Hierarchical analyses of genetic variation of samples from breeding and feeding grounds confirm the genetic partitioning of northwest Atlantic and South Atlantic populations of swordfish (Xiphias gladius L.)

J.R. Alvarado Bremer a,b,*, J. Mejuto c, J. Gómez-Márquez d, F. Boán d, P. Carpintero d, J.M. Rodríguez d, J. Viñas e,1, T.W. Greig f, B. Ely g

a Department of Marine Biology, Texas A&M University at Galveston, Galveston, TX 77551, USA
b TAMU, Department of Wildlife and Fisheries Sciences, 210 Nagle Hall, TAMU 2258, College Station, TX 77843, USA
c Instituto Español de Oceanografía, Centro Oceanográfico A Coruña, Muelle de las Animas s/n, P.O. Box 130, 15080 A Coruña, Spain
d Departamento de Bioquímica y Biología Molecular, Facultad de Biología, Universidad de Santiago de Compostela, Santiago de Compostela, Spain
e Departamento de Biología, Laboratori d’Ictiologia Genètica, Universitat de Girona, Girona, Spain
f National Ocean Service, Center for Coastal Environmental Health and Biomolecular Research at Charleston (CCEHBR), 219 Fort Johnson Road, Charleston, SC 29412, USA
g FISHTEC Genetics Laboratory, Department of Biological Sciences, University of South Carolina, 715 Sumter St., Columbia, SC 29208, USA

Received 24 January 2005; received in revised form 3 June 2005; accepted 17 June 2005

Abstract

In species with high migratory potential, the genetic signal revealing population differentiation is often obscured by population admixture. To our knowledge, the explicit comparison of genetic samples from known spawning and feeding areas has not been conducted for any highly migratory pelagic fish species. This study examines the geographic heterogeneity of swordfish mitochondrial DNA (mtDNA) lineages within the Atlantic Ocean using 330 base pairs of sequence of the control region from 480 individuals. Hierarchical analyses of sequence variation were conducted to test whether samples from areas identified as the corresponding spawning and feeding grounds for the northwest (NW) Atlantic (Caribbean and Georges Banks–US northeast) and the South Atlantic (Brazil–Uruguay and Gulf of Guinea), were more closely related to each other than to samples from any other region, including the Mediterranean Sea, the Indian Ocean, and the Pacific Ocean. Phylogeographic analyses reveal that swordfish mtDNA phylogeny is characterized by incomplete lineage sorting and secondary contact of two highly divergent clades. However, despite this complex phylogenetic signature, results from an analysis of nucleotide diversity

* Corresponding author. Department of Marine Biology, Texas A&M University at Galveston, Galveston, TX 77551, USA. Tel.: +1 409 740 4958; fax: +1 409 740 5002.
E-mail address: Jaimeab@tamu.edu (J.R. Alvarado Bremer).
1 Current address: Institut de Ciències del Mar, C.S.I.C., Passeig Marítim, 37-49, 08003 Barcelona, Spain.
1. Introduction

The genetic signal of population differentiation observed in the reproductive areas of highly migratory species is often obscured by population admixture in wintering or feeding areas (Van Wagner and Baker, 1990; Bowen et al., 1992; Wenink et al., 1994). In the marine realm, the passive dispersal potential of early life history stages facilitated by currents, or the active dispersal associated to the specific migratory behavior of a particular ontogenetic stage, or sex, creates additional complications when attempting to interpret data from samples outside of the breeding (spawning) grounds under the assumption of strict natal homing. The implications of admixture towards the management of marine resources are far reaching and the decisions for quota allocation require mixed stock analysis (Kalinowski, 2004 and references therein). There is no precedent, to our knowledge, of a genetic study of a highly migratory pelagic fish that has explicitly attempted to include samples from a population’s putative spawning and feeding areas for their comparison against the corresponding representative samples in other populations. This approach is used here for the first time to examine the level of differentiation of two Atlantic populations of swordfish (Xiphias gladius L.).

Swordfish are epipelagic scombroid fish found in all oceans around the world, most frequently between latitudes 45°N and 45°S (Palko et al., 1981). In the North Atlantic, swordfish segregate by size and by sex, with the larger individuals, most commonly females, being able to venture into colder waters at higher latitudes (Palko et al., 1981). Mature females eventually return to breeding areas in warmer waters where males appear to be more abundant. Tagging experiments in the Atlantic Ocean using both conventional tags and pop-up satellite tags suggest that swordfish seldom conduct trans-Atlantic movements (Anonymous, 1997; Sedberry and Loefer, 2001). However, the results from conventional tag-recapture experiments should be considered tentative due to the relatively low number of individuals tagged and recaptured, the potentially low survivorship of released swordfish, and until recently, a tagging coverage mostly limited to the NW Atlantic with considerable seasonal differences in tagging effort. An increase in both seasonal and geographic tag-recapture coverage suggests that swordfish migrations are more extensive and variable than previously reported (Garcia-Cortés et al., 2003). In addition, the analysis of longline capture data in three oceans suggests that swordfish can be defined as segregated into three biological regions: feeding, reproductive or spawning, and transitional (Mejuto et al., 1991, 1995, 1998; Mejuto and García-Cortés, 2003a), with each region displaying specific patterns of sex-ratio-at-size, gonadal indices (GI), abundance by sex, and size distribution.

Attempts during the last decade to elucidate the population structure of swordfish in the Atlantic have involved genetic analyses using both mitochondrial DNA (mtDNA) (Alvarado Bremer et al., 1995, 1996; Kotoulas et al., 1995; Chow et al., 1997; Chow and Takeyama, 2000) and single copy nuclear DNA (scnDNA) data (Greig et al., 2000; Chow and Takeyama, 2000; Nohara et al., 2003). Significant frequency differences between pooled North Atlantic samples and South Atlantic samples (south of 5°N) have been shown using both nucleotide sequence data (Alvarado Bremer et al., 1996) and PCR-RFLP data (Alvarado Bremer et al., 1996; Chow et al., 1997) of the D-loop or mtDNA control region. Specifically, Alvarado Bremer et al. (1996) argued that the heterogeneity in haplotype distribution between the North and the South Atlantic is partially due to frequency differences of two highly divergent mtDNA clades (Alvarado Bremer et al., 1995). Clade II haplotypes (theta subgroup), are absent from the Indo-Pacific (Alvarado Bremer et al., 1996, 1998a,b; Rosel and Block, 1996; Reeb et al., 2000), occur at low fre-
quency in the South Atlantic, increase in abundance towards the North Atlantic, and reach the highest frequency in the Mediterranean (Alvarado Bremer et al., 1995, 1996, 2005). In addition, differences between Northwest Atlantic and South Atlantic also affect the distribution of subgroups of lineages belonging to Clade I (Alvarado Bremer et al., 1996, 1998a). Clade I fish can be assigned two groups of CR-I haplotypes; the alpha lineages that contain an Rsal restriction site that rarely occurs (< 0.05%) among non-alpha lineages, and the beta lineages that lack this restriction site. Beta-lineages, which comprise the majority (90–95%) of Indo-Pacific fish, are more common in the South Atlantic than in the NW Atlantic and the Mediterranean.

On the basis of the phylogeographic association of the two clades, Alvarado Bremer et al. (2005) speculated that Clade I originated in the Pacific whereas Clade II originated in the Atlantic. The co-occurrence of these highly divergent clades in the Atlantic is explained by unidirectional gene flow from the Indo-Pacific into the South Atlantic, in a phylogeographic pattern analogous to that reported for several other pelagic fishes (Finnerty and Block, 1992; Graves and McDowell, 1995; Alvarado Bremer et al., 1998a,b; Graves, 1998; Chow et al., 2000). In addition, the higher frequencies of beta-lineages in the S. Atlantic, which are more abundant in the Indo-Pacific, provide additional support to this interpretation (Alvarado Bremer et al., 1996).

The analysis of the nuclear loci aldolase \( B \) (aldB) and lactate dehydrogenase \( A \) (ldhA) has corroborated the reported mtDNA differences between the NW Atlantic and the South Atlantic swordfish (Greig et al., 1999, 2000). However, in their PCR-RFLP study of the nuclear locus calmodulin (CaM) gene intron 4 and of the mtDNA D-loop region, Chow and Takeyama (2000) described a sharp differentiation between NW Atlantic and South Atlantic swordfish with the CaM locus, but not with the mtDNA D-loop region data. Accordingly, the authors of that study argued that the mtDNA differences reported by both Alvarado Bremer et al. (1996) and Chow et al. (1997), resulted from sampling the two extremes of a mtDNA cline in the Atlantic. This explanation is not only inaccurate for reasons explained in this study, but also because these two studies, aside from demonstrating significant differences between NW Atlantic and South Atlantic samples, provided no evidence of a mtDNA cline in the Atlantic. However, it should be noted that previous mtDNA studies of Atlantic swordfish had two shortcomings. First, the size of South Atlantic samples was small, and second, samples were constituted from punctuated sampling efforts. For instance, Alvarado Bremer et al. (1995, 1996) included 8 and 23 individuals, respectively, collected along a wide band of equatorial South Atlantic waters extending from the African coast to the coast off Brazil. These small samples, and the wide geographic coverage, may not accurately represent the genetic composition of the South Atlantic.

This study attempts to resolve the previous limitations as follows. First, the sample sizes for all Atlantic localities have been substantially increased. Second, through observer programs, samples were obtained from the North and South Atlantic regions, and the comparison of samples was limited to feeding and spawning areas that have been properly characterized (see Methods for a detailed description). For that purpose, we surveyed the patterns of nucleotide sequence variation in 330 bp of the mitochondrial control region from a sample of 480 individuals. The amount of genetic differentiation among samples was measured by comparing arrangements of regional samples using two hierarchical approaches, namely, the analysis of molecular variance (AMOVA) (Excoffier et al., 1992) and the hierarchical analysis of nucleotide diversity (Holsinger and Mason-Gamer, 1996). The advantage of these two methods over conventional \( F_{st} \)-statistics is that both incorporate the phyletic relatedness of haplotypes when calculating fixation indices, making them particularly well suited for the detection of population differentiation when using hypervariable DNA sequence data and consequently a large number of distinct haplotypes. Using these two approaches, we tested the hypothesis that representative samples of the corresponding spawning and feeding grounds of the NW Atlantic and the South Atlantic should contain lineages that are more closely related to each other than to those from any other population. We demonstrate that significant differentiation exists among all regions included in this study, and specifically between the NW Atlantic and the South Atlantic. However, within these two Atlantic regions, samples from spawning and feeding areas were not different from each other, despite their
complex phylogenetic signal that includes polyphyly and incomplete lineage sorting. These results underscore the high resolution of the mtDNA control region sequences to resolve inter- and intra-oceanic population differentiation in swordfish.

2. Methods

2.1. Criteria of sample selection

Approximately 330 base pairs of control region sequence were determined for a total of 480 swordfish, which include 87 individuals characterized previously (Alvarado Bremer et al., 1995, 1996, 2005). Samples were collected in 5 regions: Indian Ocean (n=45); Western Mediterranean (n=135); NW Atlantic: Georges Banks–US northeast (n=73), Caribbean (n=24); Pacific Ocean: Hawaii (n=29); South Atlantic: Gulf of Guinea (n=80), and Brazil–Uruguay (n=94). Sampling details for all specimens are given in Table 1. The two Atlantic regions were selected for comparison because their respective putative spawning and feeding grounds have been characterized, and because of their considerable geographic separation (Fig. 1). Delineation of spawning and feeding areas was based on three types of evidence: (1) The presence of larval stages. (2) The presence of mature females. (3) Patterns of sex-ratios-at-size and overall sex ratios, associated with the relative abundance of each sex in the longline fishery.

Swordfish larvae have been collected in the Atlantic in a very wide band that includes tropical and subtropical waters (Fig. 1). In the NW Atlantic, this area includes the Gulf of Mexico, the southeast coast of the United States, and most of the Caribbean Sea

Table 1
Sampling details of swordfish assayed in this study

<table>
<thead>
<tr>
<th>Region/locality</th>
<th>Dates of capture</th>
<th>n</th>
<th>Latitude, Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>NW Atlantic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georges Banks and US Northeast (feeding)(^{a,b})</td>
<td>August–October 1990</td>
<td>73</td>
<td>39–40°N, 68–70°W</td>
</tr>
<tr>
<td>Caribbean (spawning)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North of Puerto Rico(^{c})</td>
<td>January 1993</td>
<td>17</td>
<td>19–21°N, 59–60°W</td>
</tr>
<tr>
<td>Yucatan Channel(^{d})</td>
<td>October 1992–June 1993</td>
<td>7</td>
<td>20–28°N, 78–85°W</td>
</tr>
<tr>
<td>South Atlantic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gulf of Guinea (feeding)(^{e})</td>
<td>December 1992</td>
<td>3</td>
<td>00–05°S, 15–20°W</td>
</tr>
<tr>
<td></td>
<td>April–May 1994</td>
<td>14</td>
<td>00–05°N, 10–25°W</td>
</tr>
<tr>
<td></td>
<td>June–August 1994</td>
<td>14</td>
<td>00–05°N, 05°E–05°W</td>
</tr>
<tr>
<td></td>
<td>May–August 1995</td>
<td>49</td>
<td>05°N–05°S, 05°E–10°W</td>
</tr>
<tr>
<td>Western Mediterranean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alboran Sea</td>
<td>June–October 1995</td>
<td>111</td>
<td>35–36°N, 00–05°E</td>
</tr>
<tr>
<td>Tarifa</td>
<td>July–August 1992</td>
<td>8</td>
<td>35°N, 5–10°E</td>
</tr>
<tr>
<td>Gulf of Valencia(^{f})</td>
<td>June–October 1994</td>
<td>16</td>
<td>39.30°N, 0.10°E</td>
</tr>
<tr>
<td>Indian Ocean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-NE of Madagascar</td>
<td>February–May 1994</td>
<td>34</td>
<td>0–10°S, 40–70°E</td>
</tr>
<tr>
<td>West of Madagascar</td>
<td>March–April 1994</td>
<td>11</td>
<td>10–20°S, 40–45°E</td>
</tr>
<tr>
<td>Pacific Ocean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hawaii(^{g})</td>
<td>November–December 1993</td>
<td>29</td>
<td>30–35°N, 155–160°W</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>480</td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) Thirteen individuals from Alvarado Bremer et al. (1995).
\(^{b}\) Twenty-seven individuals from Alvarado Bremer et al. (2005).
\(^{c}\) Eight individuals from Alvarado Bremer et al. (1996).
\(^{d}\) Two individuals from Alvarado Bremer et al. (1996).
\(^{e}\) Three individuals from Alvarado Bremer et al. (1995).
\(^{f}\) Included in Alvarado Bremer et al. (2005).
\(^{g}\) Eighteen individuals from Alvarado Bremer et al. (1996).
(Markle, 1974; Nishikawa et al., 1985). However, the larvae are not available for genetic studies and the presence of larvae alone should not be considered indicative of spawning areas, primarily because in most larval studies, the size (age) of larvae is not reported (i.e. larvae could be carried from spawning areas to other regions by currents, Markle, 1974; Tibbo and Lauzier, 1969). In addition, evidence of spawning females, according to histological stage and GI (based on the ratio of the gonadal weight by the fish weight) (Kume and Joseph, 1969) data, does not correlate directly with the documented distribution of larvae in the Atlantic. As such, spawning areas defined by the presence of spawning females are more restricted than those defined by the presence of larvae. For instance, between 1990 and 1995, the US and Venezuelan fleets sampled 14,662 fish between 5 and 55°N and west of approximately 35°W, and based on the presence of high GI females (GI value ≥ 3) reported spawning to occur not only in
tropical and subtropical areas south of the Sargasso Sea, and east of the Antillean Arc, but also including the Windward Passage, the Yucatan Channel and the Southeast coast of the US (Arocha and Lee, 1996; Taylor and Murphy, 1992, Fig. 1).

The analysis of sex ratios was used as the third piece of evidence to identify spawning and feeding localities. The proportion of males to females captured in the entire Atlantic does not differ significantly from unity (Anonymous, 1997). However, the observed proportion and relative abundance of males captured in spawning regions seems to be substantially higher than in non-reproductive areas especially when comparing the cumulative frequency of certain size ranges, whereas the abundance of females (derived from nominal CPUE by sex) appears to be the same in both areas (Mejuto et al., 1995, 1998). It should be noted that in these warm waters female swordfish may have lower catchability, or be swimming in deeper layers and thus may not be accessible to the surface longline-fishing gear (Fonteneau, 1995; Mejuto et al., 1994, 1995, 1998). In sum, the overlap of areas containing fish with high GI and those with a characteristic pattern of sex-ratio-at-size bias observed in the swordfish longline fishery, circumscribed by a region where larvae are known to occur, appears to be a robust indirect indicator of spawning areas.

Thus, following the aforementioned criteria, the NW Atlantic sample consists of the Caribbean (spawning) and Georges Banks–northeastern US (feeding). Contact between these two areas is supported by substantial evidence derived from tagging programs which indicated movement from the tropical reproductive region to the feeding grounds, following the Gulf Stream along the US coast towards Georges Banks and Grand Banks off the US and Canada (Burnett et al., 1987; Jones, 1997).

The selection of the South Atlantic reproductive area is based on the presence of larvae in the area corresponding to our Brazil–Uruguay sample (Nishikawa et al., 1985), and on the analysis by Mejuto and Garcia (1997) of 13,739 fish collected between 1986 and 1996 in the area defined by 40°N and 35°S latitude and 45°W and 5°E longitude. This study concluded that spawning takes place within an equatorial band west of 10°W between 5°N and 5°S. It should be clarified that because of oceanographic regimes and characteristics of swordfish catches, the distribution of the South Atlantic swordfish stock is considered to extend to 5°N (ICCAT, 2002). Spawning in the South Atlantic also occurs in the area roughly delimited by 15–35°S and 20–40°W (Mejuto and Garcia, 1997), which corresponds to our Brazil–Uruguay sample. Although spawning was detected in all quarters of the year, the seasonality of spawning could not be characterized due to limited sampling coverage and the temporal movement of the Spanish commercial fishing fleet. It is also important to note that the apparent discreetness of two spawning areas in the South Atlantic (see Fig. 1) may simply reflect the punctuated distribution of fishing effort of the Spanish fleet. Based on this information and on the studies mentioned above, Mejuto and Garcia (1997) drafted a general overview of swordfish spawning grounds in the Atlantic, associated to warm waters with a relatively deeper thermocline in western Atlantic regions, with some additional refinements in a later work (Mejuto and García-Cortés, 2003b). For the South Atlantic region, the sample from the Gulf of Guinea in this study corresponds to feeding grounds (Fig. 1).

### 2.2. PCR amplification and sequencing

Methods for DNA extraction and PCR amplification of the mtDNA control region I (CR-I) are described in detail in Alvarado Bremer et al. (1996). Nucleotide sequencing reactions for the Caribbean, Gulf of Guinea, NW Atlantic and Pacific samples are also described in Alvarado Bremer et al. (1996).
DNA extraction and sequencing reactions of PCR-amplified products for Mediterranean, Brazil–Uruguay and Indian Ocean samples were obtained in a similar manner as follows. DNA extractions were performed using a QIAamp Tissue Extraction Kit (Qiagen, Hilden, Germany). PCR products (5 μl) were treated with alkaline phosphatase (2 units) and exonuclease I (10 units) (Amer- sham Biosciences, Buckinghamshire, England, UK) to eliminate excess nucleotides and primers. The reaction mix was incubated at 37 °C for 15 min and then at 80 °C for 15 min to inactivate the enzymes. Sequencing reactions (Amersham Biosciences) were performed in both directions using primer L15998 (Alvarado Bremer et al., 1995) and a newly designed reverse internal primer, SWF (5′-TGTCCCTCACCTTCAATGAC-3′). Sequencing reactions were analyzed by electrophoresis in denaturing gels containing 6% polyacrylamide, and 7M Urea and detected by exposure of the gel to Hyperfilm (Amersham Biosciences) for 12 h.

2.3. Sequence data alignment and character weighing

Sequence alignments of the hypervariable mtDNA control region I were optimized by eye in Bioedit (Hall, 1999) following the criteria specified in Alvarado Bremer et al. (1995). In order to incorporate insertions and deletions (indels) into the distance estimates in NUCLEODIV (see below), a substitution matrix was appended to the sequence file in which each indel was represented by a nucleotide substitution equivalent to a transversion. Each of the two or three tandemly repeated copies of the tetramer motif ‘TACA’ near the 5′ end of the L-strand of swordfish CR-I was considered as a single mutation since nucleotide sequence analysis of CR-I suggests that each tetramer gain or loss was produced by a single event (Alvarado Bremer et al., 1995). Sequences of the 272 new CR-I haplotypes were assigned GenBank accession nos. AY961097–AY961098, AY961099–AY961136, AY961138–AY961146, AY961148–AY961282, AY961284–AY961370 and AY961372. The GenBank accession nos. for the 48 previously characterized haplotypes are: AY650743–AY650747, AY650750–AY650769, AY650773, AY650775, AY650778–AY650780, AY650783–AY650792, AY650795, AY650798–AY650801, AY650804, AY650807 and AY650836.

2.4. Phylogenetic analysis

Minimum evolution (ME) trees were estimated from the matrix of Tamura–Nei distances (Tamura and Nei, 1993) in MEGA (Kumar et al., 1993) using the close-neighbor-interchange (CNI) algorithm (Fig. 2). Missing nucleotide sites were treated with the pair-deletion option and trees were rooted at midpoint. To verify the robustness of sorting Clade I lineages into alpha and beta subgroups (Alvarado Bremer et al., 1996, 2005), a NEXUS tree file was imported into McClade ver. 3.04 (Maddison and Maddison, 1992) and the character state for nucleotide position 127 of the CR-I, responsible for the presence (G) or absence (A) of the RsaI restriction site, was traced along the branches of the tree. Only one (0.5%) swordfish among the 198 alpha individuals showed a reversal to adenine, whereas, only two (0.9%) swordfish among 230 beta fish had a guanine at that site, thus independently acquiring the RsaI restriction site. The phylogeographic information contained in the ME tree was used to determine the frequency of Clade I, alpha and beta lineages, and of Clade II, theta-Atl and theta-Med lineages for each sample.

2.5. Hierarchical analysis of nucleotide diversity

The degree of genetic differentiation was estimated using a hierarchical analysis of nucleotide diversity (Holsinger and Mason-Gamer, 1996) using the program NUCLEODIV. This approach provides an analog to Wright’s $F_{st}$ to study geographically structured populations using restriction site or DNA sequence data where the variation is not independently inherited. Their measure of $g_{st}$ is used to group populations based on the average time to coalescence for pairs of haplotypes. The results are depicted in UPGMA tree diagram that shows the relationship among populations after re-sampling the data 10,000 times (Fig. 3). Significant $P$-values imply that the mean time to coalescence for two haplotypes drawn from the same node (population) of a tree is less than that for two haplotypes drawn from different nodes (populations). To determine whether the heterogeneity in the distribution of two highly divergent clades within the Atlantic was the sole factor that explains genetic differentiation, lineages were sorted within locality by clade. In addition, we tested the geographic structure of CR-I hap-
lotypes using the hierarchical analysis of the molecular variance procedure (AMOVA) (Excoffier et al., 1992) based on the matrix of Tamura–Nei distances ($\alpha = 0.761$) as implemented in ARLEQUIN (Schneider et al., 2000). The haplotypic correlation measure ($\phi_{st}$) was estimated for the following hierarchical arrangement of regions/groups: NW Atlantic/feeding and spawning; South Atlantic/feeding and spawning; Mediterranean/pooled samples from the Alboran Sea, Gulf of Valencia and Tarifa; and Indo-Pacific/Indian Ocean and Pacific Ocean. AMOVAs were conducted using the entire DNA sequence data set (i.e., Clades I and II sequences). To verify that the proportion of among-group variance was not explained exclusively by the geographically heterogeneous distribution of swordfish belonging to Clade II, an AMOVA was also conducted with only Clade I sequences. The significance level of each haplotypic correlation examined was tested with 1000 replicates of a non-parametric permutation procedure. A similar AMOVA using only Clade II sequences was not conducted because only six swordfish belonging to Clade II were sampled in the

<table>
<thead>
<tr>
<th>Population</th>
<th>Clade I</th>
<th>Clade II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pacific Ocean</td>
<td>0.02954</td>
<td></td>
</tr>
<tr>
<td>Indian Ocean</td>
<td>0.03243</td>
<td></td>
</tr>
<tr>
<td>Mediterranean</td>
<td>0.01820</td>
<td></td>
</tr>
<tr>
<td>SA Spawning</td>
<td>0.02727</td>
<td></td>
</tr>
<tr>
<td>SA Feeding</td>
<td>0.03171</td>
<td></td>
</tr>
<tr>
<td>NWA Feeding</td>
<td>0.02451</td>
<td></td>
</tr>
<tr>
<td>NWA Spawning</td>
<td>0.02671</td>
<td></td>
</tr>
<tr>
<td>NWA Feeding II</td>
<td>0.01894</td>
<td></td>
</tr>
<tr>
<td>NWA Spawning II</td>
<td>0.02082</td>
<td></td>
</tr>
<tr>
<td>SA Pooled II</td>
<td>0.02071</td>
<td></td>
</tr>
<tr>
<td>Mediterranean II</td>
<td>0.00909</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. UPGMA Tree obtained with NUCLEODIV depicting the relationship between samples based on a hierarchical analysis of nucleotide diversity (Holsinger and Mason-Gamer, 1996). Distances and associated $P$-values are depicted at each node. Values of nucleotide diversity within samples are depicted at the tips of the branches. Abbreviations for Atlantic samples and their geographical provenance can be found in Fig. 1. CR-I haplotypes for each locality (Atlantic and Mediterranean) were sorted by clade (Clade I and Clade II) and the tree depicts the relationship among localities as defined by the coalescent signal of the lineages belonging to the respective clade. Clade II lineages were found only six times in the two South Atlantic samples. Accordingly, these six lineages were pooled (SA Pooled II).
entire South Atlantic: four in the Gulf of Guinea, and two in Brazil–Uruguay. Such an analysis, given the reduced proportion of among-group variance that characterizes swordfish mtDNA data, is not informative.

3. Results

3.1. Phylogeny and phylogeography of swordfish CR-I lineages

The relationship of swordfish mtDNA CR-I lineages surveyed in this study is depicted with a phylogenetic tree (Fig. 2A). Consistent with previous studies (Alvarado Bremer et al., 1995, 1996, 2005; Rosel and Block, 1996) haplotypes can be assigned to two major clades that are about 4.1% divergent. The majority (93%) of swordfish sampled in this study belong to Clade I, and were sampled in all ocean basins. The distribution of the Clade I alpha lineages is nonrandom. Nearly half of the Mediterranean swordfish alpha lineages emerge from only two branches (Fig. 2A). The first branch groups about 6% of the Mediterranean fish, including haplotype 85 (AY961113), repeated seven times (5%). The second branch groups about 43% of all the Mediterranean swordfish, including haplotype 26 (AY650768), which is present in about 27% of the Mediterranean swordfish surveyed and is the centroid of this branch (Alvarado Bremer et al., 2005). Similarly, all of the Mediterranean beta lineages emerge from one branch which groups about 28% of the individuals, including haplotype 21 (AY650763) also a centroid (Alvarado Bremer et al., 2005). Haplotype 21 accounts for 14% of all Mediterranean observations surveyed in this study. The distribution of Clade II is nonrandom as well. Clade II lineages are absent from the Indo-Pacific, are very rare in both SA feeding (2%) and SA spawning (5%), but increase in frequency in NW Atlantic (12–22%) and the Mediterranean (23%). In addition, both NWA spawning and NWA feeding contain only Clade II theta-Atl lineages, whereas in the Mediterranean, only theta-Med lineages are present. The distribution of Clade I lineages is also heterogeneous among the samples surveyed in this study. The frequency of Rsal alpha lineages is considerably higher (43–49%) in the Atlantic and Mediterranean than in the Indo-Pacific (10–12%). Conversely, the frequency of beta Clade I lineages is lower in NW Atlantic spawning (35%) and feeding (38%), than in the S. Atlantic (50–53%) and Indo-Pacific (88–90%).

3.2. Hierarchical analysis of nucleotide diversity

The phylogenetic interpretation and the phylogeographic association depicted in Fig. 2 are corroborated by the results from the hierarchical analysis of nucleotide diversity (Fig. 3). The tree diagram shows a clear separation of two major groups, corresponding to the Clade I and Clade II lineages. The sharp differentiation between these two major groups is expected given the pronounced phylogenetic break separating them ($D_A=4.1\%$) (Alvarado Bremer et al., 2005). Mediterranean Clade II lineages are significantly differentiated ($P<0.00001$) from Atlantic Clade II lineages, including those found in the NW Atlantic feeding and spawning areas, and in the pooled sample of South Atlantic Clade II lineages. This differentiation is consistent with the reciprocally monophyletic origin of Atlantic (theta-Atl) and Mediterranean (theta-Med) Clade II lineages (Alvarado Bremer et al., 2005) depicted in Fig. 2A and B. No significant differences were observed among the Atlantic samples...
containing Clade II lineages \( (P = 0.38300) \). This lack of significance is not surprising considering the small size \( (n = 6) \) of the pooled S. Atlantic sample of Clade II haplotypes. The estimates of population differentiation derived from the comparison Clade I lineages, which account for the majority (~90%) of the individuals surveyed in this study, generated a tree diagram that depicts a well-defined global population structure. There is strong support \( (P = 0.00200) \) for the differentiation of Indo-Pacific samples from the Atlantic and the Mediterranean samples. In addition, significant differences \( (P < 0.00001) \) separate the two NW Atlantic samples from the South Atlantic samples and the Mediterranean sample. In turn, the tree diagram is consistent with lineages from the NWA feeding and NWA spawning belonging to the same population. Similarly, SA spawning and SA feeding are not different from each other, suggesting they belong to the same population.

As noted above, the signal of genetic differentiation obtained with NUCLEODIV (Fig. 3) respectively from clades I and II is not congruent. Specifically, using Clade I sequences, the Mediterranean sample clusters with the two S. Atlantic samples, whereas Clade II data suggests that the Mediterranean is different from S. Atlantic (pooled) and the two NW Atlantic samples, which in turn are not different from each other. These incongruent results can be explained by examining the complex phylogeny and phylogeography of swordfish mtDNA lineages (Fig. 2A), specifically, the paraphyletic origin of the beta and alpha Mediterranean branches of Clade I, which contrast with reciprocally monophyletic origin of theta-Med and theta-Atl Clade II lineages (Fig. 2). Accordingly, the highly significant differentiation \( (P < 0.00001) \) that separates the Mediterranean sample from the two South Atlantic samples is consistent with the phylogenetic relationship of theta-Med and theta-Atl subgroups. The lack of differentiation between NW Atlantic and S. Atlantic Clade II theta lineages can be explained in part by the small number of lineages compared. In addition, the distribution of the six SA Clade II lineages (not shown) among the branches of the theta-Atl subtree (Fig. 2A) appeared random with no particular association to any branch.

The analysis of nucleotide diversity also summarizes the amount of variation contained within each sample. The values of nucleotide diversity \( (\pi) \) estimated from Clade I lineages ranged between 0.01820 and 0.03243, with the lowest value corresponding to the Mediterranean sample (Fig. 3), and all other samples with the value of \( \pi \) above 0.02451. Similarly, \( \pi \) values for Clade II lineages were lowest in the Mediterranean \( (\pi = 0.00909) \). These results are concordant with the lower diversity values of swordfish in the Mediterranean compared to any other ocean (Alvarado Bremer et al., 2005 and references therein), and the high frequency of repeated haplotypes reported here for that basin. No substantial differences in nucleotide diversity values were observed among the NW Atlantic and South Atlantic samples for either clade.

### 3.3. AMOVA

Consistent with previous studies, the majority (91.16%) of the variation in swordfish control region is found within populations (Table 2). However, roughly 8.67% \( (P < 0.00391) \) of the variance was explained by differences among populations. The amount of variance among populations within groups was very small (0.17%) and not significant \( (P = 0.26393) \), implying that no heterogeneity is contained in the arrangement of pairs of samples. An additional AMOVA limited to North and South Atlantic samples was conducted to further test the within group differentiation between spawning and feeding areas. As expected, there was a noticeable reduction in regional differentiation, but the small proportion of variance explained (2.40%) remained significant \( (P < 0.00001) \) (Table 2). More importantly, the differentiation between populations within regions (0.04%) continued to support the homogeneity between corresponding spawning and feeding grounds \( (P = 0.39589) \). A separate AMOVA including all the samples from all oceans, but using only Clade I sequences, was conducted to test whether the significant regional differentiation in the first AMOVA was explained primarily by the globally heterogeneous distribution of the two highly divergent clades. Similar to the analysis using the entire data set, the majority of the variance in this Clade I-only AMOVA (Table 2) was contained within populations (90.47%), but with a slight and unexpected increase in the proportion of variance explained among regions (8.89%; \( P = 0.01857 \)). Also, the proportion of variance
between populations within regions (0.64%) was not significant ($P=0.13196$), reflecting the homogeneity observed between the feeding and spawning areas, within the NW Atlantic and the S. Atlantic, respectively. This AMOVA suggests that the heterogeneous distribution of alpha and beta Clade I lineages is as important in explaining the genetic differentiation of swordfish populations as is the heterogeneous distribution of the two highly divergent mtDNA clades, and is consistent with the Clade I portion of the results of the hierarchical analysis of nucleotide diversity (Fig. 3). The reciprocal analysis employing only Clade II lineages was not attempted because of the small number of Clade II mtDNAs in our combined S. Atlantic samples (two and four, respectively).

Alternative AMOVAs using sample-pair combinations of the four Atlantic localities, respectively, with each of the samples from the Mediterranean, Indian or Pacific Ocean using Clade I fish (not shown) produced similar values of among-group variance ($\approx 8\%$), but generated highly significant ($P<0.00001$) proportions of between populations within-group variance ($\sigma_b^2$) components, indicative of the heterogeneity of each of these alternative arrangements. Clade II fish were excluded from these AMOVAs to ensure that the within-group sample heterogeneity was not explained by the absence of Clade II lineages from the Indo-Pacific samples, and the lower frequency of these lineages in the S. Atlantic compared to the NW Atlantic.

**4. Discussion**

The global phylogeographic association of mtDNA lineages is consistent with the existence of at least four independent swordfish populations (Alvarado Bremer et al., 1996, 1998a; Chow et al., 1997; Nohara et al., 2003), namely the Mediterranean, NW Atlantic, South Atlantic and Indo-Pacific. Using both, the hierarchical analysis of nucleotide diversity (Holsinger and Mason-Gamer, 1996) and AMOVA (Excoffier et al., 1992), we confirmed the hypothesis that the corresponding feeding and spawning grounds within the NW Atlantic and within the South Atlantic are not different from each other but are different from any other sample and region surveyed. It should be noted that this interpretation is confirmed by hierarchical analysis of nucleotide diversity estimated with NUCLEODIV, only if we ignore Clade II data (Fig. 3). The analysis of Clade II lineages fails to detect heterogeneity between the NW Atlantic samples and the pooled S. Atlantic sample. However, only six Clade II swordfish were sampled in the S. Atlantic, a number of individuals hardly sufficient to reject the test of homogeneity, especially if we assume that the

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>% variation</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Entire data set: all regions, all lineages</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among regions</td>
<td>3</td>
<td>235.109</td>
<td>0.461200 $\sigma_a^2$</td>
<td>8.67</td>
<td>0.00391</td>
</tr>
<tr>
<td>Between populations within regions</td>
<td>3</td>
<td>21.252</td>
<td>0.01224 $\sigma_b^2$</td>
<td>0.17</td>
<td>0.26393</td>
</tr>
<tr>
<td>Within populations</td>
<td>473</td>
<td>3044.374</td>
<td>5.33675 $\sigma_c^2$</td>
<td>91.16</td>
<td>0.00000</td>
</tr>
<tr>
<td>Total</td>
<td>479</td>
<td>3300.735</td>
<td>7.06055</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NW Atlantic and S. Atlantic, all lineages</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between regions</td>
<td>1</td>
<td>26.413</td>
<td>0.15869 $\sigma_a^2$</td>
<td>2.40</td>
<td>0.00000</td>
</tr>
<tr>
<td>Between populations with regions</td>
<td>2</td>
<td>13.248</td>
<td>0.00274 $\sigma_b^2$</td>
<td>0.04</td>
<td>0.39589</td>
</tr>
<tr>
<td>Within populations</td>
<td>267</td>
<td>1723.509</td>
<td>6.45509 $\sigma_c^2$</td>
<td>97.56</td>
<td>0.00098</td>
</tr>
<tr>
<td>Total</td>
<td>270</td>
<td>1763.170</td>
<td>6.61652</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All regions, Clade I-lineages only</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among regions</td>
<td>3</td>
<td>145.082</td>
<td>0.39773 $\sigma_a^2$</td>
<td>8.89</td>
<td>0.01857</td>
</tr>
<tr>
<td>Between populations within regions</td>
<td>3</td>
<td>16.440</td>
<td>0.02855 $\sigma_b^2$</td>
<td>0.64</td>
<td>0.13196</td>
</tr>
<tr>
<td>Within populations</td>
<td>435</td>
<td>1761.065</td>
<td>4.04843 $\sigma_c^2$</td>
<td>90.47</td>
<td>0.00000</td>
</tr>
<tr>
<td>Total</td>
<td>441</td>
<td>1922.588</td>
<td>4.47470</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

See Methods for an explanation of the hierarchical regional arrangement tested.

* $P$-value corresponds to the probability of obtaining random values larger or equal than the observed values.
The proportion of among-group variance is as small as that explained between NW Atlantic and S. Atlantic Clade I lineages (Table 2). Thus, limiting the interpretation to Clade I lineages, which account for approximately 93% of the Atlantic swordfish surveyed, it becomes apparent that the NW Atlantic and the S. Atlantic are significantly different from each other, and in turn, both are different from the Indo-Pacific and the Mediterranean. These results were confirmed by the AMOVA, where the hierarchical arrangement that combined the two NW Atlantic samples and the two South Atlantic samples, respectively, explained the highest proportion of significant among-group variation. Any other combination of sample pairs within the Atlantic generated highly significant probabilities associated with the among-populations within-region component, indicating their suboptimal relationship. Combining any of the four Atlantic samples with samples from any of the other regions surveyed in this study always resulted in a reduction of among-region variance (data not shown). The results of both, the Clade I-only AMOVA and the analysis of nucleotide diversity confirm that the coalescent signal contained within Clade I is sufficient to resolve the differences between the NW Atlantic and South Atlantic swordfish populations (Alvarado Bremer et al., 1996). In fact, a slight increase in the proportion of among-group variation was obtained with the Clade I-only analysis compared to that with the entire data set.

The PCR-RFLP analysis of swordfish mtDNA CR-I developed by Alvarado Bremer et al. (1996) took advantage of restriction polymorphisms that were phylogenetically informative (cf. Quinn, 1992; Alvarado Bremer et al., 1998b) to readily sort sets of closely related composite haplotypes, and thus avoid plesiomorphic restriction site gains or losses within this hypervariable segment (Alvarado Bremer et al., 1998b). For instance, alpha lineages which appear to be a monophyletic group, are found at a higher frequency in the Atlantic (43–49%) than in the Indo-Pacific (5–12%) (Alvarado Bremer et al., 1996, 1998a, 2005; this study). Conversely, the majority (~90–95%) of Pacific swordfish correspond to the ubiquitous polyphyletic beta basal lineages, which account for the largest proportion (50–56%) of South Atlantic CR-I lineages (Alvarado Bremer et al., 1996, 2005; this study). Thus, geographic differences in the distribution of Clade I alpha and beta lineages are capable of explaining the concordance between AMOVA and the hierarchical analysis of nucleotide variance regarding the differentiation of the NW Atlantic and South Atlantic populations, even when the highly divergent Clade II haplotypes occurring at about 12–22% in the NW Atlantic and at 2–5% in the South Atlantic (Alvarado Bremer et al., 1996; this study), are excluded from the analysis.

Altogether, the results presented here indicate that the NW Atlantic and the South Atlantic samples are genetically distinct from each other, in agreement with the mtDNA studies of Alvarado Bremer et al. (1996), Chow et al. (1997) and Greig et al. (1999), but not with the frequency analysis of D-loop PCR-RFLP data of Chow and Takeyama (2000), Chow and Takeyama (2000) used PCR-RFLPs to characterize swordfish variation at both the nuclear CaM locus and at the mtDNA D-loop region, expanding the Atlantic sampling coverage surveyed by Chow et al. (1997) to include two South Atlantic samples equivalent to the Gulf of Guinea and Brazil–Uruguay in the present study, and a NW Atlantic sample equivalent to the feeding area of Georges Banks. A sharp genetic differentiation between NW Atlantic and the South Atlantic samples was detected with the CaM locus (a difference later corroborated by Nohara et al., 2003), but not with the D-loop. Chow and Takeyama (2000) concluded that both Alvarado Bremer et al. (1996) and Chow et al. (1997) might have found mtDNA differentiation because the samples compared in these two studies corresponded to the extremes of a cline of mtDNA genotype frequencies. However, this explanation is inaccurate for two reasons. First, the South Atlantic sample employed by Alvarado Bremer et al. (1996) is equivalent to the geographically intermediate tropical Atlantic samples used by Chow and Takeyama (2000), and thus cannot be considered as representative of the extreme of the range. Second, Chow et al. (1997) characterized the D-loop fragment with four restriction enzymes, AluI, HhaI, DdeI, and RsaI whereas Chow and Takeyama (2000) limited their analysis to two: AluI and RsaI. Inspection of the composite haplotype frequencies reported by Chow et al. (1997) reveal that the combined frequency of two composite haplotypes (ACBA and CCBA) accounts for 24% of the haplotypes in their South Atlantic (Brazil), but only 7% in their pooled NW Atlantic sample. By limiting their survey to two
enzymes, Chow and Takeyama (2000) removed the D-loop polymorphisms that had driven the frequency differences between the North and South Atlantic in Chow et al. (1997). For these reasons, we conclude that there is no discordance between mitochondrial DNA data (Alvarado Bremer et al., 1996; Chow et al., 1997) and nuclear DNA data (Greig et al., 1999, 2000; Chow and Takeyama, 2000; Nohara et al., 2003) regarding the differentiation of the NW Atlantic and the South Atlantic swordfish populations.

We have long recognized that the analysis of mtDNA data offers only a partial view of the behavior of swordfish populations since mitochondria are inherited maternally (Alvarado Bremer et al., 1998a,b; Greig et al., 1999). Nevertheless, the concordance of nuclear and mitochondrial DNA data with regard the differentiation of NW Atlantic, South Atlantic, Mediterranean and Indo-Pacific, suggest that there is no bias linked to the mode of inheritance, and that ongoing gene flow among swordfish populations must be extremely limited.

5. Conclusions

The subdivision of Atlantic swordfish populations into NW Atlantic and South Atlantic is supported by concordant results at both mtDNA and scnDNA data. Future studies are needed to clarify the relationship of NE Atlantic swordfish with the NW Atlantic and South Atlantic swordfish populations. The admixture of NE Atlantic and Mediterranean swordfish may confound such analysis (Kotoulas et al., 1995; Chow and Takeyama, 2000). However, preliminary surveys suggest that this contact zone is confined to a small region west of Gibraltar and south to a narrow region off the NW African Atlantic coast of Morocco and the Canary Islands (Alvarado Bremer et al., 1999). Similarly, the extent of admixture of Indian Ocean and South Atlantic swordfish around the southern tip of Africa needs to be resolved. These studies should integrate both nuclear DNA and mtDNA data.

Acknowledgments

We thank G. DeMetrio, M. Quintans, E. Alot I. González, E. Majuelos, R. Dollar, T. Geisman, L. Dixon, J.M. de la Serna and NOAA Fisheries for providing samples with special thanks to D. Lee and Cheryl Woodley. Also, we thank K. Holsinger for providing the program NUCLEODIV. E. Thelen and A. Ibarreche-Alvarado provided expert technical assistance. Thanks to Jerry Scott of NOAA fisheries for comments on an earlier version of this manuscript. This manuscript was dramatically improved thanks to the comments of two anonymous reviewers. This study was funded in part by the Texas Institute of Oceanography, and by the Cooperative Institute for Fisheries Molecular Biology (FISHTEC) NOAA-NMFS (RT/F-1) and is listed as FISHTEC contribution number FT 02-05 (number will be provided after final acceptance). Additional support was provided by IEO-4.03 and Economic Union project DGXIV, MED93-013.

References


Kumar, S., Tamura, K., Nei, M., 1993. MEGA: Molecular Evolutionary Genetics Analysis. Pennsylvania State University, University Park.


