Recent population expansion in the evolutionary history of the Californian anchovy *Engraulis mordax*

Expansión poblacional reciente en la historia evolutiva de la anchoveta de California *Engraulis mordax*

Noé Díaz-Viloria,[1] Laura Sánchez-Velasco[1] and Ricardo Pérez-Enríquez[2]

[1]Departamento de Plancton y Ecología Pesquera, Centro Interdisciplinario de Ciencias Marinas-Instituto Politécnico Nacional (CICIMAR), Av. Instituto Politécnico Nacional s/n, Col. Playa Palo de Santa Rita, La Paz, B.C.S. 23096, México
[2]Laboratorio de Genética Acuícola, Centro de Investigaciones Biológicas del Noroeste (CIBNOR), Mar Bermejo 195, Col. Playa Palo de Santa Rita, La Paz, B.C.S. 23096, México
e-mail: rperez@cibnor.mx

Díaz-Viloria N., L. Sánchez-Velasco and R. Pérez-Enríquez. 2012. Recent population expansion in the evolutionary history of the Californian anchovy *Engraulis mordax. Hidrobiológica* 22 (3): 258-266.

ABSTRACT

The Californian anchovy *Engraulis mordax*, a temperate species, may have undergone a process of population disjunction from experiencing post-glacial water heating processes around the tip of the Baja California Peninsula, Mexico about 10,000 b.p. A genetic analysis was performed to test the null hypothesis of genetic homogeneity between the Gulf of California and Southern California, U. S. A., and if this is the case, to estimate the time of haplotype emergence in terms of coalescence. A total of 80 sequences of the mtDNA hypervariable control region of *E. mordax* captured in the central Gulf of California (n = 40) and Southern California (n = 40) were analyzed. In spite of the large number of private haplotypes, no significant genetic differentiation among sites (F_{ST} = –0.0025, p = 0.686) was observed. An unimodal distribution of mismatch frequency between haplotypes indicated a model of rapid expansion in population size that, based on a mutation rate of 3.6% per million years in the control region, indicates a relatively recent nucleotide differentiation time of approximately 61,000 years. This time period corresponds to the late Pleistocene, suggesting population expansions at each locality, followed by the last episode of glaciation, which may have contributed to migration of this temperate-affinity species between two locations and the genetic homogenization. However this unique recent event of gene flow in the evolutionary history of species does not explain by itself the mismatch distribution patterns found.

Key words: Control region, gene flow, mitochondrial DNA, molecular clock, recent population expansion.

RESUMEN

La anchoveta de California *Engraulis mordax*, es una especie templada que pudo haber pasado por un proceso de disyunción poblacional, debido al proceso postglacial de calentamiento del agua alrededor de la punta de la península de Baja California, hace unos 10,000 años. Se realizó un análisis genético para probar la hipótesis nula de homogeneidad genética entre el Golfo de California, México y el sur de California, EUA y si este era el caso, estimar el tiempo de surgimiento de haplotipos en términos de coalescencia. Se analizaron en total 80 secuencias de la región control hipervariable (ADNmt) de *E. mordax*, capturadas en la región central del Golfo de California (n = 40) y el sur de California (n = 40). A pesar del gran número de haplotipos únicos, no se observó diferenciación genética significativa entre localidades (F_{ST} = –0.0025, p = 0.686). Una distribución unimodal en la frecuencia del número de diferencias entre haplotipos indica un modelo de expansión rápida en el tamaño poblacional, que basado en una tasa mutacional de

3.6% por millón de años para la región control, indicó un tiempo de diferenciación nucleotídica relativamente reciente de aproximadamente 61,000 años. Este periodo de tiempo corresponde al Pleistoceno tardío, después de la formación de la península de Baja California, sugiriendo expansiones poblacionales en cada una de las localidades, seguidas del último episodio de glaciación, el cual quizás contribuyó a la migración de esta especie de afinidad templada entre las dos localidades y a su homogenización genética. Sin embargo este único evento reciente de flujo genético en la historia evolutiva de la especie, no explica por sí solo los patrones de distribución encontrados en las frecuencias de diferencias nucleotídicas.

Palabras clave: ADN mitocondrial, expansión poblacional reciente, flujo genético, región control, reloj molecular.

INTRODUCTION

The Gulf of California is a suggestive model system for studying mechanisms of speciation of marine species by the combination of a relatively large number of endemic species and its geological history (Huang & Bernardi, 2001). The uppermost region of the Gulf of California, Mexico was formed as early as the Miocene, ~25 million years ago. The Baja California Peninsula was then restricted to a small area in the north, and the southern region of the peninsula was composed of several small and large islands, between which several natural seaways connected the Pacific Ocean to the Gulf of California. This situation lasted until ~1 million years ago when the northern peninsula and the southern islands were joined and uplifted to create the present-day Baja California Peninsula (Terry *et al*., 2000; Huang & Bernardi, 2001; Stepien *et al*., 2001). Populations of temperate fish species are thought to have been established either by migration, via extinct waterways connecting the Gulf of California to the Pacific Ocean or by migration around the southern tip of Baja California during periods of oceanic cooling associated with glaciating events (Brusca, 1973; Huang & Bernardi, 2001; Bernardi *et al*., 2003). Presently, gene flow between Pacific and Gulf populations may be limited by physical and physiological barriers (Huang & Bernardi, 2001). If this is the case, then isolation between marine populations in the Gulf of California and the Pacific, based on previous studies, should have occurred between 120,000 and 2.3 million years ago, as suggested for *Girella nigricans* Ayres 1860 (Terry *et al*., 2000), *Gillichthys mirabilis* Cooper 1864 (Huang & Bernardi, 2001), and *Paralabrax maculatofasciatus* Steindachner 1862 (Stepien *et al*., 2001).

Among the marine species that may have undergone a population vicariance or disjunction in this region is the Californian anchovy *Engraulis mordax* Girard, 1854 that is a species of temperate affinity which can be found from Vancouver Island, Canada to the southern part of the Baja California Peninsula (Whitehead *et al*., 1988), including the Gulf of California (Hammann & Cisneros-Mata, 1989; Holmgren-Urba & Baumgartner, 1993; Green-Ruiz & Hinojosa-Corona, 1997; Green-Ruiz & Cotero-Altamirano, 2009).

According to genetic variation analysis of cytochrome *b* of mitochondrial DNA (mtDNA), *E. mordax* began population expansion 290,000 years ago, and, even though a large percentage of exclusive haplotypes were found at several locations, no significant evidence of disjunction between the subpopulations inside and outside the Gulf of California were reported (Lecomte *et al*., 2004). Considering that the mtDNA control region shows larger variation than cytochrome *b* (Rocha-Olivares *et al*., 1999), it might be a suitable marker not only for genetic differentiation, but to provide an additional estimate on the time of population expansion.

The purpose of this study was to test the null hypothesis of genetic homogeneity between the Gulf of California and Southern California and to reconstruct the evolutionary history of the Californian anchovy *E. mordax* using the mtDNA hypervariable control region.

MATERIALS AND METHODS

Samples of adult *E. mordax* were collected at two sites (n = 40 each), at San Rafael in the central Gulf of California, Mexico (28°30′N,113°01′W) in autumn 2001 and in San Diego, in the southern region of California, U. S. A (~32°79′N,~117°36′W) in autumn-winter 2003 (Fig. 1). Both samples were collected through purse seine vessels. Muscle samples were preserved in 70% ethanol.

DNA was extracted (Taggart *et al*., 1992). Concentrations were standardized to 0.05 µg/µL before performing PCR. A fragment of 551 base pairs (bp) of the control region of mtDNA was amplified with primers: FCR2 (5′-ATTTCTGGCCTCTGGTTCCT-3′) and RCR2 (5′-CATCTTCAGTGCTATGCTTTG-3′) designed for the species (Díaz-Viloria *et al*., 2005). DNA amplifications were performed in a volume of 25 µL containing 1 µL diluted DNA (0.05 µg/µL), 0.48 µM of each primer, 80 µM of each dNTP, 0.025 units Taq polymerase (Invitrogen) for Taq buffer 1× (Invitrogen) and 3.5 mM $MgCl_2$. PCR thermal conditions (Progene Thermal Cycler, Techne) were: 2 min at 94 °C, 30 cycles of 1 min at 94 °C, 1 min at 56 °C, 2 min at 72 °C, and a final extension of 4 min at 72 °C.

The PCR products were separated by size on agarose gels (1%) and visualized with SybrGold™. Products with good quality and quantity were cut from the gel, purified (QIAGEN Gel Extraction Kit™) and sequenced in both directions in an automatic sequencer ABI Prism 3730XL (Macrogen, Korea), obtaining two sequences for each individual and a total of 160 sequences for the two locations.

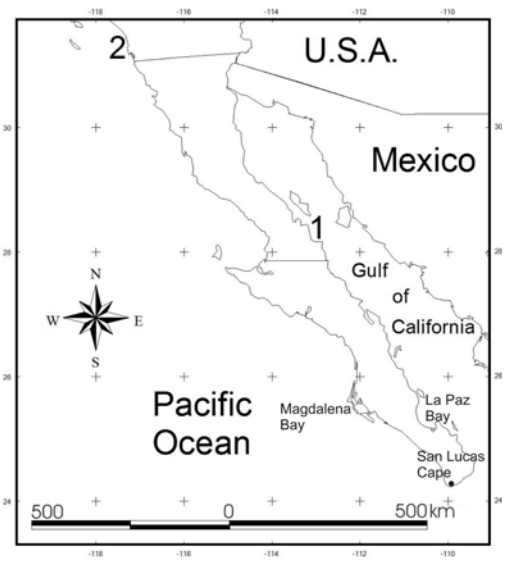

Figure 1. Sampling sites of *Engraulis mordax*. 1. San Rafael in the central Gulf of California, Mexico; 2. offshore San Diego (Southern California), USA.

The individual sequences were aligned and checked carefully with CHROMASPRO, vers. 1.34 (McCarty, 1998). The consensus sequences were obtained by reconciling the sequences of the two DNA strands of PCR products that were independently sequenced. The consensus sequences were aligned using Clustal W, vers. 1.83 (Thompson *et al.*, 1994). The sequences of 454 bp of the resulting haplotypes were deposited in GenBank (GU136650-GU136673).

Haplotype frequencies were obtained for each location. Genetic diversity was assessed by haplotype diversity (h), nucleotide diversity (π) and nucleotide polymorphism (θ).Tajima's test was performed using DnaSP, vers. 5.10. (Rozas *et al.*, 2003) to determine if mutations are selectively neutral (Tajima, 1989). Population differentiation was assessed between two locations by F_{ST} (Weir & Cockerham, 1984), based on haplotype frequencies. Significance of population differentiation was assessed by exact test (100,000 steps in Markov chain, 10,000 dememorization steps), and a global AMOVA considering one group of two populations and 50,175 permutations with ARLEQUIN, vers. 3.5 (Excoffier *et al.*, 2005). The gene flow estimate or number of migrant females (Nm_f) from DNA sequences was obtained with Equations 3 and 4 from Hudson *et al.* (1992) with DnaSP, vers. 5.10. A parsimony haplotype network using NETWORK, vers. 4.6.0.0 (Bandelt *et al.*, 1999) was obtained with the control region sequence of *Engraulis japonicus* Temminck & Schlegel, 1846 (GenBank AB040676) as outgroup. We estimated the time of differentiation between haplotypes (by nucleotide substitution) on a molecular clock of 3.6% per million years (Donaldson & Wilson, 1999). Demographic parameters (τ, θ_0, and θ_1) from mismatch distributions and fit to a model of rapid population expansion (Rogers & Harpending, 1992) were also obtained with ARLEQUIN, vers. 3.5 (Excoffier *et al.*, 2005). Values of τ from each site were used to estimate the expansion time (T) in the Gulf and Southern California, with the equation $T = \tau / 2u$ generations (Rogers & Harpending, 1992), where u was estimated with the equation $u = 2\mu k$ (Nei & Tajima, 1981), where μ is the mutation rate per nucleotide and k is the number of nucleotides that were covered in the data (454 bp). The divergence rate per nucleotide of 7.2×10^{-8} per generation (2μ) was obtained considering 3.6% per million years (3.6×10^{-8}) multiplied by two, the generation's time during which every *E. mordax* individual is sexually mature (Kucas, 1986). Finally, the population size of females before and after population expansion were obtained with the equation $N_t = \theta_i / 2u$ (Rogers & Harpending, 1992).

RESULTS

The fragment of the mtDNA control region of 80 individuals resulted in 21 variable sites that defined 24 haplotypes, of which 20 (83%) were unique haplotypes for one or another location. The three most frequent haplotypes were shared with similar frequencies between the two locations (Table 1). Tajima's test showed that all mutations were selectively neutral ($p > 0.05$).

Despite the large number of private haplotypes for each locality and the slightly higher diversity in the Gulf of California (Table 2), genetic differentiation between localities was not significant ($F_{ST} = -0.0025$, $p = 0.686$). The AMOVA test showed the highest statistical variation in component within subpopulations (100%) with a fixation index not significantly different from zero ($F_{ST} = -0.0025$, $p = 0.507$). A high estimate of gene flow ($Nm_f = 95$) was obtained between the two locations. The parsimony network showed a star-shaped phylogeny from the main haplotype (HAP6), from which several haplotypes arise, separated by one or two mutational steps (Fig. 2). The haplotype 18 (HAP18) also showed a star topology, but with a higher number of mutational steps (2-4) between this and the haplotypes that are derived from it. The relationship of HAP22 and the group associated to it (HAP18-HAP24) with the species *E. japonicus* Temminck & Schlegel 1846 indicates that such lineages are more ancestral than HAP6 (Fig. 2).

Taking into account a mutational rate of the control region of 3.6% per million years, it is probable that the appearance of haplotypes from HAP6, estimated at 0.22% (one variable site within the 454 bp fragment of the control region) took place ~61,111 (95% CI between 30,254 and 91,968) years ago.

The mismatch distribution was unimodal for the total sample and at both locations, and the observed values fitted simulated values ($p > 0.7$), indicating rapid expansion in population size, however, slight differences were observed in the modes of mismatch distributions between locations; two pairwise differences were the most frequent in the Gulf of California and one in Southern California (Fig. 3). The estimate of the time of expansion from historical demographic parameters (τ, θ_0, and θ_1) showed consis-

Table 1. Frequencies and variable sites of each haplotype (H) of the control region fragment of 454 bp of *Engraulis mordax* in the central Gulf of California (GC) and offshore Southern California (SC). The number of variable sites representing the nucleotides and their positions in the sequence. A double hyphen represents a deletion and a point represents the same nucleotide as in H6.

					Variable sites																				
H	GC		SC		58	63	81	110	174	181	187	189	232	245	263	271	274	275	304	305	308	309	334	371	396
	Freq	%	Freq	%																					
6	15	37.5	20	50.0	C	G	T	T	A	G	C	A	G	C	T	T	T	A	T	G	T	T	A	G	C
18	7	17.5	4	10.0	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	T
2	5	12.5	6	15.0	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·
5	1	2.5	2	5.0	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·
1	1	2.5	—	—	—	—	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·
3	1	2.5	—	—	·	·	·	·	G	·	·	·	·	·	·	C	·	·	·	·	·	·	·	·	·
4	—	—	1	2.5	·	·	·	·	G	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
7	—	—	1	2.5	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
8	1	2.5	—	—	·	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	·	·	·	·
9	—	—	1	2.5	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·
10	—	—	1	2.5	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·
11	1	2.5	—	—	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
12	—	—	1	2.5	·	·	·	C	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·	T	·
13	1	2.5	—	—	·	·	·	·	·	·	·	·	·	·	·	·	·	·	C	·	·	·	·	·	·
14	1	2.5	—	—	·	·	·	·	·	·	·	·	·	·	·	·	·	G	·	·	·	·	·	·	·
15	1	2.5	—	—	·	·	·	·	·	·	·	·	A	·	·	·	·	G	·	·	·	·	·	·	·
16	1	2.5	—	—	·	·	·	·	·	·	·	·	A	·	·	—	·	·	·	·	·	·	·	·	·
17	—	—	1	2.5	·	·	·	·	·	·	·	·	A	·	·	·	·	·	·	·	·	·	·	·	·
19	1	2.5	—	—	·	·	·	C	·	·	·	·	·	T	C	·	—	·	·	·	·	·	·	·	T
20	—	—	1	2.5	—	—	·	C	·	·	·	·	·	T	·	·	·	·	·	·	·	·	·	·	T
21	—	—	1	2.5	·	·	·	·	·	·	·	·	·	T	·	·	·	·	·	·	·	C	·	T	T
22	1	2.5	—	—	·	·	·	·	·	·	·	·	·	T	·	·	·	G	·	·	·	·	·	·	T
23	1	2.5	—	—	·	·	·	·	·	·	·	·	A	T	·	·	·	·	—	—	·	·	·	·	T
24	1	2.5	—	—	·	·	·	·	·	A	·	·	·	T	·	C	·	·	·	·	·	·	·	·	T

tency with the mutation rate estimate, with values of 43,441 years in Southern California and of 73,911 years in the Gulf of California. The estimate of time of expansion under assumption of two samples as a total population showed a value of 43,074 years. The estimate of female population size before rapid population growth was 9,621, and after rapid population growth, was 107,042. Calculated values are shown in Table 2.

DISCUSSION

Population Genetics. While the lack of significant genetic differentiation between the central Gulf of California and Southern California contrasted with results based on genetic markers of low variability, such as allozymes and proteins (Vrooman *et al.*, 1981; Hedgecock *et al.*, 1989; Hedgecock *et al.*, 1994; Díaz-Jaimes *et al.*, 1999), we confirm the results obtained by Lecomte *et al.* (2004) who used the mtDNA cytochrome *b* fraction.

The phylogeographic pattern observed in *E. mordax* between the Gulf of California and Southern California may result from recent historical or ongoing dispersal, with high levels of gene flow. In the case of other fish species, the topology without a separation of Pacific and Gulf of California populations into distinct clades, had already been reported in *Hermosilla azurea* Jenkins & Everman, 1889, *Halichoeres semicinctus* Ayres, 1859, *Semicossyphus pulcher* Ayres, 1854, *Sebastes macdonaldi* Eigenmann & Beeson, 1893 (Bernardi *et al.*, 2003), and has been attributed to recent historical or ongoing gene flow.

Phylogeography and Historical Gene Flow. The high haplotype diversity (h) and low nucleotide diversity (π) in *E. mordax* cor-

Table 2. Genetic diversity estimates in control region fraction (mtDNA) from *Engraulis mordax* in the localities of collection. Symbols are: sample size (*n*), percentage of private haplotypes (*ph*%), haplotype diversity (*h*), nucleotide diversity (π), nucleotide polymorphism (θ), Tajima's *D*-value (*D*) and *P*-value of Tajima's *D*-value (*P*), mutational time before present (τ), expected pair-wise differences before (θ_0) and after (θ_1) population growth, expansion time in generations (*T*) and years (*T'*) estimated with mutation rates of 4% and 6% per million years, and female population size before (N_{f0}), and after (N_{f1}) population expansion.

	Central Gulf of California	Southern California	Total
n	40	40	80
Haplotypes	16	12	24
Private haplotypes (*ph%*)	12 (75%)	8 (67%)	20 (83%)
h	0.826	0.728	0.776
π	0.003	0.002	0.003
θ	0.005	0.006	0.008
D	–1.143	–1.577	–1.638
P	>0.1	>0.05	$0.10 > p > 0.05$
Historic demographic parameters			
τ	2.416	1.420	1.408
θ_0	0.016	0.000	0.629
θ_1	8.129	9.954	6.998
T	36,955	21,720	21,537
T'	73,911	43,441	43,074
N_{f0}	245	0	9,621
N_{f1}	124,342	152,258	107,042

respond to hypothesis IV (phylogenetic continuity, lack of spatial separation) proposed by Avise *et al.* (1987), which proposes that species exhibiting this category of intraspecific phylogeography have had relatively extensive and recent historical interconnections through gene flow. This would require the historical absence of firm and longstanding zoogeographic barriers to movement, as well as life histories conducive to dispersal either as pelagic larvae, juveniles, and/or adults.

Low levels of nucleotide diversity (π), mismatch distribution, and star-shaped phylogeny of parsimony network indicate a rapid expansion from a small number of ancestors (Rogers & Harpending, 1992; Grant & Bowen, 1998), which is supported by the negative Tajima's *D* values, indicating that the population probably experienced a recent bottleneck (Tajima, 1989). According to the estimate of the molecular clock, population expansion may have occurred 61,111 years ago during the late Pleistocene, previous to the last episode of glaciation (Clayton *et al.*, 2006), with a maximum extension (Clark *et al.*, 2009). In this scenario, the rapid population growth would favor retention of new mutations and would be reflected in populations containing one or two predominant haplotypes within a cluster of haplotypes that are one or a few mutations from the central haplotype (Grant & Bowen, 1998). In *E. mordax*, haplotypes HAP6, HAP2, and HAP18 were the most successful after the bottleneck, and these haplotypes are surrounded by a large number of unique haplotypes. The HAP22 and the group associated with it (H18–H24), despite being the oldest lineages, were less successful than HAP6.

Also, the estimates of expansion time would permit hypothesizing that previous of the last glaciations, *E. mordax* showed rapid population growth in the Gulf of California 73,911 years ago and in the Southern California 43,441 years ago. During the last glaciations (25,000 years b.p.; Clayton *et al.*, 2006), environmental conditions probably contributed to the dispersal of *E. mordax* between two regions, improving the levels of gene flow and genetic homogeneity of frequent haplotypes.

The time of coalescence calculated here contrast with the ~290,000 years (95% CI between 144,000 and 447,000) obtained by Lecomte *et al.* (2004) for the same species, one half of an order of magnitude greater. Thus, the estimates of control region sequence data suggest a more recent event of rapid population growth ~61,000 years (95% CI between 30,254 and 91,968) before the present.

About recent historical gene flow, phylogenetic reconstruction (data not shown) and mismatch distribution pattern suggested that last glaciations hypothesis might not be the only event that explains gene flow in *E. mordax* between Gulf and Southern California. In previous studies where gene flow has been reestablished after divergence of populations into distinct clades, a bimodal pattern of mismatch distribution has been reported (Viñas *et al.*, 2004; Magoulas *et al.*, 2006). In *E. mordax* such bimodal mismatch distribution was not evident when both samples were analyzed as a total sample. However, differences in the mode of the number of pairwise differences of mismatch distributions between sites probably suggests a first indication of independent demographic histories, in addition to the large number of private haplotypes. But such complete divergence in populations of *E. mordax* into distinct clades was not demonstrated, probably by insufficient sample size or because the post-glacial water heating process around the tip of the Baja California Peninsula is a very recent event of population disjunction with insufficient time for lineage sorting (Avise, 2000; Freeland, 2007).

In species with high migratory potential, the genetic signal revealing population differentiation is often obscured by population admixture (Alvarado-Bremer *et al.*, 2005). Thus, alternate hypotheses to explain genetic homogeneity in *E. mordax* between the Gulf and Southern California are sporadic or ongoing gene flow.

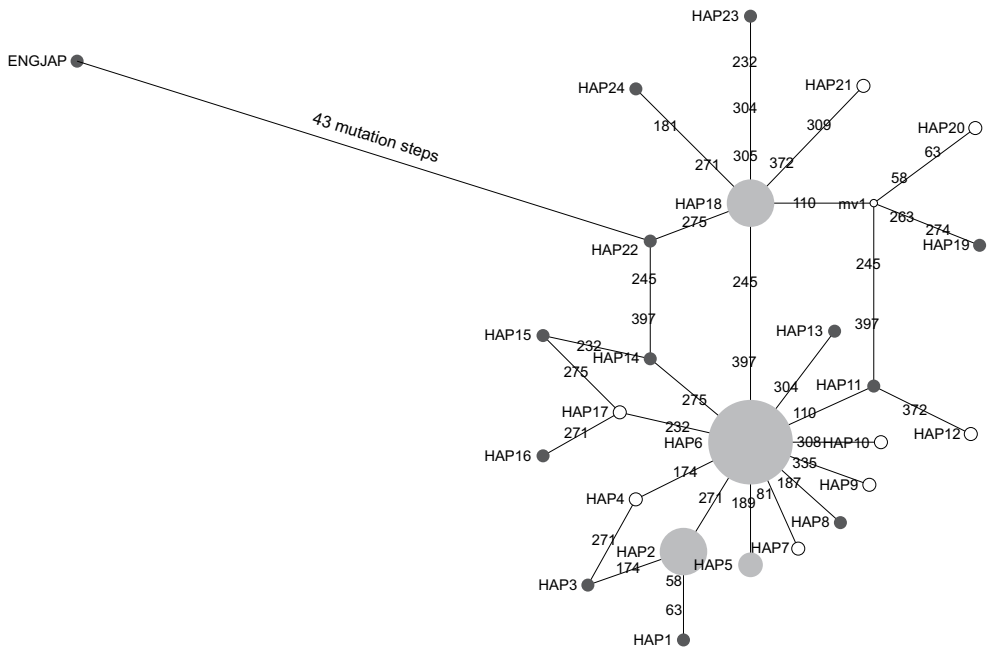

Figure 2. Network of haplotypes: haplotypes shared among sampling locations are in light gray, black are unique haplotypes (Gulf of California) or white (Southern California). ENGJAP, *Engraulis japonicus* sequence (in dark gray); HAP, Haplotype. The numbers between haplotypes represent the mutation steps; mv1 corresponds to an unsampled intermediate haplotype.

Sporadic gene flow. The presence of *E. mordax* in the Gulf of California, after 1985, probably occurred when a southern subpopulation migrated into the Gulf during anti-El Niño 1986 (Hammann & Cisneros, 1989). This suggestion implies that the Gulf of California stock is an extension of the southern subpopulation. In contrast to this idea, fish scales in anaerobic sediments of the central Gulf of California demonstrated that the anchovy was present in the gulf for at least 250 years and was abundant during the early 18th through early 19th centuries (Holmgren-Urba & Baumgartner, 1993). Those studies and the present suggest the existence of a subpopulation in the Gulf of California which could be connected with a southern subpopulation during periods of cool waters. During anti-El Niño years the southward California current turns westward off-shore typically at Magdalena Bay. At times, however its cool water extend farther south thus permitting southern movement of temperate species that are normally limited to San Lucas Cape; under extreme conditions, California current waters may even enter the Gulf of California (Robles & Marinone, 1987). Although the surface temperatures observed by Hammann & Cisneros (1989), were relatively warm (22-28 °C) for *E. mordax*, the temperature below the seasonal thermocline might have been cooler. Thus *E. mordax* appear to enter the Gulf of California during La Niña events. Since 1949 to date, there have been reported six previous intense La Niña events (http://www.esrl.noaa.gov/psd/enso/mei/#data).

Ongoing Gene Flow. Ongoing gene flow between the two regions, as described in other fish species, could be explained by two mechanisms acting together or independently. The first, eggs and larvae could be entrained around San Lucas Cape from or into the Gulf of California, and the second is by migration of adult fish via deeper water (Bernardi *et al.*, 2003).

About the first mechanism, recent zooplankton studies have shown that *E. mordax* larvae were abundant in Bahía de La Paz during February 2002 (Sánchez-Velasco *et al.*, 2006), at the central region and north of the Gulf of California, during November 1997, December 2002, and February 2006 (Avalos-García *et al.*, 2003; Peguero-Icaza *et al.*, 2008), but were absent at the mouth of Gulf of California, during November 2005 and March 2007 (León-Chávez *et al.*, 2010). In the other hand the hydrography has shown that the Gulf of California is constantly shaken by a wide spectrum of signals coming from the Pacific Ocean. Descriptions of the surface circulations in the Gulf of California suggested inflow in summer and outflow in winter. In the inner mouth there is a clear seasonal cycle, especially in temperature and in mixed layer depth, which varies from 10 m in summer to around 40 m in winter (Lavin & Marinone, 2003). All those studies point to that *E. mordax* larvae are widely distributed and abundant in the Gulf of California, from Bahía de La Paz, B.C.S. to north of Gulf of California, but it is absent in the mouth of Gulf of California, despite the existence of possible dispersal mechanism during winter. Then it might suggest no larval dispersal around San Lucas Cape, because the presence of fronts, mesoscale structures or thermal barriers (Sánchez-Velasco, pers. comm.), but this hypothesis is recently boarded.

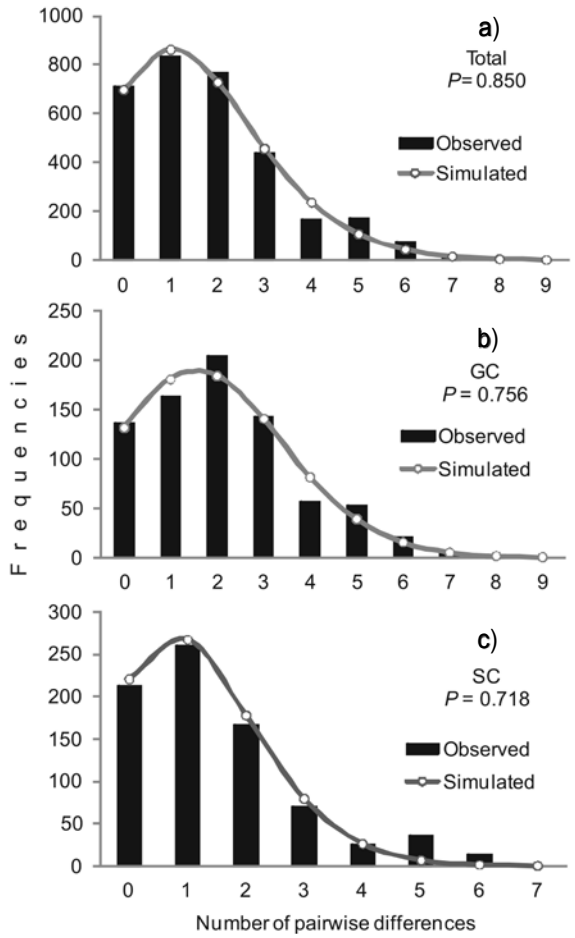

Figure 3a-c. Mismatch distribution from control region of *Engraulis mordax* for a) total (GC + SC) and two locations, b) Gulf of California and c) Southern California (SC). Bars represent observed frequencies and lines represent simulated frequencies according to the sudden expansion model. p = significance values of observed frequencies fitted to simulated frequencies. If $p < 0.05$, the hypothesis of rapid population growth is rejected.

The more plausible mechanism of ongoing gene flow is through adult fish at greater depth, perhaps through a submerged isothermal between these two regions (Rocha-Olivares *et al.*, 1999), based on the large capacity of displacement reported in *E. mordax* in the center of its distribution range (Haugen *et al.*, 1969), since it has been reported at >200 m depth as adult (Whitehead *et al.*, 1988), and because *E. mordax* spawns during spring and summer in Pacific Ocean and during autumn and winter in the Gulf of California (Nevárez-Martínez *et al.*, 2006). Under this scenario, adults of *E. mordax* could probably disperse between Pacific Ocean (the southern most subpopulation) and the Gulf of California, arrive to upwelling zones like the Midriff Archipelago region in Gulf of California (Inda-Díaz *et al.*, 2010) or Magdalena Bay (Zaytsev *et al.*, 2003), recuperate and reproduce with individuals of local population, because *E. mordax* shows different breeding seasons in Gulf of California and Pacific Ocean. Following the model of stepping stone, genetic homogeneity could be a consequence of this mechanism through upwelling zones.

Despite the genetic homogeneity observed in *E. mordax* between the Gulf of California and the Pacific, it was not possible to distinguish between historical, sporadic or ongoing gene flow events. To test the ongoing gene flow hypothesis (ecological time scale), another sampling design and more rapidly-evolving molecular markers, such as microsatellites, would be required.

Conclusions. Subpopulations of *E. mordax* in the central Gulf of California and off Southern California probably represent two subunits that are part of a larger population, with sufficient gene flow between them. Mismatch distribution patterns suggest that recent dispersal in the evolutionary history of *E. mordax* during the last glaciation could contribute to gene flow, but is not the only hypothesis which could explain genetic homogenization, sporadic or ongoing gene flow hypotheses could contribute to a better understanding of genetic similarities of *E. mordax* at both regions and these must be evaluated.

ACKNOWLEDGMENTS

We thank M. Nevárez-Martínez (CRIP-Guaymas) and J. López-Martínez (CIBNOR-Guaymas) for providing specimens from the central Gulf of California and K. Hill (NOAA) for muscle samples from the San Diego area. Editorial services were provided by I. Fogel of CIBNOR. Thanks to two anonymous referees whose suggestions improved the manuscript. L. Sánchez-Velasco received financial support from the SIP-IPN (2009-2011) and SEP-CONACYT (2008-105922) projects. This article was partly funded by CIBNOR grant EP5.2/2003 to R. Pérez-Enríquez. N. Díaz-Viloria received a fellowship from CONACYT (No. 162710).

REFERENCES

Alvarado-Bremer, J., J. Mejuto, J. Gómez-Márquez, F. Boán, P. Carpintero, J. Rodríguez, J. Viñas, T. Greig & B. Ely. 2005. Hierarchical analyses of genetic variation of samples from breeding and feeding grounds confirm the genetic partitioning of northwest Atlantic and South Atlantic populations of swordfish (*Xiphias gladius* L). *Journal of Experimental Marine Biology and Ecology* 327: 167-182.

Ávalos-García, C., L. Sánchez-Velasco & B. Shirasago. 2003. Larval fish assemblages in the Gulf of California and their relation to hydrographic variability (Autumn 1997-Summer 1998). *Bulletin of Marine Science* 72: 63-76.

Avise, J., J. Arnold, R. Ball, E. Bermingham, T. Lamb, J. Neigel, C. Reeb & N. Saunders. 1987. Intraspecific Phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18: 489-522.

Avise, J. 2000. *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, MA. 447 p.

Bandelt, H., P. Forster & A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37-48.

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

Brusca, R. 1973. *A Handbook to the Common Intertidal Invertebrates of the Gulf of California*. University Arizona Press, Tucson. 427 p.

Clayton, L., J. Attig, D. Mickelson, M. Johnson & K. Syverson. 2006. *Glaciation of Wisconsin*. Third edition. Wisconsin Geological and Natural History Survey, Madison, Wisconsin. 4 p.

Clark, P., A. Dyke, J. Shakun, A. Carlson, J. Clark, B. Wohlfarth, J. Mitrovica, S. Hostetler & A. McCabe. 2009. The Last Glacial Maximum. *Science* 325 (5941): 710-714.

Díaz-Jaimes, P., M. Uribe-Alcocer & E. Ayala-Duval. 1999. Electrophoretic variation between the central and southern populations of the northern anchovy *Engraulis mordax* Girard 1854 (Engraulidae, Pisces) from Baja California, Mexico. *Ciencias Marinas* 25: 579-595.

Díaz-Viloria, N., L. Sánchez-Velasco & R. Perez-Enriquez. 2005. Inhibition of DNA amplification in marine fish larvae preserved in formalin. *Journal of Plankton Research* 27: 787-792.

Donaldson, K. & R. Wilson Jr. 1999. Amphi-Panamic geminates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Molecular Phylogenetics and Evolution* 13: 208-213.

Excoffier, L., G. Laval & S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary bioinformatics online* 1: 47-50. Available at: www.http://lgb.unige.ch/arlequin/

Freeland, J. 2007. *Molecular Ecology*. John Wiley & Sons Ltd., West Sussex. 388 p.

Grant, W. & B. Bowen. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons of conservation. *Journal of Heredity* 89: 415-426.

Green-Ruiz, Y. & A. Hinojosa-Corona. 1997. Study of spawning area of the Northern anchovy in the Gulf of California from 1990 to 1994, using satellite images of sea surface temperatures. *Journal of Plankton Research* 8: 957-968.

Green-Ruiz, Y. & C. Cotero-Altamirano. 2009. Spawning biomass of the northern anchovy (*Engraulis mordax*) in the Gulf of California during 1992. *Ciencia Pesquera* 17: 27-36.

Hammann, M. & M. Cisneros-Mata. 1989. Range extension and commercial capture of the northern anchovy, *Engraulis mordax* Girard, in the Gulf of California, Mexico. *California Fish and Game* 75: 49-53.

Haugen, C., J. Messersmith & R. Wickwire. 1969. Progress report on anchovy tagging off California and Baja California, March 1966 through May 1969. *California Department of Fish and Game, Fishery Bulletin* 147: 75-89.

Hedgecock, D., E. Hutchinson, G. Li, F. Sly & K. Nelson. 1989. Genetic and morphometric variation in the pacific sardine, *Sardinops sagax caerulea*: comparisons and contrasts with historical data and with variability in the northern anchovy, *Engraulis mordax*. *Fishery Bulletin* 87: 653-671.

Hedgecock, D., E. Hutchinson, G. Li, F. Sly & K. Nelson. 1994. The central stock of northern anchovy (*Engraulis mordax*) is not a randomly mating population. *California Cooperative Oceanic Fisheries Investigations Reports* 35: 121-136.

Holmgren-Urba, D. & T. Baumgartner. 1993. A 250-year history of pelagic fish abundances from the anaerobic sediments of the central Gulf of California. *California Cooperative Oceanic Fisheries Investigations Reports* 34: 60-68.

Huang, D. & G. Bernardi. 2001. Disjunct Sea of Cortez-Pacific Ocean *Gillichthys mirabilis* populations and the evolutionary origin of their Sea of Cortez endemic relative *Gillichthys seta*. *Marine Biology* 138: 421-428.

Hudson, R., M. Slatkin & W. Maddison. 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics* 132: 583-589.

Inda-Díaz, E., L. Sánchez-Velasco & M. Lavín. 2010. Three-dimensional distribution of small pelagic fish larvae (*Sardinops sagax* and *Engraulis mordax*) in a tidal-mixing front surrounding waters (Gulf of California). *Journal of Plankton Research* 32: 1241-1254.

Kucas, S. T. Jr. 1986. Species profiles: life histories and environmental requirements of coastal fishes and invertebrates (Pacific Southwest)-northern anchovy. *U.S. Fish and Wildlife Service Biological Report* 82 (11.50). 11 p.

Lavin, M. & S. Marinone. 2003. An overview of the physical oceanography of the Gulf of California. *In*: Velasco, O., J. Sheimbaum & J. Ochoa de la Torre (Eds.). *Nonlinear Processes in Geophysical Fluid Dynamics*. (Kluwer Academia Publishers, Dordrecht, The Netherlands, pp. 173-204.

Lecomte, F., W. Grant, J. Dodson, R. Rodríguez-Sánchez & B. Bowen. 2004. Living with uncertainty: genetic imprints of climate shifts in east pacific anchovy (*Engraulis mordax*) and sardine (*Sardinops sagax*). *Molecular Ecology* 13: 2169-2182.

León-Chavéz, C., L. Sánchez-Velasco, E. Beier, M. Lavín, V. Godínez & J. Färber-Lorda. 2010. Larval fish assemblages and circulation in the Eastern Tropical Pacific in Autumn and Winter. *Journal of Plankton Research* 32: 397-410.

Magoulas, A., R. Castilho, S. Caetano, S. Marcato & T. Patarnello. 2006. Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy

(*Engraulis encrasicolus*). *Molecular Phylogenetics and Evolution* 39: 734-746.

McCarty, C. 1998. CHROMASPRO 1.34: Free program. Available from URL: http://www.technelysium.com.au/chromas.html.

Nei, M. & F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. *Genetics* 97: 145-163.

Nevárez–Martínez, M., M. Martínez-Zavala, C. Cotero-Altamirano, M. Jacob-Cervantes, Y. Green-Ruiz, G. Gluyas-Millán, A. Cota-Villavicencio, & J. Santos-Molina. 2006. *Peces pelágicos menores. In:* Arreguín-Sánchez, F., L. Beléndez-Moreno, I. Gómez-Humarán, R. Solana-Sansores & C. Rangel-Dávalos (Eds.). *Sustentabilidad y Pesca Responsable en México*. Instituto Nacional de Pesca, Cd. de México, pp. 263-301.

Peguero-Icaza, M., L. Sánchez-Velasco, M. Lavín & S. Marinone. 2008. Larval fish assemblages, environment and circulation in a semienclosed sea (Gulf of California, Mexico). *Estuarine, Coastal and Shelf Science* 79: 277-288.

Robles, J. & S. Marinone. 1987. Seasonal and interannual thermo-haline variability in the Guaymas basin of the Gulf of California. *Continental Shelf Research* 7: 715-733.

Rocha-Olivares, A., R. Rosenblatt & D. Vetter. 1999. Molecular evolution, systematics, and zoogeography of the rockfish subgenus *Sebastomus* (*Sebastes*, Scorpaenidae) based on mitochondrial cytochrome *b* and control region sequences. *Molecular Phylogenetics and Evolution* 11 (3): 441-458.

Rogers, A. & H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9: 552-569.

Rozas, J., J. Sánchez-Del Barrio, X. Messeguer & R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 496-2497.

Sánchez-Velasco, L., E. Beier, C. Ávalos-García & M. Lavín. 2006. Larval fish assemblages and geostrophic circulation in Bahía de La Paz and the surrounding southwestern region of the Gulf of California. *Journal of Plankton Research* 28: 1-18.

Stepien, C., R. Rosenblatt & B. Bargmeyer. 2001. Phylogeography of the spotted sand bass, *Paralabrax maculatofasciatus*: divergence of Gulf of California and Pacific coast populations. *Evolution* 9: 1852-1862.

Taggart, J., R. Hynes, P. Prodhol & A. Ferguson. 1992. A simplified protocol for routine total DNA isolation from salmonid fishes. *Journal of Fish Biology* 40: 963-965.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

Terry, A., G. Bucciarelli & G. Bernardi. 2000. Restricted gene flow and incipient speciation in disjunct Pacific Ocean and Sea of Cortés populations of a reef fish species, *Girella nigricans*. *Evolution* 2: 652-659.

Thompson, J., D. Higgins & T. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673-4680.

Viña, J., J. Alvarado-Bremer & C. Pla. 2004. Phylogeography of the Atlantic bonito (*Sarda sarda*) in the northern Mediterranean: the combined effects of historical vicariance, population expansion, secondary invasion, and isolation by distance. *Molecular Phylogenetics and Evolution* 33: 32-42.

Vrooman, A., A. Paloma & R. Zweifel. 1981. Electrophoretic, morphometric and meristic studies of subpopulations of northern anchovy, *Engraulis mordax*. *California Fish and Game* 67: 39-51.

Weir, B. & C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

Whitehead, J., G. Nelson & T. Wongratana. 1988. *Clupeoid Fishes of the World (Suborder Clupeoidei): An Annotated and Illustrated Catalogue of the Herrings, Sardines, Pilchards, Sprats, Shads, Anchovies, and Wolf Herrings. Part 2-Engraulididae*. FAO Fisheries Synopsis, Rome. 579 p.

Zaytsev, O., R. Cervantes-Duarte, O. Montante & A. Gallegos-García. 2003. Coastal Upwelling Activity on the Pacific Shelf of the Baja California Peninsula. *Journal of Oceanography* 59: 489-502.

Recibido: 31 de octubre de 2011.

Aceptado: 24 de Julio de 2012.