CCSBT-ESC/1709/12
(ESC Agenda item 7)

OCEANS AND ATMOSPHERE
www.csiro.au

SBT Close-Kin Mark-Recapture with Parent-Offspring and Half-Sibling Pairs: update on genotyping, kin-finding and model development

Mark V. Bravington, J. Paige Eveson, Peter M. Grewe, and Campbell R. Davies

CSIRO, Hobart, Australia

CCSBT-ESC/1709/12

CSIRO
Oceans & Atmosphere
Castray Esplanade, Battery Point, Hobart TAS 7000, Australia
Telephone : +61 3 6232 5222

Fax : +61 3 6232 5000

SBT CLOSE-KIN MARK-RECAPTURE WITH PARENT-OFFSPRING AND HALF-SIBLING PAIRS: UPDATE ON GENOTYPING, KIN-FINDING AND MODEL DEVELOPMENT

MARK V. BRAVINGTON, J. PAIGE EVESON, PETER M. GREWE AND CAMPBELL R. DAVIES

Abstract. Close-kin mark-recapture (CKMR) was first used to estimate the absolute abundance of adult (i.e. spawning age) SBT in 2012. The data consisted of Parent-Offspring Pairs (POPs) that were identified genetically using highly-variable microsatellites. The value of those data and of the associated "stand-alone" CKMR model for assessment and monitoring of the spawning stock have been recognised by the CCSBT; the CCSBT Scientific Research Program now incorporates the annual ongoing collection and processing of genetic samples, as well as investing in design studies. Here we report on (i) the application of a new method for identifying POPs and Half-Sibling Pairs (HSPs), based on Single Nucleotide Polymorphisms (SNPs) instead of microsatellites, and genotyped with modern Next-Generation Sequencing methods, using specifically designed DArTcap assays; and (ii) the development of a new stand-alone CKMR model that uses these new data in a population-dynamics framework that allows for length-, age-, and sex-structure among adults. A total of ~17,000 tissue samples from adult (Benoa, Indonesia) and juvenile (Port Lincoln, Australia) SBT collected over the period 2005–2015 have now been genotyped. Using the roughly 16,000 genotypes remaining after quality-control checks, we identified 77 POPs, 140 definite HSPs and 4 Full-Sibling Pairs; the true number of HSPs is estimated to be about 10% greater, because of the stringent criteria required to exclude false-positives. Examination of mitochondrial DNA indicates that about 65 of the 140 HSPs shared a mother whereas 75 shared a father, consistent with an equal sex-ratio in adult SBT. The POP and HSP data have been incorporated into the reference set of the CCSBT OMs for the 2017 stock assessment process. It has not been possible to finish the new stand-alone CKMR model in time for the ESC, due to the extremely tight schedule for the whole project, the later-than-expected completion of genotyping, and the priority placed on completing quality control and diagnostic analysis for the identified HSP and POPs. The stand-alone CKMR model will be complete by the end of 2017, and will be available for review at OMMP9 and consideration by the ESC in 2018. We include here an overview of the benefits of the combined POP-and-HSP approach, including the challenges for model development and initial considerations of solutions.

Contents

1. Introduction

CKMR uses modern genetics to find pairs of close relatives amongst large collections of tissue samples. It then embeds the number of pairs found, along with covariates such as their age and sex and date of collection, in an extended mark-recapture framework, where "recaptures" are kin of an animal rather than the animal itself. This allows direct estimation of adult abundance and other demographic parameters, without needing to rely on CPUE or catch data. Bravington et al. (2016b) provide an overview and details of the underlying statistical theory.

CKMR has proved very useful to CCSBT for providing fishery-independent monitoring of spawning-stock biomass, and for establishing key biological parameters relevant to management, such as fecundity-at-age. The initial application to SBT was based on Parent-Offspring Pairs (POPs), identified using microsatellite loci, between juveniles collected from the Great Australian Bight (GAB) and adults from Benoa, Indonesia between 2006 and 2010; it is described in Bravington et al. (2012), Hillary et al. (2013), Bravington et al. (2014), and Bravington et al. (2016a). From 2011 onwards, continued sample collection has been supported by CSIRO and CCSBT in anticipation of future CKMR analysis. However, no large-scale genotyping has been done since the completion of the original study. In 2014 and 2015, CSIRO proposed changing the genotyping technique from microsatellites to a modern Genotyping-By-Sequencing (GBS) method here referred to as"DArTcap" (TM), which uses Single Nucleotide Polymorphisms (SNP) loci rather than microsatellites; see Bravington et al. (2015) for more details. There were three independently compelling reasons for moving to SNP loci for CKMR:

(1) future-proofed
(2) cheaper
(3) able to find Half-Sibling Pairs (HSPs) as well as POPs

In the context of stock assessment, the last point is the most important, since having HSPs will permit a direct estimate of adult mortality rate, without requiring untestable assumptions. With the inclusion of HSP data, selectivity on adults can, in principle, be separated from natural (and fishing) mortality. This is a notoriously tough problem for fisheries in general, let alone for SBT.

The proposed change of genotyping method was independently reviewed (R. Waples and E. Anderson, in addendum to Bravington et al., 2015)) and agreed by the Commission (CCSBT ESC, 2015), and a design study was completed under the CCSBT Scientific Research Program (Bravington et al., 2015). The CCSBT allocated funding from the SRP to sequence and genotype samples from 2014 onwards (Anon 2014, Attachment 11. Anon 2015), CSIRO agreed to develop the new stand-alone CKMR model, and CSIRO secured funding to sequence and genotype adult and juvenile samples from 2005-2013. Specifically, the latter covered:

(1) juveniles in the original 2006–2010 samples (since their parents might still be found in post-2010 adult samples);
(2) the "back-catalogue" of as-yet-ungenotyped samples collected between 2011 and 2015.

We have now completed the identification of POPs and HSPs, following the steps outlined in Appendix D of Bravington et al. (2015). The whole process has worked well, and the results are ready for use in modeling.

1.1. **From tissue samples to POPs and HSPs.** The sample collection, genetic processing and identification of kin involves several steps. Sampling of adults takes place in Benoa, Indonesia (during processing of catches from the spawning ground fisheries, and at the same time as otolith collection); sampling of juveniles takes place in Port Lincoln, Australia (from the purse seine fishery, sampled when harvested). All samples consist of a biopsy containing ~300mg of tissue; they are stored in 2.0 mL cryovials, frozen, and transported to the CSIRO laboratories in Hobart. Tissues are held at $-80°$C until sub-sampled for DNA extraction. For each fish chosen for subsampling, a ~15mg slice of tissue is weighed and placed into an extraction chamber for tissue digestion. An Eppendorf EP motion robot completes the DNA extraction and produces two final plates: a sequencing plate, and a replica DNA archive plate. Each plate contains DNA from 92 fish, along with two blanks and two control tissue samples whose positions on the plate allow unique identification of that plate for quality control (QC) cross-checks.

The archive plates are stored frozen at $-80°$C where they remain unless required for further testing. The sequencing plates are sent for genotyping, which again involves several steps. The first part is carried out by DArT Pty Ltd (Canberra), who have developed with CSIRO a specific variant of Genotyping-By-Sequencing for close-kin purposes, known as "DArTcap". It entails: laboratory pre-processing of the plates; analysis using a high-throughput sequencer; and bioinformatic analysis of the terabytes of the resulting data, to produce simple data summaries for each fish at each SNP locus of interest. The second part of genotyping, "genotype-calling", turns those data summaries into multi-locus genotypes for each individual fish— i.e., for each fish and each locus, the pair of alleles inferred to be present. This genotype-calling entails some quite complicated algorithms developed at CSIRO specifically for DArTcap sequencing data, and also estimates the genotyping error-rates for each locus.

The final step prior to CKMR itself is kin-pair-finding, which is based on the inferred genotypes and the error-rates. For this step we have developed generic algorithms (i.e. not specific to DArTcap) from basic statistical principles, which are summarized in the Appendicesand in section 5 of Bravington et al. (2016b). Control of false-positive and false-negative rates is *crucial* to kin-finding, since ~100,000,000 comparisons might be needed to find only ~100 true kin-pairs.

Details of DArTcap setup for SBT, including choice of loci, were determined during 2016. After checking preliminary results, and finalizing the funding arrangements, large-scale sequencing of around 16,000 fish began in February 2017; the full set of sequencing-files were received by CSIRO at the end of March 2017. In parallel, CSIRO developed quality control (QC), genotype-calling and kin-finding algorithms suitable for the new type of genetic data. From April to June, these algorithms were refined and applied to deliver reliable sets of POPs and HSPs suitable for use in the 2017 reconditioning of the CCSBT OMs. The POPs were incorporated into the OM updates for OMMP8 (Hillary et al., 2017a), with the HSPs reported separately (Bravington, 2017), as there was not time to incorporate them prior to OMMP8. The HSP data were incorporated into the OMs subsequent to the OMMP8 (Hillary et

TABLE 1. Number of samples genotyped by year and origin (after some QC checks)

	Indo	Port L
2006	0	1281
2007	0	1305
2008	0	1315
2009	0	1317
2010	943	1284
2011	931	938
2012	527	844
2013	933	873
2014	904	873
2015	0	922
Total	4238	10952

al. 2017b) and the results reviewed at a special webinar meeting of the OMMP Technical Group. The OMMP agreed to include the HSP data series in the reference set of OMs for the 2017 stock assessment, and to do an extra sensitivity run (CCSBT-ESC/0817/Report of OMMP webinar).

Table 1 summarizes the available samples from DArTcap, excluding about 700 that have been rejected so far on QC grounds. (Also included in the sequencing results are about 5,000 replicate genotypes, which can be used for estimating error rates.) For the sake of economy, we did not re-genotype adults from 2006–2009, since any usable offspring would already have been found in the original microsatellite genotyping[1] (except that 2009 adults could still have undetected 3yo offspring caught in 2011). For CKMR models, therefore, the new POPs and HSPs have to be combined with the old POPs from 2006–2010 samples.

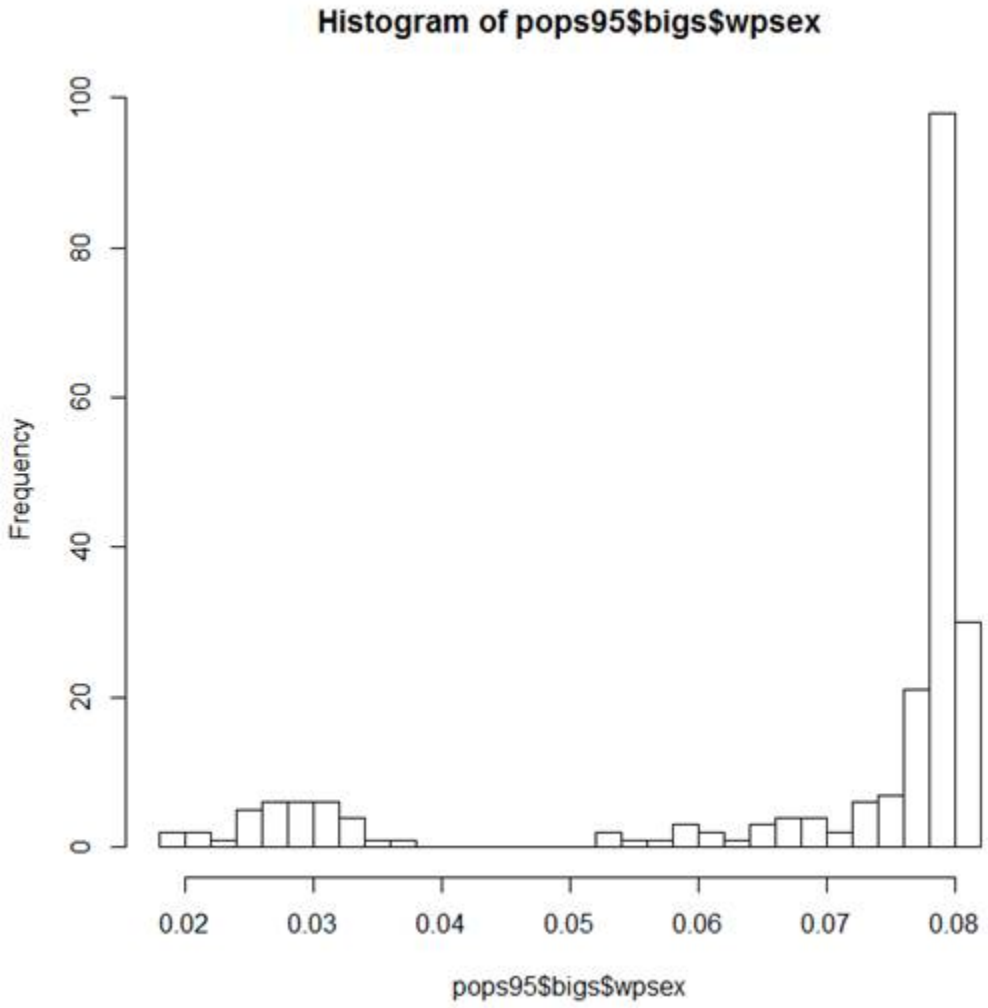
2. POP RESULTS

The microsatellites used in the first round of SBT CKMR were adequate for finding POPs using Mendelian-exclusion principles; see long appendices in Bravington et al., 2014. However, a lot of statistical processing and lengthy explanations were entailed to control false-positive rates and demonstrate that false-negatives must be rare; in short, we did have enough microsatellite loci to find POPs reliably, but only just enough. The DArTcap genotyping has been designed with the goal of identifying HSPs, which is much harder than finding POPs; consequently, finding POPs ought to be easier and clearer now.

As in 2012, we again identified POPs using a classification statistic based on Mendelian-exclusion, but some changes to the method were required to deal with the new features of DArTcap data (see Appendix B for details). Figure 2.1 shows part of the histogram of the modified exclusion statistic, referred to as the Weighted-PSeudo-EXclusion (WPSEX) statistic, across all DArTcapped adult-juvenile pairs (about 66,000,000 comparisons). The POPs are visible as a small bump on the LHS. Most of the entire histogram (to the right) has been left out here, because otherwise the true POPs are too few

[1]We deliberately do not check for POPs if the adult was caught in the season the juvenile was spawned; see previous CCSBT documents for reasons.

FIGURE 2.1. POPs via weighted-pseudo-exclusion (WPSEX) statistic; see Appendix B for details. Low values indicate POPs. X-axis truncated at 0.08 to omit the gigantic peak of UPs off to the right.



compared to the gigantic bump of UPs whose peak is around 0.116 (exactly where theory predicts it should be, based on allele frequencies of each locus). The giant bump drops off very quickly to the left of ~0.08, and the flattish tail around 0.055–0.075 will contain a number of adult/juvenile HSPs or GGPs (Grandparent-Grandoffspring Pairs), which should be somewhat rarer than true POPs on demographic grounds.The POPs are *clearly* separated from non-POPs— this is much more obvious with DArTcap data than it was with our microsatellite data. The 1500 low-information SNP loci from DArTcap are performing better than 25 high-information microsatellite loci, at about half the cost.

As per Table 1, this uses only adults from 2010 onwards, and excludes the POPs already found via microsatellites. However, as a check we also DArTcapped those particular pairs-of-samples already identified as POPs in 2012 study, and all of them clearly came up as POPs this time too. Interestingly, we also DArTcapped one curious adult/juvenile pair from 2012 which was clearly not a POP according to microsatellites but nevertheless remarkably close (just two unambiguous Mendelian exclusions in 25

TABLE 2. POP breakdown (new + old). Row = Juvenile birth-year; column = adult capture-year. From OMMP/1706/4.

	2006	2007	2008	2009	2010	2011	2012	2013	2014
2002	0	0	0	0	0	0	0	0	0
2003	0	5	1	2	0	0	0	1	0
2004	0	2	0	0	3	0	0	0	0
2005	1	4	5	4	1	0	1	1	2
2006	0	4	3	2	0	0	0	0	0
2007	0	0	3	4	1	3	2	0	2
2008	0	0	0	0	0	1	1	1	0
2009	0	0	0	0	0	1	1	1	0
2010	0	0	0	0	0	3	1	4	0
2011	0	0	0	0	0	0	1	2	1
2011	0	0	0	0	0	0	0	1	1

loci compared). The DArTcap WPSEX statistic for this pair was around 0.06, consistent with being a GGP or HSP.

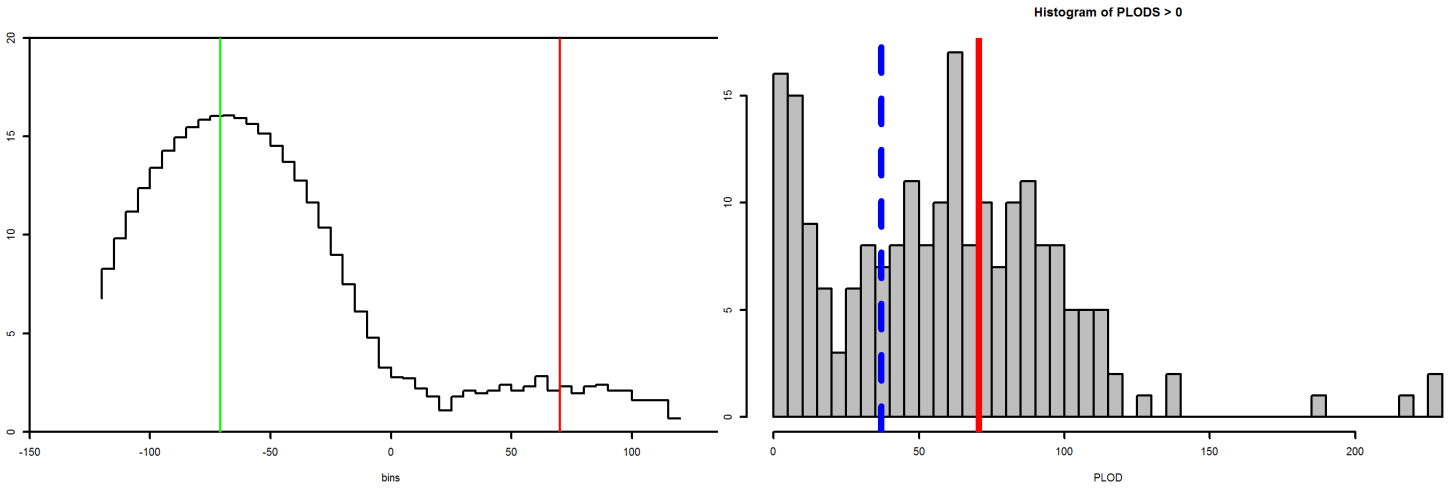
The POPs found this time appear generally consistent with previous results; see Hillary et al., 2017a from which Table 2 is taken. No POPs were found where the parent was caught in the same year as the offspring was born (such comparisons are excluded from the model anyway, and we did not check this in 2012). As in 2012, we also read the ages of all adults in POPs using otoliths collected at the same time as genetic sampling in Indonesia. The modal age of parents (at time of offspring's birth) was around 13 or 14. All bar one of the parents were 8yo or more at offspring-birth, just as in 2012; this time, though, one 7yo parent was inferred (14yo at capture, 7 years after its offspring was born). Uncertainty of ±1year in otolith-derived ages is not uncommon, so this may just be an age-reading error. We are still in the process of confirming that the parental otoliths do correspond to the genetic samples (there is at least one clear error, where samples must have been switched).

As to the apparent skip-spawning among younger adults suggested by the 2012 data: only two of the new POPs involve an adult caught at age 12 or less, but in both cases the number of years between offspring-birth and adult-capture is even, as found in 2012. Note that the occurrence of some 4yo among our newer juveniles (section 5) will make this pattern harder to detect in future.

3. HSP RESULTS

Among 10,809 juveniles, we found 140 definite HSPs (and 4 FSPs). The true number of HSPs is expected to be about 10% higher than 140, because of false-negatives— an inevitable (and expected) consequence of the statistical criteria used to ensure exclusion of all false-positives, and something that is easily allowed for in modelling. The HSPs and FSPs are quite clear graphically (Figure 3.1), and the distributions of the "PLOD" test statistic match the predictions of genetic theory, indicating that our new genotyping-and-HSP-finding processes are working reliably. The details, which are fairly complicated, may be found in Appendix C.

FIGURE 3.1. HSPs. Left: **log** histogram to show all PLODs, for every pairwise comparison of juvenile SBT (about 58,000,000 PLODs in all). Green and red lines are theoretical means for UPs and HSPs respectively. Right: actual histogram of those PLODs that are above zero. Blue dashed line is cutoff, chosen visually to exclude false-positives.



The proportion of HSPs where both were caught in the same year is somewhat higher than would be expected under a completely random breeding scenario. [2]This is evidence of "lucky litters", i.e. variable survival between spawning events[3] within each cohort— which is also the only way to explain the 4 FSPs[4]. However, SBT are overall clearly not a "sweepstake reproduction" species; the proportion of juveniles in same-cohort HSPs is still very small (<1%), confirming the conclusion of the 2012 CKMR analysis that it is a reasonable approximation to treat all POP comparisons as statistically independent. Note that same-cohort HSP comparisons are not used in our CKMR models; the HSP information comes entirely from cross-cohort comparisons.

From analysis of mtDNA, the Maternal/Paternal proportion in the HSPs is close to 50:50 (i.e. whether the shared parent is the Mother or the Father), both for same-cohort and cross-cohort HSPs. The mean number of cohorts separating each HSP is very similar for Maternal vs Paternal HSPs, so that the "SSB turnover rate" must be similar for both sexes— something which is not biologically obvious in advance. This validates an assumption underlying the exploratory use of combined-maternal-and-paternal HSPs in the 2017 OM, where adult sexes are not distinguished.

[2]In a "completely random breeding" scenario, every juvenile that we sample would have "randomly selected" its Mother *independently* from the pool of potential Mothers (weighted according to their relative fecundities), and likewise its Father. The key word is "*independently*" (i.e. between juveniles in our sample). This *does* apply to juveniles in *different* cohorts, but the HSP data show that is not entirely true for juveniles in the *same* cohort.

[3]Each SBT on the breeding ground spawns on many nights per year. Post-fertilization larval survival rates may well differ between nights.

[4]The chance of a female breeding twice independently with the same male is inverse to adult abundance, so *cross-cohort* FSPs should be about a million times rarer than HSPs. The same applies to same-cohort FSPs, unless some spawning-events (where one female and a small number of "courting" males all release eggs and sperm together) have higher post-fertilization survival than others. Unsurprisingly, all 4 FSPs that we found are same-cohort.

4. Extending CKMR models to work with HSPs as well as POPs

This section presents some of the theory behind POPs and HSPs in CKMR, in order to explain why and how HSPs and POPs should both be used for SBT CKMR. Much of this has already appeared in various CCSBT papers between 2012 and 2016, and a more systematic presentation is given in (Bravington et al., 2016b). This section attempts to give a less formal explanation that highlights the different types of information that POPs and HSPs contribute. The basic points are:

- Adult mortality does not appear in the equations for POPs[5], and so cannot be estimated from POPs alone.
- Length- or age-composition data does not normally solve the problem, because those are affected by selectivity as well as mortality.
  - For SBT, though, there are special features of reproductive biology and the fisheries which do permit adult mortality to be estimated just from POPs and length-compositions— but, this relies on an untestable assumption that may not be fully valid.
- The equations for HSPs *do* involve mortality, so HSPs provide statistical information on average adult mortality rate.
- HSPs alone are not enough (for teleosts) because they provide no way to estimate fecundity-size relationship; POPs are normally not enough because they provide no way to estimate mortality, even with length- or age-composition data. Together, POPs, HSPs and length/age-compositions provide enough information to estimate all the important parameters of adult population dynamics.

Note that some of the above points do *not* apply to species where fecundity is fairly constant throughout adults, e.g. most mammals and some sharks— for such species, it is sometimes fine to use just POPs or even just HSPs. However, the concluding dot-point is generally true for teleost fish.

Bravington et al. (2016b) sets out a general mathematical framework for reliably using CKMR, both for POPs and HSPs. In principle, every possible pair of fish is compared for each kin-type[6][7].The demographic probability of each particular pairwise comparison giving a "yes, these two fish are a kin-pair!" can be computed from formulae using (i) the known covariates for the pair of fish (age, year of sampling, etc) and (ii) working values of demographic parameters (e.g. abundance-at-age in the first year of the model, etc). In order to compute each of the pairwise probabilities (for maybe 100,000,000 comparisons, though the number of *distinct* types of comparison is much lower, because there are only a few covariates (age/year/sex etc)), the demographic parameters are first used to fill out a population dynamics model through time, just as in a standard stock assessment, as well as to

[5]Except if live-releases are used.

[6]In practice, some comparisons are *a priori* not worth making, either because they cannot occur (e.g. POPs between lethally-sampled juveniles) or because they are so unlikely to yield a kin-pair that they have very little informative power, and it is not worth the effort to extend the model to deal with them (e.g. POPs purely among adult SBT, which are fairly rare and would require extending the demographic model several decades into the past).

[7]An underlying assumption is that kin-finding from genetics is working accurately, or at least with no false-positives and predictable false-negative rate. The main point of this paper is to demonstrate that we have been able to do that successfully for SBT.

compute other necessaries such as growth curves and fecundity-at-size curves. The actual *result* of each comparison (yes or no, as determined by the genotypes) is then used to compute the log-likelihood for those particular working values of demographic parameters, treating each comparison as a single-sample Binomial random variable. This constitutes the CKMR log-likelihood. To this are added any other likelihood terms that also depend on the demographic parameters (such as from length- and age-composition samples, in the SBT stand-alone CKMR model), plus log-priors ("penalties") on random-effect-parameters such as annual recruitment deviations, plus any full-Bayes log-priors if desired. This forms the overall "objective function" in a maximum-likelihood/REML/Bayesian estimation framework, from which all the demographic parameters may be estimated in standard fashion.

The pairwise probability formulae in section 3 of (Bravington et al., 2016b) can be implemented, in a species-specific way, by considering ERRO (Expected Relative Reproductive Output). For example: in a given pair that might be a Mother and Offspring, the ERRO is:

> *how many offspring the potential Mother might have had relative to the total production from other female adults, at the time that the potential Offspring was born.*

This depends on individual-level covariates, such as when the potential-Offspring j was caught and how old it was then (since that determines when j was born). Clearly, too, i cannot be j's Mother unless she was alive and mature when j was born. In the simplest case of a species where all adults (of given sex) have similar fecundity after reaching maturity— e.g. most whales— the ERRO in words translates to the following formula in symbols:

$$\mathbb{P}\left[K_{ij} = \mathrm{MO} | z_i, z_j\right] = \frac{\mathbb{I}\left[y_i + \alpha \leqslant y_j < t_i\right]}{N_{\mathbb{Q}y_j}} \tag{4.1}$$

where $\mathbb{I}[\cdot]$ is the indicator function, t is year of capture, y is year-of-birth, and α is age-at-maturity. The notation perhaps makes the equation look more complicated than it really is. Importantly, mortality does *not* appear directly in this formula; we return to this below.

When fecundity $\beta(\cdot)$ varies with length ℓ, as it does for SBT, then (4.1) is a little more complicated, because not all adults have the same expected reproductive output (so the denominator of the ERRO is not just proportional to the *number* of adults, but also depends on their age composition), and because fish grow throughout adulthood (so i's length at capture will be different to her length at j's birth). The formula is (3.6) in (Bravington et al., 2016b); it is reproduced here, omitting the lengthy description of notation. There is some computational awkwardness in predicting i's fecundity back at j's birth based i's length when she was caught (and her age, if known), because fish have individual growth curves, and do not all follow the cohort average. However, the concept is simple.

$$\mathbb{P}\left[K_{ij} = \mathrm{MO} | z_i, z_j\right] = \frac{\mathbb{I}\left[y_i + \alpha \leqslant y_j < t_i\right] \times \beta\left(\ell_i\left(y_j\right)\right)}{\sum_{a \geqslant \alpha} N_{\mathbb{Q}ay_j} \int \beta(\ell)\, f\left(\ell | a, y_j\right) d\ell} \tag{4.2}$$

Equation (4.2) was the basis of the 2012 and 2013 CCSBT reports ((Bravington et al., 2012; Bravington et al., 2013)), and the SBT OM uses a similar age-based formula (also aggregating male and female

parents). Fecundity-at-size $\beta(\ell)$ can basically be estimated in a "model-free" way, by checking how the proportion of sampled adults that are *parents* changes with adult body size. Comparisons involving larger adults generate POPs more often than comparisons involving smaller adults. Again, though, the mortality rate z does not appear explicitly in (4.2).

The absence of z has both good and bad implications. On the plus side, since POP data can be organized as a time-series (using juvenile birth-year as the index), it would be possible to track *relative* changes in the population's total reproductive output— the denominator of (4.2)— in a "model-free" way without knowing or estimating z, by looking at how the proportion of POPs changes over time. However, the POPs themselves do not provide any information on z, so the latter cannot be estimated from POP-based CKMR alone. And the total reproductive output on the denominator depends not just on the *number* of adults, but also on the true population-level *length-* or *age-composition*, which is of course affected by z; so absolute N cannot be estimated either when fecundity depends on size. To resolve this, it is necessary to somehow estimate the adult age composition by bringing in other data[8].

The obvious source is age-composition data from a fishery. But, if selectivity is age-dependent, the sampled age-composition may be quite different to the real one. The slope of the true age-composition is basically set by the adult mortality rate (plus any trends in cohort strength measured at maturity, which generally does not have a strong effect), but the slope of the sampled age-composition is also affected by selectivity. Since selectivity is not known a priori, length/age-composition data on its own does not help. Appendix A of (Bravington, 2014) uses a simplified example to show exactly what the issue is; in every aspect of the data, the slope of selectivity-at-age and the adult mortality rate are always added together, so there is no information to separate them.

This z-versus-selectivity difficulty is in fact widespread in fish stock assessment; if total catches are known, then it manifests itself as "m cannot be estimated", often leading to unsatisfactory workarounds. Although there are situations where some information is available— e.g. strong contrasts in catch, strong contrast in cohorts— it is highly desirable to have an independent way to estimate at least the aggregate.

4.1. **Why was POP-only CKMR possible for SBT?.** In the absence of SBT CKMR, it has been necessary to assume that Indonesian selectivity (within sex) is directly proportional to residence time on the spawning grounds, i.e. that catchability-per-day is independent of length (within sex). By making this assumption, and combining it with external data on daily spawning output as a function of (female) size, we enforce a "hard link" between selectivity and fecundity. Since fecundity-at-size can in principle be estimated directly from POP data without needing to know mortality, as above, it becomes possible to estimate Indonesian selectivity indirectly. Then the effect of selectivity can be "subtracted" from the Indonesian age/length-composition curves; the remaining slope is the mortality rate. Of course, all this is actually embedded in a proper likelihood-based statistical estimation framework where parameters are estimated simultaneously, but the conceptual information flow is as just described.

While it is hard to argue with the notion that residence time on the spawning grounds must be a primary driver of selectivity for SBT caught on the spawning grounds, it is also hard to argue that it

[8]If fecundity does not depend on size, the adult age composition does not matter, and N can be estimated more directly.

must be the *only* important driver. In fact, there is evidence that there are other factors which can affect Indonesian selectivity, at least to some extent:

- strong year-to-year shifts in the Indonesian length-composition in some years. Residence time is presumably biologically-driven, so would not shift in this way.
- differences in SBT length-composition according to depth-of-hook-setting in Indonesian long-lines(Farley et al., 2015); again, this suggests some length-driven behavioural effect that could affect selectivity.

There is one additional independent piece of information on SBT adult z: the long-term slope of numbers-at-age for fish above 30 (CCSBT-OMMP2). By age 30, SBT growth has largely stopped, so *length*-based selectivity should not affect the *age*-slope for old fish[9]. However, one of the main conclusions from that data is that m must by higher for old "senescent" fish than for younger adults, so by definition the dataset is not quantitatively helpful in setting the level of m (or z) in younger adults.

4.2. **HSP principles.** HSPs in CKMR are handled using exactly the same idea as POPs: at birth, every Offspring had precisely one Mother and one Father. Consider calculating the probability that two juveniles, "Lucy" and "Peter", are Half-Siblings, and for definiteness assume that Lucy is born two years before Peter. The chance that any particular female, "Mary" say, is Lucy's mother is given by Mary's ERRO when Lucy was born, i.e.:

> *how many offspring Mary might have had relative to the total production from other female adults, when Lucy was born.*

In order for Mary to also be the Mother of Peter, Mary has to survive the next two years— this is why mortality appears directly in the HSP equations. If Mary does survive, then she will also have grown, and so she will be more fecund. Out of all the available mothers for Peter, the chance that Mary will be the one is

> *how many offspring Mary might have had two years later when Peter was born, relative to the total from other female adults then.*

Mary's ERRO at Peter's birth involves two random components (given her size etc at Lucy's birth): survival, and reproductive output conditional on being alive.

The only complication in the equation is that "Mary" could have been any of the females alive at Lucy's birth— we don't know which one. So, unlike for POPs, the probability that Lucy and Peter are (maternal) HSPs has to be summed over all possible females. The general and formal version is equation (3.9) in (Bravington et al., 2016b):

$$\mathbb{P}\left[K_{ij} = \mathrm{MHS} | y_i, y_j, R\right] = \sum_{d \in \mathcal{F}_i} \left\{ \frac{R_d\left(y_i\right)}{R_+\left(y_i\right)} \times \frac{R_d\left(y_j\right)}{R_+\left(y_j\right)} \right\} \tag{4.3}$$

See that reference for notational details, but it is worth pointing out that $R_d\left(y\right)$ is a random variable meaning "actual reproductive output" of a particular animal in year y, which will be zero if the animal

[9]For no particular reason, this dataset was actually not used in (Bravington et al., 2012; Bravington et al., 2013). However, it is included in the OM version of (Hillary et al., 2013) and subsequently.

is not alive at y. If Lucy and Peter are born in different years, then $R_d\left(y_{\text{Lucy}}\right)$ and $R_d\left(y_{\text{Peter}}\right)$ refer to different random events, and will be statistically independent *provided that* the equation is correctly specified to include all important covariates of potential mothers, such as age and length. Consequently, the expected value of their product is the product of their expected values, which makes the equation computable. But if Lucy and Peter are born in the same year, then the two R_d's refer to the same event, so that the overall probability is affected by variance as well as mean. To handle that explicitly, extra unknown parameters would be needed, so same-cohort juvenile comparisons carry no useful *demographic* information; if juvenile ages are known accurately, the simplest solution is just to omit such comparisons from a CKMR model[10].

The specific version of (4.3) for SBT, taking account of length-based fecundity as well as age, contains so many symbols that including it would confuse more than clarify. However, it is based entirely on the principles above, and the computation is not hard— most of the quantities required are already computed for the POP probabilities. The main practical difference in a full implementation is that, for HSPs, individual fecundity needs to be projected *forward* in time (from Lucy's birth to Peter's), whereas for POPs, fecundity is projected *backwards* in time (from Mary's capture backwards to Lucy/Peter's birth). There are two qualitative points to note:

- The per-comparison probability of finding an HSP, as a function of the birth-interval between the juveniles, declines because of mortality, but this is partly mitigated by individual growth of parents that do survive. Thus, the rate of decline in HSP probability with gap-length is a measure not of mortality directly, but rather of *SSB turnover rate*. In combination with fecundity-at-age from POPs, and length/age-composition data, there is enough information to disentangle the main effects of fecundity, mortality, and selectivity.
- HSPs only provide information on average adult z across all cohorts, weighted by fecundity-at-age; there is no intrinsic information on age-specific z.
- As for POPs, the overall trend over time in HSP-finding rate will change inversely to the trend in SSB (from the total reproductive output term R_+ in the denominator of (4.3)).

4.3. **Absolute abundance and HSPs.** Like POPs, HSPs also carry information about absolute adult abundance. However, slightly more assumptions are required with HSPs for the *absolute* estimates to be unbiased; for example, perpetually infertile adults would be invisible to an HSP analysis even if they turn up on the spawning ground (and if they don't even turn up on the spawning ground, then they are also invisible to POPs and to the fishery and to demography in general). One *possible* example for SBT could be if some adults persistently produce offspring that never go to the GAB. Appendix A2 of (Bravington et al., 2015) explains further, and also describes how to make a CKMR model robust against this kind of phenomenon. Basically, since POPs alone can provide perfectly good absolute estimates (once the effects of age, fecundity, mortality are accounted for by bringing together POPs, HSPs, and age/length-composition data), a safe starting point is to introduce an extra free parameter q_{HSP} that acts as a fixed multiplier on all HSP probabilities; thus, only relative changes in HSP-rate are used. If

[10]However, within-year HSPs can be useful for checking that there is no strong excess of siblings that would mess up the independence assumption for POP comparisons.

the estimate $\hat{q}_{\text{HSP}}$ turns out close to 1, then there is no evidence of any hidden peculiarities in the adults. This is the approach taken in **OM2017

4.4. **Including enough covariates.** An important principle of CKMR is that, to avoid bias, it is necessary to include all "relevant" individual-level covariates when computing the kinship probabilities— and if a covariate measurement is missing for some individuals, to integrate over its possible values. Failing to do this will lead to biased abundance estimates. The extent of any bias will depend on circumstances; the most extreme mistake would be to misapply the "cartoon" estimator $2 \times$ #comps/#POPs to a multi-year multi-cohort teleost-like setting, in which case the bias can be as large as you want. The solution is to split the probability calculations into different cases, according to the individual covariates. (Bravington et al., 2016b) section 3 gives a more precise description of what "relevant" might mean.

For POPs, one way to think about this is that there should be no"unmodelled" correlation between sampling-probability of an adult and that adult's reproductive output. For SBT (and presumably most teleosts), this means that length *as well as* age is important: bigger-than-average SBT within a cohort will tend to generate more offspring (so they will receive more "tags" from their offspring), and are also more likely to be caught (at any given age) because of length-based selectivity.

Similar considerations apply to HSPs, except that the unmodelled correlation to be avoided is between separate reproductive events from the same individual (i.e. the shared, but hidden, parent), rather than between one reproductive event and one capture event. The comments in section 4.3 can in fact be seen as a special case of this, whereby a "GABby/non-GABby" binary covariate may apply to some adults; the covariate cannot be measured directly, but parameter q_{HSP} can be seen as an estimate of the frequency of "GABby" adults. (Having said that, q_{HSP} could encompass several types of unmeasurable peculiarity, so it would be over-interpretation to infer too much about GABbiness per se based just on $\hat{q}_{\text{HSP}}$.)

The stand-alone CKMR in (Bravington et al., 2012; Bravington et al., 2013) uses a full length- and sex- and age-structured model for adults, so as to ensure that all relevant covariates are being captured in the model. (It can afford that level of complexity because it does not have to deal with the complications of ages below 8yo or of other datasets apart from Indonesian length/age-compositions.) The OM version in (Hillary et al., 2013) and subsequent iterations uses only an age-based formulation. This is a source of positive bias in the OM abundance estimates (at least from CKMR sources). Specifically, the probability that an adult of given age will feature in a POP is the mean-of-the-product (i.e. mean taken across lengths) of adult sampling-probability and expected reproductive output. This is mathematically guaranteed greater than the product-of-the-means because both sampling-prob and offspring-output are correlated in the same direction with length. The age-only CKMR equations in the OM implicitly calculate the product-of-the-means, so its computed POP probabilities are biased downwards; consequently, this is a source of positive bias in estimates of N. The bias is presumably not enormous, since the stand-alone and OM estimates of N were fairly similar, but will need to be addressed at some point.

Aggregation of sexes in the OM runs into similar issues, although the direction of bias if any is not obvious.

Once the results of a POP-and-HSP stand-alone CKMR model become available, then it should be possible to compensate for it in simpler models. For example, we could estimate the discrepancies between mean-product and product-mean based on the stand-alone results, and use them as fixed inputs to a purely-age-based OM-style model that re-estimates everything else but has effectively discarded the adult length data (so that data is not being used twice).

5. Next steps for SBT

The new POP and HSP data from 2006—2015 have only been available for a very short period— the project schedule was extremely tight, the genotyping took a little longer than expected, and we placed a high priority on thorough QC and diagnostics, which are time-consuming steps. Although it has been already been possible to update the OM to handle HSPs(Hillary et al., 2017b), we have not yet had enough time to finish a stand-alone POP-and-HSP CKMR model. However, the stand-alone CKMR model will be complete by the end of 2017, and will be available for review at OMMP9 and consideration by the ESC in 2018.

By "stand-alone CKMR model", we mean a model that:

- uses a detailed population dynamics framework for adults, involving length and sex as well as age (the OM currently uses age alone);
- uses Indonesian length- and age-composition data in addition to the POP and HSP information
- but uses no other data (e.g. no GAB indices, no CPUE, no total catch)

While the OM version is of course critically important to CCSBT, the stand-alone model also plays a key role because it is focused on the Commission's primary objective of rebuilding the SBT spawning stock. As explained in section 4.4, the detailed population dynamics model is the only way to be sure of avoiding bias in CKMR, and thus of fully checking the reliability of a simpler CKMR formulation such as the current OM.

There are three main tasks for the year ahead:

- Completion of the stand-alone CKMR model, extending (Bravington et al., 2012; Bravington et al., 2013) to include HSPs, plus a few other details mentioned below.
- Contingent on the stand-alone results, an update to the OM that can take into account qualitative and quantitative results from the stand-alone model.
- Based on the new data and models, a review of ongoing CKMR sample size requirements needed to deliver:
  (1) good precision for OM and stand-alone estimates;
  (2) reliable performance from a future MP.

5.1. **Modelling tasks for updated stand-alone CKMR with HSPs and POPs.** This is a brief checklist of the specific issues requiring attention. The model structure will be largely the same as (Bravington et al., 2013), except as noted below. The extra HSP terms explained in previous sections do not entail any change to the basic structure used for POPs— the key quantity for HSP probabilities is again ERRO (Expected Relative Reproductive Output(Bravington et al., 2016b)) and the calculations

are very similar for HSPs as for POPs. Changes to the other parts of the model, in order of decreasing importance, are:

- De-couple fecundity-at-size from selectivity-at-size in the model (these phenomena were previously hard-linked for females, as described in section 4.1). Some experimentation with functional forms will be needed.
  - Also, allow for changes in Indonesian selectivity from year-to-year. Modern developments in random-effect modelling suggest a straightforward way to do this.
- It has become clear that not all our juvenile CKMR samples are 3yo, based on inspection of length frequencies. In the first few years of CKMR up to about 2010, we were in the fortunate position of sampling close to a clear length-mode of 3yo, but in some subsequent years the modes have moved and show substantial overlap between cohorts. Since juvenile birth-year is especially important for HSPs (e.g. in excluding or allowing for same-cohort comparisons), we will need to address this properly. We are looking into ways to monitor the age-composition better in future; from a modelling perspective, it can be accommodated thru a random-effect term on ppn-3yo-by-year, but there is some cost in lost information.
- Increase the plus-group age to 30yo, for consistency with OM
  - and for 30+yo fish— which have almost stopped growing, so that age per se has little effect on selectivity above 30yo— incorporate a Z-estimate taken directly from long-term age compositions.

Acknowledgements

Thanks are due to many who made this work possible.

Bali, Indonesia: The Research Institute for Tuna Fisheries (Bali) for collaboration on sampling adults from spawning ground;

Port Lincoln, Australia: the late Tony Jones (Protec Marine P/L) and Adam Kemp at Seatec P/L, for sampling of juveniles; and the Australian Southern Bluefin Tuna Industry Association (ASBTIA) for ongoing support;

Canberra: CCSBT and FRDC for funding; Andrzej Kilian and all at Dart Pty Ltd for DArTcap genotyping;

CSIRO: Matt Landsdell, Jess Farley, Jorden Stephenson for for great efforts in logistics to complete all the processing in time.

References

Bravington, M (2014). *Close-Kin Mark-Recapture for SBT: options for the longer term*. Report for Scientific Committee CCSBT-ESC/1409/44. Commission for the Conservation of Southern Bluefin Tuna.

– (2017). *SBT kin-finding and genotyping update*. Tech. rep. CCSBT-OMMP/1706/?? Commission for the Conservation of Southern Bluefin Tuna.

Bravington, MV, PM Grewe, and CR Davies (2016a). "Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture". In: *Nature Communications* 7.

Bravington, MV, HJ Skaug, and EC Anderson (2016b). "Close-Kin Mark-Recapture". In: *Statistical Science*.

Bravington, M, P Grewe, and C Davies (2012). *Report of the Close-Kin Project: estimating the absolute spawning stock size of SBT using genetics*. Scientific Committee Report CCSBT-SC/1208/19. Commission for the Conservation of Southern Bluefin Tuna.

– (2013). *Close-kin update*. Scientific Committee report (OMMP workshop) CCSBT-OMMP/1307/Info1. Commission for the Conservation of Southern Bluefin Tuna.

Bravington, M, R Thomson, and C Davies (2017). *Genotyping issues for CKMR on Atlantic Bluefin Tuna*. Atlantic-wide research programme on Bluefin Tuna (ICCAT GBYP– PHASE 5) GBYP xxx/2016. International Commission for the Conservation of Atlantic Tuna.

Bravington, M, J Eveson, P Grewe, and C Davies (2015). *SBT Close-Kin Mark-Recapture: options for the medium term*. Tech. rep. CCSBT-ESC/1509/19. Commission for the Conservation of Southern Bluefin Tuna.

Bravington, M, P Grewe, and C Davies (2014). *Fishery-independent estimate of spawning biomass of Southern Bluefin Tuna through identification of close-kin using genetic markers*. FRDC Report 2007/034. CSIRO, Australia.

CCSBT ESC (2015). *Report of the Twentieth Meeting of the Scientific Committee*. Tech. rep. Commission for the Conservation of Southern Bluefin Tuna.

Farley, J, T Davis, M Bravington, R Andamari, and C Davies (2015). "Spawning dynamics and size related trends in reproductive parameters of Southern Bluefin Tuna, *Thunnus maccoyii*". In: *PLoS One* 10.5.

Hillary, R, A Preece, and C Davies (2013). *Updates to the CCSBT operating model including new data sources, data weighting and re-sampling of the grid*. Tech. rep. CCSBT-ESC/1309/15. Commission for the Conservation of Southern Bluefin Tuna.

Hillary, R, A Preece, and C Davies (2017a). *Updates required for new data sources and reconditioning of the CCSBT OM*. Tech. rep. CCSBT-OMMP/1706/4. Commission for the Conservation of Southern Bluefin Tuna.

Hillary, R, A Preece, and C Davies (2017b). *Summary of initial incorporation of the Half-Sibling Pair (HSP) data in the CCSBT Operating Model*. Tech. rep. CCSBT-ESC/1709/?? Commission for the Conservation of Southern Bluefin Tuna.

Appendix A. Genetic summary

Bravington et al., 2015 suggested that 1500 loci (of similar quality to the preliminary set of around 700 "focused" SNPs that we had tested at that point) should be adequate for HSP-finding in SBT. The sequencing datasets from DArTcap that we received in 2017 actually comprised nearly 2400 loci, but we discarded hundreds either because they were uselessly uninformative (one very-dominant allele) or on QC grounds, e.g. possibility of paralogs— "better safe than sorry". This left us with 1541 loci that work

as they should— close to our original plans. While people usually think of "SNP loci" as being purely biallelic, our loci are actually sequences of ~75-base-pairs, and many of the loci have several mutation sites within that sequence. To keep computations feasible, though, we chose only two "allowed alleles" (sequence variants, arbitrarily labelled A and B— also sometimes called "haplotypes") at each locus.

Null alleles[11] are very common in the SBT genome, and this would normally reduce the power to detect kin-pairs since homozygotes (XX) are conflated with single-null (XO) genotypes. However, read-depths are so high with DArTcap (typically several hundred reads per locus) that XO can usually be distinguished from XX based on the total reads of allele X for that sample-and-locus. This means that the nulls are serving as a 3rd allele, actually *increasing* the power to detect kin. Given the large number of samples needed for SBT CKMR, we would need considerably more loci (and expense) without this refinement. The idea is explained at greater length in Bravington et al., 2015.

Of the 1541 loci, we are using 1484 genotyped as just described to 6-way level (AB/AA/AO/BB/BO/OO). For the other 57 loci, the XX/XO separation did not work reliably enough, so we genotyped only to 4-way level, i.e. AB/AAO/BBO/OO where XXO means "either XX or XO". Examination of replicates indicates that error rates among the 4-way categories are very low— well under 1%. For simplicity, the POP-finding step uses only the 4-way genotypes. The HSP-finding step requires the extra information in the 6-way genotypes (for those loci where it is possible); the locus-specific error rates for XO/XX are substantially higher (10% is common), and due allowance needs to be made for the possibility of such errors.

Appendix B. POPs: a modified exclusion criterion for DartCap SNP data

Offspring inherit one allele from each parent at each locus; thus, a POP should share an allele at every locus. "Mendelian exclusion" uses this to test for POPs; in principle, failing to share an allele at any locus (here 1500 of them) means that a pair cannot be a POP. In practice, this needs softening a bit, because of null alleles, the possibility of genotyping errors and (with 1500 loci) even perhaps a mutation or two, all of which could lead to *apparent* exclusions even in genuine POPs.

The exclusion criterion we have used for the SNP loci in DArTcap is slightly different to that used for microsatellite loci in our original study, because SNP loci have only two "normal" alleles (alwass called A & B) and because null-alleles (called O) are very common in SBT. To explain further, section A provides background. Although a key part of DArTcap's lustre is the ability to resolve each genotype into the 6 possible cases AB/AA/AO/BB/BO/OO, for POP-finding we have used a simpler 4-way classification where AA/AO are merged (called "AAO") and similarly for BB/BO. Analysis of replicates shows that error rates within these 4 categories are very low— about 0.1% for most loci— and not having to worry about errors simplifies the calculations markedly.

[11]Some of the nulls in our final set of genotypes arise from our targeting only two variants at a locus, even if there are sequenceable 3rd, 4th, ... variants. Such "nulls" are easy to detect because the extra untargeted variants are visible in the dataset we receive from DArT Pty. However, most nulls in SBT are "genuine nulls"— particular sequence variant(s) that do not get sequenced at all. These are repeatable and heritable, and are presumably due to e.g. mutations in the restriction-sites targeted by ddRAD. Such nulls are not merely failures-to-observe a "normal" allele resulting from low read-depths, which can happen with some GBS methods.

If a locus has very few nulls, then finding that one fish is AAO means it is probably AA. If another fish is BBO, then it is likewise probably BB; so those two fish probably do have an exclusion at that locus. This is not definitive, because the first fish *could* be AO and the second *could* be BO and the O might then be co-inherited. Nevertheless, adding up the number of these "pseudo-exclusions" is an intuitively powerful way to separate POPs from UPs. It turns out that one refinement is necessary; pseudo-exclusions are more informative at some loci than others (depending on the frequency of nulls at each locus, and to some extent also on the frequency of A vs B alleles), and a weighted version is statistically much more powerful. In other words, an AAO/BBO pseudo-exclusion at locus ℓ receives weight w_ℓ (or 0 if no pseudo-exclusion), and for any pair of fish the Weighted-PSeudo-EXclusion (WPSEX) statistic is the sum of all w_ℓ. The weights are chosen to minimize the false-positive probability.

If there were more than two alleles, then other types of definite exclusion could be considered (e.g. AB vs CCO or AB vs CD), which is how we handled the microsatellite data— but this is by definition not an option for biallelic SNP data.

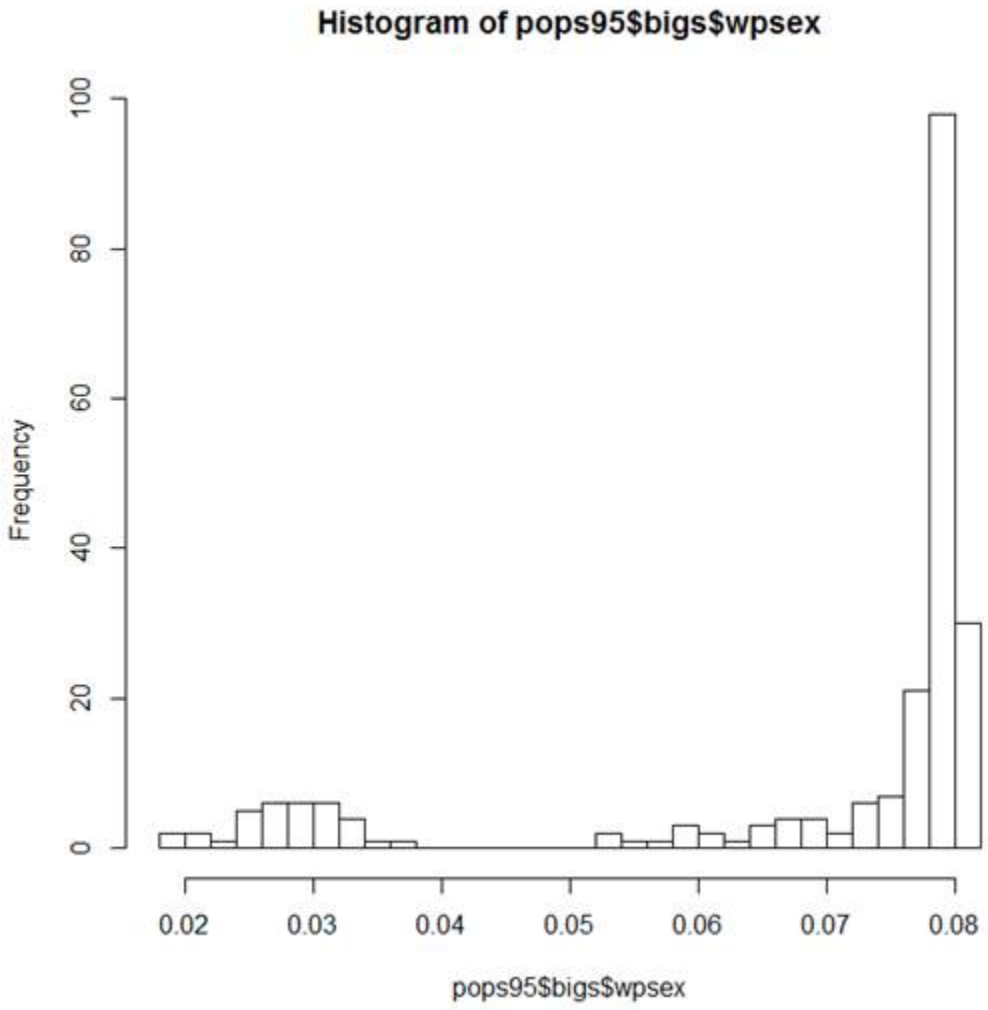
As noted in the main document, the POPs are clearly identifiable with this WPSEX statistic (Figure B.1).

It is also possible to find exclusions where one fish is AB and the other is OO. These are comparatively rare because we deliberately avoided loci with very high null-allele frequency, so on its own the number of AB-OO is not a good single criterion for POP-finding. Nevertheless, the results from n-AB-OO do nicely (and independently) back up the results from the preferred WPSEX statistic, in that POPs have many fewer n-AB-OOs (not shown). Although genotyping errors are rare overall, there are enough to generate a few AB-OO exclusions even among true POPs.

Many authors (including Bravington et al., 2016b) propose likelihood-based criterion for POP-finding, instead of exclusion. In principle, a likelihood-based criterion is more powerful (optimal, in fact) because it uses more information, even for loci where no exclusion is present; if, say, allele B is rare at one locus, then finding B in both animals of a pair increases the evidence in favour of their kinship, which is reflected in the log-likelihood but not an exclusion-based criterion. However, the problem with likelihood-based parentage is that in a perfect world exclusions should *never* happen, so even a single apparently-excluding locus will send the log-likelihood $\log \mathbb{P}\left[g_1 g_2 | K_{12} = \mathrm{POP}\right]$ to minus-infinity. In practice, though, apparent exclusions do occur occasionally because of genotyping errors and even mutation, and with 1500 loci the chance of this happening at least once per true POP is appreciable, so that there may be an unacceptably high chance of false-negatives in unadjusted likelihood-based parentage. Dealing with this would require special attention to allow for individually-rare errors, whose frequency may be hard to estimate a priori because they are rare. (Note that this problem does not apply to HSPs, where any combination of genotypes is possible at any locus, so that an error has far less impact on $\log \mathbb{P}\left[g_1 g_2 | K_{12} = \mathrm{HSP}\right]$.) On balance, we have preferred to stick with an exclusion-based criterion for SBT, for several reasons:

- simplicity;
- plenty of loci, so no need to produce an optimal method (unlike for HSPs, where every bit of data helps);
- no need to rely on estimated rates of *rare* errors.

FIGURE B.1. POPs via weighted-pseudo-exclusion (WPSEX). Low values indicate POPs. X-axis truncated to omit gigantic peak of UPs.



APPENDIX C. HSP DETAILS

Finding HSPs is more difficult than finding POPs, because the degree of kinship is weaker. In general, it is not possible to expect truly 100% reliable ID of HSPs, because the degree-of-relatedness varies randomly between different HSPs, and some pairs may be chance only be weakly related— so, no matter how thorough the genotyping, some overlap with UPs or weaker kin such as first-half-cousin-pairs (one shared grandparent) may be unavoidable. HSPs will be selected based on a statistic computed for each pair (called the PLOD; see box), with an UP likely to give a low value and an HSP likely to give a high value. What is important is to choose a cutoff value for the statistic, making it high enough to ensure that false-positives from UPs and from other weaker relatives are statistically negligible. Then, provided that the cutoff is well below the mean value expected for a true HSP, it is possible to allow for false-negatives in an unbiased way. The process follows our original plan in Bravington et al., 2015, Appendix C2; see also Bravington et al., 2016b section 5, and Bravington et al., 2017.

PLOD statistics for finding kin-pairs

Consider the genotypes $g_{1\ell}, g_{2\ell}$ of a pair of fish at just a single locus ℓ. The "LOD", or log-odds-ratio, measures the relative probability of those genotypes under two hypothetical kinships that the pair might have. In our study, the two kinships of interest are HSP and UP, and the LOD is

$$\text{LOD}_{\ell ij} = \log\left(\mathbb{P}\left[g_{\ell i}, g_{\ell j}|\text{HSP}\right]/\mathbb{P}\left[g_{\ell i}, g_{\ell j}|\text{UP}\right]\right) \tag{C.1}$$

The probabilities can be calculated from allele frequencies, under the assumption of Hardy-Weinberg Equilibrium (which we check using the entire sample of fish). If genotyping errors need to allowed for— which is certainly the case with the 6-way genotyping we need to get adequate HSP-finding power in SBT— see section A)— then the formula should be modified so that the g's are *observed* genotypes, rather than true genotypes. This modification obviously requires estimates of the error rates.

The PLOD ("Pseudo LOD") for the two fish is obtained summing all the per-locus LODs:

$$\text{PLOD}_{ij} = \sum_{\ell} \text{LOD}_{\ell ij}$$

Each LOD is statistically optimal for discriminating between the hypothesized kinships, using only the genotypes at that locus. If the loci were independent, the PLOD would in fact be a true overall LOD; however, loci are only independent when the pair is Unrelated because of genetic linkage (Bravington et al., 2016b section 5, or Wikipedia), so in general the "Pseudo" is needed. The extent of non-independence cannot be foretold, which is why the spread of the HSP bump cannot be predicted, although its location can be. For the UP bump, both the spread and location can be predicted.

In almost all of this paper, we focus on the two kinships "HSP or UP", and the term "PLOD" on its own specifically refers to that case. However, for other kinship discriminations, other versions of the PLOD can be defined accordingly, and will have better performance. In particular, we also use a "tuned PLOD" to help separate FSPs from HSPs (Figure C.2).

Note that the term "LOD score" is widespread in genetics, but often carries a more specialized meaning connected with the degree of genetic linkage between two loci. Here, we are just using the term "LOD" in its pure statistical sense.

In 2015 and 2016, we successfully implemented the approach at CSIRO for several shark species with very low abundances and a more basic version of GBS genotyping. However, to get things to work for SBT where ~100,000,000 juvenile-juvenile comparisons will eventually be made and false-positives from "lucky" UPs are thus much more of a potential risk, it is necessary to have high-precision genotyping (the 6-way null-scoring classification in section A) and a "sensitive" definition of the PLOD that allows for genotyping error. It also turns out that low-quality samples (contaminated and/or degraded DNA) can become a real problem for spurious HSPs with huge datasets, so extra care has been needed to filter them out. We are still refining that process, but have already managed to eliminated the problem in practical terms. To cut to the chase, Figure C.1 presents the PLOD statistic across all pairwise comparisons of 10,809 juveniles. Bigger PLOD values mean more relatedness; the HSP bump on the right is clearly visible, and pretty well separated from the morass of unrelateds on the left. For both bumps, the means/centres are very close to the theoretical predictions (red and green lines).

The four pairs on the far right of the RHS of Figure C.1 are Full-Sibs. In each case, both animals were caught in the same year and are thus likely to be from the same cohort; see also section C.1 for

FIGURE C.1. HSPs. Left: **log** histogram to show all PLODs. Green and red lines are theoretical means for UPs and HSPs respectively. Right: actual histogram of the PLODs that are above zero. Blue dashed line is cutoff, chosen visually to exclude false-positives.

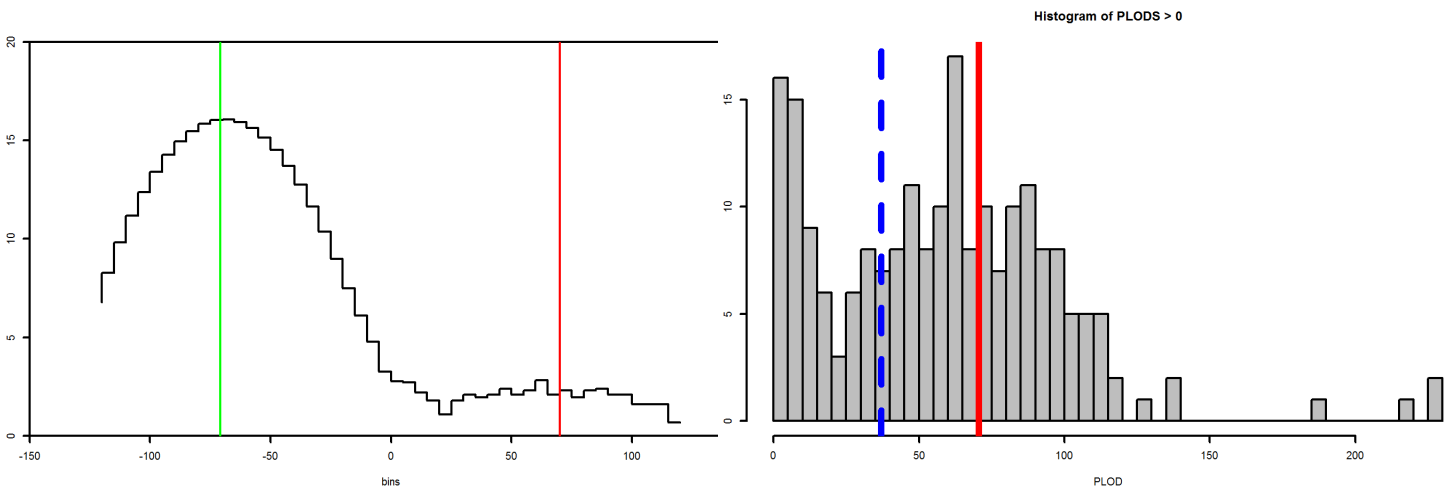
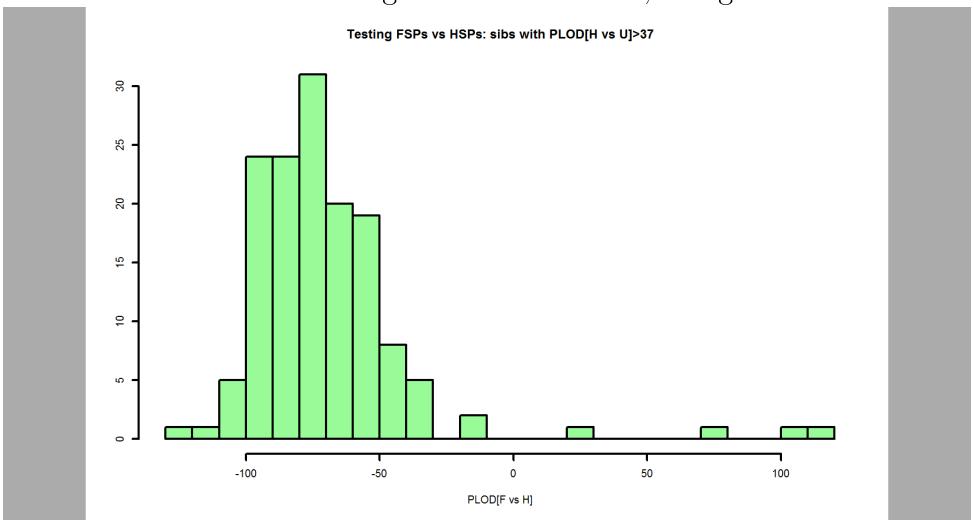


FIGURE C.2. FSPs distinguished from HSPs, using a tuned PLOD



mitochondrial evidence. Although they stand out quite clearly just using this particular PLOD statistic, which is actually optimized for distinguishing HSPs from UPs, it is also possible to develop a tuned-PLOD specifically for distinguishing FSPs from HSPs, as in Figure C.2 which shows the four FSPs very clearly. This tuned-PLOD would take the value 0 for a pair whose genotype had equal probability whether the pair was HSP or FSP. Thus, tuned-PLODs above zero are evidence in favour of FSPs, and values below zero are in favour of HSPs. Since the *a priori* proportion of HSPs is much higher, it would not be appropriate— in view of Bayes' theorem— to mindlessly take zero as a threshold. Nevertheless, zero does have a real meaning for tuned-PLODs, and the HSP/FSP distinction is clear. Note that this particular tuned-PLOD was actually calculated just using 4-way genotypes (whereas the HSP/UP PLOD uses 6-way genotypes), in order to save the coding time needed to allow for genotyping errors; hence, its discrimination power could be improved in future.

Returning to Figure C.1, the theoretical and empirical distributions of PLODs match very well. For UPs, both the theoretical mean and variance can be calculated, and match the data closely (this is obvious for the means). There is still a bump around PLOD˜=0 which comes from remaining lower-quality samples (the bump at 0 was much more prominent before applying the current bad-sample filters) and/or from less-closely-related kin-pairs. At any rate, that bump has clearly fallen away by three-quarters of the way to PLOD=50, at say PLOD=37. This might be used as a safe cutoff for false-positives. Using this choice of cutoff, the consequent false-negative probability can be estimated based on the theoretical mean (70.6, the red line) and the empirical variance of HSPs that are above the mean (see Bravington et al., 2016b, section 5; simulations show that the PLOD for true HSPs is quite close to a Gaussian distribution). This yields a false-negative estimate of 10.5% for a cutoff of PLOD=37. Actually, the false-negative rate is not critically important to inference for SBT, because the *number* of HSPs will likely not be used directly for abundance anyway (see Bravington et al., 2015 for explanation).

Using PLOD=37 as a cut-off for true HSPs, the following summary applies:

- 140 HSPs (possibly including one or two FSPs around PLOD=140, not yet checked).
- False-negative rate of 10.5%, so the true number of HSPs is probably[12] around 150–160. Note that simulations in Bravington, 2014 predicted about twice as many HSPs as POPs for SBT (in the short term), and there are a total of 84 POPs. The overall ratio depends on the numbers of juvenile and adult samples, which in practice were different to all scenarios of the 2014 simulations, for logistic and financial reasons.
- All pairs are distinct, except for three triads. In two of those triads, all the three animals are HSPs, i.e. they share the same parent; For one triad only, all three animals were caught in the same year. The third triad very clearly consists just of two pairs A-B and A-C (i.e. B and C are definitely not HSPs), so that e.g. A and B share a Mother, but A and C share a Father.
  Note that a couple of "accidental triads" are to be expected. About 280 of 10,000 juveniles are involved in HSPs, i.e. 2.8% of animals are in a pair; so, in about 2.8% of the 140 pairs found, the 2nd fish is actually likely to be from one of the other pairs— i.e. will form part of a triad.
- The pairwise tabulation of years (Table 3) shows no obvious pattern. Taking into account the number of comparisons (Table 4), though, it does look like there is a higher rate for same-year (i.e. same-cohort, mostly) HSPs than for cross-cohort HSPs, suggesting a *small* litter-effect whereby larvae from the same spawning-event may sometimes continue to associate through to the point of capture 3 years later. Overall, though, the proportion of juveniles involved is very small. This is compatible with the microsatellite results, which ruled out a *strong* litter-effect (i.e. it could not be big enough to have much impact on POP variance) but were not sensitive enough to estimate a low rate directly.
  Table 4 is the raw material for estimating adult mortality, based on changing rates of HSP-finding as the gap in years increases. However, it would be utterly wrong to do this "by eye" or even by

[12]This "likely true number" is purely for information; CKMR models actually use the actual observed number (140) together with the estimated false-negative rate, rather than guessing the true number.

TABLE 3. HSPs by years

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
2006	2	4	4	4	.	2		.	1	2
2007	.	6	3	6	2	2	2	.	.	2
2008	.	.	4	3	3	3	.	5	1	1
2009	.	.	.	8	6	1	3	7	4	.
2010	.	.	.	.	3	5	3	3	1	3
2011	.	.	.	.	.	6	1	1	2	3
2012	.	.	.	.	.	.	1	2	.	.
2013	.	.	.	.	.	.	.	2	1	2
2014	.	.	.	.	.	.	.	.	3	3
2015	.	.	.	.	.	.	.	.	.	3

TABLE 4. HSP rate per 10^7 comparisons (rounded, and based on small numbers)

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
2006	25	24	25	24	0	17	10	0	9	17
2007		72	18	36	12	17	19	0	0	17
2008			49	18	18	25	0	45	9	9
2009				94	36	8	28	62	36	0
2010					37	42	29	27	9	26
2011						141	13	12	25	35
2012							30	28	0	0
2013								53	13	25
2014									81	38
2015										72

spreadsheet; what the gaps actually show is not the mortality rate, but rather the *turnover rate* in SSB, since if Your parent does not die, it will grow instead and thereby become more fecund and more likely to generate a Sibling for You. While this can be accounted for conceptually using the information from POPs on fecundity-at-age and the length compositions, there is no reliable way to guess the results. The only thing to do is to build and fit a proper CKMR model.

C.1. **mtDNA in HSPs.** Male and female reproductive population dynamics may be different; in particular, the two sexes may differ in how fecundity changes with body size (and note that there are also growth differences between the sexes in SBT). Even if male and female adults are equally abundant, there is no guarantee that the "reproductive concentration" is the same between the sexes (consider e.g. sea-lions, where almost all females produce offspring per cycle, whereas reproduction is concentrated onto a small proportion of males). This affects the probabilities of HSPs, and the total number occurring. Hence it is essential to look not just at HSPs overall, but also to disaggregate according to Maternal or Paternal descent: i.e., does the pair share a Mother or a Father? (Full-sib pairs share both, of course.)

This can studied by genotyping not just the nuclear DNA required to establish kinship, but also the mitochondrial DNA (mtDNA) which is inherited from the Mother only. Unlike nuclear DNA which is

summarized via genotypes at many loci, mtDNA yields a single "haplotype" for each individual. Thus, if the two individuals in an HSP have different mtDNA haplotypes, then they must have different Mothers, so they must share a Father. If they have the same haplotype, then they probably share a Mother, but it could also be that they come from two different mothers who just happened to have the same haplotype. By checking the frequency of haplotypes in the population at large, the "probably-but-not-definitely-shared-Mother" for same-haplotype pairs can be accommodated explicitly via a mixing-term in the CKMR probabilities, without needing to estimate extra parameters or to make guesses about ancestry; see Bravington et al., 2016b.

To look at Maternal/Paternal descent in HSPs among SBT, we took 325 juveniles (all the definite HSPs and FSPs, i.e. with PLOD > 37, plus 13 control samples and some borderline-PLOD possible HSPs). These were genotyped over an 875-base-pair[13] fragment encompassing the control region or d-Loop of the SBT mitochondrial genome. Sequencing was done using ABI Big DYE version 3 chemitstry (Applied Biosystems) with each amplicon sequenced using two primers anchored in the tRNA regions (Proline and Phenylalanine, forward and reverse primers respectively) which flank the d-Loop. In all, we found 201 variable sites across the fragment, leading to 247 distinct haplotypes, with 172 occurring in just one fish, 72 in two fish, and 3 in three fish. Genotyping was "blind", i.e. the genotypers did not known the sib-pairings among the 325 fish.

There was little ambiguity in assigning unhaplotypes, but small discrepancies evidently do happen occasionally (see below). The very high frequency of unique SBT haplotypes[14] (in this small sample) means that there is little error in assigning Maternal-vs-Paternal descent purely based on same-vs-different haplotype, without needing the complexity of a mixing-term for now; the probability that two SBT mothers will share a haplotype by chance is low enough (~1% or less) to be safely ignored in a preliminary model.

The definition of "mtDNA haplotype" depends on which sites are examined. In our case, where we examined over 200 variable sites within the d-Loop, each haplotype corresponds to a unique combination of DNA bases at those sites. On average, two of the haplotypes chosen at random will differ at around 20 of these sites, but there are some groups of haplotypes which differ at just one or two sites. Even if two individuals really do have the same Mother, it is still possible that their *recorded* haplotypes will be (slightly) different. Such discrepancies can occur for three partly-overlapping reasons:

- heteroplasmy, whereby the original egg-cell has more than one haplotype present (all cells have many copies of the mitogenome); the most common variant is likely to be the one recorded, but which variant is most common can change between Mother and Offspring;
- genotyping error, including the possibility of recording a less-common variant when there is heteroplasmy;
- single-generation mutation.

[13] Approximately 875BP, because some fish had insertions or deletions.

[14] There are far fewer haplotypes in the four or five shark species we have examined, where in one (extreme) case only two haplotypes were present in an entire subpopulation, even though we examined a much higher portion of the mitogenome for sharks than for SBT.

TABLE 5. mtDNA comparisons among definite HSPs and FSPs

#Differing sites	0	1	3	7	11	13	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
#FSPs/HSPs	65	2	1	1	1	1	2	1	1	1	1	3	3	2	5	5	1	7	4	5	4	7	2	7	4	1	2	1

Of 140 definite HSPs and FSPs that were genotyped for mtDNA, 65–67 have identical haplotypes (i.e. ~47% Maternal) and 75–73 have different (Paternal); the ranges reflect possible discrepancies in just two cases, as explained below. The Maternal/Paternal ratio is very close to 50/50, consistent with no strong sex-difference in "reproductive concentration". Restricting comparisons to the 42 same-cohort (actually, same-capture-year; see **??**) HSPs/FSPs yields similar results (52% Maternal). For the 98 cross-cohort HSPs— the ones that are really of interest— the average recapture interval is 3.12 years for Maternal HSPs, compared with 3.06 years for Paternal HSPs— a remarkably similar figure, indicating that the rate of SSB turnover in SBT is (currently) very similar between the sexes.

Overall, there is certainly no reason to be concerned about strong bias arising from using HSPs in a simplified single-sex model (e.g. in the 2017 OM), although a full check must await a detailed sex-specific CKMR analysis..

C.1.1. *Consistency checks.* As a final detail, there are some interesting (and consistent) mtDNA results among the 4 FSPs and 3 HSP-triads mentioned earlier. All the FSPs have the same within-pair mtDNA haplotype, except for one FSP where the two haplotypes differ at just 1 site; on subsequent inspection, this looks like heteroplasmy, whereby one of the two animals seems to carry two versions of the haplotype (, differing at just one site. Among the 3 triads of HSPs, in two cases all 3 fish are HSPs (so they either all share the same Mother, or all share the same Father). And in one of those two triads, all 3 fish do have the same mtDNA (i.e. Mother) whereas in the other, all 3 have different mtDNA (i.e. Father). In the third HSP triad, the PLODs show that A & B are HSPs and A & C are HSPs, but B & C are clearly not; so the shared-parent must be Mother in one case and Father in the other. And indeed, precisely two of A/B/C do have the same haplotype.

All this demonstrates that mtDNA discrepancies are not common (observed for sure at just one site in one FSP), but do occur occasionally. The ranges shown above for Maternal/Paternal proportion reflect the possibility of discrepancies, which are only plausible in one FSP and one sib-pair (Table 5) where the haplotypes differ at just one site. On re-examination, the HSP with a 3-site difference clearly has different haplotypes. The HSP with a one-site difference is unclear (genuine or discrepancy), but its potential impact on Maternal/Paternal proportions can only be small in any case.

C.1.2. *Implications of haplotype diversity and discrepancy rates for detailed CKMR modelling.* Most HSPs are not in triads, so that mtDNA genotyping discrepancies cannot be checked directly; but when discrepancies do occur, then they will lead to a Maternal HSP being misidentified as Paternal (it is unlikely we will ever have enough data to directly estimate and allow for error rates in mtDNA haplotypes). In a comprehensive CKMR analysis, that (small) possibility of misidentification can be eliminated by pooling similar haplotypes, e.g. those with just 1- or 2-site differences. The mixing-term treatment of mtDNA haplotypes in CKMR already allows for the possibility that two haplotypes will be identical by chance (which happens often in species where mtDNA diversity is lower), so the only theoretical

downside from pooling is a minor loss of statistical information; but this is trivially small in an organism with such large haplotypic diversity as SBT. In other words, there will be no great difficulty in setting the HSP treatment up correctly in a stand-alone CKMR model, but even an approximate model that just says "same haplotype means same Mother" would work well.