OCEANS AND ATMOSPHERE
www.csiro.au

Final report: The Pilot SBT Gene-tagging Project

Ann L Preece, J Paige Eveson, Russell W Bradford, Peter M Grewe, Jorden Aulich, Matt Lansdell, Campbell R Davies, Scott Cooper, Jason Hartog, Jessica Farley, Mark Bravington and Naomi Clear.

August, 2018
Report to the Commission for the Conservation of Southern Bluefin Tuna
CCSBT-ESC/1809/06

Citation

Preece AL, Eveson JP, Bradford RW, Grewe PM, Aulich J, Lansdell M, Davies CR, Cooper S, Hartog J, Farley JH, Bravington MV and Clear NP. 2018. Final report: The Pilot SBT Gene-tagging Project. CCSBT-ESC/1808/8. CSIRO, Hobart Australia.

Contents

Acknowledgments

The Pilot SBT gene-tagging project was funded by the CCSBT and CSIRO, with additional voluntary contributions from the Australian Department of Agriculture and Water Resources to the CCSBT.

The CSIRO teams that developed the gene-tagging tool (Bradford et al., 2015), biopsy tips, the specific SNP markers and techniques needed for efficient, high quality, individual ID (Peter Grewe, Project Leader), funded through a series of strategic CSIRO research projects. These tools and techniques have made it possible and cost-effective to trial gene-tagging for SBT in this CCSBT pilot project.

2016 tagging at sea: Seaforce Marine's master and crew.

2017 tissue collection during harvest: Australian SBT Industry Association (ASBTIA) and members, and processing managers and staff of Australian Fishing Enterprises, Tony's Tuna International and Mori Seafood. Seatec for collection of tails during harvest operations.

For assistance with genetics lab work: Floriaan Devloo-Delva and Peta Hill.

Diversity Arrays Technology Pty Ltd.

The CSIRO genetics co-ordination project and database development team.

The CSIRO Science Engineering and Technical group: Andreas Marouchos, Andrew Filisetti, Dave Kube, Murray Campbell, Chris Blood and team.

The CSIRO SBT research team and research support staff.

Abstract

The southern bluefin tuna (SBT) pilot gene-tagging program commenced in 2016. The aims of the pilot study were to test the logistics and feasibility of large-scale gene-tagging of SBT and to provide a fisheries-independent estimate of absolute abundance of juveniles. A total of 3,768 fish were tagged and released in 2016. The number of fish tagged did not meet the original target of 5,000 fish, but it was possible to compensate for this by taking extra samples at harvest. A total of 16,490 tissue samples were collected during harvest in 2017, well in excess of the design study target of 10,000 samples. Protocols were refined for DNA digestion, robotic extraction and quality controls. The extracted DNA was sequenced using specifically designed SNP markers. Not all samples had DNA of sufficient quality or quantity and therefore not all samples were successfully sequenced. Fish with incomplete or poor genotype information (too few target SNP markers with good sequencing results) were excluded from the analysis.

In total, 3,456 fish were included in the tagged sample set, 15,391 fish were included in the harvest sample set, and a total of 22 recaptures were detected. The abundance of age 2 fish was estimated to be 2,417,786 with a CV of 0.21. The gene-tagging abundance estimate is close to the median estimate of age 2 fish in 2016 (2,102,853 fish aged 2 in 2016) from the 2017 stock assessment. Additional work, outside the scope of the pilot project, is underway to refine the length classes used for age 2 fish (for release samples) and age 3 fish (at harvest), and may result in revision of the data used in the analysis and final abundance estimate. The gene-tagging pilot project has demonstrated the technical feasibility and logistics of a large-scale genetic tagging program for SBT and its potential to provide an absolute abundance estimate for monitoring and management purposes. The CCSBT has commenced an on-going recruitment monitoring program using the gene-tagging method, funded by the CCSBT, CSIRO and the EU. Over 8,000 fish were tagged and released in 2017 and 2018 and tissue samples were collected from 15,000 fish harvested in June-July 2018. The recruitment estimates from this ongoing program will be used in the testing of candidate Management Procedures and future assessments of stock status.

Background

The southern bluefin tuna (SBT) pilot gene-tagging program commenced in 2016 and is now complete. The aims of the pilot study were to test the logistics and feasibility of large-scale gene-tagging for SBT and to provide a fisheries-independent estimate of absolute abundance of juveniles. The estimate of juvenile abundance will be used in the SBT operating models (OMs), candidate management procedures and to monitor rebuilding of the SBT stock. The experimental design, sample sizes, potential biases, costs and precision of estimates and examples of their use in stock assessment and management procedure models were explored in the 2015 CCSBT gene-tagging design study (Preece et al., 2015). Methods for incorporating these data in stock assessment models and a new management procedure have been further explored and developed (Hillary et al., 2016a & b). Simulation of these data has been included in the SBT projections code for Management Strategy Evaluation (Hillary et al., 2018a & b). Initial versions of candidate management procedures that include these data have been developed (Anon., 2018).

Gene-tagging SBT involves taking a very small tissue sample from a large number of 2-year-old SBT caught by pole and line fishing, releasing the fish alive, allowing 12 months for mixing with the wider untagged SBT juvenile population, and then taking tissue samples from the catch of 3-year-old fish at time of harvest from the farm operations in Port Lincoln, Australia. The two sets of samples are genotyped and then compared in order to find the number of fish with matching DNA; a match indicates that a tagged and released fish was recaptured. The abundance estimate is calculated from the number of samples in the release and harvest sets and the number of matches found.

The tagging component of the project was completed in March 2016 (Bradford et al., 2016). In 2017, tissue sampling from 3-year-old SBT from the commercial harvest and DNA extraction and sequencing was completed. A preliminary summary of harvest sampling activities was provided to the Extended Scientific Committee in 2017 (Preece et al., 2017). The data analysis and calculation of an abundance estimate was completed in April 2018. These data were provided to the CCSBT as part of the CCSBT scientific data exchange.

1 Method

The design study (Preece et al., 2015) considered several experimental design options and recommended, as the simplest design, that the pilot study should tag 2-year-old fish in the Great Australian Bight in year 1 and re-sample the same cohort at age 3 in year 2, allowing approximately 12 months for mixing. The tag and harvest sample sizes required to achieve an age 2 abundance estimate with (coefficient of variation (CV) of 25%) were calculated using estimates of the number of two-year-olds in the population from the 2014 stock assessment. The sample sizes for tagging and harvest sampling were optimised for a target CV and to minimise costs (which

are much higher for tagging fish at-sea than recapture sampling at harvest in the processing facilities in Pt Lincoln). Based on previous experience with conventional SBT tagging projects, twenty days at sea was considered the minimum viable sampling period to achieve the desired sample sizes that would allow for bad weather and poor fishing days. The design study recommended tagging and releasing 5,000 fish, and harvest sampling 10,000 fish.

The pilot project tested the methods, logistics and feasibility of the following steps:

1. Tag and release: Vessel charter and at-sea collection of tissue samples from age 2 fish in the Great Australian Bight during the summer of year 1 (2016).
2. Tissue collection during harvest: Collection of tissue sample from age 3 fish in year 2 (2017), during harvest of fish from farms which were caught by the Australian surface fishery.
3. Large-scale DNA extraction and genotyping of samples using CSIRO SNP marker panel.
4. Data analysis to identify matches and calculation of an abundance estimate. Provide the abundance estimate from the pilot gene-tagging program to the Extended Scientific committee for use in candidate Management Procedure and stock assessment models in 2018.

The design study noted extensions to the basic design that could be considered after the initial logistics had been tested and demonstrated to be cost-effective in the pilot tagging study. These extensions included tagging and resampling fish from regions outside the Great Australian Bight, tagging and resampling multiple age classes, and collection of otoliths to address uncertainties in the assignment of age classes of the fish sampled.

2 Results

2.1 Tag and release - tissue collection 2016

Results from the tagging component of the project were reported to the Extended Scientific Committee in 2016 (Bradford et al., 2016; Anon., 2016).

A total of 3,768 fish were tagged and released. Fish were tagged on 15 days in February/March 2016, across 14 locations, and from 37 schools. Some of the fish tagged were outside of the prescribed length range (Figure 1). All tagged fish with sufficient quality genotype data were included in the abundance estimate result reported here. The length range used to classify 2 year old fish is being reviewed and is discussed further below. A total of 519 fish were released without tagging because they were either outside of the target size range, presented with minor damage, or were flapping too vigorously to safely tag. These released fish were considered to be in good to medium condition and likely to survive a return to the water. There were 47 SBT assessed to be damaged and unlikely to survive, and these were euthanized. Biological samples were collected from these mortalities, including 33 sets of otoliths. A full trip report was provided to the CCSBT ESC (Bradford et al., 2016).

A fish that appeared to have been tagged several days earlier was recaptured and tagged for a second time. A photo of the first wound was taken to show the extent of healing (see Bradford et al., 2016). The genotyping has confirmed a match within the set of releases, only 1 of these 2 records is included in the abundance calculation.

The tagging component of the pilot study has demonstrated that it is possible to tag large numbers of fish by taking small tissue biopsies. The number of fish tagged did not meet the original target of 5,000 fish, but as described in the pilot design study (Preece et al., 2015), it is possible to compensate for this by taking extra samples at harvest. The CCSBT was advised that sufficient numbers had been tagged for the project to continue (Bradford et al., 2016).

The gene-tagging tool provided sufficient DNA from the tissue biopsy, even though the sample was smaller than a grain of rice (~15-20mg). The tool allowed for efficient sampling, with fish out of the water for around 20 seconds; quicker than conventional tagging methods, and considerably faster and less invasive than archival tagging methods. The gene-tagging tips allowed for collection of samples with good quality DNA, which was a major successful outcome of this large-scale trial, confirming initial results from experimental trials in farm cages (Bradford et al., 2015).

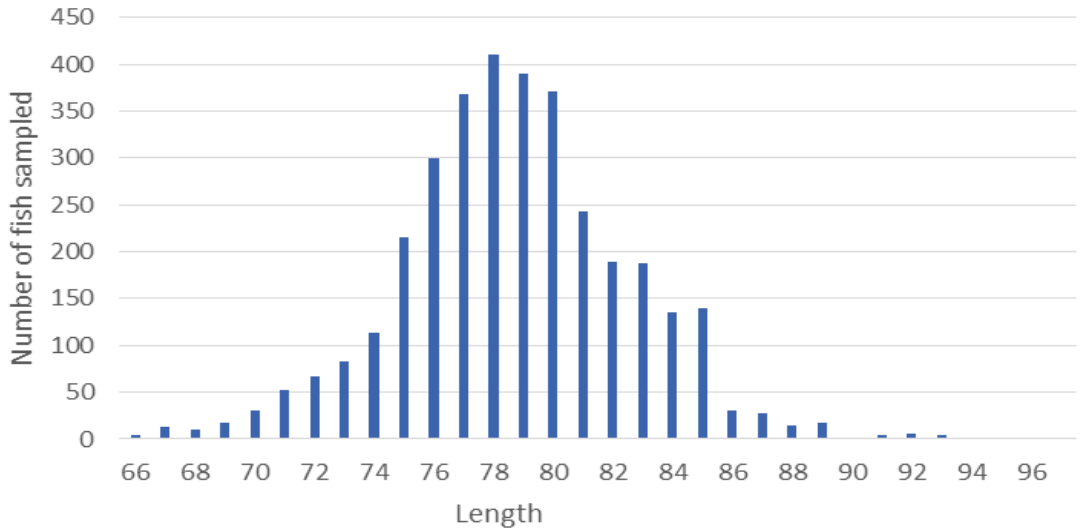

Figure 1. Length frequency distribution of tagged SBT from the pilot study in 2016 (n= 3,456). Only those with good genotype data are included.

2.2 Tissue collection during harvest in 2017

The method for collection of tissue samples during the commercial harvest from farms in Port Lincoln, South Australia, was developed during consultation with Industry representatives and members in 2016 and refined during the pilot project trials in 2017. The method needed to take into account: the speed, restricted space and Health and Safety requirements of the commercial processing activities; that on-shore processing occurs in three main factories; that a varying percentage of the fish are processed off-shore each year; the likely numbers of age 3 fish that

would be available through on-shore processing; commercial confidentiality of processing operations; and the need to not interrupt or slow the speed of commercial processing operations. Harvest sampling plans for 2017 were developed using information on lengths and numbers of fish processed in recent years, to organise the expected number of days required for sampling and the distribution of samplers across the factories. Analysis of the size distributions of fish processed onshore and offshore indicated that there was no need to stratify sampling between these two modes of processing. Hence all sampling was conducted in the three participating on-shore processing facilities. Harvest sampling was organised to collect tissue from as many fish as possible in the period available within funding constraints, and to spread the collection across all the months of the harvesting period to maximise the distribution of sampling across cages, farms, Industry members and processing plants. Tissue samples were only collected from fish in the length range 98-109 cm. This range was defined from the growth curve data used in the SBT operating models for 3 year-old fish (Eveson et al., 2004 & 2005).

The final sample collection method developed during the pilot project involved:

- Collection of tail stalks, which are removed as part of normal processing. Tails stalks were only collected from the specified size range. Date of harvest, length, collector and other details are recorded. The tail stalks were frozen and then slightly thawed at a convenient time to take the biopsy.
- Collection of tissue samples from the tail stalks. Additional data was recorded (e.g. date, collector, sample number). The tissue was collected through the skin of the tail stalk using the gene-tag tool, and loaded into individually labelled vials.

Both CSIRO scientists and staff from Seatec Pty Ltd collected tails and tissue samples from tails. All three participating factories were visited during the harvesting season. Additional biological samples were also collected.

We thank the Australian SBT Industry, factory managers and staff in the processing factories for allowing CSIRO and Seatec access to their fish tails and facilities, and for assisting with trials of different methods as the techniques were refined.

A total of 16,490 tissue samples were collected during harvest in 2017, well in excess of the design study target of 10,000 samples. A database has been developed to manage the gene-tagging sample collection data. These data were provided to the CCSBT through the 2018 data exchange.

Otoliths were collected from some harvested fish with known length to check that the length range being used was likely to be 3 year olds. In addition, vertebrae have been collected and a project to review speed and costs of vertebrae ageing has commenced. This will provide more information on the length-at-age range. Only a few samples have been analysed to date and we will report further on this prior to the 2019 data exchange and reconditioning of operating models.

The harvest sampling component of the gene-tagging projects has demonstrated that it is feasible to collect large numbers of tissue samples from age 3 fish during harvest processing in Port Lincoln in a cost effective manner.

2.3 DNA extraction and sequencing using CSIRO SNP markers

The pilot gene-tagging project has processed over 20,000 tissue samples. Protocols were developed for DNA digestion, robotic extraction and quality controls. Extracted DNA is collected into two plates, genotype and archive. The genotype plate is used for genotyping and the archive plate of DNA is stored in a -80 degree Celsius freezer at CSIRO. Data are recorded during all stages of the processing.

The extracted DNA is genotyped using a specifically designed panel of SNP markers developed by CSIRO. Each plate holds 92 gene-tagging samples, plus control samples. Not all samples had good quality or quantity of DNA and therefore not all samples were successfully sequenced.

DNA was also extracted from an additional 6,000 harvest samples that were also genotyped (funded by CSIRO). These extra samples resulted in additional recaptures and improved the precision of the abundance estimate.

The pilot gene-tagging project has demonstrated efficient high throughput processing of thousands of tissue samples.

2.4 Data analysis and calculation of an abundance estimate

The genotype data were analysed to determine which individuals were in both the release set and the harvest set (using the unique DNA fingerprint). This involved first filtering the data to exclude individuals that had: (i) significantly more heterozygous SNP markers than expected on average (suggesting they may be contaminated), and (ii) too many SNP markers with poor sequencing information (i.e., genotypes with less than 20 total counts per locus) to reliably match with other samples. Over 90% of both the release and harvest sets were retained after filtering (3,456 release samples and 15,391 harvest samples). This increase in percent retained from the December 2017 milestone report (Preece et al., 2017) is due to changes in the way that the matching of genotypes is carried out, such that matching can be done using a subset of the total of 59 SNP markers. Previously, the code used to identify matches required data on all 59 markers, which meant an individual needed to be excluded from the analysis if it had more than a few markers with uncertain genotype calls. Now, an individual can be included in the analysis as long as it has at least 30 markers with reliable genotype calls. Of course, the two individuals being compared may not have the same 30 good markers, meaning that less than 30 markers can end up being compared. However, it is usually the same markers that are poor, so out of the 3,456*15,391 = ~53.1 million pairwise comparisons, only a small percent of the total involve less than 30 markers (Figure 2). Theoretical calculations based on the observed allele frequencies for the 59 markers indicate that with 25 or more markers, the expected number of false positive matches out of 53.1 million pairwise comparisons is virtually 0 (for truly unrelated fish and half-siblings).

The final analysis identified 22 matches (recaptures) between the 3,456 release and 15,391 harvest samples that remained after filtering (Figure 3). The match was perfect for 21 of these (no SNP markers differed between those being compared), and differed by only 1 marker (out of 57) for the remaining individual. This gives an abundance estimate of age 2 fish in 2016 of 2,417,786 with a CV of 0.21. The median estimate of age 2 fish in 2016 from the 2017 stock assessment is 2,102,853.

Investigation into the 22 matches did not suggest any spatial or temporal patterns to indicate fish released in a given area, or on a given day, were more likely to have been recaptured (Figure 4, Table 1). The length frequency of the 22 matches at time of release and at-harvest are shown in Figure 5. The pilot study was not designed to clarify whether the fish in the length classes selected for releases and harvesting are all likely to be age 2 and age 3 (respectively). Direct ageing of otoliths and vertebrae may be able to help resolve this, and if necessary, the analysis can be modified to omit samples from fish that are unlikely to be age 2 (releases) or age 3 (at-harvest). We plan to provide results on the outcomes of this direct ageing project, and any potential update to the abundance estimate, prior to use of gene-tagging data in the reconditioning of operating models in 2019.

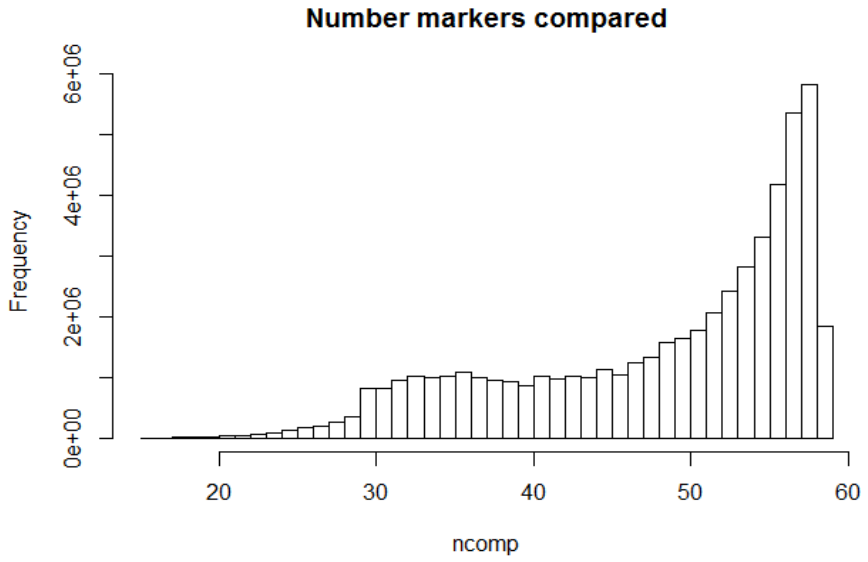

Figure 2. Number of SNP markers (out of 59) being compared between each pair of release and harvest samples (~53.1 million pairs).

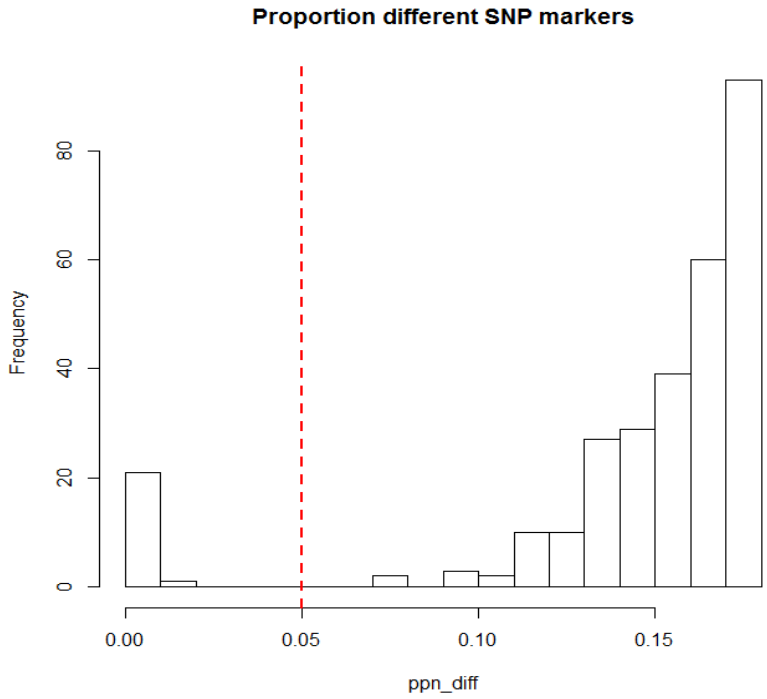

Figure 3. The proportion of SNP markers that differ between each pair of release and harvest samples (where the number being compared differs between pairs – see Fig. 1). The histogram has been right-truncated at 0.18 since the huge bump for the unrelated pairs (which has a mean around 0.50) would otherwise swamp the figure.

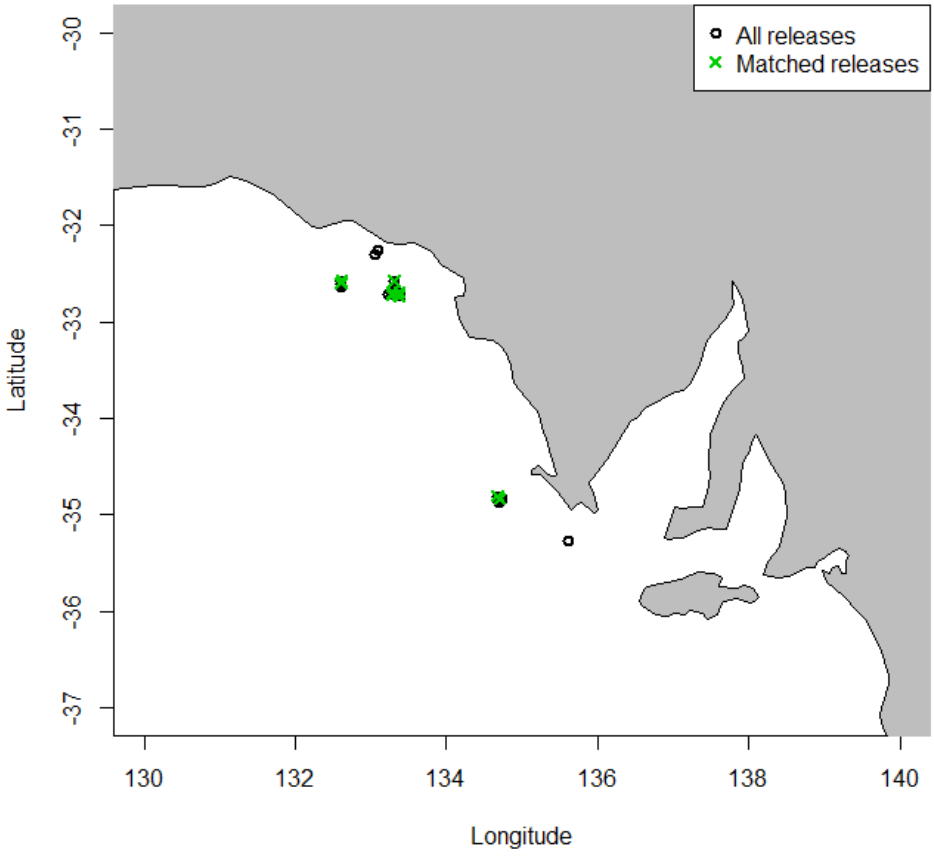

Figure 4. Map of release locations for all release samples compared to those of recaptured fish.

Table 1. Number and percent of releases and recaptures by release date.

Release Date	No. Releases	% Releases	No. Recaptures	% Recaptures
11/03/2016	38	1.1	0	0.0
12/02/2016	193	5.6	2	9.1
13/02/2016	541	15.7	3	13.6
14/02/2016	11	0.3	0	0.0
15/02/2016	73	2.1	0	0.0
16/02/2016	307	8.9	1	4.5
17/02/2016	85	2.5	1	4.5
18/02/2016	299	8.7	1	4.5
22/02/2016	92	2.7	1	4.5
23/02/2016	657	19.0	3	13.6
24/02/2016	302	8.7	2	9.1
25/02/2016	184	5.3	1	4.5
27/02/2016	306	8.9	3	13.6
28/02/2016	266	7.7	3	13.6
29/02/2016	102	3.0	1	4.5
Total	3456	100%	22	100%

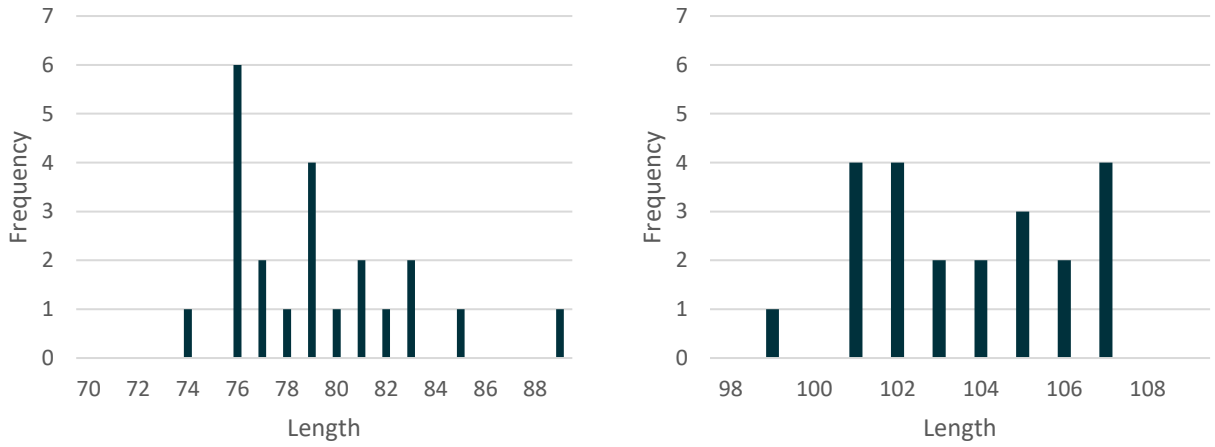

Figure 5 Length frequency of the 22 matches at time of release (left) and at harvest (right).

3 Discussion

The pilot project has demonstrated the feasibility and logistics of a large-scale genetic tagging program for SBT and its potential to provide an absolute abundance estimate for monitoring and management purposes. This includes the successful collection of samples at-sea and during commercial harvest from farms, collection of quality DNA, and high through-put processing from tissue to DNA and quality-controlled genotypes. It was planned that 5,000 tissue samples would be

collected from fish tagged and released at-sea, however bad weather resulted in 3,768 tissue samples being collected in 2016. Additional tissue samples were collected during harvest sampling in 2017 to compensate for the lower than planned number of releases, with 16,490 fish being sampled (well over the 10,000 samples initially planned). All tagging samples and harvest samples have been sequenced. The flexibility to increase sample sizes at harvest to accommodate for lower than anticipated sample sizes at release, is a strong advantage of the mark-recapture estimator being used for gene-tagging (Preece et al., 2015). This same level of adaptive sampling is not possible for conventional tagging programs as the recapture effort is controlled by the catch of the commercial sector and the comprehensiveness of the tag return program.

Quality controls for the sequencing data ensures only samples with good DNA genotypes are included in the analysis. This reduced the number of suitable samples for use in the abundance estimation. We have further optimised methods and protocols in the gene-tagging 2017 and 2018 projects to improve the percentage of samples that can be successfully genotyped.

The abundance estimate from the pilot project is similar to estimates from the 2017 stock assessment and will be used in candidate management procedures and future assessments of stock status. The design study examined, in detail, potential sources of bias and uncertainty in the estimation of abundance, including spatial dynamics, mixing and errors in ageing of recapture samples. The gene-tagging project was not designed to resolve these questions, however, the 2015 design study evaluated their likely impact on abundance estimates and recommended extensions to the pilot study design for consideration by the ESC. For example, to address the question about the percentage of the population that goes into the GAB, or mixing of 1 year old fish, fish tissue samples could be collected from a different location(s) or older age-class (e.g. age 5 fish from NZ) to calculate a second abundance estimate, which if significantly different to the original would indicate non-GAB fish had subsequently mixed with the whole population (Preece et al., 2015). Errors in ageing can be corrected for by estimating the proportion of fish in the recapture sample that are potentially aged incorrectly. Otoliths and vertebrae have been collected to examine age and length to refine the length classes sampled, and may provide correction factors for recapture samples if required.

The CCSBT has commenced an on-going recruitment monitoring program using the gene-tagging method. Around 8,000 fish were tagged and released in both 2017 and 2018 (Bradford et al., 2017; Preece and Bradford, 2018). In these subsequent years, improvements have been made based on experience and lessons learned from the pilot project. Tagging in 2017 used a refined sampling tip for tissue collection, and a different, more efficient charter vessel with a highly experienced fishing master and crew. In 2018, new “blue” sampling tips which have been mass produced on a mould were used (these were also as used for 2017 harvest sampling and had a much higher percentage of successful genotyping). Given the unpredictable nature of fishing, we recommend that the field team continue to collect more tissue samples when possible, as the harvest sampling can be modified to collect fewer or more samples as needed to achieve the target coefficient of variation for the abundance estimate.

In summary, the gene-tagging project is a new and exciting addition to recruitment monitoring of SBT and has prospective applications for other species.

References

Anon 2018. Report of the Ninth Operating Model and Management Procedure Technical Meeting. 18 - 22 June 2018, Seattle, USA

Anon 2016. Extended Scientific Committee for the Twenty First Meeting of the Scientific Committee, 5 - 10 September 2016, Kaohsiung, Taiwan

Bradford R and Preece A. 2017. Progress in stage 1 of gene-tagging 2017 and research mortality allowance request for gene-tagging 2018. CCSBT-ESC/1709/8.

Bradford R, Preece A and Davies C. 2016. Progress report on the implementation of the CCSBT gene-tagging pilot project in 2016. CCSBT-ESC/1609/7

Bradford RW, Hill P, Davies C, and Grewe P. 2015. A new tool in the toolbox for large-scale, high throughput fisheries mark-recapture studies using genetic identification. Marine and Freshwater Research http://dx.doi.org/10.1071/MF14423

Eveson, J.P., G.M. Laslett, T. Polacheck. 2004. An integrated model for growth incorporating tag-recapture, length frequency and direct aging data. Canadian Journal of Fisheries and Aquatic Science 61:292–306.

Eveson, J.P., T. Polacheck, J. Farley. 2005. Updated estimates of growth rates for juvenile SBT using tag-recapture and otolith direct ageing data up to 2005. Report CCSBT-ESC 0509/34, sixth meeting of the CCSBT Stock Assessment Group and the tenth Meeting of the Extended Scientific Committee, August 29–September 3, and September 5–8, 2005, Taipei, Taiwan.

Gunn JS, Clear NP, Carter TI, Rees AJ, Stanley CA, Farley JH, Kalish JM. 2008. Age and growth in southern bluefin tuna, Thunnus maccoyii (Castelnau): Direct estimation from otoliths, scales and vertebrae. Fisheries Research, Volume 92 (2–3) pp 207-220. https://doi.org/10.1016/j.fishres.2008.01.018

Hillary R, Preece A, and Davies C. 2018a. Data generation & changes to SBT OM. CCSBT-OMMP/1806/04.

Hillary R, Preece A, and Davies C. 2018b. Initial MP structure and performance. CCSBT-OMMP/1806/05.

Hillary R, Preece A, and Davies C. 2016a. Methods for data generation in projections. CCSBTOMMP/1609/7, CCSBT- ESC/1609/BGD-6.

Hillary R, Preece A, and Davies C. 2016b. MP results and estimation performance relative to current input CPUE and aerial survey data. CCSBT-ESC/1609/18.

Preece AL and Bradford RW. 2018. Gene-tagging recruitment monitoring in 2018. CCSBT-ESC/1809/7. CSIRO, Australia.

Preece A, Bradford R, Grewe P, Eveson P, Farley J, Davies C. 2017. Progress report on the CCSBT pilot gene-tagging program in 2017. CCSBT-ESC/0917/7.

Preece A, Eveson JP, Davies C, Grewe P, Hillary R and Bravington M. 2015. Report on gene-tagging design study. CCSBT-ESC/1509/18.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e csiroenquiries@csiro.au
w www.csiro.au

AT CSIRO, WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today – for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australia's catalyst for innovation.

CSIRO. WE IMAGINE. WE COLLABORATE. WE INNOVATE.

FOR FURTHER INFORMATION

Oceans and Atmosphere
Ann Preece
t +61 3 6232 5222
e ann.preece@csiro.au
w www.csiro.au