SCIENTIFIC COMMITTEE
FIFTEENTH REGULAR SESSION

Pohnpei, Federated States of Micronesia
12-20 August 2019

Connectivity of tuna and billfish species targeted by the Australian Eastern Tuna and Billfish Fishery with the broader Western Pacific Ocean

WCPFC-SC15-2019/SA-IP-13

Evans, K., Grewe, P., Foster. S, Gosselin, T., Gunasekera, R. and Lansdell, M.

Connectivity of tuna and billfish species targeted by the Australian Eastern Tuna and Billfish Fishery with the broader Western Pacific Ocean.

Karen Evans[1], Peter Grewe[1], Scott Foster[2], Thierry Gosselin[3], Rasanthi Gunasekera[1], Matt Lansdell[1].

[1]CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart Tasmania 7001, Australia
[2]Data 61, CSIRO, GPO Box 1538, Hobart Tasmania 7001, Australia
[3]Department of Biology, L'Université Laval, Ville de Quebec, 2325 Rue de l'Université, Québec, QC G1V 0A6, Canada

Abstract

This paper provides an update on a three-year project investigating the connectivity of tuna and billfish species caught in the fishery that operates along the east coast of Australia with those in adjacent waters and those further east in the Pacific Ocean. The project is funded through the Fisheries Research Development Corporation on behalf of the Australian Government and the CSIRO. Samples examined for this study came from three sources: 1. the Western and Central Pacific Commission's Tropical Tuna Tissue Bank (from New Caledonia, Fiji and the Marshall Islands); 2. historical samples collected by CSIRO and Indonesia (from Australia, New Zealand and the Solomon Islands); and 3. contemporary samples collected by CSIRO, independent researchers and industry from Australia and New Zealand. DNA extractions were assayed using double digest RAD sequencing techniques followed by quality assessment incorporating a multi-stage Quality Control (QC) analysis approach. A step-wise quality control analysis of albacore and yellowfin tuna sequencing data revealed varying levels of DNA cross contamination across all sampled locations. Following quality control analysis data from samples of sufficient quality (at numbers allowing for rigorous population analysis) were input into mixture models. Model results failed to provide evidence for multiple populations across the sampled regions for both species. Based on our results, we discuss potential for future analyses and suggestions for utilising tissues sampled from the WCPFC Tropical Tuna Tissue Bank in the context of genetic studies.

Background

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 yellowfin, bigeye and albacore tuna, swordfish and striped marlin. Populations of these species are known to extend well beyond the Australian Exclusive Economic Zone (EEZ) and 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. Populations are currently assessed as a single inter-connected stock distributed across the wider western and central Pacific Ocean or South Pacific Ocean and are managed at the international level under the auspices of the Western and Central Pacific Fisheries Commission (WCPFC).

Although populations are assessed as single interconnected stocks, biological information on growth rates and reproduction, movement data derived from tagging studies and spatial and temporal variability in catches of these species suggest that there is likely to be some structure to stocks throughout the WCPFC region. More recently, both traditional and next generation high throughput genotyping methods have provided evidence of population structure in yellowfin tuna across the Pacific (e.g. Aguilar et al. 2015; Grewe et al. 2015) and provide some support to the hypothesis that

yellowfin tuna fished by Australia's tuna fisheries may be a localised stock within the Coral and Tasman Sea region. If yellowfin tuna or the other principal species occurring in the ETBF do comprise localised stocks, this has implications for current consideration of species within stock assessments conducted by the WCPFC (that currently consider most species to comprise a single stock) and associated management of species both within national and regional contexts.

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. 2016). 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 three-year project funded through the Fisheries Research Development Corporation on behalf of the Australian Government and the CSIRO (see Evans et al. 2016; 2017; 2018) aims to provide an improved understanding of the population structure for five of the species caught in the ETBF (albacore, bigeye and yellowfin tunas, broadbill swordfish and striped marlin). The project also aims to establish the connectivity of the five species within the broader WCPFC region.

This project builds on previous studies conducted by the CSIRO that have documented genetic structure in yellowfin across three locations in the western, central, and eastern Pacific Ocean and is part of a broader program of work being conducted by CSIRO on the stock structure of pelagic and neritic species across the Indian and Pacific Oceans (Grewe et al. 2016; Grewe et al. 2019). Outputs from these projects are expected to provide essential information required for the assessment and management of marine species and in particular tuna and billfish species within the two ocean basins.

Methods

Sample collection

Using the output of a spatial assessment of tissue samples for tropical tuna and billfish species held in the WCPFC Tissue Bank and historical samples held by CSIRO, key areas where samples are available for stock structure analyses of yellowfin, bigeye and albacore tunas were identified and an application to access these samples submitted to the WCPFC. Where samples currently held in collections did not meet the experimental design requirements for resolving stock structure (e.g. striped marlin, swordfish), the feasibility of further sampling to resolve spatial gaps and/or inadequate numbers was explored. Within the ETBF, collection of additional samples to those held in CSIRO archives was conducted via sampling of fish during onshore processing. External to the ETBF, collection of samples has been undertaken by project collaborators as part of routine operations. 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 Table 1).

DNA extraction

Total genomic DNA 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 aliquots 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). For the purposes of the study in which the WCPFC samples were contributing to (see Evans et al. 2018), only co-dominant SNP-DArT markers were used for population analysis.

Species identification

Identification of swordfish, and striped marlin were confirmed using mitochondrial tests described by Innes et al. (1998). Identification of albacore, bigeye, and yellowfin tuna species were confirmed following restriction digestion of a mitochondrial PCR amplicon (PCR-RFLP) as described by Chow and Inoue (1993) with further modifications described by Takayama et al. (2001). Size specific banding patterns representing restriction-fragment-length-polymorphisms (RFLPs) for all five species 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 (as opposed to an admixture model) was based on the method outlined in Foster et al. (2018) and implemented in the package stockR (Foster 2018). The model assumes that each sample belongs to one of K populations ($K \geq 1$ and is an integer), and the purpose of the analysis is twofold: 1. to choose K and, 2. to assign each sample to one of the populations.

Three approaches are utilised in order to determine the best fit to the model given varying values of K:

1. Information Criteria: Two information criteria (AIC and BIC) are calculated from the fitted model with the number of stocks (K) that minimised the information criterion identified as providing the best fit.
2. Cross Validation: 5-fold cross validation was used to evaluate how quickly the predictive performance of the model diminished as more stocks were added. To obtain the cross-validation statistics B = 1000 holdout samples were used.
3. Bootstrapped Confidence Intervals: confidence intervals associated with group membership increase as the model become over-fitted and the certainty to which population a sample is assigned decreases. To obtain confidence intervals B = 1000 bootstrap samples were used.

Progress to date

The spatial sampling structure for the project and samples included in the project based on historical and contemporary collection of samples and the current state of collection and analysis is provided in Table 1. Given the nature of the collections from which samples were derived, samples comprised a mix of sexes, lengths (albacore: 48 – 100cm, bigeye: 31 - ~150 cm, yellowfin: 88 - ~150 cm; striped marlin and swordfish yet to be completed) and therefore age classes/cohorts and potentially

reproductive state. Full analysis of albacore and yellowfin tunas has been completed and will be presented here. Remaining species analysis is yet to be completed.

Genetic sequencing

DNA extraction and DNA profiling, using the DArTseq™ technique, has either been completed or is underway for all samples in hand (see also Table 1).

Quality control processes

Sample quality

Examination of gel runs on DNA extracted from yellowfin tuna samples initially received from the WCPFC Tuna Tissue Bank identified 49 samples that had degraded to the point that very little high molecular weight DNA could be extracted, which is necessary for the DArTseq™ technique. As a result, replacement of poor quality individuals by others in the WCPFC Tuna Tissue Bank was required. Degradation of DNA in tissue samples can occur for a number of reasons including from poor care of fish from which samples are collected (e.g. market fish left exposed to the sun), poor handling of samples on vessels (e.g. samples left out on the deck) or during transit from vessels to archives (e.g. thawing of samples during transit) and repeated freeze thaw cycles that may occur as a result of multiple subsampling of tissues). Metadata associated with those tissues in which DNA degradation had occurred suggest a mixture of these factors likely contributed to the poor quality of samples archived. The catch location of a further one sample was unable to be confirmed by Pacific Community staff. As a result, 50 samples did not progress to genetic sequencing (see Table 1).

Table 1. Spatial structure and status of project samples.

Species	Location	Years	Status
Albacore tuna	ETBF	2	collected and sequenced
	New Caledonia	2	collected and sequenced
	New Zealand	2	collected and sequenced
Bigeye tuna	ETBF	2	collected and sequenced y
	Marshall Islands	2	collected and sequenced
	Solomon Islands	2	collected and sequenced
Striped marlin	ETBF	2	collected, sequencing underway
	New Zealand	2	collected and sequenced
	Hawaii	2	collected, sequencing underway
Swordfish	ETBF	2	collected and sequenced
	NZ	1	collection underway
Yellowfin tuna	ETBF	2	collected and sequenced
	Fiji	2	collected and sequenced
	Marshall Islands	2	collected and sequenced

Species identification

Species identification identified six samples where species had been misidentified or mislabelled. Four samples of yellowfin tuna from the Marshall Islands were misidentified/mislabelled as bigeye tuna and two samples of bigeye tuna from the Marshall Islands were misidentified/mislabelled as yellowfin tuna. All misidentified/mislabelled samples were derived from different sampling events.

Quality control of sequencing data

The quality control steps at which samples were removed and the numbers removed at each step are detailed in Table 2.

Missing data

If an individual is missing data above a threshold as a result of poor-quality DNA, they are removed. A total of 16 albacore tuna samples from the ETBF, six albacore tuna samples from New Zealand, eight yellowfin tuna samples from the ETBF and one yellowfin tuna sample from Fiji were removed at this step.

Genome wide average heterozygosity

The position at which a SNP occurs on a chromosome is called the locus (plural loci). Because SNPs are bi-allelic, they contain two alleles at each locus. Heterozygosity is a measure of how many loci contain two different alleles (heterozygous genotype) versus how many loci have two identical alleles (homozygous genotype). On average individuals within a population will have the same level of heterozygosity as each other. However, if the heterozygosity observed for the DNA profile of an individual deviates from this average then this likely reflects sample cross contamination – introduced at the point of sampling, during handling or during subsampling – and often is the symptom of poor tissue sampling skills or inadequate cleaning protocols (e.g. not cleaning the knife or scalpel blade in between samples, not cleaning hands when handling multiple samples). Conversely, samples with lower than average heterozygosity are likely an indication of poor DNA quality that results in a homozygous excess as a result of introduced artefactual sequencing bias. An important step in assessing the quality of samples is therefore to identify samples that are either too homozygous or too heterozygous compared to the average observed level of heterozygosity. To do this, the level of genome-wide mean heterozygosity is calculated. For the current study, individual samples with a mean heterozygosity above and below statistical threshold values of higher and lower confidence limits and are filtered out of datasets for further quality control.

Of albacore tuna samples, a total of 22 samples from the ETBF, 30 samples from New Caledonia and ten samples from New Zealand were removed due to above average heterozygosity. Of yellowfin tuna samples, a total of 15 samples from the ETBF, 45 samples from Fiji and five samples from the Marshall Islands were removed at this step (Table 2).

Similar genotypes

Genetic similarity is used to identify individuals that are closely related where more closely related individuals show higher levels of genetic similarity and by extension show lower levels of genetic distance between them relative to average genetic distance between unrelated pairs. In essence, non-related individuals should have genotypes that are dissimilar (because they have no common relatives to derive their genes from). However, when cross-contamination or technical mishaps occur (e.g. labelling two samples collected from the same individual as different animals), samples with similar or almost identical genotypes can occur among individuals sampled from a population. Care needs to be taken in examining individuals with similar genotypes to determine if values of genetic distance are reflective of relatedness or the result of human error.

In addition, the sequencing process includes a number of technical replicates. These are included to examine the repeatability of sequencing results and so therefore need to be removed prior to any further analysis of sequencing results.

Table 2. Samples genotyped and removed through the species identification and quality control processes.

Species	Region/EEZ	Year	Number samples received	Number genotyped (incl. replicates)	DNA quality (qualitative)	Species ID	Missing data	Heterozygosity	Similar genotypes
Albacore	Australia	2009	50	73	—	—	16	1	15 (all technical replicates)
		2010	50	50	—	—	6	—	—
Albacore	New Caledonia	2013	25	36	—	—	—	1	8 (all technical replicates)
		2014	37	44	—	—	—	—	7 (all technical replicates)
		2016	45	64	—	—	—	30	10 (all technical replicates)
Albacore	New Zealand	2008	47	63	—	—	6	5	14 (12 technical replicates)
		2010	47	47	—	—	—	5	2
Yellowfin	Australia	2006	52	65	—	—	8	15	9 (all technical replicates)
		2013	50	77	—	—	—	—	33 (all replicate samples or technical replicates)
Yellowfin	Fiji	2014	62	77	9	—	—	30	14 (13 technical replicates)
		2015	60	39	25		1	15	3 (two technical replicates).
Yellowfin	Marshall Islands	2014	63	51	13	3	—	3	2 (one technical replicates)
		2015	52	60	2	1	—	2	10 (nine technical replicates)

Of samples sequenced, the majority of individuals identified as having similar genotypes were either technical replicates or replicate samples (included to examine potential differences caused by the two different extraction methods). Of those remaining, four albacore tuna samples from the ETBF, and four yellowfin tuna samples, two from Fiji and two from the Marshall Islands were removed. Examination of the genetic distance between pairs and the metadata associated with each identified that these were unlikely related individuals, but rather similarities were caused by either cross contamination or human error (e.g. spreadsheet/database errors or sample mislabelling).

Population modelling

For the two species that analysis has been finalised (albacore and yellowfin tuna) none of the three methods utilised could confirm the presence of more than one genetic population present in samples from ETBF and the two other sites investigated for each species (Figures 1 and 2). The two information criteria examined (AIC and BIC) were minimised at a value of $K = 1$, cross validation was capable of assigning all samples at $K = 1$ and at higher values of K rapidly declined and confidence intervals associated with group membership increased substantially at values of $K > 1$. This suggests a level of connectivity between and mixing of fish that are caught in the three areas that results in little discernible genetic differentiation by the approach utilised here. It must be noted that these results *only apply to the three sites* included in this study and therefore *cannot* be extrapolated across the wider western Pacific Ocean region.

Samples from bigeye and yellowfin tunas from a larger number of sites are currently the subject of ongoing investigations under a number of projects across the Indian and Pacific Oceans by the CSIRO and collaborating institutions and agencies (see Grewe et al. 2019). Results to date suggest that both bigeye and yellowfin data demonstrate genetic differentiation across an area extending from the central Indian to Eastern Pacific Ocean region and indicate that at these larger scales (larger than those examined in this study), some degree of restriction to genetic connectivity exists.

The mixed nature of the samples defined the questions that could be proposed and investigated by this project, namely “does the genetic signature of fish sampled from the three sites vary to the extent that they can be identified as different”. This should not be confused with questions that might be related to the investigation of distinct spawning populations and evolutionary gene flow. Rather, the samples and methods applied here provide some insights into contemporary mixing of individuals on the fishing grounds from which samples were derived.

In order to determine if multiple spawning populations for species exist, sampling would need to be structured in such a way that actively spawning fish (or those that are running ripe) from distinct locations are sampled at the same time and across a period from which temporal stability in results could be confirmed.

Next steps

Full analysis of sequencing data from bigeye tuna samples is underway. The final samples of striped marlin and swordfish are currently being collected and prepared for sequencing. Once sequenced, data will be analysed similarly to those samples presented here.

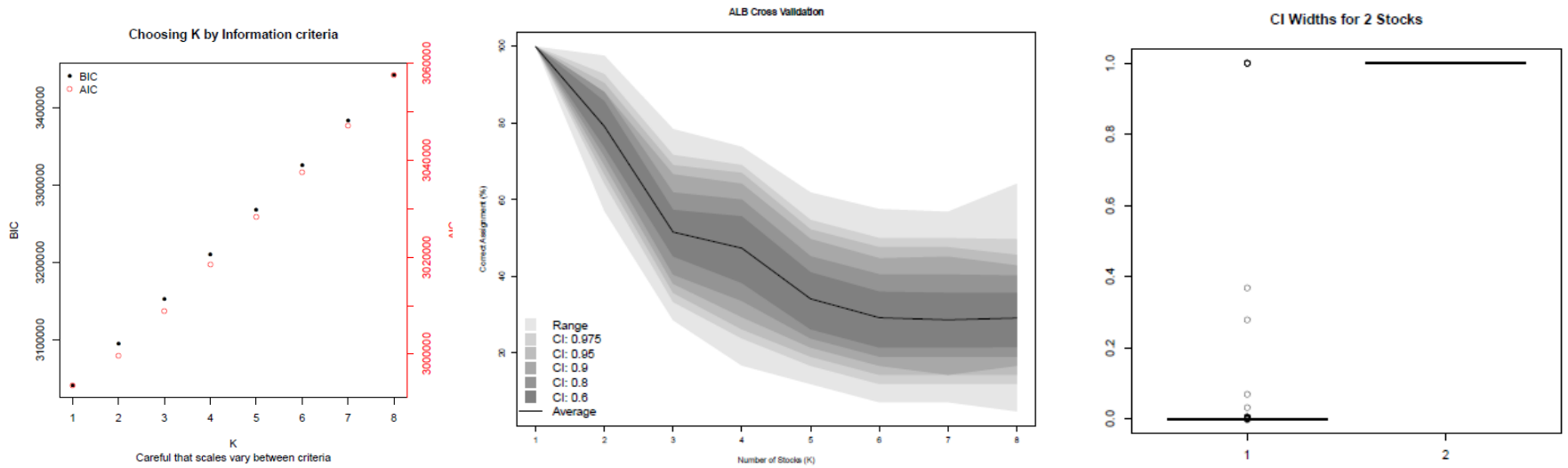

Figure 1. Model results for albacore tuna across the three approaches utilised. L-R: information criteria, cross validation, confidence intervals.

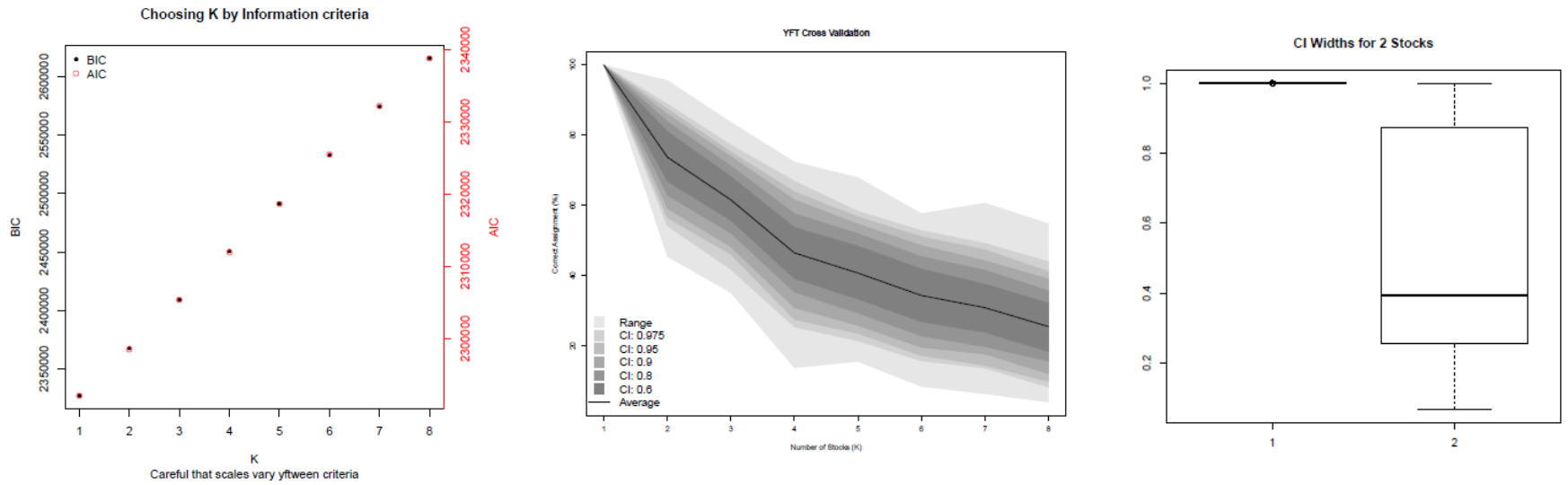

Figure 2. Model results for yellowfin tuna across the three approaches utilised. L-R: information criteria, cross validation, confidence intervals.

Reporting

Updates on the project in the form of information papers to the WCPFC scientific committee have been provided in 2016, 2017 and 2018 and a report on use of the samples from the Tissue Bank has been provided to the WCPFC Secretariat in 2018 and 2019. We anticipate providing the WCPFC scientific committee with results from the project in the second half of 2019. A final report will be produced for submission to the Australian Government Fisheries Research Corporation in the second half of 2019 and an associated peer review publication produced.

Intended outcomes

The information provided by this project will enable improved population structure and mixing considerations used in domestic and regional pelagic fisheries scientific advice and management. Conducting stock assessments and implementing management on spatial units that reflect the underlying biology of the population structure should reduce the risk of over-fishing smaller and less productive stocks, while potentially enabling higher exploitation of larger and more productive stocks. In the Australian domestic context, this will allow for the updating of the harvest strategy currently used in the management of the ETBF with operating models that have increased accuracy and precision.

Suggestions for future planning in association with the WCPFC Tuna Tissue Bank

Large scale investigations focused on establishing the presence of stock structure across fisheries based on sequencing technologies require three key requirements of samples to be met in order to ensure rigour to results:

(i) Adequate sample sizes
(ii) Establishment of temporal stability in results
(iii) Identification of the provenance of samples

Power analysis carried out by CSIRO (unpublished) suggests that in order to maximise assignment rates for stock structure discrimination, sample collections should aim for a minimum of 50 fish from each location. Furthermore, each sample collection should be obtained from two time points separated by a minimum of 12 months to ensure that any observed spatial differentiation is not a result of a random sampling artefact. Sampling across multiple years also establishes whether any observed spatial differentiation is temporally stable. Finally, the provenance of samples identified from a particular location should be ensured in order to avoid introducing “false” or additional assignments to locations not being considered by the study. This requires, particularly in the case of sampling from fish markets, a knowledge of where fishers providing fish to the market have been fishing and any tracking of transhipment processes.

The WCPFC Tropical Tuna Tissue Bank relies on samples collected under country observer programs, each with varying priorities associated with data and sample collection, aligned with each country’s fisheries management processes, plans and capacities. Establishment of the WCPFC Tropical Tuna Tissue Bank and the regular collection of samples contained and being contributed by country members of the WCPFC is a major achievement of the WCPFC - without the efforts placed into the archive to date, projects such as this would not be achievable. Spatial analysis of the tissue samples in the Tropical Tuna Tissue Bank has however, identified a number of areas that could potentially be focused on to better optimise the utility of the archive for investigations of species stock structure and population connectivity across the WCPFC:

I. Species coverage

Tissues currently contained in the WCPFC Tropical Tuna Tissue Bank largely reflect the composition and quantities of species caught across the WCPFC area. There is however, a distinct lack of samples currently archived from billfish species and some of the other species assessed under the WCPFC (e.g. shark species). Greater focus on these species and an associated increase in samples from these species would allow for the utilisation of samples in establishing currently uncertain life history parameters (e.g. age and growth) as well as building sample collections for use in investigations of stock structure and population connectivity.

II. Spatial coverage

Tissues currently contained in the WCPFC Tropical Tuna Tissue Bank to some extent reflect the distribution of the highest catches across the WCPFC Area. As a result, there are particular regions where samples are almost or completely non-existent. Greater focus on current spatial gaps in sample collection (including capacity development) would allow for more comprehensive spatial coverage of tissues archived, thereby facilitating spatial analyses of biological parameters as well as building sample collections for use in investigations of stock structure and population connectivity.

III. Sample sizes

Tissues currently contained in the WCPFC Tropical Tuna Tissue Bank, whilst impressive overall, rapidly decline in numbers once distributed on the basis of species, sample type and spatial and temporal qualifiers. In particular, the utility of the Tissue Bank declines for stock structure and population connectivity investigations, such as those being carried out by this project, where there is an aim to identify adequate samples from a defined region within a year and then across a number of years. This declines further when attempting to target samples from particular year classes or cohorts and align samples with gonad samples for the identification of spawning individuals. Greater focus on building tissue samples from a small number of regions across multiple years (these regions could vary through time) and from particular age classes and/or reproductive state would substantially increase the utility of the Tissue Bank for stock structure and population connectivity studies.

IV. Quality of samples

The quality of samples included in this study varied considerably with misidentification/mislabelling of species, samples of low DNA quality and cross-contamination identified across datasets. The incidence of these factors has flow on effects on overall sample numbers for analysis and highlights the need for first, strict sample collection and handling protocols and second, appropriate data quality control processes to be factored into studies. Without strict quality control measures, particularly those that allow for the identification of cross contamination, the potential for misinterpretation of data is increased. In this study sample sizes were reduced by as much as 53 individuals in any one year (yellowfin tuna from Fiji). Such large reductions in sample sizes reduce the robustness of any analyses, reducing confidence in results and have flow on impacts on the applicability of sample archives such as the WCPFC Tropical Tuna Tissue Bank for stock structure and population connectivity studies.

Acknowledgements

The authors would like to thank the members of WCFPC for the significant efforts placed into developing the WCPFC Tropical Tuna Tissue Bank and facilitation of this project. Pacific Community

Oceanic Fisheries Programme staff, in particular Neville Smith and Francois Roupsard are thanked for their assistance with accessing samples and associated information from the WCPFC Tuna Tissue Bank. Gary Heilman (DeBrett Seafoods) and Pavo and Heidi Walker (Walker Seafoods Australia) are thanked for their assistance in accessing samples from the ETBF, Julian Pepperell for advice and assistance in accessing samples from recreational fishers and John Holdsworth for assistance in accessing samples from New Zealand. This project is supported by funding from the Fisheries Research Development Corporation and CSIRO.

References

Aguilar, R.D., Perez, S.K.L., Catacutan, B.J.N., Lopez, G.V., Barut, N.C., Santos, M.D. (2015). Distinct yellowfin tuna (*Thunnus albacares*) stocks detected in Western and Central Pacific Ocean (WCPO) using DNA microsatellites. Working paper WCPFC-SC11-2015/SA-IP-11a prepared for the eleventh regular session of the Western and Central Pacific Fisheries Commission Scientific Committee, 5-13 August 2015, Pohnpei, Federated States of Micronesia.

Chow, S., Inoue, S. (1993). Intra- and interspecific restriction fragment length poly- morphism in mitochondrial genes of *Thunnus* tuna species. Bulletin of National Research Institute of Far Seas Fisheries 30: 207‑225.

Evans, K., Grewe, P., Kolody, D., Hillary, R., Foster, S., Timmiss, T., Lansdell, M., Hill, P. (2016). Connectivity of tuna and billfish species targeted by the Australian Eastern and Tuna Billfish Fishery with the broader western Pacific Ocean. Working paper SA-IP-15 prepared for the twelfth regular session of the Western and Central Pacific Fisheries Commission Scientific Committee, 3-11 August 2016, Bali, Indonesia.

Grewe, P.M., Feutry, P., Hill, P.L., Gunasekera, R.M., Schaefer, K.M., Itano, D.G., Fuller, D.W., Foster, S.D., Davies, C.R. (2015). Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource. Scientific Reports 5: 16916. doi: 10.1038/srep16916.

Grewe, P.M., Irianto, H.E., Proctor, C. H., Adam, M.S., Jauhary, A.R., Schaefer, K., Itano, D., Evans, K., Killian, A. Davies, C.R. (2016). Population structure ad provenance of tropical tunas: recent results from high throughput genotyping and potential implications for monitoring and assessment. Working paper SA-IP-01 prepared for the twelfth regular session of the Western and Central Pacific Fisheries Commission Scientific Committee, 3-11 August 2016, Bali, Indonesia.

Innes, H.B, 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.

Takeyama, H., Chow, S., Tsuzuki, H., Matsunaga, T. (2001). Mitochondrial DNA sequence variation within and between Thunnus tuna species and its application to species identification. Journal of Fish Biology 58: 1646‑1657. doi: 10.1006/ jfbi.2001.1575