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, PO Box 1346, Kāne'ohe, HI 96744, USA,* [b]*Department of Biology, University of Hawai'i, Mānoa, 2500 Campus Rd, Honolulu, HI 96822, USA,* [c]*University of West Florida, 11000 University Pkwy, 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, 55 Music Concourse Dr, San Francisco, CA 94118, USA,* [g]*Bernice Pauahi Bishop Museum, 1525 Bernice St, 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 hypthosis 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 monotypic 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™ (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™ (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: 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 matched 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 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 with significance determined at $P < 0.02$. A significant negative value
163 of Fu's F_S is evidence for an excess number of alleles, as would be expected from a recent
164 population expansion, whereas, a significant positive value is evidence for a deficiency of alleles,
165 as would be expected from a recent population bottleneck.
166
167 *2.3 Nuclear DNA Analysis*

168 We sequenced two nuclear loci: the recombination-activating gene 2 (RAG2) and intron
169 1 of the S7 ribosomal protein (S7). We resolved 431-bp of RAG2 using modified primers from
170 Lovejoy (1999); the forward primer is 5'-SACCTTGTGCTGCAAAGAGA-3' and reverse
171 primer is 5'-AGTGGATCCCCTTBTCATCCAGA-3'. We resolved 510-bp of S7 using primers
172 S7RPEX1F and S7RPEX2R from Chow and Hazama (1998). For each intron, PCR was
173 performed using the same reaction as described for cyt *b* but using the following temperature
174 conditions: 5 min at 94°C, 35 cycles of denaturing for 30 s at 94°C, annealing for 30 s at 58°C,
175 extension for 45 s at 72°C, and a final extension for 10 min at 72°C.
176 Allelic states with more than one heterozygous site were estimated using PHASE v.2.1
177 (Stephens and Donnelly, 2003) as implemented in DNASP. Unique sequences were deposited in
178 GenBank (Accession numbers: RAG2, KU885737 - KU885756; S7, KU885757 - KU885843)
179 Three separate runs, each of 100,000 repetitions after a 10,000 iteration burn-in, were conducted
180 for each locus; all runs returned consistent allele identities. Median-joining networks were
181 created for each nuclear dataset as outlined above. To minimize circularity between closely
182 related alleles, singletons were removed from the S7 network. However, this did not alter our
183 overall interpretation of the results. Pairwise Φ_{ST} statistics were calculated for each nuclear
184 dataset. The best-fit model of K80 and TPM1uf+I (proportion of invariable sites = 0.89) were
185 identified for RAG2 and S7, respectively as determined by JMODELTEST. Observed
186 heterozygosity (H_O) and expected heterozygosity (H_E) were calculated for each locus and an
187 exact test of Hardy-Weinberg Equilibrium (HWE) using 100,000 steps in a Markov chain was
188 performed in ARLEQUIN.
189
190 *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

214 (*Holacanthus*) is deeply divergent at cyt *b* (Alva-Campbell et al., 2010). The Pacific lineage is
215 0.6% divergent from the Red Sea lineage and 1.2% and 1.5% from Indian lineage 1 and 2,
216 respectively. The Red Sea lineage is 0.6% divergent from Indian lineage 1 and 1.0% from Indian
217 lineage 2, and the two Indian lineages are distinguished by 1.5% divergence. Coalescence
218 analysis based on cyt *b* yielded a TMRCA of 1.7 Ma for the Pacific lineage, and identified the
219 Pacific as the oldest extant lineage (Table 1). The Red Sea lineage dates to 1.4 Ma and the two
220 Indian lineages were the youngest: Indian lineage 1 at 0.7 Ma; Indian lineage 2 at 0.9 Ma.
221
222 *3.2 MtDNA Sequences*
223 Mitochondrial DNA molecular diversity indices are summarized for lineages in Table 1
224 and among populations in Table 2. Total haplotype diversity was $h = 0.817$ with 49 unique
225 haplotypes. Among lineages, the Red Sea had the highest haplotype diversity ($h = 0.701$) with
226 the lowest being observed in Indian Ocean lineage 2 ($h = 0.284$). Within populations, Indonesia
227 had the highest haplotype diversity ($h = 1.00$) followed by Okinawa *(*$h = 0.867$) and the
228 Maldives ($h = 0.808$). The lowest haplotype diversity was observed at Fiji ($h = 0.427$) and
229 Mo'orea ($h = 0.483$). Total nucleotide diversity was $\pi = 0.005$. Among lineages, the Pacific
230 Ocean and Red Sea had the higher nucleotide diversity ($\pi = 0.002$) with the lowest nucleotide
231 diversity observed in both Indian Ocean lineages ($\pi = 0.001$). Among populations, the Maldives
232 and Diego Garcia had the highest nucleotide diversity for all locations, each at $\pi = 0.009$,
233 whereas the lowest nucleotide diversity is observed at Fiji ($\pi = 0.0008$).
234 The median-joining haplotype network illustrates the low level of divergence between the
235 four evolutionary lineages recovered from the phylogenetic analysis (Fig. 3a). However, the
236 network also reveals that Red Sea haplotypes lie between the Pacific and Indian haplotypes. The

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

244 Population pairwise Φ_{ST} values for cyt *b* results are summarized in Table 3. Significance
245 was determined after controlling for false discovery rates (corrected $\alpha = 0.009$). Φ_{ST} values show
246 congruence with the haplotype network further supporting the Pacific, Indian, and Red Sea
247 groups. There was little or no population structure detected within these groups, with two
248 exceptions: Mo'orea (French Polynesia) shows significant genetic differentiation from all Pacific
249 locations, with pairwise Φ_{ST} values ranging from 0.123 with Pohnpei to 0.229 with Fiji.
250 Elsewhere in the Pacific, significant genetic structure was detected between the Marshall Islands
251 and Fiji ($\Phi_{ST} = 0.061$, $P < 0.001$). Although population level data is not reported for the single
252 location in the Sino-Japanese Province (Okinawa) due to low sample size ($N = 6$), preliminary
253 runs show no significant population structure between the Sino-Japanese Province and the
254 Pacific samples of *P. diacanthus*. The Red Sea lineage shows high levels of population
255 differentiation from all other samples (pairwise Φ_{ST}: 0.284 – 0.837). Likewise, the Indian
256 lineages show significant population differentiation from all other samples (pairwise Φ_{ST}: 0.284
257 – 0.753). The AMOVA analysis supports the Pacific, Indian, and Red Sea geographic groupings
258 based on mtDNA (Table 4) with the majority of the variation ($\Phi_{CT} = 0.66$, $P < 0.001$) existing
259 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, $P = 0.35$) aged 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

283 networks include common alleles that are observed only in the Pacific, or only in the Indian
284 Ocean locations. For S7, an Indian Ocean specific allele is also shared with a single individual
285 from Christmas Island.
286 The population genetic results for the nuclear dataset are strongly concordant with
287 mtDNA analyses for *P. diacanthus*, although they differ by degree. Genetic structure was absent
288 within the Red Sea and within the Indian Ocean. The only significant differentiation in the
289 Pacific was in 7 of 8 comparisons to Mo'orea (Society Islands, French Polynesia) with RAG2
290 (Φ_{ST} = 0.111 – 0.271; Table 6). Curiously, none of the same pairwise comparisons for Mo'orea
291 were significant with S7, however Mo'orea showed the highest differentiation from Red Sea
292 populations.
293 Both nuclear markers show high levels of genetic structure that correspond to a Pacific,
294 Indian, and Red Sea lineage. RAG2 was significant in 17 of 18 Pacific versus Indian
295 comparisons (Φ_{ST} = 0.137 – 0.343), significant in all Indian versus Red Sea comparisons (Φ_{ST} =
296 0.091 – 0.258), and significant in 15 of 18 Pacific versus Red Sea comparisons (Φ_{ST} = 0.066 –
297 0.359). The S7 differences were significant in all Pacific versus Indian comparisons (Φ_{ST} = 0.073
298 – 0.188), all Indian versus Red Sea comparisons (Φ_{ST} = 0.253 – 0.512), and all Pacific versus
299 Red Sea comparisons (Φ_{ST} = 0.159 – 0.443). The exceptions to these patterns were comparisons
300 between the Red Sea lineage and Tokelau, as well as between Saudi Arabia and Pohnpei. For S7
301 the highest genetic structure was observed between the Indian and Red Sea populations. This
302 contrasts with the RAG2 and cyt *b* comparisons, where the highest genetic structure
303 differentiated the Pacific from both Indian and Red Sea regions.
304
305 **4. Discussion**

306 *4.1 Summary of results*

307 Our data demonstrates that cryptic diversity exists within the monotypic genus *Pygoplites*
308 as evidenced by significant levels of genetic structure among three regions: the Pacific Ocean
309 (which includes a cohort at Christmas Island), the Indian Ocean (with two sympatric mtDNA
310 lineages), and the Red Sea (Table 4). This pattern of genetic structure corresponds to known
311 biogeographic provinces and phylogeographic barriers observed in other reef fishes (Rocha et al.,
312 2007; Briggs and Bowen, 2013; DiBattista et al., 2013; Eble et al., 2015; Gaither et al., 2015).
313 The Red Sea biogeographic province is distinguished by a faunal break at the Gulf of Aden, and
314 the Indo-Pacific Barrier is an intermittent terrestrial bridge between Australia and SE Asia that
315 impedes water movement between Pacific and Indian Oceans during glacial low-sea levels (see
316 Gaither & Rocha, 2013). The Sino-Japanese Province shows no genetic differentiation from the
317 Pacific population (based on $N = 6$), but Mo'orea is highly isolated, a finding we attribute to
318 prevailing oceanographic conditions (see Gaither et al., 2010). Below we discuss the
319 phylogenetic implications of cryptic lineages and examine each of these regions in light of
320 biogeographic theory

321

322 *4.2 Phylogenetic considerations*

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

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

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

351 preliminary morphological examination by L.A.R. revealed no additional morphological
352 characters that discriminate between Indian and Red Sea lineages.
353 The geographical delineation between the Pacific and Indian lineages correspond with the
354 exposure of the Sunda Shelf, which separates the Pacific and Indian Oceans during low sea level.
355 The Red Sea lineage corresponds to the Red Sea biogeographic province, which encompasses the
356 adjacent Gulf of Aden (Briggs and Bowen, 2012) and whose populations have a disjunct
357 distribution with the remainder of the range (see below). During glacial maxima the Red Sea is
358 effectively cut off from the Indian Ocean by closure of the Strait of Bab al Mandab, the only
359 natural gateway into the Red Sea, allowing sufficient time for populations to diverge into distinct
360 evolutionary lineages (DiBattista et al., 2013, 2016a).
361 The mechanisms facilitating two sympatric mtDNA lineages in the Indian Ocean are less
362 clear. Coalescence estimates indicate that lineages arose independently during roughly the same
363 period (0.72 – 0.93 Ma). As there are no known phenotypic differences within this region, the
364 unexpected recovery of two distinct lineages requires further investigation. Indian Ocean
365 samples contained similar number of each lineage (Maldives: Lineage 1, $N = 8$; Lineage 2, $N =$
366 8; Diego Garcia: Lineage 1, $N = 17$; Lineage 2, $N = 11$) indicating that the two lineages are
367 approximately equally represented.
368 The recovery of multiple evolutionary partitions within the monotypic genus *Pygoplites*
369 may not be indicative of other monotypic genera. Cryptic evolutionary partitions are routinely
370 discovered within species of marine fishes (Colborn et al., 2001; Rocha et al., 2008; DiBattista et
371 al., 2012; Fernandez Silva et al., 2015), and in this regard *P. diacanthus* is similar to the more
372 speciose inhabitants of Indo-Pacific reefs. The factors that produce a deep, monotypic lineage are
373 therefore not reflected in an unusual phylogeographic architecture. However, part of the

374 explanation for this monotype may be that five to eight million years after the divergence of
375 *Pygoplites* and *Holocanthus*, the ancestor of all modern *Pygoplites* likely radiated out of the
376 West Pacific Ocean, an extensive source of Indo-Pacific diversity (Cowman and Bellwood,
377 2013).

378 In considering the phylogenetic results through a taxonomic lens, there are several issues.
379 First, the Pacific and Indian morphs are distinguished by diagnostic differences, but they are not
380 monophyletic. The Indian Ocean contains two mtDNA lineages, each more closely related to the
381 Red Sea lineage than to each other. Second, the coloration difference between Pacific and Indian
382 forms, now matched by $d = 0.006$ divergence, could be a platform to describe them as separate
383 species. Third, the genetic divergence observed at all three loci is low in comparison to typical
384 divergences for fish species ($d = 0.03 – 0.12$; Grant and Bowen, 1998; Johns and Avise, 1998).
385 Fourth, the two morphs form mixed groups where they co-occur at Christmas Island (Hobbs and
386 Allen, 2014). Since we lack diagnostic nDNA alleles for the two morphs, we do not have the
387 power to test for hybrids between the lineages, but this is certainly a possibility. Given these
388 considerations, we believe that it is problematic to invoke species status for these three regional
389 forms and we endorse subspecies recognition distinguishing the Pacific lineage from the Indian
390 and Red Sea lineages based on shallow but diagnostic distinctions in genetics and morphology.
391 We propose the name *P. diacanthus flavescens* for the Indian Ocean and Red Sea lineages to
392 give recognition to the yellow chest coloration, a character not found in individuals from the
393 Pacific lineage (*P. d. diacanthus*).
394
395 *4.3 Red Sea isolation and refugia*
396 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 Somolia, 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

420 the Indian Ocean is through the narrow (18 km) and shallow (137 m) Strait of Bab al Mandab at
421 the southern end of the Red Sea. During periods of low sea level associated with glaciation, the
422 connection from the Indian Ocean through the strait is reduced, and the Red Sea experiences
423 extreme fluctuations in temperature and salinity (Bailey, 2009). During the last 400,000 years in
424 particular, the Red Sea has undergone at least two periods of hypersalinity (c. 19,000 and 30,000
425 years ago) that caused an aplanktonic environment in which larvae of many marine organisms
426 presumably could not survive (Siddal et al., 2003; DiBattista et al., 2016a). Coalescence analysis
427 dates the Red Sea lineage to 1.44 Ma (95% HPD = 0.51 – 2.53 Ma), which coupled with the
428 Saudi Arabian population expansion (65,000 – 130,000 years) indicates that *P. diacanthus* likely
429 survived the temperature and salinity crises that occurred during these periods, a conclusion that
430 is corroborated by other species (DiBattista et al., 2013). Our neutrality tests show no evidence
431 for changes in population size (Fu's Fs = -3.60, P = 0.035) providing evidence that refugia may
432 have existed in the Red Sea Province (possibly in the Gulf of Aden) to support a large stable
433 population of *P. diacanthus* despite the extreme environmental conditions.
434
435 *4.4 Biogeographic inferences in the Indian Ocean*
436 Christmas Island is located in the eastern Indian Ocean, a region (which includes Cocos-
437 Keeling Island) of secondary contact between Indian and Pacific species that diverged in
438 allopatry during Pleistocene glacial cycles (Gaither and Rocha, 2013). Indian and Pacific Ocean
439 phenotypes of *P. diacanthus* have both been recorded in the eastern Indian Ocean region, and
440 both Pacific and Indian Ocean mtDNA haplotypes are present at Christmas Island, indicating an
441 area of overlap (Hobbs and Allen, 2014 , Fig 3a). This region is recognized as a hybridization
442 hotspot (suture zone) with interbreeding documented between at least 27 reef fish species-pairs

443 from across eight families, and it has been suggested that Indian and Pacific *P. diacanthus*
444 lineages hybridize in this region (Hobbs and Allen, 2014). However, additional molecular work
445 will be needed to evaluate this hypothesis.
446 Genetic differences between Indian and Pacific Ocean populations are consistent with
447 Pleistocene closures of the Indo-Pacific Barrier. Despite being located in the Indian Ocean basin
448 and the presence of haplotypes that are associated with Indian Ocean lineages, our results
449 indicate that Christmas Island is genetically differentiated from other locations in the Indian
450 Ocean and instead has a stronger affiliation with the Pacific Ocean. A barrier to dispersal has
451 been previously proposed to exist west of the Cocos-Keeling Islands and east of the Chagos-
452 Laccadive ridge based on the presence of many Pacific species with distributions that extend no
453 further west than Christmas and the Cocos-Keeling Islands (Blum, 1989; Hodge and Bellwood,
454 2016).
455 Elsewhere in the Indian Ocean, the Maldives and Diego Garcia (Chagos Archipelago) are
456 genetically differentiated from the Pacific and Red Sea, but not from each other. Both
457 archipelagos are located in the central Indian Ocean, which is the western extent of the IPP,
458 although they also share faunal affinities with the Western Indian Ocean Province (Winterbottom
459 and Anderson, 1997; Gaither et al., 2010; Eble et al., 2011; Briggs and Bowen, 2012). The
460 grouping of Diego Garcia and the Maldives within the IPP is further evidenced by Pacific Ocean
461 mtDNA being found at Diego Garcia (Fig. 3a), which provides a signal that some degree of gene
462 flow occurs between the Indian and Pacific Ocean. Coalescence estimates of the two Indian
463 Ocean *P. diacanthus* lineages indicate they arose from an ancestor affiliated with the Red Sea.
464 The ability of *P. diacanthus* to persist throughout major geological and climatic shifts is
465 demonstrated by the age of expansion for all populations of *P. diacanthus* which predate the Last

Glacial Maximum, peaking at 26.5 – 19 ka (Clark et al., 2009) when global sea level 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 (*Dascyllus trimaculatus*) recruiting to reefs around Mo'orea were very close relatives, including full siblings, indicating that the larvae traveled and settled together despite a PLD of several weeks.

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

5. Conclusion

Pygoplites diacanthus is the first large angelfish to be surveyed across the Indo-Pacific. It appears to be highly dispersive, joining the ranks of smaller Pomacanthids such as the pygmy angelfish in showing little structure across ocean basins (Schultz et al., 2006; DiBattista et al., 2012). Pelagic larval duration tends to be shorter in the large angelfishes (~25 days in *Pygoplites* compared to 30 days or more in pygmy angelfishes; Thresher and Brothers, 1985), but this does not seem to restrict dispersal among the closely associated islands of the West and Central Pacific. However, this monotypic genus exhibits deep population genetic partitions between ocean basins. In every case, historical barriers existed at the junctions between observed populations, and in at least two cases (Red Sea and Mo'orea) oceanographic conditions may contribute to contemporary isolation. On the genetic continuum between isolated populations and evolutionary distinctions (Wright, 1978; Frankham et al., 2002), the deep divergences between oceans indicate that the monotypic *Pygoplites* may be on the pathway to three emerging species. The genetic and morphological divergences are certainly sufficient to recognize subspecific evolutionary (and taxonomic) partitions.

Acknowledgements

This project was supported by the Seaver Institute (to BWB), the National Science Foundation (NSF) OCE-0929031 (BWB), NOAA National Marine Sanctuaries Program MOA grant No. 2005-008/66882 (R.J. Toonen), King Abdullah University of Science and Technology (KAUST) Office of Competitive Research Funds under Award no. CRG-1-2012-BER-002 and baseline research funds to MLB, as well as National Geographic Society Grant 9024-11 to JDD. RRC was supported by NSF grant DGE-1329626 and the Dr. Nancy Foster Scholarship program under Award no. NA15NOS4290067. This paper is funded in part by a grant/cooperative agreement

535 from the National Oceanic and Atmospheric Administration, Project R/CR-14, which is
536 sponsored by the University of Hawaii Sea Grant College Program, SOEST, under Institutional
537 Grant No. NA05OAR4171048 (to BWB) from NOAA Office of Sea Grant, Department of
538 Commerce. Fieldwork at Christmas Island was supported by National Geographic Grant 8208-07
539 (M.T. Craig). We thank Eric Mason and the crew at Dream Divers in Saudi Arabia, Nicolas
540 Prévot at Dolphin Divers and the crew of the M/V Deli in Djibouti, the KAUST Coastal and
541 Marine Resources Core Lab, the KAUST Reef Ecology Lab, Amr Gusti, the Administration of
542 the British Indian Ocean Territory and Chagos Conservation Trust, as well as the University of
543 Milano-Bicocca Marine Research and High Education Centre in Magoodhoo, the Ministry of
544 Fisheries and Agriculture, Republic of Maldives, and the community of Maghoodhoo, Faafu
545 Atoll. For assistance in collection efforts we thank Alfonso Alexander, Senifa Annandale,
546 Howard Choat, Pat Colin, Lori Colin, Joshua Copus, Matthew Craig, Michelle Gaither, Brian
547 Greene, Jean-Paul Hobbs, Garrett Johnson, Stephen Karl, Randall Kosaki, Cassie Lyons, David
548 Pence, Mark Priest, Joshua Reece, D. Ross Robertson, Tane Sinclair-Taylor, Robert Thorne, and
549 Robert Whitton. We thank the HIMB EPSCoR core facility and the University of Hawai‘i’s
550 Advanced Studies in Genomics, Proteomics, and Bioinformatics facility for their assistance with
551 DNA sequencing. We also thank Michelle Gaither, Brent Snelgrove, Robert Toonen, and
552 members of the ToBo lab for their assistance, logistical support, and feedback throughout this
553 project. Thanks to Tane Sinclair-Taylor for the graphical abstract images. Thanks to editor
554 Giacomo Bernardi and one anonymous reviewer for critique and suggestions that improved the
555 paper. The views expressed herein are those of the authors and do not necessarily reflect the
556 views of NOAA or any of its subagencies. This is contribution #1653 from the Hawai‘i Institute
557 of Marine Biology, #9592 from the School of Ocean and Earth Science and Technology, and

#JC-06-03 from University of Hawai‘i Sea Grant Program.

559

560 **References**

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

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

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

569 Andrews, K.R., Williams, A., Fernandez-Silva, I., Newman, S.J., Copus, J.M., Bowen, B.W.,
570 2016. Phylogeny of deepwater snappers (Genus *Etelis*) reveals a cryptic species pair in the Indo-
571 Pacific and Pleistocene invasion of the Atlantic. Molecular Phylogenetics & Evolution
572 *(Accepted)*.

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

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

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

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

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

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

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

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

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

593 Bowen, B.W., Bass, A.L., Rocha, L.A., Grant, W.S., Robertson, D.R., 2001. Phylogeography of
594 the trumpetfishes (*Aulostomus*): Ring species complex on a global scale. Evolution 55, 1029-
595 1039.

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

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

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

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

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

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

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

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

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

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

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

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

627 Drew, J., Allen, G.R., Kaufman, L.E.S., Barber, P.H., 2008. Endemism and regional color and
628 genetic differences in five putatively cosmopolitan reef fishes. Conservation Biology 22, 965-
629 975.

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

633 Eble, J.A., Rocha, L.A., Craig, M.T., Bowen, B.W., 2011. Not all larvae stay close to home:
634 Insights into marine population connectivity with a focus on the brown surgeonfish (*Acanthurus*
635 *nigrofuscus*). Journal of Marine Biology 2011, 12.

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

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

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

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

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

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

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

655 Gaither, M.R., Rocha, L.A., 2013. Origins of species richness in the Indo-Malay-Philippine
656 biodiversity hotspot: evidence for the centre of overlap hypothesis. Journal of Biogeography 40,
657 1638-1648.

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

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

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

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

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

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

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

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

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

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

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

683 Kulbicki, M., Parravicini, V., Bellwood, D.R., Arias-Gonzàlez, E., Chabanet, P., Floeter, S.R.,
684 Friedlander, A., McPherson, J., Myers, R.E., Vigliola, L., Mouillot, D., 2013. Global
685 biogeography of reef fishes: a hierarchical quantitative delineation of regions. PLOS One
686 8:e81847.

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

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

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

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

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

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

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

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

709 McMillan, W.O., Weigt, L.A., Palumbi, S.R., 1999. Color pattern evolution, assortative mating,
710 and genetic differentiation in brightly colored butterflyfishes (Chaetodontidae). Evolution, 247-
711 260.

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

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

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

717 Planes, S., 1993. Genetic differentiation in relation to restricted larval dispersal of the convict
718 surgeonfish, *Acanthurus triostegus,* in French Polynesia. Marine Ecology Progress Series 98,
719 237-246.

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

725 Randall, J.E., 1994. Twenty-two new records of fishes from the Red Sea. Fauna of Saudi Arabia
726 14, 259-275.

727 Randall, J.E., 1998. Zoogeography of shore fishes of the Indo-Pacific region. Zoological Studies
728 37, 227-268.

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

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

733 Reece, J.S., Bowen, B.W., Smith, D.G., Larson, A.F., 2010b. Molecular phylogenetics of moray
734 eels (Murenidae) demonstrates multiple origins of a shell-crushing jaw (Gymnomuraena,
735 Echidna) and multiple colonizations of the Atlantic Ocean. Molecular Phylogenetics & Evolution
736 57, 829 – 835.

737 Robertson, D.R., Karg, F., Leao de Moura, R., Victor, B.C., Bernardi, G., 2006. Mechanisms of
738 speciation and faunal enrichment in Atlantic parrotfishes. Molecular Phylogenetics & Evolution
739 40, 795-807.

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

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

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

748 Ronquist, F., 2004. Bayesian inference of character evolution. Trends in Ecology & Evolution
749 19, 475-481.

750 Schultz, J.K., Pyle, R.L., DeMartini, E., Bowen, B.W., 2006. Genetic connectivity among color
751 morphs and Pacific archipelagos for the flame angelfish, *Centropyge loriculus*. Marine Biology
752 151, 167-175.

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

755 Song, C.B., Near, T.J., Page, L.M., 1998. Phylogenetic relations among percid fishes as inferred
756 from mitochondrial cytochrome *b* DNA sequence data. Molecular Phylogenetics & Evolution 10,
757 343-353.

758 Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype
759 reconstruction from population genotype data. American Journal of Human Genetics 73, 1162-
760 1169.

761 Szabo, Z., Snelgrove, B., Craig, M.T., Rocha, L.A., Bowen, B.W., 2014. Phylogeography of the
762 Manybar Goatfish, *Parupeneus multifasciatus* reveals moderate structure between the Central
763 and North Pacific and a cryptic endemic species in the Marquesas. Bulletin of Marine Science
764 90, 493-512.

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

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

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

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

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

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

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

782 Wright, S., 1978. Evolution and the Genetics of Populations: A Treatise in Four Volumes: Vol.
783 4: Variability Within and Among Natural Populations. University of Chicago Press, Chicago.

784 Wyrtki, K., Kilonsky, B., 1984. Mean water and current structure during the Hawaii-to-Tahiti
785 shuttle experiment. Journal of Physical Oceanography 14, 242-254.
786

787

788 **Figure Captions**

789

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

795

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

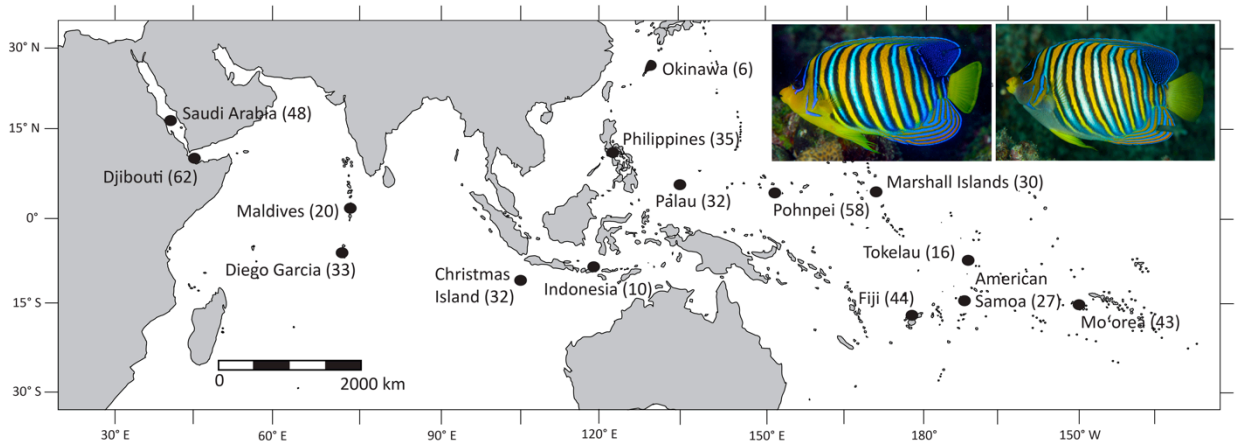
804

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

nucleotide change, and color represents collection location (see key). All singleton alleles (N = 22) were removed from the S7 analysis to minimize circularity between closely related alleles.

812

813 Figure 1.



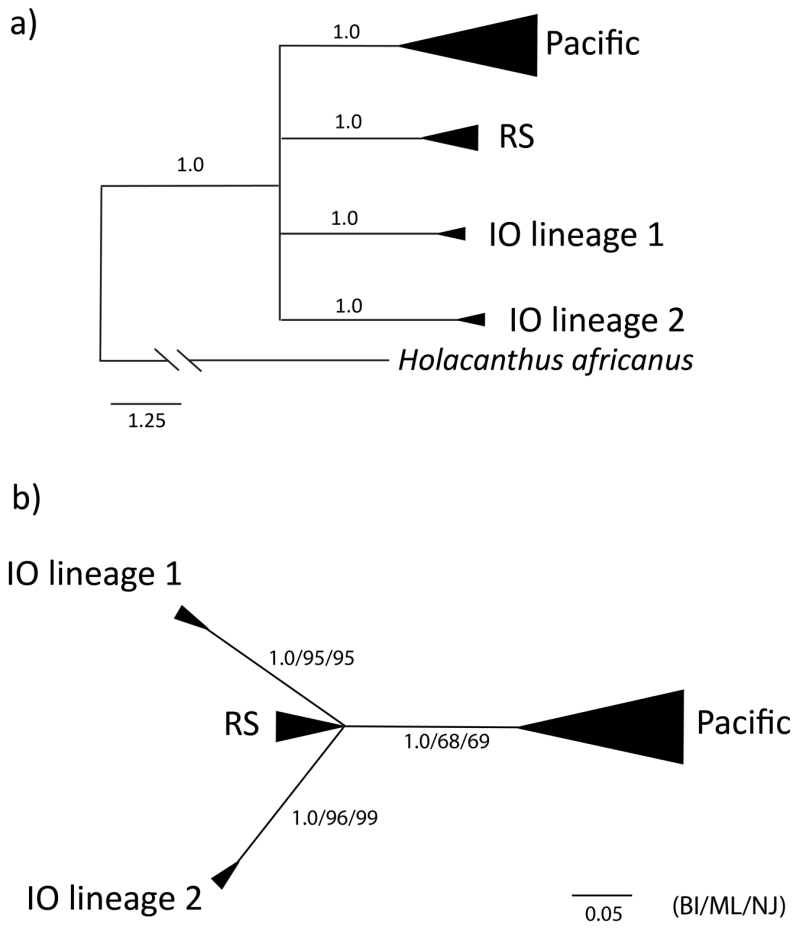
814

815

816

818 Figure 2.

819



820

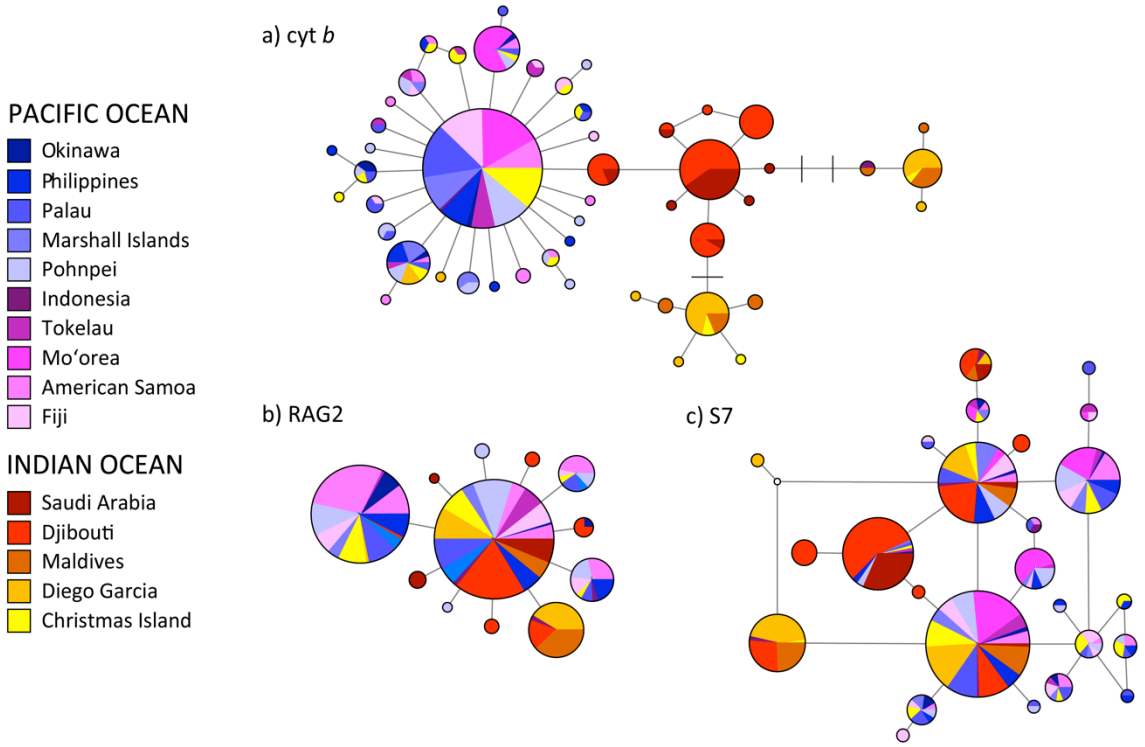
821

822

823
824

825 Figure 3.

826



827

Tables

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

828

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* ± SD	π ± 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

829

830

831

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

35
36

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.

37

38

839

840

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

841
842

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	--

843

844

Graphical Abstract