FastTrack publication

Bull Mar Sci. 94(0):000–000. 2018
https://doi.org/10.5343/bms.2017.1130

research paper

Use of rapidly evolving molecular markers to distinguish species and clarify range uncertainties in the spearfishes (Istiophoridae, *Tetrapturus*)

Virginia Institute of Marine Science, College of William & Mary, 1375 Greate Road, Gloucester Point, Virginia 23062.

** Corresponding author email: <mcdowell@vims.edu>.*

Jan R McDowell *
Nadya R Mamoozadeh
Heidi L Brightman
John E Graves

ABSTRACT.—Despite broad spatial distributions in the Atlantic, Pacific, and Indian oceans, relatively little is known about spearfishes (family Istiophoridae, genus *Tetrapturus*) due to their pelagic nature and relative scarcity. The limited biological understanding of spearfishes includes uncertain taxonomic relationships complicated by conserved morphology, thus specific identification relies heavily on geographic location of capture. Previous phylogenetic studies incorporating a limited number of loci and few representatives of each species have been unable to consistently resolve the four currently recognized species comprising *Tetrapturus*. In the present study, we surveyed 14 nuclear microsatellite loci and the mitochondrial DNA control region across relatively large numbers of samples per species to genetically discriminate spearfish species. Molecular data resolved roundscale spearfish (n = 89) (*Tetrapturus georgii* Lowe, 1841) and Mediterranean spearfish (n = 12) (*Tetrapturus belone* Rafinesque, 1810) as genetically distinct groups. Longbill spearfish (n = 79) (*Tetrapturus pfluegeri* Robins and de Sylva, 1963) and shortbill spearfish (n = 29) (*Tetrapturus angustirostris* Tanaka, 1915) were not consistently resolved. A single individual collected in the western central Atlantic Ocean off Brazil assigned to shortbill spearfish and likely represents a vagrant. Additionally, a spearfish sampled from the Indian Ocean off eastern South Africa was morphologically identified as a longbill spearfish, and the molecular profile of this specimen was consistent with the morphological identification. Our study represents the most extensive molecular evaluation of the spearfishes to date. Results reported here underscore the substantial level of genetic similarity between longbill and shortbill spearfishes, and suggest high-resolution genomic methods may be required to unambiguously resolve these species.

Date Submitted: 17 August, 2017.
Date Accepted:13 December, 2017.
Available Online: 19 February, 2018.

Effective fisheries management and conservation require the ability to accurately delineate species, as well as a thorough understanding of species' geographic ranges. Spearfishes are members of the family Istiophoridae (marlins, sailfish, and spearfishes) and comprise the genus *Tetrapturus.* These highly-migratory species display broad, continuous spatial distributions in temperate, sub-tropical, and tropical waters of the Atlantic, Pacific, and Indian oceans (Nakamura 1985). Although the spearfishes are cosmopolitan, they are the scarcest of the world's istiophorids and are primarily restricted to pelagic waters, resulting in a paucity of basic biological information (Nakamura 1985, Collette et al. 2006), including limited knowledge concerning the extent of the geographic range of each species. In addition, the species status of some members of *Tetrapturus* has been questioned. A molecular investigation based on markers better able to resolve species separated by a low level of genetic divergence, and inclusion of multiple exemplars per species, would advance the understanding of species comprising this genus.

Four species of *Tetrapturus* are currently recognized: the roundscale spearfish, *Tetrapturus georgii* Lowe, 1841, longbill spearfish, *Tetrapturus pfluegeri* Robins and de Sylva, 1963, Mediterranean spearfish, *Tetrapturus belone* Rafinesque, 1810, and shortbill spearfish, *Tetrapturus angustirostris* Tanaka, 1915. These species can be morphologically distinguished from other members of the Istiophoridae by a vent positioned far anterior to the origin of the first anal fin (Nakamura 1985). Within *Tetrapturus*, characters such as relative bill length, and dorsal and pectoral fin morphology can be used to discriminate species (Nakamura 1985); however, using these characters to identify species is challenging and, in practice, distinguishing characters are often removed when fish are processed at sea. As a result, spearfishes are generally identified based on the geographic location of capture. Specific identification is further complicated in some areas such as the eastern North Atlantic Ocean, where three species—*T. georgii, T. pfluegeri,* and *T. belone*—occur sympatrically (Robins 1974, Nakamura 1985).

The roundscale spearfish is distributed across the Atlantic Ocean, and ranges from at least 37°4′N to 28°52′S in the western Atlantic Ocean (Bernard et al. 2013). The species range in the eastern Atlantic Ocean is not well documented, but extends from Portugal south to at least Madeira, and includes the Mediterranean Sea (Robins 1974). This species can be discriminated from other spearfishes by the height and shape of the dorsal and anal fins, as well as a larger and deeper maximum body size (Nakamura 1985). Although the roundscale spearfish was originally described from a single specimen caught off the island of Madeira (Lowe 1840), the validity of this species was only recently confirmed after over a century of confusion (Shivji et al. 2006). The uncertain species status of the roundscale spearfish largely resulted from morphological similarities with the more commonly encountered white marlin, *Kajikia albida* (Poey, 1860); however, these species can be most readily distinguished by the more anterior vent placement characteristic of the genus *Tetrapturus*.

The longbill spearfish is distributed across the Atlantic Ocean from approximately 40°N to 35°S (Robins 1974, Nakamura 1985). Compared to shortbill and Mediterranean spearfishes, longbill spearfish is primarily distinguished by a bill of equal or greater length than that of the head, and by pectoral fins that are long and wide (Nakamura 1985). Though a rare-event species Atlantic-wide, longbill spearfish is more densely distributed in the western Atlantic Ocean (Nakamura 1985). The longbill spearfish was recognized as a valid species in 1963, after it was confirmed

that previously examined specimens from the western North Atlantic Ocean were distinct from spearfish specimens described from the Mediterranean Sea (Robins and de Sylva 1960, 1963).

Though limited biological information is available for the roundscale and longbill spearfishes, even less is known about the Mediterranean and shortbill spearfishes. The Mediterranean spearfish is regarded as spatially restricted to the Mediterranean Sea and is reportedly most abundant off the coast of Italy (Nakamura 1985); however, a few individuals have been reported from the Atlantic Ocean west of the Straits of Gibraltar (Di Natale et al. 2005). The Mediterranean spearfish is morphologically characterized by a short bill <18% of body length, and by short and narrow pectoral fins (Nakamura 1985). The shortbill spearfish is the only member of *Tetrapturus* reported to occur in the Pacific and Indian oceans. It is distributed throughout the Pacific Ocean from approximately 40°N–35°S, and throughout the Indian Ocean from approximately 20°N to 35°S–45°S (Nakamura 1985). Though morphologically similar to the Mediterranean spearfish, the shortbill spearfish is characterized by a bill <15% of total body length and pectoral fins that are short and narrow (Nakamura 1985).

Understanding the taxonomic relationships and documenting the geographic ranges of the spearfishes is complicated by species scarcity as well as morphological and molecular similarities. Previous molecular studies based on a limited number of loci have either been unable to resolve all members of *Tetrapturus*, failed to account for the range of diversity within species, or have not included all members of the genus (Collette et al. 2006, Shivji et al. 2006, Hanner et al. 2011, Bernard et al. 2013, 2014, Santini and Sorenson 2013). The challenge of sampling large, pelagic, rare-event species has also resulted in a lack of morphological voucher specimens associated with samples analyzed in both morphological and molecular studies; very few specimens are available in museum collections. In addition, the inability of previous genetic evaluations to consistently resolve the species currently comprising *Tetrapturus* suggests highly variable molecular markers may be necessary for species discrimination. Collectively, these factors have contributed to an inadequate understanding of the geographic distributions and genetic distinctiveness of currently recognized species within the genus *Tetrapturus*.

In the present study, we surveyed genetic variation at 14 nuclear microsatellite loci and at the mtDNA control region across sample collections representative of each currently recognized *Tetrapturus* species. The number of individuals evaluated per species in the present study is significantly greater than in previous studies that included all known species of spearfishes, providing the first comparison of levels of intra-and inter-specific diversity in *Tetrapturus* and enabling evaluation of whether highly polymorphic loci are capable of discriminating spearfish species. In addition, our inclusion of multiple sampling locations for some species facilitates an assessment of the validity of the current geographic ranges of spearfishes.

Methods and Materials

Sample Collection and DNA Isolation.—Samples analyzed in the present study were opportunistically collected during the period 2000–2016 and consisted of muscle tissue from landed spearfishes or fin clips from spearfishes that were released alive following capture. For roundscale spearfish, samples included collections from

Table 1. Sample collections evaluated in this study. The number of individuals evaluated at the 14 microsatellite loci and/or mtDNA control region surveyed in the present study is shown.

Sample collection	Microsatellites	mtDNA
Roundscale spearfish (*Tetrapturus georgii*)		
Maryland, USA	3	0
New Jersey, USA	30	10
Florida, USA	10	4
Brazil	46	2
Longbill spearfish (*Tetrapturus pfluegeri*)		
Brazil	76	18
Shortbill spearfish (*Tetrapturus angustirostris*)		
Hawaii, USA	22	28
Australia	7	9
Mediterranean spearfish (*Tetrapturus belone*)		
Mediterranean Sea	12	10
Other		
Richards Bay, South Africa (UNK_mLBS)	1	1
Brazil (SBS_Brazil)	1	1
Total	208	83

the western North Atlantic Ocean off New Jersey, Maryland, and Florida, USA, as well as a collection from the western Central Atlantic off Brazil (Table 1). For longbill spearfish, samples consisted of a collection from the western central Atlantic Ocean off Brazil. Collections of shortbill spearfish included individuals sampled off Hawaii, USA, and off the eastern coast of Australia. For Mediterranean spearfish, a single collection obtained from locations throughout the Mediterranean Sea was analyzed. In addition to the previously described sample collections, a spearfish of uncertain species identification sampled from the Indian Ocean off Richards Bay, South Africa, was included in cluster-based analyses aimed at species identification. The morphology of this individual was consistent with that of a longbill spearfish and this individual is hereafter referred to as unknown morphologically identified longbill spearfish (UNK_mLBS). All tissue samples were stored at room temperature in 95% ethanol or a 10% dimethyl sulfoxide solution (Seutin et al. 1991) until DNA isolation. Total genomic DNA was isolated from each sample using either Chelex 100 resin (Sigma Aldrich; Walsh et al. 1991) or a DNeasy Blood & Tissue Kit (Qiagen).

Microsatellite Data Generation and Analyses.—In total, 14 microsatellite loci were surveyed across all samples in the present study (n = 208; Table 1). This included nine previously published loci (Bernard et al. 2012) and five newly developed loci (Online Table S1). Isolation and characterization of microsatellite loci developed for our study were performed using samples of roundscale spearfish following methodology described in McDowell et al. (2002) and Sorenson et al. (2011). All loci were PCR amplified using a *Taq* PCR Core Kit (Qiagen) following the manufacturer's protocol and forward primers modified at the 5′ end with locus-specific tag sequences and fluorescent dye labels. PCR amplification products were sized on an ABI 3130XL Genetic Analyzer (Applied Biosystems, Inc.) and visualized using GeneMarker v2.6.0 software (SoftGenetics, LLC). Allele size calls were manually inspected for accuracy, and a subset of samples (15%) was genotyped twice across all loci to assess genotyping error.

Microsatellite genotype data were evaluated for the presence of scoring errors and null alleles using Micro-Checker v2.2 (Van Oosterhout et al. 2004). Genotype data were also evaluated for conformance to the expectations of Hardy-Weinberg equilibrium (HWE) and linkage equilibrium using Genepop v4.5 (Raymond and Rousset 1995). Expected heterozygosity was calculated in Genepop using Levene's (1949) correction for small sample sizes. Significance values were corrected for multiple pairwise comparisons using a modified false discovery rate (Benjamini and Yekutieli 2001, Narum 2006). All Genepop analyses were performed using 100 batches and 10,000 iterations of each batch. The number of alleles observed per locus, as well as rarefaction allelic richness and rarefaction private allelic richness, were determined using HP-RARE (version from February 2, 2009; Kalinowski 2004, 2005); the minimum observed sample size of 22 genes was used for rarefaction calculations.

To test whether species formed distinct genetic groups, clustering of the genotype data was evaluated using the Bayesian modeling-based algorithm implemented in Structure v2.3.4 (Pritchard et al. 2000, Falush et al. 2007, Hubisz et al. 2009). Structure simulations were performed using an admixture model of ancestry, no location prior, and a burn-in of 100,000 followed by 1,000,000 Markov chain Monte Carlo iterations. Scenarios with K equal to one through five were evaluated with 25 replicates of each K. Replicate simulations of each K scenario evaluated in Structure were summarized using Clumpp v1.1.2 (Jakobsson and Rosenberg 2007) and resulting barplots were visualized using DISTRUCT v1.1 (Rosenberg 2004). Structure analyses were performed with two data sets: one data set inclusive of sample collections for all four currently recognized *Tetrapturus* species, and a second data set with collections of roundscale spearfish excluded to improve resolution of genetic clustering among the more shallowly differentiated species of *Tetrapturus* (Waples and Gaggiotti 2006).

Multivariate analysis was used to explore diversity within and among sample collections reflective of the four currently recognized species of *Tetrapturus*. Principal component analysis (PCA) was performed using the R package ADEGENET v1.4-2 (Jombart 2008, Jombart et al. 2009). The scaleGen function of ADEGENET was used to center allele frequencies and replace missing data with mean allele frequencies for the PCA. The analysis was repeated using a data set with roundscale spearfish excluded due to the high level of genetic divergence observed between this species and all other species of *Tetrapturus*.

Genetic divergence based on the microsatellite data was assessed among all pairs of individuals regardless of species, and between sample collections of the four *Tetrapturus* species. The level of genetic divergence between all possible pairwise comparisons of individuals was determined by calculating Goldstein's distance $(\delta\mu)^2$ (Goldstein et al. 1995) and Cavalli-Sforza and Edwards' chord distance (D_C; Cavalli-Sforza and Edwards 1967) in the program Microsatellite Analyzer v4.05 (Dieringer and Schlötterer 2003). Resulting distance matrices were used to construct unrooted neighbor joining trees using the R package APE v4.1 (Paradis et al. 2004) and relationships were visualized in FigTree v1.4.3 (Rambaut 2009). Genetic divergence was also evaluated among spearfish species by calculating F_{ST} values in Arlequin v3.5 (Excoffier and Lischer 2010), and D_C and $(\delta\mu)^2$ values in Microsatellite Analyzer. While D_C and F_{ST} are based on the infinite alleles mutational model (Kimura and Crow 1964), $(\delta\mu)^2$ is based on the stepwise mutational model (Kimura and Ohta 1978) and may better represent the mutational dynamics of microsatellite markers

(Goldstein et al. 1995, Takezaki and Nei 1996). Genetic divergence values were also calculated between collections of shortbill spearfish sampled off eastern Australia and off Hawaii to evaluate the presence of intraspecific population structure; these analyses were also performed for collections of roundscale spearfish sampled off Florida, Maryland, and New Jersey, USA, and off Brazil.

Deviations from mutation-drift equilibrium indicative of recent demographic expansion or contraction for each currently recognized *Tetrapturus* species were assessed using BOTTLENECK v1.2.02 (Cornuet and Luikart 1996, Piry et al. 1999). BOTTLENECK analyses were completed using default settings, including generation of allele frequency distributions and one- and two-tailed Wilcoxon signed rank tests for heterozygosity deficiency or excess, except the number of iterations was increased to 10,000. Because BOTTLENECK analyses were performed using data generated from microsatellites, only results based on the stepwise and two-phase mutational models were considered.

Genetic assignment tests were performed with the program GeneClass2 (Piry et al. 2004) to determine the most likely species identification for every individual analyzed in the present study, including the unknown morphologically identified longbill spearfish sampled off eastern South Africa (UNK_mLBS). The probability of each individual belonging to one of four reference populations comprised of sample collections for each spearfish species was computed. GeneClass2 was also used to assess the probability that sampled individuals were first generation migrants by computing Bayesian likelihood ratios. Each individual's estimated likelihood of belonging to the species from which it was sampled was compared to the highest likelihood value over all sampled species ($L' = L_{\_home}/L_{\_max}$) based on multilocus microsatellite genotypes. All GeneClass2 analyses were performed using the Bayesian computational criteria of Rannala and Mountain (1997). Statistical significance was assessed by simulating 10,000 multilocus genotypes following Paetkau et al. (2004).

mtDNA Data Generation and Analyses.—An approximately 800-bp fragment of the mtDNA control region was sequenced for a subset of individuals inclusive of each sample collection for each spearfish species (n = 83; Table 1). Due to variable DNA quality among samples, individuals sequenced at the mtDNA control region were not always the same individuals as those genotyped at the microsatellite loci. PCR amplifications were performed using a *Taq* PCR Core Kit (Qiagen), Pro-5 forward primer (Palumbi 1996), and an internal reverse primer (Graves and McDowell 2006). Briefly, each 25 µl reaction contained 10× PCR Buffer, 200 µM of each dNTP, 0.125 µM of each primer, 0.5 units of *Taq* polymerase, and 1 µl genomic DNA. Amplifications were performed with an initial denaturation of 94 °C for 2 min followed by 35 cycles of 94 °C for 1 min, 54 °C for 1 min, and 72 °C for 2 min, and a final elongation step of 72 °C for 5 min. Amplification products were purified using a QIAquick PCR Purification Kit (Qiagen) following the manufacturer's protocol, and quantified using a NanoDrop 2000 (Thermo Scientific). Purified PCR products were sequenced in both directions using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Inc.) with the original PCR amplification primers and 1/8 the recommended concentration of Big Dye. Sequence reactions were cleaned using ethanol-sodium acetate and electrophoresed on an ABI 3130XL Genetic Analyzer (Applied Biosystems, Inc). Base calls were analyzed using Sequencing Analysis v5.2 software (Applied Biosystems, Inc.). Forward and reverse sequences were used to

construct consensus sequences in Sequencher v5.1 (Gene Codes Corporation). Sequence data were aligned using the MUSCLE multiple sequence alignment algorithm (Edgar 2004) implemented in MacVector v14.0.1 (MacVector, Inc.).

Summary statistics for the mtDNA control region sequence data including haplotype and nucleotide diversities, the number of polymorphic sites, and the mean number of pairwise differences were calculated using Arlequin. The evolutionary model that best described the sequence data (Kimura 3-parameter with unequal base frequencies, K3P; Kimura 1981) was identified using the model selection method implemented in PAUP* v4.0 beta (Swofford 2002) and based on the corrected Akaike information criterion. Genetic divergence was evaluated between all possible pairwise comparisons of the four currently recognized spearfish species by calculating Φ_{ST} values in Arlequin based on Tamura's (1992) genetic distance, the model closest to K3P implemented in Arlequin. To visualize genealogical relationships among individuals, a median joining haplotype network was generated from the mtDNA sequence data using the methods of Bandelt et al. (1999) implemented in PopArt v1.7 (Leigh and Bryant 2015). Signatures of recent population expansion were evaluated by calculating Fu's *F*s (Fu 1997) in Arlequin and Ramos-Onsins and Rozas' R2 (Ramos-Onsins and Rozas 2000) in the R package pegas v0.10 (Paradis 2010). Statistical significance of *F*s and R_2 values was assessed using 10,000 bootstrap replicates.

Results

Summary Statistics and Genetic Diversity.—Multilocus microsatellite genotypes were generated for 14 loci across 208 individuals, including 89 roundscale spearfish, 76 longbill spearfish, 29 shortbill spearfish, 12 Mediterranean spearfish, and the unknown morphologically identified longbill spearfish sampled off eastern South Africa (UNK_mLBS; Table 1). The subset of samples that were analyzed twice to assess error rates indicated that all loci were scored consistently. A number of loci (roundscale spearfish: Tge151, Tge119; longbill spearfish: Tg35, Tge79, Tg90; shortbill spearfish: Tge23, Tge79) were identified by Micro-Checker as having a general excess of homozygotes and potentially harboring null alleles. One of these loci (Tge151) was out of HWE for roundscale spearfish, and two of these loci (Tg35, Tge79) plus an additional locus (Tge139) were out of HWE for longbill spearfish (Table 2). All deviations were associated with heterozygote deficiencies, except for locus Tge139, for which neither statistical tests of heterozygote deficiency nor heterozygote excess were significant. As no locus was out of HWE across all species, all loci were retained for subsequent analyses. There was no evidence of linkage disequilibrium for any pairs of loci. The total number of alleles observed across all species combined at each locus ranged from four alleles at Tg11 to 50 alleles at Tg90 (Table 2). Within species, the average number of alleles observed at a locus ranged from four in Mediterranean spearfish to 14 in roundscale spearfish. Allelic richness and private allelic richness within species ranged from 3.45 to 7.94 and 0.22 to 2.83, respectively (Table 2). Both allelic richness and private allelic richness were lowest in Mediterranean spearfish; the highest values of allelic richness and private allelic richness were observed for shortbill spearfish and roundscale spearfish, respectively.

A 795-bp alignment of a segment of the mtDNA control region was generated for 83 individuals comprising 16 roundscale spearfish, 18 longbill spearfish, 37 shortbill spearfish, 10 Mediterranean spearfish, and UNK_mLBS (Table 1; GenBank accession

Table 2. Summary statistics for the 14 microsatellite loci surveyed in this study. a = observed number of alleles, a_R = rarefaction allelic richness, a_P = rarefaction private allelic richness, H_E = expected heterozygote frequency, H_O = observed heterozygote frequency, H_W = *P*-value associated with the comparison of observed and expected heterozygote frequencies. A critical value corrected for multiple pairwise comparisons (Benjamini and Yekutieli 2001, Narum 2006) and equal to 0.015 was used to assess statistical significance of conformance to Hardy-Weinberg equilibrium; statistically significant *P*-values are shown in bold. RSS = roundscale spearfish (*Tetrapturus georgii, n* = 89), LBS = longbill spearfish (*Tetrapturus pfluegeri, n* = 76), SBS = shortbill spearfish (*Tetrapturus angustirostris, n* = 29), MSF = Mediterranean spearfish (*Tetrapturus belone, n* = 12). Sample size for total is 206. Results displayed in this table were calculated excluding UNK_mLBS and SBS_Brazil.

	Locus														Average
Sample	Tge23	Tge76	Tge79	Tge119	Tge121	Tge135	Tge139	Tge144	Tge151	Tg10	Tg11	Tg35	Tg54	Tg90	across loci
RSS															
a	5	4	16	24	2	22	8	20	9	5	3	27	20	28	14
a_R	2.86	1.79	8.04	10.73	2.00	12.46	5.19	9.66	5.13	4.06	2.89	13.88	11.22	12.55	7.32
a_P	0.59	1.66	1.66	8.62	1.00	3.69	0.91	6.27	1.09	0.24	1.25	3.30	3.15	6.19	2.83
H_E	0.480	0.077	0.798	0.863	0.445	0.841	0.705	0.710	0.543	0.664	0.419	0.896	0.888	0.904	0.660
H_O	0.517	0.079	0.730	0.798	0.427	0.843	0.685	0.674	0.449	0.685	0.416	0.921	0.843	0.865	
H_W	0.835	1.000	0.320	0.207	0.815	0.862	0.160	0.318	**0.005**	0.020	0.934	0.948	0.106	0.047	
LBS															
a	2	5	14	9	3	15	17	6	18	6	3	26	22	34	13
a_R	1.97	3.67	8.75	7.24	2.06	9.81	9.97	4.49	10.25	3.86	1.66	12.96	11.81	14.30	7.34
a_P	0.00	0.00	0.96	1.80	0.38	1.57	3.61	0.32	2.46	0.21	0.02	2.21	1.96	3.02	1.32
H_E	0.239	0.488	0.834	0.819	0.126	0.863	0.892	0.688	0.851	0.577	0.064	0.855	0.878	0.826	0.643
H_O	0.224	0.500	0.605	0.763	0.132	0.908	0.855	0.711	0.868	0.553	0.039	0.632	0.842	0.763	
H_W	0.628	0.718	**0.000**	0.444	1.000	0.402	**0.009**	0.971	0.260	0.800	0.071	**0.000**	0.163	0.157	
SBS															
a	7	5	13	8	3	19	9	6	14	6	2	23	19	27	12
a_R	5.09	4.95	10.33	5.00	2.39	12.15	6.58	5.05	10.65	5.06	1.39	14.10	12.49	15.99	7.94
a_P	2.85	0.76	2.37	0.60	0.39	2.73	1.88	0.49	4.85	1.47	0.39	3.18	2.70	7.37	2.29
H_E	0.569	0.784	0.912	0.549	0.294	0.858	0.754	0.509	0.819	0.544	0.034	0.916	0.801	0.866	0.658
H_O	0.414	0.828	0.759	0.621	0.207	0.828	0.690	0.414	0.759	0.517	0.034	0.862	0.862	0.793	
H_W	0.037	0.760	0.108	0.199	0.265	0.250	0.016	0.209	0.831	0.106	–	0.634	0.867	0.218	

Table 2. *Continued.*

Sample	Locus														Average across loci
	Tge23	Tge76	Tge79	Tge119	Tge121	Tge135	Tge139	Tge144	Tge151	Tg10	Tg11	Tg35	Tg54	Tg90	
MSF															
a	1	4	3	4	2	2	3	4	3	3	1	7	4	8	4
a_R	1.00	3.92	3.00	3.92	2.00	2.00	2.92	4.00	3.00	3.00	1.00	6.58	4.00	8.00	3.45
a_P	0.00	0.01	0.01	0.54	0.00	0.00	0.30	0.02	0.01	0.00	0.00	0.17	1.24	0.80	0.22
H_E	0.000	0.631	0.583	0.573	0.290	0.159	0.554	0.728	0.605	0.475	0.000	0.699	0.739	0.678	0.480
H_O	0.000	0.500	0.500	0.583	0.333	0.167	0.583	0.583	0.417	0.417	0.000	0.667	0.917	0.750	
H_W	--	0.230	0.533	0.365	1.000	1.000	1.000	0.267	0.252	0.414	–	1.000	0.376	0.829	
Total															
a	8	8	20	28	5	25	17	21	23	8	4	35	30	50	20

Table 3. Diversity and neutrality test metrics for mtDNA control region sequence data. n = number of individuals sequenced, n_h = number of haplotypes, h = haplotype diversity, π = nucleotide diversity, polymorphic sites = the number of observed polymorphic sites, PWD = mean number of pairwise differences. Neutrality test metrics shown are Fu's F_S (Fu, 1997) and Ramos-Onsins and Rozas' R_2 (Ramos-Onsins and Rozas 2000); P-values are shown in parentheses and statistically significant values are bold. Final row in table refers to metrics calculated across all samples regardless of species. RSS = roundscale spearfish (*Tetrapturus georgii*), LBS = longbill spearfish (*Tetrapturus pfluegeri*), SBS = shortbill spearfish (*Tetrapturus angustirostris*), MSF = Mediterranean spearfish (*Tetrapturus belone*). Results displayed in this table were calculated excluding UNK_mLBS and SBS_Brazil.

	n	n_h	h	π	Polymorphic sites	PWD	F_S	R_2
RSS	16	16	1.00	0.024	96	18.53	**−5.17 (0.014)**	**0.052 (0.000)**
LBS	18	18	1.00	0.023	73	17.73	**−6.68 (0.006)**	**0.076 (0.009)**
SBS	37	32	0.99	0.023	96	17.44	**−11.20 (0.003)**	0.064 (0.055)
MSF	10	5	0.80	0.002	4	1.38	−1.47 (0.070)	0.168 (0.550)
All	81	71	0.99	0.071	251	55.01		

numbers MG542875–MG542957). There were 251 polymorphic sites, 204 transitions, and 48 transversions. Nucleotide composition was 34.74% A, 30.94% T, 19.48% C, and 14.84% G. Overall, 71 unique haplotypes were identified in the sequence data, and haplotype and nucleotide diversities across all individuals were 0.99 and 0.071, respectively (Table 3). Within species, haplotype and nucleotide diversities were lowest in Mediterranean spearfish (0.80 and 0.002, respectively); haplotype diversity was highest in roundscale spearfish and longbill spearfish (1.00), and nucleotide diversity was highest in roundscale spearfish (0.024). No haplotypes were shared among species.

Genetic Divergence.—Microsatellite F_{ST} values associated with the pairwise comparison of species ranged from 0.062 between longbill spearfish and shortbill spearfish to 0.247 between roundscale spearfish and Mediterranean spearfish (Table 4). This pattern was also observed in the mtDNA sequence data, where Φ_{ST} values were lowest for the pairwise comparison of longbill spearfish and shortbill spearfish (Φ_{ST} = 0.334), and greatest between roundscale spearfish and Mediterranean

Table 4. Results of pairwise divergence estimates among spearfishes. Metrics used to assess divergence based on microsatellite data consisted of F_{ST} (Wright 1951, Slatkin 1995), $(\delta\mu)^2$ (Goldstein et al. 1995), and chord distance (D_C; Cavalli-Sforza and Edwards, 1967). Φ_{ST} was used to quantify genetic divergence based on mtDNA sequence data. Permutation tests were used to assess the statistical significance of F_{ST} and Φ_{ST} values; all pairwise comparisons based on these metrics were statistically significant ($P < 0.0001$). RSS = roundscale spearfish (*Tetrapturus georgii*), LBS = longbill spearfish (*Tetrapturus pfluegeri*), SBS = shortbill spearfish (*Tetrapturus angustirostris*), MSF = Mediterranean spearfish (*Tetrapturus belone*). Results displayed in this table were calculated without UNK_mLBS and SBS_Brazil.

	F_{ST}	$(\delta\mu)^2$	D_C	Φ_{ST}
RSS vs LBS	0.172	60.86	0.501	0.877
RSS vs SBS	0.182	87.01	0.577	0.877
RSS vs MSF	0.247	71.08	0.639	0.921
LBS vs SBS	0.062	19.90	0.374	0.334
LBS vs MSF	0.150	20.38	0.445	0.372
SBS vs MSF	0.178	46.97	0.513	0.611

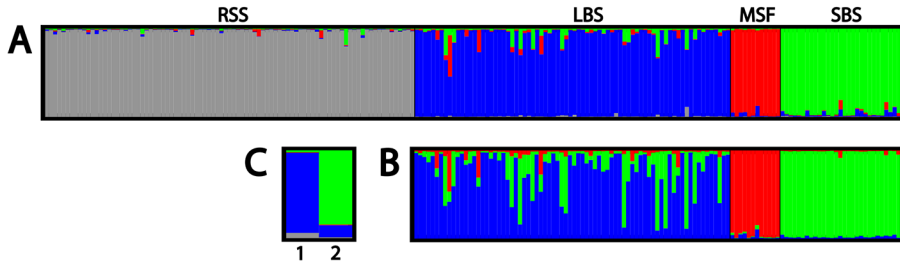

Figure 1. Barplots generated from Structure results based on microsatellite genotype data. RSS = roundscale spearfish, LBS = longbill spearfish, SBS = shortbill spearfish, MSF = Mediterranean spearfish, solid triangle = UNK_mLBS and SBS_Brazil. Bases of solid black arrowheads demarcate the last two individuals in the barplot comprising UNK_mLBS and SBS_Brazil, sequentially. *Panel A:* K = 4 barplot generated from dataset inclusive of all four *Tetrapturus* species. *Panel B:* K = 3 barplot generated from dataset with roundscale spearfish excluded. *Panel C:* Expanded view of UNK_mLBS (1) and SBS_Brazil (2) from the K = 4 analysis inclusive of all *Tetrapturus* species.

spearfish (Φ_{ST} = 0.921; Table 4). All pairwise comparisons of species based on F_{ST} and ϕ_{ST} were highly significant ($P < 0.0001$).

A similar pattern of genetic divergence was also observed for values of $(\delta\mu)^2$ and D_C calculated from the microsatellite genotype data. The lowest levels of genetic divergence were observed between longbill spearfish and shortbill spearfish [$(\delta\mu)^2$ = 19.90, D_C = 0.374; Table 4] and the higest levels of genetic divergence were observed between roundscale spearfish and shortbill spearfish [$(\delta\mu)^2$ = 87.01] and roundscale spearfish and Mediterranean spearfish (D_C = 0.639; Table 4). Intraspecific comparisons of geographically distant sample collections for shortbill spearfish and for roundscale spearfish were non-significant (results not shown).

Microsatellite Cluster Analyses.—Results from Structure using K = 4 differentiated four distinct genetic groups coterminous with the four currently recognized species of *Tetrapturus* (Fig. 1A). A low level of inferred mixed ancestry was present within each genetic group, with the exception of longbill spearfish, which primarily displayed shared ancestry with shortbill spearfish. Since roundscale spearfish was the most divergent species of *Tetrapturus* based on genetic distance metrics, Structure analyses excluding roundscale spearfish were also performed (K = 3; Fig. 1B). This analysis revealed a greater level of admixture within longbill spearfish as compared to results using the full data set. The genetic composition of UNK_mLBS was consistent with other longbill spearfish individuals (Fig. 1). Interestingly, the genetic composition of an individual sampled from the western central Atlantic Ocean off Brazil was consistent with shortbill spearfish (Fig. 1). This individual was originally presumed to represent an Atlantic species of *Tetrapturus*; however, our results suggest otherwise and this individual is hereafter referred to as SBS_Brazil. Based on the Structure results, all diversity and divergence metrics were recalculated excluding SBS_Brazil and these are the results presented in Tables 2, 3, and 4. Inclusion of this individual in diversity and divergence value calculations did not substantially alter results (data not shown).

A two-dimensional plot of the first two principal component axes associated with PCA of the microsatellite data produced three clusters of individuals (Fig. 2A). The

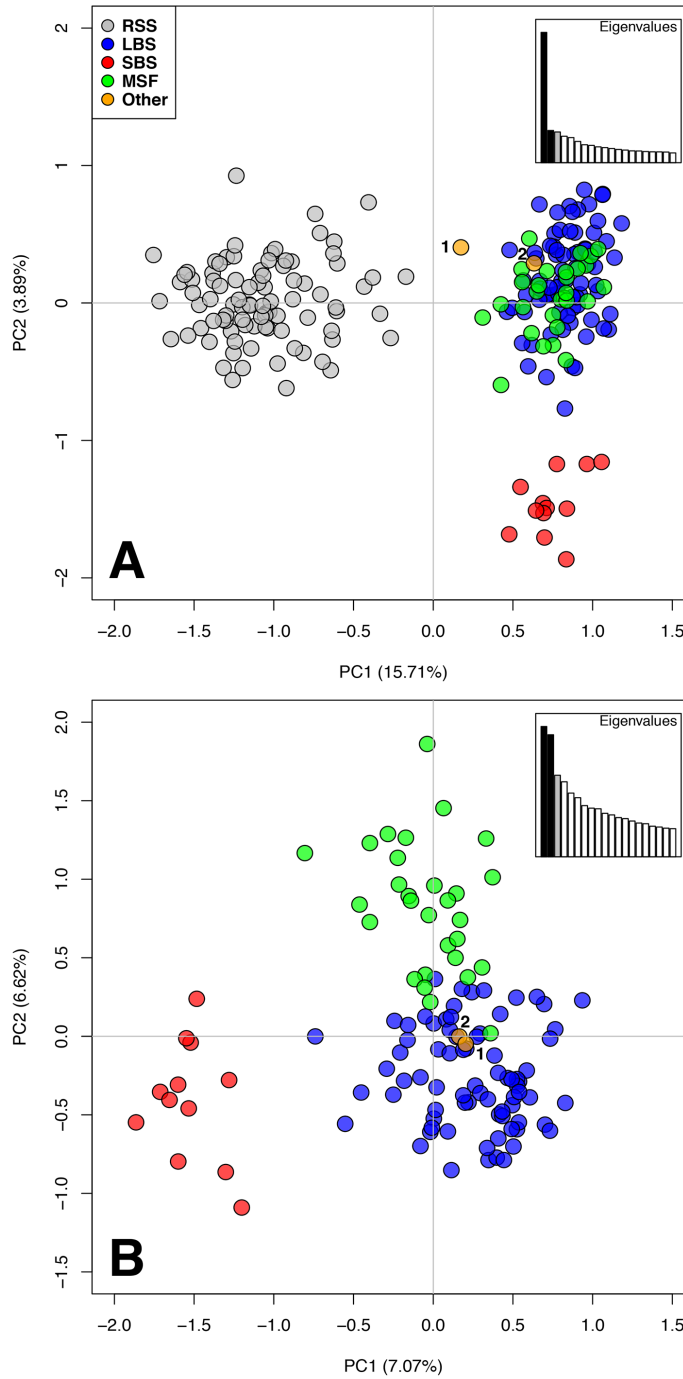

Figure 2. Two dimensional PCA plots displaying principal component axes one and two. Points are colored by species; eigen values associated with plotted principal components are shown in insets. RSS = roundscale spearfish, LBS = longbill spearfish, SBS = shortbill spearfish, MSF = Mediterranean spearfish, Other = UNK_mLBS (1) and SBS_Brazil (2). (A) Results from data set inclusive of all four *Tetrapturus* species. (B) Results from data set with roundscale spearfish excluded.

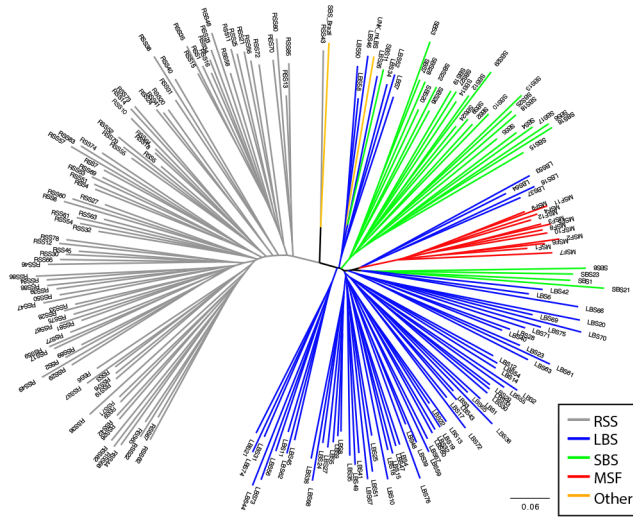

Figure 3. Unrooted neighbor joining tree based on Cavalli-Sforza and Edwards (1967) chord distances calculated for all possible pairwise comparisons of individuals. Branches are colored by species. RSS = roundscale spearfish, LBS = longbill spearfish, SBS = shortbill spearfish, MSF = Mediterranean spearfish, Other = UNK_mLBS and SBS_Brazil.

clusters resolved on these axes correspond with sample collections of roundscale spearfish, Mediterranean spearfish, and a third cluster comprising longbill spearfish + shortbill spearfish. UNK_mLBS is positioned intermediate to roundscale spearfish and shortbill spearfish + longbill spearfish, while SBS_Brazil is positioned well within the cluster comprising longbill spearfish + shortbill spearfish. Results from PCA using the data set with roundscale spearfish excluded show longbill spearfish and shortbill spearfish separated into two partially overlapping groups based on principal components one and two (Fig. 2B). In this analysis, UNK_mLBS and SBS_Brazil were resolved well within shortbill spearfish + longbill spearfish. An unrooted neighbor joining tree based on D_C genetic distances calculated for pairwise comparisons of all individuals is presented in Figure 3. Roundscale spearfish form a distinct clade with the exception of one individual (RSS43) that groups closely with SBS_Brazil. Mediterranean spearfish form a single cluster; however, this cluster is placed within a larger grouping also composed of longbill spearfish and shortbill spearfish individuals. A number of clades specific to longbill spearfish or shortbill spearfish were resolved, but these species were not reliably discriminated. The individual sampled off eastern South Africa, UNK_mLBS, was placed within a clade composed of both longbill spearfish and shortbill spearfish.

mtDNA Cluster Analyses.—A median joining haplotype network generated from the mtDNA control region sequence data revealed a high level of genetic differentiation between roundscale spearfish and all other *Tetrapturus* species (74 fixed differences; Fig. 4). Longbill spearfish and shortbill spearfish were not clearly resolved. Mediterranean spearfish cluster together within the network; however, haplotypes from this species displayed limited diversity and are separated from longbill spearfish by a single fixed mutational difference. UNK_mLBS was positioned among other shortbill spearfish and longbill spearfish haplotypes in the network. The haplotype for SBS_Brazil was placed in a portion of the network composed primarily of longbill spearfish and adjacent to Mediterranean spearfish.

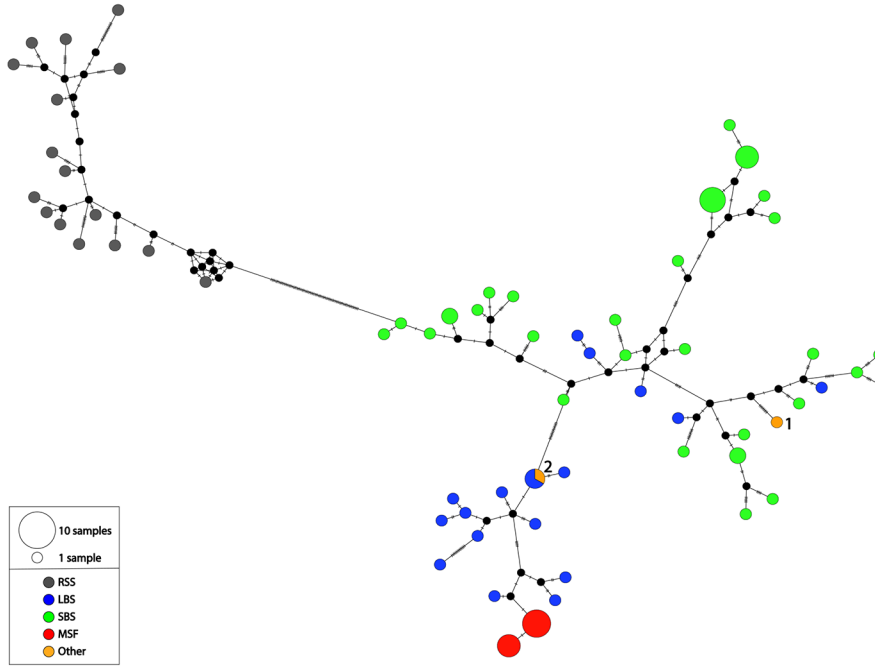

Figure 4. Median joining haplotype network based on mtDNA control region sequence data. Colored nodes represent observed haplotypes colored by species; black nodes represent inferred ancestral nodes. Hatch marks along edges indicate the number of mutational differences between nodes. RSS = roundscale spearfish, LBS = longbill spearfish, SBS = shortbill spearfish, MSF = Mediterranean spearfish, Other = UNK_mLBS (1) and SBS_Brazil (2).

Genetic Assignment Tests.—The probability of assignment to each of the four species of *Tetrapturus* for all individuals calculated using GeneClass2 (Online Table S2) indicated that all roundscale spearfish individuals assigned to roundscale spearfish (median assignment probability = 0.832), and probabilities of assignment to the other three species were insignificant ($P < 0.05$) for all individuals. Mediterranean spearfish individuals displayed the highest probabilities of assignment to Mediterranean spearfish (median = 0.739), but were not significantly excluded from longbill spearfish (median = 0.434) or shortbill spearfish (median = 0.086); probabilities of assignment to roundscale spearfish were insignificant ($P < 0.001$). Longbill spearfish individuals had insignificant probabilities of assignment to roundscale spearfish and to Mediterranean spearfish ($P \leq 0.01$); probabilities of assignment to shortbill spearfish were significant for most longbill spearfish individuals (87% using a criteria of $P > 0.01$), but assignment to longbill spearfish was most probable in all cases (0.149–0.998 (median = 0.828) for longbill spearfish vs 0.000–0.594 (median = 0.044) for shortbill spearfish). For shortbill spearfish, there were no individuals that displayed a significant probability of assignment to roundscale spearfish or to Mediterranean spearfish. Probabilities of assignment to longbill spearfish were significant for approximately 50% of shortbill spearfish individuals ($P > 0.01$); however, the highest assignment probability for shortbill spearfish individuals were all associated with assignment to shortbill spearfish [0.567–1.000 (median = 0.900) for

assignment to shortbill spearfish vs 0.000–0.596 (median = 0.012) for assignment to longbill spearfish].

Assignment probabilities were also calculated for the spearfish samples of uncertain species identification using GeneClass2. Assignment probabilities for UNK_mLBS were 0.000 for assignment to Mediterranean spearfish, 0.003 for assignment to roundscale spearfish, 0.038 for assignment to shortbill spearfish, and 0.193 for assignment to longbill spearfish (Online Table S2). For SBS_Brazil, the probability of assignment to any of the four spearfishes was low, although the highest probability was associated with assignment to shortbill spearfish (P = 0.046; Online Table S2). Four individuals were identified as probable first-generation migrants based on the results of Bayesian likelihood ratios also calculated using GeneClass2; two individuals captured off Brazil, including SBS_Brazil, were identified as migrant shortbill spearfish. The individual sampled off Richards Bay, South Africa (UNK_mLBS) was identified as a migrant longbill spearfish; a shortbill spearfish captured off Hawaii was also identified as a first generation migrant longbill spearfish.

Species Demographic History.—Values calculated for Fu's Fs and Ramos-Onsins and Rozas' R_2 based on mtDNA sequence data were statistically significant for roundscale spearfish and for longbill spearfish (Table 3). A statistically significant value of Fs (Table 3) was also calculated for shortbill spearfish; however, R_2 was not significant for this species (P = 0.055; Table 3). Values of Fs, and R_2 were not statistically significant for Mediterranean spearfish. Collectively, these results are consistent with recent population expansions for roundscale spearfish, longbill spearfish, and shortbill spearfish. There was no evidence to suggest recent reductions in effective population size for any *Tetrapturus* species, including results based on the graphical evaluation of allele frequency distributions generated in BOTTLENECK (results not shown).

Discussion

We used highly polymorphic microsatellite markers and the mtDNA control region to delineate species and evaluate geographic spatial distributions of the spearfishes. Compared to previous molecular studies, we assessed substantially larger sample sizes per species, including sample collections representative of more than one geographic location for some species, across more molecular markers with greater power to resolve low levels of genetic differentiation. Overall, Bayesian model-based clustering methods and assignment tests based on multilocus microsatellite data were consistent with the four currently recognized species of *Tetrapturus*. However, longbill spearfish and shortbill spearfish individuals could not be unambiguously discriminated using assignment tests based on either microsatellite genotypes or mtDNA sequences. Results of the present study demonstrate close genetic relationships among the spearfishes and reveal a range of genetic divergences among species and differences in levels of genetic diversity within species. Results also suggest limited inter-oceanic exchange.

Discrimination of Species and Species Relationships.—Relationships among *Tetrapturus* based on sequencing of the mtDNA control region in this study generally agree with the relationships recovered by Collette et al. (2006). Roundscale

spearfish was the most divergent member of *Tetrapturus,* Mediterranean spearfish was most closely allied with longbill spearfish, and longbill spearfish and shortbill spearfish were least divergent based on Φ_{ST} values (Table 4). Although longbill spearfish and shortbill spearfish did not share any control region haplotypes, some longbill spearfish haplotypes were more similar to shortbill spearfish haplotypes than to other longbill spearfish haplotypes in the median joining network (Fig. 4). This could be interpreted as evidence that longbill spearfish and shortbill spearfish are reflective of populations rather than distinct species. However, the level of divergence between longbill spearfish and shortbill spearfish observed in our study (Φ_{ST} = 0.334) is an order of magnitude higher than the largest values calculated among genetically distinct populations of the confamilial striped marlin (*Kajikia audax*) based on the same segment of the mtDNA control region (Φ_{ST} = 0.043 between collections of striped marlin sampled off Manta, Ecuador and Port Stephens, Australia; McDowell and Graves 2008). These results suggest that the level of divergence between longbill and shortbill spearfishes is, in fact, consistent with their status as distinct species. However, our finding using a rapidly evolving mtDNA region corroborates the conclusion of Hanner et al. (2011) that mtDNA is not sufficient for discrimination of species within *Tetrapturus*.

Relationships among *Tetrapturus* based on microsatellite loci were concordant with those using mtDNA control region sequences, and this pattern was consistent across multiple pairwise divergence estimates representative of different underlying mutation models (Table 4). Although Bayesian model-based clustering was able to separate shortbill spearfish and longbill spearfish (Fig. 1), ancestry proportions of individual longbill spearfish showed admixture with shortbill spearfish and Mediterranean spearfish. Longbill and shortbill spearfishes were not recovered as discrete groups using multivariate analysis (Fig. 2) or in a neighbor-joining tree based on Cavalli-Sforza chord distances (Fig. 3). The probability of assignment to each of the four species of *Tetrapturus* using the methods implemented in GeneClass2 indicated that all roundscale spearfish could be unequivocally assigned based on 14 microsatellite loci. While longbill spearfish and shortbill spearfish always displayed the highest probabilities of assignment to the correct species, assignment to the alternate species could not be excluded in many cases. Mediterranean spearfish could not be correctly assigned to species based on microsatellite data using allele frequency based methods, likely due to the low number of samples representative of this species; however, individuals comprising Mediterranean spearfish were well delineated using PCA, Bayesian clustering, and in the unrooted neighbor joining tree (Figs. 1–3).

Microsatellite-based comparisons between longbill and shortbill spearfishes were associated with the lowest calculated divergence values for all metrics (Table 4), but F_{ST} and Φ_{ST} comparisons between these species were significant ($P < 0.0001$). The F_{ST} value associated with the pairwise comparison of longbill spearfish and shortbill spearfish (F_{ST} = 0.062; Table 4) is nearly double the highest value calculated between populations of striped marlin using a similar number of microsatellite loci (F_{ST} = 0.0377 between collections of striped marlin sampled off eastern Australia and off the Pacific coast of Mexico based on 12 microsatellite loci; Purcell and Edmands 2011).

The analysis of highly polymorphic molecular markers in this study considerably improved the ability to resolve the spearfishes. Although a number of previous studies have utilized a molecular approach to delineate geographic distributions and

genetic relationships among *Tetrapturus*, a limited number of samples and loci were examined across species. Most previous studies have reported roundscale spearfish as most genetically divergent and thus sister to all other species of *Tetrapturus* (Collette et al. 2006, Shivji et al. 2006, Hanner et al. 2011). However, Santini and Sorenson (2013) reported longbill spearfish as sister to all other *Tetrapturus*, with roundscale spearfish sister to shortbill spearfish + Mediterranean spearfish based on the analysis of seven mitochondrial and three nuclear loci concatenated from sequences that were publicly available in GenBank. Although Collette et al. (2006) were able to resolve all four species of *Tetrapturus* using three mitochondrial and one nuclear gene region, only two exemplars of each species were included in the analysis. Hanner et al. (2011) analyzed cytochrome c oxidase subunit I (COI) sequences from 124 spearfishes to test the utility of barcoding for identification of billfishes, but were unable to separate shortbill spearfish vs. Mediterranean spearfish vs longbill spearfish using either COI or the nuclear rhodopsin gene. Collectively, these previous studies of taxonomic relationships within the spearfishes have used a variety of mitochondrial and nuclear markers, yet they have been unable to consistently resolve all members of *Tetrapturus*. In addition, none of these studies included both more than a few individuals of each species and the use of multiple independent loci.

The inability to consistently discriminate shortbill and longbill spearfishes in this and previous studies is likely due to incomplete lineage sorting attributable to a speciation event that is recent as compared to the ancestral effective population size (Pamilo and Nei 1988, Maddison 1997). Alternately, the inability to discriminate between longbill spearfish and shortbill spearfish could be attributed to introgression (i.e., secondary gene flow between species; Petit and Excoffier 2009). These scenarios are not mutually exclusive and speciation despite some level of intermittent gene flow has been demonstrated in other organisms (Hey 2006, Nosil 2008). In the case of speciation with ongoing gene flow, gene flow should preferentially occur among neighboring groups while shared polymorphisms should be evenly distributed across geographic space if incomplete lineage sorting is the cause of the observed pattern (Petit and Excoffier 2009). However, because spearfish are rare event species, this study did not have sufficient geographic coverage to distinguish between these alternatives. There are also many coalescent methods aimed at distinguishing incomplete lineage sorting from speciation with ongoing gene flow; for example, isolation with migration models (Wakeley and Hey 1998, Nielsen and Wakeley 2001, Hey 2010), likelihood ratio tests (Yang 2010), and approximate Bayesian computation (Kuhner 2009, Sunnåker et al. 2013). These methods were not implemented in this study because analyses would be more relevant if a greater number of informative characters could be included (e.g., incorporating genotyping-by-sequencing data). These alternate scenarios may be further complicated by the presence of vagrants or by an incomplete understanding of distributional ranges for these two species (*see* below). Nonetheless, the difficulty discriminating among *Tetrapturus* species indicates that the microsatellite markers used in the present study did not have sufficient power to unambiguously classify all individuals, or to discriminate among alternate evolutionary hypotheses.

Identification of Unidentified Individuals and Range Uncertainties.— Results from genetic data analyzed in the present study are consistent with previous reports of interoceanic spearfish vagrants. Bayesian model-based clustering based

on microsatellite genotypes grouped one individual (SBS_Brazil) sampled from the western Central Atlantic Ocean off Brazil and initially presumed to be longbill spearfish based on capture location with shortbill spearfish, a species considered to be confined to the Indo-Pacific (Fig. 1). These results were consistent for Bayesian clustering using data sets inclusive (K = 4) and exclusive (K = 3) of roundscale spearfish. Assignment testing using the allele frequency based method implemented in GeneClass2 could not conclusively assign this individual to any of the spearfishes and probabilities were <0.05 for all species; however, likelihood ratio tests identified this individual as a migrant shortbill spearfish with high probability. Although the mtDNA sequence of SBS_Brazil was consistent with longbill spearfish, our data indicate that the mtDNA control region is not an adequate species discriminator for the spearfishes. Taken together, the preponderance of evidence suggests that this individual may indeed be a vagrant shortbill spearfish. Interestingly, Penrith (1964) identified a shortbill spearfish from the eastern coast of South Africa based on morphological characters, extending the known range for this species. Additionally, given the occurrence of both black marlin and striped marlin off the Cape of Good Hope in late summer, Penrith concluded "that there is little reason" that shortbill spearfish should not also appear off the Cape of Good Hope (Penrith 1964, Penrith and Cram 1974). Nakamura and Nakano (1978) later reported the capture of three shortbill spearfish from the Atlantic Ocean. These specimens were taken off the western coast of Africa between August 1974 and April 1975 during a longline survey by the Japan Marine Fishery Resource Center. Based on these results, Nakamura and Nakano (1978) concluded that the presence of shortbill spearfish in the South Atlantic Ocean reflects stray individuals entering the Atlantic Ocean from around the Cape of Good Hope. The putative shortbill spearfish identified in the present study was collected off the coast of Brazil, and it is plausible that this species may not be uncommon in this area given the capture location of the specimens in Nakamura and Nakano (1978).

Although longbill and shortbill spearfishes differ in some subtle morphological characters, such as bill length, these characters are often removed during shipboard processing. Furthermore, use of bill length as a character can be problematic as bills can be damaged. In 2009, Brazil prohibited the commercial sale of white marlin (*K. albida*) and blue marlin (*Makaira nigricans* Lacépède, 1802), and consequently billfish began to be marketed headless to obscure species identification (Piva-Silva and Amorim 2014). In addition, *Tetrapturus* are often collectively identified as a group with white marlin in Brazilian landing records (Piva-Silva and Amorim 2014). Practices such as these contribute to the lack of understanding regarding the geographic ranges of these rare-event species and emphasizes the need for markers capable of genetic species discrimination.

A second individual sampled during the course of our study suggests that longbill spearfish are also capable of traversing the Cape of Good Hope. Recently, a recreational angler (J Booysen, Richards Bay Ski Boat Club, Richards Bay, South Africa, pers comm) captured a purported longbill spearfish in the Indian Ocean off Richards Bay, South Africa, and a tissue sample was sent to our lab at the Virginia Institute of Marine Science (UNK_mLBS). Although results from the control region sequence generated for this individual were equivocal, this sample clustered with other longbill spearfish in the microsatellite-based Structure analysis (Fig. 1) and was assigned to longbill spearfish and identified as a longbill spearfish migrant by GeneClass2. This

finding strengthens the suggestion that the Cape of Good Hope facilitates some degree of connectivity between spearfishes in the Indian and Atlantic oceans (Penrith 1964, Penrith and Cram 1974), and emphasizes the importance of future molecular and morphological study of spearfishes sampled from this region.

Two additional spearfishes, one sampled off Hawaii and identified as a migrant longbill spearfish and one sampled off Brazil and identified as a migrant shortbill spearfish were also identified as first generation migrants based on likelihood-ratio tests. However, Structure analyses assigned both of these fish to the expected species based on geographic location with little to no evidence of admixture. Given the lack of concordance among different analyses for these two individuals, we did not consider these fish to represent vagrants. However, use of an increased number of molecular markers may be able to resolve this issue.

The uncertainty surrounding species distributional ranges is not unique to longbill spearfish and shortbill spearfish. The range of roundscale spearfish is poorly understood due to its resemblance to white marlin (*K. albida*; Bernard et al. 2013) and its misclassification as longbill spearfish (Graves and McDowell 2012). Mediterranean spearfish, although traditionally regarded as restricted to the Mediterranean Sea, has been reported off the Straits of Gibraltar (Di Natale et al. 2005) and the all-tackle record Mediterranean spearfish was caught off Madeira Island, Portugal (IGFA 2015). Collectively, this information suggests that the distributional ranges for all species of *Tetrapturus* require additional study.

Patterns of Intraspecific Genetic Diversity and Demographic History.—Mediterranean spearfish harbor reduced levels of genetic diversity compared to all other spearfishes based on both marker classes. Average expected heterozygosity (gene diversity) and rarefaction allelic richness across microsatellite loci were markedly lower in Mediterranean spearfish (Table 2). Although all microsatellite markers used in the present study were developed for roundscale spearfish and variability is expected to be higher in the focal species due to the criteria used to select loci (Ellegren et al. 1995), diversity estimates for longbill spearfish and shortbill spearfish, the other two non-focal species, were equivalent to estimates for roundscale spearfish. This pattern of diversity was also observed in the mtDNA control region sequence data, which are not impacted by ascertainment bias. Nucleotide diversity for Mediterranean spearfish was an order of magnitude lower than that observed for other spearfishes; the lowest observed haplotype diversity was also associated with this species. Although sample sizes were smallest for Mediterranean spearfish (n = 12), this result was not an artifact of small sample size; larger sample sizes were analyzed for longbill spearfish (n = 76) and roundscale spearfish (n = 89), and control region sequences for these species were all unique. As allelic richness is a known proxy for evolutionary potential (Caballero and García-Dorado 2013), Mediterranean spearfish may be of particular conservation concern, especially given recent increases in fishing pressure (Di Natale et al. 2005, Castriota et al. 2008).

Tests of demographic history (Fu's F_S, Ramos-Onsins and Rozas' R_2) were consistent with population expansion for roundscale spearfish, longbill spearfish, and shortbill spearfish. This is consistent with previous demographic analysis of roundscale spearfish (Bernard et al. 2013) and other billfishes, including white marlin (Bernard et al. 2013), swordfish, *Xiphias gladius* Linnaeus, 1758 (Alvarado Bremmer et al. 2005), and sailfish, *Istiophorus platypterus* (Shaw, 1792) (Bagma 2006), as well

as other pelagic species including bluefin tuna, *Thunnus thynnus* (Linnaeus, 1758) (Carlsson et al. 2004), and Sierra mackerel, *Scomberomorus sierra* Jordan and Starks, 1895 (Lopez et al. 2010). The observed signal of expansion is consistent with an increase in available habitat at the end of the Pleistocene.

Concluding Remarks.—Reliance on the use of single exemplars, unvouchered specimens, and a limited number of molecular markers (e.g., mtDNA) combined with the rarity and morphological conservatism of spearfishes have collectively hindered our understanding of these species. Results from assignment tests and Bayesian model-based clustering methods based on microsatellite data generated in the present study are consistent with the four currently recognized *Tetrapturus* species. These results confirm that geographic location of capture is not a perfect indicator of species identity for the spearfishes and also confirms the predictions and observations of earlier works, which suggest that the Cape of Good Hope is not an absolute biogeographic barrier for spearfishes. Although the genetic data presented in the present study generally corroborate the findings of Collette et al. (2006), inclusion of multiple individuals of longbill and shortbill spearfishes demonstrates that these species are difficult to unequivocally discriminate, even using a relatively large number of highly polymorphic loci. As vouchering specimens of large pelagic species is often impractical, future genetic studies of *Tetrapturus* should incorporate morphometrics and photographs for each genetic sample whenever possible, especially where geographic ranges overlap. We acknowledge that the sample collections evaluated in our study were limited in geographic scope. For example, longbill spearfish were represented by a single collection of individuals sampled off Brazil, and a single exemplar of shortbill spearfish from the Indian Ocean was available for analysis. Future studies should include larger sample sizes encompassing the entire known distributional ranges of these species to account for the impact of possible population structure on the ability to discriminate species. In addition, future studies would benefit from high throughput genetic approaches capable of surveying hundreds to thousands of molecular makers across large numbers of individuals. In combination, these approaches would allow discrimination of alternative evolutionary scenarios.

Acknowledgments

We thank the recreational anglers whose willingness to provide samples made this project possible. We thank F Hazin and the Mid-Atlantic 500 tournament for providing the opportunity to collect samples. We thank Stephen Sherrer for collection of samples from Hawaii. We also thank Bruce Collette and three anonymous reviewers for comments, which helped improve our submitted manuscript. This study was supported by grant NA12NMF4690243 and the Gottwald Marine Science Endowment (VIMS) to JEG and JRM from NOAA-NMFS. This paper is Contribution No. 3717 of the Virginia Institute of Marine Science, College of William & Mary.

Literature Cited

Alvarado Bremer JR, Vinas J, Mejuto J, Ely B, 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 phylo-genies of two highly migratory pelagic fishes. Mol Phylogenet Evol. 36:169–187. https://doi.org/10.1016/j.ympev.2004.12.011

Bagma J. 2006. Contemporary population structure and historical demography of sailfish (*Istiophorus platypterus*) in the Atlantic Ocean. Master's thesis, Texas A&M University. Available from: http://hdl.handle.net/1969.1/ETD-TAMU-1876

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 16:37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Benjamini Y, Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. Ann Stat. 29:1165–1188.

Bernard A, Feldheim K, Shivji M. 2012. Development and characterization of 11 novel microsatellite loci for the roundscale spearfish *Tetrapturus georgii* and their cross-species amplification among other istiophorid species. J Fish Biol. 81:1781–1786. https://doi.org/10.1111/j.1095-8649.2012.03442.x

Bernard A, Shivji M, Domingues R, Hazin F, de Amorim A, Domingo A, Arocha F, Prince E, Hoolihan J, Wagner Silva Hilsdorf A. 2013. Broad geographic distribution of roundscale spearfish (*Tetrapturus georgii*) (Teleostei, Istiophoridae) in the Atlantic revealed by DNA analysis: implications for white marlin and roundscale spearfish management. Fish Res. 139:93–97. https://doi.org/10.1016/j.fishres.2012.10.009

Bernard A, Shivji MS, Prince ED, Hazin F, Arocha F, Domingo A, Feldheim K. 2014. Comparative population genetics and evolutionary history of two commonly misidentified billfishes of management and conservation concern. BMC Genet. 15:141. https://doi.org/10.1186/s12863-014-0141-4

Caballero A, García-Dorado A. 2013. Allelic diversity and its implications for the rate of adaptation. Genetics. 195:1373–1384. https://doi.org/10.1534/genetics.113.158410

Carlsson J, McDowell JR, Diaz-Jaimes P, Carlsson JEL, Boles S, Gold JR, Graves JE. 2004. Micrisatellite and mitochondrial DNA analysis of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) populations structure in the Mediterranean Sea. Mol Ecol. 13:3345–3356. https://doi.org/10.1111/j.1365-294X.2004.02336.x

Castriota L, Finoia M, Campagnuolo S, Romeo T, Potoschi A, Andaloro F. 2008. Diet of *Tetrapturus belone* (Istiophoridae) in the central Mediterranean Sea. J Mar Biol Assoc U K. 88:183–187. https://doi.org/10.1017/S0025315408000106

Cavalli-Sforza L, Edwards A. 1967. Phylogenetic analysis: models and estimation procedures. Evolution. 21:550–570. https://doi.org/10.1111/j.1558-5646.1967.tb03411.x

Collette B, McDowell J, Graves J. 2006. Phylogeny of recent billfishes (Xiphioidei). Bull Mar Sci. 79:455–468.

Cornuet J, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics. 144:2001–2014.

Di Natale A, Mangano A, Celona A, Valastro M. 2005. Size frequency composition of the Mediterranean spearfish (*Tetrapturus belone*, Rafinesque) catches in the Tyrrhenian Sea and in the Strait of Messina in 2003. Col Vol Sci Pap ICCAT. 55:692–709.

Dieringer D, Schlötterer C. 2003. Microsatellite analyzer (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Notes. 3:167–169. https://doi.org/10.1046/j.1471-8286.2003.00351.x

Edgar R. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797. https://doi.org/10.1093/nar/gkh340

Ellegren H, Primmmer CR, Sheldon BC. 1995. Microsatellite evolution-directionality or bias. Nat Genet. 11:360–362. https://doi.org/10.1038/ng1295-360

Excoffier L, Lischer H. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 10:564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Falush D, Stephens M, Pritchard J. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes. 7:574–578. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Fu Y. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 147:915–925.

Goldstein D, Linares A, Cavalli-Sforza L, Feldman M. 1995. An evaluation of genetic distances for use with microsatellite loci. Genetics. 139:463–471.

Graves J, McDowell J. 2006. Genetic analysis of white marlin (*Tetrapturus albidus*) stock structure. Bull Mar Sci. 79:469–482.

Graves JE, McDowell JR. 2012. Inter-annual variability in the proportion of round-scale spearfish (*Tetrapturus georgii*) and white marlin (*Kajikia albida*) in the western North Atlantic Ocean. Col Vol Sci Pap (International Commission for the Conservation of Atlantic Tunas). 68:1543–1547.

Hanner R, Floyd R, Bernard A, Collette B, Shivji M. 2011. DNA barcoding of billfishes. Mitochondrial DNA. 22:27–36.

Hey J. 2006. Recent advances in assessing gene flow between diverging populations and species. Curr Opin Genet Dev. 16:592–596. https://doi.org/10.1016/j.gde.2006.10.005

Hey J. 2010. Isolation with migration models for more than two populations. Mol Biol Evol. 27:905–920. https://doi.org/10.1093/molbev/msp296

Hubisz M, Falush D, Stephens M, Pritchard J. 2009. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour. 9:1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x

IGFA (The International Game Fish Association). 2015. Spearfish, Mediterranean (*Tetrapturus belone*). All-tackle world records. Accessed January 2018. Available from: http://wrec.igfa.org/WRecordsList.aspx?lc=AllTackle&cn=Spearfish,%20Mediterranean

Jakobsson M, Rosenberg N. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 23:1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 24:1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart T, Pontier D, Dufour A. 2009. Genetic markers in the playground of multivariate analysis. Heredity. 102:330–341. https://doi.org/10.1038/hdy.2008.130

Kalinowski S. 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. Conserv Genet. 5:539–543. https://doi.org/10.1023/B:COGE.0000041021.91777.1a

Kalinowski S. 2005. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes. 5:187–189. https://doi.org/10.1111/j.1471-8286.2004.00845.x

Kimura M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc Natl Acad Sci USA. 78:454–458. https://doi.org/10.1073/pnas.78.1.454

Kimura M, Crow J. 1964. The number of alleles that can be maintained in a finite population. Genetics. 49:725–738.

Kimura M, Ohta T. 1978. Stepwise mutation model and the distribution of allelic frequencies in a finite population. Proc Natl Acad Sci USA. 75:2868–2872. https://doi.org/10.1073/pnas.75.6.2868

Kuhner M. 2009. Coalescent genealogy samplers: windows into population history. Trends Ecol Evol. 24:86–93. https://doi.org/10.1016/j.tree.2008.09.007

Leigh J, Bryant D. 2015. Popart: full-feature software for haplotype network construction. Methods Ecol Evol. 6:1110–1116. https://doi.org/10.1111/2041-210X.12410

Levene H. 1949. On a matching problem arising in genetics. Ann Math Stat. 20:91–94. https://doi.org/10.1214/aoms/1177730093

López MD, Alcocer MU, Jaimes PD. 2010. Phylogeography and historical demography of the Pacific Sierra mackerel (*Scomberomorus sierra*) in the eastern Pacific. BMC Genet. 11:34. https://doi.org/10.1186/1471-2156-11-34

Lowe R. 1840. On new species of fishes from Madeira. Proc Zool Soc Lond. 8:36–39.

Maddison W. 1997. Gene trees in species trees. Syst Biol. 46:523–536. https://doi.org/10.1093/sysbio/46.3.523

McDowell J, Diaz-Jaimes P, Graves J. 2002. Isolation and characterization of seven tetranucleotide microsatellite loci from Atlantic northern bluefin tuna *Thunnus thynnus thynnus.* Mol Ecol Notes. 2:214–216. https://doi.org/10.1046/j.1471-8286.2002.00197.x

McDowell J, Graves J. 2008. Population structure of striped marlin (*Kajikia audax*) in the Pacific Ocean based on analysis of microsatellite and mitochondrial DNA. Can J Fish Aquat Sci. 65:1307–1320. https://doi.org/10.1139/F08-054

Nakamura I. 1985. FAO Species Catalogue Vol. 5. Billfishes of the world. An annotated and illustrated catalogue of marlins, sailfishes, spearfishes and swordfishes known to date. Rome, Italy: Food and Agriculture Organization of the United Nations. FAO Fisheries Synopsis No. 125. 65 p.

Nakamura I, Nakano S. 1978. Dispersal of the shortbill spearfish, *Tetrapturus angustirostris*, to the Atlantic Ocean. Copeia. (2):330–333. https://doi.org/10.2307/1443572

Narum S. 2006. Beyond Bonferroni: less conservative analyses for conservation genetics. Conserv Genet. 7:783–787. https://doi.org/10.1007/s10592-005-9056-y

Nielsen R, Wakeley J. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. Genetics. 158:885–896.

Nosil P. 2008. Speciation with gene flow could be common. Mol Ecol. 17(9):2103–2106. https://doi.org/10.1111/j.1365-294X.2008.03715.x

Paetkau D, Slade R, Burden M, Estoup A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol Ecol. 13:55–65. https://doi.org/10.1046/j.1365-294X.2004.02008.x

Palumbi S. 1996. Nucleic acids II. the polymerase chain reaction. *In:* Hillis D, Moritz C, Mable B, editors. Molecular Systematics. 2nd ed. Sunderland, MA: Sinauer Associates.

Pamilo P, Nei M. 1988. Relationships between gene trees and species trees. Mol Biol Evol. 5:568–583.

Paradis E. 2010. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics. 26(3):419–420. https://doi.org/10.1093/bioinformatics/btp696

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 20:289–290. https://doi.org/10.1093/bioinformatics/btg412

Penrith M. 1964. A marked extension of the known range of *Tetrapterus angustirostris* in the Indian Ocean. Copeia. 1964:231–232. https://doi.org/10.2307/1440871

Penrith M, Cram D. 1974. The Cape of Good Hope: a hidden barrier to billfishes. *In:* Shomura R, Williams F, editors. Proceedings of the International Billfish Symposium, Kailua-Kona, Hawaii, 9–12 August, 1972, Part 2. Review and contributed papers. Washington, DC: National Oceanographic and Atmospheric Association.

Petit R, Excoffier L. 2009. Gene flow and species delimitation. Trends Ecol Evol. 24:386–393. https://doi.org/10.1016/j.tree.2009.02.011

Piry S, Alapetite A, Cornuet J, Paetkau D, Baudouin L, Estoup A. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. J Hered. 95:536–539. https://doi.org/10.1093/jhered/esh074

Piry S, Luikart G, Cornuet J. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. J Hered. 90:502–503. https://doi.org/10.1093/jhered/90.4.502

Piva-Silva N, Amorim A. 2014. Fishery biology of *Tetrapturus* (Osteichthyes, Istiophoridae) caught by São Paulo longliners off southern Brazil1. Col Vol Sci Pap ICCAT. 70:2490–2498.

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

Purcell C, Edmands S. 2011. Resolving the genetic structure of striped marlin, *Kajikia audax*, in the Pacific Ocean through spatial and temporal sampling of adult and immature fish. Can J Fish Aquat Sci. 68:1861–1875. https://doi.org/10.1139/f2011-104

Rambaut A. 2009. FigTree, a graphical viewer of phylogenetic trees. http://tree.bio.ed.ac.uk.

Ramos-Onsins S, Rozas J. 2000. Statistical properties of new neutrality tests against population growth. Mol Biol Evol. 19:2092–2100. https://doi.org/10.1093/oxfordjournals.molbev.a004034

Rannala B, Mountain J. 1997. Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci USA. 94:9197–9201. https://doi.org/10.1073/pnas.94.17.9197

Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered. 86:248–249. https://doi.org/10.1093/oxfordjournals.jhered.a111573

Robins C. 1974. The validity and status of the roundscale spearfish, *Tetrapturus georgei. In:* Shomura R, Williams F, editors. Proceedings of the 1st International Billfish Symposium. Washington, DC: National Oceanic and Atmospheric Administration.

Robins C, de Sylva D. 1960. Description and relationships of the longbill spearfish, *Tetrapturus belone*, based on western North Atlantic specimens. Bull Mar Sci. 10(4):383–413.

Robins C, de Sylva D. 1963. A new western Atlantic spearfish, *Tetrapturus pfluegeri*, with a redescription of the Mediterranean spearfish *Tetrapturus belone.* Bull Mar Sci. 13:84–122.

Rosenberg N. 2004. DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes. 4:137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Santini F, Sorenson L. 2013. First molecular timetree of billfishes (Istiophoriformes: Acanthomorpha) shows a Late Miocene radiation of marlins and allies. Ital J Zool (Modena). 80(4):481–489. https://doi.org/10.1080/11250003.2013.848945

Seutin G, White B, Boag P. 1991. Preservation of avian blood and tissue samples for DNA analyses. Can J Zool. 69:82–90. https://doi.org/10.1139/z91-013

Shivji M, Magnussen J, Beerkircher L, Hinteregger G, Lee D, Serafy J, Prince E. 2006. Validity, identification, and distribution of the roundscale spearfish, *Tetrapturus georgii* (Teleostei: Istiophoridae): morphological and molecular evidence. Bull Mar Sci. 79:483–491.

Sorenson L, McDowell J, Graves J. 2011. Isolation and characterization of microsatellite markers for blue marlin, *Makaira nigricans.* Conserv Genet Resour. 3:721–723. https://doi.org/10.1007/s12686-011-9441-4

Sunnåker M, Busetto A, Numminen E. 2013. Approximate Bayesian computation. PLOS Comput Biol. 9(1):e1002803. https://doi.org/10.1371/journal.pcbi.1002803

Swofford D. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, MA: Sinauer Associates.

Takezaki N, Nei M. 1996. Genetic distances and the reconstruction of phylogenetic trees from microsatellite DNA. Genetics. 144:389–399.

Tamura K. 1992. Estimation of the nuber of nucleotide substitutions when there are strong transition-transversion and G+C content biases. Mol Biol Evol. 9:678–687.

Van Oosterhout C, Hutchinson W, Wills D, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes. 4:535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Wakeley J, Hey J. 1998. Testing speciation models with DNA sequence data. *In:* DeSalle R, Schierwater B, editors. Molecular approaches to ecology and evolution. Berlin, Germany: De Gruyter.

Walsh P, Metzger D, Higuchi R. 1991. Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. 10:506–513.

Waples R, Gaggiotti O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol. 15:1419–1439. https://doi.org/10.1111/j.1365-294X.2006.02890.x

Yang Z. 2010. A likelihood ratio test of speciation with gene flow using genomic sequence data. Genome Biol Evol. 2:200–211. https://doi.org/10.1093/gbe/evq011