e-ISSN: 1948-6596 https://escholarship.org/uc/fb doi:10.21425/F5FBG42697

RESEARCH ARTICLE

Phylogeography and conservation biogeography of the humphead wrasse, *Cheilinus undulatus*

Ka Yan Ma[1], Patrick L. Colin[2], Yvonne Sadovy de Mitcheson[3,*] and Michael N Dawson[4]

[1] School of Life Sciences, the Chinese University of Hong Kong, Shatin, Hong Kong.

[2] Coral Reef Research Foundation, PO Box 1765, Koror, PW 96940, Palau.

[3] School of Biological Sciences, Swire Institute of Marine Science, the University of Hong Kong, Pok Fu Lam Road, Hong Kong.

[4] Department of Life and Environmental Sciences, School of Natural Sciences, University of California, Merced, 5200 North Lake Road, Merced, CA 95343, USA.

* Corresponding author: yjsadovy@hku.hk

Abstract. The Humphead wrasse (also known as the Napoleon fish), *Cheilinus undulatus*, is a highly prized coral reef fish, listed on CITES Appendix II and endangered on the IUCN Red List. It is widespread across much of the Indo-Pacific region. The fish has a 4-6 week pelagic egg and larval stage, suggesting the potential for high connectivity among populations. However, its range spans important biogeographic boundaries that are associated with barriers to gene flow and deep phylogeographic structure in some marine fishes and invertebrates, raising the possibility of significant population genetic structure. We describe preliminarily the genetic structure of the Humphead wrasse across much of its range and consider the implications for effective conservation. Using mitochondrial DNA sequencing (cytochrome b and control region) coupled with microsatellite analyses, we find primarily a signal of eurymixis — i.e., low, heterogeneous population genetic differentiation across much of the species' range (F_{ST} analogs: <0.11 cytochrome b, <0.09 control region, <0.02 microsatellites) with the exception of modest differentiation primarily toward the peripheries (e.g., Pohnpei, Seychelles; F_{ST} analogs: 0.03–0.24 cytochrome b, 0.05–0.24 control region, 0.03–0.22 microsatellites) — though isolation by distance is not excluded. The general dearth of structure is consistent with population expansion, following an historical bottleneck and with high contemporary gene flow. The implications are that Humphead wrasse is a metapopulation and that its conservation status depends on successful management of a sufficient but currently unknown number and distribution of populations across a multi-national network.

Keywords: Conservation biogeography, connectivity, conservation planning, food security, gene flow, Humphead wrasse, phylogeography, population genetics, Indo-West Pacific

Introduction

The Humphead wrasse, *Cheilinus undulatus* (also known as the Napoleon fish), is a highly prized fish found on Indo-Pacific coral reefs, popular with divers and seafood gourmets alike, but fishing has seriously reduced many populations (Sadovy et al. 2003, Poh and Fanning 2012, Sadovy de Mitcheson et al. 2017). Reported to reach over 200 cm standard length, the species is the biggest species of wrasse and one of the largest reef fish species. It is an iconic component of "charismatic marine megafauna", such as whales, sharks, and manta rays, with special diving tourism value (Aw 2000, Graham et al. 2014, Briggs 2016). In some countries it has cultural significance for local consumption and an important traditional value (Sadovy et al. 2003). It also has high economic value for its taste in the international live reef fish trade (LRFT) centered around the luxury seafood sector of Hong Kong and mainland China, retailing for US$200-600 per kilogram (Sadovy de Mitcheson et al. 2017). Overexploitation for international trade, coupled with a lack of fisheries management or trade controls, resulted in *C. undulatus* receiving Convention of International Trade in Endangered Species (CITES) Appendix II listing in 2004. Humphead wrasse is the first reef food fish to receive such listing. The species is considered to be conservation-dependent (Gillett 2010), which means that conservation or management action is required wherever the species is exploited.

Several aspects of the biology of *C. undulatus* increase its vulnerability to overfishing. First, it

is long-lived (over 30 years) and matures slowly (5-7 years) (Choat et al. 2006), a life history associated with relatively low productivity and slow population growth that cannot sustain heavy fishing. Second, it is protogynous: some fish change sex as adults, from female to male at about 8 years of age (Choat et al. 2006, Sadovy de Mitcheson et al. 2010), a reproductive strategy that is sensitive to fishing which is often size-selective. Targeting large fish could lead to differential loss of males and a sex ratio bias towards females, which do not exceed about 80 cm total length (TL). Targeting small fish — the most common focus for the 'plate-sized' fish (500–1,000 g, 25–45 cm TL) export trade supplying banquets or family dinners in the Chinese seafood market — means that juveniles or small females are particularly sought-after. This potentially leads to recruitment overfishing due to reduced reproductive capacity (Kindsvater et al. 2017). Third, the species' habit of forming spawning aggregations (Colin 2010) and sleeping in shallow water reefs at night can make it particularly accessible to capture by fishers (Graham et al. 2014). Finally, it is not naturally common, nor is it farmed commercially (i.e., hatchery-produced), and declines in availability further increase its value and, hence, interest to traders in the international seafood trade. The combination of high fishing pressure on the species for export, its biological characteristics, and greater rarity leading to higher prices, can make effective management particularly challenging and has resulted also in an IUCN classification of 'threatened' (Courchamp et al. 2006, Sadovy de Mitcheson et al. 2017).

The marine biogeographic conservation challenges to sustainably managing Humphead wrasse are similar to those for other geographically widespread species with dispersive eggs and larvae that are particularly vulnerable to overexploitation and are of unknown genetic structure. Humphead wrasse is recorded from 48 countries with an east-west extent of 18,000 km (Red Sea to French Polynesia) and several-thousand kilometers north to south (Ryukyu Islands to New Caledonia) (Myers 1991, Sadovy et al. 2003). The species is closely associated with coral reef habitats but naturally uncommon in most (Fig. 1). Its range crosses open-water biogeographic barriers associated with restrictions to gene flow and deep phylogeographic structure for at least some marine fishes and invertebrates (e.g., Lindenfelser 1984, Lacson and Clark 1995, McMillan and Palumbi 1995, Lavery et al. 1996, Benzie 1998, 1999).

Field and aquaculture studies indicate that the species has a 29–42 day post-spawning pelagic larval duration (PLD; Victor 1986, Hirai et al. 2013), typical of many reef fishes. For those reef fishes with a long pelagic stage, there is high potential for dispersal and connectivity among even distant populations resulting in a largely homogeneous gene pool (Horne et al. 2008, Klanten et al. 2007), showing only 'chaotic genetic patchiness' or 'eurymixis', (i.e., complex transient spatio-temporal genetic heterogeneity, Dawson 2014a). Uncertainty about the likely scale of population structure also exists because different geographic regions may have very different potentials for larval retention and self-recruitment, which can affect genetic connectivity. Some locations, such as Palau, have some areas with high potential for retention of larvae around the archipelago while others may favor larvae becoming entrained in oceanic circulations (Hamner and Largier 2012). Other geographic locations, such as the South-China Sea, may have much lower potential of retention and little prospect of long-lived pelagic larvae returning to near their natal island or reef (McManus 1994).

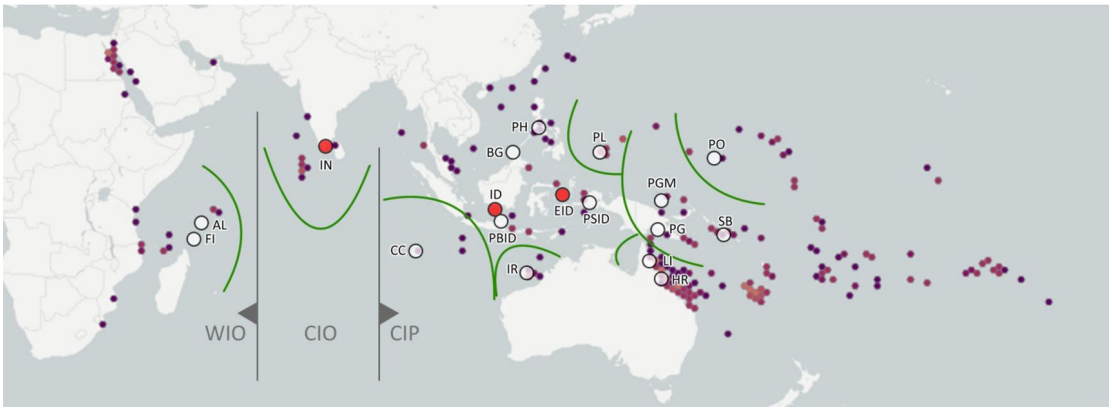

Figure 1. The geographic distribution of Humphead wrasse, *Cheilinus undulatus*, sampling localities, and study design. The distribution of *C. undulatus* is mapped using records in GBIF <https://www.gbif.org/species/2383313> and represented as small polygons; darker polygons indicate more records; map generated 14th September 2018. The sampling localities of this study are represented by large circles; the large red circles indicate government-confiscated samples that lack exact sampling data and may be sourced from a larger catchment than indicated in the map. Grey and green lines represent biogeographic regions of the seas used in AMOVA analyses: green lines delimit Large Marine Ecosystems; WIO = western Indian Ocean, CIO = central Indian Ocean, CIP = central Indo-Pacific. See Table 1 for location codes and Supplementary Documentation S1 Table S1 for additional details of sampling.

Empirical understanding of population structure and connectivity is important for the successful management and conservation of widespread marine species, such as *C. undulatus,* given the combination of species-specific biotic and region-specific abiotic factors shaping connectivity and demographic history (e.g., Lester et al. 2007, Kelly and Eernisse 2007, Schiebelhut and Dawson 2018). Moreover, for species of conservation concern that are not well protected, as is the case for this species over extensive areas, identification of key larval source areas and habitat may support more effective management and protection of populations should these reach much-reduced levels (Cabral et al. 2016). To date, although microsatellite loci (Hu et al. 2013) and whole mitochondrial genomes (Qi et al. 2013, Matthew et al. 2018) including the control region (Indriatmoko et al. 2016) have been characterized, there has been no population genetic analysis of *C. undulatus*. Our goal in this paper is to provide the first estimates of phylogeographic and population genetic structure of the Humphead wrasse across much its range and, thus, inform strategies for managing fisheries sustainably, including conserving genetic diversity in a time of rapid environmental change and increasing exploitation pressure (Rocha et al. 2007, Carvalho et al. 2017). The study will form a foundation for sustaining exploited populations and their value to source countries for food or income through sales or ecotourism.

Methods

Sampling and DNA extraction

Fin clips or muscle samples were taken from fish being sold in markets and preserved in 80% ethanol. Samples were shipped to California and stored at -20 °C until analysis. DNA was extracted using the Qiagen DNeasy purification kit following the protocol for animal tissues. See Fig. 1, Table 1, and Supplementary Document S1 for sample locations and sample sizes.

Mitochondrial DNA analyses

One microlitre of purified DNA solution was used in 50 µL polymerase chain reactions with 0.5 U AmpliTaq, 5 µL 10X buffer (Applied Biosystems, Foster City, California, USA), and final concentrations of primers at 0.5 µM (Operon Biotechnologies Inc., Huntsville, Alabama, USA), $MgCl_2$ at 2.5–3 mM (Applied Biosystems), and dNTPs at 0.2 mM (Bioline, Sydney, New South Wales, Australia) to amplify two markers: partial mitochondrial cytochrome b (*cyt b*) and partial control region (*mtCR*) (see Supplementary Document S2 for detailed protocol). Amplifications were cleaned using 47 µl PCR product with 4 µL exonuclease I and 4 µL shrimp alkaline phosphatase (USB Corporation, Cleveland, Ohio, USA) incubated at 37°C for 15 mins, 80°C for 15 mins, and cooled to 4°C until prepared for sequencing. Amplicons were sequenced by Cogenics Inc. (Houston, TX, USA), the University of Washington High-throughput Genomics Unit (Seattle, WA, USA), or the DNA Sequencing Facility of the University of California, Berkeley (CA, USA). Sequence data were edited—i.e., primers removed, terminal gaps trimmed, and base calls error-checked—using Sequencher 5.2–5.3 (Gene Codes Corp., Ann Arbor, MI, USA) and aligned using MUSCLE (Edgar 2004) with default settings. Sequences are available in GenBank (Accession numbers: *cyt b* = MK431901—MK432249, *mtCR* = MK432250—MK432593).

Haplotype diversity (*h*) and nucleotide diversity (π) (Nei 1987) of both mtDNA markers were computed using Arlequin 3.5 (Excoffier and Lischer 2010). Population genetic structure was examined by three approaches. [1] TCS networks were constructed using POPART software (Leigh and Bryant 2015) to visualize evolutionary relationships among haplotypes. [2] Pairwise Φ_{ST} statistics were computed using Arlequin with 9,999 permutations. [3] We implemented an analysis of molecular variance (AMOVA) in Arlequin to test for hierarchical genetic structure based on marine ecoregions (Spalding et al. 2007); we tested for genetic partitions (i) between the Western Indo-Pacific and the Central Indo-Pacific, (ii) between Western Indian Ocean and the rest of the Indo-Pacific, and (iii) among all marine biogeographic provinces covered in this study. To determine if significant isolation by distance (IBD) exists among populations, we used Mantel tests to test for correlation between pairwise Φ_{ST} values and geographic distance. We used the least-cost path calculated by the marmap package in R as geographic distance (Pante et al. 2013). Mantel tests were performed with 999 permutations on datasets in which negative Φ_{ST} values were converted to zeros using the vegan package in R (Oksanen et al. 2008). Mismatch distributions were constructed in Arlequin to examine population expansion in this species.

Microsatellite analyses

Microsatellite loci were developed by Genetic Identification Services (GIS: Chatsworth, CA, USA). We isolated and genotyped ten polymorphic microsatellites for the Humphead wrasse (see Supplementary Document S2 for detailed protocol). PCR products were sized on an ABI 3730 at GIS using the size standard GS500(-250) LIZ. Microsatellites were scored using Genotyper v2.0. We used Micro-Checker (Van Oosterhout et al et al. 2004) to test for null alleles, large allele dropout and scoring errors. FreeNA (Chapuis and Estoup 2007) was used to test for the impact of null alleles by estimating F_{ST} before and after ENA (excluding null alleles) correction with 10,000 pseudoreplicates to determine significance levels. The microsatellite dataset is available as Supplementary Document S3. The expected and observed heterozygosities were estimated using GenALEx 6.5 (Peakall and Smouse 2012). GENEPOP 4.0 (Rousset 2008) was used to test for significant deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) by running 10,000 Markov chain iterations. Population genetic structure was estimated by four approaches. [1] A discriminant analysis of principal components (DAPC) of genotypes using the

Table 1. Sample information and genetic diversity of Humphead wrasse, *Cheilinus undulatus*, at locations across the Indo-West Pacific. Mitochondrial **(A)** control region and **(B)** cytochrome b datasets. Results are indicated at two levels of significance: P < 0.01 (bold) and P < 0.05 (italic) after Benjamini–Hochberg correction for multiple tests. Comparisons that are significant for both markers are additionally underlined. N, sample size; Nh, number of haplotypes; h, haplotype diversity and its standard deviation (s.d.); π, nucleotide diversity and its standard deviation (s.d.).

(A) Locations	Location code	N	Nh	h	s.d.	π	s.d.	Tajima's D	Tajima P	Fu's F	Fu's F P
Pohnpei	PO	22	19	0.974	0.028	0.006	0.004	-1.005	0.167	-6.942	**0.001**
Solomons	SB	1	1	1.000	0.000	0.000	0.000	0.000	1.000	0.000	N.A.
Tulle, Helix Reef, Great Barrier Reef, Australia	HR	3	2	0.667	0.314	0.004	0.004	0.000	0.815	2.357	0.780
Lizard Isl., Great Barrier Reef, Australia	LI	3	3	1.000	0.272	0.003	0.003	0.000	0.825	-0.341	0.196
Papua New Guinea	PG	1	1	1.000	0.000	0.000	0.000	0.000	1.000	0.000	N.A.
Manus, Papua New Guinea	PGM	2	2	1.000	0.500	0.004	0.004	0.000	1.000	1.099	0.443
Eastern Indonesia	EID	58	55	0.998	0.004	0.006	0.003	-2.315	**0.002**	-25.556	**0.000**
Palau	PL	16	13	0.967	0.036	0.008	0.004	-1.400	0.063	-4.590	*0.018*
Philippines	PH	16	16	1.000	0.022	0.007	0.004	-1.759	*0.028*	-9.701	**0.000**
Banggi, Malaysia	BG	12	10	0.970	0.044	0.007	0.004	-1.276	0.121	-3.194	*0.049*
Papua Barat, Indonesia	PBID	11	9	0.946	0.066	0.004	0.002	-1.529	0.056	-4.887	**0.004**
Pulau Saredang Besar, Indonesia	PSID	13	11	0.974	0.039	0.007	0.004	-1.329	0.079	-3.868	*0.024*
Indonesia	ID	98	92	0.999	0.002	0.006	0.003	-2.323	**0.001**	-25.447	**0.000**
Imperieuse Reef, Western Australia	IR	36	32	0.989	0.012	0.008	0.004	-1.182	0.107	-6.931	*0.014*
Cocos Islands	CC	5	5	1.000	0.127	0.003	0.003	-1.146	0.076	-2.371	*0.017*
India	IN	32	32	1.000	0.008	0.005	0.003	-1.629	*0.037*	-4.862	*0.023*
Alphonse/Bijoutier, Seychelles	AL	2	2	1.000	0.500	0.002	0.003	0.000	1.000	0.693	0.372
Farquhar Is., Seychelles	FI	13	9	0.872	0.091	0.002	0.001	-1.745	*0.031*	-6.311	**0.000**
	All	**344**	**148**	**0.968**	**0.005**	**0.007**	**0.004**	**-2.426**	**0.000**	**-25.074**	**0.000**
(B) Locations	**Code**	**N**	**Nh**	**h**	**s.d.**	**π**	**s.d.**	**Tajima's D**	**Tajima P**	**Fu's F**	**Fu's F P**
Pohnpei	PO	24	8	0.743	0.074	0.003	0.002	-1.293	0.097	-3.442	**0.007**
Solomons	SB	1	1	1.000	0.000	0.000	0.000	0.000	1.000	0.000	N.A.
Tulle, Helix Reef, Great Barrier Reef, Australia	HR	3	1	0.000	0.000	0.000	0.000	0.000	1.000	0.000	N.A.
Lizard Isl., Great Barrier Reef, Australia	LI	3	3	1.000	0.272	0.003	0.003	0.000	0.943	-1.216	0.077
Papua New Guinea	PG	1	1	1.000	0.000	0.000	0.000	0.000	1.000	0.000	N.A.
Manus, Papua New Guinea	PGM	2	2	1.000	0.500	0.018	0.019	0.000	1.000	2.197	0.539
Eastern Indonesia	EID	58	8	0.534	0.060	0.001	0.001	-1.503	0.045	-4.851	**0.001**
Palau	PL	16	7	0.692	0.124	0.003	0.002	-1.611	0.035	-1.814	0.072
Philippines	PH	17	4	0.596	0.099	0.001	0.001	1.430	0.952	1.338	0.686

Table 1. Continued...

(B) Locations	All	344	148	0.968	0.005	0.007	0.004	-2.426	0.000	-25.074	0.000
	Code	N	Nh	h	s.d.	π	s.d.	Tajima's D	Tajima P	Fu's F	Fu's F P
Banggi, Malaysia	BG	11	7	0.818	0.119	0.006	0.004	-1.336	0.093	-0.671	0.318
Papua Barat, Indonesia	PBID	12	2	0.409	0.133	0.001	0.001	0.541	0.814	0.735	0.482
Pulau Saredang Besar, Indonesia	PSID	13	3	0.603	0.089	0.001	0.001	0.097	0.600	-0.021	0.382
Indonesia	ID	99	12	0.586	0.038	0.004	0.002	-1.787	0.009	-2.581	0.139
Imperieuse Reef, Western Australia	IR	37	3	0.466	0.081	0.001	0.001	0.090	0.629	0.201	0.459
Cocos Islands	CC	5	1	0.000	0.000	0.000	0.000	0.000	1.000	0.000	N.A.
India	IN	32	3	0.280	0.095	0.001	0.001	-0.655	0.271	-0.661	0.231
Alphonse/Bijoutier, Seychelles	AL	2	2	1.000	0.500	0.006	0.007	0.000	1.000	1.099	0.461
Farquhar Is., Seychelles	FI	13	8	0.808	0.113	0.004	0.003	-2.026	0.008	-3.41	0.009
	All	349	45	0.580	0.025	0.002	0.001	-2.643	**0.000**	-29.984	**0.000**

adegenet R package (Jombart et al. 2010) was used to infer the number of clusters of genetically related individuals from the microsatellites. We used all PCs for the find.clusters function and chose the number of clusters with the lowest BIC value. Generating the DAPC assignment plot with the clustering information as a prior, retained 52 PCs as determined by the xvalDpac function. [2] STRUCTURE 2.3.3 (Pritchard et al. 2000) was used to depict genetic clustering within the samples, under an admixture model without correlated allele frequencies nor sampling locations as priors. For each value of K (from K=1 to K=18), 20 replicates were run with 500,000 steps after 50,000 steps of burn-in. The best K for the dataset was determined by the Evanno method implemented in Structure Harvester web 0.6.93 (Earl 2012) and plotted using Clustering Markov Packager Across K (CLUMPAK) (Kopelman et al. 2015). [3] Pairwise F_{ST} and R_{ST} were estimated using Arlequin. [4] AMOVA was conducted using Arlequin to test for hierarchical genetic structure as described for mtDNA analyses. Locus-by-locus AMOVA was also conducted in Arlequin to reduce error in significance calculation due to missing data. IBD analyses based on F_{ST} and R_{ST} were also conducted (as described for mtDNA analyses). False discovery rates were calculated using the Benjamini–Hochberg procedure (BH).

Results

Mitochondrial DNA analyses

344 *mtCR* sequences (817 bp) and 349 *cyt b* sequences (506 bp) were obtained from 18 sampling locations (Table 1). The *mtCR* sequence alignment contained 155 variable sites, of which 78 were phylogenetically informative. The *cyt b* sequence alignment contained 70 variable sites, of which 18 were phylogenetically informative.

Cyt b and *mtCR* networks each showed a couple of predominant haplotypes with many additional singleton, doubleton or other low-frequency haplotypes (Fig. 2). The common alleles were geographically broadly distributed. With few exceptions, the rare alleles were separated from the common alleles and from each other by a single mutational step. Accordingly, mismatch distribution analysis of *cyt b* was consistent with a recent bottleneck (Harpending's Raggedness index = 0.095, p = 0.010; Fig. 3). The more variable *mtCR* did not show the Poisson-like distribution that signals a bottleneck, yielding instead a smooth normal distribution, consistent with the bottleneck being relatively recent but not persisting, suggesting the most recent population expansion is reasonably well advanced (Harpending's Raggedness index = 0.008, p = 0.830; Fig. 3). Fu's FS and Tajima's D of both *mtCR* and *cyt b*, for larger samples and when all samples were included, were significantly negative and consistent with historical population expansion (Table 1). Pairwise Φ_{ST} values indicated genetic differentiation of peripheral locations in the western Indian Ocean (Seychelles) and western Pacific (Pohnpei) although there appears to be little genetic differentiation among most samples collected across most of the range (Table 2).

Microsatellite analyses

We detected 96 alleles at the 10 polymorphic microsatellite loci in 275 Humphead wrasse from 16 areas (Supplementary Document S1 Table S2). Null alleles were detected in loci D118 and D123 in only three samples (Supplementary Document S1 Table S3). Nonetheless, F_{ST} estimations by FreeNA with and without ENA were very similar, with extensive overlap in their 95% confidence intervals (Supplementary Document S1 Table S3). Thus, genotypes of all loci in all populations were included in subsequent analyses. Observed heterozygosity (H_o) ranged from 0.000 to 1.000, while expected heterozygosity (H_e) ranged from 0.000 to 0.896 (Supplementary Document S1 Table S2). Overall, H_o and H_e were 0.670 and 0.705, respectively, and did not deviate from HWE. Significant deviation from HWE was only detected in two loci (D12 and D111) in population IDE and in one locus (D118) in population IN (Supplementary Document S1 Table S2). No significant LD was detected among any loci pair (data not shown).

DAPC analysis defined K = 6 based on the lowest BIC value. While clusters 1-4 and 6 did not show any geographical trend, almost all samples from Seychelles (AL and FI) were assigned to cluster 5 (Fig. 4). Delta K (Evanno Method) for STRUCTURE analysis was highest at K = 4 (Supplementary Document S1 Fig. S1). Similar to DAPC analysis, the STRUCTURE assignment plot indicated that the Seychelles populations were distinct from the remaining populations, among which no differentiation was evident (Fig. 5).

Congruently, pairwise F_{ST} and R_{ST} estimates suggested that the Seychelles population (FI) was genetically distinct from other populations with more than 10 individuals ($F_{ST} > 0.093$, BH corrected $P < 0.01$ in 7/8 pairwise comparisons, Table 3; $R_{ST} > 0.156$, BH corrected $P < 0.01$ in 5/8 pairwise comparisons, Table 4). Significant genetic differentiation was also detected between IR and ID, and between IR and PH based on F_{ST} (0.021, BH corrected $P < 0.01$) but not based on R_{ST} (Tables 3, 4).

Tests of biogeographic hypotheses

We found mixed support for large-scale regional genetic partitions (1) between the Western Indo-Pacific and the Central Indo-Pacific, and (2) between Western Indian Ocean and the rest of the Indo-Pacific in mitochondrial DNA, but not in microsatellite data. We found no support for (3) genetic partitions among all marine biogeographic provinces covered in this study except for locus-by-locus AMOVA of microsatellite data (Table 5). Isolation by distance tests did not reveal significant correlation between genetic and geographical distances, although the result was only marginally non-significant for microsatellite data (Table 6).

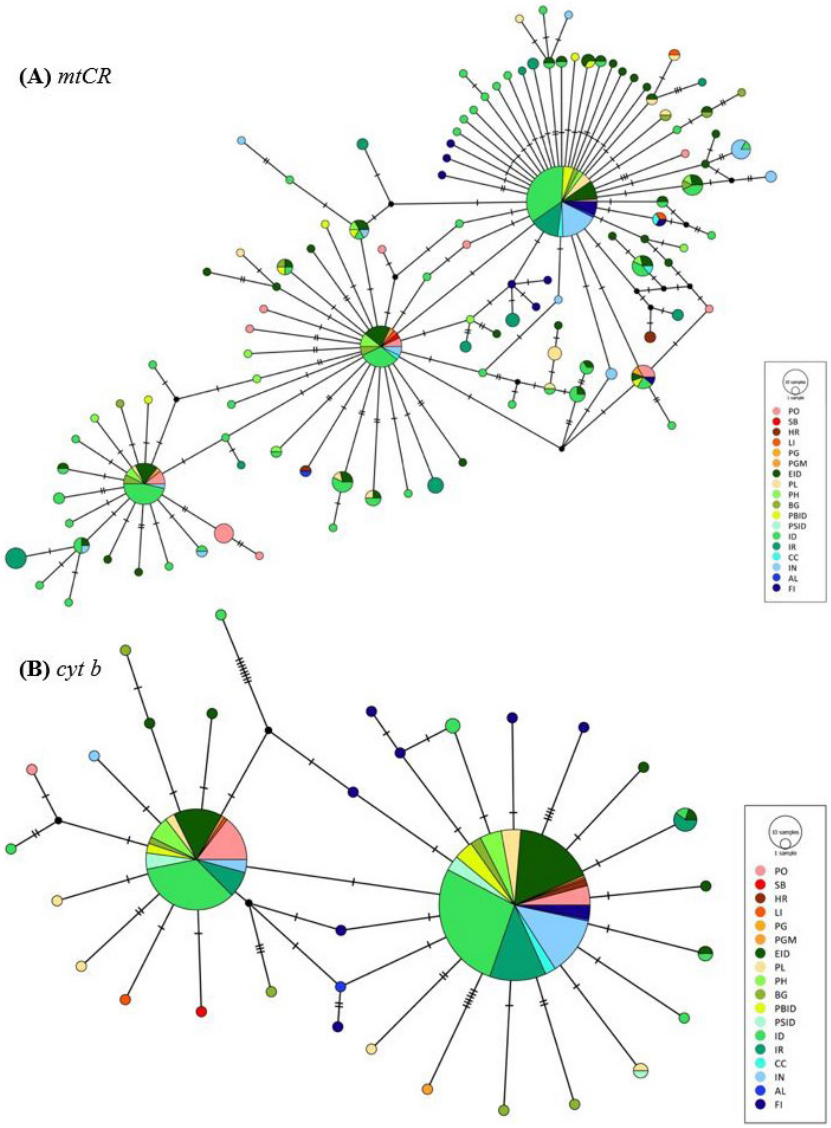

Figure 2. Haplotype networks of **(A)** control region and **(B)** cytochrome b sequences from Humphead wrasse, *Cheilinus undulatus*, across the Indo-West Pacific. Each circle denotes one haplotype, with size proportional to number of individuals and colour representing sampling locality. Each cross-hatch on a branch denotes one mutational step. See Table 1 and Fig. 1 for location codes and positions.

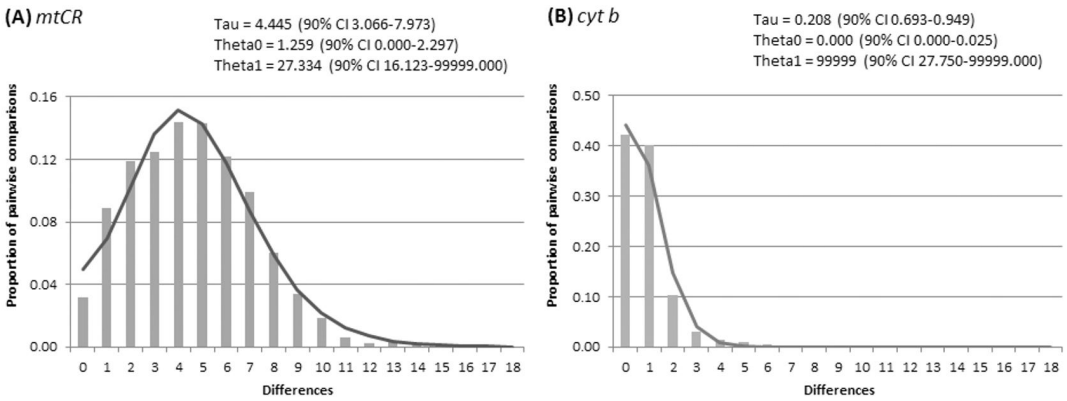

Figure 3. Empirical mismatch distributions (bars) for Humphead wrasse using **(A)** control region and **(B)** cytochrome b sequences from all samples combined. The expected distribution based on model of sudden population expansion is shown as a line. Note different y-axis scales.

Table 2. Genetic differentiation of Humphead wrasse, *Cheilinus undulatus*, at locations across the Indo-West Pacific, represented by pairwise ϕ_{ST} calculated using mitochondrial **(A)** control region and **(B)** cytochrome b datasets calculated using samples with ≥ 10 individuals. Results are indicated at two levels of significance: P < 0.01 (bold) and P < 0.05 (italic) after Benjamini–Hochberg correction for multiple tests. Comparisons that are significant for both markers are additionally underlined.

(A) *mtCR*	**PO**	**EID**	**PL**	**PH**	**BG**	**PBID**	**PSID**	**ID**	**IR**	**IN**
EID	**0.068**									
PL	**0.130**	0.033								
PH	0.064	0.003	0.058							
BG	0.045	-0.122	0.050	-0.071						
PBID	0.110	-0.024	0.021	0.024	0.002					
PSID	0.028	-0.044	0.044	-0.037	-0.010	0.039				
ID	**0.071**	0.000	*0.049*	0.002	-0.054	0.006	-0.005			
IR	**0.110**	**0.054**	**0.094**	0.040	0.002	0.064	0.059	**0.039**		
IN	**0.150**	*0.030*	*0.074*	*0.068*	-0.074	0.011	0.040	**0.041**	**0.082**	
FI	**0.241**	*0.056*	*0.109*	**0.143**	**0.118**	*0.069*	**0.182**	**0.082**	**0.118**	0.049
(B) *cyt b*	**PO**	**EID**	**PL**	**PH**	**BG**	**PBID**	**PSID**	**ID**	**IR**	**IN**
EID	*0.134*									
PL	0.058	0.007								
PH	0.035	-0.012	-0.018							
BG	0.027	0.059	-0.003	0.010						
PBID	0.105	-0.046	-0.035	-0.016	-0.001					
PSID	0.035	-0.022	-0.038	-0.060	-0.009	-0.037				
ID	*0.079*	-0.003	0.000	-0.029	0.043	-0.032	-0.033			
IR	**0.186**	-0.002	0.020	0.044	0.078	-0.041	0.023	0.014		
IN	**0.239**	0.016	0.038	0.109	0.103	-0.031	0.078	0.032	-0.003	
FI	**0.215**	*0.136*	0.054	*0.131*	0.043	0.039	0.092	*0.136*	*0.100*	*0.082*

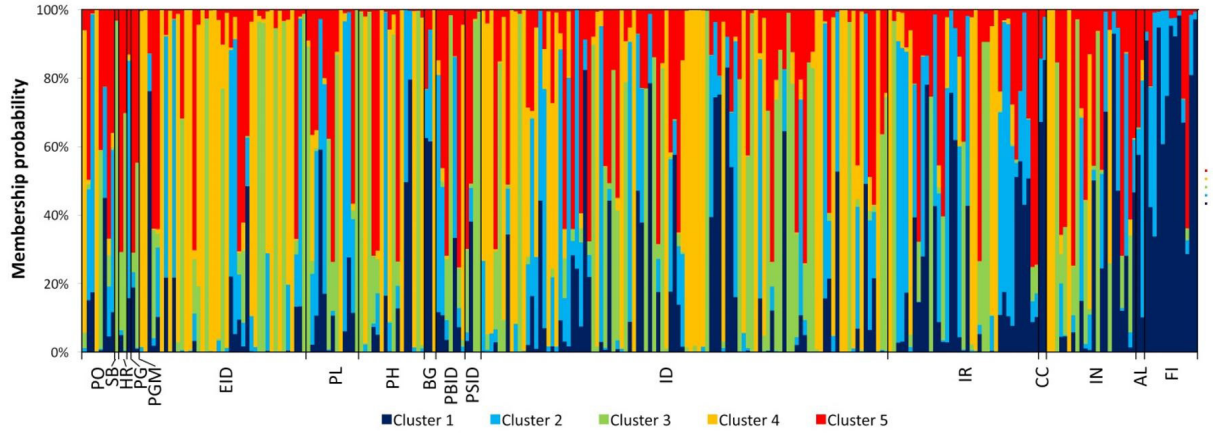

Figure 4. Discriminant analysis of principal components (DAPC) assignment plots based on 10 microsatellite loci with five clusters defined indicates eurymixis, i.e. generally extensive yet incomplete mixing, of Humphead wrasse samples across the Indo-West Pacific.

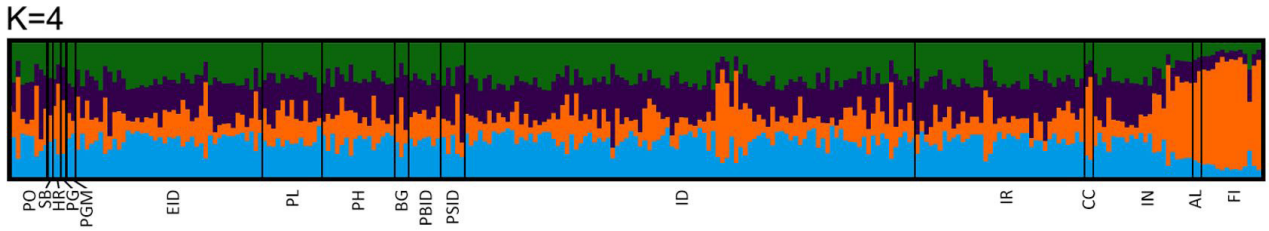

Figure 5. STRUCTURE assignment plots based on 10 microsatellite loci with K=4 indicates generally broad mixing (or lack of differentiation) of Humphead wrasse samples across the Indo-West Pacific, with the possible exception of sites in the Indian Ocean, particularly the Seychelles.

Table 3. Microsatellite-based estimates of F_{ST} calculated for all pairwise comparisons of locations for which samples of size ≥ 10 Humphead wrasse were available. Results are indicated at two levels of significance: P < 0.01 (bold) and P < 0.05 (italic) after Benjamini–Hochberg correction for multiple tests.

	EID	PL	PH	ID	IR	IN
PL	-0.034					
PH	-0.047	-0.123				
ID	-0.041	-0.114	-0.005			
IR	-0.033	-0.112	*0.021*	**0.021**		
IN	0.010	-0.046	-0.055	-0.045	-0.044	
FI	**0.122**	0.032	**0.130**	**0.117**	**0.128**	**0.061**

Table 4. Microsatellite-based estimates of R_{ST} calculated for all pairwise comparisons of locations for which samples of size ≥ 10 Humphead wrasse were available. Results are indicated at two levels of significance: P < 0.01 (bold) and P < 0.05 (italic) after Benjamini–Hochberg correction for multiple tests.

	EID	PL	PH	ID	IR	IN
PL	-0.009					
PH	-0.023	-0.037				
ID	-0.019	-0.044	-0.011			
IR	-0.053	-0.007	0.019	*0.020*		
IN	0.010	-0.006	-0.032	-0.027	-0.044	
FI	**0.169**	**0.217**	**0.215**	**0.170**	*0.069*	*0.171*

Table 5. Regional genetic differentiation of Humphead wrasse, *Cheilinus undulatus*, across the Indo-West Pacific. Analysis of Molecular Variance (AMOVA) of three hypotheses: differentiation between the Western Indo-Pacific and the Central Indo-Pacific (WIO + CIO vs CIP), between Western Indian Ocean and the rest of the Indo-Pacific (WIO vs CIO + CIP), and among all marine biogeographic provinces covered in this study (LME) (see Fig. 1). Results are indicated at two levels of significance: P < 0.01 (bold) and P < 0.05 (italic).

Partition scheme	Dataset (method)	F_{SC}	P	F_{ST}	P	F_{CT}	P
WIO + CIO vs others	Control region	0.023	**0.004**	0.055	**0.000**	0.033	0.086
	Cytochrome b	0.058	**0.009**	0.092	**0.002**	0.036	0.118
	Microsatellite Rst	-0.014	0.973	0.013	0.815	0.026	0.111
	Microsatellite Rst (lbl)	0.024	*0.045*	0.049	*0.020*	0.026	0.106
WIO vs Others	Control region	0.028	**0.000**	0.068	**0.000**	0.042	0.109
	Cytochrome b	0.055	**0.009**	0.176	**0.003**	0.129	0.062
	Microsatellite Rst	-0.028	1.000	0.127	0.814	0.151	*0.012*
	Microsatellite Rst (lbl)	0.006	0.344	0.182	*0.020*	0.177	**0.007**
LME	Control region	-0.021	0.907	0.052	**0.000**	0.072	**0.002**
	Cytochrome b	0.035	0.757	0.080	**0.002**	0.047	*0.039*
	Microsatellite Rst	-0.032	1.000	0.005	0.814	0.036	0.212
	Microsatellite Rst (lbl)	0.000	*0.016*	0.043	0.590	0.044	**0.005**

lbl: locus by locus AMOVA

F_{SC} within populations versus regions; F_{ST} within populations versus whole species; F_{CT} within groups versus whole species.

Table 6. Locus-specific tests of isolation by distance in Humphead Wrasse, comparing genetic distance against log (least cost distance).

Dataset	Mantel statistic r	Significance
Control region	0.272	0.159
Cytochrome b	0.107	0.353
Microsatellite Rst	0.608	0.055
Microsatellite Fst	0.572	0.083

Discussion

Phylogeographic structure

Despite the complex biogeographic history and modern oceanography of the Indo-Pacific region, our analyses indicate that Humphead wrasse is relatively genetically unstructured over large portions of its range. We interpret the occasional signals of regional, local, and distance-based differentiation as being most consistent with a single eurymictic phylogeographic

unit encompassing locations from the western Indian Ocean to the western Pacific Ocean, barring perhaps some geographic differentiation at the peripheries. Both the most westerly—the Seychelles (in the Indian Ocean)—and the most easterly—Pohnpei (in the Pacific)—samples appear weakly, but generally, differentiated by mitochondrial and/or nuclear markers. The species does not evince a strong genetic response to historical barriers to dispersal from low sea-level stands (e.g., Bowen et al. 2016), nor reflects major marine biogeographic partitioning (e.g., Spalding et al. 2007), in contrast to species with life-histories predisposed to lower dispersal potential (e.g., *Amphiprion ocellaris* which does not have a dispersive egg phase; Timm et al. 2008). *C. undulatus* is not unique in this regard as fishes from a variety of families—including moray eels, snappers, soldierfish, and surgeonfish, all of which have dispersive egg phases—show modest to no genetic differentiation across large distances in the Indo-West Pacific (Craig et al. 2007, Horne et al. 2008, Gaither et al. 2010, Reece et al. 2011). Some, including brown surgeonfish (*Acanthurus nigrofuscus*), like Humphead wrasse show evidence of differentiation between eastern and western Indian Ocean locations (e.g., Eble et al. 2011).

Several factors may have contributed to the observed low population genetic differentiation in *C. undulatus*, including historical demography and contemporary gene flow attributable to long pelagic duration and reproductive behavior. First, historical population expansion is indicated by the significantly negative Fu's FS and Tajima's D values for both *mtCR* and *cyt b* datasets. The mismatch distribution for *cyt b* also retains the signature of an historical bottleneck, although the mismatch distribution of the faster evolving *mtCR* suggests that population expansion is now largely complete. Thus, while historical population expansion may contribute to a lack of regionalization in *C. undulatus*, it is unlikely the sole cause. Estimates of very low pairwise population differentiation, using the rapidly evolving microsatellites, are consistent with ongoing modern high connectivity. This inferred modern dispersal appears to be connecting almost all sites studied, with the possible exception of the Seychelles and Pohnpei, the most distant island sites in our dataset.

Ongoing connectivity across the majority of the species' range is likely attributable in part to the 4–6 week pelagic larval phase of *C. undulatus*. Although there has been considerable debate calling into question whether PLD can explain genetic connectivity in marine taxa (Weersing and Toonen 2009, Selkoe and Toonen 2011), those studies were flawed conceptually (Dawson 2014b). Both modelling (Treml et al. 2012) and empirical (Dawson et al. 2014) studies show that broad-scale connectivity is strongly influenced by the length of PLD in addition to other factors influencing dispersal, such as reproductive output (as a function of abundance and fecundity of adults). Thus, adult traits such as body size, schooling behavior, and nocturnal activity also are correlates of connectivity (Luiz et al. 2013) as components of 'dispersal syndromes' (Dawson 2014a, Schiebelhut and Dawson 2018). Interestingly, the reproductive behavior of *C. undulatus*, as observed in Palau, is to form resident spawning aggregations after high tide nearly every day of the year at a large number of specific locations along the seaward edge of the barrier reef (Colin 2010) adjacent to deep ocean. Spawning at the mid-stage of falling tide favors offshore transport of eggs away from the reef where they might then be either retained near the reef for later return and settlement or advected into offshore circulation (Colin 2010). The large number of yearly spawning events, occurring over a variety of wind and sea conditions, and if typical of the species, implies ample opportunity for both local retention and episodic entrainment of eggs and larvae into oceanic current systems facilitating long distance dispersal. Other reef fishes, such as the camouflage grouper *Epinephelus polyphekadion* and the common coral trout *Plectropomus leopardus*, which spawn in a more restricted lunar and yearly time frame, i.e., short duration spawning aggregations, exhibit more marked population structure (Ma et al. 2018) than *C. undulatus*.

Thus, it is possible that local variations in adult abundance, spawning behavior, and oceanography may contribute to weak genetic differentiation — which is most notable in the Seychelles, SW Indian Ocean — across the range of *C. undulatus* despite the potential for distant larval dispersal. The two Seychelles sites were the most isolated of our study sites, separated from other suitable habitats by strong equatorial currents and upwelling off Yemen, Oman, and Somalia (Briggs and Bowen 2012, Kemp 1998). Intra-specific genetic differentiation of the Southwest Indian Ocean is also observed in some other reef fish inhabiting this region, including *Epinephelus fasciatus* (Borsa et al. 2016), *E. merra* (Muths et al. 2014), and *Acanthurus nigrofuscus* (Eble et al. 2011), and the region has high endemism of marine fauna in general (Briggs and Bowen 2012, Bowen et al. 2016).

Conservation and management implications

Our data indicate that Humphead wrasse constitutes a widespread metapopulation with extensive but incomplete mixing, leading to sporadic differentiation of local sub-units, most often toward the periphery. As such, the conservation status of the species depends on the successful management of a sufficient but currently unknown number and distribution of populations in a suitably designed, multi-national, range-wide network. While locations at the extreme edges of the geographic range of the species may harbor unique genetic diversity, from a conservation and management perspective the most urgent need is to understand dynamics towards the centre of its range, within an area from the central Indian Ocean to the western Pacific, and particularly in Southeastern Asia where its reef habitat is extensive.

The most intense fishing pressure on populations of the Humphead wrasse in recent decades appears to be from the international live reef food fish trade for which the main exporting country currently is Indonesia

(Wu and Sadovy de Mitcheson 2016). Indonesia developed an export quota under CITES Appendix II that it considers to be sustainable (Sadovy et al. 2007); this was less than 2,000 fish exported for the whole of 2017, with the same quota for 2018. Some of the trade out of Indonesia and into Hong Kong and mainland China, the major importers, however, is illegal and in excess of the quota (Wu and Sadovy de Mitcheson 2016), which implies that some of the supplying fisheries are likely to be unsustainable. Studies clearly indicate that overfishing is still occurring despite an export quota and national management measures in place for the species (Sadovy de Mitcheson 2015, Sadovy de Mitcheson et al. 2019), although there are some signs of recovery after the CITES listing. Field studies in eight locations over nine years show that, wherever fished, there are few adults, hardly any over 80 cm (i.e., males) and only small numbers of juveniles present (Sadovy de Mitcheson et al. 2019). Adults and abundance levels closer to unexploited levels were only found where the species is unfished or where fishing stopped although recruitment was noted in both lightly and medium fishing pressure sites. In intensively fished places, densities are very low and most fish are juveniles. Indonesia includes almost 20% of Indo-Pacific coral reefs and, hence, is particularly important for considering the implications of metapopulation structure on the exploitation of this coral-reef dependent fish. The current study indicates that the Indonesian populations are likely well connected with other western Pacific populations, but the degree of self-recruitment to any or a set of populations and possible source-sink relationships cannot be inferred from our analyses.

The importance of understanding population structure and larval sourcing becomes clear when considering conservation and management measures for this species. The Humphead wrasse's planktonic stages have potential to move considerable distances. This means that a protected area or population would need to be self-recruiting and/or receive recruits from elsewhere to replace its own population and to persist; recruits could travel from reefs several hundred kilometres away, or even further. If metapopulation dynamics are the norm for *C. undulatus,* with fish recruitment at least partially dependent on distant reefs, then reefs depleted of large fish can benefit from distant MPAs and sustainable management practices which may require cooperation between several countries. On the other hand, if localised recruitment is common—e.g., due to oceanography around islands—MPA's will be more likely to self-recruit but distant depleted fish stocks will not benefit from the egg supply produced by protected areas. In both cases, MPAs need to be properly sited, oceanographically, in suitable adult and juvenile habitat (which can include coral reef areas and specific settlement habitats such as *Sargassum* spp.), large enough (at least 20 km of reef is recommended; Green et al. 2015, Sadovy de Mitcheson et al. 2019), and where spawning occurs, especially if these are important sources of eggs and larvae in the region, to conserve this wide-ranging species, and to appropriately manage fishing pressure where the species is exploited.

Conclusions

Our molecular analyses indicate little genetic differentiation over a major portion of the geographic range of *C. undulatus*. While seriously overexploited across parts of its range, hence its presence on the IUCN Red List and CITES Appendix II listing, there is little indication that important endemic genetic diversity is presently at risk from local extirpation of any given population (at least considering the areas where our samples were sourced). The genetic connectivity found for the species over the Indo-West Pacific region sampled is, in this sense, comforting. However, the genetic signals of occasional local and peripheral isolation refute panmixia, indicating that safeguarding genetic diversity of this conservation-dependent species throughout its natural range will require a more detailed understanding of population structure, particularly in regions where it continues to be exploited. Given that the species has largely disappeared due to overfishing in at least one edge-of-range area—i.e., the northern sector of the South China Sea—studies from other edge-of-range sites may also be a priority if these areas are exploited or likely to become so.

Maintaining ecological connectivity, wherein adults in another region may serve to repopulate an overexploited area may be tenuous given serious population declines in some areas. Identification of key spawning and settlement/nursery areas (sources, stepping stones, and sinks) remains an important goal to ensure that suitable habitat, in the right locations, is protected for ecological connectivity. The pelagic larval duration of *C. undulatus* does not appear to differ greatly from that of many other reef fishes, including other members of the Labridae (Victor 1986). However, potentially low reproduction in many areas where males and mature females are rare or at low densities and the intensity with which the species is fished are reason for concern. This depredation, particularly evident in the central portion of its range, portends a potential future breakdown in genetic continuity if reproducing populations are lost. Numerous studies have indicated the very slow recovery of overexploited populations and therefore the need to manage the species at local levels to maintain ecologically viable populations (e.g., Hutchings and Reynolds 2004).

Thus, there is need for more detailed examination of the genetic structure of *C. undulatus* across its entire range, particularly in key central areas such as the South China Sea and Indonesia—which are under-sampled or lack geographic resolution in our study and where the species is most heavily exploited—and at all extremes of its geographic range from where this study did not obtain samples. The extent to which any population or set of populations may be self-recruiting and hence possibly isolated from range-wide collapse and a target for local management cannot be inferred from our analyses and requires higher resolution genomic markers, greater geographic sampling, and more consistently larger sample sizes that will enable robust

assignment tests on ecological-genetic time-scales. Similarly, the possible existence of important source areas, seeding surrounding or even distant regions, some of which are otherwise depleted, should be addressed in future work. Given the endangered and yet still exploited status of Humphead wrasse, such work is imperative.

Overexploitation of *C. undulatus* is having direct local impacts on food security and tourism, and it could also have indirect impacts on sustainability of *C. undulatus* populations far afield through diminished recruitment. Other information available on the condition and responses of this species in the central area of its range, Indonesia, highlight the need to control fishing levels on the species to safeguard the reproductive capacity of the population. It is also important to better understand and protect important nursery areas and implement large enough and suitably placed no-take MPAs to safeguard suitable reef habitat, spawning areas, and the biogeographic history that is recorded in the genetics of this iconic and enigmatic species.

Acknowledgements

Funding was obtained for field work in Indonesia from the CITES Secretariat and from the National Oceanographic and Atmospheric Administration of the United States to fund the work of the IUCN Groupers and Wrasses Specialist Group in relation to the CITES listing of the species, with additional support within Indonesia from The Indonesian Institute of Sciences (LIPI). Fieldwork in Palau and molecular analyses were funded by the National Fish and Wildlife Foundation (#2001-0336-015) to PLC. We thank the Bureau of Marine Resources of the Palau National Government for support and permission to carry out the work in Palau (BMR Research Permits RE-02-07 and RE-03-03). Additional samples were generously provided by J.H. Choat, D. DeMaria, T. Graham, S. Newman, N. Pilcher, and K.L. Rhodes. Many of the sequence and microsatellite data were generated by S. Abboud, J. Baumsteiger, G. Bernardi, M. Ramon, and L. Schiebelhut. Additional support was provided by the Society for the Conservation of Reef Fish Aggregations (SCRFA), which was funded by the David & Lucile Packard Foundation.

Supplementary Information

The following materials are available as part of the online article from https://escholarship.org/uc/fb:

Supplementary Document S1. Additional information on sampling and results.

Supplementary Document S2. Additional details of molecular analyses.

Supplementary Document S3. Humphead wrasse microsatellite dataset.

References

Aw, M. (2000) The vanishing Napoleon. Asia Geographic, 1, 70–81.

Benzie, J.A.H. (1998) Genetic structure of marine organisms and SE Asian biogeography. In Biogeography and Geological Evolution of SE Asia (ed. by R. Hall and J.D. Holloway). Pp. 197–209. Backhuys Publishers, Leiden.

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

Borsa, P., Durand, J.-D., Chen, W.-J., Hubert, N., Muths, D., Mou-Tham, G. & Kulbicki, M. (2016) Comparative phylogeography of the western Indian Ocean reef fauna. Acta Oecologica, 72, 72–86.

Bowen, B.W., Gaither, M.R., DiBattista, J.D., Iacchei, M., Andrews, K.R., Grant, W.S., Toonen, R.J., & Briggs, J.C. (2016) Comparative phylogeography of the ocean planet. Proceedings of the National Academy of Sciences USA, 113, 7962–7969.

Briggs, T. (2016). Where to scuba dive with Napoleon Wrasse. https://www.sportdiver.com/where-to-dive-with-napoleon-wrasse

Briggs, J.C. & Bowen, B.W. (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. Journal of Biogeography, 39, 12–30.

Cabral, R.B., Gaines, S.D, Lim, M.T., Atrigenio, M.P., Mamauag, S.S., Pedemonte, G.C. & Aliño, P.M. (2016). Siting marine protected areas based on habitat quality and extent provides the greatest benefit to spatially structured metapopulations. Ecosphere, 7, e0153315.

Carvalho, S.B., Velo-Antón, G., Tarroso, P., Portela, A.P., Barata, M., Carranza, S., Moritz, C. & Possingham, H.P. (2017) Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. Nature Ecology and Evolution, 1, 0151.

Chapuis, M.-P. & Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution, 24, 621–631.

Choat, J.H., Davies, C.R. Ackerman, J.L. & Mapstone, B.D. (2006) Age structure and growth in a large teleost, *Cheilinus undulatus*, with a review of size distribution in labrid fishes. Marine Ecology Progress Series, 318, 237–246

Colin, P.L. (2010) Aggregation and spawning of the Humphead wrasse *Cheilinus undulatus* (Pisces: Labridae): general aspects of spawning behaviour. Journal of Fish Biology, 76, 987–1007.

Courchamp, F., Angulo, E., Rivalan, P., Hall, R.J., Signoret, L., Bull, L. & Meinard, Y. (2006) Rarity value and species extinction: the anthropogenic Allee effect. PLoS Biology, 4, e415.

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

Dawson, M.N. (2014a) Biogeography and complex traits: dispersal syndromes, in the sea. Frontiers of Biogeography, 6, 11–15.

Dawson, M.N. (2014b) Natural experiments and meta-analyses in comparative phylogeography. Journal of Biogeography, 41, 52–65.

Dawson, M.N, C.G. Hays, R.K. Grosberg, & P.T. Raimondi (2014) Dispersal potential and population genetic structure in the marine intertidal of the eastern North Pacific. Ecological Monographs, 84, 435–456.

Earl, D.A. & vonHoldt, B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4, 359–361.

Eble, J.A., Rocha, L.A., Craig, M.T. & Bowen, B.W. (2011) Not all larvae stay close to home: insights into marine population connectivity with a focus on the brown surgeonfish (*Acanthurus nigrofuscus*). Journal of Marine Biology, 2011, 518516.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32, 1792–1797.

Excoffier, L. & Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10, 564–567.

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

Gillett, R. (2010) Monitoring and management of the Humphead wrasse, *Cheilinus undulatus*. FAO Fisheries and Aquaculture Circular. No. 1048. Rome, FAO. 2010. 62 pp.

Graham, K.S., Boggs, C.H., DeMartini, E.E., Schroeder, R.E., & Trianni, M. S. (2014). Status review report: humphead wrasse (*Cheilinus undulatus*). Report to National Marine Fisheries Service, Office of Protected Resources. September, 2014. 123 p. + Appendices. Available at https://www.researchgate.net/publication/271444001_Status_Review_Report_Humphead_Wrasse_Cheilinus_undulatus

Green, A.L., Maypa, A.P., Almany, G.R., Rhodes, K.L., Weeks, R., Abesamis, R.A., Gleason, M.G., Mumby, P.J. & White, A.T. (2015) Larval dispersal and movement patterns of coral reef fishes, and implications for marine reserve network design. Biological Reviews, 90, 1215–1247.

Hamner, W.M. & Largier, J.L. (2012). Oceanography of the planktonic stages of aggregation spawning reef fishes, pp. 159-190. In: Y. Sadovy de Mitcheson & P.L. Colin (eds.) Reef Fish Spawning Aggregations: Biology, Research and Management, Fish & Fisheries Series 35. Pp.621. Springer Scienced+Business Media

Hirai, N., Koiso, M., Teruya, K., Kobayashi, M., Takebe, T., Sato, T., Okuzawa, K. & Hagiwara, A. (2013). Success of seed production of Humphead wrasse *Cheilinus undulatus* with improvement of spawning induction, feeding, and rearing conditions. Hatchery Technology of High Quality Juvenile Production, Honolulu, Hawaii. 107–111.

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

Hu, J., Zhu, X. P., Luo, J., Yin, S. W., Peng, Y. H., Hu, Y. L., & Zhu, F. (2013). Development and characterization of microsatellite loci in a threatened marine fish, *Cheilinus undulatus* (humphead wrasse). Genetics and Molecular Research, 12, 2633–2636.

Hutchings, J.A. & Reynolds, J.D. (2004) Marine fish population collapses: Consequences for recovery and extinction risk. BioScience, 54, 297–300.

Indriatmoko, Syam, A.R. & Syahputra, K. (2016) Control region-mitochondrial partial DNA analysis of Humphead wrasse [*Cheilinus undulatus* (Ruppel, 1835)] from Anambas Islands, Indonesia. Aquatic Procedia, 7, 125–131.

Jombart, T., Devillard, S. & Balloux, F. (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics, 11, 94.

Kelly, R.P. & Eernisse, D.J. (2007) Southern hospitality: a latitudinal gradient in gene flow in the marine environment. Evolution, 61, 700–707.

Kemp, J.M. (1998) Zoogeography of the coral reef fishes of the Socotra Archipelago. Journal of Biogeography, 25, 919–933.

Kindsvater, H., Reynolds, J., Sadovy de Mitcheson, Y. & Mangel, M. (2017) Selectivity matters: rules of thumb for management of plate-sized, sex-changing fish in the live reef food fish trade. Fish and Fisheries, 18, 821–836.

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

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., & Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources, 15, 1179–1191.

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

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

Leigh, J.W. & Bryant, D. (2015) popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution, 6, 1110–1116.

Lester, S.E., Gaines, S.D. & Kinlan, B.P. (2007) Reproduction on the edge: large-scale patterns of individual performance in a marine invertebrate. Ecology, 88, 2229–2239.

Lindenfelser, M.E. (1984) Morphometric and allozymic congruence: evolution in the prawn *Macrobranchium rosenbergii* (Decapoda, Palaemonidae). Systematic Zoology, 33, 195–204.

Luiz, O.J., Allen, A.P., Robertson, D.R., Floeter, S.R., Kulbicki, M., Vigliola, L., Becheler, R. & Madin, J.S. (2013) Adult and larval traits as determinants of geographic range size among tropical reef fishes. Proceedings of the National Academy of Sciences USA, 110, 16498–16502.

Ma, K.Y., Van Herwerden, L., Newman, S.J., Berumen, M.L., Choat, J.H., Chu, K.H. & de Mitcheson, Y. S. (2018). Contrasting population genetic structure in three aggregating groupers (Percoidei: Epinephelidae) in the Indo-West Pacific: the importance of reproductive mode. BMC Evolutionary Biology, 18, 180.

Matthew, P., Manjaji-Matsumoto, B.M. & Rodrigues, K.F. (2018) Complete mitochondrial genome of six *Cheilinus undulatus* (Napoleon Wrasse): an endangered marine fish species from Sabah, Malaysia. Mitochondrial DNA Part B, 3, 943–944.

McManus, J.W. (1994) The Spratly Islands: A marine park. Ambio, 23, 181–186.

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

Muths, D., Tessier, E. & Bourjea, J. (2014) Genetic structure of the reef grouper *Epinephelus merra* in the West Indian Ocean appears congruent with biogeographic and oceanographic boundaries. Marine Ecology, 36, 447–461.

Myers, R.F. (1991) Micronesian reef Fishes. Coral Graphics, Guam. 298 pp.

Nei, M. (1987) Molecular Evolutionary Genetics. Columbia University Press, New York.

Oksanen J., Kindt R., Legendre P., O'Hara B., Simpson G.L., Stevens M.H.H. & Wagner H. (2008) Vegan: Community Ecology Package. R package version 1.15-0. http://cran.r-project.org/, http://vegan.r-forge.r-project.org.

Pante, E. & Simon-Bouhet, B. (2013) marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. PLoS One, 8, e73051.

Peakall, R. & Smouse, P.E. (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics, 28, 2537–2539.

Poh, T-M & Fanning, L.M. (2012). Tackling illegal, unregulated, and unreported trade towards Humphead wrasse (*Cheilinus undulatus*) recovery in Sabah, Malaysia, Marine Policy, Elsevier, 36, 696–702.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. Genetics, 155, 945–959.

Qi, X.Z., Yin, S.W., Luo, J. & Huo, R. (2013) Complete mitochondrial genome sequence of the humphead wrasse, *Cheilinus undulatus*. Genetics and Molecular Research, 12, 1095–1105.

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

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

Rousset, F. (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources, 8, 103–106.

Sadovy, Y., Kulbicki, M., Labrosse, P., Letourneur, Y., Lokani, P. & Donaldson, T.J. (2003) The Humphead wrasse, *Cheilinus undulatus*: synopsis of a threatened and poorly known giant coral reef fish. Reviews in Fish Biology and Fisheries, 13, 327–364.

Sadovy, Y., Punt, A.E., Cheung, W., Vasconcellos, M. & Suharti, S. (2007) Stock assessment approach for the Napoleon fish, *Cheilinus undulatus*, in Indonesia: a tool for quota-setting for data-poor fisheries under CITES Appendix II Non-Detriment Finding requirements. FAO Fisheries Circular. No. 1023 Rome, FAO, 71 pp.

Sadovy de Mitcheson, Y., Liu, M. & Suharti, S. (2010) Gonadal development in a giant threatened reef fish, the humphead wrasse *Cheilinus undulatus*, and its relationship to international trade. Journal of Fish Biology, 77, 706–718.

Sadovy de Mitcheson, Y., Tam, I., Muldoon, G., le Clue, S., Botsford, E. & Shea, S. (2017) The trade in live reef food fish – going, going, gone. Volume 1: Main Report. Parts I, II & III, pp.1–288. ADM Capital Foundation and The University of Hong Kong, Hong Kong Special Administrative Region.

Sadovy de Mitcheson, Y., Suharti, S. & Colin, P.C. (2019) Quantifying the Rare: baselines for the endangered Napoleon Wrasse, *Cheilinus undulatus,* and implications for conservation. Aquatic Conservation; Marine and Freshwater Ecosystems, doi:10.1002/aqc.3124

Schiebelhut, L.M. & Dawson, M.N (2018) Differences in life-history and dispersal potential correlate with differences in population differentiation in terrestrial and marine environments. Journal of Biogeography, 45, 2427–2441.

Selkoe, K.A. & Toonen, R.J. (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Marine Ecology Progress Series, 436, 291-305.

Spalding, M.D., Fox, H.E., Allen, G.R., et al. (2007) Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. Bioscience, 57, 573–583.

Timm, J., Figiel, M. & Kochzius, M. (2008) Contrasting patterns in species boundaries and evolution of anemonefishes (Amphiprioninae, Pomacentridae) in the centre of marine biodiversity. Molecular Phylogenetics and Evolution, 49, 268–276.

Treml, E.A., Roberts, J.J., Chao, Y., Halpin, P.N., Possingham, H.P. & Riginos, C. (2012) Reproductive output and duration of the pelagic larval stage determine seascape-wide connectivity of marine populations. Integrative and Comparative Biology, 52, 525–537.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., & Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535–538.

Victor, B.C. (1986) Duration of the planktonic larval stage of one hundred species of Pacific and Atlantic wrasses (family Labridae). Marine Biology, 90, 317–326.

Weersing, K. & Toonen, R.J. (2009) Population genetics, larval dispersal, and connectivity in marine systems.Marine Ecology Progress Series, 393, 1–12.

Wu, J. & Sadovy de Mitcheson, Y. (2016) Humphead (Napoleon) Wrasse *Cheilinus undulatus* trade into and through Hong Kong. TRAFFIC. Hong Kong, SAR. 42 pp.

Submitted: 07 February 2019
First decision: 08 April 2019
Accepted: 23 April 2019

Edited by Chris Burridge

