SCIENTIFIC COMMITTEE
SEVENTEENTH REGULAR SESSION

Electronic Meeting
11-19 August 2021

Connectivity of broadbill swordfish targeted by the Australian Eastern Tuna and Billfish Fishery with the broader Western Pacific Ocean

WCPFC-SC17-2021/SA-IP-12

K. Evans[1], P. Grewe[1], S. Foster[2], R. Gunasekera[1], M. Lansdell[1], S. Meredith[3], S. Sarau[4], S. Tracey[5], M. Wichman[4]

[1] CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart Tasmania 7001, Australia
[2] Data61, CSIRO, GPO Box 1538, Hobart Tasmania 7001, Australia
[3] Bay Packers, 5-11 Tyne St, Mt Maunganui, New Zealand
[4] Ministry of Marine Resources, PO Box 85, Avarua, Rarotonga, Cook Islands
[5] Institute for Marine and Antarctic Science, University of Tasmania, Private Bay 129, Hobart, Tasmania 7001, Australia

Connectivity of broadbill swordfish targeted by the Australian Eastern Tuna and Billfish Fishery with the broader Western Pacific Ocean.

Karen Evans[1], Peter Grewe[1], Scott Foster[2], Rasanthi Gunasekera[1], Matt Lansdell[1], Steve Meredith[3], Saiasi Sarau[4], Sean Tracey[5], Marino Wichman[4].

[1]CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart Tasmania 7001, Australia
[2]Data61, CSIRO, GPO Box 1538, Hobart Tasmania 7001, Australia
[3]Bay Packers, 5-11 Tyne St, Mt Maunganui, New Zealand
[4]Ministry of Marine Resources, PO Box 85, Avarua, Rarotonga, Cook Islands
[5]Institute for Marine and Antarctic Science, University of Tasmania, Private Bay 129, Hobart, Tasmania 7001, Australia

Abstract

Australia's Eastern Tuna and Billfish Fishery (ETBF) operates in waters off on the east coast of Australia and catches a number of pelagic species including broadbill swordfish. The distribution of swordfish is known to extend well beyond Australian waters and fish caught in the ETBF are considered to form part of a wider Western and Central Pacific Ocean (WCPO) population. Accordingly, regional stock assessments carried out for the Western and Central Pacific Fisheries Commission (WCPFC) assume a single interconnected stock. However, the specifics on the connectivity of swordfish caught in the ETBF with the WCPO is still a major source of uncertainty and tagging and fisheries data suggest that there may be some structuring of the species across the region. Using high throughput genetic sequencing of total genomic DNA derived from swordfish collected from the ETBF, including northern (Queensland), southern (Tasmania) and far eastern (Norfolk Island) regions, New Zealand and the Cook Islands, co-dominant single nucleotide polymorphism (SNP) markers were identified. Purpose built quality control pipelines allowed for the discrimination of poor quality and compromised (e.g. through DNA cross contamination) samples. Mixture models were then used to investigate the presence of one or more genetic populations for each species. Results suggest there is little genetic differentiation between swordfish caught at the locations sampled. However, given the low numbers of samples within years examined from eastern Tasmania and the Cook Islands, the results presented should be considered preliminary. Further sampling from these locations would be required for resolving the indicative nature of the results and further sampling from additional sites would be needed to investigate whether the results presented here extend to other locations across the western and central Pacific region.

This analysis updates preliminary analyses of swordfish population structure presented at SC16 and provides further insights into the contemporary connectivity of swordfish within the context of current approaches to the assessment of stocks under the WCPFC. The SC17 is invited to consider these results in discussions of the current stock assessment and associated future planning of stock assessments for swordfish.

Introduction

Australia's Eastern Tuna and Billfish Fishery (ETBF) operates in waters off on the east coast of Australia and catches a number of pelagic species including broadbill swordfish. Populations of swordfish are known to extend well beyond the Australian Exclusive Economic Zone (EEZ) and fish caught in the ETBF are considered to form part of at least a wider Western Pacific Ocean (WPO)

population, although specifics on connectivity between various regions is still a major source of uncertainty. In line with this assumption, swordfish are currently assessed as a single inter-connected stock distributed across the wider western and central Pacific Ocean (see Ducharme-Barth et al. 2021, SC17-SA-WP-04). Although swordfish are assessed as a single interconnected stock, movement data derived from tagging studies and catch data suggest that there may be some structuring of the species throughout the WCPFC region (Campbell and Hobday 2003, Evans et al. 2014).

The technical advances of DNA profiling used to investigate the population structure of yellowfin tuna now provide for high throughput sequencing platforms and improved power of population discrimination at much reduced cost. These methods have the potential to test the "single stock" paradigm for highly migratory stocks and provide the technical foundation for global chain of custody and provenance systems necessary to improve accuracy of catch reporting and curb Illegal, Unregulated, and Unreported (IUU) fishing (Grewe et al. 2015). Australia's national research agency, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), has invested in approximately a decade of work in developing a suite of technological advancements including DNA profiling techniques and specialised laboratory processing protocols associated with sample handling, quality control and statistical analysis methods.

Using this technology, a project funded through the Fisheries Research Development Corporation on behalf of the Australian Government and the CSIRO (see Evans et al. 2021) has provided an improved understanding of the population structure and connectivity of five of the species caught in the ETBF (albacore, bigeye and yellowfin tunas, broadbill swordfish and striped marlin). Results from analyses of swordfish from the ETBF (Queensland and Norfolk Island) and New Zealand are provided, indicating little genetic structuring of individual sampled across these three sampling locations.

This paper provides an update to this analysis by reporting on the analysis of a larger dataset including further samples from New Zealand and the inclusion of two further sampling sites, eastern Tasmania and the Cook Islands.

Methods

Sample collection

Collection of samples of approximately 1gm of muscle tissue were either directly sampled from individual fish during onshore processing (Queensland, Norfolk Island, New Zealand, Cook Islands) or sub sampled from larger samples collected by recreational fishers (eastern Tasmania) and then placed into individual vials of RNALater. Minimum sample sizes for stock assignment collection of samples aimed to achieve 50 samples from each of two years for each species. The sampling strategy for the project aimed to include three spatially restricted locations, one from the ETBF and two sites within the western Pacific Ocean (see Evans et al. 2021 for details).

DNA extraction

For DNA extraction, muscle tissues were firstly cleaned by removing all external surfaces of each muscle tissue sample using a new scalpel blade for each individual. This removed potentially surface contaminated tissue from the sample, with use of a new scalpel blade per sample avoiding any further surface cross contamination of samples. Total genomic DNA was isolated from the cleaned tissue samples (approximately 15 mg in weight) was isolated using one of two protocols; either a

Machery Nagel Nucleo-Mag bead-based DNA isolation kit or a CTAB protocol, a Phenol-Chloroform based method described by Grewe et al. (1993). The bead-based extractions were performed on an Eppendorf EP-Motion-5075 robotic liquid handling station. Gel runs were visually inspected as a first-pass qualitative check of the quality of the DNA in each sample. Samples that were qualitatively assessed as containing minimal amounts of DNA or highly denatured DNA were removed and did not progress to sequencing.

Genetic sequencing

DNA extracts were shipped to Diversity Array Technologies in Canberra where DNA complexity reduction and library construction was performed prior to sequencing.

The sequencing protocols used incorporated a DArT-Seq proprietary next generation sequencing methodology. DArTseq™ represents a combination of DArT complexity reduction methods and next generation sequencing platforms (for detailed description see Grewe et al. 2015). This represents a new implementation of sequencing complexity with reduced representations and more recent applications of this concept on the next generation sequencing platforms. Similar to DArT methods based on array hybridisations, the technology is optimized for each organism and application by selecting the most appropriate complexity reduction method (both the size of the representation and the fraction of a genome selected for assays). Four methods of complexity reduction were tested (data not presented). DNA samples were processed in digestion/ligation reactions using a single *Pst*I-compatible adaptor with two different adaptors corresponding to two different Restriction Enzyme (RE) overhangs. The *Pst*I-compatible adapter was designed to include Illumina flow cell attachment sequence, sequencing primer sequence and "staggered", varying length barcode region. The reverse adapter contained a flow cell attachment region and a *Sph*I-compatible overhang sequence.

Only "mixed fragments" (*Pst*I-*Sph*I) were effectively amplified by PCR. PCR conditions consisted of an initial denaturation at 94°C for 1 min followed by 30 cycles of 94°C for 20 sec, 58°C for 30 sec and 72°C for 45 sec, with a final extension step at 72°C for 7 min. After PCR, equimolar amounts of amplification products from each sample of the 96-well microtiter plate were bulked and applied to cBot (Illumina) bridge PCR, followed by sequencing on an Illumina Hiseq2000. The sequencing (single read) was run for 77 cycles.

Sequences generated from each lane were processed using a proprietary DArTseq analytical pipeline (DArT-Soft14 version). In the primary pipeline, the FASTQ files were first processed to filter away poor-quality sequences, applying more stringent selection criteria to the barcode region compared to the rest of the sequence. In that way the assignments of the sequences to specific samples carried in the "barcode split" step was very reliable. Approximately 2,000,000 sequences per barcode/sample were identified and used in marker calling. Finally, identical sequences were collapsed into "fastqcall files". These files were used in the secondary pipeline for DArTseq PL's proprietary single nucleotide polymorphism (SNP) and SilicoDArT (presence/absence of restriction fragments in representation) calling algorithms (DArTsoft14).

Species identification

Confirmation of species identification was conducted using polymerase chain reaction – restriction fragment length polymorphisms (PCR-RFLP) analysis of a 1400bp region of the mitochondrial DBA

molecule as described by Innes et al. (1998). Size specific banding patterns representing restriction-fragment-length-polymorphisms (RFLPs) were resolved on 1.2% agarose gels using standard lab practices.

Quality control

A step wise process for data quality control using the package radiator (Gosselin 2017) was carried out at the individual markers and sample levels. Marker filtering includes marker reproducibility, identification of monomorphic markers, identification of common markers (these are markers that are present among all individuals), minor allele counts (which eliminates sequencing artefacts), minimum and maximum read depth (which is a reliability index of DNA quality and also identifies repetitive DNA which are not single copy genes – for example junk DNA in the genome), the proportion individuals that don't have a genotype at a locus, the quality of the sequencing run, the number of SNPs at a locus (addresses whether there are SNPs from different parts of the chromosome that have similar sequences) and whether loci comply with assumption of Hardy Weinberg equilibrium (Andrews 2010). Individual samples were filtered at three key steps: 1. missing data; 2. average heterozygosity; 3.removal of highly similar/duplicate genotypes.

Population modelling

Population modelling using a mixture model was based on the method outlined in Foster et al. (2018) and implemented in the R package stockR (Foster 2018). The model assumes that each sample belongs to one of K genetic groups ($K \geq 1$ and is an integer), and for this study the purpose of the analysis is twofold: 1) To assign a probability of assignment of each sample to each of K putative genetic groups and 2) to provide information about how many genetic groups there are likely to be. For a given K, the modelling approach uses maximum likelihood to estimate the allele frequencies within each population, and then uses model summaries (posterior membership probabilities) to estimate the chance of each sample belonging to each genetic group (Foster et al. 2018).

Two approaches are utilised in order to determine an appropriate K for the data of each species:

1. Information Criteria: Two information criteria (AIC and BIC) are calculated from the fitted model with the number of groups (K) that minimised the information criterion identified as providing the best fit.
2. Cross Validation (stability): 5-fold cross validation is used to evaluate how quickly the predictive performance of the model diminishes as more stocks were added. To obtain the cross-validation statistics B = 1000 holdout samples are used. Formally, this is not a cross-validation procedure, but it is closely related. It differs from cross validation as the target of prediction, the group assignment probability, is not observed within the data themselves. In their place, the predicted assignment probabilities from the model fitted to all the data combined is used as the prediction target. So, in this sense this analysis is looking at the *stability* of the assignment probabilities due to subsampling and permutation.

Uncertainty in the assignment probabilities was quantified in the models by using the Bayesian bootstrap methods described in Foster et al. (2018). The uncertainty is graphically portrayed, along with the results, using bar plots. Individual bars represent the probability of assignment of a fish to each genetic grouping (K) plotted as a stacked bar with different colours for each group. The amount of colour saturation of the plotted colour bar is taken from the amount of uncertainty in the estimated probability for that sample in that group. If an estimated probability is highly uncertain,

then the bar is (nearly) white, whereas if it is quite certain then the bar is plotted with a solid colour. The amount of uncertainty is quantified by the width of the 95% confidence interval – with an interval of 1 being the highest possible for a probability estimate.

It is important to note that the sampling regions are not used in this analysis, only in presentation. The only information included in the modelling process are the genetic data themselves. This means that the analysis does not intentionally seek spatially consistent groupings, but if there is a real spatial signal then this should show in any case.

Results

The spatial distribution of samples and numbers of samples included in the analysis is provided in Table 1 and Figure 1.

Quality control

DNA extracts of sufficient quality were obtained from all samples with all samples progressing to sequencing. One sample from the ETBF originally identified as broadbill swordfish was identified during the species identification confirmation step as likely a striped marlin and removed from further analysis. The number of samples removed at each of the post-sequencing quality control steps (missing data, genome-wide average heterozygosity, highly similar/duplicate genotypes) are detailed in Table 1.

Genome-wide average heterozygosity values were highest in broadbill swordfish samples from the ETBF, with the highest number of samples with average heterozygosity values above the overall mean occurring in samples from New Zealand. The overall average heterozygosity of samples from New Zealand was slightly higher than the overall average heterozygosity of samples from all other sites, although samples from all locations appeared to have similar average heterozygosity values (Figure 2). Samples from the ETBF and New Zealand above the confidence limit threshold were removed from further analysis (Table 1).

The majority of sequenced individuals identified as having similar genotypes were either DArT technical replicates (repeated DNA libraries used as statistical replicate samples). The remainder were replicate tissue samples included to evaluate potential differences caused by the two different extraction methods. A lack of samples showing relative genetic distances less than 75% indicates there were no observed fish related to each other at either full-sib or half-sib levels that could potentially violate assumptions of the population genetic analysis.

Population modelling

On the basis of AIC and BIC calculated by the model, the number of genetic groups (K) that minimised the information criteria was one (Figure 3). Cross validation assigned 100 percent of all samples to K=1. This dropped to 66.9 percent at K=2 and 49.3 percent at K=3. The probability of assignment was consistent amongst years in samples from the ETBF. Given the information criteria and cross validation results, the data do not support the discrimination of more than one genetic group of broadbill swordfish across the three sampling locations (ETBF, New Norfolk, New Zealand) and amongst the years examined.

Table 1. Broadbill swordfish samples collected/received by species and region with details of the number of samples removed through the species identification and quality control processes.

Region	Year	No. samples collected/received	No. samples removed due to poor DNA quality	No. genotypes produced (incl. technical replicates)	Final no. samples (after data QC)
Australia	2015 (east Tas)	13	—	13	11
	2016 (east Tas)	7	—	7	7
	2016 (Norfolk Island)	45	—	106	76
	2017 (ETBF)	48	—	51	45
	2018 (ETBF)	31	—	31	31
Cook Islands	2019	11	—	11	10
	2020	13	—	13	12
New Zealand	2020	58	—	82	49
	2021	42	—	58	30

Discussion

The analysis presented here build on previous tagging studies to provide further insights into the connectivity of swordfish in the ETBF with adjacent waters and areas within the greater western and central Pacific Ocean (Evans et al. 2014; Tracey and Pepperell 2018).

From the perspective of the WCPFC Area, the results suggest a sufficient level of connectivity and mixing between each of the areas investigated that results in little discernible genetic differentiation observed. Consistency in results across years suggest that the groupings revealed here have some temporal stability across years across most sites (note samples collected from Norfolk Island constitute one year only). However, given the low numbers of samples within years examined from eastern Tasmania and the Cook Islands, the results presented should be considered preliminary. Further sampling from these locations would also be required for resolving the indicative nature of the results.

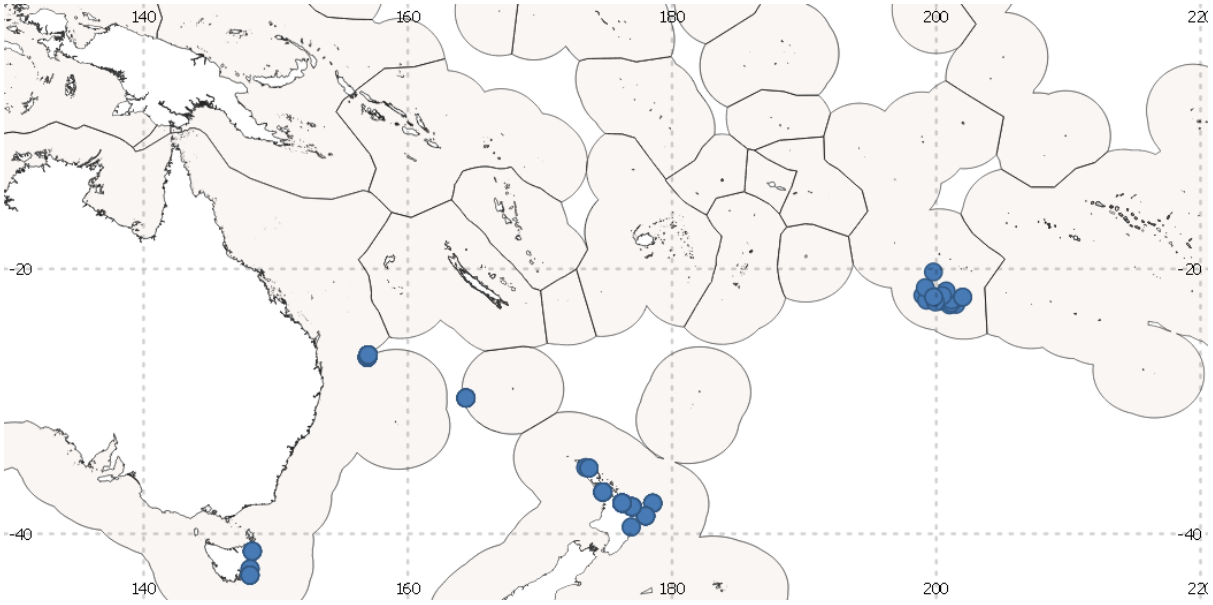

Figure 1. The spatial distribution of samples collected from broadbill swordfish. Each dot represents either the location of the fishing event in which the sample was collected (where the exact fishing location was available) or the centre of the area fished (represented by boundary coordinates provided by the fishing company) in which multiple samples were collected (where samples were collected at the processing factory on land and could not be attributed to individual fishing events).

It should be noted that these results *only apply to the sites* included for each of the species in this study and therefore *cannot* be extrapolated across the wider western and central Pacific Ocean region with any certainty. Further sampling and analysis of samples from additional sites across the western and central Pacific using the same methods used in this study would be needed to investigate whether the results presented here extend to other locations across the western and central Pacific region or whether greater genetic differentiation is discernibly present across the region.

While investigations of the population structure of broadbill swordfish globally have identified molecular variation between ocean basins (Alvarado Bremer 1996; Rosel and Block 1996; Kotoulas et al. 2007), only a small number of studies have investigated structuring within ocean basins to date. Of those that have, some indication of structuring has been observed in the Pacific Oceans at the basin-scale, with low levels of mitochondrial gene flow. This gene flow appears to have a ⊃-shaped pattern, with connectivity of animals east-west in the Northern and Southern Hemispheres and connections across the equatorial zone along the west coast of the Americas (Reeb et al. 2000). This is consistent with larval distributions (Grall et al. 1983; Nishikawa et al. 1985) and the hypothesis of separate stocks in the north and southwest Pacific Ocean (Sakagawa and Bell 1980). DNA sequence polymorphisms from swordfish collected across the Pacific Ocean identified fish from the south-east Pacific Ocean as genetically different to all other locations sampled (Alvarado-Bremer et al. 2006). In addition, fish from the north-east Pacific Ocean were observed to be different to those collected from around Hawai'i, which were in turn differentiated from those in the south-west Pacific Ocean (Alvarado-Bremer et al. 2006). Mitochondrial differentiation has been reported from samples collected from waters to the east of Australia and north of New Zealand, although this differentiation was not significant after post-hoc correction (Kasapidis et al. 2008).

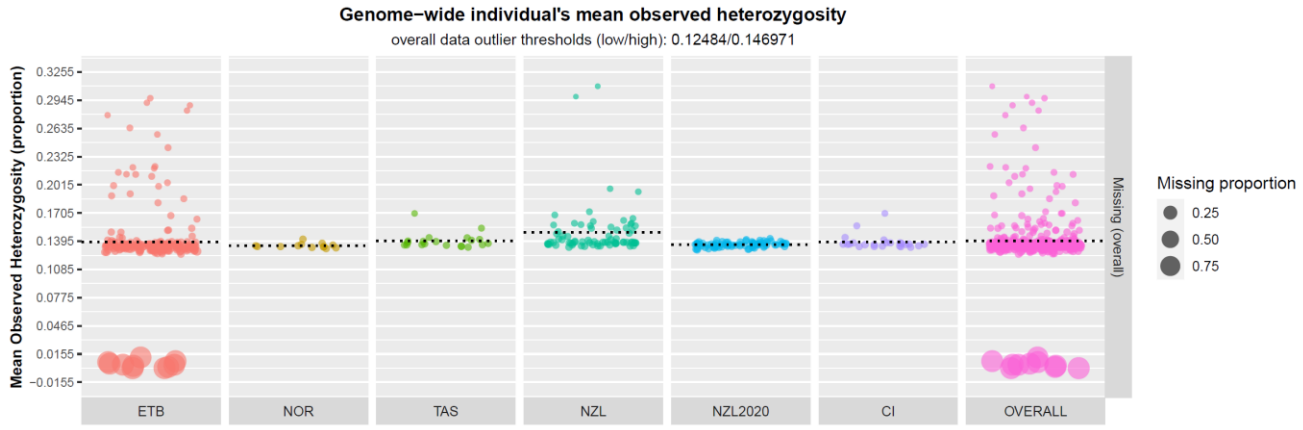

Figure 3. Genome-wide mean observed heterozygosity of samples from broadbill swordfish from the ETBF (ETB), Norfolk Island (NOR), eastern Tasmania (TAS), New Zealand (NZL, NZL2020) and the Cook Islands (CI). Dashed horizontal lines indicate the average heterozygosity for each dataset. Statistical threshold values of higher and lower confidence limits above and below which samples were removed were 0.12484 (low) and 0.146971 (high).

Our results suggest there is little genetic differentiation between swordfish caught at the locations sampled. Further sampling across the WCPFC area would be needed in order to determine a comprehensive understanding of the population structure of swordfish, and in particular if there is structuring occurring at spatial scales smaller than those observed in Reeb et al. (2000) and Alvarado Bremer et al. (2006), substantive spatial sampling across the western and central Pacific and some temporal duplication of sampling is required. This will require dedicated efforts placed towards the development, design and carrying out of such a program.

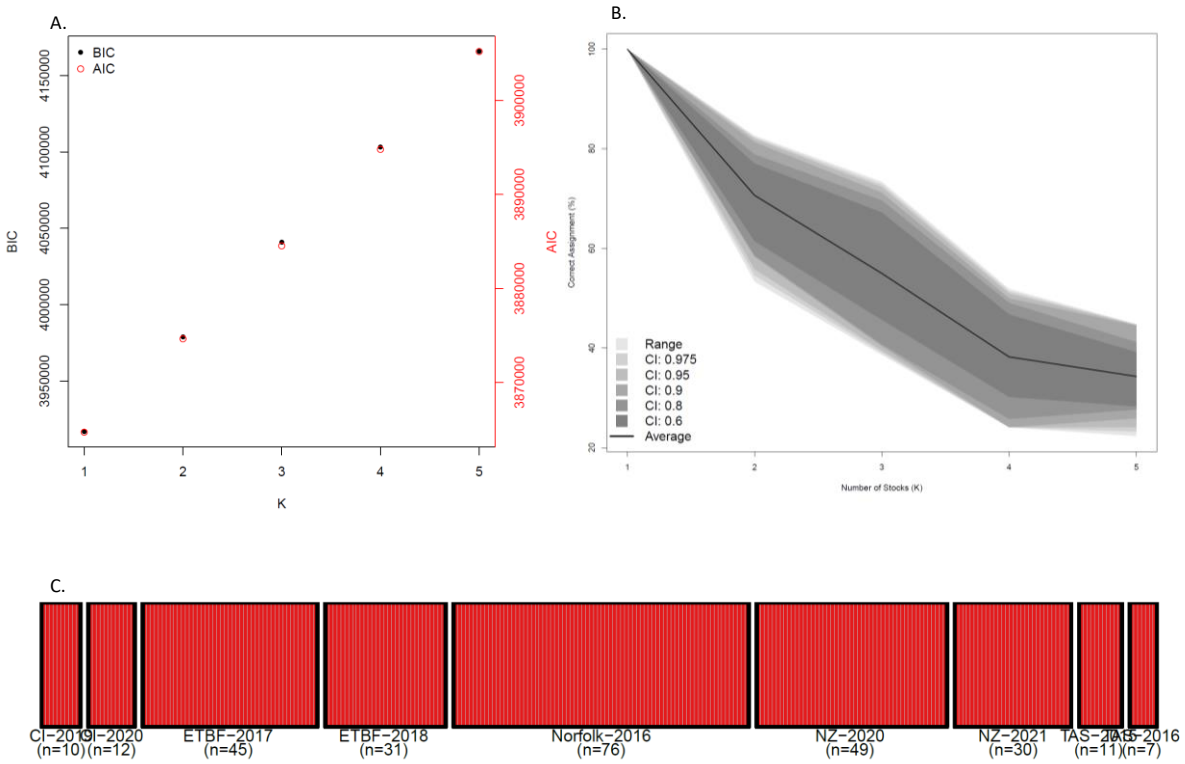

Figure 3. Output of the mixture model for broadbill swordfish. A. Information criterion given K=1 to K=5. Note the scales for each information criterion varies. B. Cross validation and percent assignment of markers on the basis of K=1 to K=5. C. The probability of individual assignment to K=1 CI: Cook Islands, ETBF: Eastern Tuna and Billfish Fishery, Norfolk: Norfolk Island, NZ: New Zealand, TAS: Tasmania.

Acknowledgements

The project team would like to thank Thierry Gosselin for his assistance in the application of the package “Radiator” on this dataset. Gary Heilman (De Brett Seafoods) and Pavo and Heidi Walker (Walker Seafoods Australia) are thanked for their assistance in accessing samples from the Eastern Tuna and Billfish Fishery. The Tropical Tuna Resource Assessment Group are thanked for providing useful feedback and suggestions to the project.

The project 2016-018 Determination of the spatial dynamics and movement rates of the principal target species within the ETBF and connectivity with the broader WCPO – beyond tagging is supported by funding from the FRDC on behalf of the Australian Government and funding from CSIRO.

References

Alvarado Bremer JR, Mejuto J, Greig TW, Ely, B. 1996. Global population structure of swordfish (*Xiphias gladius* L.) as revealed by analysis of the mitochondrial DNA control region. Journal of Experimental Marine Biology and Ecology 197: 295 – 310.

Alvarado Bremer JR, Stequert B, Robertson NW, Ely B. 1998. Genetic evidence for inter-oceanic subdivision of bigeye tuna (*Thunnus obesus*) populations. Marine Biology 132: 547 – 557.

Campbell R, Hobday A. 2003. Swordfish – Seamount – Environment – Fishery Interactions off Eastern Australia. Working Paper BBRG-3 presented to the 16th meeting of the Standing Committee on Tunas and Billfish, July 9-17, 2003 Mooloolaba, Australia.

Evans K, Abascal F, Kolody D, Sippel T, Holdsworth J, Maru P. 2014. The horizontal and vertical dynamics of swordfish in the South Pacific Ocean. Journal of Experimental Marine Biology and Ecology 450: 55 – 67. doi:10.1016/j.jembe.2013.10.025.

Evans K, Grewe P, Foster S, Gunasekara R, Lansdell M. 2021. Determination of the spatial dynamics and movement rates of the principal target species within the eastern tuna and billfish fishery and connectivity with the broader western and central Pacific Ocean - beyond tagging. CSIRO Oceans and Atmosphere and the Fisheries Research and Development Corporation. Available at: https://frdc.com.au/project/2016-018.

Foster SD. 2018. stockR: Identifying Stocks in Genetic Data. R package version 1.0.68.

Foster SD, Feutry P, Grewe PM, Berry O, Hui FKC, Davies CR. 2018. Reliably discriminating stock structure with genetic markers: Mixture models with robust and fast computation. Molecular Ecology Resources 18: 1310 – 1325.

Gosselin T. 2017. radiator: RADseq Data Exploration, Manipulation and Visualization using R.

Grall C, De Silva DP, Houde ED. 1983. Distribution, relative abundance and seasonality of swordfish larvae. Transactions of the American Fisheries Society 112: 235–246.

Grewe PM, Feutry P, Hill PL, Gunasekera RM, Schaefer KM, Itano DG, Fuller DW, Foster SD, Davies CR. 2015. Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource. Nature Scientific Reports 5: 16916.

Grewe PM, Kreuger CC, Aquadro CF, Bermingham E, Kincaid HL, May B. 1993. Mitochondrial DNA variation among lake trout (*Salvelinus namaycush*) strains stocked into Lake Ontario. Canadian Journal of Fisheries and Aquatic Science 50: 2397 – 2403.

Innes HB, Grewe P, Ward R. 1998. PCR-based genetic identification of marlin and other billfish. Marine and Freshwater Research 49: 383 – 88. doi:10.1071/MF98007.

Kasapidis P, Magoulas A, García-Cortés B, Mejuto M. 2008. Stock structure of swordfish (*Xiphias gladius*) in the Pacific Ocean using microsatellite DNA markers. Working paper WCPFC-SC4-2008/BI-WP-4 presented to the Western and Central Pacific Fisheries Commission Scientific Committee fourth regular session, 11 – 22 August 2008, Port Moresby, Papua New Guinea.

Kotoulas G, Mejuto J, Antoniou A, Kasapidis P, Tserpes G, Piccinetti C, Peristeraki P, García-Cortés B, Oikonomaki K, de la Serna JM, Magoulas A. 2006. Global genetic structure of swordfish (*Xiphias gladius*) as revealed by microsatellite DNA markers. ICCAT Collective Volume of Scientific Papers 61: 79 – 88.

Nishikawa Y, Honma M, Ueyanagi S, Kikawa S. 1985. Average distribution of larvae of oceanic species of scombrid fishes, 1956 – 1981. Far Seas Fisheries Research Laboratories S Series 12: 1 – 99.

Reeb CA, Arcangeli L, Block BA. 2000. Structure and migration corridors in Pacific populations of the swordfish *Xiphias gladius*, as inferred through analyses of mitochondrial DNA. Marine Biology 136: 123 – 131.

Rosel PE, Block BA. 1996. Mitochondrial control region variability and global population structure in the swordfish, *Xiphias gladius*. Marine Biology 125: 11 – 22.

Sakagawa GT, Bell RR. 1980. Swordfish, *Xiphias gladius*. Pages 43 – 55 in Shomura RS (ed.). Summary report of the billfish stock assessment: Pacific resources. NOAA technical Memorandum NMFS-SWC5. US Department of Commerce, Seattle.

Tracey S, Pepperell J. 2018. Understanding the movement, behaviour and post-capture survival of recreationally caught swordfish from southeast Australia – a pilot study. FRDC Project No 2015/022. University of Tasmania, Hobart.