Co-occurrence of genetically isolated groups of skipjack tuna (*Katsuwonus pelamis*) within the Indian Ocean

Naiara Rodríguez-Ezpeleta[1,*], Iraide Artetxe-Arrate[1], Iñaki Mendibil[1], Natalia Díaz-Arce[1], Iñigo Krug[1], Jon Ruiz[1], Natacha Nikolic[2,#], Anais Medieu[2], Marianne Pernak[2] Jessica Farley[3], Peter Grewe[3], Matt Lansdell[2], Jorden Aulich[2], Naomi Clear[2], Craig Proctor[2,] Wudianto[4], Toni Ruchimat[4], Zulkarnaen Fahmi[4], Fayakun Satria[4], Pratiwi Lestari[4], Muhammad Taufik[4], Asep Priatna[6], Achmad Zamroni[6], Campbell R. Davies[3], Francis Marsac[2], Igaratza Fraile[1], Hilario Murua[5]

[1]Marine Research, AZTI BRTA, Sukarrieta, Bizkaia, Spain

[2] Marbec, Univ Montpellier, CNRS, Ifremer, IRD, Montpellier, France

[3] CSIRO Oceans and Atmosphere, Hobart; Tasmania, Australia

[4] Agency for Marine and Fisheries Research and Human Resources

[5] International Seafood Sustainability Foundation, Washington DC, USA.

[#] Present address: INRAE, Saint-Pée-sur-Nivelle, France

*Correspondence: nrodriguez@azti.es

Abstract

In order to resolve the population connectivity of skipjack (*Katsuwonus pelamis*) within the Indian Ocean, we analyzed thousands of genome-wide markers of individuals from a broad geographic area of the Indian Ocean, as well as from one location in the Atlantic Ocean. Our results support a complex stock structure with multiple genetically isolated populations co-occurring in most locations, and claim for additional analyses to further understand the population structure of skipjack tuna within the Indian Ocean.

Introduction

Previous studies on the population structure of skipjack tuna have resulted in conflicting results regarding differentiation between and within oceans. For example, analyses based on the mitochondrial D-loop region (Menezes, Ikeda, Taniguchi 2006) or on nuclear microsatellite markers (Menezes et al. 2008) found, respectively, high levels of or no genetic differentiation between skipjack tuna from the Indian and Pacific oceans.

Additional analyses comparing Atlantic and Pacific Oceans skipjack tuna also revealed a high degree of genetic similarity among these two oceans (Graves, Ferris, Dizon 1984; Ely et al. 2005). Within the Indian Ocean, several studies have reported the existence of heterogeneous groups that were not related to any spatial pattern (Dammannagoda, Hurwood, Mather 2011; Menezes, Kumar, Kunal 2012; Jatmiko et al. 2019), but these studies were geographically limited preventing an understanding the population structure of skipjack of the Indian Ocean as a whole. In order to improve the assessment and management of skipjack tuna in the Indian Ocean, we further investigated the population structure of this species by analyzing thousands of genetic markers from samples obtained from almost its entire distribution range.

Materials and methods

Sampling, DNA extraction and RAD-seq library preparation

Skipjack tuna samples from the Indian and Atlantic Ocean were obtained by scientific observers on-board purse seiner vessels or by sampling fish during landings at port (Figure 1). A total of 525 samples were collected from 9 areas, named according to sampling area or sampling port (Gulf of Guinea, Great Australian Bight, Benoa, Lampulo, Maldives, Pakistan, Madagascar, Somalia and Seychelles). Samples were obtained from fish classified as i) less than 6 month-old young of the year (<35 cm), ii) more than 6 month-old young of the year (35-45 cm) or iii) adults (>45 cm), according to the age-length-key relationship described in (Eveson et al. 2015) and maturity threshold in (Grande et al. 2014). From each specimen, a ~1 cm^3 muscle tissue sample was excised and stored in 96% molecular grade ethanol at -20 °C prior to DNA extraction. Genomic DNA extraction and RAD-seq library preparation and sequencing followed the protocols described in (Rodríguez-Ezpeleta et al. 2019)

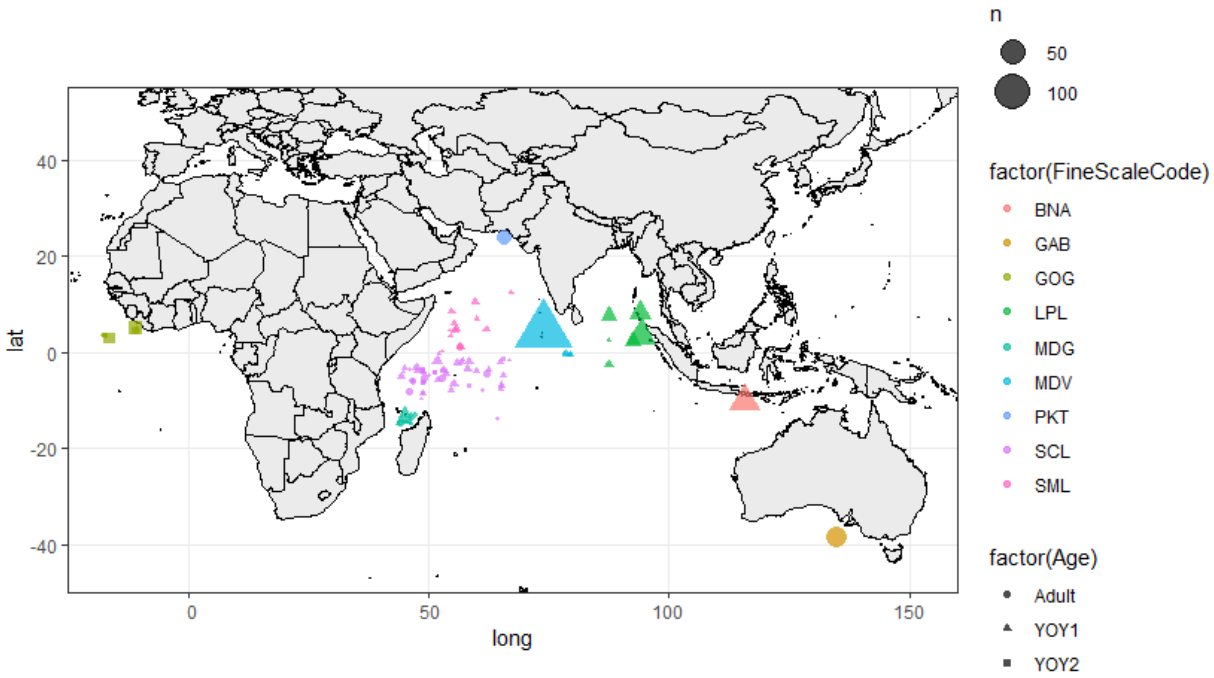


Fig. 1. Samples collected for this study. Each location is represented by one color (BNA – Benoa, GAB – Great Australian Bight, GOG – Gulf of Guinea, LPL – Lampulo, MDG – Madagascar, MDV – Maldivas, PKT – Pakistan, SCL – Seychelles, SML – Somalia) and shapes indicate if samples are less than 6 month old young of the year (YOY1), more than 6 month old young of the year (YOY2) or adults. Size of shapes are proportional to the number of samples collected per area/age.

RAD-tag assembly and SNP calling

Generated RAD-tags were analyzed using *Stacks* version 2.4 (Catchen et al. 2013) using the procedures described in (Rodríguez-Ezpeleta et al. 2019). Only samples with more than 40,000 and less than 70,000 reads were selected for further analyses. The module *populations* was used to export from the catalog, the SNPs presented in RAD loci found in at least 90% of the individuals and in the first 100 bases of each contig. Using *PLINK* version 1.07 (Purcell et al. 2007), SNPs with more than 5% missing data and a minimum allele frequency (MAF) smaller than 0.05 as well as samples with more than 25% missing data were excluded from downstream analyses. The resulting genotype tables were exported into *Structure* and *Genepop* formats.

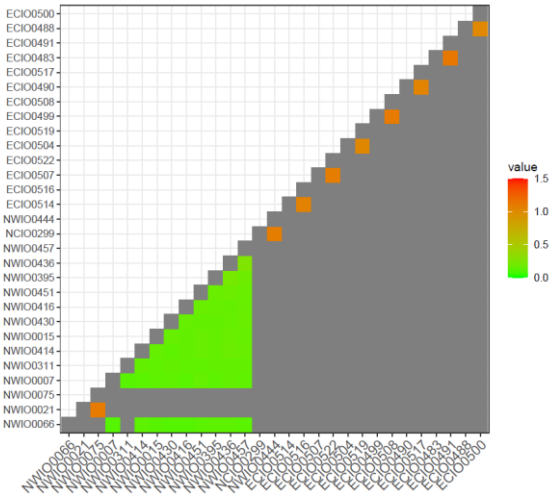
Genetic diversity and population structure analyses

Related individuals were identified using GCTA (Yang et al. 2011). Principal Component Analysis (PCA) were performed using the *adegenet* R package (Jombart, Ahmed 2011),

and ADMIXTURE (Alexander, Novembre, Lange 2009) was run assuming from 1 to 6 ancestral populations (K) setting default parameters. The value of K with lowest associated error value was identified using ADMIXTURE's cross-validation procedure.

Results

We found pairs of related individuals. Although we cannot discount that some of the individuals with very high relatedness coefficients (close to 1) were not due to contamination, we found a group of 11 individuals (10-25% relatedness) which are very unlikely due to cross-contaminations as members of groups were collected on different years, by different vessels and processed separately in the laboratory (Figure 2).



An-Code	Area	Sampling Date	Processing Date
NWIO0007	Seychelles	14/02/2018	06/06/2018
NWIO0015	Seychelles	01/02/2018	06/06/2018
NWIO0066	Seychelles	29/01/2018	06/06/2018
NWIO0311	Somalia	05/04/2019	04/06/2019
NWIO0395	Pakistan	01/09/2018	30/08/2019
NWIO0414	Pakistan	01/09/2018	30/08/2019
NWIO0416	Pakistan	01/09/2018	30/08/2019
NWIO0430	Pakistan	01/09/2018	30/08/2019
NWIO0436	Pakistan	01/09/2018	30/08/2019
NWIO0451	Pakistan	01/09/2018	30/08/2019
NWIO0457	Pakistan	01/09/2018	30/08/2019

Fig. 2. Relatedness coefficient among those pairs that show a value higher than >0.10, excluding individuals within themselves (left) and data of samples that are not identical (green) but seem related to each other (right).

Principal component analyses (Figure 3) based on the filtered dataset consisting of 368 individuals (excluding one member of each pair of related individuals) and 14346 SNPs show five groups of samples: i) a "central" group that contains individuals from all locations, ii) a group that contains individuals from all locations except from Pakistan, iii) a group that contains individuals from all locations except from Atlantic Ocean (represented here by the Gulf of Guinea) and Great Australian Bight and two small

groups that contain respectively iv) individuals from Maldives and Lampulo and v) from Gulf of Guinea and Great Australian Bight.

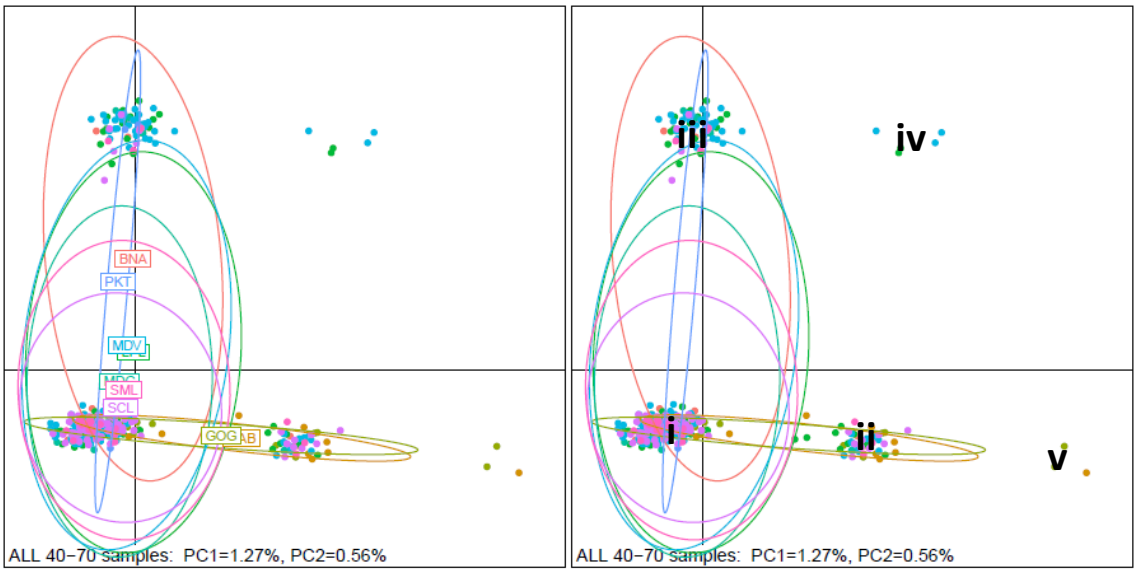


Fig. 3. Principal component analysis using the filtered datasets. Left plots are identical to right plots but with the groups mentioned in the text tagged.

ADMIXTURE analyses show that the number of assumed ancestral populations with the lowest associated error (K) is 2. Individual ancestral proportions (Figure 4) show two clusters, one including samples from all locations and a second one including individuals from all locations except from Gulf of Guinea and Great Australian Bight.

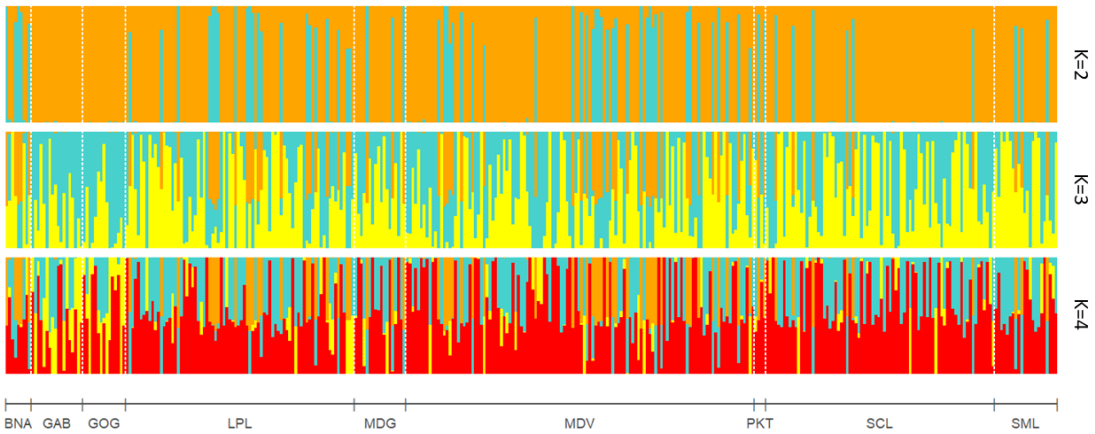


Fig. 4. Individual ancestry proportions estimated by ADMIXTURE when assuming from two to four ancestral populations (K) of samples from different locations.

Main conclusions

Our findings are congruent with previous findings and support existence of some latitudinal spatial separation from the north to the south of the Indian Ocean in the sense that two of the genetic groups are present either in Pakistan or in the Gulf of Guinea and Great Australian Bight respectively. However, all areas share at least one of the genetic groups suggesting that the population structure of the skipjack is complex and that its understanding will require testing hypothesis such as the presence of stocks spawning at different locations, the effect of currents in larvae drifting and potential co-occurrence of genetically isolated populations. Additional analyses based on adult, juvenile and larvae samples collected through several years are required to elucidate the population structure of skipjack tuna within the Indian Ocean.

Acknowledgements

This work is part of the PSTBS project supported by funding by CSIRO Oceans and Atmosphere, AZTI Basque Research and Technology Alliance, Institut de recherche pour le développement (IRD), and Research Institute for Tuna Fisheries (RITF) and financial assistance of the European Union (GCP/INT/233/EC – Population structure of IOTC species in the Indian Ocean). The views expressed herein can in no way be taken to reflect the official opinion of the European Union. We would like to thank the many vessel owners, skippers, observers, crews and processors who provided fish for sampling and access to facilities.

References

Alexander, DH, J Novembre, K Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome research 19:1655-1664.

Catchen, J, PA Hohenlohe, S Bassham, A Amores, WA Cresko. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22:3124-3140.

Dammannagoda, S, D Hurwood, P Mather. 2011. Genetic analysis reveals two stocks of skipjack tuna (Katsuwonus pelamis) in the northwestern Indian Ocean. Canadian Journal of Fisheries and Aquatic Sciences 68:210-223.

Ely, B, J Viñas, JR Alvarado Bremer, D Black, L Lucas, K Covello, AV Labrie, E Thelen. 2005. Consequences of the historical demography on the global population structure of two highly migratory cosmopolitan marine fishes: the yellowfin tuna (Thunnus albacares) and the skipjack tuna (Katsuwonus pelamis). BMC Evolutionary Biology 5:19.

Eveson, JP, J Million, F Sardenne, G Le Croizier. 2015. Estimating growth of tropical tunas in the Indian Ocean using tag-recapture data and otolith-based age estimates. Fisheries Research 163:58-68.

Grande, M, H Murua, I Zudaire, N Goñi, N Bodin. 2014. Reproductive timing and reproductive capacity of the Skipjack Tuna (Katsuwonus pelamis) in the western Indian Ocean. Fisheries Research 156:14-22.

Graves, JE, SD Ferris, AE Dizon. 1984. Close genetic similarity of Atlantic and Pacific skipjack tuna (Katsuwonus pelamis) demonstrated with restriction endonuclease analysis of mitochondrial DNA. Marine Biology 79:315-319.

Jatmiko, I, RR Zedta, M Agustina, B Setyadji. 2019. Genetic Diversity and Demography of Skipjack Tuna (Katsuwonus pelamis) In Southern and Western Part of Indonesian Waters. 2019 24:8.

Jombart, T, I Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070-3071.

Menezes, MR, M Ikeda, N Taniguchi. 2006. Genetic variation in skipjack tuna Katsuwonus pelamis(L.) using PCR-RFLP analysis of the mitochondrial DNA D-loop region. Journal of Fish Biology 68:156-161.

Menezes, MR, G Kumar, SP Kunal. 2012. Population genetic structure of skipjack tuna Katsuwonus pelamis from the Indian coast using sequence analysis of the mitochondrial DNA D-loop region. Journal of Fish Biology 80:2198-2212.

Menezes, MR, D Noguchi, M Nakajima, N Taniguchi. 2008. Microsatellite development and survey of genetic variation in skipjack tuna Katsuwonus pelamis. Journal of Fish Biology 73:463-473.

Purcell, S, B Neale, K Todd-Brown, et al. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics 81:559-575.

Rodríguez-Ezpeleta, N, N Díaz-Arce, JF Walter Iii, et al. 2019. Determining natal origin for improved management of Atlantic bluefin tuna. Frontiers in Ecology and the Environment 17:439-444.

Yang, J, SH Lee, ME Goddard, PM Visscher. 2011. GCTA: A Tool for Genome-wide Complex Trait Analysis. The American Journal of Human Genetics 88:76-82.