Original Article

Title: Regal phylogeography: Range-wide survey of the marine angelfish *Pygoplites diacanthus* reveals evolutionary partitions between the Red Sea, Indian Ocean, and Pacific Ocean

Authors: Richard R. Coleman[a, b, *], Jeffrey A. Eble[c], Joseph D. DiBattista[d,e], Luiz A. Rocha[f], John E. Randall[g], Michael L. Berumen[d], Brian W. Bowen[a]

Postal Addresses:

[a]*Hawai'i Institute of Marine Biology, University of Hawai'i, Kāne'ohe, HI 96744, USA,* [b]*Department of Biology, University of Hawai'i, Mānoa, Honolulu, HI 96822, USA,* [c]*University of West Florida, Pensacola, FL 32514, USA,* [d]*Red Sea Research Center, Division of Biological and Environmental Science and Engineering, King Abdullah University of Science and Technology, Thuwal, 23955, Saudi Arabia,* [e]*Department of Environment and Agriculture, Curtin University, PO Box U1987, Perth, WA 6845, Australia,* [f]*Section of Ichthyology, California Academy of Sciences, San Francisco, CA 94118, USA,* [g]*Bernice Pauahi Bishop Museum, Honolulu, HI 96817, USA*

Corresponding author:

Richard R. Coleman

PO Box 1346

Kāne'ohe HI 96744

Phone +1 (916) 524-3734

Fax +1 (808) 236-7433

richard.colema@gmail.com

Abstract

The regal angelfish (*Pygoplites diacanthus*; family Pomacanthidae) occupies reefs from the Red Sea to the central Pacific, with an Indian Ocean/Rea Sea color morph distinct from a Pacific Ocean morph. To assess population differentiation and evaluate the possibility of cryptic evolutionary partitions in this monotypic genus, we surveyed mtDNA cytochrome *b* and two nuclear introns (S7 and RAG2) in 547 individuals from 15 locations. Phylogeographic analyses revealed four mtDNA lineages (d = 0.006 – 0.015) corresponding to the Pacific Ocean, the Red Sea, and two admixed lineages in the Indian Ocean, a pattern consistent with known biogeographical barriers. Christmas Island in the eastern Indian Ocean had both Indian and Pacific lineages. Both S7 and RAG2 showed strong population-level differentiation between the Red Sea, Indian Ocean, and Pacific Ocean (Φ_{ST} = 0.066 – 0.512). The only consistent population sub-structure within these three regions was at the Society Islands (French Polynesia), where surrounding oceanographic conditions may reinforce isolation. Coalescence analyses indicate the Pacific (1.7 Ma) as the oldest extant lineage followed by the Red Sea lineage (1.4 Ma). Results from a median-joining network suggest radiations of two lineages from the Red Sea that currently occupy the Indian Ocean (0.7 – 0.9 Ma). Persistence of a Red Sea lineage through Pleistocene glacial cycles suggests a long-term refuge in this region. The affiliation of Pacific and Red Sea populations, apparent in cytochrome *b* and S7 (but equivocal in RAG2) raises the hypothesis that the Indian Ocean was recolonized from the Red Sea, possibly more than once. Assessing the genetic architecture of this widespread monotypic genus reveals cryptic evolutionary diversity that merits subspecific recognition.

Keywords: biogeographic barriers, coral reef fish, cryptic diversity, genetic structure, monophyly, subspecies

1. Introduction

The majority of reef fishes have a pelagic larval phase typically lasting 20 to 60 days, followed by settlement at a location where they remain through juvenile and adult phases. It is during the pelagic larval phase that nearly all dispersal is accomplished, sometimes across great distances (Leis and McCormick, 2002; Hellberg, 2009). However, even closely related species with similar life histories can show markedly different genetic structure across their respective ranges (Rocha et al., 2002; DiBattista et al., 2012). Despite these differences in realized dispersal, genetic partitions frequently align with boundaries between biogeographic provinces, which mark abrupt changes in species composition accompanied by obvious geological or oceanographic barriers (Kulbicki et al., 2013; Bowen et al., 2016). However, phylogeographic reef surveys usually examine genetic partitions both within and between congeneric species (e.g. Robertson et al., 2006; Leray et al., 2010; DiBattista et al., 2013; Gaither et al., 2014; Ahti et al . 2016; Waldrop et al., 2016). Less attention has been paid to monotypic genera, and it is unknown whether these species have evolutionary or ecological traits that promote species cohesion across time.

The family Pomacanthidae (marine angelfishes) is comprised of more than 85 species across seven genera. All of the genera have at least eight species (*Centropyge* has more than 30) except for the monotypic genus *Pygoplites*. The regal angelfish, *Pygoplites diacanthus* (Boddaert 1772), has a wide distribution from East Africa and the Red Sea to the Tuamotu Archipelago in the central Pacific. This distribution encompasses four biogeographic provinces (Fig. 2 in Briggs and Bowen, 2013): the Indo-Polynesian Province (IPP), the Sino-Japanese Province, the Western Indian Ocean Province, and the Red Sea Province (which includes the Gulf of Aden; see Briggs and Bowen, 2012). Additionally, the range of *P. diacanthus* spans the Indo-Pacific Barrier, an

episodic land bridge separating Pacific and Indian Ocean fauna during low sea levels associated with glaciation (Randall, 1998; Rocha et al., 2007). *Pygoplites* diverged from the sister genus *Holacanthus* about 7.6 – 10.2 Ma (Alva-Campbell et al., 2010), and is monotypic despite occupying a very broad range and a variety of ecological conditions.

Randall (2005) noted coloration differences between an Indian Ocean morph, with a yellow chest, and a Pacific Ocean morph with a gray chest and less yellow coloring on the head (Fig. 1), invoking the possibility of nomenclatural recognition of the two morphotypes. Historically color has been used for species delineation in reef fishes, however, coloration alone can be a deceptive foundation for taxonomical classification; molecular tools have been useful for identifying cryptic genetic partitions and resolving taxonomic uncertainty over color morphs (McMillan et al., 1999; Schultz et al., 2006; Drew et al., 2008; Drew et al., 2010; DiBattista et al., 2012; Gaither et al., 2014; Ahti et al., 2016; Andrews et al., 2016)

Here we obtained samples from across the range of *P. diacanthus* to assess genetic connectivity with mitochondrial (mtDNA) and nuclear (nDNA) markers. Our sampling allowed us to test for cryptic evolutionary partitions and evaluate the hypothesis of taxonomic distinction between Indian and Pacific morphotypes. We were further motivated to resolve the ecological and evolutionary conditions that restrict diversification within the genus *Pygoplites*, the sole genus in an otherwise speciose family of fishes.

2. Material and Methods

2.1 Sample Collections

Between 2004 and 2014, 547 tissue samples (primarily fin clips) of *P. diacanthus* were collected from 15 locations across the species distribution (Fig. 1), using nets and pole-spears

while scuba diving or snorkeling. Tissues were preserved in salt-saturated DMSO buffer (Amos and Hoelzel, 1991) and stored at room temperature. Total genomic DNA was isolated from preserved tissue following the "HotSHOT" method of Meeker et al. (2007) and stored at -20°C. Due to variation in DNA amplification and sequence resolution, not all specimens were resolved at all three loci outlined below, hence sample sizes in Fig. 1 do not match samples sizes provided in the tables.

2.2 MtDNA Analyses

A 568-base pair (bp) fragment of the mtDNA cytochrome *b* (cyt *b*) gene was resolved to identify the maternal lineage of each individual using the forward primer (5'-GTGACTTGAAAAACCACCGTTG-3') (Song et al., 1998) and reverse primer (H15573; 5'-AATAGGAAGTATCATTCGGGTTTGAT-3') (Taberlet et al., 1992). PCR was performed in 10 µl reactions containing 10-15 ng of DNA, 5 µl of premixed BioMixRed^{TM} (Bioline, Inc., Springfield, NJ, USA), 0.2 µM primer for each primer, and nanopure water (Thermo Scientific Barnstead, Dubuque, IA, USA) to volume and using the following conditions: 4 min at 94°C, 35 cycles of denaturing for 30 s at 94°C, annealing for 30 s at 50°C, extension for 45 s at 72°C, and a final extension for 10 min at 72°C.

PCR products were visualized using a 1.5% agarose gel with GelStar^{TM} (Cambrex Bio Science Rockland, Rockland MA, USA) and then purified by incubating with 0.75 units of Exonuclease and 0.5 units of Shrimp Alkaline Phosphatase (ExoSAP; USB, Cleveland, OH, USA) per 7.5 µl of PCR product for 30 min at 37°C, followed by 15 min at 85°C. DNA sequencing was performed using fluorescently-labeled dideoxy terminators on an ABI 3730XL Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) at the University of Hawai'i

Advanced Studies of Genomics, Proteomics and Bioinformatics sequencing facility.

Sequences were aligned and edited using GENEIOUS v.8.0.3 (Gene Codes, Ann Arbor, MI, USA) and unique sequences were deposited into GenBank (Accession numbers: RAG2, KU885737 - KU885756; S7, KU885757 - KU885843; cyt *b*, KU885844 - KU885892). A model for DNA sequence evolution was selected using the program JMODELTEST v.2.1 (Guindon and Gascuel, 2003; Darriba et al., 2012). The best-fit model of TIM1+G (gamma=0.0760) was identified by the Akaike Information Criterion (*AIC*) and the closest match used for subsequent analyses. Mean genetic distance between lineages was calculated in DNASP v.5.10 (Librado and Rozas, 2009). A haplotype network was constructed for each locus with NETWORK v.4.6.1.1 (http://www.fluxus-engineering.com/network_terms.htm) using a median-joining algorithm (Bandelt et al., 1999) and default settings.

To estimate the time to most recent common ancestor (TMRCA), we formatted the data with BEAUTI v.1.4.7 and used a Bayesian MCMC approach in BEAST v.2.2.0 (Drummond & Rambaut, 2007). We conducted our analysis with a strict clock of 2% per million years between lineages (Bowen et al., 2001; Reece et al., 2010a) and used a coalescent tree prior assuming exponential growth. We used default priors under the HKY+G+I model of mutation, the closest available model, and ran simulations for 10 million generations with sampling every 1000 generations. Ten independent runs were computed to ensure convergence, and log files were combined and ages averaged across runs using TRACER v.1.6 (http://tree.bio.ed.ac.uk/software/tracer/).

ARLEQUIN v.3.11 was used to generate haplotype and nucleotide diversity, as well as to test for population structure (Excoffier et al., 2005). Genetic structure among and between regions was estimated by performing an analysis of molecular variance (AMOVA). Deviations

144 from null distributions were tested with non-parametric permutation procedures (N = 9999).
145 Pairwise Φ_{ST} statistics, an analog of Wright's F_{ST} that incorporates sequence evolution and
146 divergence, were generated to assess structure and identify phylogeographic partitions. Locations
147 where samples sizes were < 8 were excluded from population genetic analyses but included in
148 overall diversity estimates. False discovery rates were controlled for and maintained at $\alpha = 0.05$
149 among all pairwise tests (Benjamini and Yekutieli, 2001; Narum, 2006).
150 Time since the most recent population expansion was estimated for each location using
151 the equation $\tau = 2\mu t$, where t is the age of the population in generations and μ is the mutation rate
152 per generation for the entire sequence (μ = number of bp x divergence rate within a lineage x
153 generation time in years). We used a sequence divergence estimate within lineages of 1-2% per
154 million years (Bowen et al., 2001; Reece et al., 2010a) to estimate population age. While
155 generation time is unknown for *P. diacanthus*, we conditionally used the equation $T = (\alpha + \omega)/2$,
156 where α is the age at first reproduction and ω is the age of last reproduction (or lifespan; Pianka,
157 1978). We obtained a generation time of 8.5 years based on an estimated reproductive age of 2
158 years and longevity of more than 15 years (Hinton, 1962). Due to the tentative nature of
159 generation time and mutation rates estimates, population age should be interpreted with caution,
160 however rank-order comparisons among populations are robust to such approximations. Fu's F_S
161 (Fu, 1997) was calculated to test for evidence of selection or (more likely) population expansion
162 using 10,000 permutations. A significant negative value of Fu's F_S is evidence for an excess
163 number of alleles, as would be expected from a recent population expansion, whereas, a
164 significant positive value is evidence for a deficiency of alleles, as would be expected from a
165 recent population bottleneck.
166

2.3 Nuclear DNA Analysis

We sequenced two nuclear loci: the recombination-activating gene 2 (RAG2) and intron 1 of the S7 ribosomal protein (S7). We resolved 431-bp of RAG2 using modified primers from Lovejoy (1999); the forward primer is 5'-SACCTTGTGCTGCAAAGAGA-3' and reverse primer is 5'-AGTGGATCCCCTTBTCATCCAGA-3'. We resolved 510-bp of S7 using primers S7RPEX1F and S7RPEX2R from Chow and Hazama (1998). For each intron, PCR was performed using the same reaction as described for cyt *b* but using the following temperature conditions: 5 min at 94°C, 35 cycles of denaturing for 30 s at 94°C, annealing for 30 s at 58°C, extension for 45 s at 72°C, and a final extension for 10 min at 72°C.

Allelic states with more than one heterozygous site were estimated using PHASE v.2.1 (Stephens and Donnelly, 2003) as implemented in DNASP. Unique sequences were deposited in GenBank (Accession numbers: XXX - XXX). Three separate runs, each of 100,000 repetitions after a 10,000 iteration burn-in, were conducted for each locus; all runs returned consistent allele identities. Median-joining networks were created for each nuclear dataset as outlined above. To minimize circularity between closely related alleles, singletons were removed from the S7 network. However, this did not alter our overall interpretation of the results. Pairwise Φ_{ST} statistics were calculated for each nuclear dataset. The best-fit model of K80 and TPM1uf+I (proportion of invariable sites = 0.89) were identified for RAG2 and S7, respectively as determined by JMODELTEST. Observed heterozygosity (H_O) and expected heterozygosity (H_E) were calculated for each locus and an exact test of Hardy-Weinberg Equilibrium (HWE) using 100,000 steps in a Markov chain was performed in ARLEQUIN.

2.4. Phylogenetic reconstruction

Phylogenetic reconstruction based on cyt *b* was rooted with *Holacanthus africanus* (family Pomacanthidae; GenBank accession number KC845351 and KC845352), as this genus is sister to *Pygoplites* (Bellwood et al., 2004; Alva-Campbell et al., 2010). Bayesian inference was conducted using MRBAYES v.3.1.2 (Huelsenbeck et al., 2001; Ronquist, 2004) running a pair of independent searches for 1 million generations, with trees saved every 1000 generations and the first 250 sampled trees of each search discarded as burn-in. Due to high divergence between *P. diacanthus* and *H. africanus* (14.7% at cyt *b*) we were unable to resolve phylogenetic relationships within the genus *Pygoplites* using an outgroup, therefore an unrooted tree was also constructed with MRBAYES based on the concatenated dataset of all loci. A maximum likelihood tree was created using PHYML v.3.0.1 (Guindon et al., 2010) as implemented in GENEIOUS with clade support assessed with 1000 non-parametric bootstrap replicates. A neighbor-joining tree was created using GENEIOUS with clade support assessed after 1000 non-parametric bootstrap replicates.

3. Results

3.1. Phylogenetic and coalescence analyses

All tree-building methods yielded identical topologies. The unrooted phylogenetic analysis recovered four lineages: a Pacific lineage that extends to Christmas Island in the eastern Indian Ocean (henceforth referred to as "Pacific lineage"), a lineage detected around Saudi Arabia and Djibouti (henceforth referred to as "Red Sea lineage"), and two lineages with overlapping ranges in the Maldives and Diego Garcia (henceforth referred to as "Indian lineage 1" and "Indian lineage 2") (Fig. 2). The phylogenetic analyses were unable to resolve branch order among these lineages using an outgroup (Fig 2a), in part because the sister genus

(*Holacanthus*) is deeply divergent at cyt *b* (Alva-Campbell et al., 2010). The Pacific lineage is 0.6% divergent from the Red Sea lineage and 1.2% and 1.5% from Indian lineage 1 and 2, respectively. The Red Sea lineage is 0.6% divergent from Indian lineage 1 and 1.0% from Indian lineage 2, and the two Indian lineages are distinguished by 1.5% divergence. Coalescence analysis based on cyt *b* yielded a TMRCA of 1.7 Ma for the Pacific lineage, and identified the Pacific as the oldest extant lineage (Table 1). The Red Sea lineage dates to 1.4 Ma and the two Indian lineages were the youngest: Indian lineage 1 at 0.7 Ma; Indian lineage 2 at 0.9 Ma.

3.2 MtDNA Sequences

Mitochondrial DNA molecular diversity indices are summarized for lineages in Table 1 and among populations in Table 2. Total haplotype diversity was $h = 0.817$ with 49 unique haplotypes. Among lineages, the Red Sea had the highest haplotype diversity ($h = 0.701$) with the lowest being observed in Indian Ocean lineage 2 ($h = 0.284$). Within populations, Indonesia had the highest haplotype diversity ($h = 1.00$) followed by Okinawa ($h = 0.867$) and the Maldives ($h = 0.808$). The lowest haplotype diversity was observed at Fiji ($h = 0.427$) and Mo'orea ($h = 0.483$). Total nucleotide diversity was $\pi = 0.005$. Among lineages, the Pacific Ocean and Red Sea had the higher nucleotide diversity ($\pi = 0.002$) with the lowest nucleotide diversity observed in both Indian Ocean lineages ($\pi = 0.001$). Among populations, the Maldives and Diego Garcia had the highest nucleotide diversity for all locations, each at $\pi = 0.009$, whereas the lowest nucleotide diversity is observed at Fiji ($\pi = 0.0008$).

The median-joining haplotype network illustrates the low level of divergence between the four evolutionary lineages recovered from the phylogenetic analysis (Fig. 3a). However, the network also reveals that Red Sea haplotypes lie between the Pacific and Indian haplotypes. The

presence of two Indian lineages radiating from the most common Red Sea haplotype provides evidence for two independent colonization events. The two Indonesia specimens are associated with Indian Ocean lineage 1; however, low samples size precludes any interpretation about lineage distribution. Christmas Island, located at the edge of the IPP, a region where Pacific and Indian Ocean fauna come into contact (Gaither and Rocha, 2013), had both Pacific and Indian lineages. In subsequent comparisons between ocean basins, Christmas Island specimens grouped with Indian and Pacific cohorts based on mtDNA identity.

Population pairwise Φ_{ST} values for cyt *b* results are summarized in Table 3. Significance was determined after controlling for false discovery rates (corrected $\alpha = 0.009$). Φ_{ST} values show congruence with the haplotype network further supporting the Pacific, Indian, and Red Sea groups. There was little or no population structure detected within these groups, with two exceptions: Mo'orea (French Polynesia) shows significant genetic differentiation from all Pacific locations, with pairwise Φ_{ST} values ranging from 0.123 with Pohnpei to 0.229 with Fiji. Elsewhere in the Pacific, significant genetic structure was detected between the Marshall Islands and Fiji ($\Phi_{ST} = 0.061$, $P < 0.001$). Although population level data is not reported for the single location in the Sino-Japanese Province (Okinawa) due to low sample size ($N = 6$), preliminary runs show no significant population structure between the Sino-Japanese Province and the Pacific samples of *P. diacanthus*. The Red Sea lineage shows high levels of population differentiation from all other samples (pairwise Φ_{ST}: 0.284 – 0.837). Likewise, the Indian lineage shows significant population differentiation from all other samples (pairwise Φ_{ST}: 0.284 – 0.753). The AMOVA analysis supports the Pacific, Indian, and Red Sea geographic groupings based on mtDNA (Table 4) with the majority of the variation ($\Phi_{CT} = 0.66$, $P < 0.001$) existing among the groups.

The demographic results for cyt *b* show indications of population expansion at every Pacific location with the exception of Okinawa, the Marshall Islands, and Mo‘orea (Table 2). Estimates of population expansion indicate that the youngest dates are in the Pacific: Fiji and Christmas Island, with estimates of 39,000 and 49,000 years, respectively. The oldest Pacific expansion dates are in Okinawa, Pohnpei, and American Samoa, at 271,000, 230,000, and 212,000 years, respectively. Within the Red Sea Province, Saudi Arabia shows evidence for a population expansion (Fu's F_S: - 4.73, $P < 0.01$) at 65,000 – 130,000 years, whereas Djibouti shows evidence for a neutral population (Fu's F_S: $P = 0.35$) coalescing at 105,000 – 209,000 years. Locations in the Indian Ocean singularly show no evidence of population expansion (Fu's F_S, $P > 0.02$) and have the oldest population expansions dates at 807,000 – 1,742,000 years. However, these estimates are shaped by the presence of two lineages that are not monophyletic (Fig. 3a). When considered individually, Indian lineage 1 has a population expansion date at 48,000 – 97,000 years (Fu's F_S: -3.70, $P < 0.001$), and Indian lineage 2 has a population expansion date at 264,000 – 528,000 years (Fu's F_S: -2.75, $P = 0.01$).

3.3 Nuclear DNA Sequences

A total of 10 variable sites yielded 12 alleles at the RAG2 locus and 31 variable sites yielded 46 alleles at the S7 locus. Samples from Palau and Tokelau were out of Hardy-Weinberg equilibrium (Palau, $P < 0.001$; Tokelau, $P = 0.04$) with excess homozygotes at the S7 locus (Table 5). Overall expected heterozygosity (H_E) was 0.43 and 0.86 for RAG2 and S7, respectively. Across all samples $H_E = 0.06 – 0.64$ for RAG2 and $H_E = 0.41 – 1.00$ for the S7 intron. The median-joining networks based on intron sequences do not show distinct lineages in the Red Sea, Indian Ocean, and Pacific Ocean (Fig. 3b, c). However, both RAG2 and S7

networks include common alleles that are observed only in the Pacific, or only in the Indian Ocean locations. For S7, an Indian Ocean specific allele is also shared with a single individual from Christmas Island.

The population genetic results for the nuclear dataset are strongly concordant with mtDNA analyses for *P. diacanthus*, although they differ by degree. Genetic structure was absent within the Red Sea and within the Indian Ocean. The only significant differentiation in the Pacific was in 7 of 8 comparisons to Mo'orea (Society Islands, French Polynesia) with RAG2 (Φ_{ST} = 0.111 – 0.271; Table 6). Curiously, none of the same pairwise comparisons for Mo'orea were significant with S7, however Mo'orea showed the highest differentiation from Red Sea populations.

Both nuclear markers show high levels of genetic structure that correspond to a Pacific, Indian, and Red Sea lineage. RAG2 was significant in 17 of 18 Pacific versus Indian comparisons (Φ_{ST} = 0.137 – 0.343), significant in all Indian versus Red Sea comparisons (Φ_{ST} = 0.091 – 0.258), and significant in 15 of 18 Pacific versus Red Sea comparisons (Φ_{ST} = 0.066 – 0.359). The S7 differences were significant in all Pacific versus Indian comparisons (Φ_{ST} = 0.073 – 0.188), all Indian versus Red Sea comparisons (Φ_{ST} = 0.253 – 0.512), and all Pacific versus Red Sea comparisons (Φ_{ST} = 0.159 – 0.443). The exceptions to these patterns were comparisons between the Red Sea lineage and Tokelau, as well as between Saudi Arabia and Pohnpei. For S7 the highest genetic structure was observed between the Indian and Red Sea populations. This contrasts with the RAG2 and cyt *b* comparisons, where the highest genetic structure differentiated the Pacific from both Indian and Red Sea regions.

4. Discussion

4.1 Summary of results

Our data demonstrates that cryptic diversity exists within the monotypic genus *Pygoplites* as evidenced by significant levels of genetic structure among three regions: the Pacific Ocean (which includes a cohort at Christmas Island), the Indian Ocean (with two sympatric mtDNA lineages), and the Red Sea (Table 4). This pattern of genetic structure corresponds to known biogeographic provinces and phylogeographic barriers observed in other reef fishes (Rocha et al., 2007; Briggs and Bowen, 2013; DiBattista et al., 2013; Eble et al., 2015; Gaither et al., 2015). The Red Sea biogeographic province is distinguished by a faunal break at the Gulf of Aden, and the Indo-Pacific Barrier is an intermittent terrestrial bridge between Australia and SE Asia that impedes water movement between Pacific and Indian Oceans during glacial low-sea levels (see Gaither & Rocha, 2013). The Sino-Japanese Province shows no genetic differentiation from the Pacific population (based on $N = 6$), but Mo'orea is highly isolated, a finding we attribute to prevailing oceanographic conditions (see Gaither et al., 2010). Below we discuss the phylogenetic implications of cryptic lineages and examine each of these regions in light of biogeographic theory

4.2 Phylogenetic considerations

Differences in coloration reviewed by Randall (2005) suggested that cryptic linages of *P. diacanthus* might exist in the Pacific and Indian Oceans. The three loci evaluated here support this Indian-Pacific distinction with diagnostic (albeit shallow) mtDNA differences and strong population genetic separations at two nuclear loci. A rooted phylogeny was unable to resolve relationships within the genus *Pygoplites* due to shallow separations and the deep divergence from the outgroup, *H. africanus*, ($d = 15.5\%$ at cyt *b*, this study), despite being the most closely

related species to *P. diacanthus* (Alva-Campbell et al., 2010). Therefore we were unable to determine the basal lineage from among the four lineages recovered. The oldest TMRCA in *P. diacanthus* is the Pacific lineage at 1.7 Ma, but the divergence between *Pygoplites* and *Holocanthus* is much older, estimated at 7.6 – 10.2 Ma (Alva-Campbell et al., 2010). Hence much of the evolutionary? history of *Pygoplites* has been erased, at least for the loci examined here.

There are two possible explanations for the lack of diversity within the genus. First, there has been no evolutionary or selective pressure for *P. diacanthus* to diversify, a feature that may be attributed to the species ability to occupy a variety of ecological niches. *P. diacanthus* is considered a generalist in that its range occupies more than half the globe in subtropical and tropical environments, its diet consists of sessile invertebrates, such as sponges and tunicates, and it appears to be a reef-habitat generalist where its range extends from the surface to depths greater than 60 m, a zone where shallow coral reef habitat is replaced by mesophotic ecosystems (Puglise et al., 2009). An alternative explanation is that other species within the genus went extinct while *P. diacanthus* persisted. However, with a poor fossil record, the evolutionary history of the marine angelfishes is poorly understood and limited to extant species. Therefore we know of no species that may have existed during the 10 million year separation between *Holacanthus* and *Pygoplites*. Nonetheless, the phylogeographic record for *Pygoplites* begins with a radiation in the last 2 MY. Although phylogenetic reconstruction was unable to determine branch order among the four lineages, the median-joining network indicates that the Red Sea lineage is basal to the two mtDNA lineages in the Indian Ocean. Coloration differences distinguish the Pacific lineage from both Indian and Red Sea lineages (Fig. 1); however, a

preliminary morphological examination by L.A.R. revealed no additional morphological characters that discriminate between Indian and Red Sea lineages.

The geographical delineation between the Pacific and Indian lineages correspond with the exposure of the Sunda Shelf, which separates the Pacific and Indian Oceans during low sea level. The Red Sea lineage corresponds to the Red Sea biogeographic province, which encompasses the adjacent Gulf of Aden (Briggs and Bowen, 2012) and whose populations have a disjunct distribution with the remainder of the range (see below). During glacial maxima the Red Sea is effectively cut off from the Indian Ocean by closure of the Strait of Bab al Mandab, the only natural gateway into the Red Sea, allowing sufficient time for populations to diverge into distinct evolutionary lineages (DiBattista et al., 2013, 2016a).

The mechanisms facilitating two sympatric mtDNA lineages in the Indian Ocean are less clear. Coalescence estimates indicate that the lineages arose independently during roughly the same period (0.72 – 0.93 Ma). As there are no known phenotypic differences within this region, the unexpected recovery of two distinct lineages requires further investigation. Indian Ocean samples contained similar numbers of each lineage (Maldives: Lineage 1, $N = 8$; Lineage 2, $N = 8$; Diego Garcia: Lineage 1, $N = 17$; Lineage 2, $N = 11$) indicating that the two lineages are approximately equally represented.

The recovery of multiple evolutionary partitions within the monotypic genus *Pygoplites* may not be indicative of other monotypic genera. Cryptic evolutionary partitions are routinely discovered within species of marine fishes (Colborn et al., 2001; Rocha et al., 2008; DiBattista et al., 2012; Fernandez Silva et al., 2015), and in this regard *P. diacanthus* is similar to the more speciose inhabitants of Indo-Pacific reefs. The factors that produce a deep, monotypic lineage are therefore not reflected in an unusual phylogeographic architecture. However, part of the

explanation for this monotype may be that five to eight million years after the divergence of *Pygoplites* and *Holocanthus*, the ancestor of all modern *Pygoplites* likely radiated out of the West Pacific Ocean, an extensive source of Indo-Pacific diversity (Cowman and Bellwood, 2013).

In considering the phylogenetic results through a taxonomic lens, there are several issues. First, the Pacific and Indian morphs are distinguished by diagnostic differences, but they are not monophyletic. The Indian Ocean contains two mtDNA lineages, each more closely related to the Red Sea lineage than to each other. Second, the coloration difference between Pacific and Indian forms, now matched by $d = 0.006$ divergence, could be a platform to describe them as separate species. Third, the genetic divergence observed at all three loci is low in comparison to typical divergences for fish species ($d = 0.03 – 0.12$; Grant and Bowen, 1998; Johns and Avise, 1998). Fourth, the two morphs form mixed groups where they co-occur at Christmas Island (Hobbs and Allen, 2014). Since we lack diagnostic nDNA alleles for the two morphs, we do not have the power to test for hybrids between the lineages, but this is certainly a possibility. Given these considerations, we believe that it is problematic to invoke species status for these three regional forms and we endorse subspecies recognition distinguishing the Pacific lineage from the Indian and Red Sea lineages based on shallow but diagnostic distinctions in genetics and morphology. We propose the name *P. diacanthus flavescens* for the Indian Ocean and Red Sea lineages to give recognition to the yellow chest coloration, a character not found in individuals from the Pacific lineage (*P. d. diacanthus*).

4.3 Red Sea isolation and refugia

The Red Sea Province is distinguished from the Indian Ocean by high levels of endemism found across a suite of taxa (Randall, 1994; Cox and Moore, 2000) as well as many fish species

whose ranges extend from the Red Sea into the Gulf of Aden (Briggs and Bowen, 2012; DiBattista et al., 2016b). This distinction is supported by our findings that show the Djibouti population of *P. diacanthus* forms a genetically homogenous population with the Red Sea, coupled with a population break separating these two locations from adjacent populations in the Indian Ocean.

Population breaks between the Red Sea and Indian Ocean have previously been documented in *P. diacanthus*, in addition to other species (Vogler et al., 2008; DiBattista et al., 2013; Fernandez Silva et al., 2015). One possible explanation for breaks across multiple species in this region is the presence of an ecological barrier. Based on differences in fish assemblages, Kemp (1998) proposed that such a barrier separated the Red Sea and western Gulf of Aden from the eastern Gulf of Aden and Indian Ocean. Furthermore, the upwelling that occurs along the Arabian coast of southern Yemen, Oman, and the Indian Ocean coast of Somalia impedes the formation of continuous reefs from Djibouti to Oman and southern Somalia, limiting opportunities for dispersal from the Gulf of Aden (for review see DiBattista et al., 2016a). Notably, we did not detect *P. diacanthus* during collection efforts in the Socotra Archipelago, Oman, and Somalia, which are located at the periphery of the Gulf of Aden and the Arabian Sea. This observation coincides with previous surveys conducted in the region indicating a gap in the distribution of *P. diacanthus* between the Gulf of Aden and the western Indian Ocean, a phenomenon found in other wide-ranging species (Kemp, 1998).

The parsimonious conclusion that a population of *P. diacanthus* has been in the Red Sea Province (including western Gulf of Aden) for over a million years implies that this population has been subjected to and survived Pleistocene glacial conditions. The only natural connection to the Indian Ocean is through the narrow (18 km) and shallow (137 m) Strait of Bab al Mandab at

419 the southern end of the Red Sea. During periods of low sea level associated with glaciation, the
420 connection from the Indian Ocean through the strait is reduced, and the Red Sea experiences
421 extreme fluctuations in temperature and salinity (Bailey, 2009). During the last 400,000 years in
422 particular, the Red Sea has undergone at least two periods of hypersalinity (c. 19,000 and 30,000
423 years ago) that caused an aplanktonic environment in which larvae of many marine organisms
424 presumably could not survive (Siddal et al., 2003; DiBattista et al., 2016a). Coalescence analysis
425 dates the Red Sea lineage to 1.44 Ma (95% HPD = 0.51 – 2.53 Ma), which coupled with the
426 Saudi Arabian population expansion (65,000 – 130,000 years) indicates that *P. diacanthus* likely
427 survived the temperature and salinity crises that occurred during these periods, a conclusion that
428 is corroborated by other species (DiBattista et al., 2013). Our neutrality tests show no evidence
429 for changes in population size (Fu's *Fs* = -3.60, *P* = 0.035) providing evidence that refugia may
430 have existed in the Red Sea Province (possibly in the Gulf of Aden) to support a large stable
431 population of *P. diacanthus* despite the extreme environmental conditions.
432
433 *4.4 Biogeographic inferences in the Indian Ocean*
434 Christmas Island is located in the eastern Indian Ocean, a region (which includes Cocos-
435 Keeling Island) of secondary contact between Indian and Pacific species that diverged in
436 allopatry during Pleistocene glacial cycles (Gaither and Rocha, 2013). Indian and Pacific Ocean
437 phenotypes of *P. diacanthus* have both been recorded in the eastern Indian Ocean region, and
438 both Pacific and Indian Ocean mtDNA haplotypes are present at Christmas Island, indicating an
439 area of overlap (Hobbs and Allen, 2014 , Fig 3a). This region is recognized as a hybridization
440 hotspot (suture zone) with interbreeding documented between at least 27 reef fish species-pairs
441 from across eight families, and it has been suggested that Indian and Pacific *P. diacanthus*

442 lineages hybridize in this region (Hobbs and Allen, 2014). However, additional molecular work
443 will be needed to evaluate this hypothesis.

444 Genetic differences between Indian and Pacific Ocean populations are consistent with
445 Pleistocene closures of the Indo-Pacific Barrier. Despite being located in the Indian Ocean basin
446 and the presence of haplotypes that are associated with Indian Ocean lineages, our results
447 indicate that Christmas Island is genetically differentiated from other locations in the Indian
448 Ocean and instead has a stronger affiliation with the Pacific Ocean. A barrier to dispersal has
449 been previously proposed to exist west of the Cocos-Keeling Islands and east of the Chagos-
450 Laccadive ridge based on the presence of many Pacific species with distributions that extend no
451 further west than Christmas and Cocos-Keeling Islands (Blum, 1989; Hodge and Bellwood,
452 2016).

453 Elsewhere in the Indian Ocean, the Maldives and Diego Garcia (Chagos Archipelago) are
454 genetically differentiated from the Pacific and Red Sea, but not from each other. Both
455 archipelagos are located in the central Indian Ocean, which is the western extent of the IPP,
456 although they also share faunal affinities with the Western Indian Ocean Province (Winterbottom
457 and Anderson, 1997; Eble et al., 2011; Gaither et al., 2011; Briggs and Bowen, 2012). The
458 grouping of Diego Garcia and the Maldives within the IPP is further evidenced by Pacific Ocean
459 mtDNA being found at Diego Garcia (Fig. 3a), which provides a signal that some degree of gene
460 flow occurs between the Indian and Pacific Ocean. Coalescence estimates of the two Indian
461 Ocean *P. diacanthus* lineages indicate they arose from an ancestor affiliated with the Red Sea.

462 The ability of *P. diacanthus* to persist throughout major geological and climatic shifts is
463 demonstrated by the age of expansion for all populations of *P. diacanthus* which predate the Last
464 Glacial Maximum, peaking at 26.5 – 19 ka (Clark et al., 2009), when global sea levels dropped

130 m below present levels (Voris, 2000). During this period, habitable shelf in the Pacific was reduced by as much as 92% from present day values and this reduction in habitat area has been linked to population bottlenecks (Ludt and Rocha, 2014), a feature not observed in *P. diacanthus*. As previously discussed, *P. diacanthus* can be considered an ecological generalist with a vertical range that extends to mesophotic depths. Thus, a reduction of shallow reef habitat due to sea level change may not have substantially reduced suitable ecological niches for this species.

4.5 Gene flow within the Pacific

Despite the wide expanse of the central and western Pacific Ocean, many species exhibit a high degree of genetic connectivity across the region (Schultz et al., 2006; Reece et al., 2010b; Gaither et al., 2011). However, populations breaks have been associated with isolated regions such as the Hawaiian Archipelago and the Marquesas, which are also known for high levels of endemism (Randall, 2005; Briggs and Bowen, 2012). Here we found population genetic differentiation of Mo'orea (Table 3, Table 6), a pattern observed in other widely distributed Pacific species (Planes, 1993; Bernardi et al., 2001; DiBattista et al., 2012; Timmers et al., 2012; Lemer and Planes, 2014).

The isolation of Mo'orea may be attributed to ocean circulation patterns. The westward flow of the Southern Equatorial Current (SEC) and eddies created in the wake of Tahiti, located approximately 17 km east of Mo'orea, contribute to a strong counterclockwise flow around the island promoting the local retention of larvae (Leichter et al., 2013). Plankton tows conducted in this region revealed that fish larvae were not recovered more than 300 km from the nearest reef (Lo-Yat et al., 2006). Additionally, Bernardi et al. (2012) found that 14% of juvenile damselfish

488 (*Dascyllus trimaculatus*) recruiting to reefs around Mo'orea were very close relatives, including
489 full siblings, indicating that the larvae traveled and settled together despite a PLD of several
490 weeks.

491 Although the counterclockwise flow surrounding Mo'orea may explain local retention of
492 larvae, it does not explain how larvae produced elsewhere in the Pacific are restricted from
493 emigrating and settling onto Moorean reefs. One possible explanation may be that the westward
494 flowing SEC restricts larvae from dispersing in an easterly direction. The SEC, located between
495 4°N and 17°S (Wyrtki and Kilonsky, 1984; Bonjean and Lagerloef, 2002), has been implicated in
496 limiting connectivity between the Marquesas, located 1300 km northeast of Mo'orea, and other
497 Pacific locales (Gaither et al., 2010; Szabo et al., 2014). Populations of *P. diacanthus* west of
498 Mo'orea, located at the southern extent of the SEC, may be restricted in easterly dispersal by the
499 strong current; however, the SEC may facilitate a western dispersal. American Samoa is the
500 closest sample location downstream from Mo'orea; it is the only sample location that is not
501 significantly differentiated from Mo'orea at RAG2 (Φ_{ST} = 0.039, P = 0.054) and has one of the
502 lowest levels of differentiation from Mo'orea at cyt *b* (Φ_{ST} = 0.128). Fine-scale sampling across
503 French Polynesia would be required to determine the extent of genetic isolation. Additionally,
504 further sampling from neighboring localities east and west of Mo'orea are needed to test our
505 hypothesis regarding the SEC. It is likely that a number of physical processes surrounding
506 Mo'orea promote local retention of larvae and prevent the recruitment of larvae from elsewhere
507 in the Pacific.
508
509 **5. Conclusion**

510 *Pygoplites diacanthus* is the first large angelfish to be surveyed across the Indo-Pacific. It
511 appears to be highly dispersive, joining the ranks of smaller Pomacanthids such as the pygmy
512 angelfish in showing little structure across ocean basins (Schultz et al., 2006; DiBattista et al.,
513 2012). Pelagic larval duration tends to be shorter in the large angelfishes (~25 days in *Pygoplites*
514 compared to 30 days or more in pygmy angelfishes; Thresher and Brothers, 1985), but this does
515 not seem to restrict dispersal among the closely associated islands of the West and Central
516 Pacific. However, this monotypic genus exhibits deep population genetic partitions between
517 ocean basins. In every case, historical barriers existed at the junctions between observed
518 populations, and in at least two cases (Red Sea and Mo'orea) oceanographic conditions may
519 contribute to contemporary isolation. On the genetic continuum between isolated populations and
520 evolutionary distinctions (Wright, 1978; Frankham et al., 2002), the deep divergences between
521 oceans indicate that the monotypic *Pygoplites* may be on the pathway to three emerging species.
522 The genetic and morphological divergences are certainly sufficient to recognize subspecific
523 evolutionary (and taxonomic) partitions.
524
525 **Acknowledgements**
526 This project was supported by the Seaver Institute (to BWB), the National Science Foundation
527 (NSF) OCE-0929031 (BWB), NOAA National Marine Sanctuaries Program MOA grant No.
528 2005-008/66882 (R.J. Toonen), University of Hawai'i Sea Grant No. NA05O AR4171048
529 (BWB), King Abdullah University of Science and Technology (KAUST) Office of Competitive
530 Research Funds under Award no. CRG-1-2012-BER-002 and baseline research funds to MLB, as
531 well as National Geographic Society Grant 9024-11 to JDD. RRC was supported by NSF grant
532 DGE-1329626 and the Dr. Nancy Foster Scholarship program under Award no.

533 NA15NOS4290067. Fieldwork at Christmas Island was supported by National Geographic Grant
534 8208-07 (M.T. Craig). We thank Eric Mason and the crew at Dream Divers in Saudi Arabia,
535 Nicolas Prévot at Dolphin Divers and the crew of the M/V Deli in Djibouti, the KAUST Coastal
536 and Marine Resources Core Lab, the KAUST Reef Ecology Lab, Amr Gusti, the Administration
537 of the British Indian Ocean Territory and Chagos Conservation Trust, as well as the University
538 of Milano-Bicocca Marine Research and High Education Centre in Magoodhoo, the Ministry of
539 Fisheries and Agriculture, Republic of Maldives, and the community of Maghoodhoo, Faafu
540 Atoll. For assistance in collection efforts we thank Alfonso Alexander, Senifa Annandale,
541 Howard Choat, Pat Colin, Lori Colin, Joshua Copus, Matthew Craig, Michelle Gaither, Brian
542 Greene, Jean-Paul Hobbs, Garrett Johnson, Stephen Karl, Randall Kosaki, Cassie Lyons, David
543 Pence, Mark Priest, Joshua Reece, D. Ross Robertson, Tane Sinclair-Taylor, Robert Thorne, and
544 Robert Whitton. We thank the HIMB EPSCoR core facility and the University of Hawai‘i’s
545 Advanced Studies in Genomics, Proteomics, and Bioinformatics facility for their assistance with
546 DNA sequencing. We also thank Michelle Gaither, Brent Snelgrove, Robert Toonen, and
547 members of the ToBo lab for their assistance, logistical support, and feedback throughout this
548 project. Thanks to Tane Sinclair-Taylor for the graphical abstract images. Thanks to editor
549 Giacomo Bernardi and one anonymous reviewer for critique and suggestions that improved the
550 paper. This is contribution #1653 from the Hawai‘i Institute of Marine Biology, #9592 from the
551 School of Ocean and Earth Science and Technology, and #XX from University of Hawai‘i Sea
552 Grant Program.

553 **References**

554 Ahti, P., Coleman, R.R., DiBattista, J., Berumen, M., Rocha, L.A., Bowen, B.W., 2016.
555 Phylogeography of Indo-Pacific reef fishes: Sister wrasses, *Coris gaimard* and *C. cuvieri*, in the
556 Red Sea, Indian Ocean, and Pacific Ocean. Journal of Biogeography *Online early*.

557 Alva-Campbell, Y., Floeter, S.R., Robertson, D.R., Bellwood, D.R., Bernardi, G., 2010.
558 Molecular phylogenetics and evolution of *Holacanthus* angelfishes (Pomacanthidae). Molecular
559 Phylogenetics & Evolution 56, 456-461.

560 Amos, B., Hoelzel, A.R., 1991. Long term preservation of whale skin for DNA analysis. Rept.
561 Intl. Whaling Commission Special Issue 13, 99-103.

562 Andrews, K.R., Williams, A., Fernandez-Silva, I., Newman, S.J., Copus, J.M., Bowen, B.W.,
563 2016. Phylogeny of deepwater snappers (Genus *Etelis*) reveals a cryptic species pair in the Indo-
564 Pacific and Pleistocene invasion of the Atlantic. Molecular Phylogenetics & Evolution
565 *(Accepted)*.

566 Bailey, G., 2009. The Red Sea, coastal landscapes, and hominin dispersals. In: Petraglia, M.D.,
567 Rose, J. (Eds.), The Evolution of Human Populations in Arabia. Springer, Dordrecht,
568 Netherlands, pp. 15-37.

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

571 Bellwood, D.R., van Herwerden, L., Konow, N., 2004. Evolution and biogeography of marine
572 angelfishes (Pisces: Pomacanthidae). Mol Phylogenet Evol 33, 140-155.

573 Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing
574 under dependency. Annals of Statistics 29, 1165-1188.

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

577 Bernardi, G., Beldade, R., Holbrook, S.J., Schmitt, R.J., 2012. Full-sibs in cohorts of newly
578 settled coral reef fishes. PLoS ONE 7, e44953.

579 Blum, S.D., 1989. Biogeography of the Chaetodontidae: an analysis of allopatry among closely
580 related species. Environmental Biology of Fishes 25, 9-31.

581 Bonjean, F., Lagerloef, G.S.E., 2002. Diagnostic model and analysis of the surface currents in
582 the tropical Pacific Ocean. Journal of Physical Oceanography 32, 2938-2954.

583 Bowen, B., Gaither, M.R., DiBattista, J.D., Iacchei, M., Andrews, K.R., Grant, W.S., Toonen,
584 R.J., Briggs, J.C., 2016. Comparative phylogeography of the ocean planet. Proceedings of the
585 National Academy of Sciences *(Accepted)*.

586 Bowen, B.W., Bass, A.L., Rocha, L.A., Grant, W.S., Robertson, D.R., 2001. Phylogeography of
587 the trumpetfishes (*Aulostomus*): Ring species complex on a global scale. Evolution 55, 1029-
588 1039.

589 Briggs, J.C., Bowen, B.W., 2012. A realignment of marine biogeographic provinces with
590 particular reference to fish distributions. Journal of Biogeography 39, 12-30.

591 Briggs, J.C., Bowen, B.W., 2013. Marine shelf habitat: Biogeography and evolution. Journal of
592 Biogeography 40, 1023-1035.

593 Chow, S., Hazama, K., 1998. Universal PCR primers for S7 ribosomal protein gene introns in
594 fish. Molecular Ecology 7, 1247-1263.

595 Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J.X.,
596 Hostetler, S.W., McCabe, A.M., 2009. The last glacial maximum. Science 325, 710-714.

597 Colborn, J., Crabtree, R.E., Shaklee, J.B., Pfeiler, E., Bowen, B.W., 2001. The evolutionary
598 enigma of bonefishes (*Albula* spp.): cryptic species and ancient separations in a globally
599 distributed shorefish. Evolution 55, 807-820.

600 Cowman, P.F., Bellwood, D.R., 2013. The historical biogeography of coral reef fishes: global
601 patterns of origination and dispersal. Journal of Biogeography 40, 209-224.

602 Cox, C.B., Moore, P.D., 2000. Biogeography: an ecological and evolutionary approach, 6th edn.
603 Blackwell Science, Oxford.

604 Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new
605 heuristics and parallel computing. Nature Methods 9, 772.

606 DiBattista, J.D., Waldrop, E., Bowen, B.W., Schultz, J.K., Gaither, M.R., Pyle, R.L., Rocha,
607 L.A., 2012. Twisted sister species of pygmy angelfishes: discordance between taxonomy,
608 coloration, and phylogenetics. Coral Reefs 31, 839-851.

609 DiBattista, J.D., Berumen, M.L., Gaither, M.R., Rocha, L.A., Eble, J.A., Choat, J.H., Craig,
610 M.T., Skillings, D.J., Bowen, B.W., 2013. After continents divide: comparative phylogeography
611 of reef fishes from the Red Sea and Indian Ocean. Journal of Biogeography 40, 1170-1181.

612 DiBattista, J.D., Choat, J.H., Gaither, M.R., Hobbs, J.P., Lozano-Cortés, D.F., Myers, R.F.,
613 Paulay, G., Rocha, L.A., Toonen, R.J., Westneat, M., Berumen, M.L., 2016a. On the origin of
614 endemic species in the Red Sea. Journal of Biogeography. 43, 13-30.

615 DiBattista, J.D., Roberts, M., Bouwmeester, J., Bowen, B.W., Coker, D.J., Lozano-Cortes, D.F.,
616 Choat, J.H., Gaither, M.R., Hobbs, J.-P.A., Khalil, M.T., Kochzius, M., Myers, R., Paulay, G.,
617 Robitzch, V.S.N., Saenz-Agudelo, P., Salas, E., Sinclair-Taylor, T.H., Toonen, R.J., Westneat,
618 M.W., Williams, S.T., Berumen, M.L., 2016b. A review of contemporary patterns of endemism
619 in the Red Sea. Journal of Biogeography. Journal of biogeography. 43, 423-439.

Drew, J., Allen, G.R., Kaufman, L.E.S., Barber, P.H., 2008. Endemism and Regional Color and Genetic Differences in Five Putatively Cosmopolitan Reef Fishes Endemismo y Color Regional y Diferencias Genéticas en Cinco Especies de Arrecife Supuestamente Cosmopolitas. Conservation Biology 22, 965-975.

Drew, J.A., Allen, G.R., Erdmann, M.V., 2010. Congruence between mitochondrial genes and color morphs in a coral reef fish: population variability in the Indo-Pacific damselfish Chrysiptera rex (Snyder, 1909). Coral Reefs 29, 439-444.

Eble, J.A., Toonen Robert J., Sorenson, L., Basch, L.V., Papastamatiou, Y., Bowen, B.W., 2011. Escaping paradise: larval export from Hawaii in an Indo-Pacific reef fish, the yellow tang Zebrasoma flavescens. Marine Ecology Progress Series 428, 245-258.

Eble, J.A., Bowen, B.W., Bernardi, G., 2015. Phylogeography of coral reef fishes. In: Mora, C. (Ed.), Ecology of Fishes on Coral Reefs. University of Hawaii Press, Honolulu, HI, pp. 64-75.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47-50.

Fernandez Silva, I., Randall, J.E., DiBattista, J., Coleman, R.R., CG, M., Rocha, L.A., Bowen, B.W., 2015. Phylogeography of the Indo-Pacific goatfish *Mulloidichthys flavolineatus*: Isolation in peripheral biogeographic provinces and a cryptic evolutionary lineage in the Red Sea. Journal of biogeography. *In press*.

Frankham, R., Briscoe, D.A., Ballou, J.D., 2002. Introduction to Conservation Genetics. Cambridge University Press, Cambridge.

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

Gaither, M.R., Toonen, R.J., Robertson, D.R., Planes, S., Bowen, B.W., 2010. Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). Journal of Biogeography 37, 133-147.

Gaither, M.R., Bowen, B.W., Bordenave, T.-R., Rocha, L.A., Newman, S.J., Gomez, J.A., van Herwerden, L., Craig, M.T., 2011. Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the indo-pacific barrier with contemporary overlap in the coral triangle. BMC Evolutionary Biology 11, 189-204.

Gaither, M.R., Rocha, L.A., 2013. Origins of species richness in the Indo-Malay-Philippine biodiversity hotspot: evidence for the centre of overlap hypothesis. Journal of Biogeography 40, 1638-1648.

Gaither, M.R., Schultz, J.K., Bellwood, D.R., Pyle, R.L., DiBattista, J.D., Rocha, L.A., Bowen, B.W., 2014. Evolution of pygmy angelfishes: Recent divergences, introgression, and the usefulness of color in taxonomy. Molecular Phylogenetics & Evolution 74, 38-47.

655 Gaither, M.R., Bernal, M.A., Coleman, R.R., Bowen, B.W., Jones, S.A., Simison, W.B., Rocha,
656 L.A., 2015. Genomic signatures of geographic isolation and natural selection in coral reef fishes.
657 Molecular Ecology 24, 1543-1557.

658 Guindon, S., Gascuel, O., 2003. A simple, fast and accurate method to estimate large
659 phylogenies by maximum-likelihood. Systematic Biology 5, 696-704.

660 Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New
661 algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the
662 performance of PhyML 3.0. Systematic Biology 59, 307-321.

663 Hellberg, M., 2009. Gene flow and isolation among populations of marine animals. Annual
664 Review of Ecology, Evolution, and Systematics 40, 291-210.

665 Hinton, S., 1962. Longevity of Fishes in Captivity as of September 1956. Zoologica 47, 105-116.

666 Hobbs, J.-P.A., Allen, G.R., 2014. Hybridisation among coral reef fishes at Christmas Island and
667 the Cocos (Keeling) Islands. Raffles Bulletin of Zoology, Supplement 30.

668 Hodge, J.R., Bellwood, D.R., 2016. The geography of speciation in coral reef fishes: the relative
669 importance of biogeographical barriers in separating sister species. Journal of Biogeography.
670 *(Online early)*.

671 Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., 2001. Bayesian inference of
672 phylogeny and its impact on evolutionary biology. Science 294, 2310-2314.

673 Johns, G.C., Avise, J.C., 1998. A comparative summary of genetic distances in the vertebrates
674 from the mitochondrial cytochrome *b* gene. Molecular Biology and Evolution 15, 1481-1490.

675 Kemp, J., 1998. Zoogeography of the coral reef fishes of the Socotra Archipelago. Journal of
676 Biogeography 25, 919-933.

677 Kulbicki, M., Parravicini, V., Bellwood, D.R., Arias-Gonzàlez, E., Chabanet, P., Floeter, S.R.,
678 Friedlander, A., McPherson, J., Myers, R.E., Vigliola, L., Mouillot, D., 2013. Global
679 biogeography of reef fishes: a hierarchical quantitative delineation of regions. PLOS One
680 8:e81847.

681 Leichter, J.J., Aildredge, A.L., Bernardi, G., Brooks, A.J., Carlson, C.A., Carpenter, R.C.,
682 Edmunds, P.J., Fewing, M.R., Hanson, K.M., Hench, J.L., 2013. Biological and physical
683 interactions on a tropical island coral reef transport and retention processes on Moorea, French
684 Polynesia. Oceanography 26, 52-63.

685 Leis, J.M., McCormick, M.I., 2002. The Biology, Behavior; and Ecology of the Pelagic, Larval
686 Stage of Coral Ree Fishes. In: Sale, P.F. (Ed.), Coral reef fishes: dynamics and diversity in a
687 complex ecosystem. Academic Press, San Diego, pp. 171-199.

688 Lemer, S., Planes, S., 2014. Effects of habitat fragmentation on the genetic structure and
689 connectivity of the black-lipped pearl oyster, *Pinctada margaritifera*, populations in French
690 Polynesia. Marine Biology 161, 2035-2049.

691 Leray, M., Beldade, R., Holbrook, S.J., Schmitt, R.J., Planes, S., Bernardi, G., 2010. Allopatric
692 divergence and speciation in coral reef fish: The three-spot dascyllus, *Dascyllus trimaculatus*,
693 species complex. Evolution 64, 1218-1230.

694 Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA
695 polymorphism data. Bioinformatics 25, 1451-1452.

696 Lo-Yat, A., Meekan, M.G., Carleton, J.H., Galzin, R., 2006. Large-scale dispersal of the larvae
697 of nearshore and pelagic fishes in the tropical oceanic waters of French Polynesia. Marine
698 Ecology Progress Series 325, 195-203.

699 Lovejoy, N., 1999. Systematics, Biogeography and Evolution of Needlefishes (Teleostei:
700 Belonidae). Cornell University, Ithaca, NY.

701 Ludt, W.B., Rocha, L.A., 2014. Shifting seas: the impacts of Pleistocene sea-level fluctuations
702 on the evolution of tropical marine taxa. Journal of Biogeography, 25-38.

703 McMillan, W.O., Weigt, L.A., Palumbi, S.R., 1999. Color pattern evolution, assortative mating,
704 and genetic differentiation in brightly colored butterflyfishes (Chaetodontidae). Evolution, 247-
705 260.

706 Meeker, N.D., Hutchinson, S.A., Ho, L., Trede, N.S., 2007. Method for isolation of PCR-ready
707 genomic DNA from zebrafish tissues. Biotechniques. 43, 610-614.

708 Narum, S.R., 2006. Beyond Bonferroni: less conservative analyses for conservation genetics.
709 Conserv Genet 7, 783-787.

710 Pianka, E.R., 1978. Evolutionary ecology. Harper and Row, New York, USA.

711 Planes, S., 1993. Genetic differentiation in relation to restricted larval dispersal of the convict
712 surgeonfish, *Acanthurus triostegus,* in French Polynesia. Marine Ecology Progress Series 98,
713 237-246.

714 Puglise, K.A., Hinderstein, L.M., Marr, J.C.A., Dowgiallo, M.J., Martinez, F.A., 2009.
715 Mesophotic coral ecosystems research strategy: International workshop to prioritize research and
716 management needs for mesophotic coral ecosystems, Jupiter, Florida, 12-15 July, 2009. US
717 Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean
718 Service.

719 Randall, J.E., 1994. Twenty-two new records of fishes from the Red Sea. Fauna of Saudi Arabia
720 14, 259-275.

721 Randall, J.E., 1998. Zoogeography of shore fishes of the Indo-Pacific region. Zoological Studies
722 37, 227-268.

723 Randall, J.E., 2005. Reef and Shore Fishes of the South Pacific: New Caledonia to Tahiti and the
724 Pitcairn Island. Unversity of Hawaii Press, Honolulu, HI.

725 Reece, J.S., Bowen, B.W., Joshi, K., Goz, V., Larson, A., 2010a. Phylogeography of two moray
726 eels indicates high dispersal throughout the Indo-Pacific. Journal of Heredity 101, 391-402.

727 Reece, J.S., Bowen, B.W., Smith, D.G., Larson, A.F., 2010b. Molecular phylogenetics of moray
728 eels (Murenidae) demonstrates multiple origins of a shell-crushing jaw (Gymnomuraena,
729 Echidna) and multiple colonizations of the Atlantic Ocean. Molecular Phylogenetics & Evolution
730 57, 829 – 835.

731 Robertson, D.R., Karg, F., Leao de Moura, R., Victor, B.C., Bernardi, G., 2006. Mechanisms of
732 speciation and faunal enrichment in Atlantic parrotfishes. Molecular Phylogenetics & Evolution
733 40, 795-807.

734 Rocha, L.A., Bass, A.L., Robertson, D.R., Bowen, B.W., 2002. Adult habitat preferences, larval
735 dispersal, and the comparative phylogeography of three Atlantic surgeonfishes (Teleostei:
736 Acanthuridae). Molecular Ecology 11, 243-251.

737 Rocha, L.A., Craig, M.T., Bowen, B.W., 2007. Phylogeography and the conservation of coral
738 reef fishes. Coral Reefs 26, 501-512.

739 Rocha, L.A., Lindeman, K.C., Rocha, C.R., Lessios, H.A., 2008. Historical biogeography and
740 speciation in the reef fish genus *Haemulon* (Teleostei: Haemulidae). Molecular Phylogenetics &
741 Evolution 48, 918-928.

742 Ronquist, F., 2004. Bayesian inference of character evolution. Trends in Ecology & Evolution
743 19, 475-481.

744 Schultz, J.K., Pyle, R.L., DeMartini, E., Bowen, B.W., 2006. Genetic connectivity among color
745 morphs and Pacific archipelagos for the flame angelfish, *Centropyge loriculus*. Marine Biology
746 151, 167-175.

747 Siddall, M., Rohling, E.J., Almogi-Labin, A., Hemleben, C., Meischner, D., Schmelzer, I.,
748 Smeed, D.A., 2003. Sea-level fluctuations during the last glacial cycle. Nature 423, 853.

749 Song, C.B., Near, T.J., Page, L.M., 1998. Phylogenetic relations among percid fishes as inferred
750 from mitochondrial cytochrome *b* DNA sequence data. Molecular Phylogenetics & Evolution 10,
751 343-353.

752 Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype
753 reconstruction from population genotype data. American Journal of Human Genetics 73, 1162-
754 1169.

755 Szabo, Z., Snelgrove, B., Craig, M.T., Rocha, L.A., Bowen, B.W., 2014. Phylogeography of the
756 Manybar Goatfish, *Parupeneus multifasciatus* reveals moderate structure between the Central
757 and North Pacific and a cryptic endemic species in the Marquesas. Bulletin of Marine Science
758 90, 493-512.

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

761 Thresher, R.E., Brothers, E.B., 1985. Reproductive ecology and biogeography of Indo-West
762 Pacific Angelfishes (Pisces: Pomacanthidae. Evolution 39, 878-887.

763 Timmers, M.A., Bird, C.E., Skillings, D.J., Smouse, P.E., Toonen, R.J., 2012. There's no place
764 like home: Crown-of-Thorns outbreaks in the central Pacific are regionally derived and
765 independent events. PLoS ONE 7, e31159.

766 Vogler, C., Benzie, J., Lessios, H., Barber, P., Worheide, G., 2008. A threat to coral reefs
767 multiplied? Four species of crown-of-thorns starfish. Biology letters 4, 696-699.

768 Voris, H.K., 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems
769 and time durations. Journal of Biogeography 27, 1153-1167.

770 Waldrop, E., Hobbs, J.P., Randall, J.E., DiBattista, J.D., Rocha, L.A., Kosaki, R.K., Berumen,
771 M.L., Bowen, B.W., 2016. Phylogeography, population structure, and evolution of coral-eating
772 butterflyfishes (Subgenus *Corallochaetodon*). Journal of Biogeography. *In press*.

773 Winterbottom, R., Anderson, R.C., 1997. A revised checklist of the epipelagic and shore fishes
774 of the Chagos Archipelago, Central Indian Ocean. Smith Institute of Ichthyology Ichthyological
775 Bulletin 66, 1-28.

776 Wright, S., 1978. Evolution and the Genetics of Populations: A Treatise in Four Volumes: Vol.
777 4: Variability Within and Among Natural Populations. University of Chicago Press, Chicago.

778 Wyrtki, K., Kilonsky, B., 1984. Mean water and current structure during the Hawaii-to-Tahiti
779 shuttle experiment. Journal of Physical Oceanography 14, 242-254.
780

Figure Captions

Figure 1. Map of collection locations, sample sizes (in parentheses), and the two recognized morphotypes of *Pygoplites diacanthus*. (*left*) Indian Ocean and Red Sea individuals are characterized by a yellow chest and head, whereas the (*right*) Pacific Ocean morph is characterized by a gray chest and head. *Photos by L. Rocha (Djibouti; Great Barrier Reef, Australia)*

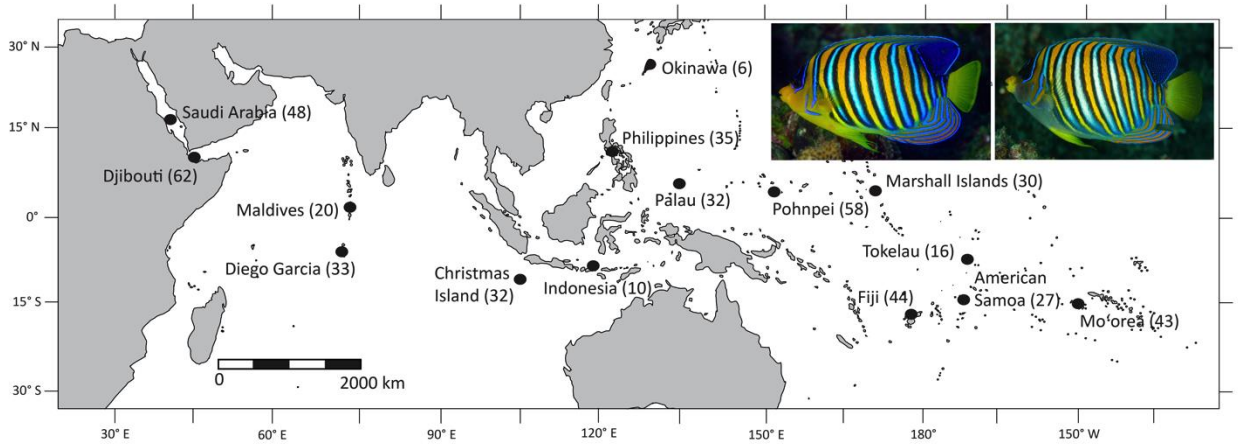
Figure 2. Molecular phylogenetic reconstruction of *Pygoplites diacanthus.* A) Rooted Bayesian tree based on mitochondrial cytochrome *b* with posterior probabilities, B) an unrooted maximum-likelihood tree based on mitochondrial and nuclear markers (cytochrome *b*, intron 1 of the S7 ribosomal protein, and the recombination-activating gene 2) with consensus values based on posterior probabilities from Bayesian inference (BI), maximum-likelihood bootstrap support (ML), and neighbor-joining bootstrap support (NJ). Percent sequence divergence is represented on the scale bar. The sizes of black triangles are proportional to the number of individuals within the lineage. Abbreviations: Red Sea Province, RS; Indian Ocean, IO.

Figure 3. Median-joining network for *Pygoplites diacanthus* constructed using NETWORK for A) cytochrome *b* sequences (568 bp) from 386 individuals, B) alleles for RAG2 (431 bp) from 366 individuals, and c) alleles for the S7 intron (510 bp) from 288 individuals. Each circle represents a unique mitochondrial haplotype or nuclear allele, with the size being proportional to the total frequency. Open circles represent unsampled alleles, branches and crossbars represent a single

803 nucleotide change, and color represents collection location (see key). All singleton alleles (*N* =

804 22) were removed from the S7 analysis to minimize circularity between closely related alleles.

805 Figure 1.



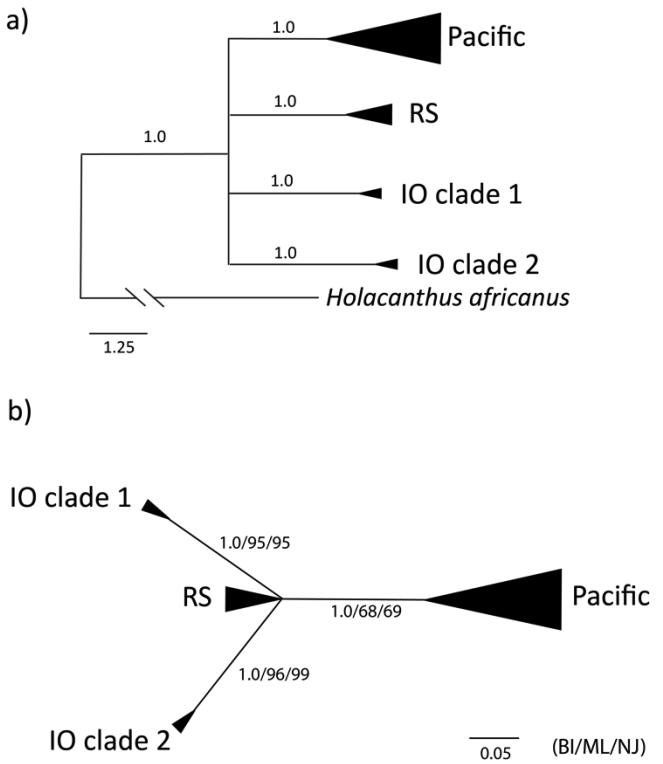
806

807

808

809 Figure 2.

810



811

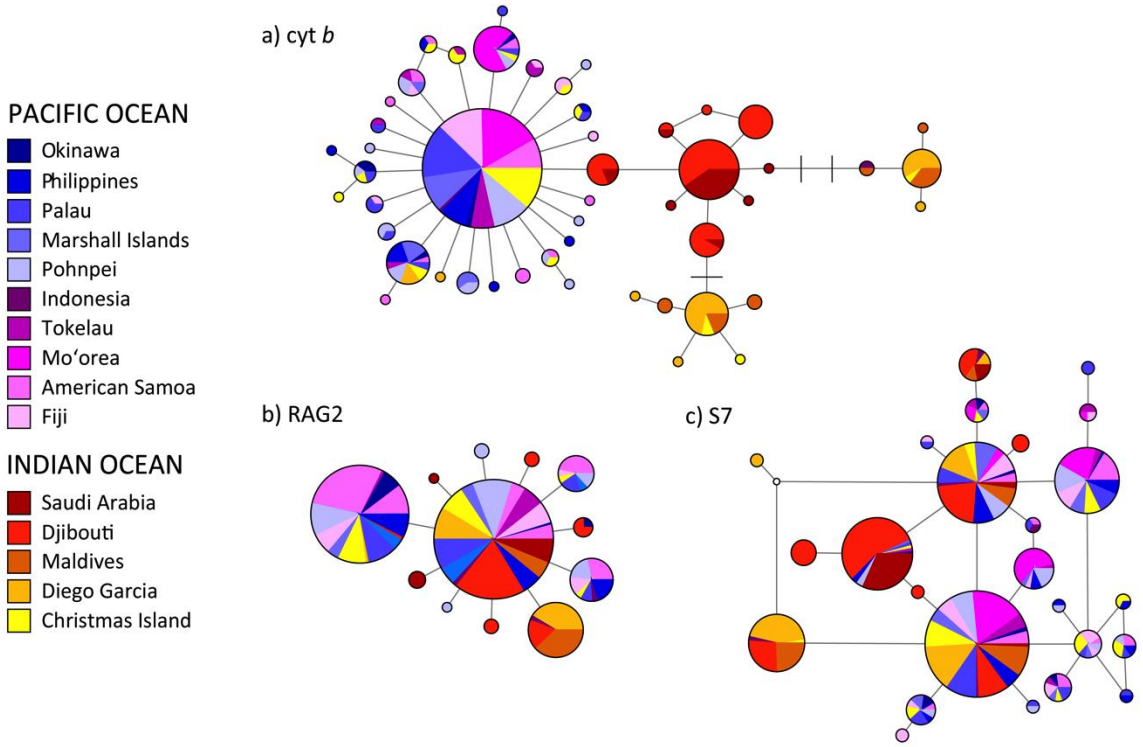
812

813

814

815 Figure 3.

816



817

Table 1. Molecular diversity indices for lineages of *Pygoplites diacanthus* based on mitochondrial DNA (cytochrome *b*, 568 bp). Number of individuals sequenced (*n*), number of haplotypes (N_h), number of segregating (polymorphic) sites (*S*), haplotype diversity (*h*), and nucleotide diversity (π) are presented. Times to most recent common ancestor (TMRCA) are presented as million years. Bolded numbers denote significance at $P < 0.02$.

Lineage	*n*	N_h	*S*	*h* ± SD	π ± SD	TMRCA (95% HPD)	Fu's F_S	Fu's F_S *P*-value
Pacific Ocean[a]	257	33	37	0.628 ± 0.034	0.002 ± 0.010	1.71 (0.91 - 2.65)	**-29.51**	<0.001
Red Sea[b]	81	9	8	0.701 ± 0.042	0.002 ± 0.003	1.44 (0.51 - 2.53)	-3.602	0.035
Indian Ocean Lineage 1	28	6	5	0.439 ± 0.114	0.001± 0.001	0.72 (0.14 - 1.52)	**-3.695**	<0.001
Indian Ocean Lineage 2	20	4	3	0.284 ± 0.128	0.001 ± 0.001	0.92 (0.27 - 1.75)	**-2.749**	0.001
All Locations	386	49	45	0.817 ± 0.018	0.005 ± 0.003	--	**-25.90**	<0.001

[a]Pacific includes all Pacific Ocean populations plus Christmas Island
[b]Red Sea includes Saudi Arabia and Djibouti

818

Table 2. Molecular diversity indices for populations of *Pygoplites diacanthus* based on mitochondrial DNA (cytochrome *b*, 568 bp) divided into phylogeographical groupings. Number of individuals sequenced (n), number of haplotypes (N_h), number of segregating (polymorphic) sites (S), haplotype diversity (h), and nucleotide diversity (π) are presented. τ is used to estimate the age of most recent population expansion (population age) using the equation $\tau = 2\mu t$ (see Material and Methods). ∞ denotes values that could not be resolved. Bolded numbers denote significance at $P < 0.02$.

Sample Location	n	N_h	S	$h \pm$ SD	$\pi \pm$ SD	τ	Population Age (years)	Fu's F_S	Fu's F_S P-value
Pacific Ocean									
Okinawa	6	4	3	0.867 ± 0.129	0.002 ± 0.002	1.54	135,000 - 271,000	-1.454	0.052
Philippines	21	7	8	0.657 ± 0.104	0.001 ± 0.001	1.04	92,000 - 183,000	**-3.473**	0.003
Palau	32	9	8	0.488 ± 0.109	0.001 ± 0.001	0.67	59,000 - 118,000	**-7.928**	< 0.001
Marshall islands	23	4	3	0.549 ± 0.105	0.001 ± 0.001	0.78	69,000 -138,000	-0.936	0.208
Pohnpei	33	12	13	0.760 ± 0.076	0.002 ± 0.001	1.30	115,000 - 230,000	**-8.754**	< 0.001
Indonesia	2	2	7	1.000 ± 0.500	0.012 ± 0.013	∞	∞	1.946	0.519
Tokelau	16	6	5	0.617 ± 0.135	0.001 ± 0.001	0.92	81,000 - 162,000	**-3.692**	< 0.001
Mo'orea	42	2	1	0.483 ± 0.039	0.001 ± 0.001	0.73	64,000 - 128,000	1.766	0.738
American Samoa	25	10	9	0.730 ± 0.094	0.002 ± 0.001	1.20	106,000 - 212,000	**-7.128**	< 0.001
Fiji	25	6	5	0.427 ± 0.122	0.001 ± 0.001	0.56	49,000 - 98,000	**-4.423**	< 0.001
Christmas Island	32	13	19	0.720 ± 0.087	0.004 ± 0.003	0.44	39,000 - 77,000	**-5.445**	0.006
Red Sea Province									
Saudi Arabia	23	7	7	0.522 ± 0.124	0.001 ± 0.001	0.74	65,000 - 130,000	**-4.731**	< 0.001
Djibouti	58	6	4	0.738 ± 0.034	0.002 ± 0.001	1.19	105,000 - 209,000	-0.879	0.349
Indian Ocean									
Maldives	16	6	11	0.808 ± 0.069	0.009 ± 0.005	9.89	871,000 - 1,742,000	1.858	0.804
Diego Garcia	32	7	17	0.692 ± 0.059	0.009 ± 0.005	9.17	807,000 - 1,614,000	3.177	0.898
Pacific Ocean	257	33	37	0.628 ± 0.034	0.002 ± 0.009	0.94	83,000 - 165,000	**-29.511**	<0.001
Red Sea Province	81	9	8	0.701 ± 0.042	0.002 ± 0.003	1.09	96,000 - 192,000	-3.602	0.035
Indian Ocean	48	11	19	0.738 +/- 0.045	0.009 ± 0.008	9.37	825,000 - 1,649,000	1.013	0.700
All Locations	386	49	45	0.817 ± 0.018	0.005 ± 0.003	3.61	318,000 - 636,000	**-25.897**	<0.001

819

820

821
822

Table 3. Matrix of pairwise Φ_{ST} statistics for 13 populations of *Pygoplites diacanthus* based on mitochondrial DNA (cytochrome *b*, 568 bp) sequences. Bolded numbers indicate significance after controlling for false discovery rates at $\alpha = 0.05$ (as per Narum, 2006). The corrected $\alpha = 0.009$. Owing to low sample size, Okinawa and Indonesia have been excluded. Abbreviations: Red Sea Province, RS; Indian Ocean, IO.

		Pacific Ocean									RS		IO
	Sample location	1	2	3	4	5	6	7	8	9	10	11	12
Pacific Ocean	1. Philippines	--											
	2. Palau	0.02286	--										
	3. Marshall Is.	-0.00690	0.04436	--									
	4. Pohnpei	-0.00113	-0.00003	-0.00602	--								
	5. Tokelau	-0.00088	0.00970	0.02593	-0.00273	--							
	6. Mo'orea	**0.21137**	**0.15115**	**0.24879**	**0.12293**	**0.21718**	--						
	7. American Samoa	0.00241	0.01337	0.01878	-0.00796	-0.00602	**0.12806**	--					
	8. Fiji	0.04077	0.00310	**0.06141**	0.00599	-0.00058	**0.22924**	0.01738	--				
	9. Christmas Is.	0.02264	0.04064	0.02876	0.02685	0.01493	**0.13152**	0.02805	0.03767	--			
RS	10. Saudi Arabia	**0.76918**	**0.80926**	**0.80834**	**0.73676**	**0.79578**	**0.83717**	**0.75536**	**0.82668**	**0.55505**	--		
	11. Djibouti	**0.73859**	**0.76222**	**0.75577**	**0.72292**	**0.74532**	**0.78731**	**0.73266**	**0.76756**	**0.58986**	0.05789	--	
IO	12. Maldives	**0.64673**	**0.71028**	**0.67118**	**0.67292**	**0.63123**	**0.75342**	**0.65734**	**0.69346**	**0.53679**	**0.42875**	**0.50323**	--
	13. Diego Garcia	**0.51036**	**0.56474**	**0.52244**	**0.53912**	**0.49451**	**0.61025**	**0.52284**	**0.54533**	**0.41311**	**0.28493**	**0.35104**	-0.00728

823

824

Table 4. Results of the analysis of molecular variance (AMOVA) based on mitochondrial DNA (cytochrome *b*) sequence data for *Pygoplites diacanthus*. Bolded values denote significance at $P < 0.05$.

	Among groups			Among populations (within groups)			Within populations		
Regions	Φ_{CT}	*P*-value	% variation	Φ_{SC}	*P*-value	% variation	Φ_{ST}	*P*-value	% variation
Pacific Ocean vs. Indian Ocean	0.60	0.058	59.91	**0.19**	< 0.001	**7.46**	**0.67**	< 0.001	**32.63**
Pacific[a] vs. Indian[b] vs. Red Sea[c]	**0.66**	< 0.001	**65.53**	**0.04**	0.017	**1.44**	**0.67**	< 0.001	**33.03**
Indian[b] vs. Red Sea[c] vs. Christmas Is.	0.44	0.078	44.09	**0.02**	< 0.001	**0.92**	0.45	0.269	54.99
Pacific[a] vs. Mo'orea	0.08	0.184	7.92	**0.05**	< 0.001	**4.82**	**0.13**	< 0.001	**87.26**

[a]Pacific includes all Pacific Ocean populations plus Christmas Island.
[b]Indian includes the Maldives and Diego Garcia.
[c]Red Sea includes Saudi Arabia and Djibouti.

825

826

827

828

Table 5. Molecular diversity indices for populations of *Pygoplites diacanthus* based on nuclear DNA (introns RAG2 and S7) for all populations. Number of individuals sequenced (n), number of alleles (N_a), number of segregating (polymorphic) sites (S), observed heterozygosity (H_O), expected heterozygosity (H_E), and the corresponding P-value

	RAG2						S7					
Sample Location	n	N_a	S	H_O	H_E	P-value	n	N_a	S	H_O	H_E	P-value
Pacific Ocean												
Okinawa	6	3	1	0.50	0.59	1.00	5	8	8	0.80	0.93	0.37
Philippines	21	3	2	0.38	0.46	0.32	15	11	9	0.73	0.85	0.19
Palau	30	4	3	0.47	0.41	0.67	22	14	12	0.88	0.82	< 0.001
Marshall islands	27	3	2	0.26	0.29	0.55	14	9	8	0.71	0.82	0.16
Pohnpei	39	5	4	0.38	0.43	0.20	21	15	16	0.76	0.87	0.24
Indonesia	4	3	2	0.50	0.46	1.00	3	6	7	1.00	1.00	1.00
Tokelau	16	2	1	0.06	0.06	1.00	8	8	8	0.63	0.81	0.04
Mo'orea	31	4	3	0.61	0.64	0.83	30	7	7	0.80	0.71	0.33
American Samoa	18	3	2	0.44	0.54	0.40	16	10	10	0.75	0.85	0.14
Fiji	21	4	3	0.43	0.43	0.83	16	12	10	0.88	0.85	0.74
Christmas Island	25	4	3	0.28	0.39	0.14	18	12	11	0.78	0.84	0.58
Red Sea Province												
Saudi Arabia	19	3	2	0.21	0.20	1.00	15	5	7	0.47	0.41	1.00
Djibouti	59	6	5	0.22	0.22	0.61	52	8	7	0.85	0.80	0.43
Indian Ocean												
Maldives	19	2	1	0.37	0.46	0.61	18	5	5	0.61	0.70	0.24
Diego Garcia	31	3	2	0.39	0.38	0.25	31	8	9	0.84	0.72	0.93
All Locations	366	12	10	0.35	0.43	<0.001	284	44	31	0.77	0.86	< 0.001

Table 6. Matrix of pairwise *F*-statistics for 13 populations of *Pygoplites diacanthus*. Φ_{ST} values for RAG2 (below diagonal) and S7 (above diagonal). Bolded numbers indicate significance after controlling for false discovery rates at α = 0.05 (as per Narum, 2006). The corrected α = 0.009. Owing to low sample size, Okinawa and Indonesia have been excluded. Abbreviations: Red Sea Province, RS; Indian Ocean, IO.

		Pacific Ocean									RS		IO	
	Sample location	1	2	3	4	5	6	7	8	9	10	11	12	13
Pacific Ocean	1. Philippines	--	-0.00920	-0.01362	-0.01360	-0.01687	0.01687	0.01128	-0.00514	0.01515	**0.34364**	**0.17142**	**0.10393**	**0.09571**
	2. Palau	-0.00538	--	0.00508	-0.00296	-0.02460	0.02700	0.00922	-0.01550	-0.00726	**0.38275**	**0.22126**	**0.09033**	**0.08852**
	3. Marshall Is.	0.00818	-0.01120	--	-0.00031	-0.00704	0.05734	0.00961	-0.00460	0.04995	**0.34476**	**0.15924**	**0.14964**	**0.13677**
	4. Pohnpei	-0.00668	-0.01082	-0.00789	--	-0.02035	0.01691	0.00989	-0.00473	0.01182	**0.30090**	**0.18403**	**0.07945**	**0.08287**
	5. Tokelau	0.07767	0.04393	0.04048	0.02592	--	0.00217	-0.01313	-0.02998	-0.01183	**0.37543**	**0.20290**	**0.09332**	**0.09858**
	6. Mo'orea	**0.11389**	**0.13061**	**0.16703**	**0.16493**	**0.27102**	--	0.07372	0.03949	0.03840	**0.44338**	**0.25787**	**0.07336**	**0.07570**
	7. American Samoa	-0.00096	0.02568	0.05152	0.03906	**0.17279**	0.03921	--	-0.01129	0.02085	**0.36682**	**0.24862**	**0.18011**	**0.18764**
	8. Fiji	-0.02068	-0.01577	-0.00644	-0.01574	0.05125	**0.12489**	0.01209	--	0.00554	**0.40033**	**0.23471**	**0.13818**	**0.13340**
	9. Christmas Is.	-0.00815	-0.01338	-0.01051	-0.00520	0.07526	**0.11113**	0.00810	-0.01339	--	**0.43065**	**0.26948**	**0.09838**	**0.10584**
RS	10. Saudi Arabia	**0.10905**	**0.08062**	**0.08407**	0.06289	0.02759	**0.29673**	**0.19761**	**0.08516**	**0.11256**	--	0.07234	**0.51195**	**0.50068**
	11. Djibouti	**0.12103**	**0.08643**	**0.07487**	**0.06583**	0.00215	**0.3587**	**0.22863**	**0.09234**	**0.11642**	0.02992	--	**0.25737**	**0.25281**
IO	12. Maldives	**0.23970**	**0.23544**	**0.26566**	**0.23347**	**0.28203**	**0.34276**	**0.27897**	**0.23017**	**0.25607**	**0.25779**	**0.22058**	--	-0.00921
	13. Diego Garcia	**0.16033**	**0.14521**	**0.15459**	**0.1368**	0.13833	**0.31590**	**0.22167**	**0.14479**	**0.16565**	**0.13958**	**0.09067**	0.01181	--