Received: 14 August 2015

IOTC–2015–WPB13–30

PRELIMINARY REPORT

Development of a novel high-throughput assay to evaluate genetic population structure in striped marlin (*Kajikia audax*)

Nadya Mamoozadeh, Jan McDowell, & John Graves

Virginia Institute of Marine Science
College of William and Mary
1375 Greate Road
Gloucester Point, Virginia 23072
United States of America

Contact: nrmamoozadeh@vims.edu

Preface

The purpose of this document is to provide a preliminary overview of a component of the research currently ongoing at the Virginia Institute of Marine Science (VIMS), and represents the latest work in a more than 20 year program focused on the genetics of istiophorid billfishes. The project described in this preliminary report is focused on evaluating the population structure of striped marlin (*Kajikia audax*) in the Indian and Pacific oceans using high-throughput genetic methodology. Results from this research will provide information on the number and geographic extent of striped marlin populations within these oceans, as well as the level of inter-oceanic connectivity. These data are important for informing management measures determined by regional fishery management organizations such as the Indian Ocean Tuna Commission.

Abstract

To date, population genetic studies of highly migratory marine fishes have generally been characterized by a small number of molecular markers that represent a limited portion of the genome, and opportunistic sampling designs that include a small number of individuals per putative population. These characteristics compromise the statistical power necessary to detect the low levels of genetic differentiation typically associated with populations of marine fishes. Additionally, unintentional sampling of mixed-population assemblages results in a noisy genetic signal that may obscure population-specific information. Although previous evaluations of genetic population structure in Pacific striped marlin have identified multiple populations, genetic differentiation has been low and likely compromised by small numbers of molecular markers and samples per population, and sampling of mixed-population assemblages. In the current study, next-generation sequencing-based methodology will be used to identify large numbers of molecular markers in samples collected using a biologically-informed sampling design to target individual populations. Utilization of low polymorphism markers will reduce the number of samples necessary per putative population. These factors will significantly improve the ability to detect low levels of genetic differentiation and facilitate the assessment of genetic population structure for striped marlin in the Pacific and, for the first time, Indian oceans.

Background

Population genetic studies of marine fishes assist the identification of population units that may correspond to distinct units for management and assessment (Allendorf et al., 1987; Ward, 2000; Ovendon et al., 2013). This information is necessary to conserve genetic diversity, and prevent the unintentional overfishing of some populations. Despite the utility of the information resulting from these studies, evaluation of genetic population structure in marine fishes is particularly challenging. Large population sizes result in the slow accumulation of a significant level of genetic differences, and a relatively homogeneous environment presumably facilitates higher levels of connectivity among populations. These characteristics result in significantly lower levels of genetic differentiation among populations of marine fishes compared to their freshwater counterparts (Ward et al., 1994). A higher level of statistical power is therefore necessary to detect genetic differentiation among populations of marine fishes, and can be increased by evaluating a larger number of molecular markers across a greater number of individuals per putative population (Waples, 1998). The recent development of low cost, high-throughput next-generation sequencing technology now enables the development of large numbers of molecular markers that can then be analyzed across sizeable numbers of samples in a high-throughput manner. With these new molecular capabilities, remaining challenges to optimizing statistical power in population genetic studies of marine fishes are associated with sampling designs appropriate for these species.

Sampling designs that incorporate large numbers of samples per putative population increase statistical power by enabling the allele frequencies of each molecular marker to be more accurately and precisely characterized in each population (Waples, 1998; Morin et al., 2009). For highly polymorphic molecular markers that may display tens of alleles per locus, large sample sizes per putative population are especially important; however, obtaining a large number of samples is particularly challenging for rare-event species. This difficulty can be alleviated by using a large number of low polymorphism molecular markers for which a smaller number of samples are necessary to characterize allele frequencies. Compared to highly polymorphic molecular markers, a larger number of low polymorphism markers are required to provide the same total number of alleles—this is important because measures of intra-specific genetic differentiation depend on the total number of alleles analyzed (Morin 2004, 2009). The need for large numbers of low polymorphism molecular markers is no longer limiting given the advent of next-generation sequencing. Population genetic studies in recent years have demonstrated greater use of low polymorphism single nucleotide polymorphism (SNP) molecular markers (typically two alleles per locus) compared to traditionally-utilized, highly polymorphic microsatellite markers (tens of alleles per locus). The utilization of SNPs–single basepair differences among individuals that may be informative of population structure–facilitates smaller sample sizes per putative population, enabling appropriate levels of statistical power in population genetic studies of rare-event species.

For highly migratory marine fishes, it is also important to consider a sampling design informed by biological characteristics of the study species (Graves et al., 1996; Bowen et al., 2005; Carlsson et al., 2007; Graves & McDowell, 2014). Due to the rare-event nature of these species, previous evaluations of genetic population structure have relied on samples opportunistically collected from a variety of geographic locations. Highly migratory marine fishes are capable of long distance dispersal (including trans- and/or inter-oceanic movements), and sampling geographically distant localities is not necessarily representative of distinct populations. In addition, many of these species display mixed-population assemblages at times

throughout the year. Opportunistic sample collections containing mixed-population assemblages result in a noisy genetic signal that obscures the discrimination of individual populations. If instead populations are sampled when naturally separated, detection of a population-specific genetic signal is possible, and individual populations can be identified. For many highly migratory marine fishes, populations are thought to naturally separate at the time of spawning (Graves et al., 1996; Carlsson et al., 2007; Graves & McDowell, 2014). Population genetic studies that focus sampling efforts on larvae and reproductively active adults enhance the ability to determine population structure in highly migratory marine fishes. To date, this strategy has been implemented in the evaluation of genetic population structure in Atlantic bluefin tuna (*Thunnus thynnus*), but not yet in similar studies of other highly migratory marine fishes. Early evaluations of genetic population structure in *T. thynnus* utilized samples collected opportunistically throughout the species range in the Atlantic Ocean, and statistically significant genetic differentiation was not detected (i.e., Ely et al., 2002; Pujolar et al., 2003). Subsequent tagging (Block et al., 2005; Wilson et al., 2005), otolith (Rooker et al., 2008), and tissue organochlorine studies (Dickhut et al., 2009; Graves et al., 2015) demonstrated the presence of eastern and western stocks of Atlantic bluefin, with distinct spawning grounds in the Mediterranean Sea and Gulf of Mexico. More recent genetic analyses that have incorporated this biological information into the sampling design have demonstrated low but statistically significant genetic differentiation between individuals sampled from the eastern and western spawning grounds (Carlsson et al., 2007; Boustany et al., 2008).

The istiophorid billfishes are characterized by long distance migrations and large-scale species distributions in tropical, sub-tropical, and temperate waters of the Atlantic, Pacific, and/or Indian oceans (Ortiz et al., 2003; Nakamura, 1985). The highly migratory nature of these species presumably facilitates even greater levels of intraspecific connectivity compared to other marine fishes. Despite this assumption, previous population genetic studies in some istiophorid species have revealed intra-oceanic population differentiation for striped marlin in the Pacific Ocean (Graves & McDowell, 1994; McDowell & Graves, 2008; Purcell & Edmands, 2011), and sailfish in the Indian and Pacific oceans (McDowell, 2002; Hoolihan et al., 2004). Inter-oceanic heterogeneity has been observed between Atlantic and Pacific populations of blue marlin and sailfish (Graves & McDowell, 1995). In general, these studies analyzed low numbers of molecular markers that represent a limited portion of the genome, and utilized relatively low numbers of samples opportunistically collected from locations throughout the species range due to the rare-event nature of istiophorids compared to more coastal species.

Throughout its distributional range, striped marlin are caught as bycatch in pelagic longline fisheries targeting tunas and swordfish, and are targeted in small-scale artisanal and primarily catch-and-release recreational sport fisheries. In the Indian Ocean, striped marlin are currently recognized by the IOTC as a single ocean-wide population. The most recent assessment of population status for striped marlin in this region indicates that the species is overfished and experiencing overfishing (IOTC, 2014). There has been no previous evaluation of genetic population structure for striped marlin in the Indian Ocean, and tagging efforts in this region have been limited. In the Pacific Ocean, genetic studies have discriminated four populations of striped marlin and this information has been beneficial in informing the regional management of this species (McDowell & Graves, 2008; Purcell & Edmands, 2011). Given the current population status of striped marlin in the Indian Ocean, it is important that the population structure of this species is understood in this region. The purpose of our research at VIMS is to evaluate the genetic population structure of striped marlin in the Indian Ocean using those

characteristics described above to optimize statistical power and the ability to detect low levels of genetic differentiation. This study is part of a larger initiative to evaluate the genetic population structure of striped marlin throughout its entire range, inclusive of the Pacific Ocean, and also enabling the evaluation of inter-oceanic connectivity.

Objectives

Objectives of this research project include the following:

1. Develop a sampling network throughout the range of striped marlin in the Indian and Pacific oceans that targets larvae and reproductively active adults as well as other seasonal assemblages (i.e., for feeding).

2. Develop a next-generation sequencing-based methodology for the *de novo* (i.e., without the use of a reference genome) discovery of SNP loci in striped marlin.

3. Generate SNP genotype data for all study samples and perform bioinformatic analyses to evaluate the null hypotheses that striped marlin constitute single, ocean-wide populations in the Indian and Pacific oceans.

Methodology

Sample Collection: Samples will consist of whole larvae and tissue from adult and juvenile fish. Efforts to develop an Indo-Pacific sampling network of anglers, agencies, and scientists have been ongoing for the past several months. Samples will be collected from locations throughout the Indo-Pacific where striped marlin display seasonal abundance, with primary focus on known spawning grounds. Indian Ocean sampling efforts are currently ongoing at those locations described in Figure 1. Sampling will occur over two consecutive years (2015-2017) so that the temporal stability of the genetic patterns observed in the study can be assessed. Existing samples located in the Fisheries Genetics Laboratory at VIMS (Table 1) will be used to provide additional geographic coverage, and demographic groups. The extended temporal coverage provided by samples in the VIMS collection may also enable the evaluation of genetic differences between generations of striped marlin, providing unique insight into the inter-generational population genetics of the species. In addition, the results of this project will be directly comparable to current and previous work on the population genetics of other istiophorid species at VIMS.

SNP Discovery: Although the reduced costs and increased availability of next-generation sequencing technology in recent years now enables the sequencing of entire genomes, the number of individual genomes required for a population genetic study is still cost-prohibitive. In addition, genetic data associated with full genomes is highly computationally expensive. For these reasons, methods have been developed to reduce the genomic content of each sample prior to sequencing (i.e., Baird et al., 2008). These ‘reduced genomic representations’ can be generated using restriction enzymes that fragment DNA by cutting at specific recognition motifs randomly distributed throughout the genome. A portion of the fragments are then sequenced using a next-generation platform, and resulting sequence data is quality filtered and searched for the presence of SNPs.

High-Throughput Genotyping: While the SNP discovery process may identify thousands of SNPs in striped marlin, a subset of SNPs will be included in a final SNP panel to be genotyped across all samples included in the study. Genotyping will then be performed in a high-throughput manner that enables a large number of samples to be genotyped in a short period of

time. Resulting SNP genotype data will be analyzed in a bioinformatic pipeline that ultimately tests for genetic differentiation among groups of samples. Various sample grouping scenarios will be tested based on spawning location, as well as length class and sex to determine if genetic differences between demographic groups exist. Inferences resulting from genetic analyses will be used in combination with information from tagging and catch data to provide comprehensive insight into striped marlin population structure.

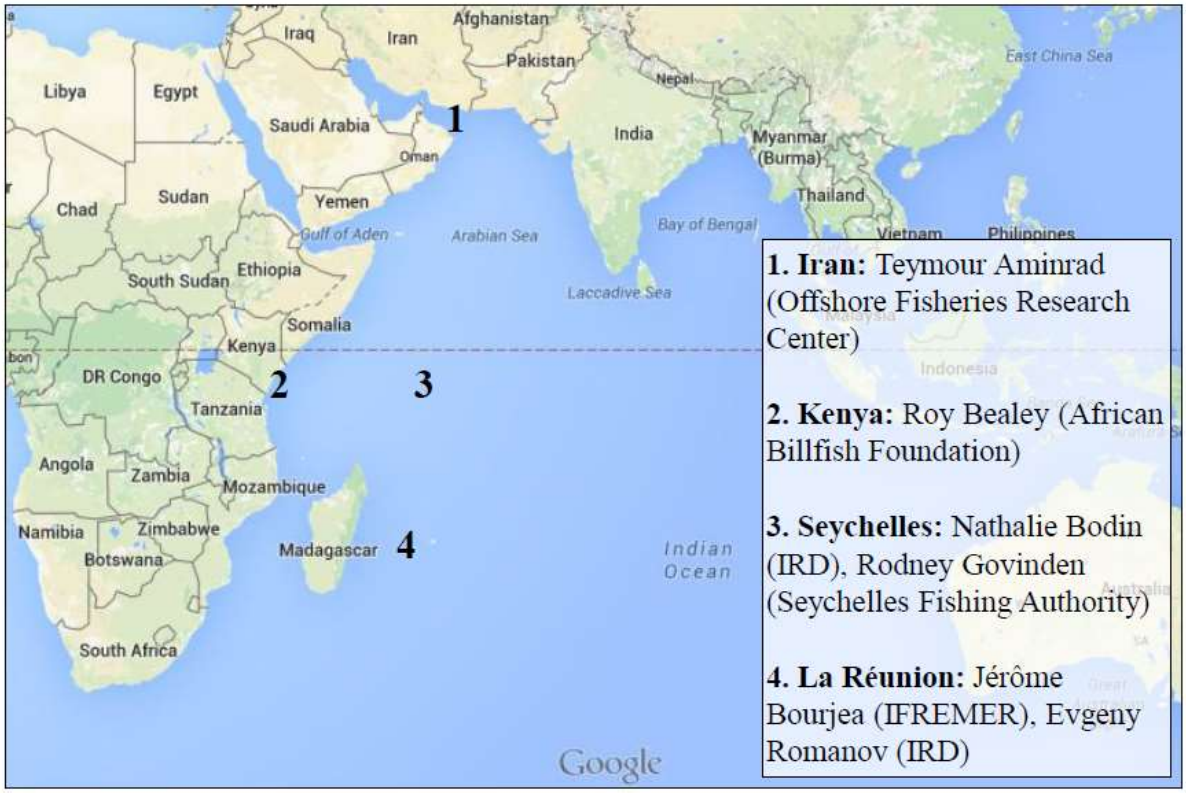
Project Significance

Information expected to result from this project includes the identification of the number and geographic extent of striped marlin populations in the Indian Ocean. This information is essential for the IOTC to determine if the current single-stock management and assessment model is appropriate for striped marlin in the Indian Ocean. Results stemming from the full project will also provide an enhanced understanding of population boundaries in Pacific striped marlin, as well as the inter-oceanic relationship between striped marlin in the Indian and Pacific oceans. Collectively, an understanding of the number and geographic extent of striped marlin populations throughout the distributional range of this species will allow managers and anglers to understand which populations they interact with. This information is especially important for Indian Ocean striped marlin considering the current population status. Additionally, the molecular methodology and biologically-informed sampling design developed in this study will provide insight for future population genetic studies of other highly migratory marine fishes. Results of this study will be made available to relevant management agencies including the IOTC and regional fishery management organizations in the Pacific Ocean.

Table 1. Tissue samples of striped marlin available for genetic analysis from the Fisheries Genetics Laboratory at VIMS.

Geographic Location	Total No. Samples
Australia (East Coast)	120
Baja California	545
Southern California, USA	47
Central America	107
Galapagos Islands	38
Taiwan	33
Kenya	33

Figure 1. Current sampling locations and agencies for the collection of striped marlin genetics samples in the Indian Ocean.



Literature Cited

Allendorf, F., Ryman, N., & Utter, F. (1987). Genetics and fishery management: past, present, and future. In N. Ryman & F. Utter (Eds.), *Population genetics and fishery management* (pp. 1-19). Seattle: University of Washington Press.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., & Johnson, E.A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE, 3*(10), e3376.

Block, B., Teo, S., Walli, A., Boustany, A., Stokesbury, M., Farwell, C., Weng, K., Dewar, H., & Williams, T. (2005). Electronic tagging and population structure of Atlantic bluefin tuna. *Nature, 434*, 1121-1127.

Boustany, A., Reeb, C., & 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., & Toonen, R. (2005). Conservation implications of complex population struture: lessons from the loggerhead turtle (*Caretta caretta*). *Molecular Ecology, 14*, 2389-2402.

Carlsson, J., McDowell, J. R., Carlsson, J. E. L., & Graves, J. E. (2007). Genetic identity of YOY bluefin tuna from the eastern and western Atlantic spawning areas. *The Journal of Heredity, 98*(1), 23-28.

Dickhut, R., Deshpande, A., Cincinelli, A., Cochran, M., Corsolini, S., Brill, R., Secor, D., & Graves, J. (2009). Atlantic bluefin tuna (*Thunnus thynnus*) population dynamics delineated by organochlorine tracers. *Environmental Science & Technology, 43,* 8522-8527.

Ely, B., Stoner, D. S., Avarado-Bremer, J. R., Dean, J. M., Addis, P., Cau, A., Thelen, E. J., Jones, W. J., Black, D. E., Smith, L., Scott, K., Naseri, I., & Quattro, J. M. (2002). Analyses of nuclear IdhA gene and mtDNA control region sequences of Atlantic northern bluefin tuna populations. *Marine Biotechnology, 4*, 583-588.

Graves, J. E., Gold, J. R., Ely, B., Quattro, J. M., Woodley, C., & Dean, J. M. (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. E. & McDowell, J. R. (1994). Genetic analysis of striped marlin (*Tetrapturus audax*) population structure in the Pacific Ocean. *Canadian Journal of Fisheries and Aquatic Sciences, 51*, 1762-1768.

Graves, J. E. & McDowell, J. R. (1995). Inter-ocean genetic divergence of Istiophorid billfishes. *Marine Biology, 122*, 193-203.

Graves, J. E. & McDowell, J. R. (2014). Population structure of istiophorid billfishes. *Fisheries Research, 166,* 21-28.

Graves, J. E., Wozniak, A., Dickhut, R., Cochran, M., MacDonald, E., Bush, E., Arrizabalaga, H., & Goni, N. (2015). Transatlantic movements of juvenile Atlantic bluefin tuna inferred from analyses of organochlorine tracers. *Canadian Journal of Fisheries and Aquatic Sciences, 72*, 1-9.

Hoolihan, J. P., Premanandh, J., D'Aloia-Palmieri, M. –A., Benzie, J. A. H. (2004). Intraspecific phylogeographic isolation of Arabian Gulf sailfish *Istiophorus platypterus* inferred from mitochondrial DNA. *Marine Biology, 145*, 465-475.

Indian Ocean Tuna Commission (IOTC). (2014). *Report of the eighteenth session of the indian ocean tuna commission*. Colombo, Sri Lanka: 1-5 June 2014. IOTC-2014-S18-R[E].

McDowell, J. R. (2002). *Genetic stock structure of the sailfish, Istiophorus platypterus, based on nuclear and mitochondrial DNA* (Doctoral dissertation). Virginia Institute of Marine Science, Gloucester Point, VA.

McDowell, J. R. & Graves, J. E. (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*(7), 1307–1320.

Morin, P. A., Luikart, G., & Wayne, R. K. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution, 19*(4), 208-216.

Morin, P. A., Martien, K. K., & Taylor, B. (2009). Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources, 9*, 66-73.

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: Food and Agriculture Organization of the United Nations.

Ovendon, J., Berry, O., Welch, D., Buckworth, R., & Dichmont, C. (2013). 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.

Pujolar, J. M., Roldan, M. I., & Pla, C. (2003). Genetic analysis of tuna populations, *Thunnus thynnus* and *T. alalunga*. *Marine Biology, 143*, 613-621.

Purcell, C. M. & 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.

Rooker, J., Simms, J., Wells, R., Holt, S., Holt, G., Graves, J., & Furey, N. (2012). Distributions and habitat associations of billfish and swordfish larvae across mesoscale features in the Gulf of Mexico. *PLoS One, 7,* e31480.

Waples, R. S. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *The Journal of Heredity*, *89*(5), 438–450.

Ward, R., Woodward, M., & Skibinski, D. (1994). A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *Journal of Fisheries Biology, 44,* 213-232.

Ward, R. (2000). Genetics in fisheries management. *Hydrobiologia, 420,* 191-201.

Wilson, S. G., Lutcavage, M. E., Brill, R. W., Genovese, M. P., Cooper, A. B., & Everly, A. (2005). Movements of bluefin tuna (*Thunnus thynnus*) in the northwestern Atlantic Ocean recorded by pop-up satellite archival tags. *Marine Biology*, *146*, 409-423.