1   For the virtual issue, "Red Sea and Western Indian Ocean Biogeography"

2   Original Article

3   LRH: M. A. Priest *et al.*

4   RRH: Barrier to larval dispersal in a range-limited grouper

5

6   **A bridge too far: dispersal barriers and cryptic speciation in an Arabian Peninsula**

7   **grouper (*Cephalopholis hemistiktos*)**

8

9   Mark A. Priest[1, 2*], Joseph D. DiBattista[2,3], Jennifer L. McIlwain[3], Brett M. Taylor[4], Nigel E.

10   Hussey[5] and Michael L. Berumen[2]

11

12   *[1]Marine Spatial Ecology Lab, School of Biological Sciences, University of Queensland, St*

13   *Lucia, QLD 4072, Australia, [2]Red Sea Research Center, Division of Biological and*

14   *Environmental Science and Engineering, King Abdullah University of Science and*

15   *Technology, Thuwal 23955-6900, Saudi Arabia, [3]Department of Environment and*

16   *Agriculture, Curtin University, GPO Box U1987, Perth, WA, Australia, [4]James Cook*

17   *University, College of Marine and Environmental Science, Townsville, QLD 4811, Australia,*

18   *[5]Great Lakes Institute for Environmental Research, University of Windsor, 401 Sunset*

19   *Avenue, Windsor, ON N9B 3P4, Canada*

20

21   *Correspondence: Mark A. Priest, Marine Spatial Ecology Lab, University of Queensland, St.

22   Lucia, QLD 4072, Australia

23

24   Email: m.priest@uq.edu.au

25

26   Word count: 6,674

27

28

29

30   **ABSTRACT**

31   **Aim** We use genetic and age-based analyses to assess the evidence for a biogeographical

32   barrier to larval dispersal in the yellowfin hind, *Cephalopholis hemistiktos,* a commercially

33   important species found across the Arabian Peninsula.

34

35   **Location** Red Sea, Gulf of Aden, Gulf of Oman and Arabian Gulf.

36

37   **Methods** Mitochondrial DNA (COI) and nuclear DNA (S7) sequences were obtained for *C.*

38   *hemistiktos* sampled throughout its distributional range. Phylogeographical and population-

39   level analyses were used to assess patterns of genetic structure and to identify barriers to

40   dispersal. Age-based demographic analyses using otoliths were also used to determine

41   differences in growth and longevity between regions.

42

43   **Results** Our analyses revealed significant genetic structure congruent with growth parameter

44   differences observed across sampling sites, suggesting cryptic speciation between populations

45   in the Red Sea and Gulf of Aden versus the Gulf of Oman and Arabian Gulf.

46

47   **Main Conclusion** Our results indicate historical disruption to gene flow and a contemporary

48   dispersal barrier in the Arabian Sea, which *C. hemistiktos* larvae are unable to effectively

49   traverse. This provides yet another example of a (cryptic) species with high dispersive

50   potential whose range is delimited by a lack of suitable habitat between locations or an

51   inability to recruit at the range edge.

52

53   **Keywords**

54   **Serranidae, coral reef fish, population genetics, demography, mitochondrial DNA,**

55   **nuclear DNA, phylogeography**

56

57   **INTRODUCTION**

58   For many coral reef fishes, the pelagic larval phase presents the only opportunity for dispersal

59   amongst fragmented and spatially patchy adult habitats (Kritzer & Sale, 2004). On a

60   demographic level, the process of dispersal determines the persistence and structure of

61   populations (Hixon *et al.*, 2002; Burgess *et al.*, 2014). This is especially important when

62   considering harvested species, which may rely on external replenishment if localized spawner

63   biomass is depleted due to overfishing. Thus, understanding patterns of connectivity can help

64   inform resource managers to preserve fishery resources and make conservation efforts more

65   effective. For example, the success of spatial management techniques such as marine reserve

66   networks relies on successful self-seeding and larval export (demographic connectivity) to

67   adjacent fished areas (Botsford *et al.*, 2003; Gaines *et al.*, 2010). Conversely, the absence of

68   connectivity between locations will define management units and result in isolated

69   populations that over time generate distinct genetic signatures (Moritz, 1994), and ultimately

70   can lead to speciation.

71

72   Genetic techniques are now routinely used to measure connectivity on multiple

73   temporal and spatial scales. Recent technological and methodological advances allow larvae

74   or recruits to be assigned back to their parents, or population of origin (e.g. Jones *et al.*, 2005;

75 Harrison *et al.*, 2012; Almany *et al.*, 2013). This process, however, is challenging as it

76 involves intensive sampling of both adults and recruits, and large numbers of genetic markers.

77 Traditionally, population genetics has been used to infer levels of connectivity and estimate

78 the number of migrants exchanged between locations, but these estimates require a number of

79 simplifying assumptions (Whitlock & McCauley, 1999) and may reflect historical rather than

80 contemporary dispersal (Hedgecock *et al.*, 2007). Thus, the absence of genetic differentiation

81 among sites does not necessarily confirm effective connectivity on demographic scales, but

82 the presence of genetic structure does suggest some level of prolonged isolation (Hellberg,

83 2007; Hellberg, 2009).

84   Limits to dispersal, and thus gene flow, in the marine environment are not always

85 reflective of larval dispersal ability (Weersing & Toonen, 2009; but see Riginos & Victor,

86 2001) and may be driven by numerous factors such as oceanographic patterns (White *et al.*,

87 2010; Simpson *et al.*, 2014), larval behaviour (Leis, 2002), ecological requirements (Rocha *et

88 al.*, 2002) and biogeographic barriers (Rocha *et al.,* 2007; Briggs & Bowen, 2012). One area

89 that may provide conditions limiting for dispersal are the seas surrounding the Arabian

90 Peninsula (Fig. 1). This area is at the intersect of several biogeographic provinces (Schils &

91 Coppejans, 2003), and contains a diverse range of habitats experiencing a wide range of

92 environmental conditions. During Pleistocene glacial cycles, the Red Sea experienced

93 repeated periods of sea level fluctuation, causing major deviations in temperature and salinity,

94 as well as putative faunal extinctions (Siddall *et al.*, 2003; DiBattista *et al.,* in reviewA). Such

95 events are thought to have driven the high levels of diversity and endemism recorded in the

96 Red Sea (14% in fishes; Randall, 1994; also see DiBattista *et al.,* in reviewB). Whilst the

97 shallow and narrow opening to the Gulf of Aden has historically restricted oceanic water and

98 propagule exchange, an ecological barrier of turbid water in the southern Red Sea has been

99 described as a contemporary constraint to dispersal (Ormond & Edwards, 1987; Roberts *et al.*,

100   1992). Outside the Gulf of Aden, the southern coast of Oman bordering the Arabian Sea is

101   subjected to seasonal cold-water upwelling events generated by southwest monsoons,

102   providing an additional isolating force (Savidge *et al.*, 1990). This results in large annual

103   fluctuations in surface water temperatures (16 – 28 °C) and impedes the development of coral

104   reefs, creating large tracts of coastline devoid of suitable settlement habitat for reef fish

105   larvae. Extensive coral reefs return in northern Oman and extend into the Gulf of Oman,

106   where the effects of upwelling events are less frequent (Coles & Wilson, 2001). The Arabian

107   Gulf (also known as the Persian Gulf), whilst containing somewhat similar benthic

108   communities to the Gulf of Oman (Feary *et al.*, 2010), has vastly different environmental

109   parameters. The Arabian Gulf is shallow and has limited water exchange with the Gulf of

110   Oman. Consequently, sea surface temperatures vary dramatically (12 – 36 °C) with

111   consistently high seasonal maximums along with elevated salinity (Sheppard *et al.*, 1992).

112   To investigate how these diverse bio-physical environments promote or impede larval

113   dispersal, we investigated the connectivity and demography of the yellowfin hind,

114   *Cephalopholis hemistiktos* (Serranindae; Rüppell 1830), a commercially important fishery

115   species (Hashim, 1993; Gladstone, 2003) restricted to the Arabian Peninsula region.

116   *Cephalopholis hemistiktos* possesses a disjunctive distribution and is absent from the western

117   Arabian Sea (southern Oman; Fig. 1) (Craig *et al.*, 2011). Additionally, previous reports have

118   suggested that these two populations differ morphologically; in pectoral fin size and ray

119   count, oblique scale rows and asymptotic size (Randall & Ben-Tuvia, 1983; see Appendix S1

120   in Supporting Information); these differences suggests isolation on evolutionary timescales.

121   Given the spatial distribution and physical differences between locations, we conducted both

122   genetic and age-based demographic assessments of *C. hemistiktos* by sampling throughout the

123   entire species range. We used one mitochondrial and one nuclear genetic marker to evaluate

124 gene flow, dispersal barriers and connectivity among populations, as well as otolith-based age

125 estimates to explore divergence in life-history characteristics.

126

127 **MATERIALS AND METHODS**

128

129 **Sampling**

130

131 *Cephalopholis hemistiktos* were collected from ten sites encompassing the majority of the

132 range of this species (Fig. 1 and Table 1). Individuals were sampled between 2005 and 2014

133 using hand spears whilst snorkeling or SCUBA diving (Gulf of Aqaba, Red Sea, Gulf of Aden

134 and Arabian Sea), or from local fish markets (Gulf of Oman and Arabian Gulf). Efforts were

135 made to sample individuals representative of the whole size range available at each site.

136 Measurements of total length (TL) were taken to the nearest millimeter. Sagittal otoliths were

137 extracted, cleaned in ethanol and stored dry until sectioning. Tissue samples for genetic

138 analysis were stored in 70% ethanol or in a saturated salt-DMSO buffer (Seutin *et al.,* 1991)

139 and stored at room temperature. As samples were collected over many years for multiple

140 research projects, we did not obtain both genetic and otolith samples from every site. Thus,

141 some sites have only genetic or otolith data (see Fig. 1 and Table 1 for details).

142

143 **Mitochondrial DNA analysis**

144

145 Genomic DNA was extracted using the 'HotSHOT' protocol (Meeker *et al.,* 2007) and stored

146 at -20°C. Samples were sequenced for a 629 base pair (bp) region of the mtDNA cytochrome-

147 c oxidase subunit-I (COI) gene using the primers FishF2 and FishR2 (Ward *et al.,* 2005).

148 Polymerase chain reactions (PCRs) were carried out in 12 µl reaction volumes using Qiagen

149 Multiplex PCR kits (Qiagen, Valencia, CA, USA) according to the manufacturer's

150 instructions. Each reaction contained 1 μl gDNA and 0.5 μl of each forward and reverse

151 primers (10 μM). PCR cycling conditions consisted of an initial denaturation step at 95°C for

152 15 min, followed by 35 cycles of 94°C denaturing for 30 sec, 50°C annealing for 60 sec and

153 72°C extension for 60 sec, with a final 72°C extension for 10 min. PCR products were

154 purified using exonuclease I and shrimp alkaline phosphatase (ExoSAP; USB, Cleveland,

155 OH, USA), sequenced in the forward direction with fluorescently labeled dye terminators

156 (BigDye version 3.1, Applied Biosystems, Foster City, CA, USA) and analysed using an ABI

157 3130xl Genetic Analyzer (Applied Biosystems). All haplotypes were deposited in GenBank

158 (Accession numbers: XXX-XXX). Sequences were aligned, trimmed and edited using

159 Geneious Pro *vers.* 7.0.6 (Drummond *et al.,* 2009). The GTR + I model for nucleotide

160 substitution was selected using JMODELTEST 2.0 *vers.* 0.1.1 (Guindon & Gascuel, 2003;

161 Darriba *et al.*, 2012) based on the Akaike Information Criterion test.

162 ARLEQUIN *vers.* 3.5 (Excoffier & Lischer, 2010) was used to calculate haplotype (h)

163 and nucleotide diversity (π), and to test for range-wide patterns of population structure for *C.*

164 *hemistiktos*. Global Φ_{ST} was estimated using analysis of molecular variance (AMOVA,

165 Excoffier *et al.,* 1992), with deviations from null distributions tested using nonparametric

166 permutations ($N = 9,999$). AMOVA was also used to test for significant genetic partitioning

167 among regions. Additionally, pairwise Φ_{ST} statistics were generated to identify sites

168 associated with genetic subdivision within regions.

169 To assess deviation from neutrality, Fu's F_S (Fu, 1997) was calculated for all sites

170 individually, all sites grouped and also sites grouped into the two regions (Red Sea and Gulf

171 of Aden [RS-GA]; Gulf of Oman and Arabian Gulf [GO-AG]) using ARLEQUIN with

172 significance assessed with 9,999 permutations. Historical demography was explored using

173 mismatch analyses for all sites together and for the two separate regions. Populations

174   experiencing recent or rapid expansion exhibit unimodal mismatch distributions and non-

175   significant raggedness indices (Harpending, 1994). The parameter τ was fitted to the two

176   regions and population ages were estimated using the equation $\tau = 2\mu t$ where t = population

177   age in generations and μ = mutation rate per generation for the sequence (Rogers &

178   Harpending, 1992). We used a mutation rate previously recorded for coral reef fish of 1.55 %

179   per million years (Myr) (Lessios, 2008). Generation time was estimated from the equation $T =$

180     $(\alpha + \omega)/2$ (Pianka, 1978), where α = age at first reproduction and ω = the age at last

181   reproduction. We therefore obtained generation times of 13 years based on our existing age-

182   based demographic and reproductive data. Given the approximations involved, we interpret

183   these coalescence estimates with caution.

184   Additional estimates of historical demographics were reconstructed using Bayesian

185   skyline plots (BSP) in BEAST *vers.* 1.8.2 (Drummond *et al.*, 2005; Drummond *et al.*, 2012)

186   and visualised using Tracer *vers.* 1.6 (Rambaut *et al.,* 2014). Markov Chain Monte Carlo

187   sampling was run for 20,000,000 iterations sampling every 20,000 iterations using the GTR +

188   I substitution model with a strict clock and a nucleotide substitution rate of 1.55 %. Due to the

189   large number of parameters estimated under the GTR + I substitution model, we specified

190   lognormal priors on substitution rates to achieve acceptable effective sample sizes. Each

191   region was analysed separately for five independent model runs after standardising sample

192   sizes between regions, with all runs being combined for the final BSP analyses.

193

194   Evolutionary relationships among COI haplotypes were constructed using a median joining

195   algorithm and default settings (as per Bandelt *et al.*, 1999) with the program NETWORK

196   *vers.* 4.5.1.0 (www.Fluxus-engineering.com/network_terms.htm). The time of divergence

197   between regions was estimated in ARLEQUIN using the corrected pairwise sequence distance

198   between regions minus the corrected pairwise sequence distance within regions (d_C; Tamura

199 & Nei, 1993). Divergence times were also estimated using a Bayesians approach in BEAST.

200 The time to most recent common ancestor (TMRCA), for both regions and all samples

201 combined, was estimated using the same model parameters used in the BSP analyses.

202

203 **Nuclear gene analysis**

204 Samples (Table 1) were sequenced for a 206 bp region of the first intron of the S7 ribosomal

205 protein gene using the primers S7RPEX2F and S7RPEX2R (Chow & Hazama 1998) to

206 provide an independent assessment of evolutionary relationships among lineages. PCR

207 cycling conditions consisted of a 95°C denaturation step for 15 min, followed by 35 cycles of

208 94°C denaturing for 30 sec, 50°C annealing for 30 sec, 72°C extension for 90 sec and a final

209 72°C extension for 10 min. PCR products were cleaned and sequenced using the same

210 protocols as outlined in the mtDNA analyses and all sequences were deposited in GenBank

211 (Accession numbers: XXX-XXX). Sequences were aligned, trimmed and edited using

212 Geneious Pro. Allelic states of nuclear sequences with more than one heterozygous site were

213 estimated with the Bayesian program PHASE *vers.* 2.1 (Stephens & Donnelly, 2003) as

214 implemented in the software DnaSP *vers.* 5.0 (Librado & Rozas, 2009). We conducted three

215 independent runs in PHASE with a burn-in of 10,000, and 100,000 iterations. All runs

216 returned consistent allele identities and PHASE was able to determine all alleles with > 96%

217 probability. Differences in allele frequencies between regions and sites were assessed using

218 AMOVA in ARLEQUIN following the mtDNA analyses. Evolutionary networks among

219 alleles were constructed using a median joining algorithm and default settings with the

220 program NETWORK.

221

222 **Age-based demographic analysis**

223 Individual fish ages were determined by otolith analysis. Otoliths were affixed by

224 thermoplastic glue (Crystalbond 509; Aremco, Valley Cottage, NY, USA) to a clear glass

225 slide and ground along the longitudinal axis to the core. The ground otolith was then removed

226 and fixed to a clean slide with the flat surface down, and then polished until a thin transverse

227 section (~150 μm) was obtained. Age was determined by counting alternating opaque and

228 translucent zones along a consistent axis on the sectioned otolith face. Blind readings were

229 made on three separate occasions, one to two weeks apart, and the final age (years) of an

230 individual was determined when two or more counts agreed. When agreement was not

231 achieved after three counts, the otolith was excluded from further analysis. Site specific

232 growth patterns were estimated using the von Bertalanffy growth function (VBGF),

233 represented by $L_t = L_\infty \left(1 - e^{-K(t-t_0)}\right)$, where L_t is the predicted TL (mm) at age t (years), L_∞

234 is the mean asymptotic LF, K is the coefficient used to describe the curvature of fish growth

235 towards L_∞, t represents age (years) and t_0 is the theoretical age at which TL is equal to zero,

236 as described by K. To enhance precision among sites and account for early growth trajectories

237 (Kritzer *et al.,* 2001; Berumen, 2005), models were constrained to a common TL at settlement

238 (50 mm). Growth parameters were compared among sites using bivariate 95% confidence

239 ellipses surrounding the K and L_∞ estimates (Kimura 1980). Plots of size and age

240 frequencies were compared among regions. Mean length, age and length at ages one through

241 five were compared using Student's t-tests with Welch's correction when unequal variance

242 was observed.

243

244 **RESULTS**

245

246 **Mitochondrial DNA analysis**

247 COI sequence data revealed 39 haplotypes with haplotype diversity ranging from 0.68 to 0.81

248 and nucleotide diversity ranging from 0.0020 to 0.0035 (Table 1). After grouping sites into

regions, AMOVA revealed that the majority of the genetic variability was explained by significant differentiation between these two regions (82 %, $\Phi_{CT} = 0.827$, $P < 0.001$; Table 2), with no shared haplotypes. The variance explained by the among-sites within-regions variance component was considerably smaller (0.1 %, $\Phi_{SC} = 0.005$, $P = 0.224$) than that between regions and not significant. Pairwise Φ_{ST} comparisons confirmed the regional patterns described above. There was no significant differentiation among sites within regions (pairwise Φ_{ST} ranged from 0.000 to 0.046); however, comparisons among sites from different regions were all statistically significant (pairwise Φ_{ST} ranged from 0.796 to 0.851).

We found limited evidence of population expansion or selective sweeps when considering sites individually. Despite negative values in seven of eight sites for Fu's F_S, only two values were statistically significant (Table 1). However, when sites were pooled by region, F_S values for both regions were significantly negative (RS-GA, $F_S = -11.71$, $P < 0.001$; GO-AG, $F_S = -6.47$, $P = 0.007$). Pairwise mismatch distributions for the two regions were bimodal but did not show significant deviation from the sudden population expansion model (Fig. 2a and b; RS-GA, Harpending's raggedness index, $r = 0.14$, $P = 0.07$; GO-AG, $r = 0.06$, $P = 0.56$), suggesting that both regions have undergone recent population expansion. Time since most recent population expansion were different between the two regions; we estimated the RS-GA population ($\tau = 2.502$) at 128,314 years, and the GO-AG population ($\tau = 1.453$) at 74,516 years. BSP provided more recent estimates of population expansion (Fig. 2c and d), suggesting both regions experienced sudden population expansion approximately 18,000 years ago, with the GO-AG region expanding at a greater rate than the RS-GA region. However, given the overlapping confidence interval between the start and end of population reconstructions for both regions, it appears the overall magnitude of recent population expansion may be modest.

273   The median joining mtDNA haplotype network supports the genetic partitioning of *C.*

274   *hemistiktos* into two distinct lineages separated by eight mutational steps, which correspond to

275   the two geographically distinct regions (Fig. 3a). Additionally, the only previously reported

276   COI sequence for *C. hemistiktos* from the Arabian Gulf (*i.e.,* Iran, Genbank accession

277   number: HQ149822; Asgharian *et al.*, 2011) grouped within the GO-AG region haplotypes.

278   Average corrected sequence divergence between these two lineages was large ($d_C = 0.0132$)

279   compared to genetic divergence within lineages (RS-GA, $d_C = 0.0027$; GO-AG, $d_C = 0.0028$),

280   suggesting a long period of isolation between them (~ 852,000 years based on a mutation rate

281   of 1.55 % per Myr; Lessios, 2008). Bayesian analyses estimated the TMRCA as 505,700

282   years (95% highest posterior density [HPD] = 249,100-787,200 years) when considering both

283   regions. Individually, TMRCA for the RS-GA region was 206,200 (95% HPD = 85,500-

284   348,200 years), and 284,600 (95% HPD = 109,100-489,200 years) for the GO-AG region.

285

286   **Nuclear gene analysis**

287   Nuclear sequences for 261 individuals at the S7 locus returned seven variable sites with nine

288   alleles. Median joining networks revealed only one shared allele between regions (Fig. 3b).

289   AMOVA determined that allele frequencies differed significantly among regions ($F_{CT} =$

290   0.287, $P = 0.03$) and among-sites within-regions ($F_{SC} = 0.287$, $P = 0.03$), suggesting that this

291   marker may be in the process of segregating by location, but at a slower rate than the mtDNA

292   COI marker. Pairwise F_{ST} comparisons were congruent with COI data. Contrasts among sites

293   from different regions were all statistically significant (pairwise F_{ST} ranged from 0.23 to

294   0.43), whereas comparisons within regions were all non-significant (pairwise F_{ST} ranged from

295   -0.02 to 0.01) with the exception of comparisons between Muscat in the Gulf of Oman and

296   Jubail in the Arabian Gulf ($F_{ST} = 0.23$).

297

298 **Demographic analysis**

299 In agreement with the observations of Randall & Ben-Tuvia (1983), the maximum lengths of

300 fish sampled from the three sites in the GO-AG region were all greater than the maximum

301 lengths of fish sampled from the RS-GA region (Table 1). Mean lengths were also

302 significantly greater in the GO-AG region (163 *vs.* 282 mm; $t = 41.77$, $P < 0.001$; Fig. 4a), but

303 mean age was not different between regions (6.1 *vs.* 6.2 years; $t = 0.47$, $P = 0.64$; Fig. 4b).

304 Comparisons of mean length at age one through five (at which point fish from the RS-GA

305 region were approaching asymptotic size) revealed that fish from the GO-AG region were

306 consistently attaining a larger size for a given age (Table 3). Subsequently, plots of bivariate

307 confidence ellipses surrounding growth parameters K and L_∞ confirmed that the fish from the

308 GO-AG region achieved greater asymptotic size but had similar growth coefficients (Fig. 5).

309 These plots also revealed significant differences in asymptotic size within regions. Muscat

310 and Musandam fish exhibited similar growth parameters, possessing much larger L_∞

311 estimates than at Abu Dhabi. However, these values were all greater than estimates for sites

312 within the RS-GA region, which showed markedly reduced L_∞ estimates for sites within the

313 central Red Sea (Al Lith and Thuwal).

314

315 **DISCUSSION**

316 Our results indicate a clear and pervasive barrier to dispersal in *C. hemistiktos* between RS-

317 GA and GO-AG regions. Moreover, the lack of shared haplotypes between locations at the

318 COI marker, and allele frequency shift for the slower segregating nuclear marker, shows that

319 these two regions have been isolated for considerable amounts of time. This historical and

320 contemporary separation has several consequences for both this species' classification and

321 management.

322   The location of the genetic break, on Southern Oman's Arabian Sea coastline,

323   corresponds with an environmentally dynamic area characterized by seasonal upwelling

324   (Savidge *et al.*, 1990). This results in highly variable seawater surface temperatures, which

325   severely restrict coral reef development along a 400 km stretch of coastline (Schils &

326   Coppejans, 2003). For *C. hemistiktos*, either traversing this distance as larvae, or subsequent

327   settlement and recruitment to establish a viable reproductive population, appears impossible.

328   Recent studies investigating individual larval trajectories have shown that larval dispersal may

329   be more restricted than first thought (*e.g.,* Jones *et al*., 2005; Almany *et al.*, 2007). Moreover,

330   estimates of dispersal in another grouper species demonstrate that over 50% of larvae are

331   retained within 14 km of the spawning site (Almany *et al*., 2013). However, we found no

332   evidence of genetic differentiation between the east and west coasts of the Red Sea, despite

333   the presence of a deep trench (> 2000 m) between coasts, indicating successful dispersal

334   occurs across this ~200 km distance (also see Fernandez-Silva *et al.,* in review). The absence

335   of connectivity between regions may be due, in part, to the reproductive timing recorded for

336   *C. hemistiktos.* In the Gulf of Aqaba spawning occurs in July and August (El-Etreby *et al.,*

337   2003) and in Oman between May and July (see Appendix S2 in Supporting Information). This

338   timing corresponds with the seasonal upwelling on the Arabian Sea coastline of Oman, when

339   sea surface temperatures can drop to 16 °C, which is well below the physiological tolerance of

340   most tropical fish larvae (McCormick & Molony, 1995; Sponaugle *et al.,* 2006). This large

341   environmental fluctuation may provide an explanatory mechanism for the continuing barrier

342   to larval dispersal in this species. Contrasting results are found when considering patterns of

343   connectivity in other coral reef fish species across the Arabian Peninsula, even within the

344   *Cephalopholis* genus. For example, *Cephalopholis argus* shares haplotypes among samples

345   from the Gulf of Oman and Red Sea (DiBattista *et al.*, 2013), indicating some level of

346   successful dispersal across the Arabian Sea, but then this species is found throughout the

347 Indo-Pacific. Furthermore, the Oman clownfish, *Amphiprion omanensis*, have larvae that are

348 capable of traversing the same distance (Simpson *et al.*, 2014) despite the reduced pelagic

349 larval duration associated with this genus (~10 days; Thresher *et al.,* 1989). Given that

350 variability in connectivity patterns can be prevalent in closely related species with similar life-

351 history characteristics (Gaither *et al.,* 2010; DiBattista *et al.,* 2012), it seems likely that the

352 existing patterns reported here may have originated via species-specific responses to historical

353 regional disturbances (DiBattista *et al.*, 2013), but are maintained by the present upwelling

354 barrier to dispersal in the Arabian Sea. Demographic reconstructions suggest that both regions

355 underwent population expansion at the end of the last glacial maximum (15 to 20 kya), whilst

356 our estimates of lineage divergence predate this period, indicating that both Red Sea and

357 Arabian Gulf populations have survived through multiple palaeoclimatic challenges after

358 initial separation. It is likely that sea-level fluctuations as a result of Pleistocene glacial cycles

359 caused multiple periods of isolation between the two regions, and after one of these events the

360 isolated populations failed to reconnect. Whilst our coalescence estimates are necessarily

361 approximate, our Bayesian estimation of lineage divergence coincides with the largest

362 Pleistocene glacial period (*c.* 450,000 years ago), during which the Red Sea endured major

363 sea level and salinity disturbance (Siddall *et al.*, 2003).Further work in this understudied

364 region is necessary to fully understand the role of such peripheral habitats in a global

365 biogeographic context (Berumen *et al.,* 2013; Bowen *et al.,* 2013).

366

367 **Management considerations**

368

369 The presence of genetic structure between RS-GA and GO-AG regions clearly shows the

370 spatial limit to dispersal in *C. hemistiktos,* as only a small number of migrants per generation

371 are needed to homogenize populations (Slatkin, 1993). Thus, the potential for external

372 recolonisation between regions after a localised extinction event is low, and implies each

373 region should be treated as a separate management unit. Within regions, the lack of genetic

374 structure suggests recent or ongoing gene flow between sites, and highlights potential

375 recolonisation pathways. However, the scale of any ongoing demographically relevant

376 connectivity between sites remains unresolved, and warrants further investigation given the

377 limited spatial ranges of the populations may put them a greater risk of extinction from natural

378 or anthropogenic disturbance (Hastings & Botsford, 2006).

379

380   Our age-based demographic assessment provides evidence of significantly different

381 growth patterns, not only among regions, but also within regions. Marked differences in the

382 growth patterns of coral reef fishes have been demonstrated in other studies over relatively

383 small spatial scales (*e.g.,* Gust *et al*., 2002; Kritzer, 2002; Taylor & McIlwain, 2010). This

384 may be a response to a number of environmental or biological factors. For *C. hemistiktos* this

385 is especially pertinent given that the two regions we surveyed differ vastly in environmental

386 and biological conditions, as well as fishing pressure. *Cephalopholis hemistiktos* is commonly

387 found on open bottom and coral rubble reefs, and as a smaller-bodied *Cephalopholis* species

388 it is dominated by its congeners during interspecific interactions (Shpigel & Fishelson, 1989).

389 Given that the GO-AG region has less well developed, shallower reef systems (Coles, 2003)

390 and reduced fish diversity (Feary *et al.,* 2010), we suspect that a combination of favourable

391 conditions, increased productivity and reduced interspecific competition may be driving the

392 differences in growth between regions. Additionally, levels of exploitation vary across

393 regions. In the RS-GA region *C. hemistiktos* is part of the artisanal catch but is generally not

394 targeted owing to its relatively small size (Gladstone, 2002; Roberts & Polunin, 1992),

395 whereas in the GO-AG region, where asymptotic size is larger, *C. hemistiktos* is actively

396 targeted as part of the commercial catch (Hashim, 1993), and in Oman this species accounts

397 for 42% of the total grouper catch (J. McIlwain, unpub. data). Whilst a full demographic

398 assessment of *C. hemistiktos* is beyond the scope of this study owing to a lack of histological

399 data, our results suggest that different management regimes may be needed within the spatial

400 boundaries demarcated by our genetic analyses.

401

402 **Taxonomic considerations**

403

404 Our genetic analyses reveal differences between populations that most likely represent cryptic

405 lineages of isolated populations, which were unable to reconnect during interglacial periods.

406 With at least five fixed mutational steps between regions (COI data), we estimate that these

407 populations have been separated for several hundred thousand years. Age-based otolith

408 analyses also confirm markedly different growth profiles between regions, matching previous

409 observation by Randall & Ben-Tuvia (1983), whom also reported morphological differences

410 in pectoral fin size and ray count, as well as oblique scale rows. Whilst we acknowledge reef

411 fish growth can be extremely plastic, limited pectoral fin ray counts during our sampling

412 study confirm Randall & Ben-Tuvia's (1983) observations. Despite the weight of congruent

413 evidence for treating the two populations as recently diverged sister species, we found no

414 obvious differences in body colour between regions. These traits are frequently used as initial

415 indicators of species determination, but in reef fishes, many species exhibit stable colour

416 morphs that do not always represent genetically different lineages (e.g., Rocha *et al.*, 2004;

417 Messmer *et al.*, 2005; DiBattista *et al.*, 2012; DiBattista *et al.,* in reviewC), and conversely,

418 cryptic speciation with no obvious visual differences is also common (e.g., Drew *et al.,* 2008;

419 DiBattista *et al.*, 2011, 2013; Fernandez-Silva *et al.,* in review). Thus, we recommend that the

420 two allopatric *C. hemistiktos* populations be considered as distinct reciprocally monophyletic

421 species, with a formal description forthcoming.

422

423   **Conclusions**

424   Our results show species-level molecular divergence congruent with the morphological

425   difference observed between two evolutionary distinct lineages of *C. hemistiktos,* signifying

426   cryptic speciation among populations in the Red Sea and Gulf of Aden versus the Gulf of

427   Oman and Arabian Gulf. Age-based demographic analyses also suggest differing management

428   regimes may be needed across the species range to provide a buffer to future environmental

429   disturbances. These findings underline the importance of molecular techniques, not only in

430   the context of connectivity, but also as a taxonomic tool for understanding the origins and

431   ongoing maintenance of biogeographic patterns within and among species.

432

447 sequencing.

448

449 **REFERENCES**

450

451 Almany, G.R., Berumen, M.L., Thorrold, S.R., Planes, S. & Jones, G.P. (2007) Local

452 replenishment of coral reef fish populations in a marine reserve. *Science*, **316**, 742-744.

453

454 Almany, G.R., Hamilton, R.J., Bode, M., Matawai, M., Potuku, T., Saenz-Agudelo, P.,

455 Planes, S., Berumen, M.L., Rhodes, K.L., Thorrold, S.R., Russ, G.R. & Jones, G.P. (2013)

456 Dispersal of grouper larvae drives local resource sharing in a coral reef fishery. *Current*

457 *Biology*, **23**, 626-30.

458

459 Asgharian, H., Sahafi, H.H., Ardalan, A.A., Shekarriz, S. & Elahi, E. (2011) Cytochrome c

460 oxidase subunit 1 barcode data of fish of the Nayband National Park in the Persian Gulf and

461 analysis using meta-data flag several cryptic species. *Molecular Ecology Resources*, **11**, 461-

462 72.

463

464 Bandelt, H.J., Forster, P. & Röhl, A. (1999) Median-joining networks for inferring

465 intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

466

467 Benjamini Y., Krieger, A.M. & Yekutieli, D. (2006) Adaptive linear step-up procedures that

468 control the false discovery rate. *Biometrika*, **93**, 491–507.

469

470 Berumen, M.L. (2005) The importance of juveniles in modelling growth: butterflyfish at

471 Lizard Island. *Environmental Biology of Fishes*, **72,** 409-413.

472

473   Berumen, M.L., Hoey, A.S., Bass, W.H., Bouwmeester, J., Catania, D., Cochran, J.E.M.,

474   Khahil, M.T., Miyake, S., Mughal, M.R., Spaet, J.L.Y. & Saenz-Agudelo, P. (2013). The

475   status of coral reef ecology research in the Red Sea. *Coral Reefs*, **32,** 737–748.

476

477   Botsford, L.W., Micheli, F. & Hastings, A. (2003) Principles for the design of marine

478   reserves. *Ecological Applications*, **13**, 25-31.

479

480   Bowen B.W., Rocha L.A., Toonen R.J., Karl S.A., Craig M.T., DiBattista J.D, Eble J.A.,

481   Gaither M.R., Skillings D., & Bird C.E. (2013) The origins of tropical marine biodiversity.

482   *Trends in Ecology & Evolution*, **28**, 359–366.

483

484   Briggs, J.C. & Bowen, B.W. (2012) A realignment of marine biogeographic provinces with

485   particular reference to fish distributions. *Journal of Biogeography*, **39,** 12-30.

486

487   Burgess, S.C., Nikols, K.J., Gresemer, C.D., Barnett, L.A.K., Derick, A.G., Satterthwaite,

488   E.V., Yamane, L., Morgan, S.G., White, J.W. & Botsford, L.W. (2014) Beyond connectivity:

489   how empirical methods can quantify population persistence to improve marine protected-area

490   design. *Ecological Applications*, **24**, 257-270.

491

492   Chow, S. & Hazama, K. (1998) Universal PCR primers for S7 ribosomal protein gene introns

493   in fish. *Molecular Ecology*, **7**, 1255-1256.

494

495   Coles, S.L. & Wilson, S.C. (2001) Environmental factors affecting reef corals in Oman: a

496   comparison to the Indo-Pacific region. In: Proceedings of the 1st International Conference on

497   Fisheries, Aquaculture and Environment in the Northwest Indian Ocean. S. Goddard, H. Al-

498   Oufi, J.L. McIlwain & M.C. Claereboudt (eds) Muscat, Oman: Sultan Qaboos University, Vol

499   2. pp. 133–145.

500

501   Coles, S.L. (2003) Coral species diversity and environmental factors in the Arabian Gulf and

502   the Gulf of Oman: a comparison to the Indo-Pacific region. *Atoll Research Bulletin*

503   **507,** 1–19.

504

505   Craig, M.T., Sadovy de Mitcheson, Y.J. & Heemstra, P.C. (2011) Groupers of the World: A

506   Field and Market Guide. NISC (Pty) Ltd, Grahamstown, South Africa.

507

508   Darriba, D., Taboada, G.L., Doallo, R. & Posada, D. (2012) jModelTest 2: more models, new

509   heuristics and parallel computing. *Nature Methods*, **9**, 772.

510

511   DiBattista, J.D., Wilcox, C., Craig, M.T., Rocha, L.a. & Bowen, B.W. (2011)

512   Phylogeography of the pacific blueline surgeonfish, acanthurus nigroris, reveals high genetic

513   connectivity and a cryptic endemic species in the hawaiian archipelago. *Journal of Marine*

514   *Biology*, **2011**, 1-17.

515

516   DiBattista, J.D., Waldrop, E., Bowen, B.W., Schultz, J.K., Gaither, M.R., Pyle, R.L. & Rocha,

517   L.A. (2012) Twisted sister species of pygmy angelfishes: discordance between taxonomy,

518   coloration, and phylogenetics. *Coral Reefs*, **31**, 839-851.

519

520   DiBattista, J.D., Berumen, M.L., Gaither, M.R., Rocha, L.A., Eble, J.A., Choat, J.H., Craig,

521   M.T., Skillings, D.J. & Bowen, B.W. (2013) After continents divide: comparative

522   phylogeography of reef fishes from the Red Sea and Indian Ocean. *Journal of Biogeography*,

523   **40**, 1170-1181.

524

525   DiBattista, J.D., Choat, J.H., Gaither, M.R., Hobbs, J.P., Lozano-Cortés, D.F., Myers, R.,

526   Paulay, G., Rocha, L.A., Toonen, R.J., Westneat, M. & Berumen, M.L. (in review

527   A) On the origin of endemic species in the Red Sea. *Journal of Biogeography*.

528   Revision submitted June 1, 2015.

529

530   DiBattista, J.D., Roberts, M., Bouwmeester, J., Bowen, B.W., Coker, D.F., Lozano-Cortés,

531   D.F., Choat, J.H., Gaither, M.R., Hobbs, J.P., Kahil, M., Kochzius, M., Myers, R., Paulay, G.,

532   Robitzch, V., Saenz-Agudelo, P., Salas, E., Sinclair-Taylor, T.H., Toonen, R.J., Westneat, M.,

533   Williams, S. & Berumen, M.L. (in review B) A review of contemporary patterns of endemism

534   for shallow water reef fauna in the Red Sea. *Journal of Biogeography*. Revision submitted

535   April 19, 2015.

536   DiBattista, J.D., Waldrop, E., Rocha, L.A., Craig, M.T., Berumen, M.L., & Bowen, B.W. (in

537   review C) Blinded by the bright: A lack of congruence between colour morph and

538   phylogeography for a cosmopolitan Indo-Pacific butterflyfish, *Chaetodon auriga*. *Journal of*

539   *Biogeography*, revision submitted May 4, 2015.

540

541   Drew, J., Allen, G.R., Kaufman, L. & Barber, P.H. (2008) Endemism and regional color and

542   genetic differences in five putatively cosmopolitan reef fishes. *Conservation Biology*, **22**,

543   965–975.

544

545   Drummond, A.J., Rambaut, A., Shapiro, B. & Pybus, O.G. (2005) Bayesian coalescent

546   inference of past population dynamics from molecular sequences. *Mol Biol Evol*, **22**, 1185-92.

547

548   Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S.,

549   522 Thierer, T. & Wilson, A. (2009) Geneious v7.0.6. Available at:

550   http://www.geneious.com/.

551

552   Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with

553   BEAUti and the BEAST 1.7. *Molecular biology and evolution*, **29**, 1969-1973.

554

555   El-Etreby, S.G., Roberts, C.M., Ghobashy, A.A., & Zyadah, M.A. (1993) Coral reef groupers

556   (family: Serranidae) in south Sinai: reproduction and hermaphroditism in the fish

557   Cephalopholis hemistiktos (Rüppell, 1830). *Journal of the Egyptian German Society of*

558   *Zoology 12(B) Anatomy and Embryology*, 1–19.

559

560   Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred

561   from metric distances among DNA haplotypes: application to human mitochondrial DNA

562   restriction data. *Genetics*, **131**, 479.

563

564   Excoffier, L. & Lischer, H.E. (2010) Arlequin suite ver 3.5: a new series of programs to

565   perform population genetics analyses under Linux and Windows. *Molecular Ecology*

566   *Resources*, **10**, 564-567.

567

568   Feary, D.A., Burt, J.A., Bauman, A.G., Usseglio, P., Sale, P.F. & Cavalcante, G.H. (2010)

569   Fish communities on the world's warmest reefs: what can they tell us about the effects of

570   climate change in the future? *Journal of Fish Biology*, **77**, 1931-47.

571

572 Fernandez-Silva I., Randall, J.E, Coleman, R.R., DiBattista, J.D., Rocha, L.A., Reimer, J.D.,

573 Meyer, C.G. & Bowen, B.W. (in review) Yellow tails in a Red Sea: Phylogeography of

574 the Indo-Pacific goatfish *Mulloidichthys flavolineatus* reveals isolation in peripheral provinces

575 and cryptic evolutionary lineages. *Journal of Biogeography*. Revision submitted June 1, 2015.

576

577 Fu, Y.-X. (1997) Statistical tests of neutrality of mutations against population growth,

578 hitchhiking and background selection. *Genetics*, **147**, 915-925.

579

580 Gaines, S.D., White, C., Carr, M.H. & Palumbi, S.R. (2010) Designing marine reserve

581 networks for both conservation and fisheries management. *Proceedings of the National*

582 *Academy of Sciences*, **107**, 18286-18293.

583

584 Gaither, M.R., Toonen, R.J., Robertson, D.R., Planes, S. & Bowen, B.W. (2010) Genetic

585 evaluation of marine biogeographical barriers: perspectives from two widespread Indo Pacific

586 snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). *Journal of Biogeography*, **37**, 133-147.

587

588 Gladstone, W. (2002) Fisheries of the Farasan Islands (Red Sea). *Naga, WorldFish Center*

589 *Quarterly*, **25**, 30-34.

590

591 Guindon, S. & Gascuel, O. (2003) A simple, fast and accurate method to estimate large

592 phylogenies by maximum-likelihood. *Systematic Biology*, **52**, 696-704.

593

594 Gust, N., Choat, J.H. & Ackerman, J.L. (2002) Demographic plasticity in tropical reef fishes.

595 *Marine Biology*, **140**, 1039-1051.

596

597 Harpending, H.C. (1994) Signature of ancient population growth in a low-resolution

598 mitochondrial DNA mismatch distribution. *Human Biology*, 591-600.

599

600 Harrison, H.B., Williamson, D.H., Evans, R.D., Almany, G.R., Thorrold, S.R., Russ, G.R.,

601 Feldheim, K.A., van Herwerden, L., Planes, S., Srinivasan, M., Berumen, M.L. & Jones, G.P.

602 (2012) Larval export from marine reserves and the recruitment benefit for fish and fisheries.

603 *Current Biology*, **22**, 1023-8.

604

605 Hashim, O. (1993). Fisheries study in the Gulf. *Marine Pollution Bulletin, 27*, 279-284.

606

607 Hastings, A. & Botsford, L.W. (2006) Persistence of spatial populations depends on returning

608 home. *Proceedings of the National Academy of Sciences of the United States of America*, **103**,

609 6067-6072.

610

611 Hedgecock, D., Barber, P.H. & Edmands, S. (2007) Genetic Approaches to Measuring

612 Connectivity. *Oceanography*, **20**, 70-79.

613

614 Hellberg, M.E. (2007) Footprints on water: the genetic wake of dispersal among reefs. *Coral*

615 *Reefs*, **26**, 463-473.

616

617 Hellberg, M.E. (2009) Gene flow and isolation among populations of marine animals. *Annual*

618 *Review of Ecology Evolution and Systematics*, **40**, 291-310.

619

620 Hixon, M.A., Pacala, S.W. & Sandin, S.A. (2002) Population regulation: historical context

621 and contemporary challenges of open vs. closed systems. *Ecology*, **83**, 1490-1508.

622

623   Jones, G.P., Planes, S. & Thorrold, S.R. (2005) Coral reef fish larvae settle close to home.

624   *Current Biology*, **15**, 1314-1318.

625

626   Kimura, D.K. (1980) Likelihood methods for the von Bertalanffy growth curve. *Fishery*

627   *Bulletin*, **77**, 765-774.

628

629   Kritzer, J.P., Davies, C.R. & Mapstone, B.D. (2001) Characterizing fish populations: effects

630   of sample size and population structure on the precision of demographic parameter estimates.

631   *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 1557-1568.

632

633   Kritzer, J.P. (2002) Variation in the population biology of stripey bass Lutjanus carponotatus

634   within and between two island groups on the Great Barrier Reef. *Marine Ecology Progress*

635   *Series*, **243**, 191-207.

636

637   Kritzer, J.P. & Sale, P.F. (2004) Metapopulation ecology in the sea: from Levins' model to

638   marine ecology and fisheries science. *Fish and Fisheries*, **5**, 131-140.

639

640   Leis, J.M. (2002) Pacific coral-reef fishes: the implications of behaviour and ecology of

641   larvae for biodiversity and conservation, and a reassessment of the open population paradigm.

642   *Environmental Biology of Fishes*, **65**, 199-208.

643

644   Lessios, H.A. (2008) The great American schism: divergence of marine organisms after the

645   rise of the Central American Isthmus. *Annual Review of Ecology, Evolution, and Systematics*,

646   **39**, 63-91.

647

648 Librado, P. & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA

649 polymorphism data. *Bioinformatics*, **25**, 1451-1452.

650

651 McCormick, M.I. & Molony, B.W. (1995) Influence of water temperature during the larval

652 stage on size, age and body condition of a tropical reef fish at settlement. *Marine Ecology*

653 *Progress Series*, **118**, 59-68.

654

655 Meeker, N.D., Hutchinson, S.A., Ho, L. & Trede, N.S. (2007) Method for isolation of PCR-

656 ready genomic DNA from zebrafish tissues. *BioTechniques*, **43**, 610-614.

657

658 Messmer, V., van Herwerden, L., Munday, P.L. & Jones, G.P. (2005) Phylogeography of

659 colour polymorphism in the coral reef fish Pseudochromis fuscus, from Papua New Guinea

660 and the Great Barrier Reef. *Coral Reefs*, **24**, 392-402.

661

662 Moritz, C. (1994) Defining "evolutionarily significant units" for conservation. *Trends in*

663 *Ecology and Evolution,* **9**, 373–375.

664

665 Ormond, R. & Edwards, A. (1987) Red Sea fishes. Key environments – Red Sea (ed. by A.J.

666 Edwards and S.M. Head), pp. 251–287. Pergamon Press, Oxford.

667

668 Pianka, E.R. (1978) Evolutionary Ecology. Harper & Row Publishers, New York.

669

670 Rambaut. A., Suchard, M.A., Xie, D. & Drummond, A.J. (2014) Tracer v1.6. Available at:

671 http://beast.bio.ed.ac.uk/Tracer.

672

673   Randall, J.E. & Ben-Tuvia, A. (1983) R review of the groupers (Pisces: Serranidae:

674   Epinephelinae) of the Red Sea, with descriptions of a new species of *Cephalopholis*. *Bulletin*

675   *of Marine Science*, **33**, 373-426.

676

677   Randall, J.E. (1994) Twenty-two new records of fishes from the Red Sea. *Fauna of Saudi*

678   *Arabia,* **14,** 259–275.

679

680   Riginos, C. & Victor, B.C. (2001) Larval spatial distributions and other early life-history

681   characteristics predict genetic differentiation in eastern Pacific blennioid fishes. *Proceedings*

682   *of the Royal Society of London Series B-Biological Sciences*, **268**, 1931-1936.

683

684   Roberts, C.M., Shepherd, A.R.D. & Ormond, R.F.G. (1992) Large-scale variation in

685   assemblage structure of Red Sea butterflyfishes and angelfishes. *Journal of Biogeography,* **19,**

686   239–250.

687

688   Rocha, L.A., Bass, A.L., Robertson, D.R. & Bowen, B.W. (2002) Adult habitat preferences,

689   larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes

690   (Teleostei : Acanthuridae). *Molecular Ecology*, **11**, 243-252.

691

692   Rocha, L.A. (2004) Mitochondrial DNA and Color Pattern Variation in Three Western

693   Atlantic *Halichoeres* (Labridae), with the Revalidation of Two Species. *Copeia*, **4**, 770-782.

694

695   Rocha, L.A., Craig, M.T. & Bowen, B.W. (2007) Phylogeography and the conservation of

696   coral reef fishes. *Coral Reefs*, **26**, 501-512.

697

698 Rogers, A.R. & Harpending, H. (1992) Population growth makes waves in the distribution of

699 pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.

700

701 Savidge, G., Lennon, J. & Matthews, J. (1990) A shore-based survey of upwelling along the

702 coast of the Dhofar region, southern Oman. *Continental Shelf Research*, **10**, 259–275.

703

704 Schils, T. & Coppejans, E. (2003) Phytogeography of upwelling areas in the Arabian Sea.

705 *Journal of Biogeography*, **30**, 1339-1356.

706

707 Sheppard, C., Price, A. & Roberts, C. (1992) *Marine Ecology of the Arabian Region: Patterns*

708 *and Processes in Extreme Tropical Environments*. Academic Press, Toronto.

709

710 Shpigel, M. & Fishelson, L. (1989) Habitat partitioning between species of the genus

711 *Cephalopholis* (Pisces,Serranidae) across the fringing reef of the Gulf of Aqaba (Red Sea).

712 *Marine Ecology Progress Series*, **58**, 17-22.

713

714 Seutin, G., White, B.N. & Boag, P.T. (1991) Preservation of avian blood and tissue samples

715 for 654 DNA analyses. *Canadian Journal of Zoology*, **69**, 82–90.

716

717 Siddall, M., Rohling, E.J., Almogi-Labin, A., Hemleben, C., Meischner, D., Schmelzer, I. &

718 Smeed, D.A. (2003) Sea-level fluctuations during the last glacial cycle. Nature, 423, 657 853–

719 858.

720

721   Simpson, S.D., Harrison, H.B., Claereboudt, M.R. & Planes, S. (2014) Long-distance

722   dispersal via ocean currents connects Omani clownfish populations throughout entire species

723   range. *PLoS One*, **9**, e107610.

724

725   Slatkin, M. (1993) Isolation by distance in equilibrium and non-equilibrium populations.

726   *Evolution*, **47**, 264-279.

727

728   Sponaugle, S., Grorud-Colvert, K. & Pinkard, D. (2006) Temperature-mediated variation in

729   early life history traits and recruitment success of the coral reef fish *Thalassoma bifasciatum*

730   in the Florida Keys. *Marine Ecology Progress Series*, **308**, 1-15.

731

732   Stephens, M. & Donnelly, P. (2003) A comparison of bayesian methods for haplotype

733   reconstruction from population genotype data. *The American Journal of Human Genetics*, **73**,

734   1162-1169.

735

736   Tamura, K. & Nei, M. (1993) Estimation of the number of nucleotide substitutions in the

737   control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and*

738   *Evolution*, **10**, 512-526.

739

740   Taylor, B.M. & McIlwain, J.L. (2010) Beyond abundance and biomass: effects of marine

741   protected areas on the demography of a highly exploited reef fish. *Marine Ecology Progress*

742   *Series*, **411**, 243-258.

743

744   Thresher, R.E., Colin, P.L. & Bell, L.J. (1989) Planktonic duration, distribution and

745   population structure of western and central Pacific Damselfishes (Pomacentridae). *Copeia*, **2,**

746   420-434.

747

748   Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R. & Hebert, P.D. (2005) DNA barcoding

749   Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological*

750   *Sciences*, **360**, 1847-1857.

751

752   Weersing, K. & Toonen, R.J. (2009) Population genetics, larval dispersal, and connectivity in

753   marine systems. *Marine Ecology-Progress Series*, **393**, 1-12.

754

755   White, C., Selkoe, K.A., Watson, J., Siegel, D.A., Zacherl, D.C. & Toonen, R.J. (2010) Ocean

756   currents help explain population genetic structure. *Proceedings of the Royal Society B:*

757   *Biological Sciences*, **277**, 1685.

758

759 **SUPPORTING INFORMATION**

760 Additional Supporting Information may be found in the online version of this article:

761

762 **Appendix S1**

763 Summary of morphometric differences between two populations of *C. hemistiktos* from

764 Randall & Ben-Tuvia (1983).

765

766 **Appendix S2**

767 Phylogenetic

768

769 **Appendix S3**

770 Reproductive timing of *C. hemistiktos* sampled from the coast of Oman.

771 **Biosketch**

772 The authors' interests are focused on illuminating the evolutionary processes that generate

773 marine biodiversity. This study builds on recently completed range-wide phylogeographic

774 surveys of over 20 Indo-Pacific reef fishes, providing a robust genetic foundation for

775 subsequent investigations of isolated archipelagos and seas, including the evolution of the

776 unique Red Sea fauna.

777

778 Author contributions: J.L.M conceived the initial idea for this study; all authors collected

779 tissue samples; M.A.P and J.D.D produced DNA sequences and analysed the data; M.L.B and

780 N.E.H. contributed reagents and materials; B.M.T and J.L.M produced and analysed age-

781 based demographic data; M.A.P led the writing with contributions from all other authors.

782

783 Editor: Michelle Gaither

Table 1 Summary statistics for genetic and demographic analyses for *Cephalopholis hemistiktos* sampled from 11 sites around the Arabian Peninsula. Given are sample sizes (*n*), haplotype diversity (*h*), nucleotide diversity (π), maximum total length sampled (TL$_{MAX}$; mm). Values in bold are statistically significant ($P < 0.02$; Fu, 1997).

		COI			Fu's	S7	Demography	
Site	Region	*n*	*h* ± SD	π ± SD	F_S	*n*	*n*	TL$_{MAX}$
Magna, KSA	Gulf of Aqaba	35	0.72 ± 0.04	0.0027 ± 0.0018	0.14	37	41	232
Thuwal, KSA	Red Sea	32	0.72 ± 0.05	0.0027 ± 0.0018	-0.89	40	57	195
Al Lith, KSA	Red Sea	13	0.69 ± 0.12	0.0020 ± 0.0015	-1.24	13	18	180
Sanganeb Atoll, Sudan	Red Sea	39	0.80 ± 0.06	0.0032 ± 0.0021	**-6.75**	33	29*	221
Farasan Islands, KSA	Red Sea	31	0.68 ± 0.06	0.0023 ± 0.0016	-1.34	27	36	215
Bay de Ghoubett, Djibouti	Gulf of Aden	49	0.73 ± 0.05	0.0027 ± 0.0017	-1.90	44	52	247
Muscat, Oman	Gulf of Oman	23	0.81 ± 0.06	0.0035 ± 0.0022	-1.59	20	493	455
Musandam, Oman	Gulf of Oman	-				-	202	443
Abu Dhabi, UAE	Arabian Gulf	-				-	31	288
Jubail, KSA	Arabian Gulf	52	0.73 ± 0.05	0.0026 ± 0.0017	**-3.79**	47	52*	342

KSA = Kingdom of Saudi Arabia; UAE = United Arab Emirates; * = not included in otolith age-based analyses

Table 2 Pairwise F-statistics for 8 populations of *Cephalopholis hemistiktos* from the Arabian Peninsula. Pairwise Φ_{ST} values based on COI data are below the diagonal and F_{ST} values based on S7 data are above the diagonal. Values in bold are significant ($P < 0.05$) after correction for multiple tests using the false discovery rate (Benjamini *et al.,* 2006).

Site	Magna	Thuwal	Al Lith	Sanganeb	Farasan Is.	Ghoubett	Muscat	Jubail
Magna	-	-0.007	-0.019	-0.009	-0.014	-0.010	**0.407**	**0.309**
Thuwal	0.015	-	-0.017	0.001	-0.014	-0.004	**0.366**	**0.287**
Al Lith	0.038	0.090	-	-0.024	-0.023	-0.009	**0.408**	**0.273**
Sanganeb	0.000	0.000	0.008	-	-0.009	-0.001	**0.399**	**0.299**
Farasan Is.	0.001	0.005	0.004	0.000	-	-0.010	**0.401**	**0.298**
Ghoubett	0.000	0.046	0.018	0.000	0.018	-	**0.427**	**0.331**
Muscat	**0.809**	**0.808**	**0.822**	**0.796**	**0.796**	**0.815**	-	**0.233**
Jubail	**0.835**	**0.835**	**0.851**	**0.846**	**0.825**	**0.836**	0.000	-

Table 3 Mean total length-at-age for age one through five for *Cephalopholis hemistiktos* among the two distinct regions. RS-GA = Red Sea and Gulf of Aden region; GO-AG = Gulf of Oman and Arabian Gulf region; n = sample size; t = test statistic; *d.f.* = degrees of freedom; P = probability value. Values in bold are significant ($P < 0.05$).

Age (years)	n	Mean TL (mm) ± SE RS-GA	n	GO-AG	t	*d.f*	P
1	12	120.9 ± 6.03	5	153.2 ± 14.49	2.48	15	**0.03**
2	18	151.1 ± 4.63	57	226.8 ± 5.04	11.06	56.9	**<0.0001**
3	19	150.3 ± 5.29	105	256.8 ± 4.11	15.90	43.6	**<0.0001**
4	23	175.7 ± 5.62	109	262.2 ± 3.99	12.54	47.5	**<0.0001**
5	21	176.5 ± 6.06	99	276.8 ± 4.23	13.57	42.2	**<0.0001**

Figure 1 Map of sampling locations for *Cephalopholis hemistiktos* across the Arabian Peninsula. Dark grey denotes extent of species range highlighting the disjunctive distribution pattern.

Figure 2 Mismatch distributions (a, b) and Bayesian skyline plots (c, d) based on mitochondrial COI sequence data (629 bp) for *Cephalopholis hemistiktos* from the Red Sea and Gulf of Aden (a, c), and Gulf of Oman and Arabian Gulf (b, d) regions. Grey lines on skyline plots represent 95% highest posterior densities.

Figure 3 Median-joining network based on (a) 629 base pairs of mitochondrial COI sequence data, and (b) a 206 base pairs of the S7 nuclear intron from *Cephalopholis hemistiktos* sampled across the Arabian Peninsula. Circles represent haplotypes and are sized in proportion to total frequency. Branches represent a single nucleotide change with crossbars indicating additional nucleotide changes. Unsampled haplotypes are unshaded and colours denote collection site as indicated by the key.

Figure 4 Length- (a) and age-frequency (b) distributions for *Cephalopholis hemistiktos* sampled from the Arabian Peninsula. Data are grouped into regions: Red Sea and Gulf of Aden (open bars), and Gulf of Oman and Arabian Gulf (black bars).

Figure 5 Comparison of growth parameters for *Cephalopholis hemistiktos* among eight sites using bivariate 95% confidence ellipses surrounding estimates of K (growth coefficient) and L_∞ (mean asymptotic length).

(a) Red Sea and Gulf of Aden

(b) Gulf of Oman and Arabian Gulf

(c)

(d)

(a)

(b)

Magna
Thuwal
Al Lith
Sanganeb
Farasan Is.
Ghoubett
Muscat
Jubail