Dammannagoda, Sudath T., Hurwood, David, & Mather, Peter (2011) Genetic analysis reveals two stocks of skipjack tuna (Katsuwonus pelamis) in the north western Indian Ocean. *Canadian Journal of Fisheries and Aquatic Sciences (CJFAS)*, *68*(2), pp. 210-223.

Genetic analysis reveals two stocks of skipjack tuna (*Katsuwonus pelamis*) in the north western Indian Ocean

Sudath T. Dammannagoda, David A. Hurwood and Peter B. Mather

Biogeosciences Discipline, Faculty of Science & Technology, Queensland University of Technology, GPO Box 2434, Brisbane Q 4001, Australia.

Correspondence: Sudath T. Dammannagoda Email: s.dammannagoda@qut.edu.au

Tel: +61 7 31385051 Fax: +61 7 31382330

Abstract

Skipjack (SJT) (*Katsuwonus pelamis*) is a medium sized, pelagic, highly dispersive tuna species that occurs widely across tropical and subtropical waters. SJT constitute the largest tuna fishery in the Indian Ocean, and are currently managed as a single stock. Patterns of genetic variation in a mtDNA gene and 6 microsatellite loci were examined to test for stock structure in the northwestern Indian Ocean. 324 individuals were sampled from five major fishing grounds around Sri Lanka, and single sites in the Maldive Islands and the Laccadive Islands. Phylogenetic reconstruction of mtDNA revealed two coexisting divergent clades in the region. AMOVA (Analysis of Molecular Variance) of mtDNA data revealed significant genetic differentiation among sites (Φ_{ST} = 0.2029, $P < 0.0001$), also supported by SAMOVA results. AMOVA of microsatellite data also showed significant differentiation among most sampled sites (F_{ST} = 0.0256, $P<0.001$) consistent with the mtDNA pattern. STRUCTURE analysis of the microsatellite data revealed two differentiated stocks. While the both two marker types examined identified two genetic groups, microsatellite analysis indicates that the sampled SJT are likely to represent individuals sourced from discrete breeding grounds that are mixed in feeding grounds in Sri Lankan waters.

Key words: Skipjack tuna, *Katsuwonus pelamis*, Indian Ocean, mtDNA, microsatellites

Introduction

A widely accepted view among many fisheries managers is that most tuna species are wide-ranging, dispersive species that given their cosmopolitan distributions, large population sizes, high fecundity and production of pelagic eggs and larvae, most probably constitute essentially homogeneous populations over oceanic spatial geographical scales. General characteristics of the open ocean environment combined with tuna life history traits have led to the expectation that most tuna species will show little, if any, population structure even at very large spatial scales (inter-oceanic) (Ward et al. 1994; Chow and Ushiama 1995; Ely et al. 2005). In this context, populations of most tuna species are currently managed as single stocks (e.g. yellowfin tuna and bigeye tuna in the Pacific, Atlantic and Indian Oceans). Given this background, skipjack tuna (*Katsuwonus pelamis*) (SJT) are currently considered, like most tuna species, to constitute a single panmictic population and hence are managed as single stocks at an ocean-wide scale in all oceans by the respective international commissions (i.e. IOTC-Indian Ocean Tuna Commission; WCPFC-Western and Central Pacific Fisheries Commission; IATTC- Inter-American Tropical Tuna Commission; and ICCAT-International Commission for the Conservation of Atlantic Tuna). This approach is not based however, on experimental stock delineation studies but largely on preconceptions about general tuna life history traits (LHTs), and fishery data.

Stock delineation studies of large tuna species have most often been based on sampling designs at oceanic scales due to an expectation that, populations are unlikely to show structure at finer spatial scales. Indeed, a population genetic study of SJT carried out at an oceanic scale in the Atlantic and Pacific Oceans that employed mtDNA RFLP data could not detect any SJT differentiation between these water bodies (Graves et al. 1984). A more recent oceanic-scale study of SJT from the Atlantic and Indian Oceans that employed mtDNA D-loop sequence data also could not differentiate SJT from the two oceans (Ely et al. 2005).

However, a recent study of SJT from the east coast of India and from the Pacific Ocean near Japan could differentiate the two samples using mtDNA RFLP markers (Menezes et al. 2005), it failed however, to differentiate the Japanese sample and an SJT sample from the west coast of India using microsatellite markers (Menezes et al. 2008). The scale of these studies, perhaps combined with a lack of sensitivity in the molecular markers used, may have hindered potential for detecting real population differentiation, if it was present. In contrast, earlier allozyme studies, that employed relatively fine scale sampling detected population structure at both inter- and intra ocean spatial scales (e.g. Sharp 1978, Fujino et al. 1981).

Evidence for heterogeneous intra-specific stocks of tuna species and other large marine pelagic fishes have increased in recent times (Atlantic cod: Ruzzante et al. 1998; Knutsen et al. 2003; Atlantic bluefin tuna: Carlsson et al. 2004, Carlsson et al. 2006, Boustany et al. 2008; yellowfin tuna: Dammannagoda et al. 2008; swordfish: Kotoulas et al. 1995; Reeb et al. 2000; flounder: Florin and Hoglund 2008; wahoo: Garber et al. 2005; sailfish: Hoolihan et al. 2004). Most have employed sampling at relatively fine geographical scales and have considered ocean current patterns, historical geomorphological factors, bathymetry of the ocean basin, salinity gradients, and specific LHTs characteristics of individual species. Thus, simplistic assumptions about a species' general life history traits (e.g. large-size, fast-swimming, highly mobile, pelagic larvae etc.) used to define the scale at which sampling is conducted, may inadvertently miss real population subdivision and hence potentially mislead management decisions.

SJT constitute the largest tuna fishery in the world and this is also the case in the Indian Ocean (FAO 2008). Substantial increases however, in the fishing effort employed by purse-seiners over the last decade in the western Indian Ocean, have suggested that the relative abundance of SJT in these areas may have declined. As the western Indian Ocean is considered to be an important SJT recruitment area, authorities are concerned about the health

78 of the fishery there (IOTC 2007). Importantly, unlike other industrial tuna fisheries around the
79 world, the SJT fishery is the principal animal protein resource for many coastal communities
80 in developing nations in the Indian Ocean and provides important employment and livelihood
81 opportunities for many poor communities. A collapse of SJT stocks in the region, could
82 compromise not only food security of many poor people, but also important employment
83 opportunities for coastal populations.

84 The IOTC currently considers that Indian Ocean SJT stocks are resilient. In addition,
85 IOTC tagging studies have shown rapid, large scale movements of SJT in the western Indian
86 Ocean and they have used these data to argue that SJT is a single stock for management
87 purposes there (IOTC 2007), and as such SJT are currently managed as a single stock in the
88 Indian Ocean. In reality however, little is known about SJT stock delineation, or maximum
89 sustainable yield, in this region (IOTC 2007).

90 Managing any fishery does not simply mean managing total fish numbers. While non-
91 genetic methods of stock assessment including morphological/meristic or fishery based
92 methods can only infer fish breeding units, a population genetics approach can directly test
93 the hypothesis that genetically different breeding units may be present (Ward 2000). If one or
94 more of the genetically distinct breeding units is at low abundance, equal fishing pressure on
95 both groups, can lead to extinction of one or more of the low abundance groups. Furthermore,
96 a number of marine fisheries around the world consist of mixtures of genetically discrete
97 stocks in commercial fishing grounds, where individuals have originated from discrete
98 spawning sites (e.g. salmon fisheries in the northwest Pacific (Shaklee et al. 1999)). Evidence
99 has also increased that some tuna fisheries consist of mixed stocks in fishing grounds (e.g.
100 Atlantic bluefin tuna (ABT) in the Mediterranean Sea (Carlsson et al. 2006) and the Gulf of
101 Mexico (Boustany et al. 2008).

102 Uncontrolled harvesting of wild populations can perturb natural genetic subdivision
103 among populations and reduce overall stock productivity. This may not be recognised
104 however, unless subpopulations have been identified independently, and individuals from
105 population mixtures are assigned to their correct subpopulations (Allendorf et al. 2008).
106 Consequently, to manage populations in a sustainable way, we need to understand 'real' stock
107 structure (Waples et al. 2006; Palsboll et al. 2007). It is likely therefore, that a lack of detailed
108 population genetic studies of Indian Ocean SJT stocks has constrained development of
109 appropriate and effective scientific management strategies for this species.
110 Here we examine the extent of genetic differentiation in both mtDNA and
111 microsatellite markers among SJT samples collected between 2001 and 2004 in waters of the
112 north western Indian Ocean around Sri Lanka including the Maldive Islands and Laccadive
113 Islands. The aim was to test for population structure among major fishing grounds around Sri
114 Lanka and to evaluate whether SJT catches in this region could represent heterogeneous
115 stocks.
116

Materials and methods

Sampling

SJT samples were collected from commercial fishing operations from five sites around Sri Lanka and single sites in the Maldive Islands and the Laccadive Islands (Figure 1). Sampling sites were selected to represent major fishing grounds in Sri Lankan waters, and two collections (Maldive Islands and the Laccadive Islands) as outgroups to compare the levels of genetic differentiation in SJT populations from geographically remote regions in the northwestern Indian Ocean. Samples were collected between 2001 and 2004, with white muscle tissue collected from approximately 50 individuals per site (Table 1). Muscle tissue samples were removed from fish and stored in 95% ethanol for later genetic analyses.

(insert Figure 1)

DNA extraction, PCR and screening for variation

Total genomic DNA was extracted from each sample using either a phenol-chloroform (Sambrook et al. 1989) or a modified salt extraction method (Miller et al. 1988). The whole mtDNA ATP-*synthase* subunits 6 and 8 region (ATP*ase* 6, 8) was targeted using the ATP 8.2 L (5'AAA GCR TYR GCC TTT TAA GC 3') and COIII.2H (5' GTT AGT GGT CAK GGG CTT GGR TC 3') primers (http://striweb.si.edu/bermingham/research/primers/index.html) resulting in a ~950bp fragment. Subsequently, two ATP*ase* internal primers were designed to produce a 540bp fragment, a size more appropriate for Temperature Gradient Gel Electrophoresis (TGGE) analysis (Forward primer: 5' CCT AGT GCT AAT GGT GCG ATA AA 3'; Reverse primer: 5' TTC CTC CAA AAG TTA TAG CCC AC 3') that was used in further analyses. MtDNA PCR reactions consisted of 2.5µLRoche 10X buffer, 0.5µL 25mM Fisher $MgCl_2$, 1µL Roche deoxynucleotide triphosphate (dNTP), 1µL 10 mM primer, 0.2µL Roche Taq DNA polymerase, 1µL DNA template (~200ng) adjusted to a final volume of 25µL with 16µL ddH_2O. MtDNA PCR conditions were, 5 minutes at 94^0C for initial

142 denaturation, then 30 cycles of 40 seconds at 94^0C, 40 seconds at 52^0C, 40 seconds at 72^0C,
143 with the final extension step 8 minutes at 72^0C. MtDNA variation was assayed using
144 Temperature Gradient Gel Electrophoresis with outgroup heteroduplex analysis
145 (TGGE/OGHA) as described in Lessa and Applebaum (1993) and Campbell et al. (1995).
146 Each unique haplotype identified using TGGE/OGHA was sequenced. A few unique
147 haplotypes were sequenced in both directions, and as forward and reverse sequences were
148 perfectly complementary, remaining haplotypes were then sequenced only in the forward
149 direction. All sequencing was performed at the Australian Genome Research Facility (AGRF)
150 using an ABI 3730xl sequencing platform. PCR products from each unique haplotype were
151 purified using ethanol precipitation, and sequenced using the forward primer and the BigDye
152 terminator Sequencing Ready Reaction v3.1 kit (Applied Biosystems, California, USA)
153 following manufacturer's specifications. Sequences were uploaded to GenBank under
154 accession numbers FJ481378-FJ481426.

155 SJT genomic libraries were developed specifically for the current study using a
156 radioisotopic method as described in Chand et al. (2005). Microsatellite markers were
157 targeted, isolated and trialed to produce polymorphic loci that were optimised to screen for
158 variation in the sampled populations. Two tri- (UTD328, UTD203) and four tetra-nucleotide
159 microsatellite loci (UTD73, UTD535, UTD523, and UTD172) were amplified and analysed
160 here (Electronic Appendix S3). Microsatellite PCR reaction mix consisted of ~50ng/µL DNA
161 1µL, 1.25µL of 10X PCR buffer (Roche), 0.25µL of 25mM $MgCl_2$, 0.5µL of 10mM dNTP
162 (Roche), 0.5µL of each 10mM forward and reverse primers, 0.1µL of Taq (Roche) and
163 ddH_2O to a final volume of 10µL. PCR conditions were; 5 minutes at 94^0C for initial
164 denaturation, then 30 cycles of 30 seconds at 95^0C, 30 seconds at relevant annealing
165 temperature (Electronic Appendix S3), 30 seconds at 72^0C and a final extension step at 72^0C
166 for 8 minutes. Microsatellite polymorphisms were analysed on a Gelscan2000 System

(Corbett Research) in 5% acrylamide gels run according to the manufacturer's instructions. A 50-350bp size standard (Tamra 350) was run at both ends of each gel and in two additional lanes to confirm allele size. In addition to the size standard, an allele reference standard was generated for known SJT allele sizes and this standard was run in two additional lanes to ensure consistent allele scoring across all gel runs. Microsatellite allele sizes were scored using One D-scan 2.05 (Scanalytics, Inc., 1998).

Statistical analysis

mtDNA variation

MtDNA haplotype sequences were edited and aligned in BioEdit version 7.0.1 (Hall 1999) and sequence data then tested for deviation from neutral expectations using both Tajima's D (Tajima 1989) and Fu's F_S (Fu 1997) tests implemented in Arlequin version 2.00 (Schneider et al. 2005). Significance for both tests was determined using the coalescent simulation process in DnaSP 4.10 (Rozas et al. 2003) with 1000 replicates. *P* values of multiple neutrality tests were adjusted using the Bonferroni correction (Rice 1989).

A mtDNA parsimony cladogram of haplotypes was constructed (at 95% level connectivity) using TCS version 1.18 (Clement et al. 2000). Haplotype networks reconstruct the genealogical history of haplotypic variation and illustrate the evolutionary relationship among unique haplotypes. Under coalescent principles, internal haplotypes in a network are assumed to be ancestral, while tip haplotypes are considered younger, more recently derived types (Templeton et al. 1987; Templeton and Sing 1993; Crandall 1996). Frequency and site information were incorporated into the SJT network here to illustrate the distribution of haplotypes among locations.

Genetic variation was examined using several standard diversity indices including; the number of segregating (polymorphic) sites (S), haplotype diversity (H_d), and molecular

192 diversity indices including nucleotide diversity (π) (Nei, 1987), the average number of
193 pairwise nucleotide differences (k) (Tajima, 1983), and expected heterozygosity based on
194 number of segregating sites (θ_s) (Watterson, 1975) implemented in Arlequin for the total SJT
195 sample collection and for individual samples.

196 Population genetic analyses were performed using Arlequin and DnaSP. Spatial
197 differentiation was estimated using the fixation index (Φ_{ST}) (Excoffier et al. 1992) that
198 includes information on mitochondrial haplotype frequency (Weir and Cockerham 1984), and
199 genetic divergence among unique haplotypes. For all Φ_{ST} analyses, the Tamura and Nei
200 (1993) distance method was used as it accounts for mutational rate heterogeneity among
201 nucleotides within the gene fragment. Using Φ statistics, analysis of molecular variance
202 (AMOVA) was performed to examine the amount of genetic variability partitioned within and
203 among SJT populations (Excoffier et al. 1992). Hierarchical AMOVA was used to investigate
204 the effects of temporal and spatial sampling. Hierarchical AMOVA partitions total genetic
205 differentiation (Φ_{ST}) among pre-defined groups, and yields two measures of genetic
206 differentiation; Φ_{SC} that describes variation among sample populations within groups, and at a
207 higher level in the hierarchy, Φ_{CT} that describes differentiation among groups (Excoffier et al.
208 1992). In the current study, first, the entire sample collection was tested for genetic
209 differentiation to test the null hypothesis of panmixia among collections. Secondly,
210 hierarchical AMOVA analysis was undertaken to determine if any observed structure was
211 partitioned temporally or spatially. Hierarchical groups were organised as year groups. For
212 year groups, sites were grouped based on year of sampling (i.e. 4 groups - 2001, 2002, 2003,
213 and 2004). Genetic variation was therefore partitioned among year (Φ_{CT}), among sites within
214 years (Φ_{SC}), and within sites. Using this hierarchical grouping, we tested the stability of any
215 inferred structure among sampling sites that remained stable over time, and simultaneously,
216 assessed genetic differentiation among sites within years. All other mtDNA analyses were

carried out for pooled temporal samples from each site as there was no temporal variation. The significances of variance components for each hierarchical comparison were tested using a nonparametric permutation procedure incorporating 1000 iterations.

Spatial population structure was also examined by estimating genetic differentiation between all pairs of sites (pair wise Φ_{ST} analysis). Significance of pair site comparisons were tested using the permutation process, as above. In all instances with multiple tests, *P* values were adjusted using the Bonferroni correction (Rice 1989).

Spatial structure was investigated further via Spatial Analysis of Molecular Variance (SAMOVA) (Dupanloup et al. 2002). This analysis identifies groups of sample sites that are most similar and that are geographically meaningful. SAMOVA uses the statistics derived from an AMOVA, and incorporates geographical information on sampling sites (i.e. geographical distances among sites) with a simulated annealing approach to maximise the Φ_{CT} among groups of populations as well as identifying possible genetic barriers between them, without pre-defining populations (Dupanloup et al. 2002). Thus, SAMOVA defines groups of samples that are geographically homogeneous and also maximally differentiated from each other (Dupanloup et al. 2002).

microsatellite variation

Microsatellite data were checked for presence of null alleles, large allele dropout and errors in scoring due to stutter bands using Micro-Checker software version 2.2.3 (Oosterhout et al. 2004). Each locus and each site was also tested for deviation from Hardy-Weinberg equilibrium (HWE) in Arlequin with significance of deviations in observed vs expected heterozygosity tested using Exact tests (Guo and Thompson 1992). Inbreeding coefficient

(F_{IS}) for each locus and each site were tested using FSTAT version 2.9.3.2 (Goudet 1995). The possibility of linkage disequilibrium (LD) among loci was investigated using the method of Slatkin and Excoffier (1996) in Arlequin with 1000 permutations ($\alpha = 0.05$). *P* values were adjusted using the Bonferroni correction. Measures of genetic variation including; number of alleles and expected and observed heterozygosities were calculated in Arlequin. An analysis of molecular variance (AMOVA) was used to examine the amount of genetic variation partitioned within and among populations (Excoffier et al. 1992). AMOVA and hierarchical AMOVA tests for the microsatellite data set were organised in a similar way to the mtDNA analysis described above. Further analyses of microsatellite data were carried out for the pooled temporal samples from each site as there was also no temporal variation for microsatellite data also.

We assessed the microsatellite data set for sufficient statistical power to detect genetic heterogeneity (i.e. estimating the potential for Type II error) at various true levels of divergence using POWSIM software as described by Ryman et al. (2006). This program uses sample sizes, number of loci, and allele frequencies, and simulates genetic sampling from multiple populations that have drifted apart to a predefined expected degree of divergence defined as F_{ST}. Samples from these populations are used to test the hypothesis of genetic homogeneity and to estimate the α error at each locus separately, using both Fisher's exact test and traditional chi-square approaches, respectively.

Presence of multiple SJT breeding units was first tested using the multi-locus microsatellite data set in STRUCTURE version 2.2 (Pritchard et al. 2000). STRUCTURE uses a model-based full Bayesian Markov Chain Monte Carlo (MCMC) approach that clusters individuals to minimise Hardy-Weinberg disequilibrium and gametic phase disequilibrium between loci within groups. For each value of *K*, (i.e. the number of genetically distinct populations), the MCMC scheme was run with a burn-in period of 100,000 steps and a chain

length of 1,000,000 replicates following the admixture model. Multiple runs were performed to evaluate the reliability of the results, and the number of populations was then determined from posterior probabilities of K estimated, assuming that they originated from 1 to 7 populations (i.e. K = 1 to 7). In the case of migration rates between populations not being equal, K values may be incorrectly estimated, so values of ΔK following Evanno et al. (2005) were also calculated for each value of K. Individuals were assumed to have been correctly assigned to a population when their q-value (i.e. their posterior probability of belonging to an original population) was at least 80% for the population (Pritchard et al. 2000).

Results

Mitochondrial DNA variation

Genetic variation

Genetic analyses were conducted on 324 individuals sampled from five fishing grounds around Sri Lanka (NE, WE, TA, KM and TR), and single sites from the Maldive Islands (MD) and Laccadive Islands (LC) (Figure 1 and Table 1). MtDNA haplotype sequence data produced alignment of a 488 bp fragment that included partial ATP*ase* 6 and the ATP*ase* 8 gene regions. A total of 52 nucleotide sites were variable (segregating sites) (Table S1, Electronic Appendix). Statistical tests of neutrality, Fu's F_S and Tajima's D tests for each sample population, showed that sample populations did not deviate from the expected neutral model of evolution (data not shown).

Polymorphic sites defined a total of 49 unique haplotypes (Table S1 and S2 in the Electronic Appendix). Analysis of variable sites showed that all mutations were synonymous. 29 haplotypes were singletons, and the most abundant haplotype (haplotype 6, frequency 23.46%) was present at 6 out of 7 sample sites (absent from site MD) (Table S2, Electronic Appendix). The second most abundant haplotype (haplotype 37, frequency 8.33%) was present at 5 sites. Overall nucleotide diversity, and the average number of pairwise nucleotide differences were 0.007 and 3.855, respectively. Population genetic summary statistics are presented in Table 1.

(Insert Table 1)

Phylogenetic relationships

Haplotypes grouped into two distinct, divergent clades (Tamura and Nei corrected mean divergence among clades was 0.015). Haplotypes belonging to both clades were found at all sample sites. Clade I contained 37 of the 49 haplotypes identified while only 12 haplotypes

were present in Clade II. The parsimony network (Figure 2) shows that the most common haplotype (Ht6), was ancestral in Clade I (occurs in the centre of network) while Ht37 was ancestral in Clade II. Uncorrected pairwise divergence among haplotypes in the parsimony cladogram ranged from 0 to 3.7%. Collections at all sample sites represent mixtures of haplotypes from both clades that are independent of year of collection or relative sample size at a site.

(Insert Figure 2)

Population differentiation

The pattern of SJT mtDNA haplotype diversity observed at all sites (Electronic Appendix S2 and Figure 2) consisted of a single haplotype (sites TA and MD) or two haplotypes (sites NE, WE, KM, TR, and LC) present at high frequency. Common haplotype frequencies however, varied widely among sites. A large number of singleton haplotypes were also present at individual sites (29 of 49 possible haplotypes).

Hierarchical analysis that incorporate Tamura and Nei's genetic distance in AMOVA is summarised in Table 2. Across the total sample collection, there was significant genetic differentiation among sites (global Φ_{ST} = 0.2029, P<0.001) when collections over years within sites were pooled. The entire sample collection, when grouped into year classes to assess the impact of temporal collections, showed no significant genetic differentiation among groups for years (among year collections in 2001, 2002, 2003, and 2004) (Φ_{CT} = 0.0002, P = 0.8739). As no significant genetic variation was evident among year total collections (i.e. among 2001, 2002, 2003 and 2004 collections, irrespective of sampling site), this implied that overall genetic composition of sampled SJT populations remained temporally stable across the study period. Significant genetic differentiation was evident however, among sites within all sample years indicating that sampled populations showed consistent spatial heterogeneity for genetic variation (Φ_{SC} = 0.2030, P < 0.001) over years. Clade-wise hierarchical AMOVA,

325 while predictably showing significant differentiation among clades (Φ_{CT} = 0.563, P < 0.001),
326 also detected significant structure among samples within clades (Φ_{SC} = 0.295, P < 0.001).
327 (Insert Table 2)
328 Further analysis of individual clades revealed that among site genetic differentiation
329 was limited only to Clade I individuals (Φ_{SC} = 0.2136, P < 0.001) (data not shown), while no
330 significant genetic variation was evident for Clade II individuals. Thus, overall genetic
331 differentiation among sites resulted essentially from spatial variation in the distribution of
332 Clade I individuals among sites. This may have been influenced however, by the relatively
333 low numbers of Clade II (n = 43) individuals in the total sample and hence there may have
334 been insufficient power to reject the null hypothesis of panmixia for this Clade.
335 Spatial genetic variation among SJT sampling sites was tested at higher resolution
336 using pairwise Φ_{ST} analysis. Pairwise Φ_{ST} estimates for the entire collection are shown in
337 Table 3. Highly significant genetic differentiation was evident between the majority of pairs
338 of sites. After Bonferroni correction for multiple comparisons, 16 pairs of sites were
339 significantly differentiated. AMOVA analyses overall, showed that while significant spatial
340 genetic differentiation was evident among sampled SJT populations, in general, the entire SJT
341 collection remained genetically stable over time (i.e. the same haplotypes remained in similar
342 frequencies within sites across years).
343 (Insert Table 3)
344 Genetic differentiation among sampled sites not only resulted from relative admixture
345 of the two clades at individual sites, but was also influenced by divergent frequencies of
346 Clade I haplotypes among sites. As a consequence, Spatial Analysis of Molecular Variance
347 (SAMOVA) analysis was employed for the total SJT sample and Clade I to assess whether
348 genetically homogeneous SJT groups were present across the study area.

As the best significant homogeneous grouping, SAMOVA indicated three genetically differentiated SJT population groups (Table 4). Specifically, the analysis recognised; KM and MD as discrete populations, and all remaining sites (NE, WE, TA, TR, LC) together as a third discrete population. Similarly, SAMOVA analysis of Clade I individuals also showed additional structuring within the clade, with three genetically differentiated groups identified as i) TR, ii) MD, and iii) all other sites (NE, WE, TA, KM, and LC), that appear to lack any strong geographical pattern.

(Insert Table 4)

microsatellite variation

Genetic variability estimates, Hardy-Weinberg and linkage equilibrium

No null alleles, large allele drop out or error scoring were detected using Micro-Checker (95% confidence interval) for the six SJT microsatellite loci (UTD73, UTD203, UTD328, UTD535, UTD523, and UTD172), except at locus UTD328. The analysis showed that locus UTD328 results could have been affected by null alleles. Subsequent analyses in AMOVA both including, then excluding locus UTD328 data however, produced similar outcomes. All six loci were therefore included in all further analyses. Descriptive statistics for the six loci are summarised in Table 5. A small number of individuals could not be scored at specific loci due to amplification problems.

Sample populations were then tested for conformation to Hardy-Weinberg equilibrium. Significant heterozygote deficiencies (P<0.001) were observed in some comparisons particularly at UTD328 (in all 7 collections) supporting the hypothesis of the presence of null alleles at this locus. Other loci (except for locus UTD203) also showed significant heterozygote deficiencies at two to six sites (Table 5).

(Insert Table 5)

374 In addition to heterozygote deficiencies, some loci showed linkage disequilibrium but
375 after Bonferroni correction this constituted only two out of 15 comparisons (data not shown).
376 Heterozygote deficiencies with evidence for linkage disequilibrium, in combination, is a
377 pattern consistent with admixture of genetically heterogeneous SJT groups among the
378 sampled populations (e.g. So et al. 2006).
379
380 ***Population structure***
381 Hierarchical AMOVA analysis results for SJT microsatellite data are summarised in Table 6.
382 The result for the entire data set showed significant genetic differentiation among sites (F_{ST} =
383 0.025, P<0.001). Pairwise F_{ST} analysis of microsatellite data identified that NE, TR, KM, and
384 LC populations were significantly genetically differentiated from most other populations, a
385 result in concord with mtDNA pairwise Φ_{ST} data (Table 3 and Table 7).
386 (Insert Table 6)
387 Hierarchical AMOVA was conducted to determine temporal and spatial variation
388 among sites. No significant temporal genetic differentiation was evident among year groups
389 sampled from 2001 to 2004 (F_{CT} = -0.0096, P = 0.861). Significant spatial genetic variation
390 was detected however, among sites within years (F_{SC} = 0.0326, P<0.001) (Table 6). These
391 results were also consistent with mtDNA hierarchical AMOVA results.
392 (Insert Table 7)
393 As two divergent clades were identified in the mtDNA analysis, microsatellite data
394 were tested to assess the pattern of differentiation for nDNA among clades. No significant
395 differentiation was observed when individuals were grouped by their mtDNA clade type,
396 although significant differentiation was evident among populations within clades (Table 6).
397 The very high within population variation (>97%) and moderate among population within
398 clade variation (2%) could have contributed to a lack of significant differentiation among

399 clades. Therefore, we carried out analysis of the microsatellite data using POWSIM, to test
400 whether the microsatellite data set possessed sufficient statistical power to detect true genetic
401 differentiation, if it were present, that may have gone undetected in hierarchical AMOVA
402 analysis.

403 Separate POWSIM analyses of the seven sampled sites and two mtDNA clades
404 showed clearly that our nDNA data set had high statistical power for detecting true genetic
405 differentiation among populations, but insufficient power to differentiate the two
406 (microsatellite) clades. The microsatellite data set from 7 sample sites was capable of
407 detecting a true F_{ST} estimate of as low as 0.0025 with a probability of 98%, and a true
408 differentiation of 0.005 with a probability of 100%. However, the probability of detecting a
409 true F_{ST} estimate of as low as 0.0005 (from AMOVA analysis) of microsatellite data among
410 mtDNA clades was >42.0%. The α error (corresponding with the probability of obtaining
411 false significance when the true $F_{ST} = 0$) was close to 5% in all cases.

412 We further tested the nDNA data for admixture using the program STRUCTURE
413 (Pritchard et al. 2000). STRUCTURE analysis recognised two distinct groups among the
414 samples using only the 'Log probability of data [$L(K)$]' approach (as described in Pritchard et
415 al. 2000). The number of clusters that fitted our data best was not clear because the $L(K)$ did
416 not reach a plateau after $K = 1$ (Table 8). The ΔK statistic described by Evanno et al. (2005)
417 however, showed clearly presence of two distinct groups (the highest ΔK was obtained with K
418 = 2). All seven sample sites were mixtures of the two identified groups in different
419 proportions (Figure 3a) with the vast majority of individuals assigned with 80% or greater
420 probability to one microsatellite group or the other.

421 (Insert Table 8)

422 The fact that two well-supported and genetically identifiable groups were detected
423 using both mtDNA and nDNA markers would strongly suggest that the samples from each

site represented a mixture of two separate breeding units. However, for this to be true, cyto-nuclear disequilibrium among the two marker systems should be evident. We therefore tested this hypothesis by repeating the microsatellite assignment analysis in STRUCTURE for K = 2 but separating samples into groups based on their respective mtDNA clades. In this analysis, Clade I individuals had an approximate 50/50 split of confidently assigned individuals from both microsatellite groups. The pattern was different for Clade II individuals where the split was approximately 75/25 (Figure 3b). From this result, it appears that there is little evidence to suggest that the two groups identified by both marker systems have resulted from the same evolutionary/demographic processes.

(Insert Figure 3)

Discussion

Harvesting fish populations has the potential to cause three types of genetic change: alter population subdivision, cause genetic variation levels to decline, and produce selective genetic changes (Allendorf et al. 2008). To sustain the productivity of any harvested fish population, it is crucial to incorporate genetic considerations into their management. Management plans should be developed by applying basic genetic principles combined with molecular genetic monitoring to minimise any harmful genetic change (Allendorf et al. 2008). Understanding the genetic changes and evolutionary responses of exploited populations is also crucial for design of management strategies aimed at sustainable exploitation of natural biological resources (Walsh et al. 2006). SJT in the Indian Ocean are currently managed as a single stock yet results presented in this study strongly suggest that mixed stocks exist in waters surrounding Sri Lanka.

Phylogenetic relationships

Phylogenetic analyses provide strong support for presence of two divergent mtDNA clades in the north western region of the Indian Ocean. Pairwise divergence among haplotypes in the parsimony cladogram (based on the ATP*ase* region) ranged from 0 - 3.7% while Tamura and Nei corrected mean divergence among clades was 1.5%. Recent studies of other tuna and billfish species have also reported multiple highly divergent intra-specific mtDNA lineages both within or among ocean basins notably for, Atlantic bigeye tuna (Martinez et al. 2005), Atlantic bonito (Vinas et al. 2004), and several billfishes (Scombroidei: Xiphidae) including blue marlin (Buonnacorsi et al. 2001), sailfish (Graves and McDowell 2003), and swordfish (Buonnacorsi et al. 2001; Graves and McDowell 2003; Alvarado Bremer et al. 2005). Two sympatric clades have also been recognised for Atlantic bonito in the Mediterranean Sea (Vinas et al. 2004) while for some billfishes (Buonnacorsi et al. 2001; Graves and McDowell

2003; Alvarado Bremer et al. 2005) and bigeye tuna (Alvarado Bremer et al. 1998; Martinez et al. 2005), multiple sympatric clades were reported in the Atlantic Ocean. MtDNA heterogeneity observed in these studies and in the present study may have been caused by common vicariance events during Pleistocence glacial maxima that resulted in isolation of populations by reduction in availability of tropical marine habitats due to decreases in ocean water temperature (e.g.Graves and McDowell 2003; Vinas et al. 2004; Alvarado Bremer et al. 2005). Further, potential unidirectional gene flow of formerly allopatric populations during interglacial periods providing secondary contacts may have resulted in asymmetrical distribution of clades (Alvarado Bremer et al. 2005). While few studies have focused their attention on SJT, two divergent mtDNA lineages of SJT were detected in the Atlantic and Pacific Oceans (Ely et al. 2005), although not specifically reported at the time. The level of divergence shown by the Ely et al. (2005) study for mtDNA control region sequence is consistent with that of ATP*ase* divergence reported here.

Population structure

Both mtDNA and nDNA microsatellite data show strong evidence for spatial genetic heterogeneity among the SJT sample sites around Sri Lanka and adjacent areas in the north western Indian Ocean. In addition, evidence for fine geographical scale genetic heterogeneity was confirmed by pairwise analysis of both mtDNA and microsatellite data. It appears however, that fine scale structuring was influenced primarily by differing frequencies of Clade I and II individuals and nDNA genetic group individuals identified in the mtDNA and microsatellite analyses, respectively. It is possible that this is a random sampling effect.

From a fisheries management perspective, it was important to determine whether individuals belonging to the two mtDNA clades, interbreed. This however, requires analysis of nDNA because information is only available from the maternal parent for mtDNA.

485 STRUCTURE analysis of nDNA microsatellite data identified the existence of two
486 genetically distinct groups and hence this suggests that individuals belonging to the two
487 mtDNA clades may not interbreed. This depends however, on their being a relatively close
488 association between an individual's mtDNA clade membership and their nDNA group
489 identified in the cluster analysis. If we assume that the divergent mtDNA clades have arisen
490 in alloparty (i.e. there are two geographically distinct breeding grounds and there is a high
491 degree of natal site philoparty over a long evolutionary period), and that individuals of two
492 distinct breeding units mix only after reproduction, then there should be a strong correlation
493 between their individual mtDNA clade and the alleles they carry at nDNA loci. This was not
494 the case here. A simple explanation for this pattern is that there was some historical secondary
495 contact thereby admixing the mtDNA clades. More recently however, there has been a
496 constraint on gene flow between breeding grounds resulting in nuclear allele frequencies
497 diverging due to genetic drift which has provided the signature of differentiated stocks seen in
498 the microsatellite analysis.

499 The two different stocks may spawn in distant areas of the Indian Ocean, following
500 which, juveniles may disperse towards Sri Lanka because this area is a highly fertile feeding
501 ground. Satellite images have shown very high primary productivity/chlorophyll 'a'
502 concentrations around Sri Lanka during the southwest monsoon period (Wiggert et al. 2006).
503 High primary productivity will lead to high zooplankton levels and large schools of bait fish
504 on which SJT feed. As evidence for this relationship, the Wadge Bank near to Sri Lanka is
505 recognised as an important tuna fishing ground. Other explanations for physical mixture of
506 the two SJT clades at sampling sites include; differential passive transport of larvae belonging
507 to the different clades to fishing grounds in Sri Lanka by monsoonal currents in the Indian
508 Ocean.

It should be recognized however, that there was no strong pattern in the geographical distribution of the groups to provide insight into the regions for respective breeding grounds (i.e. there is no discernable genetic group frequency cline for either group from east to west). Therefore, while we can be confident that all samples in this study are a mixture of two discrete breeding units, there is no evidence to suggest that the fishing grounds around Sri Lanka represent a zone of admixture *per se*. Furthermore, as the present study only analysed data on SJT samples from the north western Indian Ocean, the results do not preclude presence of additional SJT spawning stocks and spawning areas in other regions of the Indian Ocean. More work is warranted therefore, aimed at confirming the natal homing of the two SJT stocks, identification of potential spawning sites in the Indian Ocean and more extensive genetic stock structure analysis of SJT populations across the region.

It is obvious from the results presented here that SJT stock structure is more complex than had originally been thought, i.e. that SJT populations within ocean basins are not necessarily homogeneous. Given the importance of SJT as a regional food resource to millions of people in the Indian Ocean, effective management of this resource in the future, will require developing a better understanding of the scale at which populations in the region should be managed and secondly, conserving unique areas where discrete stocks may reproduce (spawning grounds).

Table 1 Descriptive statistics for SJT samples. No. of haplotypes (h), No. of polymorphic sites (*S),* Gene diversity (H_d), mean pairwise nucleotide difference (*k*), Nucleotide diversity (π), Equilibrium heterozygosity per site estimated from *S* (θ_s)

Population	Location	n	h	S	H_d	*k*	π	θ_s
Negombo (NE)	$79^0$18`, $6^0$057`	53	11	22	0.706	2.506	0.005	1.646
Weligama (WE)	$80^0$18`, $5^0$034`	52	12	18	0.800	3.165	0.006	1.409
Tangalle (TA)	$81^0$14`, $5^0$042`	41	10	20	0.431	1.404	0.002	1.667
Kalmunei (KM)	$82^0$29`, $7^0$008`	54	15	20	0.797	5.420	0.011	1.513
Trincomalee (TR)	$81^0$51`, $8^0$058`	49	7	14	0.669	2.622	0.005	1.181
Laccadive (LC)	$72^0$31`, $11^0$01`	48	7	17	0.733	3.531	0.007	1.383
Maldives (MD)	$73^0$09`, $4^0$ 20`	27	9	23	0.649	5.096	0.010	2.207
Clade I		281	37	38	0.959	2.639	0.005	6.115
Clade II		43	12	28	0.773	4.686	0.009	6.471
Total collection		324	49	52	0.965	3.855	0.007	8.180

549 **Table 2** Results of AMOVA testing genetic structuring of SJT samples based on
550 mitochondrial ATP*ase* region sequence data.

	Structure tested	Observed partition		Φ statistics
		Variance	% of Total	
1	Total collection (2001,2002,2003,2004) – one gene pool			
	Among populations	0.6286	20.29	Φ_{ST} = 0.2029***
	Within populations	2.4689	79.71	
2	Among years			
	Among years (TEMPORAL)	0.0007	0.02	Φ_{CT} = 0.0002
	Among populations within years (SPATIAL)	0.6280	20.30	Φ_{SC} = 0.2030***
	Within populations	2.4647	79.68	Φ_{ST} = 0.2032
3	Clade-wise			
	Between two clades	1.9336	56.35	Φ_{CT} = 0.5635***
	Among samples within clades	0.4457	12.86	Φ_{SC} = 0.2945***
	Within populations	1.0677	30.80	Φ_{ST} = 0.6920

551 (***, P<0.001)
552
553
554
555
556
557
558

559 **Table 3** MtDNA pairwise Φ_{ST} among SJT sampling sites after Bonferroni correction for

560 entire collection. ($\alpha = 0.05/21 = 0.002$)

	NE	WE	TA	KM	TR	LC	MD
NE							
WE	0.116**						
TA	0.081**	0.059					
KM	0.271**	0.186**	0.237**				
TR	0.171**	0.069*	0.117*	0.229**			
LC	0.187**	0.070	0.142**	0.267**	0.111**		
MD	0.142**	0.065	0.071	0.116	0.112**	0.116**	

561 (*, $P< 0.002$; **, $P<0.0001$)

562

563

564

565

566

567

568

569

570

571

572

573

574

Table 4 Population structure based on mtDNA differentiation of SJT (in SAMOVA). The row in bold type indicates the details of geographically meaningful groups with maximum genetic differentiation

No. of Groups	Structure	Variation among groups	Variation %	Φ_{CT}	*P*
2	(KM) (NE,WE,TA,TR,LC,MD)	0.328	14.95	0.149	0.135
3	**(KM) (MD) (NE,WE,TA,TR,LC)**	**0.269**	**12.68**	**0.126**	**0.048**
4	(KM) (MD) (NE,TA) (WE,TR,LC)	0.234	11.54	0.115	0.003
5	(KM) (MD) (TR) (NE,TA) (WE, LC)	0.230	11.52	0.115	0.006
6	(KM) (MD) (TR) (NE) (LC) (WE,TA)	0.261	13.13	0.131	0.046

Table 5 Descriptive statistics for six microsatellite loci among SJT collections. Significant probability values after Bonferroni correction (α = 0.05/42 = 0.0011). Number of samples (*n*), Number of alleles (*a*).

Sample		Locus						Average across loci
		UTD535	**UTD523**	**UTD172**	**UTD328**	**UTD203**	**UTD73**	
NE	***n***	36	36	40	45	43	37	39
	a	14	21	7	22	5	7	12.67
	He	0.916	0.949	0.725	0.935	0.344	0.727	0.77
	F_{is}	0.122	0.124	-0.139	0.140*	0.343	-0.041	0.09
WE	***n***	44	52	43	52	39	41	45
	a	26	25	9	16	8	12	16.00
	He	0.962	0.951	0.725	0.904	0.423	0.809	0.80
	F_{is}	0.175*	0.028*	-0.287*	0.189*	-0.049	0.143	0.03
TA	***n***	33	33	38	44	35	24	34
	a	15	23	11	18	8	6	13.50
	He	0.914	0.951	0.845	0.899	0.459	0.738	0.80
	F_{is}	0.107	0.045	-0.027*	0.101*	0.05	0.291*	0.09
KM	***n***	45	36	42	47	47	45	43
	a	14	21	8	16	5	13	12.83
	He	0.912	0.931	0.681	0.861	0.464	0.845	0.78
	F_{is}	0.075	0.046	-0.37*	0.263*	-0.245	0.145*	-0.01
TR	***n***	35	32	32	46	46	33	37
	a	15	20	6	20	3	15	13.17
	He	0.913	0.938	0.558	0.903	0.231	0.881	0.74
	F_{is}	0.094	0.43*	0.088	0.420*	0.291	0.000*	0.22
LC	***n***	33	35	44	48	48	47	42
	a	16	22	7	24	7	8	14.00
	He	0.916	0.946	0.672	0.932	0.409	0.751	0.77
	F_{is}	0.176*	0.157*	-0.323*	0.213*	0.072	-0.02	0.05
MD	***n***	41	43	45	49	51	51	46
	a	16	27	10	13	7	11	14.00
	He	0.914	0.961	0.826	0.853	0.463	0.72	0.79
	F_{is}	0.175	0.422*	-0.131*	0.224*	-0.006	0.047	0.12

(*, P<0.0001)

Table 6 AMOVA results testing spatial and temporal genetic structuring of SJT samples based on microsatellite data.

	Structure tested	Observed partition		*F* statistics
		Variance	% Total	
1	Total collection (2001,2002,2003,2004)			
	Among populations	0.045	2.56	F_{ST} = 0.0256***
	Within populations	1.728	97.44	
2	Year-wise collection (2001, 2002, 2003, 2004)			
	Among groups	-0.0171	-0.97	F_{CT} = -0.0096
	Among populations (of the same year) within groups	0.0583	3.30	F_{SC} = 0.0326***
	Within populations	1.7278	97.67	F_{ST} = 0.0233
3	Clade-wise			
	Between two clades	0.0009	0.05	
	Among samples within clades	0.0404	2.28	F_{CT} = 0.0005
	Within populations	1.7321	97.66	F_{SC} = 0.0228***

(*** P< 0. 001)

Table 7 Microsatellite pairwise F_{ST} among SJT collections after Bonferroni correction (initial $\alpha = 0.05/21 = 0.002$)

	NE	WE	TA	KM	TR	LC	MD
NE							
WE	0.0080**						
TA	0.0193***	0.0062					
KM	0.0266***	0.0018	0.0048				
TR	0.0722***	0.0605***	0.0618***	0.0744***			
LC	0.0101	0.0003	0.0127***	0.0132***	0.0778***		
MD	0.0218***	0.0046	-0.0034	0.0174***	0.0559***	0.0159***	

(*** $P<0.001$)

635 **Table 8** Log probability and Δ*K* for clusters using Bayesian assignment test in STRUCTURE

636

K (Number of clusters)	*L(K)* Average of log probability	Standard deviation	*L'(K)*	| *L''(K)* |	Δ*K*
1	-8412.7	8.94			
2	-8246.5	21.20	166.2	161.2	7.60
3	-8085.3	25.25	161.2	102.2	4.05
4	-7983.1	28.79	102.2	22	0.76
5	-8005.1	33.82	-22	135.7	4.01
6	-7869.4	33.83	135.7	30.3	0.90
7	-7899.7	37.64			

637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654

Figure 1 SJT sampling sites. Abbreviations correspond to sampling localities; (NE), Negombo; (WE), Weligama; (TA), Tangalle; (KM), Kalmunei; (TR), Trincomalee; (LC), Laccadive; (MD), Maldives

Figure 2 Parsimony cladogram of SJT haplotypes showing the evolutionary relationships among haplotypes. Each circle represents a unique haplotype in the sample, and the size of each circle represents the relative frequency of each haplotype (see legend). Patterns and their percentage in each circle represent the presence of each haplotype at individual sites and their relative abundance respectively. Black dots between haplotypes represent haplotypes that were not sampled in the study

Figure 3(a) A Bar plot of Bayesian clustering analysis (K = 2) for SJT microsatellite data set (364 individuals; 6 loci) performed using STRUCTURE. Each bar represents proportional probability of assignment to each genetic group. Abbreviations correspond to sampling localities; (NE), Negombo; (WE), Weligama; (TA), Tangalle; (KM), Kalmunei; (TR), Trincomalee; (LC), Laccadive; (MD), Maldives

Figure 3(b) A Bar plot of Bayesian clustering analysis (K = 2) for SJT two mtDNA clades (364 individuals; 6 loci) performed using STRUCTURE.

Acknowledgements

This paper forms a component of the PhD research of Sudath T. Dammannagoda on population genetic structure and migratory patterns of yellowfin tuna and skipjack tuna in Sri Lankan waters (Indian Ocean). We acknowledge Dr. Robert D. Ward, Marine and Atmospheric Research Division, CSIRO, Hobart, Australia for helpful comments and support as the external associate supervisor for this PhD program. We thank Maria Menezes and an anonymous reviewer for comments and suggestions that greatly improved the manuscript. All samples were collected and all experiments were performed in compliance with relevant laws and requirements.

References

Allendorf , F.W., England, P.R., Luikart, G., and Ritchie, P.A. 2008. Genetic effects of harvest on wild animal populations. Trends in Ecol. Evol. **23**: 327-337.

Alvarado Bremer, J.R., Stequert, B., Robertson, N.W., and Ely, B. 1998. Genetic evidence for inter-oceanic subdivision of bigeye tuna (*Thunnus obesus*) populations. Mar. Biol. **132**: 547-557.

Alvarado Bremer, J.R., Vinas, J., Mejuto, J., Ely, B., and Pla, C. 2005. Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. Mole. Phylo. and Evol. **36**: 169-187.

Boustany, A.M., Reeb, C.A., and Block, B.A. 2008. Mitochondrial DNA and electronic tracking reveal population structure of Atlantic bluefin tuna (*Thunnus thynnus*). J. Mar. Biol. **156**: 13-24.

Buonaccorsi, V.P., McDowell, J.R., and Graves, J.E. 2001. Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). Mole. Ecol. **10**: 1179-1196.

Campbell, N.J.H., Harris, F.C., Elphinstone, M.S., and Baverstock, P.R. 1995. Outgroup heteroduplex analysis using temperature gradient gel electrophoresis: high resolution, large scale, screening of DNA variation in the mitochondrial control region. Mole. Ecol. **4**: 407-418.

Carlsson, J., McDowell, J.R., Diaz-James, P., Carlsson, J.E.L., Boles, S.B., Gold, J.R., and Graves, J.E. 2004. Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranian Sea. Mole. Ecol. **10**: 1179-1196.

751 Carlsson, J., McDowell, J.R., Carlsson, J.E.L., Olafsdottir, D., and Graves, J.E. 2006.
752 Genetic heterogeneity of Atlantic bluefin tuna caught in the eastern North Atlantic
753 Ocean south of Iceland. J. Mar. Sci. **63**: 1111-1117.
754 Chand, V., de Bruyn, M., and Mather, P.B. 2005. Microsatellite loci in the eastern form of
755 the giant freshwater prawn (*Macrobrachium rosenbergii*). Mole. Ecol. Notes **5**: 308-
756 310.
757 Chow, S., and Ushiyama, S. 1995. Global population structure of albacore (*Thunnus*
758 *alalunga*) inferred by RFLP analysis of the mitochondrial ATPase gene. Mar. Biol. **123**:
759 39-45.
760 Clement, M., Posada, D., and Crandall, K.A. 2000. TCS: a computer programme to estimate
761 gene genealogies. Mole. Ecol. **9**: 1657-1660.
762 Crandall, K.A. 1996. Multiple inter species transmissions of human and simion T-cell
763 leukemia/ Lymphoma virus type I sequences. Mole. Biol. and Evol. **13**: 115-131.
764 Dammannagoda, S.T., Hurwood, D.A., and Mather, P.B. 2008. Evidence for fine
765 geographical scale heterogeneity in gene frequencies in yellowfin tuna (*Thunnus*
766 *albacares*) from the north Indian Ocean around Sri Lanka. Fish. Res. **90**: 147-157.
767 Dupanloup, I., Schneider, S., and Excoffier, L. 2002. A simulated annealing approach to
768 define the genetic structure of populations. Mole. Ecol. **11**: 2571-258.
769 Ely, B., Vinas, J., Alvarado Bremer, J., Black, B., Lucas, L., Covello, K., Labrie, and A.V.,
770 Thelen, E. 2005. Consequences of the historical demography in the global population
771 structure of two highly migratory cosmopolitan marine fishes: the yellowfin tuna
772 (*Thunnus albacares*) and skipjack tuna (*Katsuwonus pelamis*). BMC Evol. Biol. **5**: 19.
773 Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of
774 individuals using the software STRUCTURE: a simulation study. Mol. Ecol. **14**: 2611-
775 2620.

776 Excoffier, L., Smouse, P.E., and Quattro, J.M. 1992. Analysis of molecular variance inferred
777 from metric distances among DNA haplotypes: Application to human mitochondrial
778 DNA restriction data. Genetics **131**: 479- 491.
779 FAO 2008. FAO fishery statistical collections: Global tuna nominal catches-online query.
780 http://www.fao.org/fishery/statistics/tuna-catches.
781 Florin, A.B., and Hoglund, J. 2008. Population structure of flounder (*Platichthys flesus*) in
782 the Baltic Sea: differences among demersal and pelagic spawners. Heredity **101**: 27-38.
783 Fu, Y.X. 1997. Statistical tests for neutrality of mutations against population growth,
784 hitchhiking and background selection. Genetics **147**: 915-925.
785 Fujino, K., Sassaki, K., and Okumura, S. 1981. Genetic diversity of skipjack tuna in the
786 Atlantic, Indian and Pacific Oceans. Bull. Jap. Society Sci. Fish. **47**: 215-222.
787 Garber, A.F., Triangali, M.D., and Franks, J.S. 2005. Population genetic and
788 phylogeographic structure of wahoo, *Acanthocybium solandri*, from the western central
789 Atlantic and central Pacific Oceans. Mar. Biol. **147**: 205-214.
790 Goudet, J. 1995. FSTAT (vers. 2.9.3.2): a computer program to calculate F-statistics. J.
791 Heredity. 86: 485-486.
792 Graves, J.E., Ferris, S.D., and Dizon, A.E. 1984. High genetic similarity of Atlantic and
793 pacific skipjack tuna demonstrated with restriction endonuclease analysis of
794 mitochondrial DNA. Mar. Biol. **79**: 315-319.
795 Graves, J.E., and McDowell, J.R. 2003. Stock structure of the worlds Istio-phorid billfishes;
796 a genetic perspective. Mar. Fresh. Res. **54**: 287-298.
797 Guo, S., and Thompson, E. 1992. Performing the exact test of Hardy-Weinberg proportion
798 for multiple alleles. Biometrics **48**: 361-372.
799 Hall, T.A. 1999. BioEdit: A user friendly biological sequence alignment editor and analysis
800 programme for Windows 95/98/NT. Nuc. Acid. Symp. **41**: 95-98.

801 Hoolihan, J.P., Premanandh, J., D'Aloia-Palmeiri, M.A., and Benzie, J.A.H. 2004.
802 Intraspecific phylogeographic isolation of Arabian Gulf sailfish *Istiophorus platypterus*
803 inferred from mitochondrial DNA. Mar. Biol. **145**: 465-475.
804 Indian Ocean Tuna Commission 2007. Report of the tenth session of the scientific
805 committee. Indian Ocean Tuna Commission.Victoria, Seychelles, 5-9 November, 2007.
806 Available from
807 http://www.iotc.org/English/documents/doc_proceed_details.php?docid=1255.
808 Knutsen, H., Jorde, P.E., Andre, C., and Stenseth, C.H.R. 2003. Fine-scaled geographical
809 population structuring in a highly mobile marine species: the Atlantic cod. Mole. Ecol.
810 **12**: 385-394.
811 Kotoulas, G., Magoulas, A., Tsimenides, N., and Zouros, E. 1995. Marked mitochondrial
812 DNA differences between Mediterranean and Atlantic populations of the swordfish,
813 *Xiphias gladius*. Mol. Ecol. **4**: 473-481.
814 Lessa, E.P., and Applebaum, G. 1993. Screening techniques for detecting allelic variation in
815 DNA sequences. Mole. Ecol. **2**: 119-129.
816 Martinez, P., Gonzalz, E.G., Castilho, R., and Zardoya, R. 2005. Genetic diversity and
817 historical demography of Atlantic bigeye tuna (*Thunnus obesus*). Mole. Phylo. Evol. **39**:
818 404-416.
819 Menezes, M.R., Ikeda, M., and Taniguchi, N. 2005. Genetic variation in skipjack tuna
820 *Katsuwonus pelamis* (L.) using PCR-RFLP analysis of the mtDNA D-loop region. J.
821 Fish Biol. **68**: Supplement A, 156-161.
822 Menezes, M.R., Noguchi, D., Nakajima, M., and Taniguchi, M. 2008. Microsatellite
823 development and survey of genetic variation in skipjack tuna *Katsuwonus pelamis*. J.
824 Fish Biol. **73**: 463-473.

825 Miller, S.A., Dykes, D.D., and Polesky, H.F. 1988. A simple salting out procedure for
826 extracting DNA from nucleated cells. Nuc. Acid Res. **16**: 1215.
827 Nei, M. 1987. Molecular Evolutionary Genetics. New York, Columbia University Press.
828 Oosterhout, C., Hutchinson, W.F., Wills, P.M.D., and Shipley, P. 2004. Microchecker:
829 software for identifying and correcting genotyping errors in microsatellite data. Mole.
830 Ecol. Notes **4**: 535-538.
831 Palsboll, P.J., Berube, M., and Allendorf, F.W. 2007. Identification of management units
832 using population genetic data. Trend. Ecol. Evol. **22**: 11-16.
833 Pritchard, J.K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using
834 multi locus genotype data. Genetics **155**: 945-959.
835 Reeb, C.A., Arcangeli, L., and Block, B.A. 2000. Structure and migration corridors in
836 Pacific populations of the swordfish *Xiphius gladius* as inferred through analyses of
837 mitochondrial DNA. Mar. Biol. **136**: 1123-1131.
838 Rice, W.J. 1989. Analysing tables of statistical tests. Evolution **43**: 223-225.
839 Rozas, J., Sanchez-DelBarro, J.C., Messeguer, X., and Rozas, R. 2003. DnaSp, DNA
840 polymorphism analyses by the coalescent and other methods. Bioinformatics **19**: 2496-
841 2497.
842 Ruzzante, D.E., Taggart, C.T., Cook, D. 1998. A nuclear DNA basis for shelf- and bank-
843 scale population structure in northwest Atlantic cod (*Gadus morhua*): Labrador to
844 Georges bank. Mole. Ecol. **7**: 1633-1680.
845 Ryman, N., Palm, S., Andre, C., Carvalho, G.R., Dahlgren, T.G., Jorde, P.E., Laikre, L.,
846 Larsson, L.C., Palme, A., and Ruzzante, D.E. 2006. Power for detecting genetic
847 divergence; difference between statistical method and marker loci. Mole. Ecol. **15**:
848 2031-2045.

849 Sambrook, J., Fritsch, E.J., and Maniatis, T. 1989. Molecular Cloning. A Laboratory
850 Manual, Second ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New
851 York.
852 Schneider, S., Roesslli, D., and Excoffier, L. 2005. Arlequin version 2.00: a software for
853 population genetics data analysis. Genetics and Biometry Laboratory. University of
854 Geneva, Geneva.
855 Shaklee, J.B., Beacham, T.D., Seeb, L., and White, B.A. 1999. Managing fisheries using
856 genetic data: case studies from four species of Pacific salmon. Fish. Res. **43**: 45-78.
857 Sharp, G.D. 1978. Behavioural and physiological properties of tunas and their effects on
858 vulnerability to fishing gear. *In* The Physiological Ecology of Tunas. *Edited by* Sharp,
859 G.D., and Dizon, A.E. Academic Press, New York. pp. 397-449.
860 Slatkin, M., and Excoffier, L. 1996. Testing for linkage disequilibrium in genotypic data
861 using the expectation-maximization algorithm. Heredity **76**: 377-383.
862 So, N., Maes, G.E., and Volckaert, F.A.M. 2006. High genetic diversity in cryptic
863 populations of the migratory sutchi catfish *Pangasianodon hypophthalmus* in the
864 Mekong River. Heredity **96**: 166-174.
865 Tajima, F. 1983. Evolutionary relationship of DNA sequences to finite populations. Genetics
866 **105**: 437-460.
867 Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA
868 polymorphism. Genetics **123**: 585-595.
869 Tamura, K., and Nei, M. 1993. Estimation of the number of nucleotide substitutions in the
870 control region of mitochondrial DNA in humans and chimpanzees. Mole. Biol. Evol.
871 **10:** 512-526.
872 Templeton, A.R., Boerwinkle, E., and Sing, C.F. 1987. A cladistic analysis of phenotypic
873 associations with haplotypes inferred from restriction endonuclease mapping-I. Basic

874 theory and an analysis of alcohol dehydrogenase activity in *Drosophila.* Genetics **117**:
875 343-351.
876 Templeton, A.R., Sing, C.F. 1993. A cladistic analysis of phenotypic associations with
877 haplotypes inferred from restriction endonuclease mapping-IV. Nested analyses with
878 cladogram uncertainty and recombination. Genetics **134**: 659-669.
879 Vinas, J., Alvarado Bremer, J., and Pla, C. 2004. Phylogeography of the Atlantic bonito
880 (*Sarda sarda*) in the northern mediterranian: the combined effects of historical
881 vicariance, population expansion, secondary invasion and isolation by distance. Mole.
882 Phylo. Evol. **33**: 32-42.
883 Walsh, M.R., Munch, S.B., Chiba, S., and Conover, D.O. 2006. Maladaptive changes in
884 multiple traits caused by fishing: impediments to population recovery. Ecol. Lett. **9**:
885 142-148.
886 Waples, R.S., and Gaggiotti, O. 2006. What is a population? An empirical evaluation of
887 some genetic methods for identifying the number of gene pools and their degree of
888 connectivity. Mol. Ecol. **15**: 1419-1439.
889 Ward, R.D. 2000. Genetics of fish populations. *In* Handbook of Fish Biology and Fishries
890 Management. *Edited by* Hart, P.J.B., Reynolds, J.D. **1**: Blackwell publishing. United
891 Kingdom. pp. 200-224.
892 Ward, R.D., Elliotte, N.G., Grewe, P.M., and Smolenski, A.J. 1994. Allozyme and
893 mitochondrial DNA variation in yellowfin tuna (*Thunnus albacares*) from the Pacific
894 Ocean. Mar. Biol. **118**: 531-539.
895 Watterson, G.A. 1975. On the number of segregating sites in genetical models without
896 recombination. Theor. Pop. Biol. **7**: 256-276.
897 Weir, B.S., and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population
898 structure. Evolution **38**: 1358-1370.

899 Wiggert, J.D., Murtugudde, R.G., and Christian, J.R. 2006. Annual ecosystem variability in
900 the tropical Indian Ocean: Results of a coupled bio-physical ocean general circulation
901 model. Deep-Sea Res. II **53**: 644-676.
902

903 **Electronic Appendix**

904

905 **Table S1** Variable nucleotide sites in the SJT mtDNA ATP region

	1111 1111111222 2222233333 3333333334 4444444444 44	GenBank
	25580245 5666789255 5667901122 4556788990 0112335477 88	Accession
	3946967532 3147051714 7035684806 1092739564 7065781369 01	Number
Ht1	TATGACCAAT TTAAACACCC ATTTGCAAGT CCGGCATACA TTCCAGCATT AT	FJ481378
Ht2	....C .T........ CC....T... ..	FJ481379
Ht3	......T... ....T...C .T........ .C....T... ..	FJ481380
Ht4	.........C ....C .T........ .C....T... ..	FJ481381
Ht5	....G... ....A..... .T........ .CT....... ..	FJ481382
Ht6	.....T.... ....C .T....G... .C....T... ..	FJ481383
Ht7	.... .T....C... .C........ ..	FJ481384
Ht8	....C .T........ AC....T... ..	FJ481385
Ht9	.... TT........	FJ481386
Ht10	....T. .....T...C .T........ .C....T... ..	FJ481387
Ht11	....T. ....AT...C .T........ .C....T... ..	FJ481388
Ht12	.... .C........ .T........	FJ481389
Ht13	.G........ ....C .T........ ......T... .C	FJ481390
Ht14	.G........ ....C .T........ ......T... ..	FJ481391
Ht15	.... C....T...T ..C....... .T......T. CC....T.A. ..	FJ481392
Ht16	.... C....T...T ..C....... .T....A.T. CC....T... ..	FJ481393
Ht17	C........C .... G......... .TA....... .C........ ..	FJ481394
Ht18	....C .T........	FJ481395
Ht19	.... CC...T...T ..CC...... .T......T. CC....T.A. ..	FJ481396
Ht20	....C .T........ .C....T..C ..	FJ481397
Ht21	....C..... ....C .T........ .C....T... ..	FJ481398
Ht22	.... C....A...T ..C....... .T......T. CC....T.A. ..	FJ481399
Ht23	....G..T.. .... .T........	FJ481400
Ht24	.... .T........ .C........ ..	FJ481401
Ht25	....T...T ..C....... .T......T. CC....T.A. ..	FJ481402

Ht26 ....G..... .....G... .TA..... .C..... ..	FJ481403
Ht27 ..... C....T...T ..C.....A. .TA.....T. CC....T.A. ..	FJ481404
Ht28 ....T...C .T..... .C....T... ..	FJ481405
Ht29 .... .T.....	FJ481406
Ht30 ....T. .....C .T..... .C....T... ..	FJ481407
Ht31 ....C .T..... .C....T... ..	FJ481408
Ht32 ..... C....T...T ..C..... .T.....T. CC....T.A. T.	FJ481409
Ht33 ..... C....T...T ..C..... .T...G..TG CC....T.A. ..	FJ481410
Ht34 .... G.....C .T..... .C....T... ..	FJ481411
Ht35 ....C..... .T..... .C..... ..	FJ481412
Ht36 .....G. ....C .T.....G.. .C...AT..C ..	FJ481413
Ht37 ...C..... C..G.T...T ..C..... .TA.....T. CC.T..T.A. ..	FJ481414
Ht38 ..... C....T...T ..C..... .TA..... CC....T.A. ..	FJ481415
Ht39 ..... C....T...T ..C..... .TA.....T. CC....T.A. ..	FJ481416
Ht40 ..... C....T...T ..C..... .T.....T. CC..T.T.A. ..	FJ481417
Ht41 ..C..... .... .T.....	FJ481418
Ht42 ..... .C..... ..... .T.....	FJ481419
Ht43 .... .T..... .....G.. ..	FJ481420
Ht44 ....G.. .T.....	FJ481421
Ht45 .... .T.....G.. .C..... ..	FJ481422
Ht46 ....C .T.A..... .C....T... ..	FJ481423
Ht47 ....C .T..T..... .C....T... ..	FJ481424
Ht48 .....G.. ....C .T..... .C....T... ..	FJ481425
Ht49 ..C..... ....C .T..... .C....T... ..	FJ481426

912 **Table S2** SJT Haplotype distribution among sampling sites
913

Haplotype	Site							Total	Haplotype
	NE	WE	TA	KM	TR	LC	MD		Frequency %
Ht1	1	0	0	0	0	0	0	1	0.31
Ht2	0	0	0	1	0	0	0	1	0.31
Ht3	0	0	0	0	0	1	0	1	0.31
Ht4	0	0	0	0	0	5	0	5	1.54
Ht5	0	0	2	0	0	0	16	18	5.56
Ht6	3	17	31	3	3	19	0	76	23.46
Ht7	0	15	0	0	0	0	0	15	4.63
Ht8	21	0	0	1	0	0	0	22	6.79
Ht9	0	0	0	0	0	0	1	1	0.31
Ht10	0	1	0	0	0	0	0	1	0.31
Ht11	0	0	0	0	22	0	0	22	6.79
Ht12	20	1	0	2	0	0	0	23	7.10
Ht13	0	0	1	0	0	0	0	1	0.31
Ht14	0	0	1	0	0	0	0	1	0.31
Ht15	0	0	0	2	0	0	0	2	0.62
Ht16	0	0	0	1	0	0	0	1	0.31
Ht17	0	0	0	3	0	0	0	3	0.93
Ht18	3	5	1	1	0	0	0	10	3.09
Ht19	0	5	0	0	0	0	0	5	1.54
Ht20	0	0	1	0	0	15	1	17	5.25
Ht21	0	0	0	0	1	0	0	1	0.31
Ht22	0	1	0	0	0	0	0	1	0.31
Ht23	0	0	1	0	0	0	0	1	0.31
Ht24	1	0	0	0	0	0	0	1	0.31
Ht25	0	1	0	0	0	0	0	1	0.31
Ht26	0	0	0	2	18	0	0	20	6.17
Ht27	1	0	0	0	0	0	0	1	0.31
Ht28	0	1	0	0	0	0	0	1	0.31
Ht29	0	0	0	1	0	0	0	1	0.31

Ht30	0	0	0	0	1	0	0	1	0.31
Ht31	0	0	0	0	0	1	0	1	0.31
Ht32	0	0	0	0	0	0	1	1	0.31
Ht33	0	0	0	0	0	0	1	1	0.31
Ht34	0	0	0	0	0	0	1	1	0.31
Ht35	0	1	0	0	0	0	0	1	0.31
Ht36	0	0	0	0	0	0	2	2	0.62
Ht37	1	3	0	20	2	1	0	27	8.33
Ht38	1	0	0	0	0	0	0	1	0.31
Ht39	0	0	0	1	0	0	0	1	0.31
Ht40	0	0	1	0	0	0	0	1	0.31
Ht41	1	0	0	1	0	0	2	4	1.23
Ht42	0	0	0	0	0	6	0	6	1.85
Ht43	0	0	0	0	0	0	2	2	0.62
Ht44	0	0	0	0	2	0	0	2	0.62
Ht45	0	1	0	0	0	0	0	1	0.31
Ht46	0	0	0	14	0	0	0	14	4.32
Ht47	0	0	1	0	0	0	0	1	0.31
Ht48	0	0	0	1	0	0	0	1	0.31
Ht49	0	0	1	0	0	0	0	1	0.31
Samples	**53**	**52**	**41**	**54**	**49**	**48**	**27**	**324**	

914

915 **Table S3** Characteristics of six SJT microsatellite loci

Locus	**Repeat motif**		**primer sequence (5'-3')**	**Expected product size**	**Anneal. Temp. ^{0}C**
UTD203	$(GAA)_7CT(GAA)_2$	F	CCC TGT GCT GTC TGT GAA G	157	50
		R	TTG AAT CAA TGG CAA CTG GA		
UTD73	$(AACT)_6$	F	TGT GTG ATG AAG CTA AAG	135	50
		R	CAA AAA TAT AGC CTT CGT		
UTD328	$(GCT)_8$	F	GAG AGA GAA GCG GAC AGG ATA GG	143	50

		R	TGA GTA ATA GAG AGT GGG AAT GG		
UTD535	$(AGAT)_9$	F	CAC TGA AGA TAT AGG CAG CCT TG	193	55
		R	TTT CTC CAG CGG CAT TAC AT		
UTD523	$(GATA)_{18}$	F	TTT GAA TGG GAG ACA TGC AG	247	55
		R	TGT CCT GCA CTT GTG TTC ACT		
UTD172	$(GACT)_5$	F	GTT GTG TAT TTT TGG CTG GAC C	145	55
		R	CAA CAG CTA ACG GGC AAA TTC C		