinction. We find no significant differences between temporal samples, but small significant differences among all populations of *L. tenuis*, and unequivocal support for the separate species distinction of *L. sardina*. Genetic data suggest that the Monterey Bay population of *L. tenuis* near the species' most northern range likely represents a relatively recent colonization event from populations along the species' more traditional range south of Point Conception, California. We conclude that both the topographic features of the California and Baja California coastlines and the grunions' unique reproductive behavior have influenced the genetic structure of the populations.

Key words: *gene flow, microsatellites, phylogeography, population structure*

A wealth of population structure literature has shown that genetic breaks can be difficult to predict for California marine species. Some taxa show great intraspecific genetic similarity along the coast (Addison et al. 2008), some exhibit isolation by distance (Buonaccorsi et al. 2004; Lewallen et al. 2007), and others show significant phylogeographic structure concordant with recognized biogeographic breaks, such as Point Conception and Monterey Bay (reviewed by Burton 1998; Dawson 2001). Such results indicate that in addition to geographic factors in the environment, the pattern of population structure may be strongly influenced by larval dispersal and behavioral traits (Sivasundar and Palumbi 2010). Some fish species, such as the grunions (genus *Leuresthes)*, do not take baited hooks and are rarely collected in trawls making direct observation of movement at sea difficult or rare. For such taxa, genetic monitoring provides the most informative insights into migration and dispersal over time.

The genus *Leuresthes* is composed of two species of marine fish: the Gulf grunion, *Leuresthes sardina* (Jenkins and Evermann 1889), and the California grunion, *Leuresthes tenuis* (Ayres 1860). Together they represent the only two species of marine fishes known to constitutively leave the water to spawn on beach land (Thompson 1919; Walker 1952), and the only marine fishes whose egg incubation is completely terrestrial. Although this unique mode of reproduction has stimulated interest in ecological and molecular genetic studies, it has also raised some concern over the stability of the populations on a temporal scale; the fishes are vulnerable to the alteration of beach habitat, which may render areas unsuitable for egg laying, and also to human predation while mating. In fact, the recreational

American Genetic Association
The Journal of Heredity
www.jhered.oxfordjournals.org

fishery was reportedly becoming depleted during the 1920s, and although population sizes of the California grunion have never been known, it is not thought to be an abundant species (Gregory 2001). These concerns have led to limited legal protection for *L. tenuis* in the state of California, where a closed season is observed from April through May and take is restricted during the other months of the year (Gregory 2001).

With regard to spatial arrangement, the grunions are allopatric with *L. sardina* endemic to the northern region of the Gulf of California from Bahía Concepción, Baja California Sur, and Guaymas, Sonora, to the Río Colorado Delta (Moffatt and Thomson 1975), whereas *L. tenuis* typically ranges from south of Point Conception, California, to Bahía Magdalena, Baja California Sur (Miller and Lea 1972), although a small and physically disjunct population exists in Monterey Bay (Phillips 1943; Yoklavich et al. 2002, Figure 1). Neither of the species is known to occur around the southern end of Baja California, and *L. tenuis* has been infrequently reported between Point Conception and Monterey Bay in California (Straughan 1982). It is thought that these regions do not possess suitable beach habitat for spawning (Reynolds et al. 1976) and may exhibit unsuitable shoreline temperatures for the adult fish and/or the developing embryos. In addition, a limited amount of tagging data indicate that the California grunion may exhibit spawning site fidelity (Walker 1952) and hatched larvae are thought to remain abundant close to the coastline (Gruber et al. 1982); they are not known to join ichthyoplankton assemblages offshore (McGowan 1993). These factors may prohibit or severely limit migration between southern and northern California waters.

Currently, the California grunion is considered to be a single management unit by the California Department of Fish and Wildlife, although it has not been subject to full genetic monitoring along its range. A few studies have provided some insight on the genetic population structure of the species. An allozyme study conducted by Gaida et al. (2003) analyzing southern California populations from Santa Barbara to San Diego was inconclusive with the data unable to discern between isolation by distance and panmixia. In 2001, a northern extension of the distribution of *L. tenuis* was reported in San Francisco Bay (Jahn 2004), a region where the species had not been recorded since 1860. This population is composed of significantly more diminutive fish than their conspecific counterparts south of Point Conception, both in total length (TL) and in mass (Johnson et al. 2009). However, a recent genetic study using microsatellite and mitochondrial markers indicates that this reportedly new San Francisco Bay population does not show significant genetic differentiation from the southern California populations, despite showing significant differences in size morphology (Johnson et al. 2009). The authors suggest that these findings may represent a range expansion too recent to result in genetic differentiation and they posit that phenotypic plasticity in response to cooler San Francisco Bay water temperatures may be responsible for the notable size differences.

Unlike the San Francisco Bay population, the Monterey Bay population has been known to be small since its presence was first recorded in 1940 (Phillips 1943). This population has not been genetically monitored or compared with the southern California populations. If this population has been genetically isolated from the stocks south of Point Conception for a period of time sufficient to allow assortment of genetic diversity, then there may be reason to consider this population a separate management unit for conservation purposes. Currently, information regarding *L. tenuis* populations in northern Baja California is extremely sparse. It is unknown whether these populations represent ones that are genetically isolated from the southern California populations or whether between-population mating or migration events result in panmixia with respect to genetic population structure.

In addition, existing studies of the California grunion are based on samples taken at a single point in time, and therefore indirectly assume temporal genetic stability of the local populations. Such assumptions may not hold because in finite populations allele frequencies change over time due to genetic drift and, therefore, such analyses cannot attest to the temporal microevolutionary changes of genetic diversity. In order to monitor natural evolutionary genetic diversity, as well as the impacts of anthropogenic forces, temporal samples must be utilized (Schwartz et al. 2007).

In this study, we use mitochondrial DNA (mtDNA) and microsatellite markers to complement previous survey work assessing the pattern of gene flow in spatially varying populations of *L. tenuis* by utilizing new genetic markers, more extensive range coverage (including samples from the previously unstudied Monterey Bay, California, and Ensenada, Baja California, Mexico, populations), and temporally diverse sampling. We also evaluate genetic divergence between *L. tenuis* and *L. sardina* to investigate the molecular evidence for the separate species distinction within the genus.

Materials and Methods

Study Sites and Sampling

Fin clippings or whole animal samples were taken from a total of 391 fishes from 4 locations along the coasts of California, United States of America, and Baja California, Mexico (Figure 1; Table 1), either by hand at random during beach spawning events or offshore by beach seine. A total of 189 samples representing a northern California population of *L. tenuis* (which is physically disjunct from the *L. tenuis* population south of Point Conception) were collected from Monterey Bay between 1999–2002 and 2004–2007. A total of 132 samples from Laguna Beach collected from 2006 to 2008 and 42 samples collected from Ensenada in 2009 represent the southern contingency of the species. Also included are 28 samples of *L. sardina* collected in 1998 from Puerto Peñasco, Mexico (Gulf of California).

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

American Genetic Association
The Journal of Heredity
www.jhered.oxfordjournals.org

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

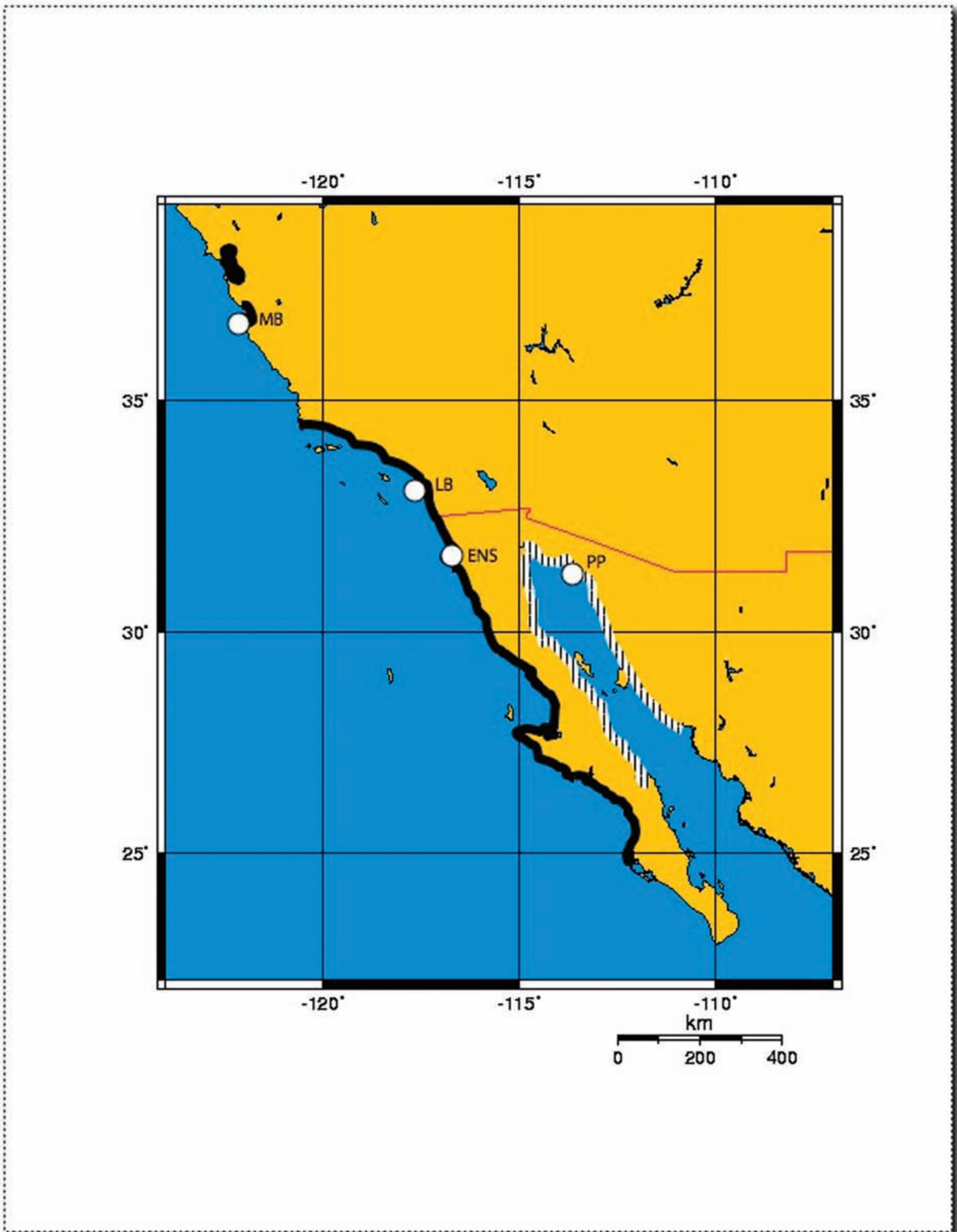

Figure 1. Map of the native ranges of the California grunion (solid lines) and the Gulf grunion (hatched line). Collection locales (Monterey Bay [MB], Laguna Beach [LB], Ensenada [ENS], and Puerto Peñasco [PP]) are indicated with white dots. This map was created using an online map creator at http://www.aquarius.geomar.de/omc.

American Genetic Association

The Journal of Heredity

www.jhered.oxfordjournals.org

Table 1 Population abbreviations and number of fish genotyped at 4 microsatellite loci (N_g) and sequenced at the Cyt b (N_{Cyt}) and CR (N_{CR}) loci for each population of *Leuresthes tenuis* and *Leuresthes sardina*

Species	Sampling location	Collection year	Population abbreviation	N_g	N_{Cyt}	N_{CR}
L. tenuis	Monterey Bay (California, United States)	1999	MB99	38	24	27
		2000	MB00	36		
		2001	MB01	8		
		2002	MB02	8		
		2004	MB04	31		
		2005	MB05	29		
		2006	MB06	29	29	28
		2007	MB07	10		
	Laguna Beach (California, United States)	2006	LB06	54	14	37
		2007	LB07	20	10	18
		2008	LB08	58	15	34
	Ensenada (Baja California, Mexico)	2009	Ens09	42	30	26
L. sardina	Puerto Peñasco (Sonora, Mexico)	1998	PP98	28	28	15

Genetic Data Collection

Samples were preserved in the field by either immediate placement on ice or in 95% ethanol, and all samples were later stored at −20 °C in the laboratory. Genomic DNA was extracted from fin clippings or from muscle tissue (when collected) following the animal tissue protocol of Qiagen DNeasy kits. Polymerase chain reaction (PCR) was used to amplify 4 polymorphic microsatellite loci in all 391 samples using primers previously developed specifically for *L. tenuis*: B18, B19, B39, and B82 (Byrne and Avise 2009). One of the primers for each locus was fluorescently labeled with either FAM (B18 and B82) or HEX (B19 and B39). PCR conditions and characteristics of the loci are described by Byrne and Avise (2009). Following PCR amplification, products were diluted and mixed in 2 multiplex groups (B18 with B19; B39 with B82) for genotyping. Mixtures containing 1 µL of the diluted PCR product mix, 9.55 µL Hi-Di formamide (Applied Biosystems), and 0.45 µL of GeneScan-ROX 500 size standard (Applied Biosystems) were denatured for 3 min at 95 °C and electrophoresed on an ABI 3100 Genetic Analyzer. Scoring of microsatellite alleles was performed automatically by GeneMapper 4.0 software (Applied Biosystems) and verified by eye for each sample. Genotyping errors may greatly bias the results of population genetic studies and therefore, an effort should be made to limit and quantitatively estimate the extent of errors (Bonin et al. 2004). Here, a random subset of the samples (~10%) were rescored for a second time by eye and compared with the previous results to ensure reproducibility of genotyping and allele calling.

A total of 149 and 185 individuals were also PCR amplified at the mitochondrial cytochrome b (Cyt b) and control region (CR) loci, respectively (Table 1). Three sets of previously designed primers were used: two sets to amplify two overlapping fragments of the CR (Johnson et al. 2009) resulting in a combined 392-bp segment and one set to amplify the Cyt b gene (Palumbi 1996) resulting in a 1073-bp segment. Each amplification was performed in a 15-µL mixture of the following: 0.2 mM of each deoxyribonucleotide triphosphate, 0.2 µM each of the forward and reverse primers, 1.5 mM $MgCl_2$, 0.75 units GoTaq DNA Polymerase (Promega), and 3.0 µL 5× GoTaq Reaction Buffer (Promega). A 2-min denaturation at 95 °C began the amplifications, followed by 35 cycles of 30 s at 95 °C, 1 min at annealing temperature (50 °C for CR; 53 °C for Cyt b), and 2 min at 72 °C, followed by a final extension period of 5 min at 72 °C. Each PCR product was sequenced in both directions by Agencourt Bioscience.

Microsatellite Analysis

We utilized the program MICRO-CHECKER (Van Oosterhout et al. 2004) to check for the presence of null alleles, large allele drop out, and other genotyping errors. GENEPOP (Raymond and Rousset 1995) was then employed to calculate expected and observed heterozygosities, as well as deviations from Hardy–Weinberg equilibrium (F_{IS}, as an indicator of heterozygote excess or deficit; Weir and Cockerham 1984), for each population (specific to each location and year) per locus and per locus across populations. Deviations from Hardy–Weinberg equilibrium were also calculated for each sampling location, with data between years combined. Allelic richness estimates and F_{ST} analyses (Weir and Cockerham 1984) were calculated using FSTAT (v. 2.9.3.2; Goudet 1995). F_{ST} analyses were conducted first with 13 populations indicated (accounting for both sampling location and year), and then with 4 populations indicated (accounting only for the area of collection with years collapsed: Monterey Bay [MB], Laguna Beach [LB], Ensenada [ENS], and Puerto Peñasco [PP]). We used POWSIM (Ryman and Palm 2006) to assess the statistical power of our tests for genetic homogeneity. Finally, Nei's genetic distance

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

American Genetic Association

The Journal of Heredity

www.jhered.oxfordjournals.org

(Nei 1978) was calculated between these 4 populations using the program GDA (Lewis and Zaykin 2001). The sequential Bonferroni technique was used to correct *P* values for all multiple tests (Rice 1989).

To assess genetic population structure, a Bayesian genotype clustering analysis was performed with all 391 individuals genotyped at 4 loci using STRUCTURE v. 2.2 (Pritchard et al. 2000). We indicated the number of populations (*K*) in the range 1–13 (with 13 being the maximum possible number of populations counting different years as independent populations), and we assumed admixture in the model and correlated allele frequencies between populations, as is considered best in cases of subtle population structure (Falush et al. 2003). Ten independent runs for each value of *K* were conducted with a 50 000 iteration burn-in period followed by 100 000 Markov chain Monte Carlo repetitions to obtain for each *K* the mean estimated logarithmic probability of data, ln*P*(*D*). In initial trials with STRUCTURE, no meaningful differences were seen when we indicated potential populations of origin versus not indicating origin. For the analyses reported here, the population of origin was indicated in the data file. This process was then repeated with *K* in the range of 1–12, omitting the *L. sardina* samples to allow greater sensitivity to a small level of variability that may exist within *L. tenuis*. Finally, although the maximum value of ln*P*(*D*) is often referred to as the true number of populations, simulations have shown that this method can result in an estimation of *K* larger than the true number of populations. To correct for this phenomenon, we calculated the statistic ΔK, the second-order rate of change of the likelihood function with respect to *K*, as described by Evanno et al. (2005).

Mitochondrial DNA Sequencing Analysis

Sequences were edited and assembled using MegAlign software (DNASTAR) and aligned in MEGA-4 (Tamura et al. 2007). The number of haplotypes (*h*), haplotype diversity (H_d), and nucleotide diversity (π) were calculated in DnaSP 4.10.9 (Rozas et al. 2003) in 2 ways. First, the number of populations was assumed to be the number of original sampling sites and years ($n = 7$; Table 1), and then it was assumed to be equal to the number of sampling locations ($n = 4$), with data from multiple years within 1 sampling location combined. Collapsed data between these 4 locales were then used to estimate the number of migrants between populations and K_{ST} values were calculated with probability obtained by the permutation test with 1000 replicates.

Two separate networks describing haplotype relationships were constructed for the CR and the Cyt b region sequences in TCS 1.21 (Clement et al. 2000) implementing the statistical parsimony algorithm described by Templeton et al. (1992) with a default parsimony connection limit of 95%. Minimum spanning networks were also described using ARLEQUIN v. 3.5.1.2 (Excoffier et al. 2005).

Results

Microsatellite Analysis

Analysis of the genotyping data set via MICRO-CHECKER did not show evidence of null alleles, large allele drop out, or other scoring errors. F_{IS} values were all nonsignificant after Bonferroni corrections and allelic richness (Table 2) was high across all populations for all microsatellite loci. Observed heterozygosities calculated via GENEPOP did not depart from Hardy–Weinberg expectations at any locus, or within any population. The power of our microsatellite markers to detect significant population differentiation was high with simulations in POWSIM suggesting a probability of at least 0.92 to detect a true differentiation of $F_{ST} = 0.0025$ under different scenarios of N_e and numbers of generations of drift (*t*) with 1000 replications.

Pairwise F_{ST} values across all sampling locations and years were low and nonsignificant (Table 3). However, pairwise F_{ST} values between the 4 sampling locations, with years collapsed within the sites, were significant after 120 permutations (Table 4). Calculations of Nei's genetic distance between these 4 populations were low for intraspecific (*L. tenuis*) comparisons, ranging from 0.015 to 0.085, but were high between *L. tenuis* and *L. sardina* populations, ranging from 0.397 to 0.531 (Table 4).

Analysis of the genotyping data in the program STRUCTURE with *K* in the range of 1–12 (only including *L. tenuis* individuals) resulted in the highest mean estimated logarithm of likelihood for $K = 1$, which also exhibited the smallest standard deviation (SD; Figure 2). The distribution of ΔK supported a mode of 2, although the height of this modal value was extremely low ($\Delta K = 9.5$), indicating that the strength of the signal detected by STRUCTURE was weak (Evanno et al. 2005). Individuals did not cluster in any meaningful way with respect to geographic or temporal space. When *K* was set in the range of 1–13 (*L. tenuis* and *L. sardina* samples), the highest mean estimated logarithm of likelihood

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

Table 2 Allelic richness per locus and population (sampling location and year), and per locus across all populations (RT), and observed (H_O) and expected (H_E) heterozygosity estimates for all microsatellite loci

	Allelic richness													
Locus	MB99	MB00	MB04	MB05	MB06	MB07	LB06	LB07	LB08	Ens09	PP98	RT	H_O	H_E
B18	8.59	7.22	7.36	8.52	8.14	8.00	6.78	5.63	8.17	7.77	8.26	7.90	0.88	0.85
B19	10.20	10.79	10.54	11.16	10.17	9.00	10.34	9.49	10.17	9.63	8.51	10.36	0.92	0.91
B39	11.62	10.88	11.76	11.54	12.50	14.00	12.82	11.78	12.59	13.61	15.04	12.84	0.95	0.93
B82	13.26	12.23	12.79	11.03	13.32	11.00	12.56	14.49	13.91	12.34	10.98	13.43	0.95	0.95

Allelic richness calculations are based on a minimum sample size of 10 diploid individuals.

Table 3 Pairwise F_{ST} values between the sampling locations and years using microsatellites

	MB99	MB00	MB01	MB02	MB04	MB05	MB06	MB07	LB06	LB07	LB08	Ens09	PP98
MB99													
MB00	0.004												
MB01	−0.002	−0.003											
MB02	0.108	0.121	0.122										
MB04	0.004	0.009	0.018	0.123									
MB05	0.005	0.010	0.011	0.123	0.009								
MB06	−0.002	0.002	0.006	0.115	−0.001	0.001							
MB07	−0.002	0.001	−0.019	0.119	0.012	0.012	0.001						
LB06	0.005	0.002	0.004	0.123	0.005	0.011	0.002	−0.002					
LB07	0.009	0.001	−0.003	0.133	0.011	0.022	0.010	−0.000	0.007				
LB08	−0.000	0.006	0.005	0.113	−0.001	0.009	−0.001	0.002	0.002	0.007			
Ens09	0.006	0.014	−0.003	0.114	0.010	0.013	0.006	0.004	0.012	0.010	0.007		
PP98	0.033	0.045	0.033	0.117	0.041	0.037	0.031	0.031	0.035	0.049	0.028	0.041	

No values are significantly different from 0 after sequential Bonferroni correction.

Table 4 Pairwise F_{ST} (below diagonal) and Nei's genetic distance (Nei 1978; above diagonal) values between the sampling locations using microsatellites with years collapsed within sites

	MB	LB	ENS	PP
MB		0.015	0.073	0.437
LB	0.001*		0.085	0.397
ENS	0.008*	0.009*		0.531
PP	0.036*	0.034*	0.041*	

*Significance after sequential Bonferroni correction (α = 0.0083) for F_{ST} values.

was also for $K = 1$, however when the data were constrained to fit $K = 2$, the 363 *L. tenuis* samples consistently clustered separately from the 28 *L. sardina* individuals.

Mitochondrial DNA Analyses

Of the 392 sites analyzed at the Cyt b locus, 39 polymorphic sites were identified including 13 unique haplotypes. Hd ranged from 0.075 to 0.270 with the lowest value for the Monterey *L. tenuis* samples and the highest value for the Gulf of California *L. sardina* samples (Table 5). Nucleotide

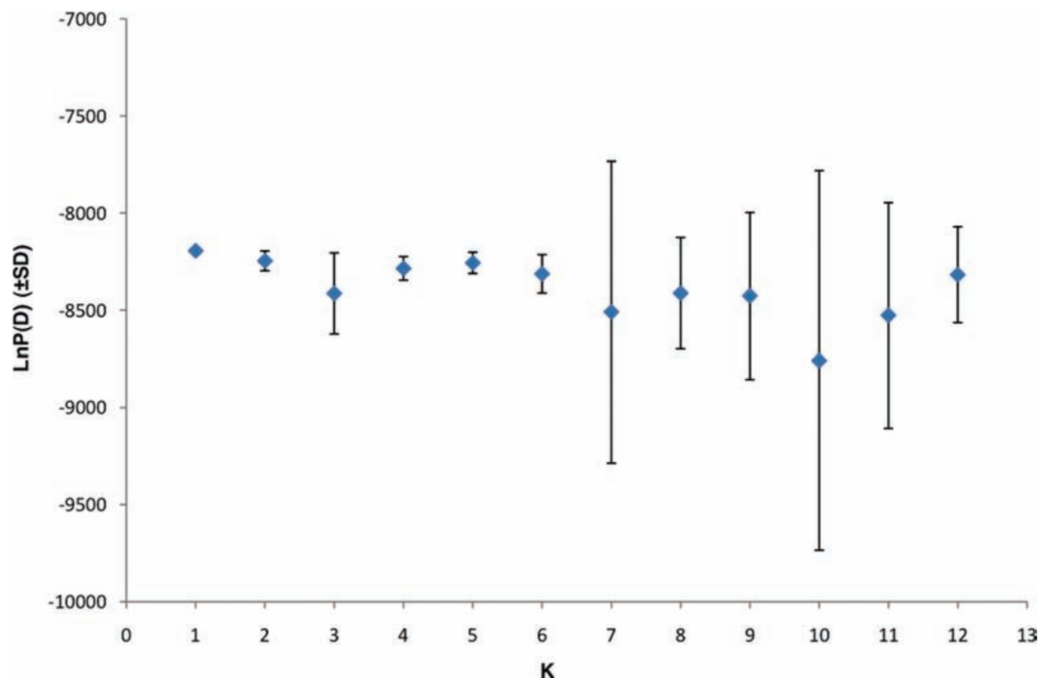

Figure 2. Mean log probability of data [ln$P(D)$] and standard deviation (SD) of 10 replicates for each value of K calculated from the multilocus genotyping data.

American Genetic Association
The Journal of Heredity
www.jhered.oxfordjournals.org
Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

American Genetic Association
The Journal of Heredity
www.jhered.oxfordjournals.org

Table 5 Haplotype (H_d) and nucleotide (π) diversity, sampling sizes (*n*), and number of haplotypes (*h*) of Cyt b and CR sequences

Sampling location	H_d	π	*n*	*h*
MB	0.075/0.935	0.0002/0.0033	53/55	3/27
LB	0.154/0.958	0.0004/0.0032	38/89	4/55
ENS	0.131/0.963	0.0003/0.0030	30/26	3/20
PP	0.270/0.971	0.0009/0.0175	28/15	5/13

diversity (π) was low and ranged from 0.00019 to 0.0009, again with the lowest values for the more northern Monterey population and the highest value for the *L. sardina* population. Overall haplotype and nucleotide diversities were 0.391 and 0.026, respectively.

Haplotype 1 was by far the most common and was represented by 114 individuals represented by each of the *L. tenuis* populations and years sampled (but not by *L. sardina*). The second most common haplotype, haplotype 9, was represented by 24 individuals, all *L. sardina*. Haplotypes 1 and 9 differed by 31 substitutions. The remaining 11 haplotypes were represented by single individuals: 2 in the MB samples, 2 in ENS, 3 from the LB samples, and the remaining 4 in *L. sardina*. These rare haplotypes differed from the nearest common haplotype by 2 substitutions at most.

A total of 1073 sites were analyzed at the mitochondrial CR, resulting in the identification of 78 polymorphic sites and 94 unique haplotypes. Hd ranged from 0.935 to 0.971 and nucleotide diversity ranged from 0.00326 to 0.01752 (Table 5). Similarly to the Cyt b results, Hd was lowest in the more northern Monterey population and highest in the Gulf of Mexico. Nucleotide diversities were similar across *L. tenuis* populations and highest in the *L. sardina* population. Overall haplotype and nucleotide diversities were 0.9514 and 0.007, respectively.

Haplotypes 4 and 18 were the most common; they were found in 35 and 12 individuals respectively and were each represented by all 3 *L. tenuis* populations. Of the remaining haplotypes, 10 were unique to *L. sardina*, 43 to LB, 12 to ENS, 18 to MB, and 9 were shared between *L. tenuis* sites (but not found in *L. sardina*). Single individuals represented 75 of the 94 haplotypes.

Pairwise K_{ST} values were highly significant between *L. sardina* samples and each population of *L. tenuis* at both the CR and the Cyt b loci ($P < 0.001$; Table 6). Values were not significant between *L. tenuis* localities.

Estimated number of migrants was high between all populations of *L. tenuis*, ranging from 17 to 20 (mean = 18) when estimated from Cyt b data, or from 17 to 57 when CR data were used (mean = 31.3). For both loci, the estimated number of migrants between *L. sardina* and each *L. tenuis* population was 0.

No defined clusters resulted from the statistical parsimony networks constructed in TCS with either Cyt b or CR sequences, except for a distinct separation of the *L. sardina* samples from all *L. tenuis* samples in each network (data not shown). The minimum spanning network described using

Table 6 Pairwise K_{ST} values between the sampling locations with years collapsed within sites

	MB	LB	ENS	PP
MB		0.0000	0.0000	0.9888***
LB	0.0055		0.0000	0.9855***
ENS	0.0022	−0.0041		0.9859***
PP	0.4962***	0.4689***	0.5119***	

CR values are below diagonal and Cyt b values are above diagonal.
***$P < 0.001$.
Overall Cyt b: K_{ST} = 0.9844***.
Overall CR: K_{ST} = 0.3908***.

Cyt b sequences in ARLEQUIN showed 2 dominant haplotypes, one which is composed only of *L. tenuis* samples and another composed solely of *L. sardina* samples. Eleven haplotypes were represented by singletons and each was just 1 substitution away from the dominant haplotype for its species. A minimum of 31 substitutions separates the sister species (Figure 3). The minimum spanning network utilizing CR sequences exhibited star-like projections of haplotypes radiating outward from several haplotypes common to each *L. tenuis* locale (Figure 3). Similarly to the Cyt b haplotype network results, no clear distinctions are seen between *L. tenuis* haplotypes between locales, but strong clustering is seen between the species with a minimum of 40 substitutions.

Discussion

Due to their high mutation rates, microsatellite markers are useful for elucidating contemporary fine-scale population structure but may be considered undesirable for larger scale estimates of structure due to high potential levels of homoplasy (Hewitt 2004a, 2004b). Frequently, studies of population structure have been based on one or more mtDNA genes (Moore 1995; Avise 2000), which are typically marked by slower mutation rates. Although these loci might prove satisfactory for describing historical events, they may lack the resolution for recent microevolutionary processes. Our current study on the grunions, genus *Leuresthes*, combines the analysis of both microsatellite genotyping and mtDNA sequencing data in an attempt to elucidate both historical and recent genetic population structuring.

Spatiotemporal Population Structure and Migration of the California Grunion

Despite large geographic distances between the MB population of *L. tenuis* and its conspecifics south of Point Conception, as well as a limited amount of tagging study data that may indicate spawning site fidelity (Walker 1952), the molecular genetic data suggest that recent migration or colonization between the populations has limited the effects of genetic drift and subsequent divergence. Results from mtDNA sequencing data analyses support 1 panmictic *L. tenuis* population. The Cyt b locus is characterized by 1 dominant haplotype that is shared between each of the 3 sampled localities, as well as a few uninformative haplotypes held by

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

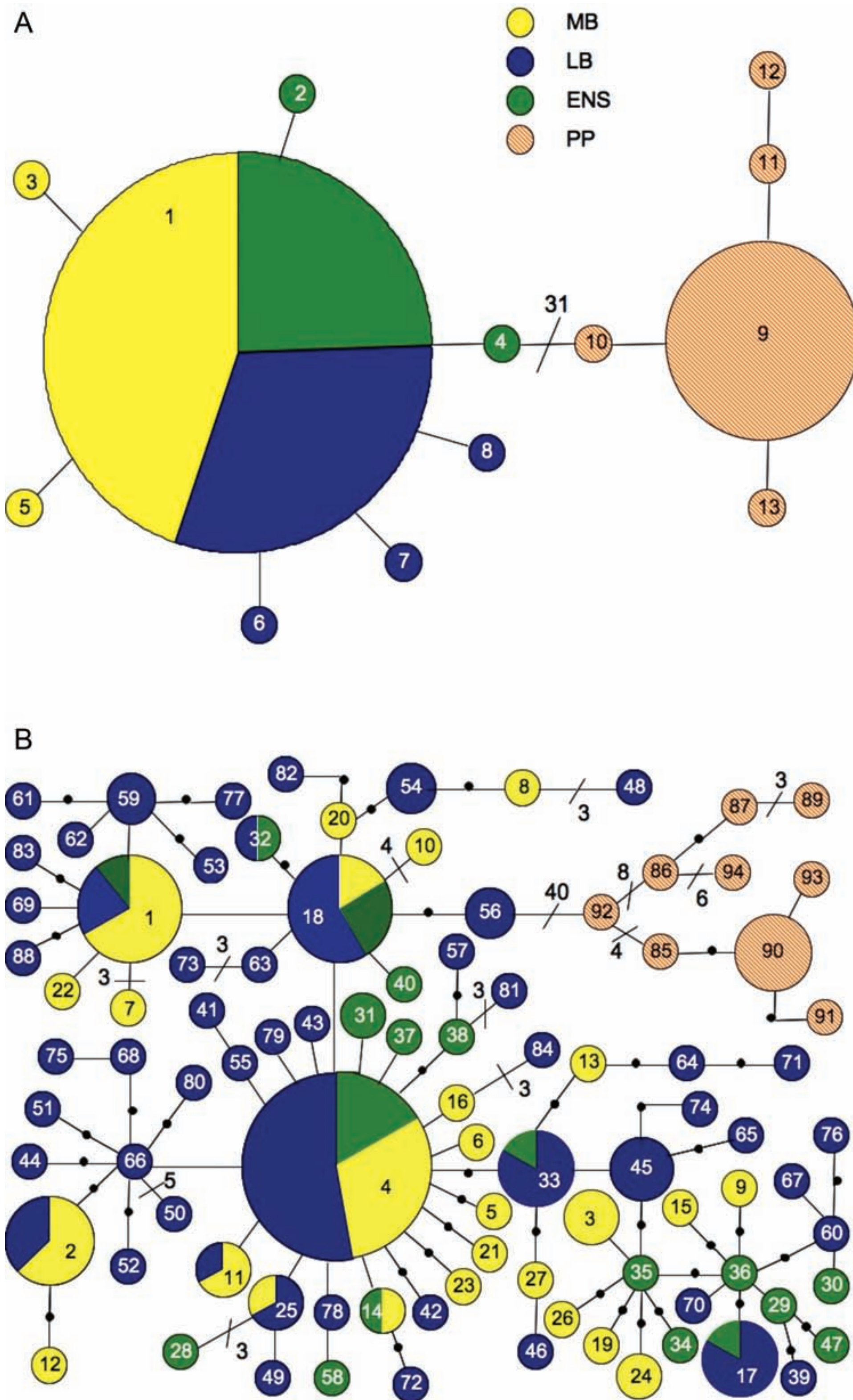

Figure 3. Minimum spanning networks of haplotypes observed in the genus *Leuresthes* from Cyt b (**a**) and CR (**b**) sequencing data (haplotype number indicated within circles). Haplotypes separated by more than one substitution are indicated by slashes and the number of substitutions. Small unnumbered circles indicate hypothetical haplotypes not observed in this study. Shading indicates locale of origin and size is proportional to the frequency of the haplotype.

American Genetic Association
The Journal of Heredity
www.jhered.oxfordjournals.org
Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

American Genetic Association

The Journal of Heredity

www.jhered.oxfordjournals.org

single individuals (Figure 3a). The more quickly evolving CR locus is marked by several shared haplotypes between populations and numerous haplotypes exhibited in one or a few individuals that do not group by geographic location in any meaningful way (Figure 3b). Pairwise K_{ST} analyses also failed to detect genetic differentiation using mitochondrial sequencing data (Table 6). In addition, estimates of the number of migrants between populations of *L. tenuis* are high.

Results from the program STRUCTURE using microsatellite data also did not reveal genetic differentiation between populations and all pairwise F_{ST} values between sampling locations and years (uncollapsed data; Table 3) were nonsignificant, indicating panmixia. Conversely, all pairwise F_{ST} values from microsatellite data between sampling locales (with multiple years at individual sites collapsed) were significant (Table 4), but low, implying limited subdivision of the populations. The observed discrepancies between these F_{ST} analyses are likely the result of increased power to detect small genetic differences in the larger sample sizes of the collapsed data sets.

Microsatellite loci are characterized by high mutation rates and are therefore typically useful in elucidating fine scale and recently derived population structure. Nonrecombining mtDNA, on the other hand, evolves more slowly and is generally appreciated for usefulness in less-contemporary historical genetic processes (Avise 2000). This general observation may help explain the somewhat contrasting results between nuclear and mtDNA data described here. The slight genetic differences implied by microsatellite F_{ST} analyses most likely are indicative of a colonization event too recent to produce observable levels of divergence in the more slowly evolving mtDNA.

These results are consistent with findings from other populations of the species using allozyme markers (Gaida et al. 2003) and DNA (Johnson et al. 2009), where low levels of genetic differentiation between populations were discovered. However, although Gaida et al. (2003) found that an isolation by distance model best supported their allozyme data analyses between 4 southern California populations, Johnson et al. (2009) found that geographic distance was not a good estimator of genetic differentiation between populations north and south of Point Conception. Our data are concordant with the latter; our microsatellite data indicate that the LB population is more similar to the MB population than it is to the less geographically distant ENS population (approximately 546 vs. 397 km). Interestingly, these data also indicate that the ENS population is almost equally genetically dissimilar from the LB population as it is from the MB population, despite a more than 2-fold difference in geographic distance (approximately 397 vs. 943 km; Table 4).

These findings may be the result of continuous or occasional migration events from the southern California populations toward the northern locales, or it may be possible that the populations are in fact genetically isolated but that colonization occurred recently enough such that not enough time has passed to result in strong genetic divergence between them. For either case, we may conclude that the molecular genetic evidence at hand do not support a genetically distinct population in MB that has been isolated from its conspecifics for a long period of time. Notwithstanding this conclusion, microsatellite results do indicate marginal molecular differences between each population sampled, which may describe the very early stages of microevolutionary divergence.

Although our findings here do not indicate strong genetic population structuring, this may be an artifact of the loci investigated. Selection that acts on some loci, but not others, can lead to heterogeneity among loci with regard to genetic differentiation. For example, Koehn et al. (1976, 1984) found little genetic differentiation between populations of the blue mussel, *Mytilus edulis*, between populations to the north and south of Cape Cod along the eastern coast of North America; however, other studies have shown a definite allele frequency cline in the aminopeptidase-1 (leucine aminopeptidase) allozyme locus (Koehn et al. 1980; Hilbish and Koehn 1985). The authors suggest that the differences in environmental salinity between populations can explain the natural selection directed on this locus. In the case of the California grunion, it is possible that the more northern MB population is affected by local selection pressures in response to the cooler waters or to differences in sand temperatures where eggs develop, which differ from those of its more southern conspecifics. Such differential selection pressures may create a heterogenic genetic environment that would not be recognized by the presumably neutral molecular markers used here.

Species Distinctions of the Grunions

The mtDNA analyses at both the Cyt b (1073 nucleotide fragment) and at the CR (392 nucleotide fragment) loci strongly support the separate species distinction of the California and Gulf grunions. The lack of any shared haplotypes, as well as the high mutational distance between the most similar haplotypes between the species, unequivocally support their separate species status. Interspecific pairwise F_{ST} analyses using microsatellite data resulted in significant, but low values, which are most likely explained by the high degree of polymorphism of the loci. This is supported by the high values of Nei's genetic distance (Table 4) and results from analysis in STRUCTURE. Also supporting these findings are previously described differences in behavior and salinity tolerance between the species (Reynolds et al. 1976), morphometric and meristic characteristics (Moffatt and Thomson 1975; Dyer and Chernoff 1996), and a limited amount of molecular genetic data (Crabtree 1987; Bernardi et al. 2003).

Both haplotype and nucleotide diversity were highest in the *L. sardina* samples in the Gulf of California, lower in the northern Baja California and southern California samples, and with one minor exception were lowest in the most northern sampling locale, MB (Table 5). These trends are consistent with a south-to-north expansion of the genus, perhaps achieved by adult movement during episodes of warm water pulses as suggested by Roberts et al. (2007), which would allow individuals to travel northward along the coast into the thermal refuges of the northern coastal bays (Johnson et al. 2009).

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

American Genetic Association

The Journal of Heredity

www.jhered.oxfordjournals.org

Conservation Implications

It is likely that both the unique reproductive behavior of the California grunion and the topographic features of the California and Baja California coastlines have strongly influenced the distribution and structure of the populations of *L. tenuis* and both may create specific challenges regarding the conservation management of these fish. Most notably, their tendency to spawn only on slowly sloping sandy beaches limits the area of spawning habitat suitable to their needs, developing embryos are negatively affected by anthropogenic alterations of these beaches (Martin et al. 2006; Matsumoto and Martin 2008), and adults are vulnerable to depredation by humans while spawning (Gregory 2001). Possibly confounding these challenges is the observation that the population in MB spawns less frequently and in smaller numbers than the populations south of Point Conception (Bernardi, personal communication), indicating small population size. These findings are consistent with those for the other northern population in San Francisco Bay (Johnson et al. 2009). We, therefore, support recommendations by Johnson et al. (2009) that areas north of Point Conception be classified as "no take" zones for California grunion throughout their entire spawning season.

Funding

University of California at Irvine and National Science Foundation (NSF grant DGE-0638751).

Acknowledgments

Animals were collected with permission from the California Department of Fish and Game with Scientific Collecting Permit ID Number SC-008834 granted to R.B.K. and SC-000189 granted to G.B. We are grateful to Hannes Baumann for providing tissue samples from Ensenada, Mexico, and to Andrey Tatarenkov and Jin Xian Liu for thoughtful review of the manuscript.

References

Addison JA, Ort BS, Mesa KA, Pogson GH. 2008. Range-wide genetic homogeneity in the California sea mussel (*Mytilus californianus*): a comparison of allozymes, nuclear DNA markers, and mitochondrial DNA sequences. Mol Ecol. 17:4222–4232.

Avise JC. 2000. Phylogeography. The history and formation of species. Cambridge (MA): Harvard University Press.

Ayres WO. 1860. Descriptions of fishes. Proc Calif Acad Sci. 1:73–77.

Bernardi G, Findley L, Rocha-Olivares A. 2003. Vicariance and dispersal across Baja California in disjunct marine fish populations. Evolution. 57:1599–1609.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P. 2004. How to track and assess genotyping errors in population genetics studies. Mol Ecol. 13:3261–3273.

Buonaccorsi VP, Westerman M, Stannard J, Kimbrell C, Lynn E, Vetter RD. 2004. Molecular genetic structure suggests limited larval dispersal in grass rockfish, *Sebastes rastrelliger*. Mar Biol. 145:779–788.

Burton RS. 1998. Intraspecific phylogeography across the Point Conception biogeographic boundary. Evolution. 52:734–745.

Byrne RJ, Avise JC. 2009. Multiple paternity and extra-group fertilizations in a natural population of California grunion (*Leuresthes tenuis*), a beach-spawning marine fish. Mar Biol. 156:1681–1690.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. Mol Ecol. 9:1657–1659.

Crabtree CB. 1987. Allozyme evidence for the phylogenetic relationships within the silverside subfamily Atheinopsinae. Copeia. 4:860–867.

Dawson MN. 2001. Phylogeography in coastal marine animals: a solution from California? J Biogeogr. 28:723–736.

Dyer BS, Chernoff B. 1996. Phylogenetic relationships among atheriniform fishes (Teleostei: Atherinomorpha). Zool J Linnean Soc. 117:1–69.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 14:2611–2620.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online. 1:47–50.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 164:1567–1587.

Gaida IH, Buth DG, Matthews SD, Snow AL, Luo SB, Kutsuna S. 2003. Allozymic variation and population structure of the California grunion, *Leuresthes tenuis* (Atheriniformes: Atherinopsidae). Copeia. 3:594–600.

Goudet J. 1995. Fstat version 1.2: a computer program to calculate F statistics. J Hered. 86:485–486.

Gregory PA. 2001. Grunion. In: Leet WS, Dewees CM, Klingbeill R, Larson EJ, editors. California's living marine resources: a status report. Sacramento (CA): California Department of Fish and Game. p. 246–247.

Gruber D, Ahlstrom EH, Mullin MM. 1982 Distribution of ichthyoplankton in the Southern California Bight. Cal Coop Ocean Fish. 23:172–179.

Hewitt GM. 2004a. Genetic consequences of climatic oscillations in the Quaternary. Philos Trans R Soc Lond, B, Biol Sci. 359:183–95; discussion 195.

Hewitt GM. 2004b. The structure of biodiversity—insights from molecular phylogeography. Front Zool. 1:4.

Hilbish TJ, Koehn RK. 1985. Dominance in physiological phenotypes and fitness at an enzyme locus. Science. 229:52–54.

Jahn A. 2004. Summertime distribution of three species of atherinopsid fishes in east-central San Francisco Bay. Bull South Calif Acad Sci. 103:34–35.

Jenkins OP, Evermann BW. 1889. Description of eighteen new species of fishes from the Gulf of California. Proc U. S. Nat Mus. 11:137–158.

Johnson PB, Martin KL, Vandergon TL, Honeycutt RL, Burton RS, Fry A. 2009. Microsatellite and mitochondrial genetic comparisons between northern and southern populations of California grunion (*Leuresthes tenuis*). Copeia. 2009:467–476.

Koehn RK, Hall JG, Innes DJ, Zera AJ. 1984. Genetic differentiation of *Mytilus edulis* in eastern North America. Mar Biol. 79:117–126.

Koehn RK, Milkman R, Mitton JB. 1976. Population genetics of marine pelecypods. IV. Selection, migration and genetic differentiation in the blue mussel *Mytilus edulis*. Evolution. 30:2–32.

Koehn RK, Newell RI, Immermann F. 1980. Maintenance of an aminopeptidase allele frequency cline by natural selection. Proc Natl Acad Sci USA. 77:5385–5389.

Lewallen EA, Anderson TW, Bohonak AJ. 2007. Genetic structure of leopard shark (*Triakis semifasciata*) populations in California waters. Mar Biol. 152:599–609.

Lewis PO, Zaykin D. 2001. Genetic Data Analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). Available from: http://lewis.eeb.uconn.edu/lewishome.software.html

Martin K, Speer-Blank T, Pommerening R, Flanner J, Carpenter K. 2006. Does beach grooming harm grunion eggs? Shore Beach. 74:17–22.

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013

American Genetic Association

Matsumoto JK, Martin KLM. 2008. Lethal and sublethal effects of altered sand salinity on embryos of beach-spawning California grunion. Copeia. 2008:483–490.

McGowan GE. 1993. Coastal ichthyofaunal assemblages, with emphasis on the Southern California Bight. B Mar Sci. 53:692–722.

Miller DJ, Lea RN. 1972. Guide to the coastal marine fishes of California. Fish B—Calif Fish Game. 157:1–235.

Moffatt NM, Thomson DA. 1975. Taxonomic status of the Gulf grunion (*Leuresthes sardina*) and its relationship to the Californa grunion (*L. tenuis*). Trans San Diego Soc Nat Hist. 18:75–84.

Moore WS. 1995. Inferring phylogenies from mtDNA variation—mitochondrial-gene trees versus nuclear-gene trees. Evolution. 49:718–726.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 89:583–590.

Palumbi SR. 1996. Nucleic acids II: the polymerase chain reaction. In: Hillis DM, Moritz C, Mable BK, editors. Molecular systematics. Sunderland (MA): Sinauer & Associates. p. 205–247.

Phillips JB. 1943. Grunion in Monterey Bay. Calif Fish Game. 29:82.

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

Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered. 86:248–249.

Reynolds WW, Thomson DA, Casterlin ME. 1976. Temperature and salinity tolerances of larval California grunion, *Leuresthes tenuis*, (Ayres): a comparison with Gulf grunion, *L. sardina* (Jenkins & Evermann). J Exp Mar Biol Ecol. 24:73–82.

Rice WR. 1989. Analyzing tables of statistical tests. Evolution. 43:223–225.

Roberts D, Lea RN, Martin KLM. 2007. First record of the occurrence of the California grunion, *Leuresthes tenuis*, in Tomales Bay, California; a northern extension of the species. Calif Fish Game. 93:107–110.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 15:174–175.

Ryman N, Palm S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Mol Ecol. 6:600–602.

Schwartz MK, Luikart G, Waples RS. 2007. Genetic monitoring as a promising tool for conservation and management. Trends Ecol Evol (Amst). 22:25–33.

Sivasundar A, Palumbi SR. 2010. Life history, ecology and the biogeography of strong genetic breaks among 15 species of Pacific rockfish, *Sebastes*. Mar Biol. 157:1433–1452.

Straughan D, 1982. Inventory of the natural resources of sandy beaches in Southern California. Tech Rep Allan Hancock Found. 6:312–328.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics. 132:619–633.

Thompson WF. 1919. The spawning of the grunion (*Leuresthes tenuis*). Fish B—Calif Fish Game. 3:1–29.

Van Oosterhout C, Hutchinson WF, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes. 4:535–538.

Walker BW. 1952. A guide to the grunion. California Fish Game. 38:409–420.

Weir BS, Cockerham CC. 1984. Estimating F-Statistics for the analysis of population structure. Evolution. 38:1358–1370.

Yoklavich MA, Cailliet GM, Oxman DS, Barry JP, Lindquist DC. 2002. Fishes. In: Caffrey J, Brown M, Tyler WB, Silberstein M, editors. Changes in a California estuary: a profile of Elkhorn Slough. Moss Landing (CA): Elkhorn Slough Foundation. p. 163–185.

Received June 29, 2012; First decision August 19, 2012; Accepted March 18, 2013

Corresponding Editor: Jose V. Lopez

The Journal of Heredity

www.jhered.oxfordjournals.org

Downloaded from http://jhered.oxfordjournals.org/ at University of California, Santa Cruz on August 13, 2013