X. Specification for the Close-Kin Mark-Recapture data used in the MP

Close-Kin Mark Recapture (CKMR) uses modern genetics to identify close relatives (parent-offspring-pairs (POPs) and half-sibling-pairs (HSPs)) amongst large sample sizes of fish, in order to estimate adult abundance and make demographic inferences about the adult stock (Bravington et al. 2016). As part of the CKMR program for SBT, genetic samples have been collected annually since 2006 from adults on the Indonesian spawning grounds and from juveniles (3-year-olds) in the Great Australian Bight (Davies et al. 2018). Each year, updated numbers of POPs and HSPs, along with the numbers of comparisons made in identifying these kin pairs, are provided to the CCSBT data exchange. In the Cape Town Procedure, these data get used in a population dynamics model to provide an index of abundance of reproductive adults (or total reproductive output, TRO), which is then used to modify the TAC (Hillary et al., 2019).

In Indonesia, tissue samples are collected from adult SBT of all sizes at the Benoa Fishing Port each spawning season during processing of catches from the longline fishery. In Australia, tissue samples are collected from juvenile SBT each June-July at the tuna processors during harvest in Port Lincoln; samples are obtained from fish ranging from 98 to 109 cm fork length to ensure 3-year-olds are being sampled. In both sampling locations, sample collection is spread as evenly as practical throughout the harvest season.

DNA is extracted from the tissue samples selected for genotyping. Archived plates of extracted DNA are shipped to Diversity Arrays Technology (DArT) in Canberra for genotype sequencing, referred to as “DArTcap”, and when completed, the sequencing data are provided to CSIRO Hobart. These data are used to call the genotype (i.e., to infer the pair of alleles present) for each fish and locus in the data set using sophisticated algorithms developed at CSIRO specifically for DArTcap sequencing data. The genotyping error rate is also estimated for each locus (of which ~1500 are used in kin-finding), which is important in the identification of HSPs. A series of quality control (QC) steps are applied to the genotyped data to remove fish with unreliable genotype calls and provide a final data set for kin-finding. Note that the QC steps have evolved (and may continue to) over the course of the program, so the exact sample sizes used in kin-finding can change; Table 1 gives the sample sizes used in the 2020 analysis.

POPs are identified across all genotyped adult-juvenile pairs using a modified Mendelian-exclusion statistic referred to as the Weighted-PSeudo-EXclusion (WPSEX) statistic (see Appendix B of Bravington et al. 2017). The numbers of POPs obtained from the 2020 analysis, broken down by juvenile birth year and adult capture year, are given in Table 2 (note this includes POPs that were identified using microsatellites prior to the genotyping method changing in 2015 to DArTcap sequencing; see Bravington et al. 2015, 2017).

HSPs are identified among all genotyped juvenile pairs using a pseudo-log-odds-ratio (PLOD) statistic, which measures the relative probability of a pair of fish having their observed genotypes if they are HSPs compared to if they are unrelated (see Appendix C of Bravington et al. 2017). Unlike the WPSEX statistic for identifying POPs, the PLOD statistic does not give a clear separation between HSPs and unrelated/less-related fish (see Figures 3 and 4 of Farley et al. 2019). Thus, the theoretical means and approximate variances of the PLOD distributions for HSPs and unrelated/less-related pairs are used to determine a lower cut-off PLOD value that minimises the number of false positive HSPs whilst still maintaining a large enough number of HSPs for the estimate to have good precision. An inevitable consequence of ensuring that false positives are rare is that a reasonable number of false negatives will be present; the false-negative rate is

estimated using the expected PLOD distribution for HSPs, and is allowed for in modelling (Bravington et al. 2017). Note that the division between PLOD values for HSPs and more related fish (i.e., full-sibling-pairs) is clear. The numbers of high-confidence HSPs identified from the 2020 analysis, broken down by birth years of siblings, are given in Table 3.

Table 1. Number of fish available for kin-finding analyses in 2020 after quality control checks. For the adults, samples were collected from Indonesia in the fishing season ending in the year shown (i.e., samples collected over the 2005/06 fishing season are referred to as year 2006).

Year	Adults	Juveniles
2006	0	1317
2007	0	1325
2008	0	1356
2009	0	1347
2010	972	1315
2011	958	963
2012	536	876
2013	959	903
2014	922	899
2015	0	953
2016	951	854
2017	971	948
2018	700	777
Total	6969	13,833

Table 2. Number of POPs identified in the 2020 analysis (including those identified using microsatellites; see Bravington et al. 2016) broken down by juvenile birth year (rows) and adult capture year (columns). Note: The exact number of POPs identified, and the total number of comparisons made, may vary between each year's analysis, as the entire updated data set is quality controlled and re-analysed.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2016	2017	2018
2002	0	0	0	0	0	NA	NA	NA	NA	NA	NA	NA
2003	0	5	1	2	0	0	0	1	0	1	0	0
2004	0	2	0	0	3	0	0	0	0	0	0	0
2005	1	4	5	4	1	0	0	1	2	0	0	0
2006	NA	4	3	2	0	0	0	0	0	0	0	0
2007	NA	NA	3	4	1	3	2	0	2	0	1	0
2008	NA	NA	NA	NA	0	1	1	1	0	0	0	2
2009	NA	NA	NA	NA	0	1	1	1	0	0	0	0
2010	NA	NA	NA	NA	0	0	1	4	0	2	0	0
2011	NA	NA	NA	NA	0	0	1	2	1	2	0	0
2012	NA	NA	NA	NA	0	0	0	1	1	0	0	1
2013	NA	NA	NA	NA	0	0	0	0	0	1	1	3
2014	NA	NA	NA	NA	0	0	0	0	0	0	1	0
2015	NA	NA	NA	NA	0	0	0	0	0	1	0	0

Table 3. Number of HSPs identified in the 2020 analysis broken down by birth year of younger sibling (rows) and older sibling (columns). Note: The exact number of HSPs identified, and the total number of comparisons made, may vary between each year's analysis, as the entire updated data set is quality controlled and re-analysed.

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
2003	2	4	2	1	0	0	1	0	0	2	0	2	1
2004		6	3	6	2	2	1	0	0	2	0	0	0
2005			5	3	3	3	0	5	1	1	0	2	0
2006				8	4	1	3	5	3	0	1	1	1
2007					3	3	2	2	2	2	2	1	2
2008						5	1	1	2	3	0	1	0
2009							1	2	1	0	0	0	0
2010								2	1	2	1	0	1
2011									3	2	1	0	3
2012										3	2	1	1
2013											2	4	1
2014												2	2
2015													4

References

Bravington M, Eveson P, Grewe P, Davies C. 2015. SBT Close-Kin Mark-Recapture: options for the medium term. CCSBT-ESC/1509/19.

Bravington MV, Grewe PM, Davies CR. 2016. Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. Nature Communications 7:13162. https://doi.org:10.1038/ncomms13162

Bravington MV, Eveson JP, Grewe PM, Davies CR. 2017. SBT close-kin mark-recapture with parent-offspring and half-sibling pairs: update on genotyping, kin-finding and model development. CCSBT-ESC/1709/12.

Davies CR, Bravington MV, Grewe PD, Eveson JP, Lansdell M, Hill P, Aulich J. 2018. Close-kin project report. CCSBT-ESC/1809/14.

Farley J, Eveson P, Bravington M, Grewe P. 2019. Update on the SBT close-kin tissue sampling, processing and kin finding. CCSBT-ESC/1909/08.

Hillary R, Preece A, Davies C. 2019. Performance of a revised candidate MP using all 3 data sources. CCSBT-ESC/1909/16.