Report on the potential and feasibility of genetic tagging of SBT

Campbell Davies
Andy Moore
Peter Grewe
Russ Bradford
Marinelle Basson

Prepared for the CCSBT 5th Management Procedure Workshop 2-7 September and the 13th Meeting of the Extended Scientific Committee 8-12 September 2008
Rotorua, New Zealand

Table of Contents

Abstract

The conventional tagging program of the CCSBT, which formed a core component of the CCSBT Scientific Research Program (SRP) was suspended in 2007, primarily due to concerns about difficulties in estimating reliable reporting rates. This led to the CCSBT Scientific Committee (SC) identifying two potential alternative tagging methods which could potentially overcome the reporting rate problem. In this paper we discuss the potential for using genetic tagging.

Genetic mark-recapture methods rely on a DNA molecular marker rather than a physical tag and therefore address two of the major limitations of conventional tagging programs: tag shedding and non-reporting of tag recaptures. We discuss the applicability and feasibility of genetic tagging to SBT. We note that a set of markers have already been developed for SBT as part of existing research at the CSIRO, and this set can be used as a starting point for application in genetic tagging.

The different components of the costs involved in such a program are identified. The cost involved in obtaining a tag at sea would be similar to costs involved in placing a conventional tag. The cost of tag recovery will depend on how and where sampling takes place, though there should be no need for 'reward' costs as is the case for the return of conventional tags. The genetic analysis of both the 'tagged' samples and 'recaptured' samples depend on the number of loci considered. The costs presented here are very approximate, and we estimate that further efficiencies should bring total costs down substantially.

Introduction

The conventional tagging program of the CCSBT, which was established as a key component of the Scientific Research Program (SRP) in 2001, was aimed at improving our ability to estimate fishing mortality rates and growth rates. It was hoped that it would also improve our understanding of the movement dynamics of SBT. The success of the program in meeting all its objectives, however, relies on recaptured tags being returned with the relevant recapture information, and the ability to reliably estimate the associated reporting rate for each fishery. Unfortunately, the full potential of the conventional tagging program has not been achieved because of low returns and lack of information to reliably estimate reporting rates for some fisheries. The SRP conventional tagging was in fact suspended in 2007, partly because of concerns over the inability to estimate reliable reporting rates (Anon. 2007).

The mark-recapture approach (e.g. tag release and recapture) is still, however, a potentially powerful tool if the associated issue of reporting rates can be overcome. The CCSBT SC therefore considered, at its meeting in 2007, which alternative tagging techniques and/or tag types could be used. Two approaches were highlighted as possible alternatives to conventional (external dart) tags: Passive Integrated Transponder (PIT) tags and genetic tagging (or "DNA fingerprinting"). The potential use of PIT tags is discussed in Harley et al. 2008 (CCSBT-ESC/0809/14). In this paper, the potential use of genetic tagging is discussed.

Standard indicators of abundance

Conventional tagging programs continue to be a major tool in assessing fish stocks especially pelagic species. Traditional fishery dependant indicators of abundance such as catch-per-unit-effort and catch size and age distribution have inherent limitations related to the accurate reporting of catch and standardisation of fishing effort for behaviour and technological efficiency gains. Properly designed and implemented tagging studies can reduce reliance on fishery dependent data as an index of stock abundance and, or, harvest rate. However, conventional tagging programs have several important drawbacks that can reduce the accuracy of the data they provide. These limitations are typically tag shedding, non-reporting of recaptures, mortality due to capture and adequate sampling at spatial and temporal scales. Currently the conventional tagging program for Southern Bluefin Tuna (SBT) has been temporarily suspended because of low reporting rates for components of the longline fleet and uncertainty in reporting rates for the surface fishery (Anon, 2007).

Aerial surveys based on line transect methods over consistent temporal and spatial scales have been used as an estimator of stock abundance for juvenile (age 2-4) SBT in Australia (Hartog et al., 2007). While aerial surveys can provide fishery independent indices of relative abundance of juveniles; they cannot be used to estimate fishing mortality or catch, and they provide no information on the older reproductive age classes taken by the longline fisheries. Alternatives such as archival tags can in principle provide estimates of exploitation rates but are affected by the

same limitations as conventional tagging (e.g. non reporting of tags) and the costs per tag can be expensive.

Gene tagging

Modern gene tagging relies on the genotyping of individuals using a suite of microsatellite markers (Woods et al., 1999). These individual genotypes from the initial sample are compared to fish that are subsequently sampled and the number of DNA matches provide a measure of recapture rate. This recapture rate can then be used to estimate population size and mortality rates, depending on the design of the program and underlying assumptions. Re-sampling can be done either at, or post, capture. This allows for recapture sampling to be done at fish processing facilities. Alternatively, recapture-samples could be taken at sea using a standard sampling protocol to provide a consistent sub-sample of the catch. This would have the advantage of providing recapture location. As the "tags" are invisible, there is no need for estimates of reporting rates. Removing the need to estimate reporting rates provides the potential for less uncertain results than those currently available, particularly for the high seas components of the fishery.

In addition to removing the need to estimate reporting rates, any handling effects of the tagging process may also be reduced through reducing both the physical trauma of tag insertion and carrying the conventional tags when at liberty.

Feasibility of gene tagging for SBT

Gene tagging has been used in various forms for fisheries research for some time. However, it was the wide-scale adoption of highly polymorphic microsatellite markers and technological advances in high throughput screening systems that have allowed gene tagging to become more cost-effective and widely used. For example, the technique has been used in Northern Australia to estimate fishing mortality rates in narrow-barred Spanish mackerel (*Scomberomorus commerson*) (Buckworth et al., 2006). The research on Spanish mackerel provided insights into the effectiveness of gene tagging for marine pelagic fish. The applicability of this research to other Scombrids, such as Southern Bluefin Tuna (SBT), is likely to be high. Both species share similar life history traits and an estimate of fishing mortality can be determined from harvested specimens to overcome previous issues with under reporting of catch. The technique is also currently being used in North America to monitor populations of rock cod (J. Ovenden *pers. comm.*).

Potential caveats of gene tagging

In order for gene tagging to be effective a suite of microsatellite markers must be developed for the species under study. The time and cost of developing these markers can be considerable and includes not only isolation of microsatellite containing sequences but also groundtruthing individual locus assays. Groundtruthing of individual loci entails examination of amplification products from a suitable random sample of the population to ensure PCRs produce consitent amplification products that can be easily and reliably scored. In the case of SBT, CSIRO have already developed and tested a suite of DNA microsatellite markers for SBT as part of existing research on the species. From a set of 100 unique microsatellite containing

fragments, a subset of 40 were chosen for further development and testing. Of these 16 have been determined to produce consistent PCR amplification products free from null alleles and all conform to Hardy-Weinberg expectations for loci segregating and a randomly mixing population. All loci easily facilitate automated scoring on ABI capillary sequencing machines using Genemapper software.

A secondary issue of concern in terms of quality of marker repeatability is lack of amplification product at one or more loci for any given individual. DNA quality (i.e. how well the tissue has been preserved) is the most significant problem affecting consistency of successful PCR amplification of individual samples. Poor DNA quality can lead to PCR failure. In many studies samples are collected, stored, and transported under varying conditions. However, CSIRO is currently developing standard protocols to facilitate high quality sampling at sea and in processing plants. Null alleles (where DNA samples consistently fail to amplify) (Dakin & Avise, 2004) can also lead to PCR failure when the priming site of the template and PCR primer lack significant homology to produce a PCR product. However, two tissue samples from the same individual should produce a similar result. In this respect, independent assays and therefore the presence of either single allele (heterozygous for a potential null) or no product (homozygous for null allele) should be reproducible. In any event a few alternative markers should be included as redundant loci to further aid in confirmation of identity to avoid this potential issue during full scale genetic screening. The current suite of 16 SBT loci will be sufficient to avoid any of these issues and if null alleles are present they are at extremely low (undetected) frequency. Further development of additional loci from our exisiting libraries is also possible for a fraction of the cost of starting from scratch.

The ability to run multiple samples in a single run (multiplexing) will reduce time and running costs significantly. The trade-off between using enough loci (genetic makers) to provide sufficient power to detect genotype differences and the costs and complexity of managing the vast amount of data generated (possibly millions of comparisons) will also need to be considered. Software and database development in this area has increased in recent years and programs that can manage these vast and complex data sets as well as resolve genetic differences will be a key element to the success of such research. These analysis packages mirror those developed for human forensic sciences which is addressing a similar question of matching two DNA samples, one from a crime scene and one from a suspect.

Costs of monitoring using gene tagging

The cost estimate for DNA profiling of a tissue sample for the purposes of gene tagging depends on several factors. For the purpose of this paper these factors can be described in four general steps of the process: i) obtaining the tissue sample (ii) DNA extraction; (iii) PCR amplification and electrophoretic analysis of the DNA profile; (iv) review and quality control analysis of the DNA profiles to provide a final mark-recapture data set for analysis (Table 1). In comparison to conventional tagging, obtaining tissue samples for a gene tagging program requires a two-step process: the first being the taking of the tag sample from fish at sea (similar to the process of conventional tag release); and the second being recovery of the tag at capture, a market place or processing facility (return of the tag). The cost involved in obtaining a

tag at sea is similar to costs (and logistics) involved in placing a conventional tag and is not included in this costing.

In contrast to a conventional tagging program, the cost of tag recovery is different for gene tagging. This requires the collection of a second tissue sample from a sub-sample of fish. For our purposes this usually ends up coasting about AUD$3.00 per sample if the sampling effort required is only obtaining the sample from a processing facility. There is a cost associated with shipping the sampled tissue that can increase the cost of tissues sampled and shipped to the lab to about AUD$3.50 per sample. In addition, prior to DNA extraction it is necessary to sub-sample the tissue, which adds approximately AUD$0.50 per sample. Therefore there is an approximate unit cost of AUD$4.00 per sample for the collection of the "release" and potential "recapture" samples.

The second and third steps of the process (ii and iii above) can be sub-contracted to a commercial laboratory. Current rates for DNA extractions in Australia are AUD $5.50 with PCR and electrophoresis AUD$1.50 per locus per individual. Once electrophoresis is completed the final phase of the genotyping process requires data files to be checked for quality control of automated genotype calling and correct DNA profile assignment (phase iv). This last stage requires personnel experienced with specific software and specific training in DNA profile pattern recognition. Once the initial setup has been complete and the software has been trained to recognise and automatically identify subtle genotypic variations, the process of checking the automated calls is fairly straight forward. Initially, each locus requires approximately 15 to 30 minutes to quality control around 100 individuals or about AUD$0.10 per locus per individual (at an estimated wage of AUD$20.00/hour).

It should be noted that these costs are approximate and provided for the purposes of considering the feasibility and relative (to conventional tagging) cost-effectiveness of gene tagging. Further efficiencies would be expected once standardised protocols have been developed and if larger numbers of individuals are required to be processed. We estimate that further efficiencies (developed for phase ii, iii, and iv) should be able to reduce total costs by 40 - 50%. These cost estimates further assume that analysis of six loci will provide a genotype that can uniquely identify an individual as a recapture. Additional loci are required when either variability is low within the target population (i.e. inbred and or critically endangered populations) or for kinship studies examining parent / offspring relationships (e.g. Bravington and Grewe 2008). For some studies a specific number of individuals (albeit a rare number, i.e. less than 10% of the sample) will need to be examined for a total of 18 loci. This will likely only be required for kinship studies examining parent / offspring or brother / sister relationships (e.g. Bravington and Grewe 2008) and is not expected to be necessary in the case of standard mark-recapture models for individuals.

The potential cost of the additional ancillary data, particularly size and/or age data of the fish sampled for genetic analysis should also be considered. However, this cost should be similar to what it is for conventional tags and therefore not important in a relative comparison between the two approaches. A final consideration when deciding on how and where to collect samples, is the issue of data on time and location of capture. If that information is carried through from the vessels to the processors, then sampling at the processing plants should be sufficient. If this is not the case or there is uncertainty in the reliability of this information, then consideration should be given to sampling at harvest/capture. Alternatively, fish for genetic sampling could possibly

be identified on board and necessary data recorded, with the tissue sample from the identified fish taken at processors. It is unlikely that these different options would substantially affect the costing, but would be important in considering the design of the tagging program and associated field logistics and sampling protocols.

Table 1. Cost estimates for "gene-tagging" of SBT based on analysis of a single locus per individual analysis, analysis of six loci per individual, analysis of 12 loci per individual, and analysis of 18 loci per individual. All values in Australian dollars and assume the availability of a suitable optimised DNA microsatellite assays for SBT (see text for details).

Activity	Cost per unit	Costing per 6 loci	Costing per 12 loci	Costing per 18 loci
Market/processor sample of tissue	3.50	3.50	3.50	3.50
Sub-sample tissue in lab	0.50	0.50	0.50	0.50
DNA extraction	5.50	5.50	5.50	5.50
PCR and electrophoresis	1.50	9.00	18.00	27.00
Quality Control	0.10	0.60	1.20	1.80
Total	**$11.10**	**$19.10**	**$28.70**	**$38.30**

Conclusions

Genetic mark-recapture methods address two of the significant uncertainties and limitations associated with conventional "physical" tagging programs: tag shedding and rate of reporting of tag recaptures. This is because the technique relies upon a permanent DNA molecular marker rather than a physical tag. The wide-scale adoption of highly polymorphic microsatellite markers and technological advances in high through-put screening systems has seen gene tagging being more widely used in fisheries research and monitoring. Research on Spanish mackerel provides insights into the effectiveness of gene tagging for marine pelagic fish and the applicability of this research to other Scombrids such as Southern Bluefin Tuna (SBT) is likely to be significant.

In order for gene tagging to be effective a suite of microsatellite markers must be developed for the species under study. The CSIRO have developed a set of markers for SBT as part of existing research on the species. These markers can be used as a starting point; their ability to amplify DNA from samples of varying quality can be assessed, and if necessary further work to identify additional or alternative markers can be undertaken. The time and resources required to do any additional work requires are small, relative to the investment already made in the development of the DNA microsatellite library.

As in the case of conventional tagging, a genetic tagging program would include taking tissue samples from live fish at sea (and releasing them, i.e. like tagging with a conventional tag) and taking tissue samples from a sample of harvested fish (potential recapture of a 'tagged' fish). The cost involved in obtaining a tag at sea is similar to

costs (and field logistics) involved in placing a conventional tag. The cost of tag recovery will depend on where the sample is obtained, for example, at the processors, or at sea. However, with genetic tagging there should be no need for 'reward' costs as is the case for the return of conventional tags.

The genetic analysis of both the 'tagged' samples and 'recaptured' samples depend on the number of loci considered. The costs presented here are approximate and intended to provide for a relative comparison with the costs of the conventional tagging program under the SRP. We estimate that further efficiencies would be found in the implementation of a large-scale program, which would further reduce the total costs associated with a "gene-tag" program for SBT.

References

Bravington, M. and P. Grewe. 2008. Update on SBT close-kin abundance estimation CCSBT-ESC/0809/29.

Buckworth, R.C., Ovenden, J.R., Broderick, D., McPherson, G.R., Street, R. and McHale, M. 2006 Genetag: Monitoring fishing mortality rates and catchability using remote biopsy and genetic mark-recapture. Lyle, J.M., Furlani, D.M. & Buxton, C.D. (Eds.). *Cutting-edge technologies in fish and fisheries science.* Australian Society for Fish Biology Workshop Proceedings, Hobart, Tasmania, August 2006, Australian Society for Fish Biology.

Dakin, E.E. and Avise, J.C. 2004 Microsatellite null alleles in parentage analysis. *Heredity*. 93: 504-509.

Harley, S., Bradford, R., and C.R. Davies. 2008 Using passive integrated transponder (PIT) technology to improve performance of CCSBT's conventional tagging programme. CCSBT-ESC/0809/14.

Hartog, J., Preece, A. and Kolody, D. 2007 Fishery indicators for the SBT stock 2006/07. *CCSBT 8th Meeting of the Stock Assessment Group and the 12th meeting of the Extended Scientific Committee*, 4-8 September and 10-14 September 2007, Hobart, Australia.

McKelvey, K.S. and Schwartz, M.K. 2004 Providing reliable and accurate genetic capture-mark-recapture estimates in a cost-effective way. *The Journal of Wildlife Management* . 68**(3)**: 453-456.

Woods, J.G., Paetkau, D., Lewis, D., McLellan, B.N., Proctor, M. and Strobeck, C. 1999 Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin*. 27**(3)**: 616-627.