Molecular Ecology (2005) **14**, 4085–4096 doi: 10.1111/j.1365-294X.2005.02729.x

Tempo and mode of speciation in the Baja California disjunct fish species *Anisotremus davidsonii*

GIACOMO BERNARDI and JENNIFER LAPE*
Department of Ecology and Evolutionary Biology, University of California Santa Cruz, 100 Shaffer Road, Santa Cruz, CA 95060, USA

Abstract

The Baja California region provides a natural setting for studying the early mechanisms of allopatric speciation in marine systems. Disjunct fish populations from several species that occur in the northern Gulf of California and northern Pacific coast of Baja California, but are absent from its southern shores, were previously shown to be genetically isolated, making them excellent candidates for studying allopatry. In addition, one of these species, the sargo *Anisotremus davidsonii*, has two pairs of congeneric Panamic trans-isthmian geminate species that allow for internal molecular clock calibration. Phylogeographic and demographic approaches based on mitochondrial (cytochrome *b*) and nuclear (S7 ribosomal protein) sequences showed that *A. davidsonii* entered the gulf from the south, and later colonized the Pacific coast, approximately 0.6–0.16 million years ago. Pacific coast colonization may have used a route either around the southern cape of Baja California or across the peninsula through a natural seaway. However, while several seaways have been described from different geological times, none matches the dates of population disjunction, yet much geological work remains to be done in that area. At the present time, there is no evidence for dispersal around the southern tip of the Baja California Peninsula. Signatures of incipient allopatric speciation were observed, such as the reciprocal monophyly of disjunct populations for the mitochondrial marker. However, other characteristics were lacking, such as a strong difference in divergence and coalescence times. Taken together, these results suggest that disjunct populations of *A. davidsonii* may be consistent with the earliest stages of allopatric speciation.

Keywords: allopatry, *Anisotremus*, Baja California, demography, phylogeography, speciation

Received 17 May 2005; revision received 19 July 2005; accepted 4 August 2005

Introduction

Natural population disjunctions provide an opportunity to study the early characteristics of allopatric speciation (Mayr 1963; Endler 1977; Coyne & Orr 2004; Wiens 2004). In marine systems, population disjunction may be triggered by vicariant events, such as the rise of the Isthmus of Panama, which occurred 3.1–3.5 million years ago (Ma), and resulted in large numbers of geminate species (Jordan 1908). In the absence of a land barrier, gene flow between disjunct marine populations may be limited by oceanographic features (currents, salinity, temperature) combined with reduced dispersal capabilities (Palumbi 1992, 1994; Taylor & Hellberg 2005). While the role of population disjunctions is pivotal in understanding the early stages of marine allopatric speciation, relatively few studies have focused on this topic (Avise 1994, 2000; Randall 1998; Lessios *et al*. 2001; Muss *et al*. 2001; Collin 2003). The Isthmus of Panama (Bermingham *et al*. 1997; Lessios 1998), and the Atlantic–Gulf of Mexico regions (reviewed in Avise 1992) have been thoroughly investigated and are notable exceptions. Recently, the Baja California region was shown to be an ideal system to study marine allopatric speciation (Present 1987; Terry *et al*. 2000; Huang & Bernardi 2001; Riginos & Nachman 2001; Stepien *et al*. 2001; Bernardi *et al*. 2003).

Over the past 12 million years (Myr), the Baja California Peninsula has experienced geological rearrangements

Correspondence: Giacomo Bernardi, Fax: 831 459 3383; E-mail: bernardi@biology.ucsc.edu.
*Present address: Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106, USA

including spreading, subduction, and lifting (Riddle *et al*. 2000a). This geological activity has affected the marine fauna and flora of the surrounding region, the western Pacific coast, and to the east the Gulf of California, also called Sea of Cortez (from hereon referred to as the Gulf). The geology of the Baja California Peninsula resulted in three postulated vicariant events. Approximately 4 Ma, the southern Gulf of California was established and separated from the Pacific Ocean by a small peninsula. A second vicariant event occurred approximately 3 Ma: two seaways were created in the north (northern Gulf) and in the south (Isthmus of La Paz) of a 'proto' Baja California Peninsula. Finally, approximately 1 to 1.6 Ma, a mid-peninsular seaway connected again the Gulf with the Pacific Ocean (Riddle *et al*. 2000a). These historical events shaped phylogeographic patterns in both terrestrial (Upton & Murphy 1997; Grismer 1999, 2000; Riddle *et al*. 2000a, b, c) and marine systems (Bernardi *et al*. 2003). In addition, colder water temperatures during glacial periods probably also allowed temperate Pacific species to migrate along the southern part of Baja California and enter the Gulf (Walker 1960). The combination of geological and climatic events resulted in a group of disjunct marine species that is now confined to the northern Gulf and to the northern Pacific coast of the Baja California Peninsula, but is rare or absent from the warmer southern waters of the cape region (Cabo San Lucas, Fig. 1). In the case of fishes, 19 species from 14 families were identified as disjunct (Bernardi *et al*. 2003). For these species, the Baja California Peninsula acts as a land barrier, but unlike the impassable Isthmus of Panama, dispersal around its southern end is still possible. This is a region where gene flow between geographically disjunct populations may be nil, very high, or variable in the course of time (Bernardi *et al*. 2003).

In a study based on mitochondrial DNA (mtDNA) markers, Bernardi *et al*. (2003) showed that out of 12 disjunct species, four exhibited high levels of gene flow between the Pacific Ocean and the Sea of Cortez. In contrast, eight species showed restricted gene flow with fixed differences between these regions, making them excellent candidates for studying incipient allopatric speciation. The timing of disjunction, the historical directionality of gene flow, and the mode of speciation, however, were not fully addressed in that study due to the approach used and the lack of internal molecular clock calibration. One species from that study, the sargo *Anisotremus davidsonii*, was chosen to explore these questions.

The genus *Anisotremus (*family Haemulidae, grunts) comprises 10 marine species of fishes that are found in a restricted geographic region surrounding the Isthmus of Panama: the eastern tropical Pacific (seven species) and the western tropical Atlantic (three species). In addition to being responsible for the disconnected geographic distribution of the genus, the rise of the Isthmus of Panama was the likely vicariant event that resulted in the formation of two pairs of geminate *Anisotremus* species, which in turn can be used to estimate mutation rates (Bermingham & Lessios 1993; Knowlton *et al*. 1993; Bermingham *et al*. 1997; Lessios 1998; Donaldson & Wilson 1999; Marko 2002). The Caribbean porkfish, *Anisotremus virginicus*, and the black margate, *Anisotremus surinamensis*, are the presumed sister species of the eastern tropical Pacific Panamic porkfish, *Anisotremus taeniatus* and burrito grunt, *Anisotremus interruptus*, respectively (Bermingham *et al*. 1997; Thomson

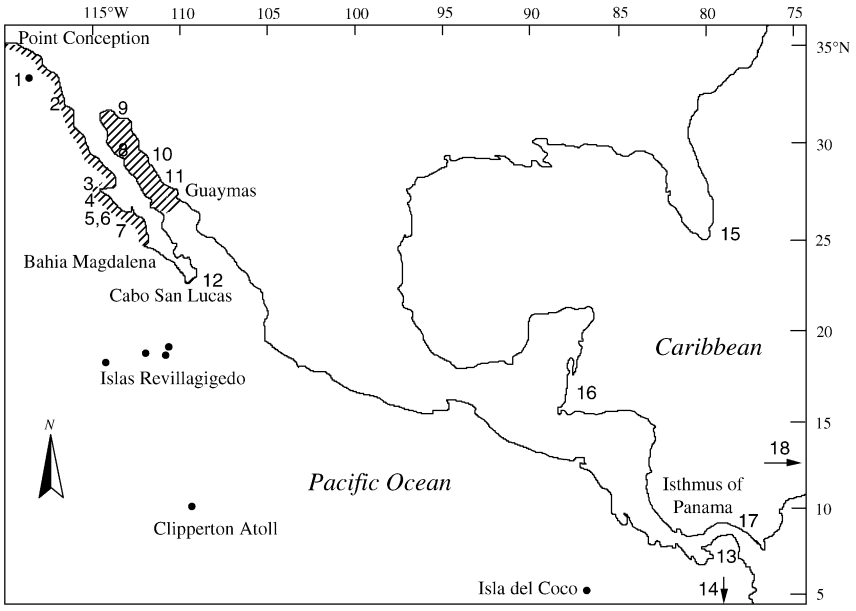

Fig. 1 Distribution map and sampling locations of *Anisotremus* and *Haemulon*. Numbers refer to sampling locations described in Table 1. Shaded area corresponds to the distribution range of the focal species *Anisotremus davidsonii*.

et al. 2000). These geminate species provide an opportunity for understanding the early processes of speciation in the closely related Baja California disjunct *A. davidsonii*.

The focal species of this work, the sargo *A. davidsonii*, is a rocky reef species that lives in shallow water down to depths of 60 m. Sargos live to approximately 15 years of age, reach sexual maturity on average at 3 years of age, and spawn April to August (Cailliet *et al*. 2000). They produce pelagic larvae that stay in the water column for approximately 40–50 days (Watson & Walker 1992, personal communication), a longer time than its congeneric warmer-water species, 15–22 days for *A. virginicus* and *A. surinamensis* (Lindeman *et al*. 2000). *A. davidsonii* is a disjunct species that is found from Point Conception, California to Magdalena Bay, on the Pacific coast of the Baja California Peninsula (Miller & Lea 1972), and in the northern Sea of Cortez (Thomson *et al*. 2000), but is absent from southern Baja California waters (Fig. 1).

In the present study, we used a combined phylogeographic and demographic approach based on mitochondrial and nuclear markers to evaluate (i) the historical fluctuations in population size and movements between disjunct populations of sargo *A. davidsonii*, (ii) the level and timing of separation of those populations, and (iii) if phylogeographic and demographic characteristics are consistent with incipient speciation models.

Materials and methods

Collections and DNA samples

Sampling sizes and locations are listed in Table 1. Six out of 10 *Anisotremus* species were obtained. These included individuals from the two *Anisotremus* geminate pairs (*Anisotremus virginicus*, *Anisotremus taeniatus*, *Anisotremus surinamensis*, *Anisotremus interruptus*), as well as Peruvian grunt individuals (*Anisotremus scapularis*) (Table 1). The closely related blue striped grunt, *Haemulon sciurus* and sailor's choice, *Haemulon parra*, collected at Turneffe Atoll, Belize, were used as outgroups. After collection, samples were immediately placed in 95% ethanol and stored at ambient temperature in the field, and then at 4 °C in the laboratory. Muscle or liver tissue was later dissected from these samples. Total genomic DNA was prepared from 75 to 150 mg of muscle or liver tissue by proteinase K digestion in lysis buffer (10 mM Tris, 400 mM NaCl, 2 mM EDTA, 1% SDS) overnight at 55 °C. This was followed by purification using chloroform extractions and alcohol precipitation (Sambrook *et al.* 1989).

PCR amplification and sequencing

Amplification of the mitochondrial cytochrome *b* region followed Bernardi *et al*. (2003). Twenty-six (out of 78) cytochrome *b* sequences were from Bernardi *et al*. (2003). Amplification of the first intron of the nuclear S7 ribosomal protein used the primers and protocols of Chow & Hazama (1998). After purification following the manufacturer's protocol (ABI, PerkinElmer), sequencing was performed in both directions with the primers used in the polymerase chain reaction (PCR) amplification on an ABI 3100 automated sequencer (Applied Biosystems). Heterozygous individuals were found to be very rare, and when present, only one allele was scored per individual.

Table 1 Collection localities for *Anisotremus* spp. and *Haemulon* spp. Columns represent the number of individuals included in the study, the locality abbreviations as in Fig. 1, and the acronyms used in Fig. 2

Species Sampling site	*n*	Fig. 1 label	Fig. 2 label
Anisotremus davidsoni			ADA
Pacific Coast	33		
California, USA			
Catalina Island	3	1	CAT
San Diego	9	2	SD
Baja California, Mexico			
Punta Eugenia	5	3	PEU
Punta San Roque	8	4	PSR
Isla San Roque	4	5	ISR
Bahia Asunsion	3	6	ASU
Bahia Tortugas	1	7	BT
Gulf of California, Mexico	45		
Bahia de Los Angeles	14	8	BLA
Punta Cholla	14	9	PC
Bahia Kino	16	10	BK
Venicia	1	11	VE
Anisotremus interruptus			AIN
Mexico, B. de Los Angeles	2	8	BLA
Anisotremus scapularis			ASC
Peru, Lima	2	14	PER
Anisotremus surinamensis			ASU
USA, Florida	3	15	FLO
Panama	1	17	PAN
Venezuela, Los Roques	1	18	VEN
Anisotremus taeniatus			ATA
Mexico, Cabo Pulmo,	2	12	CPU
Panama	2	13	PAN
Anisotremus virginicus			AVI
USA, Florida (commercial)	2	15	CA
Belize, Turneffe Atoll	5	16	BEL
Panama, San Blas Islands	3	17	PAN
Outgroups			
Haemulon parra			HPA
Belize, Turneffe Atoll	3	16	BEL
Haemulon sciurus			HSC
Belize, Turneffe Atoll	1	16	BEL

Phylogenetic analyses and population structure

We used the computer program CLUSTAL V implemented by Sequence Navigator (Applied Biosystems) to align the DNA sequences. Phylogenetic relationships were assessed by unweighted maximum-parsimony (MP), neighbour-joining (NJ), and Bayesian methods implemented by the software package PAUP (Phylogenetic Analyses Using Parsimony, version 4.0, Swofford 2003) and MRBAYES (version 2.1, Huelsenbeck & Ronquist 2001). Nucleotide sequence evolution models were evaluated using likelihood-ratio tests implemented by MODELTEST version 3.6 (Posada & Crandall 1998). Most-parsimonious trees were obtained using a branch-and-bound search. Neighbour-joining reconstructions were based on substitution models obtained with MODELTEST (HKY + G). Statistical confidence in nodes was evaluated using 2000 nonparametric bootstrap replicates (Felsenstein 1985). MRBAYES default settings for the likelihood analysis were adopted including the GTR model (unequal base frequencies and six substitution rates). Stationarity of tree likelihood, sampled every 100 cycles, was consistently achieved after 3000 generations and all sampled trees preceding stationarity were discarded. Topological differences were tested using a Shimodaira and Hasegawa test (Shimodaira & Hasegawa 1999) implemented by PAUP.

Genetic divergence and gene flow

The time of divergence between species or populations can be estimated using genetic divergence, or an estimate of coalescence time (Edwards & Beerli 2000). In both cases, mutations are assumed to proceed randomly and uniformly (molecular clock). Molecular clock enforcements were tested using a Shimodaira and Hasegawa test (Shimodaira & Hasegawa 1999) implemented by PAUP. Genetic divergence was estimated using distances based on substitution models obtained with MODELTEST (HKY + G). In order to account for polymorphism in each population (or species), divergence was estimated as the average pairwise distance between populations (or species) minus the average pairwise distance within a population (or species). Genetic divergence (as a proxy for population divergence) estimates, however, may differ from the actual time of separation between populations (see Edwards & Beerli 2000). Gene flow (F_{ST} and Nm, where N is the effective population size and m the migration rate), haplotype numbers (Hn), and haplotype diversity (Hd) were calculated using the software package DNASP version 4.0.6 (Rozas *et al*. 2003).

Historical demography

Historical demography (population fluctuations based on coalescent models) was evaluated separately for mitochondrial and nuclear markers. *A. davidsonii* populations were separated in Pacific and Sea of Cortez groups. Population specifications and sample numbers are given in Table 1 and Table 2. Population parameters Θ theta = $2N\mu$, where μ is the mutation rate for mtDNA and g (the exponential growth parameter in units of μ) were estimated using a coalescent approach with FLUCTUATE 1.4 (Kuhner *et al*. 1998). The parameter Θ was estimated with population growth (parameters are estimated jointly) or with growth kept constant (g = 0). Both estimates were obtained by running 10 replicates, which generated a mean value and its associated standard deviation. Analysis of each data set was done with 10 short Monte Carlo chains of 4000 steps each and 5 long chains of length 20 000, with a sampling increment of 20. FLUCTUATE generated a random topology for initial searching. Evidence for gene flow between Pacific Ocean and Sea of Cortez was only found in nuclear data (see Results section); thus, migration between these regions could only be analysed for nuclear data with the software MIGRATE 1.7.3 (Beerli 2003), which is a maximum-likelihood estimator based on the coalescent theory. It uses a Markov chain Monte Carlo approach to investigate possible genealogies with migration events. Analysis of each data set was done with 10 short Monte Carlo chains of 4000 steps each and 5 long chains of length 20 000, with a sampling increment of 20. For each estimate, 10 replicates were

Table 2 Historical demography of *Anisotremus davidsonii*. Columns represent the regions investigated, sample numbers, number of haplotypes, haplotype diversity, theta with no growth, theta with growth, growth, and immigration (when applicable). The latter four columns are averages of 10 replicates, between parentheses are their standard deviations. Immigration is given in *Nm*, the number of migrants per generation into the populations from the other population (e.g. there were 126 migrants from the Pacific into the Gulf)

		N	nH	HD	Theta (c)	Theta (v)	g	Immigration
All samples	mt	78	37	0.90	0.026 (± 0.002)	0.081 (± 0.027)	694.5 (± 169.1)	N/A
	nuc.	57	29	0.93	0.032 (± 0.003)	0.397 (± 0.176)	1146.2 (± 206.8)	N/A
Pacific coast	mt	33	19	0.92	0.014 (± 0.001)	0.081 (± 0.018)	1675.3 (± 323.1)	N/A
	nuc.	25	14	0.90	0.016 (± 0.001)	0.282 (± 0.065)	1810.8 (± 372.9)	2094.7 (± 640.9)
Gulf of California	mt	45	18	0.74	0.011 (± 0.000)	0.029 (± 0.011)	828.6 (± 256.6)	N/A
	nuc.	32	19	0.94	0.021 (± 0.001)	0.342 (± 0.206)	1801.1 (± 333.9)	126.0 (± 176.3)

used to generate a mean value of migration (*Nm*) and its associated standard deviation.

Anisotremus davidsonii populations' coalescence times were also determined. The time of coalescence was estimated by assuming that coalescence was reached when the population size was reduced to 1% of its present-day value, following Wares & Cunningham (2001). In order to estimate coalescence time, we estimated the mutation rate (μ) as μ = substitutions per site per generation. Generation time, a value necessary to estimate coalescence time, was estimated at 3 years, the average time for sexual maturity for *A. davidsonii* (Cailliet *et al*. 2000), as well as other grunts (2.4–3.6 years, Froese & Pauly 2000; Pajuelo *et al*. 2003). Thus, mutation rate was computed by considering half the number of substitutions between species or populations (only one branch is taken into account) divided by the total number of nucleotides, divided by the number of generations. The number of generations was the time of divergence divided by generation time, in this case 3 years.

Results

Sequences

The 5'-end portion of the mitochondrial cytochrome *b* was sequenced for 105 individuals (Table 1). Among *Anisotremus davidsonii* individuals, out of 692 aligned base pairs (bp), 44 bp were variable and 23 bp were phylogenetically informative. Among *Anisotremus* species, 174 bp were variable and 160 bp were informative. As expected, the first intron of the nuclear S7 ribosomal protein was less variable. Among *A. davidsonii*, out of 415 bp, 23 bp were variable and 8 bp were informative (46 bp and 36 bp, respectively, among *Anisotremus* species).

Phylogenetic relationships

Phylogenetic relationships between *Anisotremus* species using mitochondrial and nuclear molecular markers were not found to be significantly different (Shimodaira–Hasegawa test, $P = 0.25$) (Figs 2 and 3). The presumed geminate species grouped together with high bootstrap support (both markers and with all methods): *Anisotremus virginicus* with *Anisotremus taeniatus* and *Anisotremus surinamensis* with *Anisotremus interruptus*. The Peruvian grunt, *Anisotremus scapularis* was found to be the sister taxon to the other sampled species in the genus. The focal species, *A. davidsonii* was found to be the sister taxon to the *A. virginicus*/*A. taeniatus* clade. These latter two relationships, however, were only weakly supported by our data (Figs 2 and 3).

In contrast, mitochondrial and nuclear markers differed in the phylogenetic arrangements of *A. davidsonii* populations. While the mitochondrial cytochrome *b* sequences partitioned Sea of Cortez and Pacific samples in two distinct clades (Fig. 2), the nuclear intronic sequences did not (Fig. 3). Indeed, four fixed differences were found between cytochrome *b* sequences of Sea of Cortez and Pacific individuals, while no fixed differences were found in the nuclear sequences.

Gene flow levels between disjunct populations

Gene flow analysis reflected the phylogenetic relationships described above. Gene flow between Pacific and Sea of Cortez populations of *A. davidsonii* was not detectable at the mitochondrial level (fixed differences), and was also low for the nuclear marker ($F_{ST} = 0.03$, $Nm = 7.5$). When using a coalescent approach, historical directionality of gene flow, based on the nuclear marker, could be estimated (Table 2). Gene flow was found to be highly asymmetrical with 16 times more migrants going from the Sea of Cortez into the Pacific than the reverse (Table 2).

Genetic divergence and temporal divergence

Our data did not show any departure from a molecular clock ($P = 0.209$ for mtDNA and $P = 0.189$ for nuclear DNA) indicating that a molecular clock could not be rejected. In addition to a molecular clock, prior knowledge of mutation rates is necessary to estimate divergence times. In this study, rate calibration was obtained using two congeneric trans-isthmian geminate species pairs. Based on mitochondrial cytochrome *b* sequences, and using the evolution model obtained through MODELTEST (HKY + G), the divergence between *A. virginicus* and its geminate species *A. taeniatus* was found to be 1.7 times higher than the divergence between the other pair of geminate species, *A. surinamensis* and *A. interruptus* (Table 3). Divergence of *A. taeniatus* and *A. virginicus* was also found to be greater than for *A. surinamensis* and *A. interruptus* for the nuclear marker (ratio 24:1, model selected, F81 + G). However, the very low level of nuclear divergence in the latter species pair (only one fixed difference) prevents an accurate estimate of divergence and is likely responsible for this very high value.

The genetic divergence between *A. davidsonii* and its sister clade (*A. taeniatus* + *A. virginicus*) was 10.76% and 2.66% based on the mitochondrial and nuclear DNA markers, respectively. The genetic divergence between *A. davidsonii* populations was lower than between trans-isthmian geminate species. Mitochondrial genetic divergence between *A. davidsonii* Sea of Cortez and Pacific populations (0.40%) was found to be 8.7 and 5.0 times less than the divergence of *A. virginicus*/*A. taeniatus* (3.48%) and *A. surinamensis*/*A. interruptus* (2.02%), respectively.

Based on mitochondrial cytochrome *b* sequences, if we assume that the separation of *A. surinamensis* and *A. interruptus* occurred at the closure of the Isthmus of Panama, 3.5–3.1 Ma, the other geminates (*A. virginicus* and *A. taeniatus*)

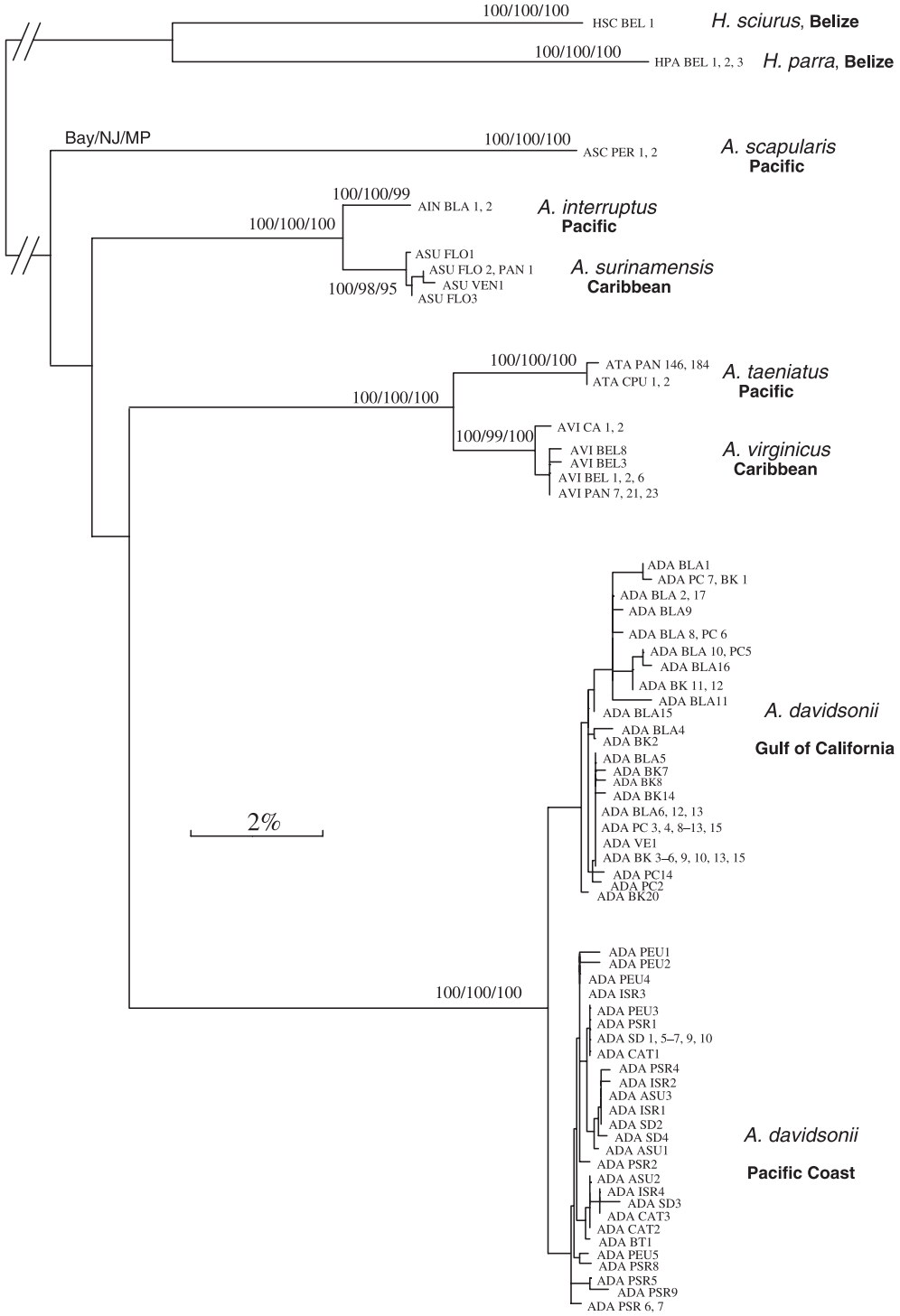

Fig. 2 Molecular phylogeny of *Anisotremus* based on the mitochondrial control region marker using the neighbour-joining method. Labels are described in Table 2. Bootstrap support is shown when above 50%, for the three methods used, Bayesian, neighbour joining, and maximum parsimony, in that order. Scale bar represents 2% sequence divergence. *Haemulon sciurus* and *Haemulon parra* were used as outgroups.

would have diverged 6.0–5.3 Ma. In this case, *A. davidsonii* would have diverged from its sister clade between 18.6 and 16.5 Ma, and Sea of Cortez and Pacific populations of *A. davidsonii* would have diverged from each other between 693 000 and 614 000 years ago (Table 3). In contrast, if we consider *A. virginicus*/*A. taeniatus* to have diverged 3.5–3.1 Ma, then we find that the divergence between *A. davidsonii* and its sister clade to have occurred between 10.7 and 9.6

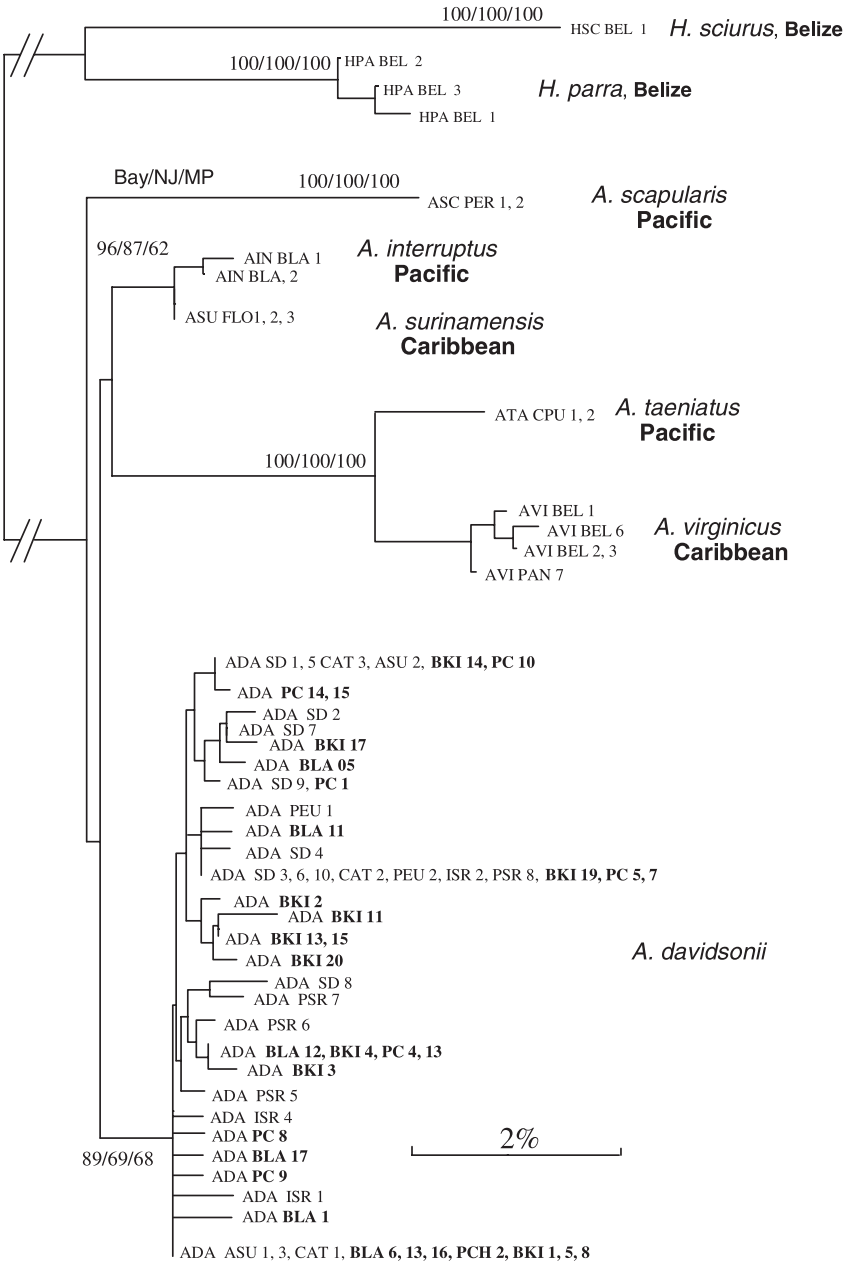

Fig. 3 Molecular phylogeny of *Anisotremus* based on the first intron of the nuclear ribosomal protein S7 using the neighbour-joining method. Labels are described in Table 2. Bootstrap support is shown when above 50%, for the three methods used, Bayesian, neighbour joining, and maximum parsimony, in that order. Scale bar represents 2% sequence divergence. *Haemulon sciurus* and *Haemulon parra* were used as outgroups.

Ma, and the separation of the Gulf and Pacific populations to have occurred between 400 000 and 350 000 years ago (Table 3). Based on nuclear markers, the divergence time between *A. davidsonii* and its sister clade was estimated to have occurred between 6.5 and 5.7 Ma.

Coalescence and temporal divergence

Demographic factors were estimated for both population size (Θ) as well as population growth (g) (Table 2). These two estimates were found to be consistent among replicates (very low standard deviations), and between mitochondrial and nuclear markers (Table 2). Both population size and growth were similar in Sea of Cortez and Pacific populations of *A. davidsonii*.

Coalescence for *A. davidsonii* was reached between 760 000 and 400 000 years ago (Table 3). Coalescence for populations of *A. davidsonii* ranged between 640 000 and 160 000 years ago based on mitochondrial markers and 426 000 and 363 000 years ago when using nuclear markers (Table 3). Where comparisons are possible (mitochondrial markers), coalescence times are similar (expectedly slightly smaller) to estimates of times of population divergence based on sequence divergence (690 000–350 000 years ago).

Table 3 Percentage sequence divergence (substitution model is described in the text) of geminate species *Anisotremus virginicus/taeniatus* and *Anisotremus surinamensis/Anisotremus interruptus*, based on mitochondrial cytochrome *b* and nuclear S7 ribosomal protein first intron sequences, and their associated mutation rate μ (mutations/site/generation) assuming a split between species occurring 3.1–3.5 million years ago (Ma) (left two columns). Divergence and coalescence times (right 4 columns, in thousands of years) for *Anisotremus davidsonii* are based on values from columns 1 and 2, respectively. Cytochrome *b* sequence divergence between Gulf and Pacific populations of *A. davidsonii* was 0.40%

	Geminate species		*Anisotremus davidsonii*			
					Within regions	
	% divergence	μ (10^{-8})	All samples coalescence	Between regions divergence	Gulf of California coalescence	Pacific coast coalescence
mtDNA — cyt *b*						
A. virginicus/A. taeniatus	3.48	1.7–1.5	445–394	402–356	372–330	184–163
A. surinamensis/A. interruptus	2.02	1.0–0.9	765–678	693–614	641–568	317–281
Nuclear DNA — S7						
A. virginicus/A. taeniatus	1.44	0.7–0.6	669–573	N/A	426–365	423–363
A. surinamensis/A. interruptus	0.12	N/A	N/A	N/A	N/A	N/A

Mitochondrial markers indicated that coalescence was reached more recently in Pacific populations (320 000–160 000 years ago) than in Gulf populations (640 000–330 000 Ma), which is consistent with the outward migration trend from the Gulf into the Pacific described above (including a possible signature of founder effect).

Discussion

Phylogeography

Alternative historical scenarios, which predict different divergence times between disjunct populations, may explain the present distribution of *Anisotremus davidsonii*. In our study, two approaches (sequence divergence and coalescence) were used to estimate divergence times between populations of *A. davidsonii*. Both estimates were based on mutation rates obtained from congeneric trans-isthmian geminates. However, the two pairs that were used to calibrate the divergence between *A. davidsonii* populations (*A. virginicus/A. taeniatus* and *A. surinamensis/A. interruptus*) showed different genetic divergences. This mirrors previously published results for these two species pairs where the ratio of divergences based on the mitochondrial cytochrome oxydase I (COI) was found to be 2.6 (Bermingham *et al*. 1997). These results seem inconsistent with a single vicariant event (namely the rise of the Isthmus of Panama, which occurred 3.1–3.5 Ma) being responsible for original allopatry in these species. This observation, however, is not unique. It has been shown in other fishes (Bermingham *et al*. 1997; Lessios 1998), and invertebrates (Marko 2002). It was most striking in closely related congeneric species of *Alpheus* snapping shrimp (Knowlton & Weigt 1998), where differences in genetic divergence were attributed to ecological differences. Some *Alpheus* species live in deeper water; others live in mangroves, a likely habitat for the last phase of the closure of the Isthmus. Thus, population disjunction in these species is likely to have occurred at different times, resulting in different genetic divergences. In the case of *Anisotremus* (and other grunts), mangrove habitat is also very important both as a nursery ground and as adult habitat (Mumby *et al*. 2004). It is therefore possible that the two *Anisotremus* geminate pairs diverged at different times based on different mangrove habitat use, with *A. virginicus* diverging from *A. taeniatus* earlier than *A. surinamensis* from *A. interruptus*. One alternative hypothesis requires unequal substitution rates in these two clades. This is possible, but less likely considering their close phylogenetic relationship and the fact that our molecular data are consistent with a molecular clock and do not show any sign of unequal substitution rates (such as very different branch lengths). Finally, a third possible explanation is that the closest pair, *A. surinamensis/A. interruptus* diverged more recently than 3.1 Ma. A Pleistocene breach of the Isthmus of Panama that would have occurred approximately 2 Ma has been proposed (Coates & Obando 1996; Cronin & Dowset 1996; Banford *et al*. 2004). If the genetic separation of *A. surinamensis/A. interruptus* dates from that event, then the back-calculated separation of *A. virginicus* and *A. taeniatus* would have happened approximately 3.4 Ma, a date that is consistent with the time of full closure of the Isthmus of Panama (3.5–3.1 Ma).

If we consider the complete range of divergence times obtained, Gulf and Pacific *A. davidsonii* populations would have diverged between 690 000 and 350 000 years ago based on mitochondrial sequence divergences, and between 640 000 and 160 000 years ago based on coalescence times (430 000–360 000 years ago when using the nuclear marker). This

range becomes narrower when one only uses the early separation hypothesis (3.5–3.1 Ma) for *A. surinamensis–A. interruptus* geminate pair for clock calibration: 690 000 to 610 000 years ago based on divergences, and 640 000 to 570 000 years ago based on the oldest coalescence time (Gulf population), or if one uses the Isthmus breach hypothesis (2 Ma) 400 000–360 000 years ago based on divergences, and 370 000–330 000 years ago based on coalescence. While the range of divergences is fairly large, it still allows for the testing of specific phylogeographic hypotheses. The three main seaways across the Baja California Peninsula opened and then closed approximately 4 Ma, 3 Ma, and 1 Ma, potentially creating vicariant events. These seaways predate the disjunction between Gulf and Pacific populations of *A. davidsonii*, the most recent mid-peninsular closure (1 Ma) being still 310 000 years older than our highest estimate, and pre-dates by approximately 400 000–600 000 years the other and possibly more likely estimate.

These findings suggest that disjunct populations were either established by migration around the cape long after the closure of all seaways, or alternatively that more recent seaways did exist but have not yet been uncovered. Indeed, recent geological studies on marine incursions in the Baja California Peninsula have underscored the paucity of knowledge in this region and the potential for the presence of more seaways yet to be characterized (Ledesma-Vázquez 2002; Oskin & Stock 2003).

Fixed differences between Gulf and Pacific populations indicate that, once established, genetic separation between populations was later maintained, including during the subsequent glaciating events, indicating a lack of recent migration around the cape.

Few Baja California disjunct fish species show affinities with North Pacific fishes (Bernardi *et al*. 2003). For example, *Sebastes macdonaldi*, the Mexican rockfish, is the southernmost species of the northeast Pacific rockfishes (Rocha-Olivares *et al*. 2003). In contrast, *A. davidsonii* is the northernmost species of its genus. In addition, historical migration trends of *A. davidsonii* individuals were shown to be predominantly directed out of the Gulf of California towards the Pacific coast, and coalescent estimates consistently show that the Pacific population is younger than Gulf population. Thus, a likely scenario is for *A. davidsonii* to have first invaded the Sea of Cortez from the south between 0.7 and 0.4 Ma, soon after having colonized the Pacific coast.

Speciation

Did population disjunction in *A. davidsonii* result in speciation? Several species concepts have been proposed, yet all of them require some level of breeding barrier (Avise 2000; Bowen *et al*. 2001). The biological species concept (BSC) (Mayr 1942) identifies allopatry and sexual incompatibility as characteristic of incipient species. *A. davidsonii* populations are likely to be sexually compatible, even if at the present time mtDNA markers suggest that populations are not interbreeding. The other main branch of species concepts grouped under the phylogenetic species concept (PSC) (Eldredge & Cracraft 1980) requires individuals to cluster in reciprocally monophyletic clades. This definition should be fulfilled at all loci, a requirement rarely tested (Taylor *et al*. 2000). Considering that lineage-sorting time is vastly different for mitochondrial (N_e generations) and nuclear (4 N_e generations) loci, this definition is overly restrictive when one deals with the very early stages of speciation. In the case of *A. davidsonii* populations, mitochondrial sequences do cluster in reciprocally monophyletic clades, while nuclear sequences do not, a result obtained in several other systems (Shaw 1992). In order to accommodate for these results, a coalescent component was included in more recent species concepts such as the genealogical species concepts (GSC) (Baum & Shaw 1995; Shaw 1998; Hudson & Coyne 2002). In true species, gene divergence should precede coalescence (Edwards & Beerli 2000). In our study, gene divergence was very similar to (or slightly older than) coalescent estimation (Table 3), indicating that speciation was either not observed or that we were at the very early stages of the speciation process.

Phenetic considerations resulted in similar conclusions. Disjunct populations of *A. davidsonii* diverged between 640 000 and 160 000 years ago, a time long enough for them to diverge into separate species (e.g. McCune & Lovejoy 1998; Seehausen 2000). However, cytochrome *b* sequence divergence between disjunct populations (0.40%) is lower than divergence levels observed in allopatrically derived sister species of fishes (2–8%), and is consistent with either sympatrically derived sister species (0–1.25%) or with geographically isolated populations (0–5.6%) (McCune & Lovejoy 1998). A scenario of sympatric speciation is unlikely in the context of *A. davidsonii*; thus, our results seem more consistent with the isolated-populations hypothesis.

Conclusion

The Baja California disjunct distribution of *Anisotremus davidsonii*, combined with the possibility of a molecular clock calibration in closely related congeneric trans-isthmian geminate species, provides a unique opportunity to study the early stages of allopatric speciation and to rigorously test phylogeographic hypotheses. Nuclear and mitochondrial markers allowed us to reconstruct the evolutionary history of the species. The species invaded the Gulf of California from the south, an event followed by the migration by some individuals out of the Gulf towards the Pacific coast (between 160 000 and 640 000 years ago). This migration occurred either around the southern tip of the Baja California Peninsula (Cabo San Lucas), or through a yet-to-be-discovered seaway across the peninsula.

Coalescent analysis, sequence divergence estimates, and cladistic analysis of the disjunct populations bear some of the characteristics of incipient speciation (reciprocal monophyly, time of divergence), but also lack some of the requirements of bona fide species (difference between coalescent times and divergence times, reciprocal monophyly at all loci). We consider these conflicting genetic signatures as evidence of the very early stages of allopatric speciation. As such, the disjunct populations of *A. davidsonii* may provide key to the understanding of the early mechanisms of speciation.

Acknowledgements

This paper is respectfully dedicated to the memory of Ernst Mayr. Several people generously donated samples for this study: H. Lessios, Y. Alva, S. Aalbers, E. Azzurro, C. Mora, and M. Edwards. Help in the field was provided by S. Alesandrini, E. Azzurro, D. Canestro, N. Crane, K. Crow, P. DalFerro, M. Edwards, J. Figurski, P. Hastings, D. Huang, M. Kenner, E. Maloney, A. Marchesi, J. Portocarrero, M. Ramon, M. Readdie, D. Steller, and A. Terry. We would like to thank P.T. Raimondi, M. Dawson, H. Lessios, B. Riddle, H.J. Walker, W. Watson, and P. Hastings for discussions and anonymous reviewers for comments. Ivano Aiello and Matt Forrest provided insightful comments on the geology of Baja California. This work was funded by UCMEXUS, and the David and Lucille Packard Foundation PISCO programme.

References

Avise JC (1992) Molecular population structure and biogeographic history of a regional fauna- a case history with lessons for conservative biology. *Oikos*, **63**, 62–76.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Banford HM, Bermingham E, Collette BB (2004) Molecular phylogenetics and biogeography of transisthmian and amphi-Atlantic needlefishes (Belonidae: *Strongylura* and *Tylosurus*): perspectives on New World marine speciation. *Molecular Phylogenetics and Evolution*, **31**, 833–851.

Baum D, Shaw KL (1995) Genealogical perspectives on the species problem. In: *Experimental and Molecular Approaches to Plant Biosystematics* (eds Hoch PC, Stephenson AC), pp. 289–303. Missouri Botanical Garden, St. Louis, Missouri.

Beerli P (2003) MIGRATE — *A maximum likelihood program to estimate gene flow using the coalescent, Tallahassee/Seattle*. http://evolution.gs.washington.edu/lamarc/migrate/html.

Bermingham E, Lessios HA (1993) Rate variation of protein and mitochondrial DNA evolution as revealed by sea urchins separated by the Isthmus of Panama. *Proceedings of the National Academy of Sciences, USA*, **90**, 2734–2738.

Bermingham E, McCafferty SS, Martin AP (1997) Fish biogeography and molecular clocks: perspectives from the Panamian Isthmus. In: *Molecular Systematics of Fishes* (eds Kocher TD, Stepien CA), pp. 113–128. Academic Press, San Diego, California.

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

Bowen BW, Bass AL, Rocha LA, Grant WS, Robertson DR (2001) Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global scale. *Evolution*, **55**, 1029–1039.

Cailliet GM, Burton EJ, Cope JM, Kerr LA (2000) *Biological Characteristics of Nearshore Fishes of California: A Review of Existing Knowledge*. Pacific States Marine Fisheries Commission, Sacramento, CA (www.dfg.ca.gov/mrd/lifehistories/).

Chow S, Hazama K (1998) Universal PCR primers for S7 ribosomal protein gene introns in fish. *Molecular Ecology*, **7**, 1255–1256.

Coates AG, Obando, JA (1996) The geologic evolution of the Central American Isthmus. In: *Evolution and Environment in Tropical America* (eds Jackson J, Budd A, Coates A), pp. 21–56. University of Chicago Press, Chicago, Illinois.

Collin R (2003) Phylogenetic relationships among calyptraeid gastropods and their implications for the biogeography of marine speciation. *Systematic Biology*, **52**, 618–640.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.

Cronin TM, Dowsett HF (1996) Biotic and oceanographic response to the Pliocene closing of the Central American Isthmus. In: *Evolution and Environment in Tropical America* (eds Jackson J, Budd A, Coates A), pp. 76–104. University of Chicago Press, Chicago and London.

Donaldson KA, Wilson RR (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.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Eldredge N, Cracraft J (1980) *Phylogenetic Patterns and the Evolutionary Process*. Columbia University Press, New York.

Endler JA (1977) *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, New Jersey.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.

Froese R, Pauly D (2000) *FishBase: Concepts, Design and Data Sources*, p. 344. ICLARM, Los Baños, Laguna, Philippines.

Grismer LL (1999) Phylogeny, taxonomy, and biogeography of *Cnemidophorus hyperythrus* and *C. ceralbensis* (Squamata: Teiidae) in Baja California, Mexico. *Herpetology*, **55**, 28–42.

Grismer LL (2000) Evolutionary biogeography on Mexico's Baja California Peninsula: a synthesis of molecules and historical geology. *Proceedings of the National Academy of Sciences, USA*, **97**, 14017–14018.

Huang D, Bernardi G (2001) Disjunct Sea of Cortez–Pacific Ocean *Gillichthys mirabilis* populations and the evolutionary origin of their paedomorphic relative, *Gillichthys seta*. *Marine Biology*, **138**, 421–428.

Hudson RR, Coyne JA (2002) Mathematical consequences of the genealogical species concept. *Evolution*, **56**, 1557–1565.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics*, **17**, 754–755.

Jordan DS (1908) The law of geminate species. *American Naturalist*, **42**, 73–80.

Knowlton N, Weigt LA (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **265**, 2257–2263.

Knowlton N, Weight LA, Solorzano LA, Mills DK (1993) Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. *Science*, **260**, 1629–1632.

Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, **149**, 429–434.

Ledesma-Vázquez J (2002) A gap in the Pliocene invasion of seawater to the Gulf of California. *Revista Mexicana de Ciencias Geologicas*, **19**, 145–151.

Lessios HA (1998) The first stage of speciation as seen in organisms separated by the Isthmus of Panama. In: *Endless Forms: Species and Speciation* (eds Howard DJ, Berlocher SH), pp. 186–201. Oxford University Press, New York.

Lessios HA, Kessing BD, Pearse JS (2001) Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. *Evolution*, **55**, 955–975.

Lindeman KC, Pugliese R, Waugh GT, Ault JS (2000) Developmental patterns within a multispecies reef fishery: management applications for essential fish habitats and protected areas. *Bulletin of Marine Science*, **66**, 929–956.

Marko PB (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evology*, **19**, 2005–2021.

Mayr E (1942) *Systematics and the Origin of Species*, pp. 334. Columbia University Press, New York.

Mayr E (1963) *Animal Species and Evolution*. Harvard University Press, Cambridge, Massachusetts.

McCune AR, Lovejoy NR (1998) The relative rate of sympatric and allopatric speciation in fishes: tests using DNA sequence divergences between sister species and among clades. In: *Endless Forms: Species and Speciation* (eds Howard DJ, Berlocher SH), pp. 172–185. Oxford University Press, New York.

Miller DJ, Lea RN (1972) *Guide to Coastal Marine Fishes of California*. Fish Bullettin no. 157. Department of Fish and Game, University of California Press, Berkeley.

Mumby PJ, Edwards AJ, Arias-Gonzalez JE *et al.* (2004) Mangroves enhance the biomass of coral reef fish communities in the Caribbean. *Nature*, **427**, 533–536.

Muss A, Robertson DR, Stepien CA, Wirtz P, Bowen BW (2001) Phylogeography of *Ophioblennius*: the role of ocean currents and geography in reef fish evolution. *Evolution*, **55**, 561–572.

Oskin M, Stock J (2003) Marine incursion synchronous with plate boundary localization in the Gulf of California. *Geology*, **31**, 23–26.

Pajuelo JG, Lorenzo JM, Gregoire M, Dominguez-Seoane R (2003) Life history of *Pomadasys Incisus* (Osteichthyes: Haemulidae) in the Canarian Archipelago. *Scientia Marina* **67**, 241–248.

Palumbi SR (1992) Marine speciation on a small planet. *Trends in Ecology & Evolution*, **7**, 114–118.

Palumbi SR (1994) Genetic divergence, reproductive isolation and marine speciation. *Annual Review of Ecology and Systematic*, **25**, 547–572.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Present T (1987) Genetic differentiation of disjunct Gulf of California and Pacific outer coast populations of *Hypsoblennius jenkinsi*. *Copeia*, **1987**, 1010–1024.

Randall JE (1998) Zoogeography of shore fishes of the Indo-Pacific region. *Zoological Studies*, **37**, 227–268.

Riddle BR, Hafner DJ, Alexander LF (2000b) Phylogeography and systematics of the *Peromyscus eremicus* species group and the historical biogeography of North American warm regional deserts. *Molecular Phylogenetics and Evolution*, **17**, 145–160.

Riddle BR, Hafner DJ, Alexander LF (2000c) Comparative phylogeography of baileys' pocket mouse (*Chaetodipus bailey*) and the *Peromyscus eremicus* species group: historical vicariance of the Baja California Peninsular desert. *Molecular Phylogenetics and Evolution*, **17**, 161–172.

Riddle BR, Hafner DJ, Alexander LF, Jaeger JR (2000a) Cryptic vicariance in the historical assembly of a Baja California Peninsular desert biota. *Proceedings of the National Academy of Sciences, USA*, **97**, 14438–14443.

Riginos C, Nachman MW (2001) Population subdivision in marine environments: the contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. *Molecular Ecology*, **10**, 1439–1453.

Rocha-Olivares A, Leal-Navarro RA, Kimbrell C, Lynn EA, Vetter RD (2003) Microsatellite variation in the Mexican rockfish *Sebastes macdonaldi*. *Scientia Marina* **67**, 451–460.

Rozas J, Sánchez-De JC, Barrio I, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Seehausen O (2000) Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. *Advances in Ecological Research*, **31**, 237–274.

Shaw KL (1992) Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proceedings of the National Academy of Sciences, USA*, **99**, 16122–16127.

Shaw KL (1998) Species and the diversity of natural groups. In: *Endless Forms: Species and Speciation* (eds Howard DJ, Berlocher SJ), pp. 44–56. Oxford University Press, Oxford, UK.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biological Evology*, **16**, 1114–1116.

Stepien CA, Rosenblatt RH, Bargmeyer BA (2001) Phylogeography of the spotted sand bass, *Paralabrax maculatofasciatus*: divergence of Gulf of California and Pacific coast populations. *Evolution*, **55**, 1852–1862.

Swofford DL (2003) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Taylor MS, Hellberg ME (2005) Marine radiations at small geographic scales: speciation in neotropical reef gobies (*Elacatinus*). *Evolution*, **59**, 374–385.

Taylor JW, Jacobson DJ, Kroken S *et al.* (2000) Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology*, **31**, 21–32.

Terry A, Bucciarelli G, Bernardi G (2000) Restricted gene flow and incipient speciation in disjunct Pacific Ocean and Sea of Cortez populations of a reef fish species, *Girella nigricans*. *Evolution*, **54**, 652–659.

Thomson DA, Findley LT, Kersitch AN (2000) *Reef Fishes of the Sea of Cortez. The Rocky Shore Fishes of the Gulf of California*. University of Texas Press, Austin, Texas.

Upton DE, Murphy RW (1997) Phylogeny of the side-blotched lizards (Phrynosomatidae: *Uta*) based on mtDNA sequences: support for a midpeninsular seaway in Baja California. *Molecular Phylogenetics and Evolution*, **8**, 104–113.

Walker BW (1960) The distribution and affinities of the marine fish fauna of the Gulf of California: the biogeography of Baja California and adjacent seas. *Systematic Zoology*, **9**, 123–133.

Wares JP, Cunningham CW (2001) Phylogeography and historical ecology of the north Atlantic intertidal. *Evolution*, **55**, 2455–2469.

Watson W, Walker HJ (1992) Larval development of sargo (*Anisotremus davidsonii*) and salema (*Xenistius californiensis*) (Pisces, Haemulidae) from the southern California bight. *Bulletin of Marine Science*, **51**, 360–406.

Wiens JJ (2004) What is speciation and how should we study it? *American Naturalist*, **163**, 914–923.

Giacomo Bernardi is interested in the ecology, evolution, and speciation of fishes, with particular focus on the Baja California disjuncts. Jen Lape is interested in fish ecology and community structure.