Update on SBT close-kin abundance estimation, 2008

Mark Bravington
Peter Grewe

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

Abstract

We describe progress on estimating SBT spawner abundance using close-kin data, following on from the study proposed last year. Further samples have been collected, and more fish have now been genotyped, allowing us to examine the quality of the genetic data. Results are promising, and we expect to deliver preliminary estimates by CCSBT 2009.

Update on SBT close-kin abundance estimation

This paper is a short update on progress with SBT close-kin abundance estimation, following on from the study proposed last year in CCSBT-SC/0709/18 (Bravington and Grewe, 2007).

Project arrangements

Funding has now been provisionally agreed between CSIRO and FRDC, and the project will be overseen by a steering committee including international experts on genetics, mark-recapture, and tuna assessment. Funding began in July 2008, so there has only been limited time for further genetic analysis (see next section). The revised schedule for the project is described below.

CSIRO has continued to collect samples, as listed in Table 1; we are still receiving samples from Port Lincoln in 2008. Most of the Port Lincoln samples are from age-3 fish, with a substantial proportion of age-2s (based on length measurements). Otoliths are available for almost all the Indonesian samples from 2005-6 and 2006-7 (2007-8 data not available yet), and a number have been aged as part of the standard Indonesian ageing programme (Farley and Proctor, 2008).

Year (Jul-Jun)	Place	Samples held	DNA extracted
2005-6	Indo	216	216
	PL	4000	200
2006-7	Indo	1520	1069
	PL	4000	200
2007-8	Indo	1594	0
	PL	800+	0

Table 1: Samples collected and stored up to August 2008

Genetic progress

Bravington and Grewe, 2007, included preliminary estimates of how many loci per fish would need to be scored to exclude false positive matches (i.e. a juvenile and an adult that are actually not a parent-offspring pair, but that by chance happen to have at least one allele in common at every locus examined). Those estimates indicated that only a modest number of loci would need to be scored, thus making the costs feasible. However, the available data for those calculations came from only 16 fish, so the allele frequency estimates which underpin the calculation were inevitably uncertain. Further, it was not possible to check some aspects of locus reliability (e.g. null alleles) because of the small sample sizes.

We have now genotyped 96 adult fish (all so far from Indonesia 2005-6) at 18 polymorphic loci (mostly the same loci used last year, with some changes for technical reasons). The larger sample size gives us better estimates of allele frequency, and allows us to check for null alleles. All genotyping for this larger batch was done by the Australian Genome Research Facility, using primers and amplification protocols developed at CSIRO.

With the larger sample size, the allele frequency estimates for individual loci do change somewhat but without affecting the overall number of loci required for exclusion *if* genotyping is assumed exact (i.e. no scoring error). A few loci show some evidence of scoring error, in the sense that they have an excess of apparent homozygotic fish (departure from Hardy-Weinberg equilibrium, HWE), likely through inability to score the other allele on that fish— i.e. showing non-amplifying or null alleles. Table 2 shows the sorted p-values for HWE as produced by the GENEPOP program (Raymond and Rousset, 1995); small values indicate possible null allele issues. Given the number of loci being tested, some p-values will turn out small by chance, so in fact only the first 4 or 5 loci are of any conceivable concern; based on the p-values, at least 13 of the 18 loci show no evidence of null alleles.

It is important to note that, even if a locus does exhibit null alleles, the locus may still be useful for DNA fingerprinting; the presence of null alleles simply means that relatedness between a pair of fish cannot be ruled out based on that locus if either fish is an apparent homozygote. This less stringent criterion is less powerful statistically, but robust to null alleles. Depending on the p-value, we can decide whether to use the more stringent or less stringent criterion. If we use the 10 most powerful of the 18 loci in Table 2, and use the less stringent exclusion criterion for the 3 of the 10 with p-values below 5%, then we should still eliminate about 99% of possible false positive juveniles and adults (see section 6.0.4 in last year's paper for basis of calculation). The remaining small number of potential matches— most of which will in fact be true parent-offspring pairs— can be checked by examining a small number of extra loci, at minimal extra cost.

We also examined how often (i.e. in what proportion of fish) each locus failed to amplify

Locus	D111	D232	D139a	D11b	D225	B5	D201	D4D6	B232a
%p-value	0	0	0.14	0.98	2.3	5.0	14.8	22.7	22.9
Locus	3D4	D115	D211	D122	D10	D3	D135	D235	D12
%p-value	24.0	26.6	32.1	42.3	43.6	47.1	58.0	79.3	85.7

Table 2: Testing departures from Hardy-Weinberg equilibrium

at all; this is related to, for example, quality of tissue preservation, and is a serious problem in some genetic tagging studies (e.g. for Spanish mackerel in the Northern Territory, where small pieces of tissue remain uncollected for hours in warm water). In our Indonesian samples, though, there does not seem to be a problem. Depending on the locus, between 0 and 7% of the fish failed to score at all at that locus. However, most of these failures arose from just 3 fish (a small proportion of the 96— and even these may succeed if DNA is re-extracted). If those three fish are excluded, the average unscored genotype frequency per locus is about 1.5% and the maximum in any "top-ten" locus is about 4.5%. Hence total dropout does not seem to be a concern for the Indonesian samples, which are collected under very good conditions. Conditions for Port Lincoln samples should be even better.

The plan from here

Our immediate plans are to:

- genotype a larger set (300 juvenile fish) from the GAB, to examine incidence of siblings and half-siblings;
- finalise locus choice & protocols for mass genotyping;
- present results to date to the project Steering Committee in September 2008, to demonstrate feasibility and plan the next steps.

Although a high proportion (say >30%) of siblings or half-siblings would not bias our abundance estimates (see last year's paper), it would cause problems for precision, so the first check above is important for assessing whether our sample sizes are adequate. Assuming all is well, we will then proceed with genotyping all existing samples and with data analysis as described in last year's paper, presenting a preliminary report to CCSBT 2009. Data analysis is rarely a one-step process, so it is likely that further statistical analysis will be required before a more final report can be presented to CCSBT 2010. Final analysis and write-up will be completed by July 2011. There will also be a further year of sampling in Indonesia (summer 2008-9) and Port Lincoln (harvest 2009), and data from those samples will be available in time for CCSBT 2010.

Acknowledgements

Particular thanks are due to Craig Proctor (CSIRO), Ms Retno Andamari (Research Institute for Mariculture, Gondol, Bali), Mr Kiroan Siregar, and Mr Rusjas Mashar (sampling enumerators) for their diligent efforts in ensuring the successful collection of Indonesian tissue samples.

We appreciate the assistance of the the Australian SBT industry (Port Lincoln), the Indonesian tuna fishing industry (Benoa), and the Research Centre for Capture Fisheries (Jakarta).

Funding for this work is provided by CSIRO Wealth from Oceans flagship and by the Fisheries Research and Development Corporation.

References

[Bravington and Grewe, 2007] Bravington, M. and Grewe, P. (2007). A method for estimating the absolute spawning stock size of sbt, using close-kin genetics. Scientific Committee Report CCSBT-SC/0709/18, Commission for the Conservation of Southern Bluefin Tuna.

[Farley and Proctor, 2008] Farley, J., A. R. and Proctor, C. (2008). Update on the length and age distribution of sbt in the indonesian longline catch. Scientific Committee Report CCSBT-ESC/0809/27, Commission for the Conservation of Southern Bluefin Tuna.

[Raymond and Rousset, 1995] Raymond, M. and Rousset, F. (1995). GENEPOP (v.1.2): population genetics software for exact tests and ecumenicism. *Fisheries Science*, 86:248–249.