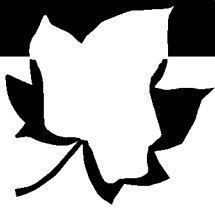ORIGINAL ARTICLE

Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*)

Michelle R. Gaither[1]*, Robert J. Toonen[1], D. Ross Robertson[2], Serge Planes[3] and Brian W. Bowen[1]

[1]*Hawaii Institute of Marine Biology, University of Hawaii, Kaneohe, HI,* [2]*Smithsonian Tropical Research Institute, Panamá, APO AA, USA and* [3]*UMS 2978 CNRS-EPHE, Centre de Recherches Insulaires et Observatoire de l'Environnement (CRIOBE), BP 1013, 98729 Moorea, French Polynesia*

ABSTRACT

Aim In the Indo-Pacific, the mass of islands of the Indonesian archipelago constitute a major biogeographical barrier (the Indo-Pacific Barrier, IPB) separating the Pacific and Indian oceans. Evidence for other, more localized barriers include high rates of endemism at the Marquesas and other isolated peripheral islands in the Pacific. Here we use mitochondrial-sequence comparisons to evaluate the efficacy of biogeographical barriers on populations of the snappers *Lutjanus kasmira* and *Lutjanus fulvus* across their natural ranges.

Location Pacific and Indian oceans.

Methods Mitochondrial cytochrome *b* sequence data were obtained from 370 individuals of *L. kasmira* and 203 individuals of *L. fulvus* collected from across each species' range. Allele frequency data for two nuclear introns were collected from *L. kasmira*. Phylogenetic and population-level analyses were used to determine patterns of population structure in these species and to identify barriers to dispersal.

Results *Lutjanus kasmira* lacks genetic structure across the IPB and throughout 12,000 km of its central Indo-Pacific range. In contrast, *L. fulvus* demonstrates high levels of population structure at all geographical scales. In both species, highly significant population structure results primarily from the phylogenetic distinctiveness of their Marquesas Islands populations (*L. kasmira*, $d = 0.50$–0.53%; *L. fulvus*, $d = 0.87$–1.50%). Coalescence analyses of the *L. kasmira* data indicate that populations at opposite ends of its range (western Indian Ocean and the Marquesas) are the oldest. Coalescence analyses for *L. fulvus* are less robust but also indicate colonization from the Indian to the Pacific Ocean.

Main conclusions The IPB does not act as a biogeographical barrier to *L. kasmira*, and, in *L. fulvus*, its effects are no stronger than isolating mechanisms elsewhere. Both species demonstrate a strong genetic break at the Marquesas. Population divergence and high endemism in that archipelago may be a product of geographical isolation enhanced by oceanographic currents that limit gene flow to and from those islands, and adaptation to unusual ecological conditions. *Lutjanus kasmira* shows evidence of Pleistocene population expansion throughout the Indo-central Pacific that originated in the western Indian Ocean rather than the Marquesas, further demonstrating a strong barrier at the latter location.

Keywords

Cytochrome *b*, endemism, Indo-Pacific Barrier, marine fish, Marquesas, mtDNA, nuclear intron, phylogeography.

*Correspondence: Michelle R. Gaither, Hawaii Institute of Marine Biology, University of Hawaii, PO Box 1346, Kaneohe, HI 96744, USA. E-mail: gaither@hawaii.edu

INTRODUCTION

Biogeographical provinces are defined by level of endemism and recognized habitat discontinuities (Briggs, 1974; Brown & Lomolino, 1998; Bellwood & Wainwright, 2002; Robertson et al., 2004; Floeter et al., 2008). In the open ocean, biogeographical barriers can shape species distributions by physically restricting dispersal (strong ocean currents) or by limiting dispersal success owing to the distance between areas of suitable habitat. Examples of tropical marine biogeographical barriers include the expanses of open ocean that separate the central and eastern Pacific (Robertson et al., 2004; Lessios & Robertson, 2006) and the eastern and western Atlantic (Banford et al., 1999), the Amazon–Orinoco outflow that separates the Brazilian and Caribbean provinces (Rocha et al., 2002), and the large cold upwelling area off southwest Africa that separates the tropical Atlantic and Indian oceans (Briggs, 1974; Rocha et al., 2007).

The Indo-Pacific Barrier (IPB) is a widely recognized partition (based on faunal distributions) that separates the Pacific and Indian Ocean provinces (Briggs, 1974). Although the location of the boundary between these provinces is debated (see Hobbs et al., 2009), this barrier is most often associated with the Sunda Shelf between Asia and Australia (Fleminger, 1986; Barber et al., 2006). During Pleistocene glacial cycles the repeated lowering of sea levels (as low as 120 m below present levels) imposed a nearly complete barrier between the two oceans. Strong upwelling in the region probably enhanced the effectiveness of the IPB by reducing the availability of suitable habitat for tropical marine organisms (Galloway & Kemp, 1981; Fleminger, 1986; Voris, 2000; Naish et al., 2009). Historical and contemporary restrictions to dispersal between the Pacific and Indian oceans are indicated by the confinement of many demersal species primarily to one ocean or the other (Woodland, 1983; McMillan & Palumbi, 1995; Randall, 1998; Briggs, 1999; but see Hobbs et al., 2009). More recently, the effects of the IPB and other barriers have been assessed using genetic data. Taxa that disperse as adults show genetic continuity between the Pacific and Indian oceans, indicating that they cross the IPB with regularity (e.g. whale shark, Rhincodon typus, Castro et al., 2007; wahoo, Acanthocybium solandri, Theisen et al., 2008). Studies of demersal organisms that lack vagile adults have found intraspecific genetic differentiation across the IPB in many fishes (Lacson & Clark, 1995; Planes & Fauvelot, 2002; Bay et al., 2004; Lu et al., 2006; Menezes et al., 2006; Craig et al., 2007) and invertebrates (Lavery et al., 1995, 1996; Williams & Benzie, 1998; Benzie, 1999; Duda & Palumbi, 1999; Barber et al., 2000; Lessios et al., 2001, 2003) with a few exceptions (Bowen et al., 2001; Lessios et al., 2001; Horne et al., 2008).

In addition to the biogeographical barriers that separate tropical marine provinces, barriers within provinces exist and can promote more localized isolation. Randall (1998, 2001, 2007) discussed five areas in Oceania that have high levels of endemism among shore fishes: Hawaii (25.0%), Easter Island (22.2%), the Marquesas (11.6%), Lord Howe Island and Norfolk Island (7.2%), and Rapa (5.5%). These peripheral island groups are all geographically isolated, being 500–1500 km from the nearest neighbouring archipelago (Randall, 1998). Genetic studies of widely distributed species have demonstrated genetic divergences among these five insular areas in milkfish (Chanos chanos; Winans, 1980), convict surgeonfish (Acanthurus triostegus; Planes & Fauvelot, 2002), yellowfin goatfish (Mulloidichthys vanicolensis; Stepien et al., 1994), crown-of-thorns starfish (Acanthaster planci; Benzie & Stoddart, 1992), and four species of coral (Ayre & Hughes, 2004). However, only one genetic study has involved comparison of Marquesan populations (Acanthurus triostegus; Planes & Fauvelot, 2002).

The common bluestripe snapper, Lutjanus kasmira (Forsskål, 1775), and the blacktail snapper, Lutjanus fulvus (Schneider, 1801) (Lutjanidae), are close relatives (Miller & Cribb, 2007) that occupy the same geographical range, from the Marquesas Islands in the central Pacific to the east coast of Africa (Fig. 1). Lutjanus kasmira inhabits a wide depth range from shallow water to at least 265 m (Allen & Talbot, 1985; Randall, 1987; Friedlander et al., 2002). Little has been published on the ecology of L. fulvus, but anecdotal reports from fishermen indicate that L. fulvus occupies a narrower depth range than L. kasmira (1–40 m). Both species are nocturnal predators. Data from two studies indicate that L. kasmira reaches sexual maturity at 1–2 years (Rangarajan, 1971; Morales-Nin & Ralston, 1990), and engages in mass spawning (Suzuki & Hioki, 1979). The pelagic larval duration (PLD) is unknown for both L. kasmira and L. fulvus, although other lutjanids have PLDs of 20–44 days (Zapata & Herron, 2002; Denit & Sponaugle, 2004).

Here we analysed DNA sequence data from L. kasmira and L. fulvus to assess biogeographical partitions in the Indo-Pacific. We resolved patterns of genetic isolation among reefs and archipelagos of the Indo-Pacific to address the following questions: (1) Are patterns of genetic structure concordant with known biogeographical boundaries? (2) Specifically, is there evidence of genetic structure across the Indo-Pacific Barrier? (3) What processes might account for geographical patterns of genetic structure in these fishes?

MATERIALS AND METHODS

A total of 370 individuals of L. kasmira from 12 locations and 203 individuals of L. fulvus from six locations were collected from across each species' range in the Pacific and Indian oceans (Fig. 1) by scuba divers using polespears. Tissue samples (fin clips or gill filaments) were preserved in either 95% ethanol (EtOH) or saturated NaCl solution (Seutin et al., 1991) and stored at room temperature. DNA was isolated using DNeasy Tissue kits (Qiagen, Inc., Valencia, CA, USA) following the manufacturer's protocol for animal tissues, and stored at −20° C. In specimens of L. kasmira and L. fulvus, c. 560 bp of mitochondrial cytochrome b (cyt b) were amplified using the primers H15020 (Meyer, 1994) and Cytb-07L (Taberlet et al., 1992). In specimens of L. kasmira, we also
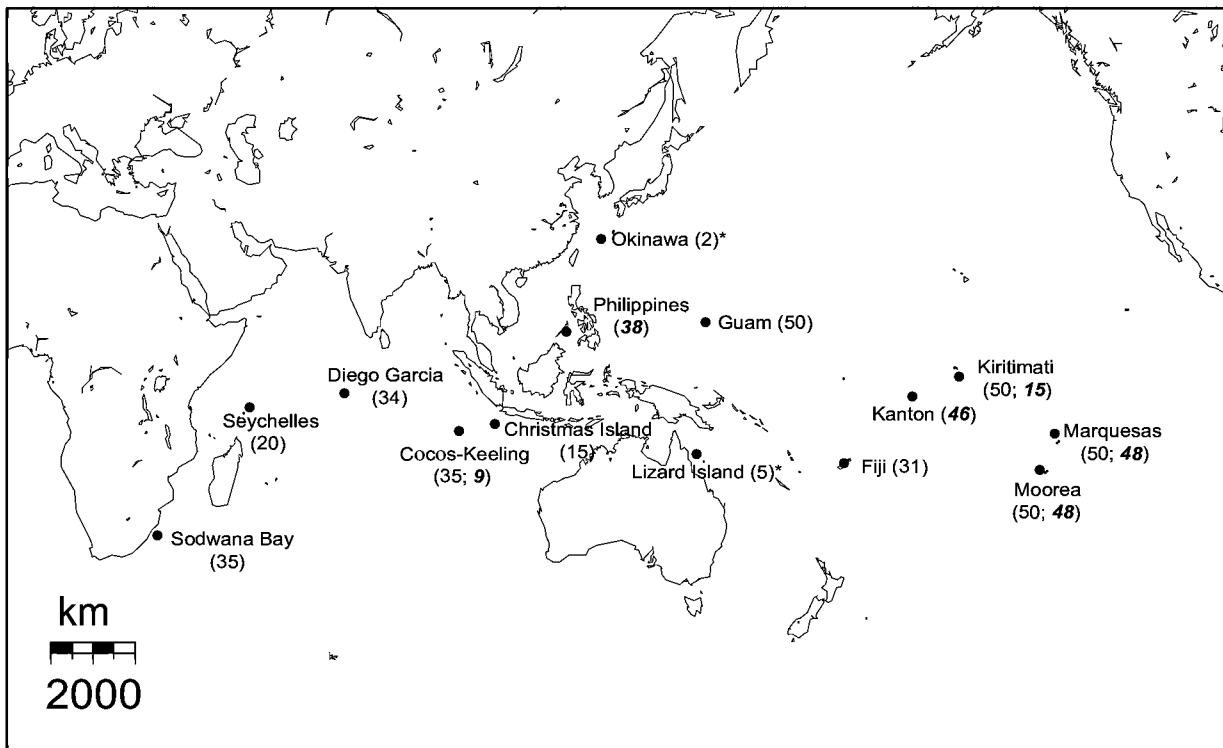
Figure 1 Collection locations and sample sizes for *Lutjanus kasmira* (numbers in roman) and *Lutjanus fulvus* (numbers in bold). Owing to low sample size, locations marked with an asterisk (*) were excluded from population-level analyses.

used the primers Gh5F and Gh6R (Hassan *et al.*, 2002), to amplify *c.* 650 bp of intron 5 of the nuclear growth hormone (GH) gene, and the primers ANTf1 and ANTr1 (Jarman *et al.*, 2002), to amplify *c.* 395 bp of the nuclear adenine nucleotide transporter translocase (ANT) intron. Polymerase chain reactions (PCRs) for all three markers were carried out in a 20-μL volume containing 5–50 ng of template DNA, 0.25–0.5 μM of each primer, 1.5 mM MgCl$_2$, 2.5 mM of each dNTP, 0.5 units IMMOLASE™ DNA polymerase (Bioline Inc., Springfield, NJ, USA), 2.0 μL of 10× ImmoBuffer (Bioline Inc.) and deionized sterile water to volume. PCRs utilized the following cycling parameters: initial denaturation at 95°C and final extension at 72°C (10 min each), with an intervening 35 cycles of 30 s at 94°C, 30 s at the annealing temperature (48°C for cyt *b*; 60°C for GH; 58°C for ANT), and 1 min at 72°C. Amplification products were purified using 0.75 units of Exonuclease I: 0.5 units of Shrimp Alkaline Phosphatase (ExoSAP; USB, Cleveland, OH, USA) per 7.5 μL of PCR products at 37°C for 60 min, followed by deactivation at 80°C for 10 min. DNA sequencing was performed with fluorescently labelled dideoxy terminators on an ABI 3130XL Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) at the Hawaii Institute of Marine Biology EPSCoR Sequencing Facility. All samples were initially sequenced in the forward direction. Unique genotypes were confirmed by subsequent sequencing in the reverse direction.

Individuals of *L. kasmira* whose allelic states at the nuclear GH and ANT loci could not be resolved unambiguously from

direct sequences were cloned using a TA cloning method to identify the alleles. Prior to cloning, PCR products were purified using the Qiagen QIAquick PCR Purification kit following the manufacturer's protocol and eluting in 30 μL of elution buffer per 20 μL PCR reaction. All products were cloned using the T-tailed vector pZErO-2 (Invitrogen, Carlsbad, CA, USA) in *Escherichia coli* strain DB3.1. Ligation reactions included 1.6 μL 2× ligation buffer (Promega, Madison, WI, USA), 5.0 ng of EcoRV-digested and T-tailed pZErO-2 vector, 1.0 ng of PCR product and 0.4 μL of T4 DNA ligase (Promega) in 4 μL total volume. Heat shock transformation was conducted using 10 μL of α-Select Chemically Competent cells (Bioline Inc.) and 1 μL of ligated vector following the manufacturer's protocol. LB agar plates with 50 μg mL^{-1} kanamycin were spread with 40 μL of transformation mixture. Following an overnight incubation at 37°C, cells from individual colonies were suspended in 20 μL of sterile water. Clones were amplified in 20-μL PCRs containing 0.1 μL of the cell suspension, 0.2 μM of each primer (M13F-GTAAAACGACGGCCAG and M13R-CAGGAAACAGCTAT-GAC), 2.0 mM MgCl$_2$, 0.15 mM dNTPs, 0.5 μg BSA, 0.5 units IMMOLASE™ DNA polymerase, and 2.0 μL of 10× PCR buffer (Bioline, Inc.). PCR cycling conditions included an initial denaturing step at 94°C for 3 min followed by 35 cycles of 94°C for 30 s, 60°C for 30 s and 72°C for 30 s, followed by a final extension at 72°C for 10 min. To verify the identity of alleles, clones were sequenced until at least two copies of each allele were observed.

Sequences for each locus were aligned, edited, and trimmed to a common length using the DNA sequence assembly and analysis software SEQUENCHER 4.6 (Gene Codes, Ann Arbor, MI, USA). In all cases, alignment was unambiguous with no indels or frameshift mutations at any locus. Unique haplotypes and alleles were identified with the merge taxa option in the phylogenetic analysis software MacCLADE 4.05 (Maddison & Maddison, 2002), and deposited in GenBank [accession numbers: *L. kasmira* FJ754049–FJ754133 (cyt *b*), FJ754178–FJ754184 (GH intron), FJ754157–FJ754177 (ANT intron); *L. fulvus* FJ754134–FJ754156 (cyt *b*)].

Data analyses

Mitochondrial DNA

Summary statistics for each species, including haplotype diversity (h) and nucleotide diversity (π), were estimated with algorithms in Nei (1987) as implemented in the statistical software package ARLEQUIN 3.11 (Excoffier et al., 2005). Fu's F_S (Fu, 1997) was calculated to test for evidence of population expansion using 10,000 permutations and excluding the two locations with small sample size (Lizard Island and Okinawa; Table 1). Significant negative values of F_S indicate an excess of low-frequency haplotypes, a signature characteristic of either selection or a recent demographic expansion (Fu, 1997). Haplotype frequencies and distance matrixes were used to estimate gene genealogies, and statistical parsimony networks were constructed using the program TCS 2.21 (Clement et al., 2000). We used mismatch analyses to determine whether the number of pairwise differences among all DNA sequences reflected expanding or stable populations (Harpending, 1994; Schneider et al., 2000). Effective female population size (N_{ef}) was calculated from the cyt *b* data using the equation $\theta = 2N_{ef}v$, where v is the mutation rate per generation for the entire sequence (v = number of bp × divergence rate within a lineage × generation time in years). Population age was calculated using the equation $\tau = 2vt$, where t is age in generations. We use a sequence divergence estimate of 1–2% per Myr between lineages to estimate coalescence times (see Discussion). Generation times for most snappers are unknown, including *L. kasmira* and *L. fulvus*. Combining data for size at first maturity for *L. kasmira* from the Andaman Sea (Rangarajan, 1971) with age–length relationships of *L. kasmira* from Hawaii (Morales-Nin & Ralston, 1990), we estimate generation time to be 3.7 years for this species. Because *L. kasmira* and *L. fulvus* are closely related congeners of similar size (Miller & Cribb, 2007) we provisionally apply this estimation of generation time to *L. fulvus* as well. Given the tentative nature of the estimates of generation time and mutation rate, the corresponding population age and N_{ef} values should be interpreted as first-order approximations (see Hudson & Turelli, 2003).

A likelihood approach, implemented in MODELTEST 3.7 (Posada & Crandall, 1998), was used to determine the mutational model that best fit the cyt *b* data. The GTR+I and the TIM+G models were found to be the best-fit models by the Akaike information criterion for *L. kasmira* and *L. fulvus*, respectively. Because neither model is available in ARLEQUIN 3.11 (Excoffier et al., 2005), we used the most similar model available (Tamura & Nei, 1993). To test for hierarchical population genetic structure in each species, an analysis of molecular variance (AMOVA) was performed in ARLEQUIN using 20,000 permutations. An analogue of Wright's F_{ST}, which incorporates a model of sequence evolution (Φ_{ST}), was calculated for each data set and for pairwise comparisons among all locations with more than five specimens. The average number of nucleotides (corrected; Tamura & Nei, 1993) that differ between populations was calculated in ARLEQUIN. From these values we calculated average percentage difference between populations, which we report here as sequence divergence (d).

Nuclear introns

Observed heterozygosity (H_O) and expected heterozygosity (H_E) were calculated for each locus, and an exact test of Hardy–Weinberg equilibrium using 100,000 steps in a Markov chain was performed using ARLEQUIN. We tested for linkage disequilibrium between the two nuclear loci using the likelihood ratio test with 20,000 permutations as implemented in ARLEQUIN. Statistical parsimony networks for alleles at each locus were constructed using the program TCS. The multi-locus data set includes only those individuals that amplified at both loci. Using this data set we calculated mean H_E and F_{ST} for the entire data set and for pairwise comparisons between samples with more than five specimens.

RESULTS

Lutjanus kasmira

Mitochondrial DNA

We resolved 475 bp of cyt *b* in 370 individuals of *L. kasmira* (Fig. 1). Owing to low sample size, Lizard Island (n = 5) and Okinawa (n = 2) were excluded from all inter-population analyses. We observed 83 haplotypes, with 56 of these haplotypes occurring in single individuals. The number of individuals sequenced (n), number of haplotypes (N_h), haplotype diversity (h), and nucleotide diversity (π) for each location are provided in Table 1. Overall nucleotide diversity in *L. kasmira* was low (π = 0.004), whereas the corresponding haplotype diversity was high (h = 0.74). Across all samples, π = 0.000–0.005 and h = 0.00–0.88. The most common haplotype (51.2% of specimens) was found at every location except the Marquesas (Fig. 2a). In stark contrast, the Marquesas shared no haplotypes with any other location and formed an isolated cluster in the parsimony network (Fig. 2a).

Overall Φ_{ST} was 0.30 ($P < 0.001$). When we grouped samples by ocean basin (Pacific Ocean: Marquesas, Moorea, Kiritimati, Fiji, Guam, Okinawa and Lizard Island; Indian

Table 1 Molecular diversity indices and coalescence times for cytochrome b sequences from 12 populations of *Lutjanus kasmira*. Sample location, number of individuals sequenced (n), number of haplotypes (N_h), haplotype diversity (h), nucleotide diversity (π), Fu's F-value and corresponding P-value in parentheses, Harpending's raggedness indices (r) with corresponding P-value in parentheses, and mismatch distribution parameters τ, θ_0 and θ_1 as reported by ARLEQUIN 3.11 are listed. The age of each population in years was calculated using the equation $\tau = 2vt$, where t is age in generations and v is the mutation rate per generation for the entire sequence. Effective female population size (N_{ef}) was calculated for time zero and present day using the equation $\theta = 2N_{ef} v$. Coalescence values that could not be resolved are designated by infinity signs (∞). The divergence rate between lineages was estimated to be 1–2% per 10^6 years (Bowen *et al.*, 2001; Lessios, 2008) and generation time was estimated to be 3.7 years (Rangarajan, 1971; Morales-Nin & Ralston, 1990).

| Sample location	n	N_h	h	π	Fu's F_s	r	τ	Population age (years)	θ_0	N_{ef0} ($\times 10^3$)	θ_1	N_{ef1} ($\times 10^3$)
Pacific Ocean												
Marquesas	47	19	0.84 ± 0.04	0.004 ± 0.003	−12.22 (< 0.001)	0.046 (0.54)	1.75	180,000–370,000	0.00	0	824.40	23,000–47,000
Moorea	49	11	0.54 ± 0.08	0.002 ± 0.001	−5.87 (< 0.001)	0.075 (0.85)	1.07	110,000–230,000	0.33	9–19	1.23	40–70
Kiritimati	49	14	0.69 ± 0.07	0.003 ± 0.002	−7.88 (< 0.001)	0.038 (0.90)	1.29	140,000–270,000	0.00	0	7.62	220–430
Fiji	31	9	0.50 ± 0.11	0.001 ± 0.001	−4.03 (< 0.001)	0.086 (0.83)	0.77	80,000–160,000	0.25	7–14	1.52	40–90
Lizard Island	5	1	0.00 ± 0.00	0.000 ± 0.000	–	–	–	–	–	–	–	–
Guam	50	15	0.59 ± 0.08	0.002 ± 0.002	−8.13 (< 0.001)	0.046 (0.91)	0.30	30,000–60,000	1.02	29–58	2.66	80–150
Okinawa	2	1	0.00 ± 0.00	0.000 ± 0.000	–	–	–	–	–	–	–	–
Indian Ocean												
Cocos-Keeling	35	12	0.54 ± 0.10	0.002 ± 0.002	−5.03 (< 0.001)	0.092 (1.00)	∞	∞	0.00	0	∞	∞
Christmas Island	15	8	0.84 ± 0.09	0.003 ± 0.002	−0.84 (0.004)	0.051 (0.75)	0.91	96,000–190,000	0.90	26–51	∞	∞
Diego Garcia	34	15	0.88 ± 0.05	0.005 ± 0.003	−4.39 (< 0.001)	0.069 (0.19)	2.05	210,000–430,000	0.71	19–40	15.45	440–880
Seychelles	19	10	0.74 ± 0.11	0.004 ± 0.003	−3.35 (< 0.001)	0.065 (0.64)	3.06	320,000–640,000	0.00	0	3.45	100–200
Sodwana Bay	34	14	0.82 ± 0.06	0.004 ± 0.003	−3.81 (< 0.001)	0.088 (0.22)	2.81	300,000–590,000	0.01	0	6.50	180–400
All samples	370	83	0.74 ± 0.03	0.004 ± 0.002	−26.15 (< 0.001)	0.013 (0.99)	3.23	340,000–680,000	0.00	0	3.05	90–180
All samples except the Marquesas	323	64	0.67 ± 0.03	0.003 ± 0.002	−26.77 (< 0.001)	0.034 (0.91)	2.27	240,000–480,000	0.00	0	2.14	60–120
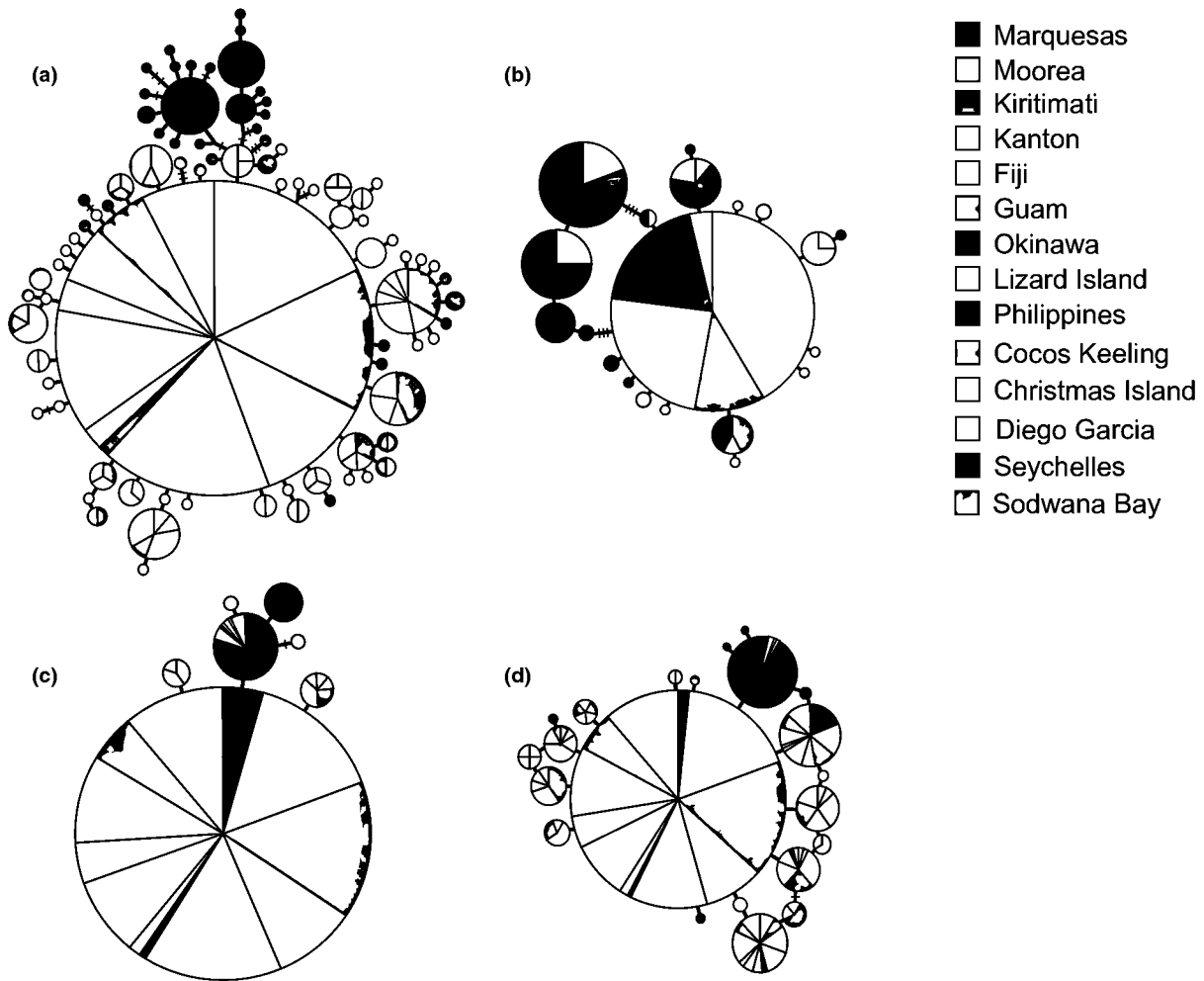
Figure 2 Statistical parsimony networks for *Lutjanus kasmira* (a,c,d) and *Lutjanus fulvus* (b) constructed using TCS 2.21 (Clement *et al.*, 2000) for (a) 370 cytochrome *b* sequences, (b) 203 cytochrome *b* sequences, (c) alleles at the growth hormone (GH) intron from 369 individuals, and (d) alleles at the adenine nucleotide transporter translocase (ANT) intron for 336 individuals. Each circle represents one mitochondrial haplotype or nuclear allele, with the area of each circle proportional to the number of that particular haplotype or allele in the data set; dashes represent hypothetical haplotypes or alleles; colours represent collection location (see key).
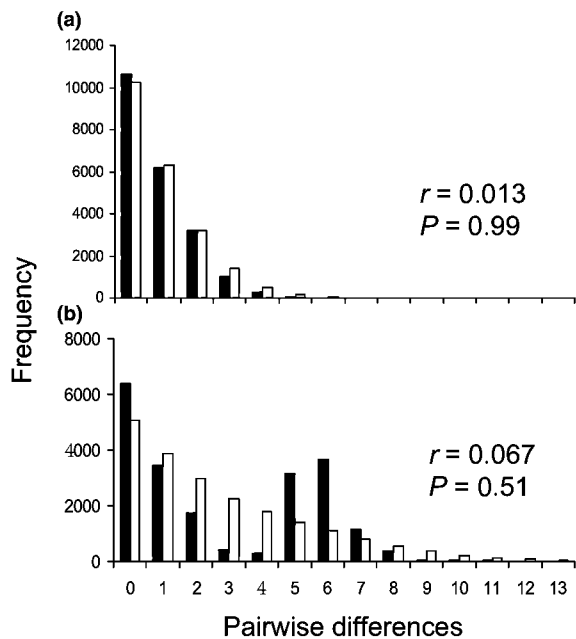
Ocean: Cocos-Keeling, Christmas Island, Diego Garcia, Seychelles and Sodwana Bay) we found no significant structure between the Pacific and Indian oceans ($\Phi_{CT} = -0.003$, $P < 0.59$). Within oceans we found significant structure in the Pacific Ocean ($\Phi_{ST} = 0.45$, $P < 0.001$) but not in the Indian Ocean ($\Phi_{ST} = 0.002$, $P = 0.33$). Pairwise comparisons indicate high levels of population structure between the Marquesas and every other sample location, with significant pairwise $\Phi_{ST} = 0.53–0.63$ ($P < 0.001$) (Table 2). When the phylogenetically distinct Marquesas sample ($d = 0.50–0.53\%$) was removed from the analysis, the overall Φ_{ST} dropped to 0.012 ($P = 0.02$), we found low but significant structure between the Pacific and Indian oceans ($\Phi_{CT} = 0.013$, $P = 0.05$), and Φ_{ST} within the Pacific Ocean was no longer significant ($\Phi_{ST} = 0.012$, $P = 0.12$). Within this data set, Moorea showed low but significant levels of structure when compared with all other sample locations ($\Phi_{ST} = 0.032–0.062$; Table 2). The Indian Ocean locations of Diego Garcia and

Sodwana Bay showed low but significant levels of population structure when compared with all Pacific Ocean locations ($\Phi_{ST} = 0.022–0.045$, $\Phi_{ST} = 0.014–0.043$, respectively; Table 2). The only other significant pairwise comparison was between Kiritimati and Cocos-Keeling ($\Phi_{ST} = 0.022$, $P = 0.015$).

Fu's F_S for the overall data set was -26.15 ($P < 0.001$), indicating an excess of low-frequency haplotypes (Table 1). The mismatch distribution (Fig. 3a) for the overall data set was unimodal (Harpending's raggedness index $r = 0.013$, $P = 0.99$) and resulted in values of $\tau = 3.23$, $\theta_0 = 0.00$ and $\theta_1 = 3.05$. Based on a generation time of 3.7 years and a molecular clock estimate of 1–2% divergence per 10^6 years between species (Lessios, 2008; Bowen *et al.*, 2001; also see Discussion), we calculated a coalescence time of roughly 340,000–680,000 years. The initial female effective population estimate is $N_{ef0} = 0$ and the current effective population estimate is $N_{ef1} = 90,000–180,000$. Coalescence times and

Table 2 Pairwise *F* statistics for 10 populations of *Lutjanus kasmira*. Pairwise Φ_{ST} values for cytochrome *b* data are below diagonal and pairwise F_{ST} values for the multi-locus nuclear data set are above diagonal. Values in bold are significant ($P < 0.05$).

Sample location	Marquesas	Moorea	Kiritimati	Fiji	Guam	Cocos-Keeling	Christmas Island	Diego Garcia	Seychelles	Sodwana Bay
Marquesas	–	**0.559**	**0.549**	**0.492**	**0.533**	**0.496**	**0.498**	**0.527**	**0.441**	**0.530**
Moorea	**0.634**	–	−0.001	0.007	0.005	0.010	**0.048**	0.010	0.026	0.005
Kiritimati	**0.590**	0.062	–	−0.003	0.000	0.002	0.023	0.006	**0.026**	−0.001
Fiji	**0.618**	0.049	−0.006	–	−0.002	0.000	0.015	0.002	0.008	−0.002
Guam	**0.609**	0.049	−0.007	−0.011	–	0.003	0.014	−0.004	0.010	−0.009
Cocos-Keeling	**0.610**	0.032	**0.022**	0.002	0.004	–	0.010	−0.008	0.010	−0.004
Christmas Island	**0.564**	0.051	−0.007	−0.008	−0.005	0.017	–	0.016	0.021	0.018
Diego Garcia	**0.534**	0.045	0.036	0.022	0.024	0.001	0.006	–	0.013	−0.008
Seychelles	**0.543**	0.045	0.018	0.004	0.011	0.001	−0.005	−0.015	–	0.007
Sodwana Bay	**0.538**	0.043	**0.032**	**0.018**	**0.014**	0.002	0.008	−0.003	−0.005	–

(a)

$r = 0.013$
$P = 0.99$

(b)

$r = 0.067$
$P = 0.51$

Frequency

0 1 2 3 4 5 6 7 8 9 10 11 12 13

Pairwise differences

Figure 3 Mismatch distribution based on (a) 370 cytochrome *b* sequences from 12 populations of *Lutjanus kasmira*, (b) 203 cytochrome *b* sequences from six populations of *Lutjanus fulvus*. Black bars are the observed and white bars are the simulated pairwise differences as reported by Arlequin 3.11. Harpending's raggedness indices and corresponding *P*-values for the complete data sets are shown.

demographic parameter estimates for each population are reported in Table 1.

Nuclear introns

We resolved 148 bp of the GH intron in 369 specimens and 168 bp of the ANT intron in 336 specimens of *L. kasmira* (Table 3). Seven polymorphic sites yielded seven alleles at the GH locus, and 16 polymorphic sites yielded 22 alleles at the ANT locus. Statistical parsimony networks showed that alleles at each locus are closely related (Fig. 2c,d). The most common allele at each locus was observed at all sample locations across the range. The second most common allele at both loci was observed primarily in the Marquesas Islands, but also at low frequencies elsewhere. The number of individuals sequenced (n), number of alleles (N_a), observed heterozygosity (H_O), expected heterozygosity (H_E), and the corresponding *P*-value for the exact test for Hardy–Weinberg equilibrium for each locus are listed in Table 3. Number of genotypes (N_g) and mean H_E are also listed for the multi-locus data set (Table 3). The samples from Diego Garcia were found to be out of Hardy–Weinberg equilibrium, with an excess of homozygotes at both the GH and ANT loci ($P = 0.03$ and 0.03 respectively), whereas the samples from the Marquesas, Moorea and Guam were found to have an excess of homozygotes only at the ANT locus ($P = 0.04$, 0.03 and > 0.01 respectively) (Table 3). Overall H_E was 0.25, 0.61 and 0.43 for the GH, ANT and multi-locus data sets respectively. Across all samples, $H_E = 0.00$–0.59 for the GH intron, $H_E = 0.38$–0.83 for the ANT intron, and mean $H_E = 0.21$–0.48 for the multi-locus nuclear data set. There was no indication of linkage disequilibrium between the two loci ($P > 0.05$).

Overall F_{ST} for the GH, ANT and the multi-locus data set were $F_{ST} = 0.41$ ($P < 0.001$), $F_{ST} = 0.21$ ($P < 0.001$) and $F_{ST} = 0.28$ ($P < 0.001$), respectively. Using the multi-locus data set we grouped samples by ocean basin (Pacific Ocean: Marquesas, Moorea, Kiritimati, Fiji, Guam, Okinawa and Lizard Island; Indian Ocean: Cocos-Keeling, Christmas Island, Diego Garcia, Seychelles and Sodwana Bay). Similar to the case for cyt *b* data, we found no significant structure between the Pacific and Indian oceans ($F_{CT} = -0.014$, $P = 0.77$), significant structure within the Pacific Ocean ($F_{ST} = 0.38$, $P < 0.001$), and no structure within the Indian Ocean ($F_{ST} = 0.004$, $P = 0.32$). Pairwise F_{ST} values for the multi-locus data set are reported in Table 2. Population comparisons indicate highly significant population structure between the Marquesas and every other sample location, with pairwise $F_{ST} = 0.44$–0.56. When the Marquesas was excluded from the analysis, overall F_{ST} dropped to 0.004 ($P = 0.23$), we found low but significant structure between the Pacific and Indian oceans ($F_{CT} = 0.004$, $P = 0.02$), and F_{ST} within the Pacific

Table 3 Molecular diversity indices for 12 populations of *Lutjanus kasmira* for the growth hormone (GH) and adenine nucleotide transporter translocase (ANT) nuclear introns. Sample location, number of individuals sequenced (n), number of alleles (N_a), observed heterozygosity (H_O) and expected heterozygosity (H_E) are listed for each intron. P-values are the result of exact tests for Hardy–Weinberg equilibrium using a Markov chain with 100,000 steps in Arlequin 3.11. Number of genotypes (N_g) and mean H_E are listed for the multi-locus data set, which includes only those individuals that amplified at both loci.

	GH intron					ANT intron					Multi-locus		
Sample location	n	N_a	H_O	H_E	P-value	n	N_a	H_O	H_E	P-value	n	N_g	H_E
Pacific Ocean													
Marquesas	49	3	0.66	0.59	0.58	49	6	0.29	0.38	0.04	48	12	0.48
Moorea	50	3	0.10	0.10	1.00	47	5	0.32	0.38	0.03	47	8	0.24
Kiritimati	50	3	0.06	0.06	1.00	48	10	0.44	0.46	0.07	48	12	0.26
Fiji	31	3	0.13	0.12	1.00	29	8	0.64	0.55	0.90	28	9	0.34
Lizard Island	5	1	0.00	0.00	1.00	4	2	0.50	0.43	1.00	4	2	0.21
Guam	49	2	0.04	0.04	1.00	35	9	0.40	0.52	< 0.01	35	10	0.27
Okinawa	2	1	0.00	0.00	1.00	2	3	1.00	0.83	1.00	2	2	0.42
Indian Ocean													
Cocos-Keeling	30	5	0.13	0.19	0.06	25	10	0.40	0.51	0.10	25	13	0.33
Christmas Island	14	1	0.00	0.00	1.00	15	8	0.40	0.58	0.08	14	8	0.31
Diego Garcia	33	3	0.06	0.12	0.03	32	10	0.48	0.55	0.03	25	8	0.27
Seychelles	20	2	0.15	0.22	0.25	19	8	0.68	0.59	0.64	19	10	0.41
Sodwana Bay	36	3	0.06	0.06	1.00	34	11	0.47	0.51	0.51	34	12	0.29
All samples	369	7	0.16	0.25	< 0.001	336	22	0.43	0.61	< 0.001	329	34	0.43

Ocean was no longer significant ($F_{ST} = -0.0001$, $P = 0.50$). The only other significant pairwise comparisons were between Moorea and Christmas Island ($F_{ST} = 0.048$, $P = 0.04$) and between Kiritimati and the Seychelles ($F_{ST} = 0.026$, $P = 0.034$).

Concordance between mtDNA and nuclear markers

We found strong agreement between our mitochondrial cyt *b* and multi-locus nuclear intron data sets for *L. kasmira*. The two marker types show the same pattern of the Marquesas population being highly divergent from all other populations, with only low levels of structure elsewhere. The marker types differ only in the degree of genetic differentiation demonstrated, with the Marquesas sharing no mitochondrial haplotypes with any other population (Fig. 2a), whereas the nuclear markers show strong shifts in allele frequency between populations with the most common allele at each locus being found in every population (Fig. 2c,d). These differences may result from the fourfold lower effective population size of the mitochondrial genome (Avise, 2004), although other factors such as differing mutation rates probably contribute as well.

Lutjanus fulvus

Mitochondrial DNA

We resolved 480 bp of cyt *b* in 203 individuals of *L. fulvus* (Fig. 1), yielding 23 haplotypes, with eight of these haplotypes observed in single individuals. Sample sizes and diversity estimates for each location are provided in Table 4. Similar to

the case for *L. kasmira*, overall nucleotide diversity in *L. fulvus* was low ($\pi = 0.006$) whereas the corresponding haplotype diversity was high ($h = 0.69$). Across all samples, $\pi = 0.0003$–0.006 and $h = 0.12$–0.78. Similar to the case for *L. kasmira*, the most common haplotype (53.2% of specimens) was found at every location except the Marquesas (Fig. 2b).

Overall population structure was much higher in this species than in *L. kasmira* ($\Phi_{ST} = 0.64$, $P < 0.001$). All but three of the 15 population-level pairwise comparisons were significant ($P < 0.05$; Table 5). The three non-significant comparisons were between the Philippines and Kiritimati, the Philippines and Cocos-Keeling, and Cocos-Keeling and Kanton. Similar to the findings for the *L. kasmira* data, the highest levels of structure were found at the Marquesas, with significant pairwise values ranging from $\Phi_{ST} = 0.66$–0.91 ($P < 0.001$) (Table 5). When the Marquesas sample ($d = 0.87$–1.50%) was removed from the analysis, the overall Φ_{ST} dropped to 0.11 ($P < 0.001$).

Fu's F_S for the overall data set was -5.85 ($P < 0.001$), indicating an excess of low-frequency haplotypes (Table 4). The mismatch distribution (Fig. 3b) for the overall data set was roughly unimodal (Harpending's raggedness index $r = 0.067$, $P = 0.51$) and resulted in values of $\tau = 7.00$, $\theta_0 = 0.00$ and $\theta_1 = 3.02$. Based on a generation time calculated for *L. kasmira* (3.7 years) and a molecular clock estimate of 1–2% divergence per 10^6 years between lineages, we calculated a coalescence time of roughly 730,000–1,460,000 years. The initial female effective population estimate is $N_{ef0} = 0$ and the current effective population estimate is $N_{ef1} = 85,000$–170,000. Coalescence times and demographic parameter estimates for each population are reported in Table 4.

Table 4 Molecular diversity indices and coalescence times for cytochrome b sequences from six populations of *Lutjanus fulvus*. Sample location, number of individuals sequenced (n), number of haplotypes (N_h), haplotype diversity (h), nucleotide diversity (π), and mismatch distribution parameters τ, θ_0 and θ_1 as reported by ARLEQUIN 3.11 are listed. The age of each population in years was calculated using the equation $\tau = 2vt$, where t is age in generations and v is the the mutation rate per generation for the entire sequence. Effective female population size (N_{ef}) was calculated for time zero and present day using the equation $\theta = 2N_{ef}v$. Coalescence values that could not be resolved are designated by infinity signs (∞). The divergence rate between lineages was estimated to be 1–2% per 10^6 years (Bowen *et al.*, 2001; Lessios, 2008) and generation time was estimated to be 3.7 years (Rangarajan, 1971; Morales-Nin & Ralston, 1990).

Sample location	n	N_h	h	π	Fu's F_s	r	τ	Population age (years)	θ_0	N_{ef0} ($\times 10^3$)	θ_1	N_{ef1} ($\times 10^3$)
Pacific Ocean												
Marquesas	48	7	0.72 ± 0.04	0.002 ± 0.002	−2.06 (0.003)	0.142 (0.02)	1.15	120,000–240,000	0.00	0	∞	∞
Moorea	48	4	0.12 ± 0.06	0.0003 ± 0.001	−4.22 (< 0.001)	0.588 (0.72)	3.00	300,000–630,000	0.00	0	0.15	4–8
Kiritimati	15	2	0.34 ± 0.13	0.001 ± 0.001	0.60 (0.94)	0.216 (0.39)	0.46	50,000–100,000	0.00	0	∞	∞
Kanton	46	10	0.66 ± 0.07	0.006 ± 0.003	−1.31 (0.002)	0.070 (0.68)	7.11	740,000–1,480,000	0.00	0	2.16	60–120
Philippines	37	9	0.66 ± 0.08	0.003 ± 0.002	−2.95 (< 0.001)	0.042 (0.88)	2.10	220,000–440,000	0.00	0	2.34	70–130
Indian Ocean												
Cocos-Keeling	9	4	0.78 ± 0.11	0.003 ± 0.002	2.57 (0.139)	0.015 (0.44)	12.08	1,200,000–2,500,000	0.00	0	4.46	125–250
All samples	203	23	0.69 ± 0.03	0.006 ± 0.003	−5.85 (< 0.001)	0.067 (0.51)	7.00	730,000–1,460,000	0.00	0	3.02	85–170
All samples except the Marquesas	155	18	0.51 ± 0.05	0.002 ± 0.002	−8.59 (< 0.001)	0.010 (1.00)	∞	∞	0.00	0	533.95	15,000–30,000

DISCUSSION

Our survey of the snappers *L. kasmira* and *L. fulvus* revealed significant levels of genetic structure across their ranges. However, pairwise population comparisons indicated that most of this structure results from the extreme isolation of the Marquesan population in each species. This location is divergent from all other samples (*L. kasmira*, $d = 0.50$–0.53%; *L. fulvus*, $d = 0.87$–1.50%). Once the phylogenetically distinct Marquesan population was removed from the analyses, population structure dropped dramatically in both taxa, revealing contrasting patterns of population structure between the two species. Our results indicate that *L. kasmira* is a highly dispersive fish that displays low but significant genetic structure between the Pacific and Indian oceans, but no population structure across 12,000 km of its central range from Kiritimati in the central Pacific to Cocos-Keeling in the eastern Indian Ocean. In contrast, the congeneric *L. fulvus* demonstrated significant population structure at every geographical scale we examined across the same 12,000 km, including a break across the IPB that was of similar magnitude to other pairwise comparisons. Prior to dissecting these results, we address two caveats.

1. Several authors have calibrated molecular clocks for cyt b in marine fishes based on the closure of the Isthmus of Panama. Bowen *et al.* (2001) found a divergence rate of 1.8–2.2% Myr^{-1} for trumpetfishes (*Aulostomus* spp.). Lessios (2008) summarized data for 16 pairs of sister taxa distributed across the Isthmus and found sequence divergences of 1.7–21.6%. In the five pairs of sister taxa that Lessios (2008) included for which divergence is assumed to have been initiated at final closure of the Isthmus (3 Ma), divergence rates were 1.1–1.6% Myr^{-1} between lineages. Based on these studies, and to convey an appropriate caution, we choose the approximation of 1–2% Myr^{-1} between lineages (0.5–1.0% Myr^{-1} within lineages). Regardless of the specific approximation we use, our data always indicate a Pleistocene expansion. Thus for data interpretation we use the rank order of the coalescence times and do not attempt to make conclusions based on precise values.

2. Cases in which mitochondrial and nuclear data sets reveal fundamentally different patterns typically involve taxa with sex-biased dispersal (sea turtles, Bowen *et al.*, 2005; sharks, Pardini *et al.*, 2001; dolphins, Möller & Beheregaray, 2004). However, this pattern has never been documented in reef fishes. Here we found strong genealogical concordance among three independent loci in *L. kasmira*, and concordance (regarding the isolation of the Marquesas) across two co-distributed species (see Avise, 2004). In light of these findings we did not sequence the two nuclear introns in *L. fulvus*.

Effects of the Indo-Pacific Barrier

The shallow Sunda Shelf defines the corridor between the tropical Pacific and Indian oceans. Lowering of sea level, to as much as 120 m below present levels, occurred at least three

Table 5 Pairwise F statistics for six populations of *Lutjanus fulvus*. Pairwise Φ_{ST} values for cytochrome b data are below the diagonal and corresponding P-values are above the diagonal. Values in bold are significant ($P < 0.05$).

Sample location	Marquesas	Moorea	Kiritimati	Kanton	Philippines	Cocos-Keeling
Marquesas	–	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Moorea	**0.912**	–	0.011	< 0.001	0.002	< 0.001
Kiritimati	**0.873**	0.157	–	0.073	0.205	0.010
Kanton	**0.660**	0.153	0.086	–	0.004	0.050
Philippines	**0.825**	0.071	0.022	**0.104**	–	0.533
Cocos-Keeling	**0.856**	**0.300**	0.153	0.114	−0.014	–

times during the Pleistocene (Voris, 2000; Naish *et al.*, 2009), greatly reducing shallow-water habitat in tropical and subtropical seas and probably causing dramatic declines in populations of shallow-depth fauna. In many species that are found in both oceans, a signature of historical isolation can be detected in the intraspecific genetic differentiation between oceans. A search of the literature uncovered studies of 18 species of marine fishes and invertebrates that have been sampled across the IPB (Table 6). Of these, 83% (15 of 18) exhibited significant structure across the IPB. To date, the only

exceptions are the sea urchin *Diadema savignyi* (Lessios *et al.*, 2001), the trumpetfish *Aulostomus chinensis* (Bowen *et al.*, 2001), and the surgeonfish *Naso brevirostris* (Horne *et al.*, 2008). Here we report that *L. kasmira* also shows no evidence of a significant genetic break across the IPB. Although we did find significant structure between the Pacific and Indian Ocean basins in *L. kasmira*, most of that divergence results from the geographically remote island of Diego Garcia and Sodwana Bay at the end of the species' range. The other three Indian Ocean locations, including the Seychelles in the western

Table 6 Genetic surveys of population structure in reef organisms across the Indo-Pacific Barrier. Species scientific name, marker type, genetic structure between the Pacific and Indian oceans (as reported) and references are listed.

Species	Marker	Genetic structure	Reference
Soldierfish			
Myripristies berndti	mtDNA	$\Phi_{ST} = 0.58$ ($P < 0.001$)	Craig *et al.* (2007)
Trumpetfish			
Aulostomus chinensis	mtDNA	$\Phi_{ST} = -0.027$ ($P = 0.805$) not reported in reference	Bowen *et al.* (2001)
Surgeonfish			
Acanthurus triostegus	Allozymes	$F_{ST} = 0.21-0.25$ ($P < 0.05$)	Planes & Fauvelot (2002)
Naso vlamingii	mtDNA	$\Phi_{ST} = 0.077$ ($P < 0.05$)	Klanten *et al.* (2007)
Naso unicornis	mtDNA	$\Phi_{ST} = 0.018$ ($P = 0.02$)	Horne *et al.* (2008)
Naso brevirostris	mtDNA	$\Phi_{ST} = 0.030$ ($P = 0.08$)	
Damselfish			
Stegastes nigricans	Allozymes	Fixed allele differences	Lacson & Clark (1995)
Chrysiptera biocellata			
Chrysiptera glauca			
Chrysiptera eucopoma			
Parrotfish			
Chlorurus sordidus	mtDNA	Fixed haplotype differences	Bay *et al.* (2004)
Seastar			
Linkia laevigata	mtDNA	$\Phi_{ST} = 0.332$ ($P < 0.001$)	Williams & Benzie (1998)
	Allozymes	$F_{ST} = 0.083$ ($P < 0.001$)	
Acanthaster planci	Allozymes	$F_{ST} = 0.273$ ($P < 0.001$)	Benzie (1999)
Sea urchins			
Tripneustes gratilla	mtDNA	$\Phi_{ST} = -0.01-0.40$ ($P < 0.05$ for 15 out of 32 pairwise comparisons)	Lessios *et al.* (2003)
Diadema savignyi	Isozymes	$F_{ST} = -0.01-0.19$ ($P < 0.05$ for 2 out of 21 pairwise comparisons)	Lessios *et al.* (2001)
Diadema setosum	Isozymes	$F_{ST} = 0.36-0.73$ ($P < 0.05$ for 16 out of 16 pairwise comparisons)	
Coconut crab			
Birgus latro	Allozymes	$F_{ST} = 0.082$ ($P < 0.001$)	Lavery *et al.* (1995, 1996)
Tiger prawn			
Penaeus monodon	Nuclear intron	$F_{ST} = 0.51$ ($P < 0.001$)	Duda & Palumbi (1999)

Indian Ocean, demonstrate no significant population differentiation when compared with the Pacific Ocean. *Lutjanus fulvus* has a significant population partition at the IPB, but the level of that differentiation is no higher than detected elsewhere (excluding the Marquesas). In both species, levels of divergence occurring across the IPB are similar to those among other populations, indicating that this barrier does not have a strong impact on either species.

Isolation of the Marquesas Islands

The Marquesas Islands are volcanic in origin and lie between 8° S, 140° W and 10° S, 138° W on the south-eastern edge of Oceania, 475 km north of the Tuamotu Islands, 1050 km east of the Caroline Islands, and 4700 km west of the American mainland. The shore-fish fauna has a distinctly lower diversity than in the rest of Polynesia (Randall, 2001). In contrast to the widely recognized IPB, which did not impede dispersal of either species of *Lutjanus*, the barrier at the Marquesas proved to be substantial. With no obvious geographical obstacles between the Marquesas and other South Pacific archipelagos, what could have produced the high level of population differentiation in the Marquesas indicated by its 11.6% endemism rate, and the occurrence of strong genetic differentiation in three of five species of widely distributed reef fishes?

There have been three previous genetic studies of Marquesan shore fish, with one indicating strong differentiation of the Marquesan population (Planes & Fauvelot, 2002), and the two others indicating no differentiation (Craig *et al.*, 2007; Schultz *et al.*, 2007). Our data show a strong phylogeographic break between the Marquesas and populations scattered throughout the rest of the range in both *Lutjanus* species. These shared breaks stand in contrast to the very different patterns of intraspecific population structure across the central west Pacific in these species: very little structure in *L. kasmira* versus strong structure in *L. fulvus*. Yet both strongly and weakly dispersive snappers demonstrate an evolutionary genetic break at the Marquesas.

Randall (2001) attributed the endemism of the Marquesas to a combination of geographical isolation (a biogeographical barrier to dispersal) and unusually variable sea temperatures for an equatorial archipelago (ecological distinctiveness). However, the extensive Tuamotu archipelago is < 500 km away from the Marquesas, a relatively small distance for a highly dispersive species such as *L. kasmira*, which shows strong genetic connectivity across much greater distances elsewhere throughout the Indo-central Pacific.

Directionality of the prevailing ocean current may amplify the effects of distance in the case of the Marquesas. The Southern Equatorial Current (SEC) flows from east to west between 4° N and 17° S (Wyrtki & Kilonsky, 1983; Bonjean & Lagerloef, 2002), with the Marquesas located on the edge of the strongest portion of that current (Fig. 4). The SEC originates 6000 km to the east of the Marquesas, where it draws water from an area with a temperate fish fauna. This combination of vast distance and inappropriate source fauna effectively isolates
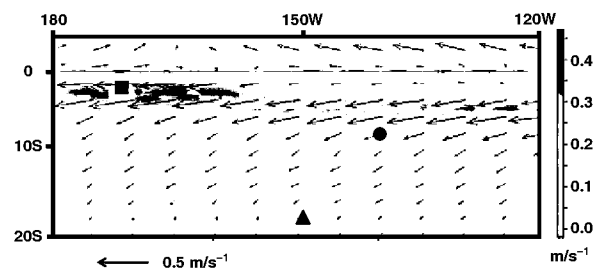

Figure 4 Mean ocean-surface current vectors from May to September, when spawning in snappers is most likely to occur, for 1997–2007. The orientation of the arrow indicates the direction of current and the length of the arrow indicates the current speed (0.514 m s^{-1} = 1 knot). Coloured scale bar represents mean current speed (m s^{-1}). Data obtained from http://www.oscar.noaa.gov; accessed 10 June 2009 (Bonjean & Lagerloef, 2002). Circle, Marquesas Island of Nuku Hiva; triangle, Moorea; square, Kanton Island.

the Marquesas from the east. The most likely direction of gene flow is from the Marquesas westwards. Our data indicate such directionality in gene flow in *L. fulvus*: the single population downstream from the Marquesas that we were able to sample (Kanton Island, 3500 km from the Marquesas) was the only location that shared haplotypes with the Marquesan population. The lack of Marquesan haplotypes of both *L. fulvus* and *L. kasmira* at two sites that are much closer to but not downstream from the Marquesas (Moorea to the south and Kiritimati to the north) is consistent with such a pattern of dispersal. Unfortunately we were unable to sample *L. kasmira* at Kanton. Further sampling of both species is needed at the reef systems nearest the Marquesas (the Tuamotus to the south and the Caroline Islands to the west) to test our hypothesis about the SEC, and establish the extent of any 'leakage' from the Marquesas.

Unusual local environmental conditions may reinforce the biogeographical isolation of the Marquesan fish fauna (see Rocha & Bowen, 2008). Unlike the nearest Polynesian islands to the south-west, the Marquesas include an upwelling zone, have highly variable seawater temperatures, and little coral reef development. Those conditions may reduce the viability of colonists leaving or entering the Marquesas. Such a pattern of varying success has been documented in Atlantic reef fishes (Rocha *et al.*, 2005). Thus, both a significant, although localized biogeographical barrier (distance enhanced by current directionality) and strongly divergent environmental conditions on either side of that barrier may have contributed to the genetic distinctiveness of Marquesan shore-fish populations and the high endemism at these islands.

Glacial refugia and rapid expansions

Our cyt *b* data for both species exhibit low nucleotide diversity, high to moderate haplotype diversity (Tables 1 and 4) and a star-shaped parsimony network (Fig. 2a,b), a common mtDNA pattern in marine fishes (Grant & Bowen, 1998).

These patterns are characteristics of a historical expansion in population size, a hypothesis supported by the unimodal mismatch distribution (Fig. 3a,b), the negative Fu's F-values and the θ_0 and θ_1 values reported in Tables 1 and 4. This pattern may be a signature of severe population reductions attributed to the dramatic decreases in sea level during the Pleistocene period (Grant & Bowen, 1998). Rapid re-colonization of Oceania following sea-level rise is evidenced by the low level of endemism in the region (Briggs, 1999) and a lack of intraspecific population structure across the central west Pacific in a variety of marine organisms including snappers (*L. kasmira*; this study, excluding the Marquesas), trumpetfish (*Aulostomus chinensis*; Bowen *et al.*, 2001), soldierfish (*Myripristis berndti*; Craig *et al.*, 2007), angelfish (*Centropyge loriculus*; Schultz *et al.*, 2007), parrotfish (*Chlorurus sordidus*; Bay *et al.*, 2004), surgeonfish (*Naso* spp.; Klanten *et al.*, 2007; Horne *et al.*, 2008), sea urchins (Lessios *et al.*, 2001) and sea stars (*Linkia laevigata*; Williams & Benzie, 1998). These data indicate that for many marine species the stretches of open water throughout the central and western Pacific Ocean are not significant barriers to gene flow. Dispersal may be facilitated by the relatively short distances between islands in that area (Keeney & Heist, 2006; Schultz *et al.*, 2008). Departures from this pattern of high connectivity include snappers (*L. fulvus*; this study), damselfish (*Dascyllus trimaculatus*; Bernardi *et al.*, 2001; *Dascyllus albisella*; Ramon *et al.*, 2008), surgeonfish (*Acanthurus triostegus*; Planes & Fauvelot, 2002) and sea urchins (*Tripneustes gratilla*; Lessios *et al.*, 2003; *Echinometra mathaei*; Palumbi *et al.*, 1997), and are probably the result of (unidentified) taxon-specific factors.

Our *L. kasmira* data set shows no evidence of genetic partitioning within the Indian Ocean, a result similar to that for populations of *M. berndti* (Craig *et al.*, 2007) and *Naso unicornis* (Horne *et al.*, 2008), in which no significant structure in the mitochondrial genome was found between Cocos-Keeling and the Seychelles. In contrast, populations of *N. brevirostris* from these two locations demonstrated low but significant genetic structure (Horne *et al.*, 2008). Because so few studies have sampled more than a few locations in the Indian Ocean, it is difficult to generalize regarding dispersal patterns in this ocean, but initial data sets indicate high levels of gene flow, at least among reef fish populations.

Coalescence times for *L. kasmira* (Population age, Table 1) show a trend of decreasing age from west to east, with the oldest and most genetically diverse populations at the opposite ends of the species' range (the Seychelles and Sodwana Bay in the western Indian Ocean and the Marquesas in the central Pacific). Although our data are not conclusive, lines of evidence including coalescence times (Table 1), genetic diversity indices (Table 1 and multi-locus data from Table 3) and the mtDNA statistical parsimony network (Fig. 2a) all support a scenario of glacial refugia at the ends of the range, with populations elsewhere undergoing dramatic reductions or extirpations. As the glaciers retreated and sea level rose, the eastern Indian and Pacific oceans seem to have been replenished from the western Indian Ocean. Mitochondrial sequence divergences indicate that the Marquesan population remained isolated from the rest of the species' range throughout this period of geographical expansion. This pattern is of great interest because it indicates: (1) that *L. kasmira* was able to expand eastwards through the Indo-central Pacific (i.e. against the prevailing westward flows of major ocean currents); and (2) the mechanisms isolating the Marquesan population were strong enough to prevent it from acting as a source for repopulation of depleted reefs elsewhere in the central Pacific.

CONCLUSIONS

Our genetic survey of the snappers *L. kasmira* and *L. fulvus* revealed a strongly contrasting pattern of population subdivision in these two species. Whereas *L. kasmira* demonstrated no population structure across the IPB, the Indian Ocean and (most of the) Pacific Ocean, *L. fulvus* proved to be a highly structured species throughout the same range, with population structure across the IPB concordant with other pairwise comparisons. Despite differences in population structure, both fishes demonstrate a remarkably strong phylogeographic break at the Marquesas Islands. The isolation of Marquesan populations may have arisen through a biogeographical barrier to inward dispersal (distance and contrarian ocean currents) and an unusual local environment that inhibits the survival of propagules from outside the archipelago. Coalescence analyses for *L. kasmira* prompt the hypothesis of glacial refugia in the western Indian Ocean and the Marquesas, with range expansion from eastern Africa through the Indo-central Pacific reinforcing the view that strong and enduring mechanisms have isolated the Marquesan population over evolutionary time.

Mindy Mizobe of the HIMB EPSCoR core facility for their assistance with DNA sequencing, and Shelley Jones and all the members of the ToBo lab for their assistance, support and feedback throughout this project. We thank Jenny Schultz, David Bellwood and two anonymous referees, whose comments substantially improved the manuscript. This is contribution no. 1352 from the Hawaii Institute of Marine Biology and no. 7763 from the School of Ocean and Earth Science and Technology.

REFERENCES

Allen, G.R. & Talbot, J.H. (1985) Review of the snappers of the genus *Lutjanus* (Pisces: Lutjanidae) from the Indo-Pacific, with the description of a new species. *Indo-Pacific Fishes*, No. 11. Bernice Pauahi Bishop Museum, Honolulu.

Avise, J.C. (2004) *Molecular markers, natural history, and evolution*, 2nd edn. Sinauer Associates, Sunderland, MA.

Ayre, D.J. & Hughes, T.P. (2004) Climate change, genotypic diversity and gene flow in reef-building corals. *Ecology Letters*, 7, 273–278.

Banford, H.M., Bermingham, E., Collette, B.B. & McCafferty, S.S. (1999) Phylogenetic systematics of the *Scomberomorus regalis* (Teleostei: Scombridae) species group: molecules, morphology and biogeography of Spanish mackerels. *Copeia*, 1999, 596–613.

Barber, P.H., Palumbi, S.R., Erdmann, M.V. & Moosa, M.K. (2000) A marine Wallace's line? *Nature*, 406, 692–693.

Barber, P.H., Erdmann, M.V. & Palumbi, S.R. (2006) Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the Coral Triangle. *Evolution*, 60, 1825–1839.

Bay, L.K., Choat, J.H., van Herwerden, L. & Robertson, D.R. (2004) High genetic diversities and complex genetic structure in an Indo-Pacific tropical reef fish (*Chlorurus sordidus*): evidence of an unstable evolutionary past? *Marine Biology*, 144, 757–767.

Bellwood, D.R. & Wainwright, P.C. (2002) The history and biogeography of fishes on coral reefs. *Coral reef fishes: dynamics and diversity in a complex ecosystem* (ed. by P.F. Sale), pp. 5–32. Academic Press, San Diego.

Benzie, J.A.H. (1999) Major genetic differences between crown-of-thorns starfish (*Acanthaster planci*) populations in the Indian and Pacific Oceans. *Evolution*, 53, 1782–1795.

Benzie, J.A.H. & Stoddart, J.A. (1992) Genetic structure of crown-of-thorns starfish *Acanthaster planci* in Australia. *Marine Biology*, 112, 631–639.

Bernardi, G., Holbrook, S.J. & Schmitt, R.J. (2001) Gene flow at three spatial scales in a coral reef fish, the three-spot dascyllus, *Dascyllus trimaculatus*. *Marine Biology*, 138, 457–465.

Bonjean, F. & Lagerloef, G.S.E. (2002) Diagnostic model and analysis of the surface currents in the tropical Pacific Ocean. *Journal of Physical Oceanography*, 32, 2938–2954.

Bowen, B.W., Bass, A.L., Rocha, L.A., Grant, W.S. & Robertson, D.R. (2001) Phylogeography of the trumpetfishes

(*Aulostomus*): ring species complex on a global scale. *Evolution*, 55, 1029–1039.

Bowen, B.W., Bass, A.L., Soares, L. & Toonen, R.J. (2005) Conservation implications of complex population structure: lessons from the loggerhead turtle (*Caretta caretta*). *Molecular Ecology*, 14, 2389–2402.

Briggs, J.C. (1974) *Marine zoogeography*. McGraw-Hill, New York.

Briggs, J.C. (1999) Extinction and replacement in the Indo-West Pacific Ocean. *Journal of Biogeography*, 26, 777–783.

Brown, J.H. & Lomolino, M.V. (1998) *Biogeography*, 2nd edn. Sinauer Associates, Sunderland, MA.

Castro, A.L.F., Stewart, B.S., Wilson, S.G., Hueter, R.E., Meekan, M.G., Motta, P.J., Bowen, B.W. & Karl, S.A. (2007) Population genetic structure of Earth's largest fish, the whale shark (*Rhincodon typus*). *Molecular Ecology*, 16, 5183–5192.

Clement, M., Posada, D. & Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, 9, 1657–1659.

Craig, M.T., Eble, J.A., Bowen, B.W. & Robertson, D.R. (2007) High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Marine Ecology Progress Series*, 334, 245–254.

Denit, K. & Sponaugle, S. (2004) Growth variation, settlement, and spawning of gray snapper across a latitudinal gradient. *Transactions of the American Fisheries Society*, 133, 1339–1355.

Duda, T.F., Jr & Palumbi, S.R. (1999) Population structure of the black tiger prawn, *Penaeus monodon*, among western Indian Ocean and western Pacific populations. *Marine Biology*, 134, 705–710.

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

Fleminger, A. (1986) The Pleistocene equatorial barrier between the Indian and Pacific Oceans and a likely cause for Wallace's line. *UNESCO Technical Papers in Marine Science*, 49, 84–97.

Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., Edwards, A.J., Barreiros, J.P., Ferreira, C.E.L., Gasparini, J.L., Brito, A., Falcón, J.M., Bowen, B.W. & Bernardi, G. (2008) Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, 35, 22–47.

Friedlander, A.M., Parrish, J.D. & DeFelice, R.C. (2002) Ecology of the introduced snapper *Lutjanus kasmira* (Forsskal) in the reef fish assemblage of a Hawaiian bay. *Journal of Fish Biology*, 60, 28–48.

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

Galloway, R.W. & Kemp, E.M. (1981) Late Cenozoic environments in Australia. *Ecological biogeography of Australia* (ed. by A. Kearst), pp. 51–80. Junk, Hague.

Grant, W.S. & Bowen, B.W. (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights

from sardines and anchovies and lessons for conservation. *Journal of Heredity*, **89**, 415–426.

Harpending, H.C. (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Hassan, M., Lemaire, C., Fauvelot, C. & Bonhomme, F. (2002) Seventeen new exon-primed intron-crossing polymerase chain reaction amplifiable introns in fish. *Molecular Ecology Notes*, **2**, 334–340.

Hobbs, J.-P.A., Frisch, A.J., Allen, G.R. & Van Herwerden, L. (2009) Marine hybrid hotspot at Indo-Pacific biogeographic border. *Biology Letters*, **5**, 258–261.

Horne, J.B., van Herwerden, L., Choat, H.J. & Robertson, D.R. (2008) High population connectivity across the Indo-Pacific: congruent lack of phylogeographic structure in three reef fish congeners. *Molecular Phylogenetics and Evolution*, **49**, 629–638.

Hudson, R.R. & Turelli, M. (2003) Stochasticity overrules the "three-times rule": genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, **51**, 182–190.

Jarman, S.N., Ward, R.D. & Elliott, N.G. (2002) Oligonucleotide primers for PCR amplification of coelomate introns. *Marine Biotechnology*, **4**, 347–355.

Keeney, D.B. & Heist, E.J. (2006) Worldwide phylogeography of the blacktip shark (*Carcharhinus limbatus*) inferred from mitochondrial DNA reveals isolation of western Atlantic populations coupled with recent Pacific dispersal. *Molecular Ecology*, **15**, 3669–3679.

Klanten, O.S., Choat, J.H. & Van Herwerden, L. (2007) Extreme genetic diversity and temporal rather than spatial partitioning in a widely distributed coral reef fish. *Marine Biology*, **150**, 659–670.

Lacson, J.M. & Clark, S. (1995) Genetic divergence of Maldivian and Micronesian demes of the damselfishes *Stegastes nigricans*, *Chrysiptera biocellata*, *C. glauca* and *C. leucopoma* (Pomacentridae). *Marine Biology*, **121**, 585–590.

Lavery, S., Moritz, C. & Fielder, D.R. (1995) Changing patterns of population structure and gene flow at different spatial scales in *Birgus latro* (the coconut crab). *Heredity*, **74**, 531–541.

Lavery, S., Mortiz, C. & Fielder, D.R. (1996) Indo-Pacific population structure and evolutionary history of the coconut crab *Birgus latro*. *Molecular Ecology*, **5**, 557–570.

Lessios, H.A. (2008) The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annual Review of Evology, Evolution, and Systematics*, **39**, 63–91.

Lessios, H.A. & Robertson, D.R. (2006) Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2201–2208.

Lessios, H.A., Kessing, B.D. & Pearse, J.S. (2001) Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. *Evolution*, **55**, 955–975.

Lessios, H.A., Kane, J. & Robertson, D.R. (2003) Phylogeography of the pantropical sea urchin *Tripneustes*: contrasting patterns of population structure between oceans. *Evolution*, **57**, 2026–2036.

Lu, C.-P., Chen, C.A., Hui, C.-F., Tzeng, T.-D. & Yeh, S.-Y. (2006) Population genetic structure of the swordfish, *Xiphias gladius* (Linnaeus, 1758), in the Indian Ocean and West Pacific inferred from the complete DNA sequence of the mitochondrial control region. *Zoological Studies*, **45**, 269–279.

Maddison, D.R. & Maddison, W.P. (2002) *MacClade 4: analysis of phylogeny and character evolution*. Sinauer Associates, Sunderland, MA.

McMillan, W.O. & Palumbi, S.R. (1995) Concordant evolutionary patterns among Indo-West Pacific butterflyfishes. *Proceedings of the Royal Society B: Biological Sciences*, **260**, 229–236.

Menezes, M.R., Ikeda, M. & Taniguchi, N. (2006) Genetic variation in skipjack tuna *Katsuwonus pelamis* (L.) using PCR-RFLP analysis of the mitochondrial DNA D-loop region. *Journal of Fish Biology*, **68**, 156–161.

Meyer, A. (1994) Shortcomings of the cytochrome *b* gene as a molecular marker. *Trends in Ecology and Evolution*, **9**, 278–280.

Miller, T.L. & Cribb, T.H. (2007) Phylogenetic relationships of some common Indo-Pacific snappers (Perciformes: Lutjanidae) based on mitochondrial DNA sequences, with comments on the taxonomic position of the Caesioninae. *Molecular Phylogenetics and Evolution*, **44**, 450–460.

Möller, L.M. & Beheregaray, L.B. (2004) Genetic evidence for sex-biased dispersal in resident bottlenose dolphins (*Tursiops aduncus*). *Molecular Ecology*, **13**, 1607–1612.

Morales-Nin, B. & Ralston, S. (1990) Age and growth of *Lutjanus kasmira* (Forskål) in Hawaiian waters. *Journal of Fish Biology*, **36**, 191–203.

Naish, T., Powell, R., Levy, R. *et al.* (2009) Obliquity-paced Pliocene West Atlantic ice sheet oscillations. *Nature*, **458**, 322–328.

Nei, M. (1987) *Molecular evolutionary genetics*. Columbia University Press, New York.

Palumbi, S.R., Grabowsky, G., Duda, T., Geyer, L. & Tachino, N. (1997) Speciation and population genetic structure in tropical Pacific sea urchins. *Evolution*, **51**, 1506–1517.

Pardini, A.T., Jones, C.S., Noble, L.R., Kreiser, B., Hamish, M., Bruce, B.D., Stevens, J.D., Cliff, G., Scholl, M.C., Francis, M., Duffy, C.A.J. & Martin, A.P. (2001) Sex-biased dispersal of great white sharks. *Nature*, **412**, 139–140.

Planes, S. & Fauvelot, C. (2002) Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific Ocean. *Evolution*, **56**, 378–399.

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

Ramon, M.L., Nelson, P.A., De Martini, E., Walsh, W.J. & Bernardi, G. (2008) Phylogeography, historical demography, and the role of post-settlement ecology in two Hawaiian damselfish species. *Marine Biology*, **153**, 1207–1217.

Randall, J.E. (1987) Introductions of marine fishes to the Hawaiian Islands. *Bulletin of Marine Science*, **41**, 490–502.

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

Randall, J.E. (2001) Four new cardinalfishes (Perciformes: Apogonidae) from the Marquesas Islands. *Pacific Science*, **55**, 47–64.

Randall, J.E. (2007) *Reef and shore fishes of the Hawaiian Islands*. Sea Grant College Program, University of Hawaii, Honolulu.

Rangarajan, K. (1971) Maturity and spawning of the snapper, *Lutjanus kasmira* (Forskål) from the Andaman Sea. *Indian Journal of Fisheries*, **18**, 114–125.

Robertson, D.R., Grove, J.S. & McCosker, J.E. (2004) Tropical transpacific shorefishes. *Pacific Science*, **58**, 507–565.

Rocha, L.A. & Bowen, B.W. (2008) Speciation in coral-reef fishes. *Journal of Fish Biology*, **72**, 1101–1121.

Rocha, L.A., Bass, A.L., Robertson, D.R. & Bowen, B.W. (2002) Adult habitat preferences, larval dispersal, and comparative phylogeography of three Atlantic surgeonfishes (Teleostei: Acanthuridae). *Molecular Ecology*, **11**, 243–252.

Rocha, L.A., Robertson, D.R., Roman, J. & Bowen, B.W. (2005) Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 573–579.

Rocha, L.A., Craig, M.T. & Bowen, B.W. (2007) Phylogeography and the conservation of coral reef fishes. *Coral Reefs*, **26**, 501–512.

Schneider, S., Roessli, D. & Excoffier, L. (2000) *Arlequin: a software for population genetics data analysis ver 2.000*. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Geneva.

Schultz, J.K., Pyle, R.L., DeMartini, E. & Bowen, B.W. (2007) Genetic connectivity among color morphs and Pacific archipelagos for the flame angelfish, *Centropyge loriculus*. *Marine Biology*, **151**, 167–175.

Schultz, J.K., Feldheim, K.A., Gruber, S.H., Ashley, M.V., McGovern, T.M. & Bowen, B.W. (2008) Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). *Molecular Ecology*, **17**, 5336–5348.

Seutin, G., White, B.N. & Boag, P.T. (1991) Preservation of avian and blood tissue samples for DNA analyses. *Canadian Journal of Zoology*, **69**, 82–92.

Stepien, C.A., Randall, J.E. & Rosenblatt, R.H. (1994) Genetic and morphological divergence of a circumtropical complex of goatfishes: *Mulloidichthys vanicolensis*, *M. dentatus* and *M. martinicus*. *Pacific Science*, **48**, 44–56.

Suzuki, K. & Hioki, S. (1979) Spawning behavior, eggs, and larvae of the lutjanid fish, *Lutjanus kasmira*, in an aquarium. *Japanese Journal of Ichthyology*, **26**, 161–165.

Taberlet, P., Meyer, A. & Bouvet, J. (1992) Unusual mitochondrial DNA polymorphism in two local populations of blue tit *Parus caeruleus*. *Molecular Ecology*, **1**, 27–36.

Tamura, K. & Nei, M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Theisen, T.C., Bowen, B.W., Lanier, W. & Baldwin, J.D. (2008) High connectivity on a global scale in the pelagic wahoo, *Acanthacybrium solandri* (tuna family Scombridae). *Molecular Ecology*, **17**, 4233–4247.

Voris, H.K. (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**, 1153–1167.

Williams, S.T. & Benzie, J.A.H. (1998) Evidence of a biogeographic break between populations of a high dispersal starfish: congruent regions within the Indo-West Pacific defined by color morphs, mtDNA, and allozyme data. *Evolution*, **52**, 87–99.

Winans, G.A. (1980) Geographic variation in the milkfish *Chanos chanos* I. Biochemical evidence. *Evolution*, **34**, 558–574.

Woodland, D.J. (1983) Zoogeography of the Siganidae (Pisces): an interpretation of distribution and richness patterns. *Bulletin of Marine Science*, **33**, 713–717.

Wyrtki, K. & Kilonsky, B. (1983) Mean water and current structure during the Hawaii-to-Tahiti shuttle experiment. *Journal of Physical Oceanography*, **14**, 242–254.

Zapata, F.A. & Herron, P.A. (2002) Pelagic larval duration and geographic distribution of tropical eastern Pacific snappers (Pisces: Lutjanidae). *Marine Ecology Progress Series*, **230**, 295–300.

BIOSKETCH

Michelle R. Gaither is a PhD candidate at the University of Hawaii at Manoa. The data described here were presented at the 4th International Conference of the International Biogeography Society held in Merida, Mexico, 8–12 January 2009.

Editor: David Bellwood