**Supplementary Information for Close-kin mark recapture abundance estimation: practical insights and lessons learned**

Verena M. Trenkel1, Grégory Charrier2, Pascal Lorance1, Mark Bravington3

1Ifremer, Nantes, France

2Univ. Brest, CNRS, IRD, Ifremer, LEMAR, Plouzané, France

3CSIRO, Hobarth, Australia

[S1 Material 2](#_Toc91000119)

[S1.1 Samples 2](#_Toc91000120)

[S1.2 SNPs and genotyping 2](#_Toc91000121)

[S1.3 SNP and individual filtering 3](#_Toc91000122)

[S2 Methods 4](#_Toc91000123)

[S2.1 Estimating length, age and egg laying year 4](#_Toc91000124)

[S2.1.1 Estimating length from disk width or tail measurements: P(length) 4](#_Toc91000125)

[S2.1.2 Estimating age: P(age) 6](#_Toc91000126)

[S2.2 Maturity at length and age 8](#_Toc91000127)

[S2.3 Kinfinding 9](#_Toc91000128)

[S2.3.1 Identifying parent-offspring pairs using WPSEX 9](#_Toc91000129)

[S2.3.2 Distinguishing full-sibling pairs from POPs 10](#_Toc91000130)

[S2.4 TMB code for abundance estimation 11](#_Toc91000131)

[S3 Results 14](#_Toc91000132)

[S3.1 Characteristics of parent-offspring pairs 14](#_Toc91000133)

[S3.3 Identification of related individuals using PLOD statistics 14](#_Toc91000134)

[S3.4 Absence of related individuals between local populations 15](#_Toc91000135)

[S3.5 MAF for local populations 15](#_Toc91000136)

[S3.6 Model parameter estimates 15](#_Toc91000137)

[S3.7 Sample sex ratio investigation 16](#_Toc91000138)

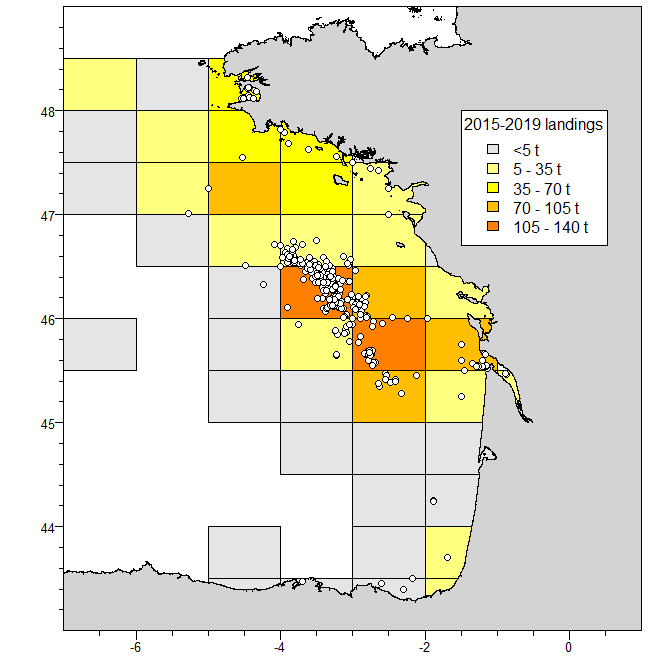
[References 16](#_Toc91000139)

# S1 Material

## S1.1 Samples

Overall 7451 individuals were sampled, of which 216 individuals were genotyped twice, 77 individuals three times and one individual four times for quality control. Among the sampled individuals 7039 came from commercial landings, hence these individual were removed from the population by the fishery (lethal sampling). Based on the agreement between genotypes one individual was later identified as having been sampled twice, once by a survey and later when landed by the fishery.

For individuals sampled from small vessels (<12 m) fishing in coastal waters, sampling position was allocated based on fishers information collected at harbour or according to logbook data, where fishing locations are reported as ICES rectangle (one degree in longitude and half a degree in latitude). The fishing location was not assumed to be the centre of the ICES rectangle, but a more coastal location as these vessels operate mostly in the 12 nm band along the coast. Individuals sampled from landings of larger vessels equipped with VMS (vessel monitoring by satellite), daily fishing locations were estimated as the centre of gravity of daily landings allocated to a 3' x 3' grid, which was derived combining VMS and logbook data (see Ifremer. Système d'Informations Halieutiques (2021) for a description of the method). Figure S1.1 provides sampling positions overlaid onto cumulated French landings by ICES rectangles, during the five years of sampling (2015-2019).



**Fig. S1.1** Map of sampling locations for thornback ray in the Bay of Biscay (white spots) and cumulated French landings in years 2015-2019.

## S1.2 SNPs and genotyping

The SNP data set was obtained in the following way. For SNP development, a RADseq (Restriction site Associated DNA sequencing) protocol was applied to 225 individuals sampled in the Bay of Biscay and elsewhere (Northeast Atlantic and Mediterranean Sea) as described in Marandel *et al.* (2020) using DNA extracted from fin-clips. From this, the most polymorphic 9120 loci (allele frequency > 0.08 for individuals from the Bay of Biscay) were selected (Le Cam *et al.* 2019).

In the second step, all samples were genotyped at these SNP loci using an Infinium® XT iSelect-96 SNP-array. Individual genotypes were scored using the clustering algorithm implemented in the Illumina® GenomeStudio Genotyping Analysis Module v2.0.3. The chip was created with DesignStudio Microarray Assay Designer. The high-throughput solution Infinium XT requires integrated systems that streamline sample preparation and analysis, so the Illumina Automation Control software for the Tecan liquid handling robot was used. Genotyping reports were analysed with GenomeStudio, which normalizes the intensities of signals for each locus and assigns a cluster position to each sample. The GenCall score was then calculated for each genotype. A no-call threshold of 0.15 was used to not call individuals too far away from the cluster centre. During chip development, 987 SNPs had been duplicated as there was a second SNP in the 50 nucleotide bases flanking sequence (Le Cam *et al.* 2019). These duplicated SNPs were used for filtering individuals and SNP (section S1.3).

## S1.3 SNP and individual filtering

Several filtering steps were applied to SNPs and individuals to obtain a reliable data set for further analyses with only neutral unlinked SNPs. First, contaminated individuals identified by their unusual number of common SNPs and their neighbourhood on the genotyping plate were removed (16 individuals). Second, monoallelic SNPs and those SNP showing large differences for replicated individuals or duplicated SNPs were removed (2258 SNPs). Third, the genotypes of duplicated SNPs described above were combined (944 SNPs). For the rare cases of disagreement, an individual was assumed heterozygote if at least one of the duplicated SNP pair was heterozygous or both were homozygous but for different alleles (0.2% of SNP calls). Fourth, SNPs with minor allele frequency <0.1 and SNPs far from Hardy-Weinberg equilibrium (ratio observed/expected proportion <0.5 or >1.5) were removed (1363 SNPs). This removes all SNPs identified in Trenkel *et al.* (2020) as being on the X chromosome. Fifth, to remove correlated SNPs, only one SNP was retained in pairs with linkage disequilibrium based pairwise correlation >0.1 (127 SNPs removed). Sixth, low scoring SNPs were removed, retaining those with a call frequency of at least 98%. Seventh, only individuals with a call rate of 98% of SNPs were kept. The seven filtering steps lead to a data set with 3668 SNPs for 6555 individuals (table S1.1).

**Table S1.1** Number of retained thornback ray individuals by sampling year and subarea in the Bay of Biscay. For abundance estimation data before 2015 were removed.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Subarea** | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** | **2020** | **Total** |
| **Gironde** | 0 | 2 | 7 | 0 | 386 | 212 | 328 | 14 | 125 | 213 | 1287 |
| **Offshore** | 0 | 2 | 0 | 7 | 232 | 3191 | 1533 | 25 | 3 | 0 | 4993 |
| **Other** | 10 | 0 | 0 | 45 | 84 | 122 | 6 | 5 | 3 | 0 | 275 |
| **Total** | 10 | 4 | 7 | 52 | 702 | 3525 | 1867 | 44 | 131 | 213 | 6555 |

To explore the potential distribution of the final set of 3668 filtered SNPs throughout the genomea BLAST search using the executable BLAST+ 2.6.0 package (Altschul *et al.* 1990) optimized for short sequences was carried for the sequences including each SNP. As the genome of thornback ray is currently not available, the whole genome assembly of another ray species was used (starry ray *Amblyraja radiata*, male adult, testis and liver tissues, GenBank accession number GCA\_010909765.1. Starry ray and thornback ray have both 49 chromosomes (Stingo & Rocco 2001). All but 32 SNPs were significantly matched to starry ray chromosomes, the later were only matched scaffolds that have not be positioned on chromosomes (Fig S1.2).



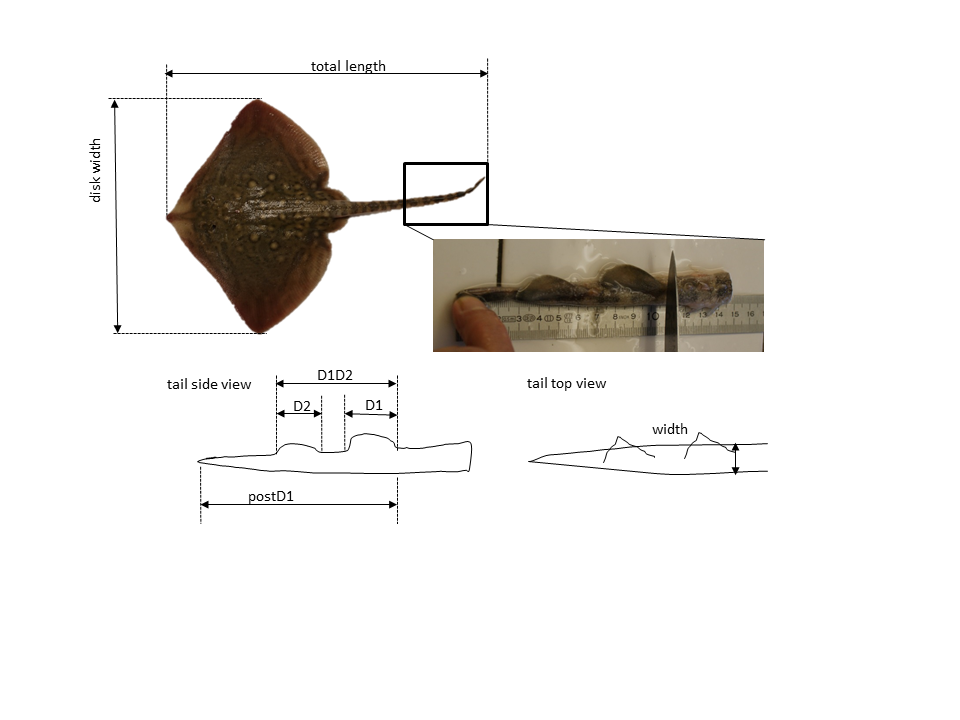
**Fig. S1.2** Position of final set of thornback ray SNPs on chromosomes of starry ray.

# S2 Methods

## S2.1 Estimating length, age and egg laying year

### S2.1.1 Estimating length from disk width or tail measurements: P(length)

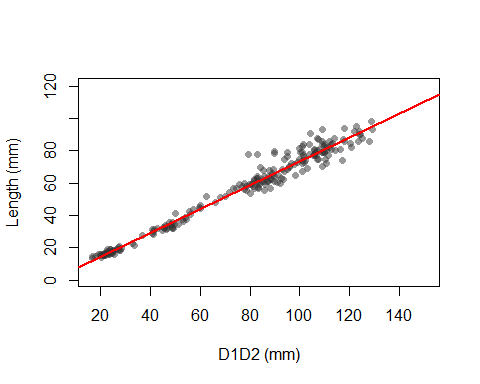
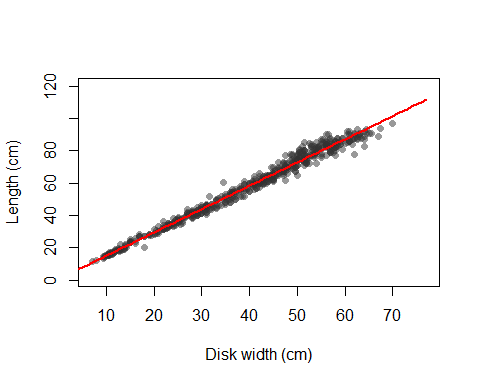
For 5449 individuals total length was estimated from allometric relationships between disk width or tail measurements and total length. The measurements are represented in figure S2.1. The fitted linear or polynomial regressions are shown in figure S2.2 and model parameters are summarised in table S2.1.

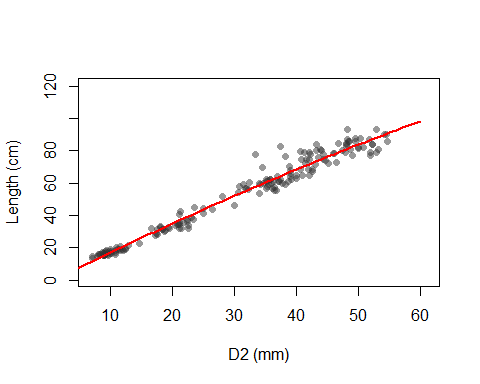
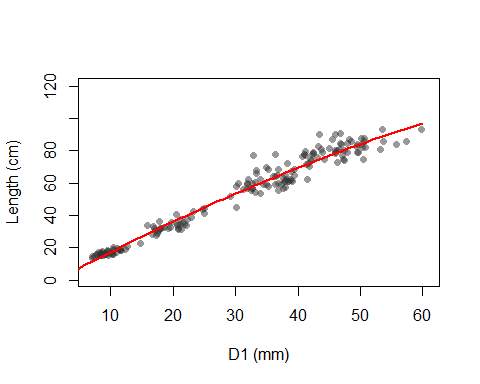


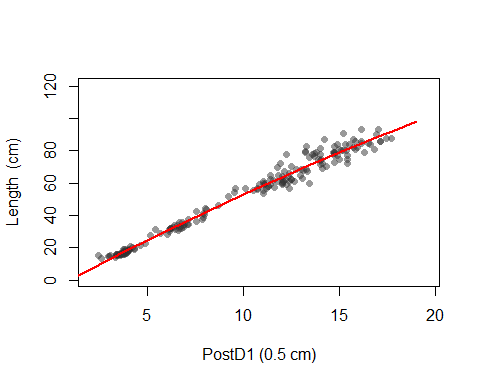
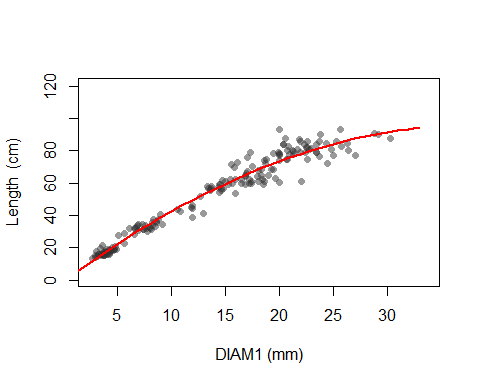
**Fig. S2.1** Schematic representation of body and tail measurements used for estimating total length.

**Table S2.1** Allometric relationships between body and tail measurements and total length (cm) used to predict missing total lengths. N sample size, L total length, σresidual standard error.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Name** | **Description** | **N** | **Model** | **σ** | **Adj. r2** |
| **Disk width** | disk width | 518 |  | 2.84 | 0.9839 |
| **D1D2** | distance between origin of first dorsal fin and insertion of second dorsal fin | 216 |  | 4.283 | 0.9708 |
| **D1** | length of first dorsal fin | 179 |  | 4.766 | 0.9645 |
| **D2** | length of second dorsal fin | 175 |  | 4.457 | 0.9689 |
| **DIAM** | diameter of tail at the first dorsal fin origin | 181 |  | 4.963 | 0.9615 |
| **PostD1** | tail length from first dorsal fin origin to caudal fin tip | 185 |  | 3.792 | 0.97743 |





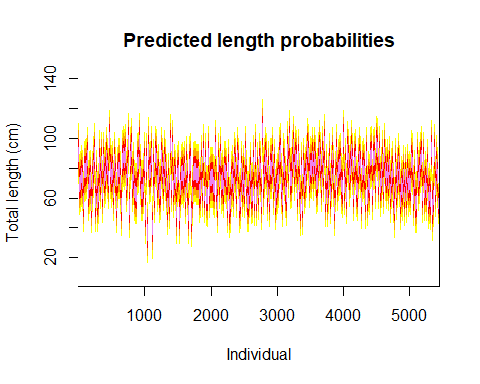


**Fig. S2.2** Allometric relationships between disk width or tail measurements and total length. For definitions and fitted relationships see table S2.1.

Missing total lengths were predicted using the fitted relationships in table S2.1 with the equation depending on which data were available, starting at the top of the table, i.e. disk width was used if available, otherwise D1D2 if available, etc. To account for uncertainty in the relationship, discretised length probability distributions were estimated for each individual *i* from its disk width or tail measurements

(1)

where *l* are 1-cm length categories, the point estimate for individual *i* obtained using the equation in table S1.1 and the corresponding residual variance. The resulting predicted length probability distributions are shown in Fig. S2.3. For individuals with measured total lengths (to lower cm), the probability is set to one for this length class and to 0 otherwise.



**Fig. S2.3.** Predicted length probability distributions for individuals with missing total length.

### S2.1.2 Estimating age: P(age)

For individual *i,* its age probability distribution at the time of sampling was derived in several steps following Hillary *et al.* (2018). It is defined as

(2)

where *P*(*a*|*l*) is the probability distribution of age-at-length in the sample and is length distribution obtained in section S2.1.1. Applying Bayes’s rule it is decomposed

(3)

with *P*(*l*|*a*) the probability distribution of length given age in the population, *P*(*a*) is the probability distribution of age *a* in the sample and P(l) is the probability distribution of length *l*, also in the sample.

*P*(*l*|*a*) was derived combining mean length-at-age estimates from a von Bertalanffy growth model with a normal variability distribution around this mean, assuming constant variance across ages

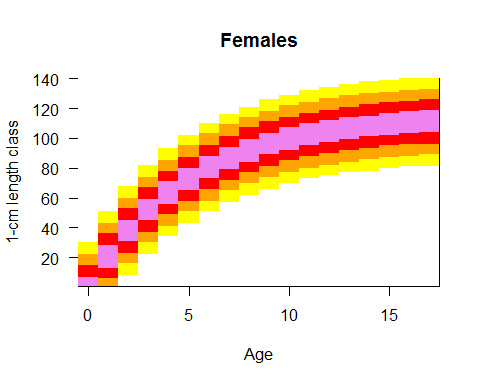
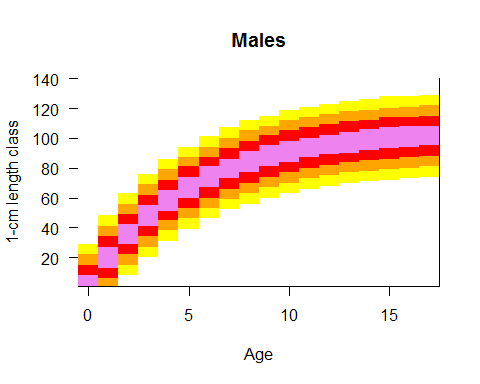
(4)

(5)

As thornback ray displays growth dimorphism, sex specific growth parameters were used (table S2.2). As the growth curve refers to the time of egg hatching, all ages were shifted by -0.5 years to make age count from egg laying which occurs around 6 months earlier. Reference of age to the time of egg laying instead of hatching is chosen as CKMR abundance estimation is for the time of sexual mating which occurs prior to egg laying by the mother. The resulting length-at-age probability distributions *P*(*l|a*) are shown in Fig. S2.4.

**Table S2.2** Parameter values for von Bertalanffy growth function (eq. 4) and length variance (eq. 5). *L*0 length-at-birth.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sex** | **L∞**  **(cm)** | **k**  **(year-1)** | **L0**  **(cm)** | **(cm2)** |
| **female** | 115 | 0.2 | 11 | 40 |
| **male** | 105 | 0.2 | 11 | 34 |

**Fig. S2.4**. Length-at-age probability distributions *P*(*l|a*) by sex for thornback ray. Age counts from egg laying.

Next, the prior age distribution P(*a*) in the sample is estimated from the equation

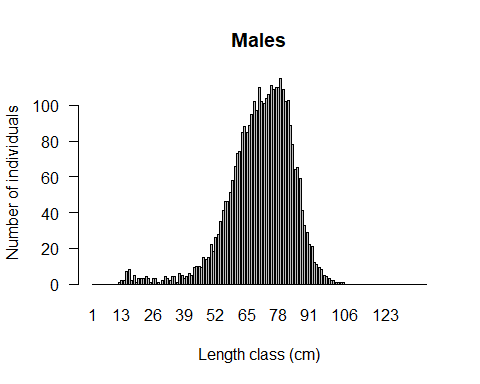
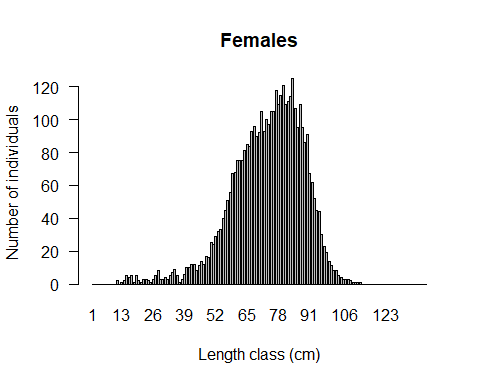
(6)

assuming this age distribution is constant in space and time. In reality cohort variations and spatial sampling effects will make it vary somewhat. *P*(*a*) was estimated by maximum likelihood from (section S2.1.2) and *P*i(*l*) (section S2.1.1). The multinomial likelihood function is defined as

(7)

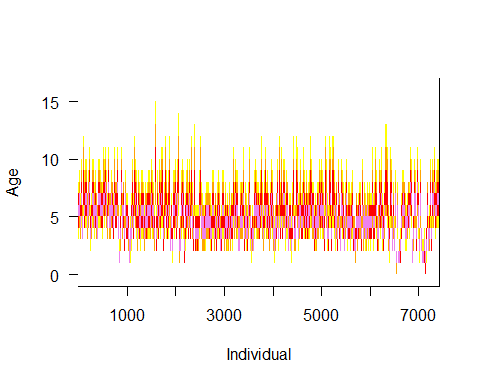
with *N* the total number of samples individuals and *N*l the number of sampled individuals per length class, obtained by summing predicted and observed individual length probabilities *P*i(*l*) (Fig. S1.5)

(8)



**Fig. S2.5.** Number of sampled individuals by length class (observed or predicted).

Combining P(a) with estimates using eq. 3 the probability distribution of age-at-length *P*(a|l) is obtained (Fig. S2.6). Plugging this estimate into eq. 2 provides the required age probability distributions *P*i(*a*) of the sampled individuals at the time of sampling (Fig. S1.6).



**Fig. S2.6.** Predicted age probability distributions at the time of sampling for all sampled individuals. Age counts from egg laying.

## S2.2 Maturity at length and age

Maturity as a function of length (in cm) was calculated as

using parameter values in table S2.3. Combining this with the distribution of age-at-length (eq. 3), the proportion mature by age was calculated (Fig. S2.7).

**Table S2.3** Parameters of length-based maturity ogive from Whittamore and McCarthy (2005).

|  |  |  |
| --- | --- | --- |
| Sex | r | L50 (cm) |
| Females | 0.4338 | 70.5 |
| Males | 0.73 | 58.8 |



**Fig. S2.7.** Maturity at length (left) and at age (right) for thornback ray.

## S2.3 Kinfinding

### S2.3.1 Identifying parent-offspring pairs using WPSEX

In the complete absence of null alleles and of genotyping errors, the POPhood of a pair of samples could be assessed simply by looking across loci for Mendelian exclusions, where one individual scores as AA and the other as BB. Since all loci in a POP should share at least one allele through inheritance, finding even one AA/BB genopair would eliminate POPhood— whereas the chance of a false-positive POP (where no AA/BB genopairs are found) becomes negligible with many loci. However, this would be far too stringent in the presence of nulls and errors, leading to false-negative rejection of many true POPs. The WPSEX (Weighted PSeudo-EXclusion) statistic is designed to robustly identify parent-offspring pairs from biallelic SNP (A/B) data where (i) many loci may have (heritable) null alleles, and (ii) rare genotyping errors may occasionally lead to (non-heritable) undetected alleles. The idea is that per locus null-allele frequency can be estimated reliably in advance by maximum-likelihood, based on sample genotype frequencies in the categories AB/AAO/BBO/OO where AAO means either AA or AO (i.e. where single-null cannot reliably be distinguished from homozygote) assuming Hardy-Weinberg equilibrium. However, no estimates of genotyping error rate are required (though it is assumed to affect only a small proportion of called genotypes). The complications around reliably estimating very-low-but-non-zero locus-specific genotyping error rates have driven us away from likelihood-based approaches to POP-finding.

The overall WPSEX is a weighted sum across loci of the number of pseudo-exclusions where one genotype is AAO and the other BBO, which could either be a true exclusion in a non-POP, a coinherited null, or a genotyping error where AB is mis-called)

(9)

where the indicator function is 1 if its condition is met and 0 otherwise, and is a locus-specific weight discussed below. On average, POPs will have fewer pseudo-exclusions (because no true exclusions) and thus lower WPSEX’s than will Unrelated Pairs (UPs). Given some set of , the expected values and variances can be calculated for POPs and UPs based on estimated allele frequencies including that of nulls, but excluding genotyping errors. The weights are then chosen to minimize false-positive probability, i.e. that an UP might have a WPSEX as low as the average WPSEX for a true POP: specifically, to maximize . Loci with higher null frequency will receive lower weights, because they are more likely to give a pseudo-exclusion due to coinheritance of a null. Since any locus in a UP can generate a pseudo-exclusion simply by chance, the weights are not heavily skewed towards particular loci.

WPSEX was designed for ddRAD datasets with high read-depths (thus very low genotyping error rate) but with heritable nulls too common to permit discarding all null-encumbered loci, for example, Southern Bluefin Tuna (Bravington, 2017). In fact, given the different genotyping method and different genome in this paper, estimated null-allele frequencies are very low and so the null-robust property of WPSEX is not crucial; nevertheless, it is simple to calculate and seems (from various CKMR projects) to work well whether or not nulls are common. Note that a full log-likelihood approach for POPs (such as we have used for HSPs), while theoretically even better than WPSEX if good estimates of all parameters are available, would require explicit estimation of genotyping error rates.

### S2.3.2 Distinguishing full-sibling pairs from POPs

The PLOD statistic (designed for distinguishing between HSPs and UPs) is also very effective in separating first-order kin (FSPs and POPs) from second-order (HSPs etc.), but does not help much in splitting FSPs from POPs. The WPSEX statistic does better at that, but since on average 3/4 of the loci in an FSP will share at least one allele, it too is of limited effectiveness. Discrimination between FSPs and POPs can be improved by also considering the proportion of identical genotype pairs, since on average 1/4 of loci will coinherit both alleles in an FSP, but only one allele per locus is coinherited in a POP (though an allele may of course also be shared by chance rather than coinheritance). Typically, FSPs will tend to have more identical genotype pairs than POPs. We thus also compute

(10)

where the weights are again chosen to maximize the power of discrimination, this time defined as . Again, the optimal weights can be calculated from minor-allele and null-allele frequency estimates. The combination of WSAME and WPSEX seems to give a reasonably effective and certainly robust (to nulls and errors) toolkit for discriminating first-order kin, without the error-rate complications entailed by fully-likelihood-based methods. Nevertheless, POP/FSP discrimination on genetic grounds alone is still challenging with the number of loci required for affordable CKMR studies, so it is also important to make use of individual demographic information— age in particular— if available.

## S2.4 TMB code for abundance estimation

#include <TMB.hpp>

#include <iostream>

/\* Parameter transform \*/

template<class Type>

Type ilgt(Type x){return( exp(x)/(Type(1.0)+exp(x)));} // return( inv\_logit( x))

template<class Type>

Type objective\_function<Type>::operator() ()

{

/\*data section \*/

DATA\_INTEGER(first\_y); // first year (=earliest birth year) (calendar year - constant);

DATA\_INTEGER(last\_y); // last year (=latest birth year) (calendar year - constant);

DATA\_INTEGER(first\_yS); // first year of sampling years as relative index (calendar year - constant)

DATA\_INTEGER(last\_yS); // last year of (tissue ) sampling years

DATA\_INTEGER(n\_l); // number of length categories (columns) in cm for matrices F\_age\_length & M\_age\_length

DATA\_IARRAY(MPOPobs); // table for MOP, same categories as MSamp & JSamp: Length Par, Length Off, Sampling year Par, Sampling year Off, number

DATA\_IARRAY(PPOPobs); // table for FOP, same categories as PSamp & JSamp: Length Par, Length Off, Sampling year Par, Sampling year Off, number

DATA\_INTEGER(Nm); // number of rows in MSamp

DATA\_INTEGER(Np); // number of rows in PSamp

DATA\_INTEGER(Nj); // number of rows in JSamp

DATA\_IARRAY(MSamp); // table for number of sampled individuals that are potential mothers: Length id, Year id, N

DATA\_IARRAY(PSamp); // table for number of sampled individuals that are potential fathers : Length id, Year id, N

DATA\_IARRAY(JSamp); // table for number of potential offspring : Length id, Year id, N

DATA\_ARRAY(is\_mature\_F); // probability matrix for female i being mature (1 mature, 0 immature) at birth of offspring j dimension (birth years i x birth years j)

DATA\_ARRAY(is\_mature\_M); // probability matrix for male i being mature at birth of offspring j dimension (birth year i x birth year j)

DATA\_IARRAY(is\_alive); // id matrix for individual i being alive at birth of j dimension (first\_yS:last\_yS x first\_y:last\_y)

DATA\_IARRAY(is\_born); // id matrix for individual with birth year i being born at birth j dimension (first\_y:last\_y) x (first\_y:last\_y)

DATA\_INTEGER(Agemax); // number of age classes in PAgeLength.F/M matrices

DATA\_ARRAY(F\_age\_length); // matrix (age classes x length classes) with conditional probability of age given length for females

DATA\_ARRAY(M\_age\_length); // matrix (age classes x length classes) with conditional probability of age given length for males

/\*parameter section\*/

PARAMETER(n0\_parvec); // log(N0) number in year0 which is the same for each sex [female, male]

PARAMETER(logSdGamma); // log of standard deviation of random effect growth parameter U

PARAMETER(gammainit); // starting value for random walk of gamma

PARAMETER\_VECTOR(U); // Latent random variable for random walk of growth rate

int n\_y=last\_y+1; // number of years for model, ie birth years

int n\_yS=last\_yS+1; // number of sampling years, including missing years

int female=0; // index for mature females

int male=1; // index for mature males

/\*Transform parameters to ensure correct range \*/

//N0 abundance by sex

vector<Type> n\_s0(2);

n\_s0(0)=exp(n0\_parvec);

n\_s0(1)=exp(n0\_parvec);

// growth parameter random walk

Type SdGamma=exp(logSdGamma);

vector<Type> gamma(n\_y);

gamma(0)=gammainit;

for (int y=first\_y;y<=last\_y;y++){

gamma(y)=gamma(y-1)+U(y-1);

}

/\*Population dynamics models\*/

array<Type> n\_sy(2,n\_y); //numbers by sex

vector<Type> N\_y(n\_y); // total number of reproducing adults

for(int s = 0; s <2; s++){ //s =0 female, s=1 male

n\_sy(s,0)=n\_s0(s); // year 0

N\_y(0)+=n\_sy(s,0);

for(int y = first\_y; y <= last\_y; y++){ //subsequent years starting with first\_y

n\_sy( s, y) = abs(n\_sy( s, y-1) \* exp(gamma(y)));

N\_y(y)+=n\_sy(s,y);

}

}

/\*Calculate total reproductive output RO \*/

//calculate inverse of numbers in year y for sex s

array<Type> inv\_n\_sy(2,n\_y);

for(int s = 0; s < 2; s++) { //

for (int y = first\_y; y <= last\_y; y++) { // allows for sampling in any year - historical or future

if(n\_sy(s,y)>1) inv\_n\_sy(s,y) = Type(1.0) / n\_sy(s,y); //assumes all indiviuals have same reproductive output

else inv\_n\_sy(s,y)=0;

} // y

} // s

/\*Calculate kin probabilities \*/

// create matrices for probabilities of POP

array<Type> Pr\_MPOP(n\_y+1,n\_yS+1,n\_y+1); //mother-offspring pair

array<Type> Pr\_PPOP(n\_y+1,n\_yS+1,n\_y+1); //father-offspring pair

for(int bi= first\_y; bi<= last\_y;bi++){ // i's birth year; parent

for(int ti= first\_yS; ti<=last\_yS;ti++){// i's capture year; parent

for(int bj= first\_y;bj<=last\_y;bj++){ // j's birth year; offspring

// Was i mature & was i still alive?

Pr\_MPOP(bi, ti, bj)=is\_born(bi,bj)\*is\_alive(ti,bj)\*is\_mature\_F(bi,bj)\*inv\_n\_sy(female,bj);// mother-offspring pair

// fathers assume random mating

Pr\_PPOP(bi, ti, bj)= is\_born(bi,bj)\*is\_alive(ti,bj)\*is\_mature\_M(bi,bj)\*inv\_n\_sy(male,bj);// father-offspring pair

}

}

}

/\*Transform kin probabilities at age into probalities at length\*/

//use female age-length probabilities for juveniles

array<Type> Pr\_MPOP\_L(n\_l,n\_l,n\_yS+1,n\_yS+1); //mother-offspring pair

array<Type> Pr\_PPOP\_L(n\_l,n\_l,n\_yS+1,n\_yS+1); //father-offspring pair

for (int li=0; li<n\_l;li++ ){ //F\_age\_length matrix : there are n\_l classes

for(int lj=0; lj<n\_l;lj++ ){ //n\_l

for(int ti= first\_yS; ti<=last\_yS;ti++){// i's capture year; parent

for(int tj=first\_yS; tj<=last\_yS;tj++){ //j's capture year; offspring

for(int bi= first\_y; bi<=last\_y;bi++){ // i's birth year; parent

for(int bj= first\_y; bj<=last\_y;bj++){ // j's birth year; offspring

// sum over age=capture year - birth year

int agei=ti-bi;

int agej=tj-bj;

if(agei<=Agemax && agej<=Agemax && agei>=0 && agej>=0){

Pr\_MPOP\_L(li,lj,ti,tj)+=Pr\_MPOP(bi, ti, bj)\*F\_age\_length(agei,li)\*F\_age\_length(agej,lj);// mother-offspring pair

Pr\_PPOP\_L(li,lj,ti,tj)+=Pr\_PPOP(bi, ti, bj)\*M\_age\_length(agei,li)\*F\_age\_length(agej,lj);// father-offspring pair

}}

}

}

}

}

}

/\* Likelihood function for POP \*/

Type tot\_lglk = 0.0;

//for MOP mother-offspring pairs

for(int i= 0; i<Nm;i++){ // number of length-year classes for potential mothers

for(int j= 0; j<Nj;j++){// number of length-year classes for potential offspring

tot\_lglk+=MPOPobs(MSamp(i,0),JSamp(j,0),MSamp(i,1),JSamp(j,1))\*log(MSamp(i,2)\*JSamp(j,2)\*Pr\_MPOP\_L(MSamp(i,0),JSamp(j,0),MSamp(i,1),JSamp(j,1)))-MSamp(i,2)\*JSamp(j,2)\*Pr\_MPOP\_L(MSamp(i,0),JSamp(j,0),MSamp(i,1),JSamp(j,1));

}

}

//for FOP father-offspring pairs

for(int i= 0; i<Np;i++){ // number of length-year classes for potential fathers

for(int j= 0; j<Nj;j++){// number of length-year classes for potential offspring

tot\_lglk+=PPOPobs(PSamp(i,0),JSamp(i,0),PSamp(i,1),JSamp(j,1))\*log(PSamp(i,2)\*JSamp(j,2)\*Pr\_PPOP\_L(PSamp(i,0),JSamp(j,0),PSamp(i,1),JSamp(j,1)))-PSamp(i,2)\*JSamp(j,2)\*Pr\_PPOP\_L(PSamp(i,0),JSamp(j,0),PSamp(i,1),JSamp(j,1));

}

}

//random effect for variable growth, mean 0, sd SdGamma

for(int y = 0; y < last\_y; y++){ //all years

tot\_lglk += -logSdGamma - 0.5\*pow(U(y)/SdGamma,2); // normal random effect for difference between gamma

}

tot\_lglk\*=Type(-1.0); //return negative lglk

ADREPORT(n\_sy); //pop numbers by sex and year

ADREPORT(SdGamma);

ADREPORT(gamma);

ADREPORT(U);

ADREPORT(N\_y);

ADREPORT(tot\_lglk); //negative log likelihood

return tot\_lglk;

}

# S3 Results

## S3.1 Characteristics of parent-offspring pairs

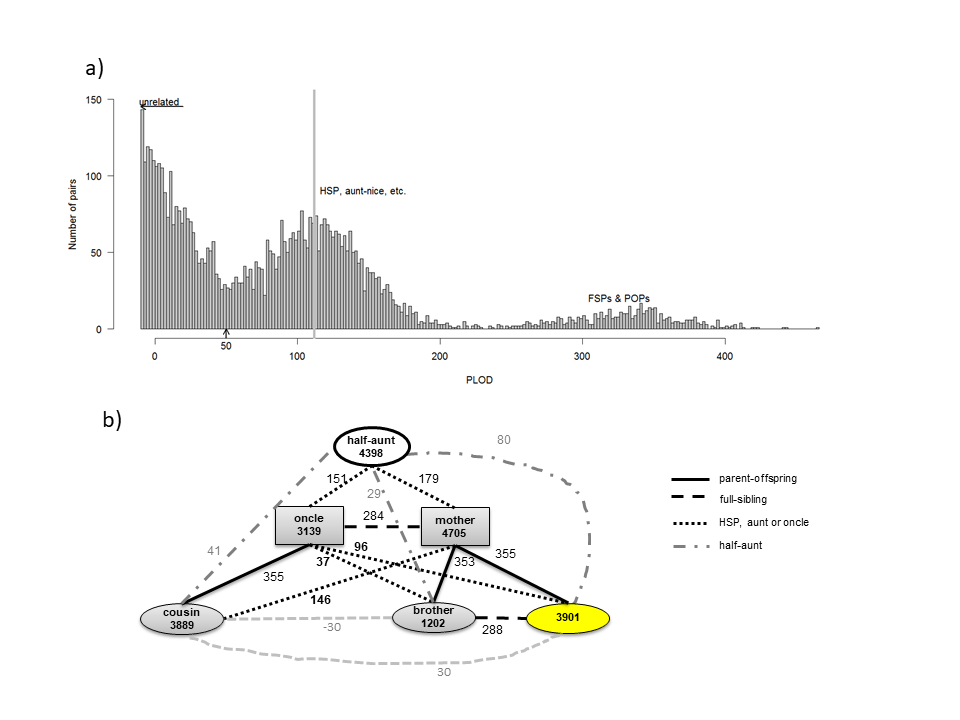
**Table S3.1**. Sex of parents and offspring in identified POPs of thornback ray, all individuals.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Female offspring | Male offspring | Sum | % |
| Mother | 35 | 23 | 58 | 59 |
| Father | 20 | 21 | 41 | 41 |
| Sum | 55 | 44 | **99** |  |
| % | 56 | 44 |  |  |

**Table S3.2.** Sample size, number of observed parent-offspring pairs, and number of comparisons corresponding to potential number of parent-offspring pairs in samples from the Gironde estuary and offshore populations of thornback ray in the Bay of Biscay used for abundance estimation. MPOP maternal parent-offspring pairs, PPOP paternal parent-offspring pairs. juveniles <75 cm ; adults ≥75 cm at capture and mature in year of birth of juvenile (according to maturity ogive in Fig S2.7).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Gironde adults n♀= 427, n♂=317** | | **Offshore area adults n♀= 1250, n♂=942** | |
| **MPOP/# comp.** | **PPOP/# comp.** | **MPOP/# comp.** | **PPOP/# comp.** |
| **Gironde juveniles**  **n=534** | **12** / 160 000 | **4** / 121 000 | **0** / 442 000 | **0** / 362 000 |
| **0ffshore area**  **juveniles**  **n=2799** | **0** / 837 000 | **0** / 600 000 | **31** /2 744 000 | **26** /2 318 000 |

## S3.3 Identification of related individuals using PLOD statistics



**Fig. S3.1.** Example family. Kinship names are for focal individual in yellow. Values on edges are PLOD values between connected individuals. Numbers in boxes and ovals are sample ID numbers.

## S3.4 Absence of related individuals between local populations

**Table S3.3** Number of related pairs/number of comparisons for Gironde estuary and offshore populations of thornback ray in the Bay of Biscay. First and 2nd-order (all) related individuals have PLOD statistics >60. All individuals sampled.

|  |  |  |  |
| --- | --- | --- | --- |
| Kinship | Gironde-Gironde | Gironde-offshore | Offshore-offshore |
| Parent-offspring | **25**/911 223 | **0**/5 594 774 | **73**/14 576 731 |
| All | **245**/827 541 | **0**/6 425 991 | **3146**/12 462 528 |

## S3.5 MAF for local populations



**Fig. S3.2** Comparison of minor allele frequencies of SNPs between local populations of thornback ray in the Bay of Biscay.

## S3.6 Model parameter estimates

**Table S3.4** Parameter estimates (standard deviations in brackets) of population dynamics models with and fitted to local populations of thornback ray in the Bay of Biscay; index *s* stands for sex and *t* for year. Estimated was close to zero in all cases.

|  |  |  |
| --- | --- | --- |
| Population | *N*0 |  |
| Gironde | 7.58 (3.02) | 0.12 (0.21) |
| offshore | 10.09 (3.87) | 0.06 (0.25) |

## S3.7 S***ample sex ratio investigation***

To investigate whether the unequal sex ratio of potential parents in the sample had any effect, we carried out abundance estimates for multiple data sets using the combined data set (Gironde and offshore). We compared results for the observed sex ratio of 1.33 (females:males) with those obtained with equal sex ratios. For this we created ten data sets with balanced sex ratios for the potential parents (individuals >75 cm) by subsampling (without replacement) females and ten data sets with 1.33:1 sex ratios. All 20 data sets had 2518 potential parents to remove the sample size effect. The resulting abundance estimates differed more between random data sets than between the two sex ratios (Fig. S3.3).



**Fig. S3.3** Abundance estimates for data sets with balanced (1:1) and unbalanced (1.33 females : 1 male) sex ratios for the potential parents.

# References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology,* **215,** 403-410.

Bravington, M. 2017. SBT kin-finding and genotyping update. ICES Document Tech. rep. CCSBT-OMMP 8-12. Commission for the Conservation of Southern Bluefin Tuna.

Hillary, R.M., Bravington, M.V., Patterson, T.A., Grewe, P., Bradford, R., Feutry, P., Gunasekera, R., Peddemors, V., Werry, J., Francis, M.P., Duffy, C.A.J. & Bruce, B.D. (2018) Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Scientific Reports,* **8**.Bravington, M. 2017. SBT kin-finding and genotyping update. ICES Document Tech. rep. CCSBT-OMMP 8-12. Commission for the Conservation of Southern Bluefin Tuna. .

Ifremer. Système d'Informations Halieutiques 2021. Algorithme de traitement de données de géolocalisation ALGOPESCA. Note synthétique. <https://archimer.ifremer.fr/doc/00682/79405/>.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., and Prodohl, P. A. 2013. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods in Ecology and Evolution, 4: 782-788.

Whittamore, J. M., and McCarthy, I. D. 2005. The population biology of the thornback ray, Raja clavata in Caernarfon Bay, north Wales. Journal of the Marine Biological Association of the United Kingdom, 85: 1089-1094.

[Le Cam, S., Bidault, A., Charrier, G., Cornette, F., Lamy, J.-B., Lapegue, S., Lorance, P., Marandel, F. & Trenkel, V. (2019) SNPs for thornback ray](https://archimer.ifremer.fr/doc/00682/79405/) *Raja clavata*, validated from genotyped individuals. *SEANOE.* [*https://doi.org/10.17882/70546*](https://doi.org/10.17882/70546).

Marandel, F., Charrier, G., Lamy, J.-B., Le Cam, S., Lorance, P. & Trenkel, V.M. (2020) Estimating effective population size using RADseq: effects of SNP selection and sample size. *Ecology and Evolution,* **10,** 1929-1937.

Serra-Pereira, B., Figueiredo, I. & Gordo, L.S. (2011) Maturation, fecundity, and spawning strategy of the thornback ray, *Raja clavata*: do reproductive characteristics vary regionally? *Marine Biology,* **158,** 2187-2197.

Stingo, V. & Rocco, L. (2001) Selachian cytogenetics: a review. *Genetica,* **111,** 329-347.

Trenkel, V.M., Boudry, P., Verrez-Bagnis, V. & Lorance, P. (2020) Methods for identifying and interpreting sex-linked SNP markers and carrying out sex assignment: application to thornback ray (*Raja clavata*). *Molecular Ecology Resources,* **20,** 1610-1619.

Whittamore, J. M., and McCarthy, I. D. 2005. The population biology of the thornback ray, Raja clavata in Caernarfon Bay, north Wales. Journal of the Marine Biological Association of the United Kingdom, 85: 1089-1094.