Local genetic patchiness but no regional differences between Indo-West Pacific populations of the dogtooth tuna Gymnosarda unicolor

B. P. Bentley¹, E. S. Harvey², S. J. Newman²,³, D. J. Welch⁴,⁵, A. K. Smith⁶, W. J. Kennington¹,*

¹Centre for Evolutionary Biology, School of Animal Biology (M092), The University of Western Australia, Crawley, WA 6009, Australia
²Department of Environment and Agriculture, Curtin University, Bentley, WA 6102, Australia
³Western Australian Fisheries and Marine Research Laboratories, Department of Fisheries, Government of Western Australia, PO Box 20, North Beach, WA 6920, Australia
⁴Centre for Sustainable Tropical Fisheries and Aquaculture, James Cook University, Townsville, QLD 4811, Australia
⁵C2O Fisheries, Woolgoolga, NSW 2456, Australia
⁶PO Box 1379, Townsville, QLD 4810, Australia

ABSTRACT: Physical barriers can have profound impacts on dispersal in marine species. Here, we investigate population structure and levels of relatedness among individuals of the coral reef associated fish, the dogtooth tuna Gymnosarda unicolor, collected from 15 sites across the Indo-West Pacific region. We screened 92 individuals for genetic variation at 13 nuclear microsatellite loci and the cytochrome c oxidase subunit 1 (COI) mitochondrial DNA (mtDNA) gene. We detected no genetic differentiation between ocean basins or between sites within ocean basins, suggesting G. unicolor possesses a highly mobile larval or juvenile stage. In addition, the lack of deep evolutionary mtDNA divergences suggests gene flow was also not limited historically. However, comparisons of relatedness between pairs of individuals revealed that individuals collected from the same site were more related to each other, on average, than individuals collected from different sites. Such patterns are consistent with chaotic genetic patchiness, a possible consequence of high variance in reproductive success and patchy, local recruitment.

KEY WORDS: Gymnosarda unicolor · Spatial genetic structure · Microsatellites · mtDNA · Ephemeral genetic patchiness

INTRODUCTION

Many coral reef fishes have a pelagic early life-history stage, followed by a benthic, reef-associated adult stage (Leis 1991). Because the pelagic stages are highly mobile, the pelagic larval duration (PLD) is often used as a proxy for dispersal capability, with a longer PLD expected to result in higher dispersal and lower levels of genetic subdivision (Waples 1987). Consistent with this prediction, species with an extended PLD typically have lower levels of population structure than those with a short PLD (Waples 1987, Doherty et al. 1995, Riginos & Victor 2001), although recent studies have failed to detect a significant relationship between PLD and genetic divergence (Bay et al. 2006, Galarza et al. 2009, Weersing & Toonen 2009). Indeed, there are numerous examples where species with a long PLD have pronounced population structure (Barber et al. 2002, Taylor & Hellberg 2003, Galarza et al. 2009), highlighting the...

Physical barriers can also have profound impacts on dispersal and population structure. For example, the Indo-Pacific barrier (IPB) is thought to have contributed to strong genetic breaks in many species distributed across the Indian and Pacific Oceans (Craig et al. 2007, Rocha et al. 2007, Gaither et al. 2010). Although movement across the IPB is currently possible, sea levels were ~130 m lower than contemporary levels during glacial maxima in the Pleistocene. In those times, the Sunda Shelf would have been almost entirely above sea level, resulting in a near complete barrier to gene flow between the Indian and Pacific Oceans (Voris 2000, Naish et al. 2009, Gaither et al. 2010).

The dogtooth tuna Gymnosarda unicolor Rüppell, 1838, is an epipelagic, coral reef-associated fish endemic to the Indo-Pacific region. The common name of dogtooth tuna is derived from its large, conical teeth and tuna-like appearance in association with its reputation as a voracious predator (Collette & Nauen 1983). The species is distributed in tropical waters from the east coast of Africa and the Red Sea, up to Japan, and across to the islands of the central Pacific (Collette & Nauen 1983, Collette 2001). While few fisheries specifically target G. unicolor, it is often by-caught with heavily exploited species such as yellowfin, bluefin and bigeye tuna. Most catches occur in the Indian Ocean, with global catch rates rising from around 150 t yr⁻¹ in the 1970s to >600 t yr⁻¹ in 2005 (Joshi et al. 2012).

Despite these rising catch rates, very little is known about the biology of G. unicolor. It is almost exclusively a solitary species, associated with coral reef structures, and is rarely found in schools or in the open water (Collette 2001). The PLD of the species is currently unknown. The aim of this study was to evaluate population genetic structure in G. unicolor to gain insights about its dispersal capability. By comparing microsatellite and mitochondrial DNA (mtDNA) sequence variation in samples collected from the Indian and Pacific Oceans, we also seek to evaluate whether the IPB has imposed limits to dispersal. These limits may be contemporary or historical. Finally, we estimate effective population size (Nₑ) and test for evidence of recent genetic bottlenecks that may have arisen through overexploitation.

MATERIALS AND METHODS

Sample collection and DNA extraction

A total of 92 tissue samples (fin clip or muscle) from Gymnosarda unicolor were obtained from 5 sites in the Indian Ocean and 10 sites in the Pacific Ocean between 2007 and 2010 (Fig. 1). Samples were collected by line or spearfishing, with samples provided from numerous collaborators across Australia. Samples were preserved in 100% ethanol and stored at −20°C until utilized for extraction and analyses.

DNA was extracted for polymerase chain reaction (PCR) using a standard ‘salting out’ procedure as described by Sunnucks & Hales (1996) with the following alterations. Tissue samples (1 cm x 1 cm) were subdivided into smaller, more easily digested fragments and were then incubated at 56°C overnight. Extracted DNA was resuspended in a volume of 50 µl sterile water. DNA concentrations and protein levels were analysed with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and stored at −20°C until required for use.

Fig. 1. Sampling locations of Gymnosarda unicolor across the eastern Indian (blue) and western Pacific (red) Oceans. Location IDs are provided in Table 1
Mitochondrial DNA sequencing and microsatellite genotyping

A 680 bp portion of the mitochondrial cytochrome c oxidase subunit 1 (COI) gene was amplified using the primers FishF2 and FishR2 (Ward et al. 2005). This gene was selected because it has shown to be effective in detecting deep genetic divergences between populations in marine fish (Zemlak et al. 2009). PCRs were carried out in 25 µl reactions containing 2 µl of neat DNA or a 1/10 dilution if neat concentrations were larger than 500 ng µl⁻¹. The remaining reaction volume contained 1 × PCR reaction buffer, 4 mM MgCl₂, 0.2 mM dNTPs, 1.2 µM forward primer, 1.2 µM reverse primer and 55 units of Taq polymerase (Fisher Biotech’s Tth Plus DNA Polymerase). PCRs were run at the following cycling parameters: an initial denaturation at 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 54°C for 30 s and 72°C for 45 s then a final extension step of 72°C for 10 min. Following a purification step, PCR products were sequenced on an ABI 3730 Sequencer by the Australian Genome Research Facility (Brisbane, Australia). The resulting sequences were aligned and edited using SEQUENCHER v.5 (Gene Codes) and trimmed to a common length using MEGA v.5.05 (Tamura et al. 2011).

Genotypes at 13 microsatellite loci (GUN3−11, 13, 15, 18 and 20) were determined for each individual using the primers and PCR conditions described by Bentley et al. (2013). PCR products were analysed on an ABI 3700 sequencer using a GeneScan-500 LIZ internal size-standard. Genotypes were scored using GENEMARKER (SoftGenetics) software.

Data analysis

Mitochondrial DNA (mtDNA) variation within each site was quantified by calculating the number of haplotypes (k), number of polymorphic sites (PS), haplotype diversity (H) and nucleotide diversity (π) using ARLEQUIN v.3.11 (Excoffier et al. 2005). We also used ARLEQUIN to calculate the summary statistics D (Tajima 1989) and Fₜ (Fu 1997) to test for deviations from neutral predictions for a constant-sized population. Significant negative values of D or Fₜ indicate an excess of rare haplotypes, a signature of either selection or recent demographic expansion. Mismatch distributions, which are the pairwise nucleotide differences between sequences, were also calculated with ARLEQUIN to investigate the possibility of demographic changes. The evolutionary relationships and geographic distribution of mtDNA haplotypes were visualised using a median-joining haplotype network created with the program NETWORK v.4.1 (Bandelt et al. 1999).

Microsatellite data were checked for null alleles, large allele dropouts and scoring errors at each locus using the software package MICROCHECKER (van Oosterhout et al. 2004). Genetic variation at each locus and site was quantified by calculating allelic richness (a measure of the number of alleles independent of sample size) and gene diversity. Analyses for a deficit or excess of heterozygotes within each site were conducted using randomization tests, with results characterized with the Fₛ statistic. Significantly positive Fₛ values indicate a deficit of heterozygotes relative to a random mating model, while negative results indicate an excess of heterozygotes. Tests for differences in genetic diversity and Fₛ among sites were performed using Friedman’s ANOVA and Wilcoxon rank tests with the R v.2.14 statistical package (R Development Core Team 2011). Linkage disequilibrium between microsatellite loci was assessed by testing the significance of association between genotypes. Estimates of genetic variation and Fₛ and tests for deficits in heterozygotes, genetic differentiation between pairs of sites and genotypic disequilibrium were calculated using the FSTAT v.2.9.3 (Goudet 2001) software package.

Population structure was assessed using a hierarchical analysis of molecular variance (AMOVA), with sites consisting of <5 individuals excluded from the analysis. Due to the low sample sizes at many sites, population structure was also assessed using 2 individual-based methods. The first method involved comparing relatedness values between individuals at different spatial scales. Three spatial scales were compared: (1) between pairs of individuals within the same site, (2) between individuals from different sites within the same ocean basin and (3) between individuals from different sites in different ocean basins. Pairwise relatedness values (R) were calculated using the method of Queller & Goodnight (1989) with the KINGROUP software package (Konovalov et al. 2004). Tests for differences in pairwise relatedness values were carried out using Mann-Whitney U-tests in R. The second method was a Bayesian clustering analysis performed using the STRUCTURE v.2.3.3 software package (Pritchard et al. 2000). This method identifies genetically distinct clusters (K) based on allele frequencies across all loci. Analyses were based on an ancestry model that assumes admixture and correlated allele frequencies. No prior information about the origin of the samples was utilized in the analysis. Datasets were explored with initial burn-ins of 10 000 followed by 100 000 Markov Chain Monte Carlo iterations. Ten replicates for each value of K ranging...
from 1 to 20 for the entire data set were run. STRUCTURE HARVESTER (http://taylor0.biology.ucla.edu/struct_harvest/) was used to determine the most likely number of clusters by comparing the likelihood of the data for different values of K and using the ΔK method of Evanno et al. (2005).

Tests to determine whether populations had undergone a recent decrease in \( N_e \) were conducted using BOTTLENECK v.1.2.02 (Piry et al. 1999). Two different methods of analysis were used. The first method was based on the principle that the number of alleles decreases faster than the expected heterozygosity after a population bottleneck. In this case, the expected heterozygosity should be higher than the equilibrium heterozygosity predicted in a stable population from the observed number of alleles. Analyses were run under the 2-phase model (TPM) with 95% single-step mutation, 5% multiple-step mutations and a variance of 12 among multiple steps. A Wilcoxon signed rank test was used to determine whether each site had an excess of heterozygosity. The second method was a qualitative test based on allele frequency distributions. This test distinguishes between a common L-shaped distribution, in which a full range of rare and common frequency alleles are present, and a shifted distribution, which has lost rare alleles and is indicative of a population bottleneck. Analyses were run using sites with >10 individuals and using 1000 iterations.

Finally, estimates of the contemporary \( N_e \) were calculated using the single sample method described by Waples (2006) using the LDNE software (Waples & Do 2008). Estimates of \( N_e \) were based on all samples as well as samples pooled into respective ocean basins for analysis. Three critical values of minimum allele frequency (0.01, 0.02 and 0.05) were used in this analysis.

Detecting genetic differentiation between ocean basins

To determine whether the sample sizes and number of loci used in this study were sufficient to detect genetic differentiation between ocean basins, simulated data sets were created using the EASYPOP v.2.0.1 program (Balloux 2001). The simulations were based on a simple model consisting of 2 groups of subpopulations (\( n = 2 \) and \( 4 \) in each group, as with the actual data in this study) with no gene flow between groups and complete mixing between subpopulations within each group. Simulations based on models with low levels of gene flow between archipelago groups (number of migrants moving between populations per generation, \( N_m = 1−2 \) and \( 2−4 \)) were also run. Population sizes for the subpopulations were set to 1000 (500 males and 500 females). The simulations ran for 1000 generations and were based on 12 loci each with 8 alleles (the average number of alleles per locus in this study), free recombination and a maximum initial variability at each locus. A stepwise mutation model was used with a mutation rate of \( 1 \times 10^{-4} \) per generation. Increasing the mutation rate decreased the level of divergence between archipelago groups, a parameter we investigated by changing levels of gene flow, so the mutation rate was kept constant in all the simulation models. After 1000 generations, 10 individuals were randomly selected from each subpopulation (the average sample size in this study) and used to test for genetic differentiation between groups using a hierarchical AMOVA with ARLEQUIN, as undertaken with the original data. One hundred simulated data sets were generated and tested for genetic differentiation for each model.

RESULTS

Genetic variation within sites

In total, 23 mtDNA haplotypes, including 14 singletons, were obtained from the 92 individuals that were sequenced for the COI gene. These haplotype sequences have been deposited in GenBank (accession nos.: KJ534606 to KJ534628). The median-joining network constructed using these haplotypes was relatively simple with no deep divergences (Fig. 2). None of the observed haplotypes differed by >4 bp. Additionally, none of the non-singleton haplotypes appeared to be geographically restricted, except haplotype 6, which was only found in the Pacific Ocean. The number of haplotypes and estimates of haplotype diversity for each site are presented in Table 1. All sites with >5 samples contained at least 4 haplotypes, with the exception of Imperieuse Reef (\( n = 7 \)), which contained only 2 haplotypes. No haplotype was shared across all sites. The most abundant haplotype (Hap1) was found in 28% of sequenced individuals, with the exception of Imperieuse Reef (\( n = 7 \)), which contained only 2 haplotypes. No haplotype was shared across all sites. The most abundant haplotype (Hap1) was found in 28% of sequenced individuals, with the second most abundant haplotype (Hap3) found in 20% of individuals.

One microsatellite locus (GUN4) was found to have null alleles using MICROCHECKER and was therefore excluded from subsequent analyses. Genetic diversity estimates for the remaining 12 loci are presented in Table 2. No significant differences in diversity estimates were detected between sites with \( \geq 10 \) individ-
Bentley et al.: Spatial genetic structure in dogtooth tuna

Genetic differentiation among sites

The AMOVA analyses failed to detect any significant differentiation between ocean basins or between sites within ocean basins, with almost all genetic variation, both mtDNA and nuclear, occurring within sites (Table 3). Similarly, the fluctuations in the log probability estimates and ΔK values in the Bayesian clustering analysis suggest that the most probable number of genetic clusters is 1. By contrast, evidence of population structure was found when comparing pairwise relatedness at different spatial scales, with significantly higher relatedness, on average, between individuals within sites compared to individuals from different sites (Fig. 3; p < 0.001 for both within site vs. between site within ocean basin and within site vs. between ocean basin comparisons). However, there was no significant difference in pairwise

![Median-joining haplotype network for the COI gene in 92 individuals of Gymnosarda unicolor from 15 populations across the Indo-West Pacific. Each circle represents a unique haplotype, the area of each circle represents the relative proportion of individuals with that haplotype, colours represent collection site (Pacific: red, Indian: blue), and each line represents 1 mutation](image)

Table 1. Sample size (n), number of haplotypes (N_HAP), number of polymorphic sites (PS), haplotype diversity (h) and nucleotide diversity (π) in the mitochondrial COI gene at 15 sampling locations in the Indian and Pacific Oceans

<table>
<thead>
<tr>
<th>Population</th>
<th>n</th>
<th>N_HAP</th>
<th>PS</th>
<th>h (SE)</th>
<th>π (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indian Ocean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christmas Island (CI)</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>1.00 (0.27)</td>
<td>0.006 (0.005)</td>
</tr>
<tr>
<td>Ashmore Reef (AR)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.00 (0.00)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Scott Reef (SR)</td>
<td>17</td>
<td>9</td>
<td>8</td>
<td>0.87 (0.07)</td>
<td>0.003 (0.002)</td>
</tr>
<tr>
<td>Cocos (Keeling) Island (CoI)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.00 (0.00)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Imperieuse Reef (IR)</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>0.48 (0.17)</td>
<td>0.002 (0.002)</td>
</tr>
<tr>
<td>All samples</td>
<td>29</td>
<td>10</td>
<td>9</td>
<td>0.84 (0.04)</td>
<td>0.003 (0.002)</td>
</tr>
<tr>
<td><strong>Pacific Ocean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vema Bank (VB)</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>0.94 (0.07)</td>
<td>0.005 (0.003)</td>
</tr>
<tr>
<td>Bougainville Reef (BR)</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1.00 (0.50)</td>
<td>0.004 (0.005)</td>
</tr>
<tr>
<td>Escape Reef (ER)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.00 (0.00)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Linden Bank (LB)</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>1.00 (0.50)</td>
<td>0.007 (0.008)</td>
</tr>
<tr>
<td>Holmes Reef (HR)</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>1.00 (0.27)</td>
<td>0.004 (0.004)</td>
</tr>
<tr>
<td>Faraday Shoals (FS)</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>0.90 (0.16)</td>
<td>0.004 (0.003)</td>
</tr>
<tr>
<td>Flinders Reef (FR)</td>
<td>20</td>
<td>9</td>
<td>11</td>
<td>0.86 (0.06)</td>
<td>0.004 (0.002)</td>
</tr>
<tr>
<td>Thimble Reef (TR)</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>0.67 (0.31)</td>
<td>0.003 (0.003)</td>
</tr>
<tr>
<td>Knuckle Reef (KR)</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>1.00 (0.27)</td>
<td>0.004 (0.004)</td>
</tr>
<tr>
<td>Marion Reef (MR)</td>
<td>15</td>
<td>10</td>
<td>11</td>
<td>0.94 (0.04)</td>
<td>0.004 (0.003)</td>
</tr>
<tr>
<td>All samples</td>
<td>63</td>
<td>19</td>
<td>19</td>
<td>0.88 (0.03)</td>
<td>0.004 (0.002)</td>
</tr>
</tbody>
</table>
relatedness between individuals from different sites within the same ocean basin and individuals from different ocean basins (p = 0.970). While the $F_{ST}$ value based on the microsatellite data was significantly different from zero (Table 2), none of the pairwise $F_{ST}$ values were significant after adjusting for multiple comparisons.

Neutrality tests, demographic history and estimates of effective population size

Fu’s $F_S$, Tajima’s $D$ and mismatch analysis statistics for each site and for samples pooled across sites within oceans are presented in Table 4. Tajima’s $D$ was not significantly different from zero at any site or either of the ocean basins. However, significantly negative $F_S$ values were found at Scott Reef, Flinders Reef and Marion Reef as well as in the Pacific Ocean when all individuals were pooled, suggesting population expansions or genetic hitchhiking at these sites. The mismatch distribution also provided evidence for population expansion and growth. In all samples except Flinders Reef, sum of squared deviations (SSD) and raggedness values did not differ significantly from a distribution simulated under a model of rapid expansion.

None of the sites showed evidence of recent genetic bottlenecks. No significant heterozygosity excesses were detected, and all sites exhibited normal L-shaped allele frequency distributions. Estimates of contemporary $N_e$ were negative or infinite

<table>
<thead>
<tr>
<th>Site</th>
<th>Sample size</th>
<th>$A_r$</th>
<th>$H$</th>
<th>$F_{IS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian Ocean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>17.0 (0.0)</td>
<td>3.8 (0.4)</td>
<td>0.66 (0.07)</td>
<td>−0.03 (0.04)</td>
</tr>
<tr>
<td>IR</td>
<td>7.0 (0.0)</td>
<td>3.5 (0.4)</td>
<td>0.58 (0.09)</td>
<td>−0.05 (0.05)</td>
</tr>
<tr>
<td>Mean</td>
<td>12.0 (5.0)</td>
<td>3.7 (0.2)</td>
<td>0.62 (0.04)</td>
<td>−0.04 (0.01)</td>
</tr>
<tr>
<td>Pacific Ocean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VB</td>
<td>7.5 (0.3)</td>
<td>3.3 (0.4)</td>
<td>0.61 (0.08)</td>
<td>0.05 (0.04)</td>
</tr>
<tr>
<td>FS</td>
<td>4.8 (0.1)</td>
<td>3.8 (0.4)</td>
<td>0.66 (0.07)</td>
<td>−0.01 (0.07)</td>
</tr>
<tr>
<td>FR</td>
<td>19.8 (0.1)</td>
<td>3.9 (0.4)</td>
<td>0.67 (0.07)</td>
<td>0.05 (0.03)</td>
</tr>
<tr>
<td>MR</td>
<td>14.9 (0.1)</td>
<td>3.7 (0.4)</td>
<td>0.65 (0.08)</td>
<td>0.04 (0.03)</td>
</tr>
<tr>
<td>Mean</td>
<td>11.7 (3.4)</td>
<td>3.7 (0.1)</td>
<td>0.65 (0.01)</td>
<td>0.03 (0.01)</td>
</tr>
</tbody>
</table>

Table 3. Hierarchical AMOVA based on mtDNA and nuclear microsatellite DNA data from sites with at least 5 individuals. *F-statistic significantly different from zero at p < 0.05; NS not significantly different from zero

<table>
<thead>
<tr>
<th>Total variation (%)</th>
<th>$F$ or $\Phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>mtDNA (haplotype frequency)</td>
<td></td>
</tr>
<tr>
<td>Between oceans</td>
<td>−1.3</td>
</tr>
<tr>
<td>Among sites within oceans</td>
<td>2.4</td>
</tr>
<tr>
<td>Within sites</td>
<td>98.9</td>
</tr>
<tr>
<td>mtDNA (pairwise distance)</td>
<td></td>
</tr>
<tr>
<td>Between oceans</td>
<td>0.9</td>
</tr>
<tr>
<td>Among sites within oceans</td>
<td>−0.7</td>
</tr>
<tr>
<td>Within sites</td>
<td>99.8</td>
</tr>
<tr>
<td>Microsatellites</td>
<td></td>
</tr>
<tr>
<td>Between oceans</td>
<td>0.1</td>
</tr>
<tr>
<td>Among sites within oceans</td>
<td>1.0</td>
</tr>
<tr>
<td>Within sites</td>
<td>98.9</td>
</tr>
</tbody>
</table>

Table 4. Tajima’s $D$ and Fu’s $F_S$ neutrality tests and mismatch analyses based on a 680 bp sequence of the COI mitochondrial gene. Asterisks represent the significance of tests for neutrality ($F_S$ and $D$) and model of population expansion (mismatch analysis). SSD: sum of squared deviations, RAG: raggedness. *p < 0.05, **p < 0.01
for all sites and when sites from the Indian Ocean were pooled together. This indicates that the signal caused by genetic drift was less than the noise created by the sampling error. An exception was the pooled Pacific Ocean sample with estimates of $N_e$ ranging from 128 to 288 for different minimum allele frequency limits. The 95% confidence levels for these values were infinite, 1266 and 228 respectively.

Detecting genetic differentiation between ocean basins

The simulations demonstrate that the sample sizes and number of loci used in our study were sufficient to detect low to moderate levels of genetic differentiation between ocean basins. In simulations where there was no gene flow between archipelago groups, levels of genetic divergence between groups were moderately high ($F_{CT}$ ranged from 0.10 to 0.20), and significant between-group divergences were detected in all 100 simulated data sets ($p < 0.001$ in all cases). When gene flow between archipelago groups was introduced, levels of divergence between groups were, predictably, much lower ($F_{CT} = 0.02$ to 0.06 when there were 1 to 2 migrants per generation, and $F_{CT} = 0.01$ to 0.03 when there were 2 to 4 migrants per generation), but significant between-group divergences were still detected in all simulated data sets (maximum $p$ value = 0.004).

DISCUSSION

Insights into dispersal

The absence of genetic differentiation between ocean basins and similar pairwise genetic relatedness values between sites irrespective of their geographical separation suggests Gymnosarda unicolor possesses a highly mobile pelagic larval stage with a prolonged duration or, alternatively, a highly mobile juvenile stage, or perhaps both. Indeed, the high levels of genetic diversity and sharing of haplotypes over vast geographical distances that were observed in this study are more consistent with a single, widespread pelagic taxa than a reef-associated species such as G. unicolor (Bowen et al. 2001). It should be noted, however, that while some larvae may be able to traverse vast distances, the average realized dispersal may be much lower because only a few migrants per generation are required for genetic homogeneity (Slatkin 1987).

Effect of the Indo-Pacific barrier

In contrast to the AMOVA and Bayesian clustering results, which suggested no broad-scale genetic structure, the pairwise relatedness values indicate a significantly higher degree of relatedness between individuals collected at the same site than among individuals from different sites. This pattern of fine-scale genetic heterogeneity, coupled with broad-scale genetic homogeneity, is consistent with chaotic genetic patchiness (Johnson & Black 1982). Chaotic genetic patchiness can arise from natural selection acting after settlement. Alternatively, it can arise from temporal variation in the genetic composition of recruits coupled with patchy local recruitment (Johnson & Black 1984, Hedgecock 1994). Under the ‘sweepstakes reproductive success’ hypothesis (Hedgecock 1994), temporal genetic variance in recruits arises through large differences in the reproductive success of individuals due to chance matching of reproductive activity and oceanographic conditions that are conducive for larval survival.

The patchy settlement of cohorts limits the development of broad-scale population structure because recruitment is generally random and dependent on a number of biological and environmental factors, such as ocean currents. The low estimate of $N_e$ for the pooled Pacific Ocean sample is consistent with sweepstake reproductive success, as a large variance in the reproductive success contributes to low $N_e$ (Hedrick 2005). Fluctuations in population size and uneven sex ratios also affect $N_e$. While the estimated $N_e$ was surprising low for this sample, it is consistent with sizes found in other marine fishes (Hauser & Carvalho 2008).

Fine-scale heterogeneity

In contrast to the AMOVA and Bayesian clustering results, which suggested no broad-scale genetic structure, the pairwise relatedness values indicate a significantly higher degree of relatedness between individuals collected at the same site than among individuals from different sites. This pattern of fine-scale genetic heterogeneity, coupled with broad-scale genetic homogeneity, is consistent with chaotic genetic patchiness (Johnson & Black 1982). Chaotic genetic patchiness can arise from natural selection acting after settlement. Alternatively, it can arise from temporal variation in the genetic composition of recruits coupled with patchy local recruitment (Johnson & Black 1984, Hedgecock 1994). Under the ‘sweepstakes reproductive success’ hypothesis (Hedgecock 1994), temporal genetic variance in recruits arises through large differences in the reproductive success of individuals due to chance matching of reproductive activity and oceanographic conditions that are conducive for larval survival.

The patchy settlement of cohorts limits the development of broad-scale population structure because recruitment is generally random and dependent on a number of biological and environmental factors, such as ocean currents. The low estimate of $N_e$ for the pooled Pacific Ocean sample is consistent with sweepstake reproductive success, as a large variance in the reproductive success contributes to low $N_e$ (Hedrick 2005). Fluctuations in population size and uneven sex ratios also affect $N_e$. While the estimated $N_e$ was surprising low for this sample, it is consistent with sizes found in other marine fishes (Hauser & Carvalho 2008).
Menezes et al. 2006, Craig et al. 2007, Klanten et al. 2007, Fitzpatrick et al. 2011). However, a few species have shown genetic homogeneity across the IPB (Bowen et al. 2001, Lessios et al. 2001, Horne et al. 2008, Gaither et al. 2010), similar to the results presented here. *Gymnosarda unicolor* is therefore one of only a limited number of species that show high gene flow across the IPB in the past, deep genetic divergences, revealed by the presence of distinct genetic clades, would be expected to be present, as with species such as the mantis shrimp *Haptosquilla pulcella* (Barber et al. 2002). However, this was not observed in *G. unicolor*, expanding the evidence for continual gene flow over a barrier that has impeded dispersal in numerous other species. If gene flow was restricted by the IPB in the past, deep genetic divergences across the Indo-West Pacific, the interpretation of the results, particularly with respect to sustainable fisheries management, should be considered with caution. As previously stated, in large populations, migration rates may be large enough to maintain genetic homogeneity between populations but insufficient to facilitate the recovery of the population in the area that has undergone the depletion (Waples 1998). It is also important to recognise that a lack of evidence of population genetic structure in neutral molecular markers does not necessarily reflect a lack of adaptive variation. Neutral and adaptive genetic loci are generally uncoupled (Reed & Frankham 2001), and genetic homogeneity in neutral markers may be underlain with strong heterogeneity in adaptive loci (Larsen et al. 2007). Identification of adaptive genetic differences is imperative, as locally adapted populations should be considered single, separate management units (Waples 1998). It may well be that the populations used in this study are distinct in their adaptive variation with adaptations for their specific local environment; however, this remains to be examined in further studies. Managing this species as a single panmictic population may lead not only to a decline in population numbers but also a reduction in genetic diversity, leaving the species vulnerable to environmental perturbations and disease. As such, fishery managers need to consider a precautionary approach to the sustainable management and harvest of *G. unicolor* that seeks to maintain adequate levels of total spawner biomass by developing and implementing harvest strategies that limit the potential for localised depletion events.

**Implications for management**

Although this study provides evidence for high dispersal across the Indo-West Pacific, the interpretation of the results, particularly with respect to sustainable fisheries management, should be considered with caution. As previously stated, in large populations, migration rates may be large enough to maintain genetic homogeneity between populations but insufficient to facilitate the recovery of the population in the area that has undergone the depletion (Waples 1998). It is also important to recognise that a lack of evidence of population genetic structure in neutral molecular markers does not necessarily reflect a lack of adaptive variation. Neutral and adaptive genetic loci are generally uncoupled (Reed & Frankham 2001), and genetic homogeneity in neutral markers may be underlain with strong heterogeneity in adaptive loci (Larsen et al. 2007). Identification of adaptive genetic differences is imperative, as locally adapted populations should be considered single, separate management units (Waples 1998). It may well be that the populations used in this study are distinct in their adaptive variation with adaptations for their specific local environment; however, this remains to be examined in further studies. Managing this species as a single panmictic population may lead not only to a decline in population numbers but also a reduction in genetic diversity, leaving the species vulnerable to environmental perturbations and disease. As such, fishery managers need to consider a precautionary approach to the sustainable management and harvest of *G. unicolor* that seeks to maintain adequate levels of total spawner biomass by developing and implementing harvest strategies that limit the potential for localised depletion events.

**CONCLUSIONS**

Overall, this study shows no evidence of genetic differentiation between populations of *Gymnosarda unicolor* distributed across a 7500 km range in the Indo-West Pacific from the Cocos (Keeling) Islands (Indian Ocean) to Marion Reef (Pacific Ocean). Such observations suggest that this species has high dispersal capabilities, with the IPB observed to have no contemporary or historic effect on gene flow between ocean basins. No evidence of recent population bottlenecks was observed for this species, but evidence of a population expansion was detected, particularly in the Pacific Ocean. Estimates of $N_e$ were relatively low in a pooled sample and, in conjunction with evidence of higher pairwise relatedness between individuals within sites than between sites, support the idea that sweepstake reproductive success may be resulting in the development of fine-scale genetic heterogeneity.

Further studies should aim to investigate the role of sweepstake reproductive success in generating fine-scale genetic heterogeneity in *G. unicolor* by measuring temporal variation in recruits or by estimating the
Acknowledgements. Support for this project was provided by The University of Western Australia, the Department of Fisheries, Western Australia and James Cook University. We sincerely thank the spearfishing community of eastern Australia who provided the Pacific Ocean samples used in the study, particularly Mick de Rooy and members of the Townsville Skindiving Club. The Australian Underwater Federation was also pivotal in procuring samples that are unusually difficult to obtain. In addition, we thank Ross Newton (Reel Teaser Fishing Adventures) for collecting samples from Scott Reef (WA), Mark Rochfort (Shorefire) for assistance in collecting samples from Christmas Island and our colleagues Craig Skepper, Ben Rome and Gabby Mitsopoulos for coordinating sample collections. Yvette Hitchen and Sherralee Lukehurst provided valuable technical assistance.

LITERATURE CITED


Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915−925


Liu JX, Gao TX, Zhuang ZM, Jin XS, Yokogawa K, Zhang YP (2006b) Late Pleistocene divergence and subsequent population expansion of two closely related fish species, Japanese anchovy (Engraulis japonicus) and Australia anchovy (Engraulis australis). Mol Phylogenet Evol 40:712−723
R Development Core Team (2011) R: a language and environment for statistical computing (v. 2.14.0). R Founda-


Editorial responsibility: Karen Miller, Hobart, Australia

Submitted: September 23, 2013; Accepted: April 4, 2014
Proofs received from author(s): June 4, 2014