

3-2018

## Genetic evaluation of population structure in white marlin (*Kajikia albida*): the importance of statistical power

Nadya Mamoozadeh

*College of William and Mary - Virginia Institute of Marine Science*

Jan McDowell

*Virginia Institute of Marine Science*

Jay R. Rooker

John E. Graves

*Virginia Institute of Marine Science*

Follow this and additional works at: <https://scholarworks.wm.edu/vimsarticles>



Part of the [Aquaculture and Fisheries Commons](#)

---

### Recommended Citation

Mamoozadeh, Nadya; McDowell, Jan; Rooker, Jay R.; and Graves, John E., "Genetic evaluation of population structure in white marlin (*Kajikia albida*): the importance of statistical power" (2018). *VIMS Articles*. 681.

<https://scholarworks.wm.edu/vimsarticles/681>

This Article is brought to you for free and open access by the Virginia Institute of Marine Science at W&M ScholarWorks. It has been accepted for inclusion in VIMS Articles by an authorized administrator of W&M ScholarWorks. For more information, please contact [scholarworks@wm.edu](mailto:scholarworks@wm.edu).



## Contribution to the Themed Section: 'International Billfish Conference'

### Original Article

# Genetic evaluation of population structure in white marlin (*Kajikia albida*): the importance of statistical power

Nadya R. Mamoozadeh<sup>1\*</sup>, Jan R. McDowell<sup>1</sup>, Jay R. Rooker<sup>2</sup>, and John E. Graves<sup>1</sup>

<sup>1</sup>Department of Fisheries Science, Virginia Institute of Marine Science, College of William & Mary, 1375 Greate Road, Gloucester Point, VA 23062, USA

<sup>2</sup>Department of Marine Biology, Texas A&M University, 1001 Texas Clipper Road, Galveston, TX 77554, USA

\*Corresponding author: tel: +1 804 684 7434; fax: +1 804 684 7097; e-mail: [nrmamoozadeh@vims.edu](mailto:nrmamoozadeh@vims.edu)

Mamoozadeh, N. R., McDowell, J. R., Rooker, J. R., and Graves, J. E. Genetic evaluation of population structure in white marlin (*Kajikia albida*): the importance of statistical power. – ICES Journal of Marine Science, 75: 892–902.

Received 18 November 2016; revised 22 February 2017; accepted 23 February 2017; advance access publication 26 April 2017.

The genetic basis of population structure in white marlin (*Kajikia albida*) is not well understood. Previous evaluation of genetic population structure in this species utilized a small number of molecular markers to survey genetic variation across opportunistically collected samples of adults, resulting in statistically significant levels of genetic differentiation for some pairwise comparisons and global levels of genetic differentiation that approached statistical significance. This study increased statistical power to improve resolution of genetic population structure in white marlin by surveying a larger number of molecular markers across sample collections of increased size, including collections from additional geographic locations and a robust collection of larvae. Increased statistical power resulted in lower levels of genetic heterogeneity compared with the previous study, and results were consistent with the presence of a single genetic stock of white marlin in the Atlantic Ocean. These results indicate that when statistical power is low, the ability to distinguish noise from a true signal of population structure is compromised. This relationship is especially important for population genetic assessments of marine fishes where genetic differentiation, if it exists, is expected to be low.

**Keywords:** microsatellite, population genetics, statistical power, white marlin.

## Introduction

Population genetic studies of marine fishes provide information to fisheries managers useful for the identification of biological units relevant for assessment and management, and for maintaining unique genetic variation (Allendorf *et al.*, 1987; Ward, 2000; Ovenden *et al.*, 2015). Despite their utility, these studies are particularly challenging for many marine fishes because the large effective population sizes of these species limit genetic drift, resulting in low levels of genetic differentiation among populations (Ward *et al.*, 1994; Waples, 1998). A high level of statistical power is required to detect genetic differentiation in population genetic studies of marine fishes, and can be achieved by evaluating larger numbers of molecular markers and/or samples per putative population (Waples, 1998; Ryman *et al.*, 2006). Although the former strategy is facilitated by currently available laboratory methodologies, obtaining appropriate sample sizes for

population genetic studies of relatively rare-event species can be difficult.

Additional considerations for population genetic studies of highly migratory marine fishes are associated with sampling designs appropriate for these species. Highly migratory marine fishes are capable of long distance movements, and opportunistic experimental designs based on the sampling of geographically distant locations may not be informative because individuals of the same stock could be sampled from multiple locations depending on the time of year (Graves *et al.*, 1996; Carlsson *et al.*, 2007; Graves and McDowell, 2015). Many highly migratory marine fishes also display seasonal mixed assemblages, and opportunistic sampling may lead to sample collections representative of more than one stock, resulting in a noisy genetic signal that may mask genetic differentiation (Waples, 1998; Bowen *et al.*, 2005). For highly migratory marine fishes that display population structure,

populations are most likely to separate at the time of spawning (Graves *et al.*, 1996; Carlsson *et al.*, 2007; Graves and McDowell, 2015). Replacing opportunistic sampling with a biologically informed design that targets larvae and/or reproductively active adults improves the ability to detect population subdivision in these species, if it exists. This concept is illustrated in Atlantic bluefin tuna (*Thunnus thynnus*) for which eastern and western Atlantic stocks were initially recognized based on the presence of distinct spawning grounds in the Gulf of Mexico (GOM) and Mediterranean Sea as well as differences in biological characteristics between fish from these regions (Fromentin and Powers, 2005; Rooker *et al.*, 2007). Early genetic studies based on opportunistic sample collections were unable to detect population subdivision between putative stocks; however, subsequent genetic analyses that incorporated a biologically informed sampling design targeting larvae and mature adults on spawning grounds during the spawning season detected statistically significant population subdivision consistent with inferences from non-genetic data (Carlsson *et al.*, 2007; Boustany *et al.*, 2008).

Population genetic studies that incorporate sampling designs also inclusive of temporal replicates provide the ability to determine if an observation of statistically significant differentiation is stable over time and unlikely to result from artifacts such as the non-random sampling of populations, stochastic fluctuation in allele frequencies, or variation in reproductive success (Allendorf and Phelps, 1981; Waples and Teel, 1990; Hedgecock, 1994; Waples, 1998; Tessier and Bernatchez, 1999). Temporal replicates are particularly important for marine fishes because the level of genetic divergence among populations, if present, is expected to be low. In such cases, it becomes increasingly challenging to distinguish between noise and a weak but meaningful level of heterogeneity (Waples, 1998). Obtaining temporal replicates for some highly migratory marine fishes is especially challenging due to the relatively rare event nature of a number of these species. Frequently, these challenges result in very low sample sizes per sampling location in any given year, and necessitate the use of sample collections that are pooled across years for each location.

White marlin (*Kajikia albida*) is a highly migratory marine fish distributed throughout the Atlantic Ocean in temperate, subtropical, and tropical waters (Nakamura, 1985). This species is managed along with other pelagic fishes in the Atlantic Ocean by the member nations of the International Commission for the Conservation of Atlantic Tunas (ICCAT). Prior to the year 2000, ICCAT utilized a two stock management model for white marlin that recognized distinct populations in the North Atlantic and South Atlantic ocean. This model was based on the presence of seasonally displaced spawning grounds north and south of the equator, fishery dependent data that reflected limited catches in equatorial waters, and a lack of trans-equatorial movements reported for tagged fish (ICCAT, 1994). ICCAT adopted a single stock management model in 2000 in response to new fisheries data that indicated continuous catches across the equator, trans-oceanic and trans-equatorial tag returns for a few individuals, and a population genetic study that was unable to detect statistically significant intra-oceanic heterogeneity (Graves and McDowell, 2001).

Since the adoption of the single stock management model for white marlin, additional genetic study has suggested the possibility of stock structure in this species. Analysis of variation at five nuclear microsatellite markers revealed statistically significant genetic differentiation between collections of white marlin from

the western North Atlantic and western South Atlantic (WSA) ocean, and global levels of genetic differentiation that approached statistical significance based on the microsatellite data and on mitochondrial (mt) DNA control region sequence data (Graves and McDowell, 2006). The authors concluded that while the null hypothesis of genetic homogeneity could not be rejected, genetic population structure could exist but may not have been detected due to low statistical power and/or the opportunistic sampling design employed in the study (Graves and McDowell, 2006). Thus, there is considerable uncertainty regarding the suitability of a single stock management model for white marlin, and management of this species would benefit from a more thorough understanding of population structure.

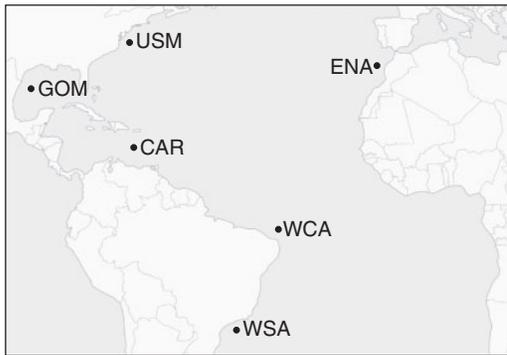
In this study, the null hypothesis of a single Atlantic-wide stock of white marlin was evaluated by surveying genetic variation at 24 microsatellite loci and at the mtDNA control region in multiple collections of white marlin sampled from locations throughout the Atlantic Ocean. Relative to Graves and McDowell (2006), the ability to detect genetic population structure was improved by analyzing a larger number of microsatellite markers, increasing the number of sampling locations and sample sizes for some locations, and including a robust collection of larvae from the GOM.

## Material and methods

### Sample collection and generation of nuclear genotype data

Samples consisting of muscle tissue from landed adult white marlin or fin clips from live-released adult white marlin were opportunistically collected between 1992 and 2014 from the following locations: United States mid-Atlantic coast (USM,  $n = 263$ ) off Cape May, NJ and Ocean City, MD; Caribbean Sea (CAR,  $n = 40$ ) off the Dominican Republic and Cumaná, Venezuela; GOM (Adults,  $n = 49$ ) off Veracruz, Mexico and from National Marine Fisheries Service pelagic longline survey stations throughout the GOM; western Central Atlantic (WCA,  $n = 55$ ) off the northern and central coasts of Brazil; WSA ( $n = 39$ ) off Santos, Brazil; and eastern North Atlantic (ENA,  $n = 33$ ) off Morocco (Figure 1). Samples were stored at room temperature in 95% ethanol or a 10% DMSO solution. Biological information including sex, length, and weight was available for a limited number of samples; spawning condition was unknown for most specimens. White marlin larvae sampled from the northern GOM (Larvae,  $n = 75$ ) between 2006 and 2013 were also evaluated; ichthyoplankton sampling methods are described in Rooker *et al.* (2012). Morphological species identification of larval white marlin was confirmed using the microsatellite genotype and mtDNA control region sequence data generated in this study, where larvae of species other than white marlin appeared as extreme outliers when compared with genotype and sequence data for adult white marlin.

Isolation of total genomic DNA was performed using standardized kits including the DNeasy Blood and Tissue Kit (Qiagen) and the ZR Genomic DNA Tissue MiniPrep Kit (Zymo Research). A subset of microsatellite loci previously developed for other istiophorid billfishes were optimized for use in white marlin (Buonaccorsi and Graves, 2000; Purcell *et al.*, 2009; Sorenson *et al.*, 2011; Bernard *et al.*, 2012; Williams *et al.*, 2015a). Twenty-four loci with consistent PCR amplification in white marlin (Supplementary Table S1) were developed into multiplex amplification reactions using forward primers modified at the 5' end



**Figure 1.** Geographic sampling locations for sample collections. Points reflect representative sampling location for each region. USM, United States mid-Atlantic; GOM, Gulf of Mexico; CAR, Caribbean; WCA, western Central Atlantic; WSA, western South Atlantic; ENA, eastern North Atlantic.

with locus-specific tag sequences and fluorescent dye labels. Amplifications were performed using the Type-It Microsatellite PCR Kit (Qiagen) following the manufacturer's protocol. 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 quality and a subset of samples (20%) were genotyped a second time to verify call consistency.

#### Generation of mtDNA sequence data

A segment of the mtDNA control region (~875 bp) was sequenced for a random subset of individuals from each sample collection and represented  $\geq 29\%$  of the total number of individuals in each collection (Table 1). PCR amplifications were performed using Pro-5 forward primer (Palumbi, 1996), an internal reverse primer (Graves and McDowell, 2006), and a PCR Core Reagents Kit (Qiagen). Amplification products were purified using the QIAquick PCR Purification Kit (Qiagen). Sequencing and consensus sequence construction were performed according to Portnoy *et al.* (2010), except with use of the MUSCLE multiple sequence alignment algorithm (Edgar, 2004). Control region sequence data previously generated for white marlin and reported in Graves and McDowell (2006) were downloaded from GenBank (accession numbers DQ835191–DQ835281) and included in the final alignment.

#### Analyses of microsatellite genotype data

Microsatellite genotype data were evaluated for the presence of scoring errors, null alleles, and large allele dropout using MicroChecker v2.2.3 (Van Oosterhout *et al.*, 2004, 2006). Loci were tested for selective neutrality using the  $F_{st}$  outlier detection method implemented in Lositan v1.0 (Antao *et al.*, 2008) using 500 000 simulations, a stepwise mutation model, and neutral and forced mean  $F_{st}$ . Conformance to the expectations of Hardy-Weinberg and linkage equilibria were evaluated for each locus using Arlequin v3.5 (10 000 iterations; Excoffier and Lischer, 2010) and Genepop v4.5 (10 000 iterations; Raymond and Rousset, 1995; Rousset, 2008), respectively. Allelic richness and the presence of private alleles were determined using the rarefaction methodology implemented in HP-RARE (version issued February 2009; Kalinowski, 2004, 2005); a sample size of 62 genes

was utilized to reflect the minimum number of genes observed in a sample collection. To account for the large difference in sample sizes across sample collections, mean expected heterozygosity corrected for uneven sample sizes among sample collections was calculated using GenClone v2.0 (Arnaud-Haond and Belkhir, 2007). Resampling was performed based on a corrected sample size of 33 individuals to reflect the smallest sample collection (ENA), and using 1000 permutations. Mean expected heterozygosity was also calculated as a ratio based on the comparison of the smallest sample collection (ENA) to every other sample collection using the R package *diveRsity* and 1000 bootstrap replicates (Keenan *et al.*, 2013). The standard error for each mean expected heterozygosity ratio was also calculated in *diveRsity*.

Genetic differentiation across all samples and pairwise differentiation between sample collections was assessed by calculating  $F_{st}$  values in Arlequin v3.5 (Excoffier and Lischer, 2010). Significance of  $F_{st}$  values was assessed based on 10 000 permutations of the data and a critical value corrected for multiple pairwise comparisons using a modified false discovery rate (Benjamini and Yekutieli, 2001; Narum, 2006). Bayesian model-based cluster analyses were performed in STRUCTURE v2.3.4 (Pritchard *et al.*, 2000) using 500 000 MCMC repetitions, 25 iterations of each K, and an admixture ancestry model (Falush *et al.*, 2007; Hubisz *et al.*, 2009). STRUCTURE simulations were performed with and without the use of sampling location as a prior, and results were compiled in STRUCTURE Harvester v0.6.94 (Evanno *et al.*, 2005; Earl and VonHoldt, 2012). The scaleGen function in the R package *adegenet* v1.4-2 was used to perform principal component analysis (PCA) with centered allele frequencies and missing data replaced with mean allele frequencies (Jombart, 2008; Jombart *et al.*, 2009, 2010).

Power analyses were executed in POWSIM v4.1 (Ryman and Palm, 2006) using empirical allele frequency data generated in this study and in Graves and McDowell (2006). Two scenarios based on the ability to detect the presence of two populations with a low ( $F_{st} = 0.005$ ) or high ( $F_{st} = 0.05$ ) level of genetic differentiation were evaluated; these levels were chosen based on the range of genetic differentiation reported in previous studies of genetic population structure in istiophorid billfishes. POWSIM simulates genetic drift in populations of effective size  $N_e$  over a specified number of generations  $t$  to produce a level of genetic differentiation described by  $F_{st} = 1 - (1 - 1/2N_e)^t$  (Nei, 1987). Many combinations of  $N_e$  and  $t$  will produce a particular  $F_{st}$  (Ryman and Palm, 2006). In this study,  $t$  was altered between scenarios of low and high genetic differentiation to produce the desired  $F_{st}$ . Both simulation scenarios were completed with 1000 replications, effective population sizes of 1500, and 15 (low scenario) or 155 (high scenario) generations of drift. The proportion of significant results (defined as  $p < 0.05$ ) was used to represent statistical power and was reported using Fisher's exact test (Ryman and Palm, 2006). The probability of committing a Type I statistical error was also evaluated in each simulation.

Temporal replicates consisting of  $\geq 10$  individuals sampled from the same geographic location in more than 1 year were available for some sampling locations (Supplementary Table S3). These replicates were used to evaluate the consistency of a statistically significant population structure signal over time within a geographic region. Pairwise  $F_{st}$  values were generated for temporal replicates within a sampling location using Arlequin. Significance of  $F_{st}$  values was assessed based on 10 000 permutations of the data and a critical value corrected for multiple

**Table 1.** Summary information for mtDNA control region sequence data.

Sample collection	No. sequences	No. haplotypes	No. polymorphic sites	<i>h</i>	$\pi$
WSA	17	17	134	1.00	0.036
WCA	40	39	166	0.99	0.031
USM	75	73	225	0.99	0.014
CAR	37	34	178	0.99	0.016
ENA	32	32	165	1.00	0.032
GOM Adults	44	41	185	0.99	0.032
GOM Larvae	31	31	176	1.00	0.032
Pooled Adults	245	216	347	0.99	0.031
All samples	276	242	354	0.99	0.031

*h*, haplotype diversity;  $\pi$ , nucleotide diversity.

pairwise comparisons (Benjamini and Yekutieli, 2001; Narum, 2006). The minimum number of samples required for a temporal replicate to be considered in this analysis ( $n = 10$ ) was chosen based on the minimum number of samples associated with an acceptable ( $\geq 95\%$  probability of detection) level of statistical power in power analyses performed using the high  $F_{st}$  scenario.

### Analyses of mtDNA sequence data

Arlequin was used to determine the number of haplotypes and associated measures of haplotype and nucleotide diversity. Arlequin was also used to evaluate genetic differentiation globally across all samples, and pairwise between sample collections. The statistical significance of global and pairwise levels of genetic differentiation was assessed with 10 000 permutations of the data, and by correcting for multiple pairwise comparisons using a modified false discovery rate method (Benjamini and Yekutieli, 2001; Narum, 2006). A median joining haplotype network was generated in PopArt v1.7 (Leigh and Bryant, 2015).

## Results

### Summary statistics and genetic diversity

A total of 554 individuals representative of six geographic regions and two demographic groups were genotyped at 24 microsatellite loci (Supplementary Table S1). There was no indication of null alleles or large allele dropout in the genotype data. Genotype calls for the subset of samples genotyped twice were consistent between runs. One microsatellite locus (Isin29) demonstrated statistically significant deviation from the expectations of Hardy-Weinberg equilibrium in the WSA sample collection ( $p = 0.004$ ). This deviation was associated with a deficiency of heterozygotes. As there was no pattern of deviation across multiple sample collections, Isin29 was retained for subsequent analyses. Observed and expected heterozygosity, including mean expected heterozygosity corrected for uneven sample sizes, were similar among sample collections (Supplementary Tables S1 and S2). No loci were identified as experiencing selection. The total number of alleles per locus ranged from 3 at Ta149 to 35 at Isin1 (mean  $a = 16.33$ ); mean allelic richness within sample collections ranged from 9.96 to 10.39 and the mean number of private alleles per sample collection ranged from 0.23 to 0.46. Allelic richness for GOM Larvae ( $a_R = 10.37$ ; Supplementary Table S1) did not differ from allelic richness in any other sample collection ( $a_R = 9.96 - 10.39$ ), nor from the pooled collection of all adult samples ( $a_R = 10.27$ ).

An 858 bp alignment of the mtDNA control region was analysed for a total of 276 individuals, including 185 newly

generated sequences (GenBank accession numbers K595230–K595414) and 91 sequences downloaded from GenBank (Table 1). Nucleotide composition consisted of 31.09% A, 29.37% T, 22.70% C, and 16.84% G. There were 354 polymorphic sites, 311 transitions, and 18 transversions. A total of 242 haplotypes were identified. Haplotype and nucleotide diversities across all samples were 0.99 and 0.031, respectively. Within sample collections, nucleotide diversity ranged from 0.014 to 0.036, and was lowest for USM (0.014) and CAR (0.016). Haplotype and nucleotide diversities in the larval collection ( $h = 1.00$ ,  $\pi = 0.032$ ; Table 1) were similar to that in all other sample collections ( $h = 0.99-1.00$ ,  $\pi = 0.014-0.036$ ) and to the pooled collection of all adult samples ( $h = 0.99$ ,  $\pi = 0.031$ ).

### Population structure inferred from genetic data

Overall genetic differentiation based on the microsatellite genotype data were not statistically significant (global  $F_{st} = -0.00009$ ,  $p = 0.525$ ).  $F_{st}$  values associated with the pairwise comparison of sample collections ranged from 0 to 0.003 and were not statistically significant, with the exception of the comparison between the larval and adult sample collections from the GOM ( $F_{st} = 0.003$ ,  $p = 0.010$ ; Table 2). Results from STRUCTURE did not indicate the presence of more than one genetic group, and iterations with  $K = 1$  were associated with the highest log likelihood values (Figure 2). STRUCTURE results were similar between admixture models that did or did not utilize sampling information as a prior to inform clustering. PCA also demonstrated a single grouping of individuals (Supplementary Figure S1). The eigenvalues associated with this analysis were similar across principal components.

Overall genetic differentiation based on the mtDNA sequence data were not statistically significant (global  $\phi_{st} = -0.00190$ ,  $p = 0.614$ ). Levels of genetic differentiation associated with pairwise comparisons of sample collections were not statistically significant (Table 2). The high haplotype diversity of the mtDNA control region was reflected in a midpoint spanning haplotype network (Figure 3). The three clusters of haplotypes apparent in this network were separated by 12 and 17 mutational differences.

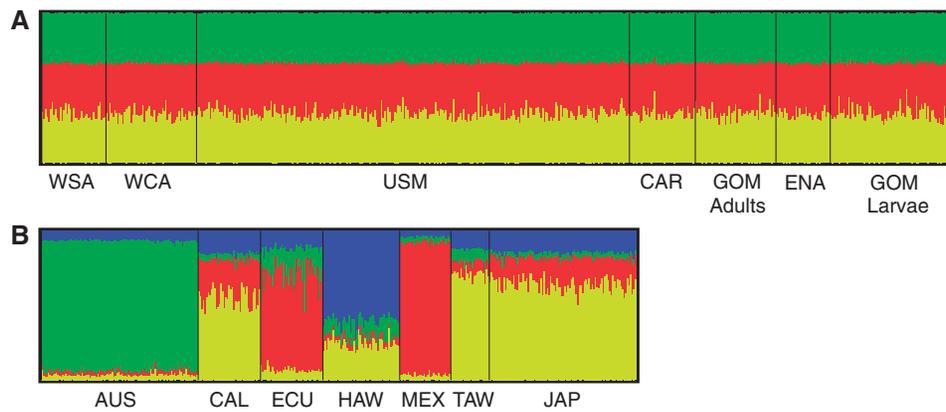
### Statistical power

Results from power simulations based on empirical allele frequency data generated in this study and in Graves and McDowell (2006) indicate that when the level of genetic differentiation between populations is relatively high ( $F_{st} = 0.05$ ), an acceptable ( $\geq 95\%$  probability of detection) level of statistical power is possible with low sample sizes ( $n \geq 10$ ) per population and either a small ( $n = 5$ ) or large ( $n = 24$ ) number of microsatellite markers

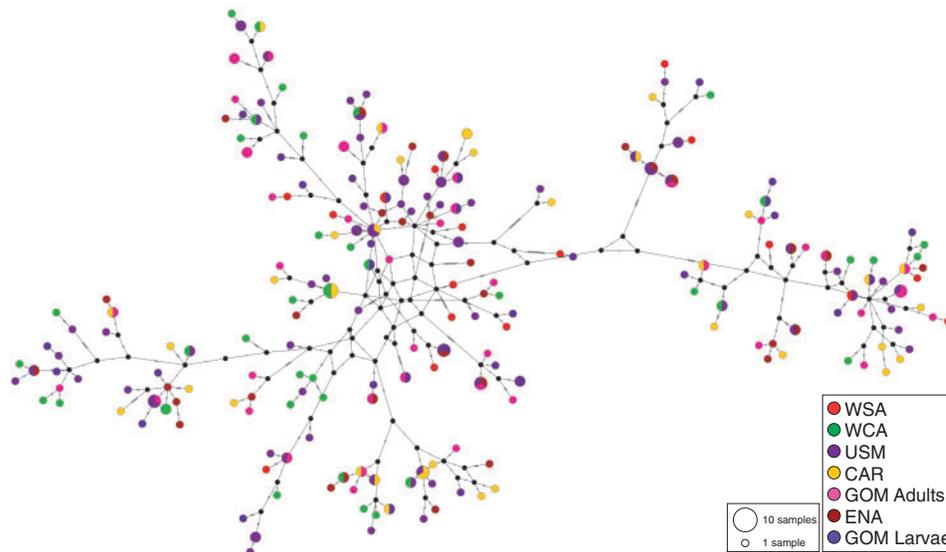
**Table 2.** Matrix of pairwise  $F_{st}$  values based on microsatellite data (below diagonal) and pairwise  $\phi_{st}$  values based on mtDNA sequence data (above diagonal).

	WSA	WCA	USM	CAR	ENA	GOM Adults	GOM Larvae
WSA	–	0.01796 (0.126)	0.01175 (0.167)	–0.00206 (0.411)	–0.00170 (0.407)	–0.00139 (0.398)	–0.01356 (0.760)
WCA	–0.00015 (0.447)	–	–0.00142 (0.471)	0.00561 (0.221)	–0.00386 (0.544)	–0.00445 (0.578)	–0.00912 (0.800)
USM	0.00006 (0.444)	–0.00125 (0.964)	–	0.01025 (0.109)	–0.00676 (0.769)	0.00285 (0.253)	–0.00421 (0.613)
CAR	–0.00118 (0.733)	–0.00066 (0.607)	–0.00053 (0.690)	–	–0.01313 (0.923)	–0.00283 (0.474)	–0.00704 (0.651)
ENA	0.00008 (0.440)	–0.00160 (0.783)	–0.00049 (0.643)	0.00018 (0.448)	–	–0.00542 (0.596)	–0.01261 (0.904)
GOM Adults	0.00253 (0.059)	–0.00025 (0.478)	0.00128 (0.069)	0.00283 (0.041)	0.00134 (0.209)	–	–0.01496 (0.987)
GOM Larvae	–0.00092 (0.688)	–0.00166 (0.915)	–0.00035 (0.698)	–0.00054 (0.601)	0.00102 (0.222)	0.00308 (0.010)	–

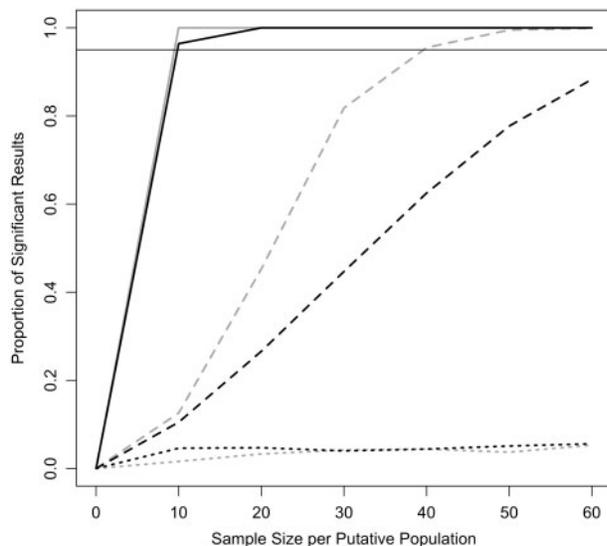
Significance values are shown in parentheses, statistically significant comparisons are italicized. A critical value corrected for multiple pairwise comparisons and equal to 0.014 was used.



**Figure 2.** STRUCTURE results based on microsatellite genotype data generated for 24 loci in white marlin in this study (A; plot shown for  $K = 3$ ) and for 5 loci in striped marlin in *McDowell and Graves (2008)* (B; plot shown for  $K = 4$ ). STRUCTURE results were similar with or without the use of sampling location as a prior to inform clustering. Individuals are organized by sample collection for both species. Sample collection labels for striped marlin are as follows: AUS, Port Stephens, Australia; CAL, San Diego, California; ECU, Manta, Ecuador; HAW, Kona, Hawaii; MEX, Cabo San Lucas, Mexico; TAW, Taiwan; and JAP, Japan.



**Figure 3.** Median joining haplotype network generated from mtDNA control region sequence data. Ancestral nodes inferred from observed data are shown in black. The number of mutational differences between nodes are represented by hatch marks along edges. Nodes are coloured according to sample collection, and node size corresponds to the number of occurrences for a particular haplotype.



**Figure 4.** Results from power analyses based on empirical allele frequency data generated in this study (gray lines) and in Graves and McDowell (2006; black lines). Solid lines reflect scenario with high level of genetic differentiation between simulated populations ( $F_{st} = 0.05$ ), dashed lines reflect low level of genetic differentiation between simulated populations ( $F_{st} = 0.005$ ). Dotted lines reflect alpha level of statistical error, and horizontal line represents 0.95 probability of detection.

(Figure 4). When the level of genetic differentiation between populations is low ( $F_{st} = 0.005$ ), considerably larger numbers of samples and microsatellite markers are necessary to provide a suitable level of statistical power. In simulations where  $F_{st} = 0.05$ , both the present study and Graves and McDowell (2006) displayed similarly high levels of statistical power; however, the five microsatellite loci surveyed in the 2006 study did not facilitate a  $\geq 95\%$  probability of detection at any of the sample sizes explored with simulations when  $F_{st} = 0.005$ . In comparison, the 24 microsatellite loci surveyed in the present study provided an acceptable level of statistical power with sample sizes  $\geq 40$  individuals per putative population in the low differentiation scenario. Two sample collections in this study included  $< 40$  individuals: WSA ( $n = 39$ ) and ENA ( $n = 33$ ). The Type I error rate remained low for all simulation scenarios based on both empirical datasets.

### Temporal replicates

Pairwise  $F_{st}$  values were not statistically significant for the two temporal replicates available for the WCA and for the GOM Adults sample collections, and were also not significant for the three temporal replicates available for the GOM Larvae sample collection (results not shown). Fourteen temporal replicates with sample sizes  $\geq 10$  were available for the USM sample collection (Supplementary Table S3); sample sizes for these replicates ranged from 10 to 42 individuals per year. Statistically significant levels of genetic differentiation were observed for five pairwise comparisons between USM temporal replicates and were consistently associated with replicates from the year 1998 or 2012 (Supplementary Table S4). Sample sizes for the 1998 and 2012 temporal replicates were 11 and 14, respectively. All microsatellite loci conformed to the assumptions of Hardy-Weinberg

equilibrium within all USM temporal replicates, with the exception of MnE in the 1995 replicate ( $p = 0.008$ ) and MnI in the 1998 replicate ( $p = 0.011$ ).

### Discussion

The objective of this study was to evaluate genetic population structure in white marlin using a higher level of statistical power compared with previous studies. Increased statistical power was achieved by surveying a larger number of molecular markers across greater numbers of samples, including sample collections from additional geographic locations and a collection of larvae. With the increase in statistical power, levels of genetic heterogeneity in the present study decreased compared with those observed in previous studies with lower power. Graves and McDowell (2006) reported global levels of genetic differentiation that approached statistical significance ( $F_{st} = 0.0022$ ,  $p = 0.057$ ;  $\phi_{st} = 0.0163$ ,  $p = 0.069$ ); however, in this study global levels of differentiation were considerably reduced ( $F_{st} = -0.00009$ ,  $p = 0.525$ ;  $\phi_{st} = -0.00190$ ,  $p = 0.614$ ). The null hypothesis of a single genetic stock of white marlin could not be rejected based on the lack of genetic heterogeneity observed in this study.

This study represents the first population genetic analysis of an istiophorid billfish species to include a collection of larvae. Ideally, the inclusion of two or more collections of tissue from spawning adults or larvae sampled from geographically distinct spawning grounds would enable the direct evaluation of genetic differentiation between putative source populations of white marlin. However, confirmation of spawning status in adult fish requires direct observation of the release of gametes or the histological examination of gonadal tissue. Although the former approach may be possible for actively spawning females that may release hydrated oocytes when external pressure is applied to the ovaries, the latter approach requires the sacrifice of individual fish in order to inspect gonadal tissue. White marlin is an overfished species and the collection of fin clips as in this study prevents the need to sacrifice fish for the collection of tissues for genetic analysis. The collection of larvae from istiophorid billfishes is also challenging due to limited information on spawning in these species, and due to the logistics of obtaining robust sample sizes. In this study, sample collections of adult fish were not known to include actively spawning individuals, and only a single larval collection was available for genetic analysis.

Nevertheless, evaluation of a single collection of larvae can be informative of population structure when compared with collections of adults from throughout the species' range. In the case of genetically discrete populations, a larval collection would be expected to display a statistically significant level of genetic differentiation when compared with collections of adults representative of another source population. In this study, there were no statistically significant levels of genetic differentiation between sample collections based on either the microsatellite or the mtDNA sequence data, with the exception of the microsatellite-based comparison of GOM Larvae and GOM Adults. The statistically significant heterogeneity observed between the GOM sample collections was driven by two loci (Ta149 and Isin29; Supplementary Figure S2) and resulted in an  $F_{st}$  for which the statistical significance was just below the critical value ( $p = 0.010$ ,  $p_{crit} = 0.014$ ). This result was likely due to random sampling error and is not likely to be biologically meaningful (Waples, 1998). The lack of support for more than one genetic cluster in the STRUCTURE results is also consistent with the observed lack of genetic

differentiation among sample collections, although the ability of STRUCTURE to estimate the true number of populations is reduced when genetic differentiation between populations is low (Waples and Gaggiotti, 2006). In addition, observed levels of genetic differentiation between sample collections based on the mtDNA control region sequence data may be influenced by homoplasy at this gene region (Reeb *et al.*, 2010; Bradman *et al.*, 2011); however, support for population subdivision based on the nuclear genotype data is lacking.

If genetically discrete populations exist, a larval collection may also display lower genetic diversity relative to that of a pooled collection of adult samples. Additionally, genetic diversity in the larval collection may also be lower than non-pooled geographically distant sample collections if those collections comprised mixed stocks. In this study, levels of genetic diversity based on the microsatellite and the mtDNA sequence data were similar between the GOM larval collection and the pooled collection of all adult samples. Levels of genetic diversity based on both marker types were also similar between the larval collection and all individual (non-pooled) sample collections. Collectively, the absence of genetic differentiation and similar levels of genetic diversity among sample collections, particularly between the larval collection and the pooled collection of all adult samples, do not provide sufficient evidence to reject the null hypothesis of a single genetic stock of white marlin in the Atlantic Ocean.

The lack of statistically significant genetic heterogeneity observed for white marlin in this study differs from results reported by Graves and McDowell (2006). The increase in statistical power associated with this study resulted in decreased levels of global genetic differentiation compared with the earlier study, and reduced genetic differentiation between sample collections from the WSA and the USM. Power analysis simulations based on empirical allele frequency data indicate that the number of loci and samples analyzed in the present study facilitated relatively high statistical power even at a low level of genetic differentiation between simulated populations. For example, with sample sizes of 35 individuals and an  $F_{st}$  of 0.005, this study was associated with a nearly 90% probability of detecting genetic differentiation. In comparison, power analysis simulations based on empirical allele frequency data from the five microsatellite markers surveyed by Graves and McDowell (2006) indicate a <60% probability of detecting genetic differentiation with sample sizes of 35 individuals and an  $F_{st}$  of 0.005. For istiophorid billfishes in which the analysis of microsatellite markers has previously revealed population subdivision, statistically significant  $F_{st}$  values ranged from 0.007 to 0.047 (McDowell and Graves, 2008; Purcell and Edmands, 2011; Lu *et al.*, 2015; Williams *et al.*, 2015b). This suggests that statistical power associated with this study was high enough to facilitate the detection of genetic population subdivision in white marlin, if it exists. Three of the loci surveyed by Graves and McDowell (2006) were not included in this study, and additional analysis of the 2006 data indicates that the genetic heterogeneity observed in that study was not driven by one locus in particular. The lack of genetic heterogeneity observed in this study compared with Graves and McDowell (2006) is likely due to the large difference in statistical power resulting from the limited number of microsatellite markers surveyed in the earlier study.

The temporal replicates evaluated in this study also provide useful insights into the importance of statistical power in interpreting the results of population genetic studies. Collectively,

results of this study lack any evidence to suggest the presence of more than one genetic stock of white marlin; however, a number of statistically significant pairwise comparisons were observed between temporal replicates from the USM sample collection. These statistically significant comparisons are presumed to reflect random noise rather than a true population structuring signal. The USM temporal replicates consisted of sample sizes ranging from 10 to 42 individuals per replicate; however, only replicates of very small sample sizes (1998,  $n = 11$ ; 2012,  $n = 14$ ) were associated with statistically significant pairwise comparisons. These results suggest that when statistical power is low due to small sample sizes and/or limited numbers of molecular markers, the ability to distinguish between noise and a true population structuring signal is diminished, resulting in a higher probability of detecting a false signal and committing a Type I statistical error (Waples, 1998). Conversely, low statistical power may also result in failure to detect a low but biologically meaningful population structuring signal and lead to a Type II statistical error.

The apparent lack of genetic population structure in white marlin has also been observed for other istiophorid billfishes. Previous evaluations of genetic population structure for blue marlin (*Makaira nigricans*) in the Atlantic Ocean have failed to detect population subdivision despite the analysis of multiple types of nuclear and mt markers (Buonaccorsi *et al.*, 1999, 2001; McDowell *et al.*, 2007). Genetic population subdivision was also not detected for sailfish (*Istiophorus platypterus*) in the Atlantic Ocean based on the analysis of mtDNA control region sequence data and a small number ( $n = 5$ ) of microsatellite loci (McDowell, 2002; McDowell and Graves, 2002). Analysis of mtDNA control region sequence data for blue marlin and sailfish has demonstrated the presence of two distinct mt lineages for both species in the Atlantic Ocean (McDowell and Graves, 2002; McDowell *et al.*, 2007). The existence of these lineages is attributed to historical isolation between Atlantic and Indo-Pacific populations of these species, followed by subsequent re-introduction of Indo-Pacific fish to the Atlantic Ocean (Graves and McDowell, 1995). In this study, haplotype diversity at the mtDNA control region is high, but distinct lineages are not present in white marlin. This result is consistent with a lack of historical isolation in white marlin compared with blue marlin and sailfish, perhaps due to differences in thermal tolerance and/or the spatial distribution of these species.

Despite the apparent lack of genetic population structure for some istiophorid billfish species in the Atlantic Ocean, genetic heterogeneity has been demonstrated for other highly migratory marine fishes in this ocean. Analyses of mtDNA control region sequence data and a number of nuclear loci, including microsatellite markers, revealed distinct populations of swordfish (*Xiphias gladius*) in the North Atlantic and South Atlantic (Chow and Takeyama, 2000; Alvarado-Bremer *et al.*, 2005; Kasapidis *et al.*, 2006). These populations correspond with the presence of distinct spawning grounds in the northern and southern hemispheres (Alvarado-Bremer *et al.*, 2005). In the Pacific Ocean, genetic population structure has been reported for sailfish and black marlin (*Istiompax indica*; McDowell, 2002; Lu *et al.*, 2015; Williams *et al.*, 2015b). In addition, the presence of at least four stocks has been genetically determined for striped marlin (*Kajikia audax*), the closely related sister species of white marlin, in the Pacific Ocean (McDowell and Graves, 2008; Purcell and Edmands, 2011). STRUCTURE analysis of genotype data previously generated for striped marlin (McDowell and Graves, 2008) and based

on only five microsatellite loci clearly demonstrates the presence of four genetic clusters (Figure 2). Although Figure 2 provides an interesting comparison of the level of genetic subdivision observed for congeneric species that inhabit different ocean basins, the lack of populations resolved by STRUCTURE for white marlin may only reflect the limitations of this analysis in resolving subtle population structure in this species.

The genetic homogeneity observed for white marlin in this study may be consistent with what we currently know about the biology of this species. The distribution of catches based on fisheries-dependent data suggests a continuous species distribution in the Atlantic Ocean (ICCAT, 2015). Movements of white marlin inferred from the tagging of individual fish include a small number of conventional tag recaptures representing trans-oceanic ( $n = 7$ ) or trans-equatorial ( $n = 3$ ) dispersal (Snodgrass *et al.*, 2013). Information from electronic tags reflects highly inconsistent directionality of movements by individuals tagged from seasonal assemblages of white marlin (Loose, 2014; Schlenker, 2014). In addition, seasonally displaced spawning grounds have been confirmed in the northern and southern hemispheres, including locations in the western North Atlantic, GOM, and Caribbean Sea for which spawning appears to primarily occur in the second quarter of the year (de Sylva and Breder, 1997; Luthy *et al.*, 2005; Prince *et al.*, 2005; Arocha and Barrios, 2009; Richardson *et al.*, 2010; Rooker *et al.*, 2012), and in the WSA where spawning primarily occurs in the fourth quarter (Ueyanagi *et al.*, 1970; Arfelli *et al.*, 1986; Amorim and Arfelli, 2003; Schmidt *et al.*, 2015). Because the spawning season in these regions is not concurrent, it is possible that white marlin may spawn in multiple locations within a year, facilitating gene flow among geographically distant regions and resulting in the lack of genetic subdivision observed in this study.

The inability to reject the null hypothesis of a single genetic stock of white marlin in this study does not necessarily mean that ecological stocks of this species do not exist. Results of this study suggest that ocean-wide connectivity in white marlin is sufficient to prohibit the detection of genetic heterogeneity with the molecular markers surveyed in this study, but the magnitude of intraspecific connectivity beyond this threshold is unknown and could reflect dispersal in a small to large proportion of individuals. A 365-day geolocation track from Loose (2014) provides evidence that some white marlin complete cyclical migrations to seasonal foraging and spawning grounds on an annual time scale. Rooker *et al.* (2013, Supplementary Figure S5) also report cyclical movement for white marlin in the GOM based on a year-long geolocation track. Conventional tag recapture data also suggest annual periodicity associated with feeding and spawning assemblages of white marlin (Ortiz *et al.*, 2003). If annual migrations are associated with some level of fidelity to discrete spawning grounds, this could result in the presence of ecological stocks; however, whether white marlin display spawning site fidelity is currently unknown. Atlantic-wide electronic tagging efforts that incorporate multi-year deployment periods would provide an informative approach for inferring ocean-wide connectivity in white marlin; however, tag deployment periods of this duration have not been previously successful.

Recently developed next-generation sequencing technology now facilitates the discovery of thousands of molecular markers representative of the whole genome, providing an unprecedented ability to detect genetic population structure (Davey and Blaxter, 2010; Narum *et al.*, 2013). For example, application of this

methodology to yellowfin tuna (*Thunnus albacares*; Grewe *et al.*, 2015; Pecoraro *et al.*, 2016) revealed inter- and intra-oceanic genetic heterogeneity previously unresolved by molecular markers such as allozymes, mtDNA, and microsatellites (Ward *et al.*, 1994; Diaz-James and Uribe-Alcocer, 2006). Elucidation of unresolved or cryptic genetic heterogeneity can have significant consequences for fisheries management, particularly for stocks currently threatened by unsustainable fishing practices. The evaluation of genome-wide molecular markers may significantly improve the ability to detect genetic population structure, especially for weakly differentiated populations of marine species. In addition, genome-wide methodologies facilitate the identification of putative adaptive loci, which may provide insight on localized adaptation and regional genetic diversity, even if the level of intraspecific connectivity is enough to obscure heterogeneity at neutral markers (Allendorf *et al.*, 2008; Andrews *et al.*, 2016).

The ability to detect genetic population structure in white marlin was considerably improved in this study by increasing the number of molecular markers surveyed across a larger number of samples, including collections from additional geographic locations and larvae from the GOM. This increase in statistical power resulted in lower levels of genetic heterogeneity relative to those reported in a previous study that utilized fewer molecular markers and lower numbers of opportunistically collected samples. These results highlight the importance of statistical power in population genetic assessments of species for which population subdivision, if present, is expected to be shallow. Results from this study are consistent with the presence of a single genetic stock of white marlin in the Atlantic Ocean, and with the single-stock assessment and management model currently utilized by ICCAT. However, these inferences are based on molecular markers that may be limited in their ability to resolve low levels of genetic differentiation, and the degree of Atlantic-wide connectivity beyond that required to mask heterogeneity using the molecular markers surveyed in this study is unknown. It is especially important that Atlantic-wide connectivity in white marlin is understood considering the most recent stock assessment identified this species as overfished, with biomass in 2010 at half of that necessary to produce maximum sustainable yield (ICCAT, 2012). Additional investigation using genome-wide molecular markers and collections of larvae and/or reproductively active adults from additional spawning locations to inform inferences on ocean-wide genetic connectivity, as well as studies to determine the biological significance of the genetic results, are warranted.

### Supplementary data

Supplementary materials are available at the ICESJMS online version of the article.

### Acknowledgements

We thank a number of recreational anglers throughout the Atlantic region for providing samples for this study. We also thank the Mid-Atlantic Tournament for facilitating the opportunity to collect samples over a number of years.

### Funding

This study was supported by a grant from NOAA-NMFS and is Virginia Institute of Marine Science, College of William and Mary Contribution No. 3627.

## References

- Allendorf, F., England, P., Luikart, G., Ritchie, P., and Ryman, N. 2008. Genetic effects of harvest on wild animal populations. *Trends in Ecology and Evolution*, 23: 327–337.
- Allendorf, F., and Phelps, S. 1981. Use of allelic frequencies to describe population structure. *Canadian Journal of Fisheries and Aquatic Sciences*, 38: 1507–1514.
- Allendorf, F., Ryman, N., and Utter, F. 1987. Genetics and fishery management: past, present, and future. In *Population Genetics and Fishery Management*, pp. 1–19. Ed. by N. Ryman and F. Utter, University of Washington Press, Seattle, WA. 440 pp.
- Alvarado Bremer, J., Vinas, J., Mejuto, J., Ely, B., and Pla, C. 2005. Comparative phylogeography of Atlantic bluefin tuna and swordfish: the combined effects of vicariance, secondary contact, introgression, and population expansion on the regional phylogenies of two highly migratory pelagic fishes. *Molecular Phylogenetics and Evolution*, 36: 169–187.
- Amorim, A., and Arfelli, C. 2003. Review of white marlin (*Tetrapturus albidus*) fishery biology off the southern Brazilian coast (1971–2001). ICCAT Collective Volume of Scientific Papers, 55: 467–474.
- Andrews, K., Good, J., Miller, M., Luikart, G., and Hohenlohe, P. 2016. Harnessing the power of RADseq for ecological and evolutionary genetics. *Nature Reviews Genetics*, 17: 81–92.
- Antao, T., Lopes, A., Lopes, R., Beja-Pereira, A., and Luikart, G. 2008. LOSITAN: a workbench to detect molecular adaptation based on a *Fst*-outlier method. *BMC Bioinformatics*, 9: 323.
- Arfelli, C., Amorim, A., and Galhardo-Amado, J. 1986. Analysis on *Tetrapturus albidus* Poey (1861) caught off south and southeast of Brazil (1971–1984). ICCAT Collective Volume of Scientific Papers, 25: 202–217.
- Arnaud-Haond, S., and Belkhir, K. 2007. Genclone: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes*, 7: 15–17.
- Arocha, F., and Barrios, A. 2009. Sex ratios, spawning seasonality, sexual maturity, and fecundity of white marlin (*Tetrapturus albidus*) from the western central Atlantic. *Fisheries Research*, 95: 98–111.
- Benjamini, Y., and Yekutieli, D. 2001. The control of false discovery rate under dependency. *Annals of Statistics*, 29: 1165–1188.
- Bernard, A., Feldheim, K., and 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. *Journal of Fish Biology*, 81: 1781–1786.
- Boustany, A., Reeb, C., and Block, B. 2008. Mitochondrial DNA and electronic tracking reveal population structure of Atlantic bluefin tuna (*Thunnus thynnus*). *Marine Biology*, 156: 13–24.
- Bowen, B., Bass, A., Soares, L., and Toonen, R. 2005. Conservation implications of complex population structure: lessons from the loggerhead turtle (*Caretta caretta*). *Molecular Ecology*, 14: 2389–2402.
- Bradman, H., Grewe, P., and Appleton, B. 2011. Direct comparison of mitochondrial markers for the analysis of swordfish population structure. *Fisheries Research*, 109: 95–99.
- Buonaccorsi, V., and Graves, J. 2000. Isolation and characterization of novel polymorphic tetra-nucleotide microsatellite markers from the blue marlin, *Makaira nigricans*. *Molecular Ecology*, 9: 820–821.
- Buonaccorsi, V., McDowell, J., and Graves, J. 2001. Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). *Molecular Ecology*, 10: 1179–1196.
- Buonaccorsi, V., Reece, K., Morgan, L., and Graves, J. 1999. Geographic distribution of molecular variance within the blue marlin (*Makaira nigricans*): a hierarchical analysis of allozyme, single-copy nuclear DNA, and mitochondrial DNA markers. *Evolution*, 53: 568–579.
- Carlsson, J., McDowell, J., Carlsson, J., and Graves, J. 2007. Genetic identity of YOY bluefin tuna from the eastern and western Atlantic spawning areas. *The Journal of Heredity*, 98: 23–28.
- Chow, S., and Takeyama, H. 2000. Nuclear and mitochondrial DNA analyses reveal four genetically separated breeding units of swordfish. *Journal of Fish Biology*, 56: 1087–1098.
- Davey, J., and Blaxter, M. 2010. RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, 9: 416–423.
- de Sylva, D., and Breder, P. 1997. Reproduction, gonad histology, and spawning cycles of North Atlantic billfishes (Istiophoridae). *Bulletin of Marine Science*, 60: 668–697.
- Diaz-Jaimes, P., and Uribe-Alcocer, M. 2006. Spatial differentiation in the eastern Pacific yellowfin tuna revealed by microsatellite variation. *Fisheries Science*, 72: 590–596.
- Earl, D., and VonHoldt, B. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4: 359–361.
- Edgar, R. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32: 1792–1797.
- Evanno, G., Regnaut, S., and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14: 2611–2620.
- Excoffier, L., and Lischer, H. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10: 564–567.
- Falush, D., Stephens, M., and Pritchard, J. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, 7: 574–578.
- Fromentin, J. –M., and Powers, J. E. 2005. Atlantic bluefin tuna: population dynamics, ecology, fisheries and management. *Fish and Fisheries*, 6: 281–306.
- Graves, J., Gold, J., Ely, B., Quattro, J., Woodley, C., and Dean, J. 1996. Population genetic structure of bluefin tuna in the North Atlantic Ocean. Identification of variable genetic markers. ICCAT Collective Volume of Scientific Papers, 45: 155–157.
- Graves, J., and McDowell, J. 1995. Inter-ocean genetic divergence of istiophorid billfishes. *Marine Biology*, 122: 193–203.
- Graves, J., and McDowell, J. 2001. A genetic perspective on the stock structures of blue marlin and white marlin in the Atlantic Ocean. ICCAT Collective Volume of Scientific Papers, 53: 180–187.
- Graves, J., and McDowell, J. 2006. Genetic analysis of white marlin (*Tetrapturus albidus*) stock structure. *Bulletin of Marine Science*, 79: 469–482.
- Graves, J., and McDowell, J. 2015. Population structure of istiophorid billfishes. *Fisheries Research*, 166: 21–28.
- Grewe, P., Feutry, P., Hill, P., Gunasekera, R., Schaefer, K., Itano, D., Fuller, D., Foster, S., and Davies, C. 2015. Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource. *Scientific Reports*, 5: 16916.
- Hedgecock, D. 1994. Does variance in reproductive success limit effective population sizes of marine organisms? In *Genetics and Evolution of Aquatic Organisms*, pp. 122–134. Ed. By A. Beaumont. Chapman and Hall, London.
- Hubisz, M., Falush, D., Stephens, M., and Pritchard, J. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9: 1322–1332.
- ICCAT. 1994. Report of the second ICCAT billfish workshop. ICCAT Collective Volume of Scientific Papers, 41: 13–166.
- ICCAT. 2012. Report of the 2012 white marlin stock assessment meeting. ICCAT Collective Volume of Scientific Papers, 55: 1–69.
- ICCAT. 2015. Statistical Bulletin: Volume 42(2) (1960–2013). Madrid, Spain.

- Jombart, T. 2008. ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24: 1403–1405.
- Jombart, T., Devillard, S., and Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structure populations. *BMC Genetics*, 11: 1471–2156.
- Jombart, T., Pontier, D., and Dufour, A. –B. 2009. Genetic markers in the playground of multivariate analysis. *Heredity*, 102: 330–341.
- Kalinowski, S. 2004. Counting alleles with rarefaction: private alleles and hierarchical sampling designs. *Conservation Genetics*, 5: 539–543.
- Kalinowski, S. 2005. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, 5: 187–189.
- Kasapidis, P., Mejuto, J., Tserpes, G., Antoniou, A., Garcia-Cortes, B., Peristeraki, P., Oikonomaki, K., Kotoulas, G., and Magoulas, A. 2006. Genetic structure of the swordfish (*Xiphias gladius*) stocks in the Atlantic using microsatellite DNA analysis. ICCAT Collective Volume of Scientific Papers, 61: 89–98.
- Keenan, K., McGinnity, P., Cross, T., Crozier, W., and Prodohl, P. 2013. diveRsiTY: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4: 782–788.
- Leigh, J., and Bryant, D. 2015. POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6: 1110–1116.
- Loose, E. 2014. Seasonal movements, habitat utilization, and comparative scale morphology of whitemarlin (*Kajikia albida*) and roundscale spearfish (*Tetrapturus georgii*). Unpublished master's thesis, Virginia Institute of Marine Science, Gloucester Point, VA.
- Lu, C. –P., Alvarado Bremer, J., McKenzie, J., and Chiang, W. –C. 2015. Analysis of sailfish (*Istiophorus platypterus*) population structure in the North Pacific Ocean. *Fisheries Research*, 166: 33–38.
- Luthy, S., Cowen, R., Serafy, J., and McDowell, J. 2005. Toward identification of larval sailfish (*Istiophorus platypterus*), white marlin (*Tetrapturus albidus*), and blue marlin (*Makaira nigricans*) in the western North Atlantic Ocean. *Fisheries Bulletin*, 103: 588–600.
- McDowell, J. 2002. Genetic stock structure of the sailfish, *Istiophorus platypterus*, based on nuclear and mitochondrial DNA. Doctoral dissertation, Virginia Institute of Marine Science, College of William & Mary.
- McDowell, J., Carlsson, J., and Graves, J. 2007. Genetic analysis of blue marlin (*Makaira nigricans*) stock structure in the Atlantic Ocean. *Gulf and Caribbean Research*, 19: 75–82.
- McDowell, J., and Graves, J. 2002. A genetic perspective on Atlantic sailfish stock structure. ICCAT Collective Volume of Scientific Papers, 54: 805–810.
- McDowell, J., and Graves, J. 2008. Population structure of striped marlin (*Kajikia audax*) in the Pacific Ocean based on analysis of microsatellite and mitochondrial DNA. *Canadian Journal of Fisheries and Aquatic Sciences*, 65: 1307–1320.
- Nakamura, I. 1985. Billfishes of the World: An Annotated and Illustrated Catalogue of Marlins, Sailfishes, Spearfishes, and Swordfishes Known to Date. Rome, Italy: Food and Agricultural Organization.
- Narum, S. 2006. Beyond Bonferroni: less conservative analyses for conservation genetics. *Conservation Genetics*, 7: 783–787.
- Narum, S., Buerkle, C., Davey, J., Miller, M., and Hohenlohe, P. 2013. Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, 22: 2841–2847.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Ortiz, M., Prince, E., Serafy, J., Holts, D., Davy, K., Pepperell, J., Lowry, M., and Holdsworth, J. C. 2003. Global overview of the major constituent-based billfish tagging programs and their results since 1954. *Marine and Freshwater Research*, 54: 489–507.
- Ovenden, J., Berry, O., Welch, D., Buckworth, R., and Dichmont, C. 2015. Ocean's eleven: a critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. *Fish and Fisheries*, 16: 125–159.
- Palumbi, S. 1996. Nucleic acids II. the polymerase chain reaction. In *Molecular systematics*, 2nd edn, pp. 205–247. Ed. by D. Hillis, C. Moritz and B. Mable. Sinauer Associates, Sunderland, MA. 655 pp.
- Pecoraro, C., Babbucci, M., Villamor, A., Franch, R., Papetti, C., Leroy, B., Ortega-Garcia, S. et al. 2016. Methodological assessment of 2b-RAD genotyping technique for population structure inferences in yellowfin tuna (*Thunnus albacares*). *Marine Genomics*, 25: 43–48.
- Portnoy, D., McDowell, J., Heist, E., Musick, J., and Graves, J. 2010. World phylogeography and male-mediated gene flow in the sandbar shark, *Carcharhinus plumbeus*. *Molecular Ecology*, 19: 1994.
- Prince, E., Cowen, R., Orbesen, E., Luthy, S., Llopiz, J., Richardson, D., and Serafy, J. 2005. Movements and spawning of white marlin (*Tetrapturus albidus*) and blue marlin (*Makaira nigricans*) off Punta Cana, Dominican Republic. *Fisheries Bulletin*, 103: 659–669.
- Pritchard, J., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155: 945–959.
- Purcell, C., and 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. *Canadian Journal of Fisheries and Aquatic Sciences*, 68: 1861–1875.
- Purcell, C., Harrison, J., and Edmands, S. 2009. Isolation and characterization of 10 polymorphic microsatellite markers from striped marlin, *Tetrapturus audax*. *Molecular Ecology Resources*, 9: 1556–1559.
- Raymond, M., and Rousset, F. 1995. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86: 248–249.
- Reeb, C., Arcangeli, L., and Block, B. 2000. Structure and migration corridors in Pacific populations of the swordfish *Xiphias gladius*, as inferred through analyses of mitochondrial DNA. *Marine Biology*, 136: 1123–1131.
- Richardson, D., Llopiz, J., Guigand, C., and Cowen, R. 2010. Larval assemblages of large and medium-sized pelagic species in the Straits of Florida. *Progress in Oceanography*, 86: 8–20.
- Rooker, J., Alvarado-Bremer, J., Block, B., Dewar, H., De Metrio, G., Corriero, A., Kraus, R. et al. 2007. Life history and stock structure of Atlantic bluefin tuna (*Thunnus thynnus*). *Reviews in Fisheries Science*, 15: 265–310.
- Rooker, J., Kitchens, L., Dance, M., Wells, R., Falterman, B., and Cornic, M. 2013. Spatial, temporal, and habitat-related variation in abundance of pelagic fishes in the Gulf of Mexico: potential implications of the Deepwater Horizon oil spill. *PLoS One*, 8: e76080.
- Rooker, J., Simms, J., Wells, R., Holt, S., Holt, G., Graves, J., and Furey, N. 2012. Distribution and habitat associations of billfish and swordfish larvae across mesoscale features in the Gulf of Mexico. *PLoS One*, 7: e34180.
- Rousset, F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8: 103–106.
- Ryman, N., and Palm, S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes*, 6: 600–602.
- Ryman, N., Palm, S., Andre, C., Carvalho, G., Dahlgren, T., Jorde, P., Laikre, L., Larsson, L., Palme, A., and Ruzzante, D. 2006. Power for detecting genetic divergence: differences between statistical methods and marker loci. *Molecular Ecology*, 15: 2031–2045.
- Schlenker, L. 2014. Physiological stress and post-release mortality of white marlin (*Kajikia albida*) caught in the U.S. recreational

- fishery. Master's thesis, Virginia Institute of Marine Science, College of William & Mary.
- Schmidt, R., Rodrigues, T., Pimenta, E., Hilsdorf, A., and Amorim, A. 2015. Preliminary occurrence of Istiophoridae larvae (Perciformes, Xiphioidei) in southern Brazil. ICCAT Collective Volume of Scientific Papers, 71: 2256–2261.
- Snodgrass, D., Orbesen, E., Hoolihan, J., and Prince, E. 2011. The U.S. conventional tagging database updates for Atlantic white marlin (1954–2008). ICCAT Collective Volume of Scientific Papers, 66: 1760–1766.
- Sorenson, L., McDowell, J., and Graves, J. 2011. Isolation and characterization of microsatellite markers for blue marlin, *Makaira nigricans*. Conservation Genetics Resources, 3: 721–723.
- Tessier, N., and Bernatchez, L. 1999. Stability of population structure and genetic diversity across generations assessed by microsatellites among sympatric populations of landlocked Atlantic salmon (*Salmo salar* L.). Molecular Ecology, 8: 169–179.
- Ueyanagi, S., Kikawa, S., Uto, M., and Nishikawa, Y. 1970. Distribution, spawning, and relative abundance of billfishes in the Atlantic Ocean. Bulletin of the Far Seas Fisheries Research Laboratory, 3: 15–25.
- Van Oosterhout, C., Hutchinson, W., Wills, D., and Shipley, P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4: 535–538.
- Van Oosterhout, C., Weetman, D., and Hutchinson, W. 2006. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. Molecular Ecology Notes, 6: 255–256.
- Waples, R. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. The Journal of Heredity, 89: 438–450.
- Waples, R., and 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. Molecular Ecology, 15: 1419–1439.
- Waples, R., and Teel, D. 1990. Conservation genetics of Pacific salmon I. Temporal changes in allele frequency. Conservation Biology, 4: 144–156.
- Ward, R. 2000. Genetics in fisheries management. Hydrobiologia, 420: 191–201.
- Ward, R., Elliott, N., Grewe, P., and Smolenski, A. 1994. Allozyme and mitochondrial DNA variation in yellowfin tuna (*Thunnus albacares*) from the Pacific Ocean. Marine Biology, 118: 531–539.
- Ward, R., Woodwark, M., and Skibinski, D. 1994. A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. Journal of Fish Biology, 44: 213–232.
- Williams, S., Bennett, M., Pepperell, J., Morgan, J., and Ovenden, J. 2015b. Spatial genetic subdivision among populations of the highly migratory black marlin *Istiompax indica* within the central Indo-Pacific. Marine and Freshwater Research, 166: 29–32.
- Williams, S., Pepperell, J., Corley, S., and Ovenden, J. 2015a. Isolation and characterisation of 18 polymorphic microsatellite loci for black marlin (*Istiompax indica*) and their utility for Pacific billfish species. Fisheries Research, 166: 29–32.

Handling editor: Lorenz Hauser