THE ATLANTIC-WIDE RESEARCH PROGRAMME FOR BFT (GBYP Phase 12)

SHORT-TERM CONTRACT (ICCAT GBYP 02/2023)

PILOT STUDY ON EPIGENETIC AGEING TECHNIQUE FOR AGE ESTIMATION OF ATLANTIC BLUEFIN TUNA – GBYP 2/2023.

Final Report

February 2024

THE COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION (CSIRO)

Campbell Davies, Ben Mayne, Peter Grewe, Luke-Lloyd Jones, Nicola Potter, Chloe Anderson and Jessica Farley

INSTITUTO ESPAÑOL DE OCENOGRAFIA

Enrique Rodriguez-Marin

This project is co-funded by the European Union

PILOT STUDY ON EPIGENETIC AGEING TECHNIQUE FOR AGE ESTIMATION OF ATLANTIC BLUEFIN TUNA

The ICCAT GBYP program is investigating the feasibility of applying Close-kin Mark Recapture (CKMR) to the eastern stock of Atlantic bluefin tuna. A key consideration in this investigation is the ability to obtain accurate age information for use in the construction of the kin probabilities and population modelling. This paper reports on progress of a pilot study to assess the suitability of epigenetic ageing - estimating the age of individuals from the level of DNA methylation from analysis of tissue samples - for this purpose. The aims of the pilot study were to: i) calibrate an epigenetic age model using the approach developed by Mayne et al 2020 and tissue samples from eastern and western stocks with age data from standardized ICCAT otolith reading methods, ii) examine the influence of stock and sex on epigenetic age calibration, and iii) assess the relative cost-effectiveness with the use of otolith-based age estimation for use in CKMR for eastern BFT (which would also be applicable to western BFT). This report summarises the initial results from this study. Three hundred and sixty-one samples were successfully processed for methylation scores from a total of 768 samples sourced from the western Atlantic (n=173), Mediterranean Sea (n=177) and Eastern Atlantic (n=11) collections of matching otolith and tissues and decimal age estimated based on ICCAT protocols. Predictive performance was high for the range of epigenetic age calibration models fitted to the age and methylation data with $R^2 > 0.96$ for the three highest performing models. The multiple regression model was marginally the best performing of these three, with $R^2 > 0.973$ and mean-squared error of 0.49 years. The largest deviations were 4.7, 3.8, and 6.8 years for the best performing models and, in each case, were for the highest age in the sample (~27 years). For ages less than, or equal to 20, the maximum deviations were 1.98, 2.11 and 1.78. The high quality of the fits and predictive performance for the best performing models suggests that the effects of sex and geographical area on the calibrations are unlikely to be substantial. These results clearly demonstrate that estimating age using calibrated epigenetic age models and high-quality tissue samples is feasible for Atlantic bluefin and, based on these preliminary results, a single calibration models could be used for samples sourced from both eastern and western stocks. Future work should focus on: i) more detailed investigation of the two identified types of contamination (organic and DNA cross-contamination) that reduced the sample size available for the model calibration; ii) improving the representation of age classes at the tails of the distribution to improve predictive performance for these ages, and: iii) assessing options for reducing cost of data collection throughout the whole workflow from tissue collection to sequencing.

KEYWORDS

Epigenetic ageing, calibration, Close-kin mark–recapture (CKMR), Atlantic bluefin tuna, genetic methodologies, stock structure, genetic sex determination

1. Introduction

The ICCAT GBYP program is investigating the feasibility of applying Close-kin Mark Recapture (CKMR) to the eastern stock of Atlantic bluefin tuna (BFT) (Anon., 2023a). A key consideration in this investigation is the ability to obtain accurate age information for use in the construction of the kin probabilities (Anon., 2023a). This paper reports on a pilot study to assess the suitability of epigenetic ageing - estimating the age of individuals by analysing the degree of DNA methylation in their tissue - for this purpose.

Development of an epigenetic clock for BFT will follow the approach used on other fish species (Mayne et al 2021 and 2022). In short, an epigenetic clock works by measuring DNA methylation at multiple cytosine-phosphate-guanine (CpG) sites and uses machine learning methods to translate DNA methylation to age. We have previously identified 1,311 CpG sites, known to be age associated in zebrafish (*Danio rerio*), from a total of several hundred thousand markers from genome-wide scans (Mayne et al., 2020). These sites are conserved over a wide range of teleost taxa and have been used to successfully calibrate epigenetic age models for a range of teleost species, including tuna (Davies et al, 2023; Mayne et al, In prep). Age associated sites from this set that are conserved in BFT genome will be targeted for amplification by multiplex PCR. The DNA is then sequenced using an Illumina MiSeq sequencer with high coverage to measure the degree of methylation at each site. This data is used in combination with the "known age" data, estimated from otolith readings, to construct a statistically calibrated epigenetic age model that can be used to independently predict age directly from the level of DNA methylation.

The aims of the pilot study are to: i) calibrate an epigenetic age model, using the approach developed by Mayne et al 2021, and tissue samples from eastern and western stocks of ABT and age data from standardized ICCAT otolith reading methods; ii) examine the influence of stock and sex on epigenetic age calibration; and iii) assess the relative cost-effectiveness for use in CKMR relative to the routine use of otolith-based age estimation for eastern BFT. In addition to the use in CKMR, if successful, epigenetic age would also be valuable for providing catch-at age data, which would be a significant improvement on current methods used in the regular stock assessment.

This report summarises the results from the full design and analyses from Davies et al. (2023b), including activities completed within GBYP Phase 12 and that additional work that was reported to the Bluefin Species Group meeting of the SCRS in September 2023 (Davies et al., 2023c) and addresses the additional questions that we raised at that meeting. It includes recommendations on the use of the method in CKMR for ABT and future work that should be completed to refine this preliminary calibration and use epigenetic age efficiently in a future implementation project for CKMR, should that proceed.

2. Methods

Source of samples

The study design aimed to examine the potential effects of stock of origin (eastern vs western) and sex on DNA methylation and the epigenetic age calibration. Accordingly, tissue samples and age data for the eastern stock were provided by AZTI, from the GBYP tissue bank, and samples and age data for the western stock were provided through the existing NOAA-CSIRO close-kin project for western BFT (Grewe et al, 2023 and Lauretta et al, 2023 in CKMR workshop report) via the sampling conducted by Department of Fisheries and Oceans, Canada, the University of Maine and NOAA Southeast Fisheries Centre, along the east coast of the USA. Table 1 provides a summary of the samples available to this pilot study by sampling region and age class.

The western samples were available as: i) previously extracted DNA stored in archive plates at -80 degrees Celsius, as they had previously been extracted for CKMR kin identification and provenance as part of the NOAA-CSIRO western ABT CKMR study. For these western samples, the extracted DNA was used for the sex determination and epigenetic age analyses, and the existing DArTseq files were used for the provenance determination for this study.

DNA extraction and QC

DNA extractions were prepared from approximately 15mg of tissue subsampled from individual biopsies. Samples were extracted on an Eppedorf EP motion 5057 liquid robotic handler using a modification of the

QIAamp® 96 DNA QIAcube HT Kit (QIAGEN, Hilden, Germany). This extraction includes a lysis step in the presence of Proteinase K followed by bind-wash-elute QIAGEN technology. Low quality/degraded samples were re-extracted using the modified CTAB method following Grewe et al. (1993).

Provenance determination

DArT-CAP sequence data were examined to confirm individual provenance by examination of stock specific markers obtained from a training set sampled from Gulf of Mexico (Grewe et al., 2021). Only samples with high degree of confidence for GOM stock of origin were designated western samples. The eastern (GBYP samples) were assumed to be of eastern stock origin for the purposes of the epigenetic age analysis.

For provenance analysis, DNA aliquots of the 250 eastern samples were shipped to Diversity Array Technologies (DArT) in Canberra for DNA complexity reduction and library construction prior to sequencing to generate genotype data for each individual. Identical markers and protocols to those used to generate the existing genotype data for the western samples were used for the eastern samples.

Individual DArT-SEQ library preps were sequenced on an Illumina platform that produced raw FASTQ sequence data files for each individual. The DArT-Soft14 (DS14) was used to generate multi-locus SNP genotype profiles from the raw FASTQ files. Individual genotypes were QC filtered using two separate pipelines, Kinference (Baylis et al, In prep) and RADIATOR (Gosselin, 2020) to delete poor quality (e.g., paralogous) loci and individuals exhibiting overall poor-quality DNA profiles (*e.g.* DNA cross-contamination and sequencing dropout). Data-cleaning and quality control workflows were completed prior to provenance analysis (see Davies et al., 2020, Grewe et al., 2021). DArT-CAP sequence data were examined to confirm individual provenance by examination of stock specific markers obtained from training set data sampled from Gulf of Mexico and Mediterranean larvae (Grewe et al., 2021). Only samples with high degree of confidence for stock of origin were chosen to represent Western and Eastern Stocks used for stock specific age analysis.

Sex determination

Genetic sex determination was completed using a set of 22 fluorescently labelled primer pairs in an 11-plex PCR assay that targets three male specific amplicons plus 8 autosomal DNA microsatellite loci. The assay produces sex-specific fragments that were separated and visualised on an ABI 3730 DNA sequencer. DNA microsatellite loci provide confirmation of a positive PCR resulting from good quality DNA and are also used to assess the degree of DNA cross-contamination for each individual tissue sample. Files for each individual are then genotyped using GeneMapper 4 softeware package (Applied Biosystems Inc.) following protocols developed by CSIRO (Grewe et al, unpublished) as part of the original southern bluefin tuna CKMR project (Bravington et al., 2016).

Epigenetic age calibration and prediction

Age-associated CpG sites were identified through genome pairwise alignment with known age-associated sites in zebrafish. Briefly, by targeting conserved and age-associated sites from another species, this reduces the expense in biomarker development. This process has been carried out previously in other species, including fish (Polanowski, A.M., et al.,2014, Mayne et al., 2021). CpG sites that were both known to be age-associated in zebrafish and conserved between both species were targeted for primer design in the multiplex PCR assay specifically for BFT.

Genomic DNA was extracted using the QIAamp 96 DNA QIAcube HT Kit (QIAGEN) as instructed by the manufacturer's protocol. DNA was bisulfite treated using a modified version of a previous protocol (Lu, A., et al., 2021). Multiplex PCR was designed for one pool of primers using PrimerSuite (Li, H., et al.2019). Each amplicon was tested in singleplex with annealing temperatures between 55-60°C. Primer pairs that produced multiple or no bands were excluded from the multiplex PCR reaction. Barcoding was carried out using the Fluidigm 384 set of barcodes (Cat. 100-4876). Barcode reactions were pooled together in equal volumes and were sequenced using an Illumina MiSeq Reagent Kit v2 with custom sequencing primers (300 cycle; PN MS-102-2002) at the Australian Genome Research Facility (AGRF).

Sequencing data were aligned to a representative genome for each species based on the amplicon sequences. DNA methylation for each CpG site was called as a percentage, where 0% means the site is unmethylated and

100% the site is fully methylated. Samples that had no sequencing coverage for any amplicon were removed from the analysis.

Epigenetic model fitting and validation.

An elastic net regression model and a new approach developed by CSIRO (Mayne et al., In prep.) that better accounts for the characteristics of the data were used to generate calibration models using all targeted CpG sites (Engebretsen, S. and J. Bohlin, 2019). The performance of the fitted models was assessed using five-fold cross-validation with the available dataset for each species divided into five equal-sized and disjoint subsets that were randomised to preserve the original distribution of otolith age as much as possible. For each of the five iterations, four subsets were used to train the five models, and then all models were evaluated for predicting otolith-age in the held-out fifth subset. For each iteration, the performance of the model is measured using the prediction R2, mean-squared error (MSE) and the mean absolute error (MAD). Averages over the five folds for each performance measure were used to summarise the results.

SCAM solution for stock differences

If differences in rate/pattern of methylation with age between stocks were present, then this would manifest in the individual methylation probe otolith age relationships. The expectation being that at some sites in the ABT genome the deviation in genetic alleles at a locus would lead to different relationships between age and rates of methylation at that locus. There is also the potential for variation in environments between stocks to influence the rates of methylation across the ABT genome.

We tested this by fitting a SCAM model that models both a mean difference in methylation score by stock as well as a by-stock smooth-faction interaction model for each methylation locus. Differences in mean levels can then be tested along with the difference between the smooth functions estimated for each stock tested by investigating whether there is evidence for deviations in the difference between the two smooth functions.

3. Results

Sample characteristics

More than two-hundred and fifty tissue samples were available from both sources (eastern and western) with otolith ages ranging from 2-27 years (Table 1). There were 15 year-classes in the eastern stock samples (2-16 years) with 17-20 samples per each age class between 2 and 15, providing good replication across this age range. There was approximately a 50:50 sex ratio for the selected eastern samples based on the phenotypic sex provided with those samples.

The western samples ranged from 4-27 years, with 9-33 samples per year class between ages 6-17 and fewer than 2 per age class above 17 years old. Again, this provided for sufficient replication out to 17 years, but low or no replication in the older year classes. Phenotypic sex was available for only a small subset of the western samples. Hence, it was not possible, *a priori,* to balance the selection the subset of samples for the western stock based on sex and, therefore, examination of the effect of sex on the epigenetic age estimation for the western samples relied on the success of the genetic sex determination (see below) to provide this covariate.

The age distribution of the available samples should provide for a high-quality calibration (5-10+ individuals) for ages 2-16 for the combined data set. Initial QC of the DNA extracted from the eastern samples indicated it was of high quality, with 246 (89%) of the 277 eastern samples passing the initial QC requirements for each of the downstream genetic analysis. Those that did not pass initial QC were re-extracted and included in the subsequent processing.

QC protocols and cross-contamination between individuals

Results of an early run of sex identification assays of the western samples indicated the potential presence of cross-contamination of DNA between individuals. This was evident as a low incidence of male DNA present in results for individuals samples that were phenotypically identified as female in the results from these initial runs. Following further detailed investigation, it was decided to use only samples from the Gulf of St Lawrence sampling program for the western stock, as there was a lower incidence of individual cross-contamination for samples from this source in the initial runs. The primary concern in this case was the potential for individual

DNA cross-contamination to potentially impact on the success/accuracy of the epigenetic age assay, not so much the impact on the sex assay, where this form of cross-contamination was generally evident in the resulting assay scores, if present. Although identification of cross contamination between two male individuals is more difficult than between male and female, given the design of the assay.

As a result of this decision to limit the western samples to the Gulf of St Lawrence, the final age distribution of samples sourced from each population and processed for provenance, sex and epigenetic age is presented in Table 2.

Provenance determination

As a result of delays incurred with the installation of a new genetics robot and the interaction with sample selection for age and sex, it was not possible to complete the sequencing and provenance analyses for the eastern samples in time for reporting to the September 2023 SCRS meeting, or as part of this project. Given this, it was necessary for the analysis of the effect of stock on the epigenetic age to assume that the samples provided by GBYP were, in fact, eastern origin fish. Given that most of the samples were captured within the Mediterranean, this is most likely the case. However, the presence of some western origin fish in the eastern samples, particularly those from the NE or SE Atlantic, cannot be discounted. The eastern samples were sequenced for provenance and the data is now available for future analysis. Notwithstanding this, the results of the current analysis of the effects of stock on the epigenetic age calibration indicate that the effect of stock of origin on methylation is small (if present at all) and not significant, at least for the sample sets analysed in this pilot study (see epigenetic age calibration below).

Sex determination

The sex determination assay was completed for 500 samples (250 eastern and 250 western), as per the revised sample set summarised in table 2. The results indicated that there was a significant degree of cross contamination of individual DNA among samples. This result stemmed from the allele specific nature of the PCR assay and was most apparent where low quantities of male markers were detected in samples that were otherwise scored as female. Due to this allele specific bias of the PCR assay, detecting instances of male individuals that have low levels of cross-contaminating female DNA is more challenging among true male individuals. Similarly, the assay is also less sensitive to detecting low levels of additional male DNA cross-contamination. As a result, 361 of the 500 samples were able to be confidently scored a male or female. Females with observed male DNA cross-contamination presented as ambiguous or "mixed" amplification profiles (n = 139 samples) and were excluded from the epigenetic age calibration (Table 3).

Epigenetic age: model selection, calibration and prediction

DNA extractions for the epigenetic age calibration was completed for the complete set of samples presented in Table 2. Post quality control data consisted of methylation measured at 43 CpG sites for 361 Atlantic bluefin tuna (Table 3). Most of the samples lost from the total sequenced ~500 were lost to quality control thresholds at the bisulphite treating stage. The reasons for these QC failures are multitude, and require further detailed investigation to be specific, but most likely due to organic contamination of the original sample (i.e. organic by-products of the extraction process that cause the bisulphite treatment to fail). Methylation at the 43 CpG sites ranged from 0 (unmethylated) to 100 (methylated) with these values representing the percentage of reads that were methylated at each CpG site.

Age estimates for Atlantic bluefin samples were obtained directly from the GBYP and NOOA-DFO ageing programs completed with consistent counting methods and partial age calculations. Decimal otolith age estimates ranged from 2.373 to 27.710 with a mean of ~10.8 years (See Figure 1 for otolith age distribution). Sample numbers appeared reasonably balanced between eastern and western Atlantic samples (Table 3) with most samples derived from the Canadian 2018 samples (Table 3) and the sex of samples was ~70% male (Table 3).

Patterns of methylation with age

Scatterplots of the methylation percentage on otolith age for each CpG site showed highly non-linear relationships (Figure 2). Across the probes, asymptotic behaviour was observed at the tails of the otolith age distribution (Figure 2). A diverse set of relationships were observed across the CpG sites with varying asymptotes and increasing variance at varying ranges of the age distribution (Figure 2). Generally, asymptotic behaviour was observed for

ages less than 5 and greater than 20 years. These plots indicate that fitting non-linear relationships at an individual probe scale could improve predictions of otolith age. We compared the non-linear models versus standard linear models and the elastic-net, which is the standard model fitted for tuning epigenetic clocks.

Calibration models

Mayne et al. (In prep) compared a series of alternative modelling approaches for developing calibration and prediction models for bigeye (*Thunnus obesus*), yellowfin (*Thunnus albacares*), and southern bluefin tuna (*Thunnus maccoyii*) for comparison with the elastic-net approach that is conventionally used to calibrate epigenetic age clocks. Those methods were refined for application to the Atlantic bluefin methylation data in this project. Differences for the fitting of the nonlinear shape-constrained generalised additive models (SCAMs) used by Mayne et al., (In prep), included the addition of a larger 'gamma' parameter in the model fitting, which controls the smoothing of the SCAM fits with larger values producing smoother functions. The goal is to avoid over smoothing at the edges of the age distribution preferring a monotonic linear function at the edges of the modelled otolith age range. A gamma = 30 was chosen and is an arbitrary choice at the initial stage of model comparisons. Cross-validation was used at a later stage to choose the best value (see last section of results).

For a large set of CpG sites, the nonlinear models showed asymptotes past age 20 (e.g., Figure 2). When combining all probes this manifests as poor predictions at the tails of the age distribution. We therefore chose probes that showed less asymptotic behaviour and low noise at the tails of the age distribution, which left five probes to build a predictor from the nonlinear SCAM modelling. We again compared this model against multiple regression with just these probes and with all probes (models m1-m3) and with a log transformation on age. We also ran the elastic-net with alpha set at 0.5 with not log transformation on age (model m4). Five-fold cross validation with a set of prediction accuracy metrics were assessed.

Across the five models compared, prediction accuracy metrics from the cross-validation analysis showed strong predictive capacity (e.g., $R2 > 0.96$) for the nonlinear SCAM model, multiple regression using all probes, and elastic-net models (Table 4 and Figure 3). The other models showed poorer but still reasonable accuracy.

The multiple regression model had, marginally, the best predictive performance, with all three models showing prediction deviations near 0.5 years on average. The largest deviations were 4.7, 3.8, and 6.8 years for the SCAM, multiple regression and elastic-net models, respectively, and were all for the largest recorded age of ~27 years. For ages less than or equal to 20, the maximum deviations were 1.98, 2.11 and 1.78, respectively for these three best performing models.

Effects of stock on epigenetic clock age prediction

The age ranges for the samples from GBYP for the eastern stock showed different ranges to those of the western stock (Table 2). On median Western ages were nearly three years older (12.7 versus 8.87 years) and contained the oldest individuals in the samples analysed (Table 5). Eastern samples had individuals with the lowest ages and covered the ages less than five more comprehensively. Visually, scatterplots of methylation percentages on otolith age appeared different between samples from the two assigned stocks. When adjusting for the portion of the age range that the data are representing in each stock it is more convincing that for most probes the relationships are quite consistent between the stocks.

Across all methylation sites instances of differences between smooths are present (e.g. Figure 4). However, from inspection of these results it is most likely that these differences are driven by the non-linear relationships between age and methylation percentages and the lack of coverage of age for low ages by the western stock and the higher ages for the eastern stocks. For some probes this lack of coverage leads to a difference in the fitted curves. For those methylation probes that show evidence for differences these differences are nearly always in areas of the age range that do not overlap between the stocks. From these observations we conclude that the best predictor of future otolith age from this pilot study should be based off the combined-stock SCAM model. If only those age ranges were measured in the future for each stock, then a by-stock predictor may perform better than the combined model.

Influence of heteroscedasticity and selection of probes on epigenetic clock performance

We observed evidence for heteroscedasticity for the probes with variance in methylation score higher for higher methylation values (see Figure 2 for representative sample). From the SCAM fit we tested the evidence of heteroscedasticity in the residuals using the Breusch-Pagan test. Visual inspection of residuals versus fitted plot

coupled with the test results showed evidence for heteroscedasticity for most probes. To mitigate this, we implemented a weighted regression approach.

We estimated the conditional variance function, by fitting a GAM to the squared residuals from the SCAM fit against the otolith age for each probe and used the fit at each point on the age curve as an estimate of the variance. The inverse variance was then used as weights in a weighted SCAM model. The variance in the weighted predictor across probes now included a weight at each point estimated from the conditional variance function. This had the benefit that predictions from parts of the otolith age distribution that show higher variance in methylation percentage are given less weight in the combined predictor than those with low variance.

In the previous model selection analyses, we chose the best set of probes using an ad-hoc choice of best fit probes that had low variance at different parts of the age range and were from unique regions of the ABT genome. We had also chosen a value of 'gamma' parameter in the model fitting, which controls the smoothing of the SCAM fits with larger values producing smoother functions. In this analysis, we chose gamma by running the updated heteroscedastic model described above, using all probes and five-fold cross validation. A grid of gamma values ranging from 1-30 we investigated. For higher values of gamma, the fits are less 'flat' in the tails, which performs better for some probes and eliminates extreme predictions for some probes. A gamma=10 showed the best average MSE across the five-folds and was fixed in the following analysis.

This approach for the choice of gamma facilitated the implementation a forward selection process that starts with the best probe from running five-fold cross validation using single probes to predict otolith age. The process then runs all two-probe models including the best probe from iteration one and all other probes. The process proceeds for all three probe models and so forth until the average MSE over the five folds does not decrease any further. This iterative selection process showed that 11 probes was optimal with no improvement adding probes past this (Figure 5). Marginal improvements in prediction parameters were observed with these tuned gamma and best-set of-probes version of the INV model (Table 6), relative to those presented in Table 4 for the original comparison.

4. Discussion

The aim of this project was to evaluate the feasibility of applying epigenetic ageing to Atlantic bluefin tuna to facilitate the application of CKMR to the eastern stock of this species. The results presented here demonstrate it is possible to calibrate an epigenetic clock for BFT with high predictive performance and relative low average error. They also demonstrate the value of exploring the pattern of methylation and probe selection for each new species at the calibration step in order to select the most appropriate forms of calibration model, given the nature of the data, and maximise predictive performance. The results of the analyses presented here show the considerable promise of the method for BFT and are consistent with the results for other tuna species (Mayne et al., In prep). They demonstrate that epigenetic age is a viable alternative to conventional otoliths to estimate age for CKMR studies for either stock, but particularly for the eastern stock where large scale collection of otoliths is expensive and logistically challenging.

The combined issues of i) organic contamination in the bisulphite step of preparing the samples to measure the level of methylation, and ii) the large number of incidents of cross-contamination of DNA among female individuals in the sex assay, meant that the final data set available for analysis in this pilot study was not as balanced, or powerful, as originally intended. However, the high correlation coefficients and close fits of to the data demonstrate there was no strong evidence to suggest methylation rates varied between sexes or between samples from the eastern or western stocks. The results of the multi-model comparison indicate that the final INV model, with an R^2=0.982 and MSE = 0.341 would be the most appropriate calibration model to use for age prediction for both stocks BFT, if necessary, until an updated calibration is available that addresses the imbalance in old/young fish for the eastern and western stocks, respectively.

Notwithstanding this, as highlighted, there remain a number of issues that require resolution before initiating large-scale sampling and processing of samples for epigenetic age. The most important of these, in the context of the feasibility of large-scale sampling for CKMR for eastern BFT, are: i) to eliminate, or minimise, the cross-contamination of individual DNA and ii) the source of the residual organic contamination from the extraction step that caused a larger proportion of samples from both sources to be excluded from the final epigenetic clock analyses. The latter of these should be relatively straight-forward as it is likely to be a technical issue associated with the tissue of this species and/or the extraction and preparatory protocols used. While not common, this has been encountered for other scombrid species and has been successfully addressed by isolating the issue and refining the lab protocols accordingly. The former may be more challenging to isolate with complete certainty,

given the number of points at which individual DNA cross contamination can occur. This is best done by those directly involved in the sampling and sample management programs. In the context of large-scale sampling for CKMR, one option is to use purpose designed, one use, biopsy tools for the CKMR tissue sampling program. These not only have the advantage of eliminating/reducing to acceptable levels DNA cross-contamination between individual samples, but they can substantially reduce the time and cost associated with large-scale genetic programs when integrated into sampling, sample management and DNA extraction work-flows (Bradford et al 2016; Preece et al 2023; Anon, 2023b).

The other key issue is the cost associated with using epigenetic ageing. This has two components: i) calibration and ii) routine ageing of samples once an epigenetic clock has been calibrated (Mayne, 2023). The former is a more expensive, generally, one-off investment. It requires a genome of the species (or a closely related relative) to identify the appropriate age-related markers and a "calibration set" of matched high quality otolith ages and tissues samples from the same individuals, which are used to calibrate the level of methylation to "known" age. The results of this study provide a calibrated epigenetic clock that can be used for either stock, albeit with identified weaknesses that can be addressed in the future when the required samples are available. The extent to which this is required is best addressed as part of the detailed design study for the CKMR, as it will depend to some extent on the age ranges required for the eastern stock. For example, if the CKMR design for eastern uses larvae (known age 0+) and adults less than 15 years old, then there may be limited value in pursuing an improved calibration for ages 15 and above.

The second, more important, cost consideration is the unit cost for routine processing of what is likely to be 10,000's of samples. In the case of this project, where the entire process was done "in-house" in a research agency, the unit cost was in the order of ~USD65. This is certainly higher than what it will cost in the future from an established high through-put commercial service[1]. The uncertainty is how much higher this current research cost is relative to a commercial service now, or in 2 to 5 years' time. This uncertainty arises because of a range of factors associated with the relative novelty of the technology, including, the need to develop high-through put workflows and associated QC protocols and costs associated with doing so, and the scale of the commercial market. We consider that, in the short-term, it is reasonable to expect that the unit cost will decline by at least a third over the coming 2-3 years, assuming a similar level of performance (accuracy and consistency) can be obtained from high through-put work flows as has been demonstrated here.

Conclusions

This pilot study has demonstrated that epigenetic ageing is feasible for Atlantic bluefin tuna and is likely to be more cost-effective than ages obtained through conventional otolith ageing, particularly in the context of the sampling (and sample sizes) required for Close-kin Mark Recapture. From the samples that were successfully analysed for methylation, 361 of the ~500 processed, and a range of alternative calibration models examined, the INV model provided that best performance, although three other variants also had high predictive performance. For this sample set, there was no strong evidence to suggest that either sex or stock of origin had a measurable effect on the epigenetic age relationship.

A large proportion of the total samples available were either i) excluded from the analysis due to direct indication of DNA cross-contamination between individuals, or ii) failed to reliably score for the level of methylation, due to high levels of residual organic contamination from the preparatory workflows. Addressing these two separate issues should be a priority for future work, should ICCAT decide to pursue CKMR for E-BFT and use epigenetic ageing for that purpose. In the context of i), single use biopsy tools should be considered as a means to a) reduce/eliminate this source of DNA cross-contamination and b) increase the efficiency of field and laboratory protocols for the high-throughput work-flows required for CKMR and, in doing so, reduce the costs.

5. Acknowledgements

We thank Francisco Alemany and Igaratza Fraile for their cooperation in arranging transport of the tissue samples and associated data from the GBYP Tissue Bank. Nicola Potter and Chloe Anderson provided excellent technical

[1] **Note: potential for a conflict of interest**. CSIRO, as the developer of this form of epigenetic ageing, is actively seeking to licence this technology to commercial laboratories so that it is available more widely as a high-through-put commercial service.

assistance with the genetics and Shane Bayliss assisted with the retrieval of the DArT-Cap genotype and meta-data for the western samples.

This work has been carried out under the ICCAT Atlantic-Wide Research Programme for BFT (GBYP), which is funded by the European Union, several ICCAT CPCs, the ICCAT Secretariat, and other entities (see https://www.iccat.int/gbyp/en/overview.asp). The content of this paper does not necessarily reflect ICCAT's point of view or that of any of the other sponsors, who carry no responsibility. In addition, it does not indicate the Commission's future policy in this area.

6. References

Anon., 2023. Report of the workshop on Bluefin Tuna close-kin mark recapture (CKMR), including biological sampling coordination. Madrid, Spain, hybrid, 1-6 March, 2023.

Anon b, 2023. Progress towards a Close-kin Mark-Recapture Application to South Pacific Albacore (Project 100C). Information paper (WCPFC-SC19-2023/SA-WP-13) to the 19th meeting of the Scientific Committee of the WCPFC, August 2023, https://meetings.wcpfc.int/node/19361.

Bradford, R.W., Hill, P., Davies, C. and Grewe, P. 2016. A new tool in the toolbox for large-scale, high-throughput fisheries mark-recapture studies using genetic identification. Marine and Freshwater Research 67: 1081-1089.

Davies C, Marsac F, Murua H, Fraile I, Fahmi Z,et al., 2020. Population Structure of IOTC species and sharks of interest in the Indian Ocean: Estimation with Next Generation Sequencing Technologies and Otolith Micro-chemistry. Final project Report to IOTC. 249p.

Engebretsen, S. and J. Bohlin, *Statistical predictions with glmnet.* Clinical Epigenetics, 2019. **11**(1): p. 123.

Grewe, P.M., Baylis, S., Bravington, M.V., Aulich, J., Devloo-Delva, F., Reglero, P., Alvarez-Berastegui, D., Martin, M., Gosselin, T., McDowell, J. Lauretta, M., Walter, J. and Davies, C.R. 2021. Genomic assessment of Atlantic bluefin tuna larvae collected from the Mediterranean Sea: Refinement of SNP markers for stock of origin assignment (Western, Gulf of Mexico versus Eastern, Mediterranean Sea) of adult fish used for genetic Close-Kin Mark-Recapture. Report to the Pew Charitable Trust. May 2021, 12p.

Li, H., et al., *Cost-reduction strategies in massive genomics experiments.* Marine Life Science & Technology, 2019. **1**(1): p. 15-21

Lu, A., et al., *Universal DNA methylation age across mammalian tissues.* bioRxiv, 2021: p. 2021.01.18.426733.

Mayne, B., et al., 2021. *Nonlethal age estimation of three threatened fish species using DNA methylation: Australian lungfish, Murray cod and Mary River cod.* Molecular Ecology Resources, 2021. **21**(7): p. 2324-2332.

Mayne, B., et al., 2022. *Age prediction of green turtles with an epigenetic clock.* Molecular Ecology Resources, 2022. **00**: p. 1– 10

Mayne. B., et al., In prep. A multi-species epigenetic clock for bigeye (*Thunnus obesus*), yellowfin (*Thunnus albacares*) and southern bluefin tuna (*Thunnus maccoyii*).

McDowell, J.R., Bravington, M., Grewe, P.M.. et al., 2022. Low levels of sibship encourage use of larvae in western Atlantic bluefin tuna abundance estimation by close-kin mark-recapture (2022). *Scientific Reports*, 12(1), 18606.DOI: 10.1038/s41598-022-20862-9.

Polanowski, A.M., et al., Epigenetic estimation of age in humpback whales. Molecular Ecology Resources, 2014. 14(5): p. 976-987.

Preece A.L. and Bradford R.2023. Update on the Gene- tagging program 2023 and RMA request. Working- paper submitted to the Extended Scientific Committee, for the Twenty Eighth Meeting of the Scientific Committee, 28th August – 2nd September 2023. https://www.ccsbt.org/system/files/2023-08/ESC28_09_CCSBT_GT2023.pdf.

TABLES

Table 1: Summary of source, sampling region and age of tissue samples/DNA available for eastern and western stocks of BFT for pilot study on the suitability of epigenetic ageing. GBYP = Grand Bluefin Year Program administered by ICCAT. MED = Mediterranean, NALT = North Atlantic, SALT = South Atlantic, WALT = Western Atlantic. NOAA/UM/DFO – collaborative sampling program between US-NOAA, University of Maine and Department of Fisheries and Oceans-Canada. CAN= Canada, MAI = Maine.

		Age class (years)																											
Source	Region	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**Total**
GBYP	MED	17	15	19	18	14	17	16	15	17	11	12	15	17	5	1	-	-	-	-	-	-	-	-	-	-	-	209
(Eastern)	NATL	0	0	0	0	1	0	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4
	SATL	0	1	1	2	3	4	2	3	-	-	6	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	33
	WATL	-	-	-	-	-	-	1	-	-	1	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	4
	Total	17	16	20	20	18	21	20	19	19	20	18	18	18	5	1	0	0	0	0	0	0	0	0	0	0	0	**250**
		2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**Total**
NOAA/UM/DFO	CAN	-	-	1	-	-	8	24	24	26	15	31	33	21	14	8	9	2	2	1	2	2	-	-	-	-	1	231
(Western)	MAI	-	-	-	-	3	4	6	1	-	1	2	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	19
	Total	0	0	1	0	9	12	30	25	26	16	33	33	22	14	9	9	2	2	1	2	2	0	0	0	1	1	**250**

Table 2: Summary of samples of BFT processed by sample source for pilot study on the suitability of epigenetic ageing. GBYP = Grand Bluefin Year Program administered by ICCAT. MED = Mediterranean, ALT = Atlantic, NOAA/UM/DFO – collaborative sampling program between US-NOAA, University of Maine and Department of Fisheries and Oceans-Canada. CAN= Canada.

		Age class (years)																										
Source	Region	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**Total**
GBYP (Eastern)	MED	20	23	19	21	18	18	17	15	14	13	13	15	15	4	1	-	-	-	-	-	-	-	-	-	-	-	226
	ATL	1		2	3	5	5	5	3	4	6	7	2	1	-	-	-	-	-	-	-	-	-	-	-	-	-	44
NOAA/UM/DFO (Western)	CAN	-	-	2	2	12	30	49	38	36	34	56	56	32	11	8	9	2	-	2	4	2	-	-	-	1	1	387
	Total	**21**	**23**	**23**	**26**	**35**	**53**	**71**	**56**	**54**	**53**	**76**	**73**	**48**	**15**	**9**	**9**	**2**	**0**	**2**	**4**	**2**	**0**	**0**	**0**	**1**	**1**	**657**

Table 3: Summary of results of sex assay to determine genetic sex (female, male) of samples for epigenetic age calibration of Atlantic Bluefin Tuna. "Mixed" samples represent samples for which it was possible to confidently identify the likely sex of the individual and the presence of cross-contamination with DNA from another individual(s).

Sample Source	**Female**	**Male**	**Mixed**	**Total**
Atlantic	3	7	1	**11**
Mediterranean	31	115	31	**177**
Western	6	113	54	**173**
Total	**40**	**235**	**86**	**361**

Table 4: Predictive model performance metrics averaged over five-folds from cross-validation for five models tested. The first column corresponds to the inverse-variance weighted SCAM model; Model 1, M1) multiple linear regression with otolith-age as the dependent variable and the lowest-variance probes from the SCAM analysis as predictor variables, M2) model 1) but with otolith-age log-transformed, 3) multiple linear regression with log-transformed otolith-age as the dependent variable and all 54 probes as predictor variables, and 4) elastic-net regression model with otolith-age not transformed.

MSE					MAD					R^2				
IVW	M1	M2	M3	M4	IVW	M1	M2	M3	M4	IVW	M1	M2	M3	M4
0.626	0.830	2.570	0.434	0.562	0.606	0.672	1.060	0.494	0.496	0.962	0.948	0.859	0.973	0.969

Table 5: Summary statistics for otolith age distributions of samples from eastern and western ABT stocks.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
East	2.373	6.707	8.874	9.486	12.79	21.71
West	4.87	9.62	12.71	12.36	13.87	27.71

Table 6: Summary of prediction performance metrics for non-linear inverse variance weighted regression with forward selection and heteroscedasticity adjustment.

MSE	**MAD**	$\boldsymbol{R^2}$
0.341	0.401	0.982

FIGURES

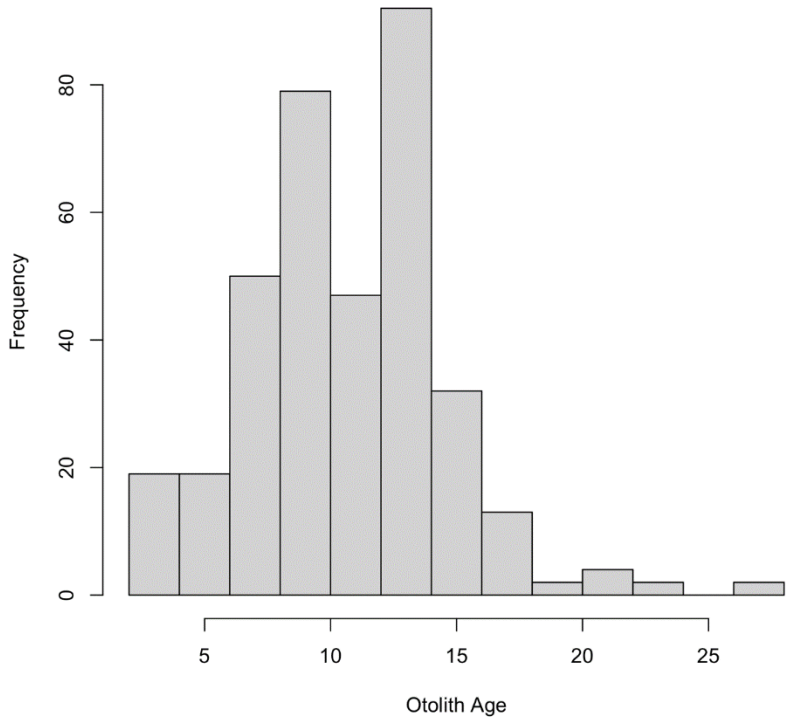

Figure 1: Histogram of otolith ages from 361 Atlantic bluefin tuna used in epigenetic clock analysis.

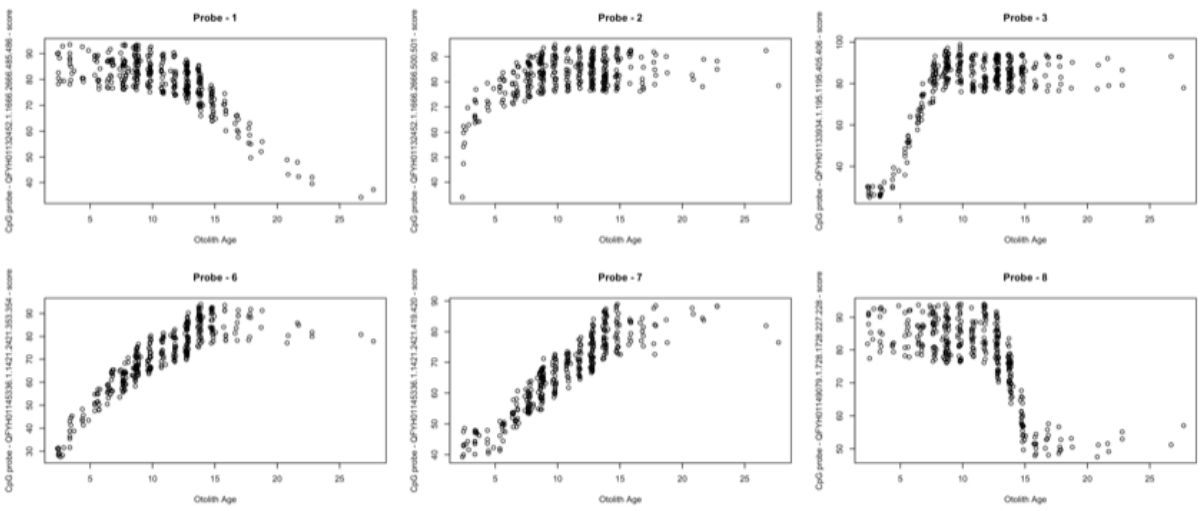

Figure 2: Raw scatter plots for a representative sample of probes to highlight heteroscedasticity of percentage methylation by age.

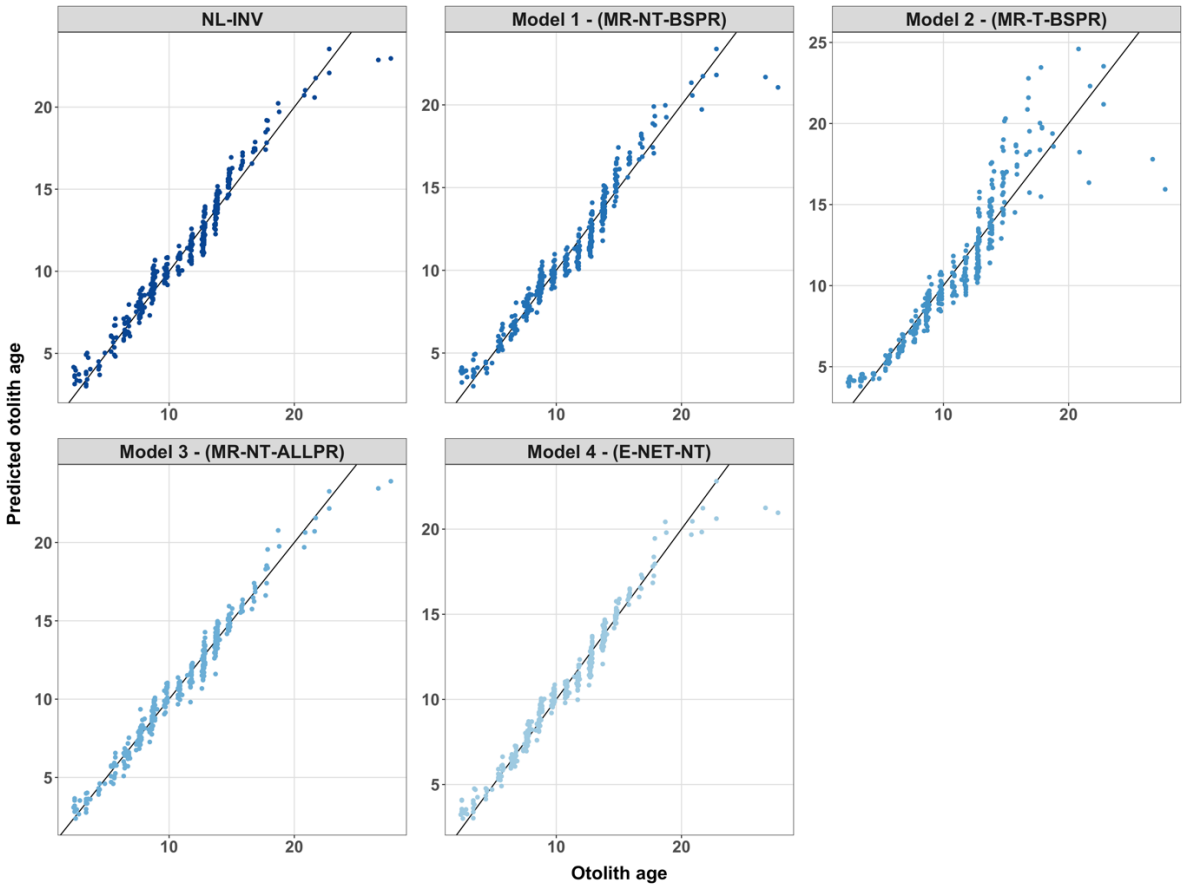

Figure 3: Visualisation of epigenetic clock predictive model performance from five-folds cross-validation for Atlantic bluefin tuna. The predictions from the five cross-validation test sets are presented as an aggregated plot for each of the models. The solid black line is the y = x line. Panel headings corresponds to the nonlinear inverse-variance weighted SCAM model (NL-INV), Model 1) multiple linear regression with otolith-age as the dependent variable and the lowest-variance CpGs from the SCAM analysis as predictor variables, 2) model 1) but with otolith-age log-transformed, 3) multiple linear regression with not transformed otolith-age as the dependent variable and all 43 CpGs as predictor variables, and 4) elastic net regression model with otolith-age not transformed as the data showed less evidence for requiring a transformation.

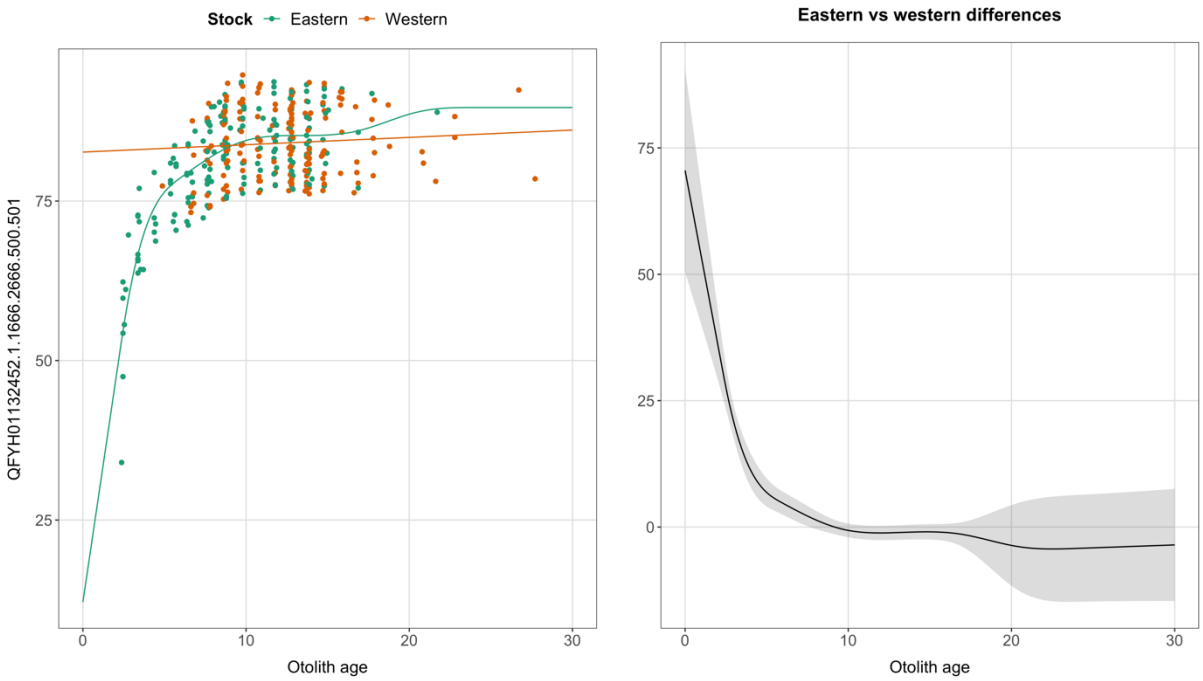

Figure 4: Fitted SCAM stock-smooth interaction model and differencing test. Left panel shows the fit of the SCAM smooth-stock interaction model and the right panel the difference between the smooths along with the 95% confidence interval in grey. This methylation site is an instance in which there is a presence of a difference between the two smooths. The difference is likely driven by the lack of age coverage for lower ages for the western stock rather than a difference in how the ABT genome is methylated at this locus.

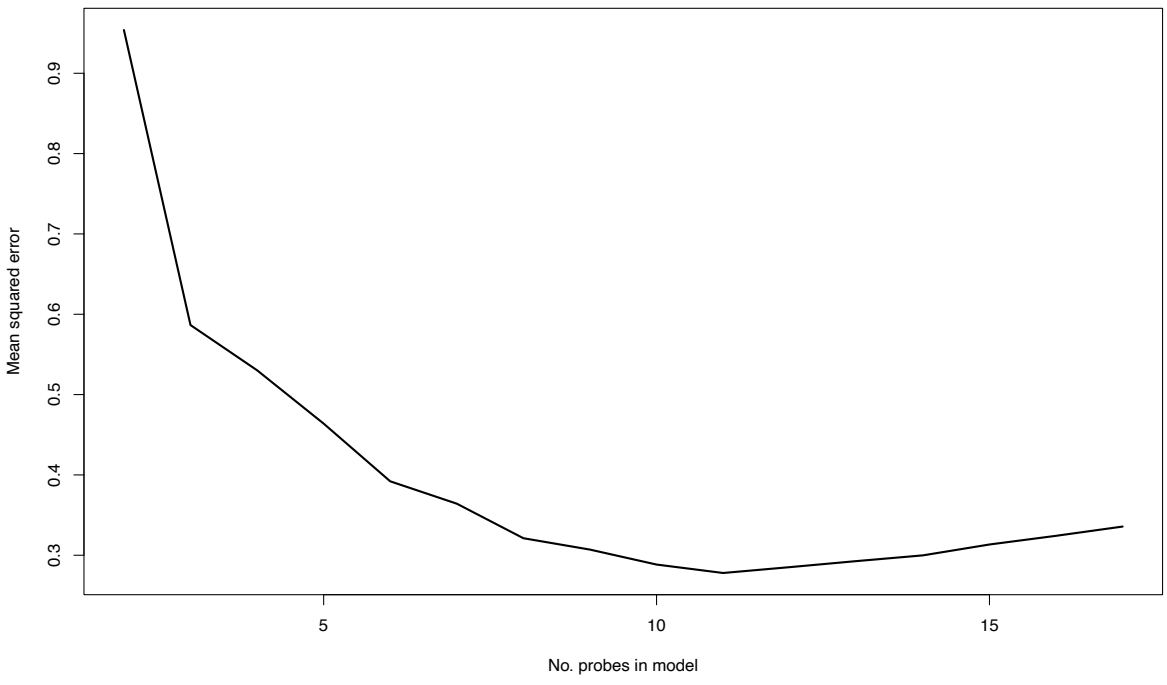

Figure 5: Average mean squared error change as more probes are added in the forward selection procedure.

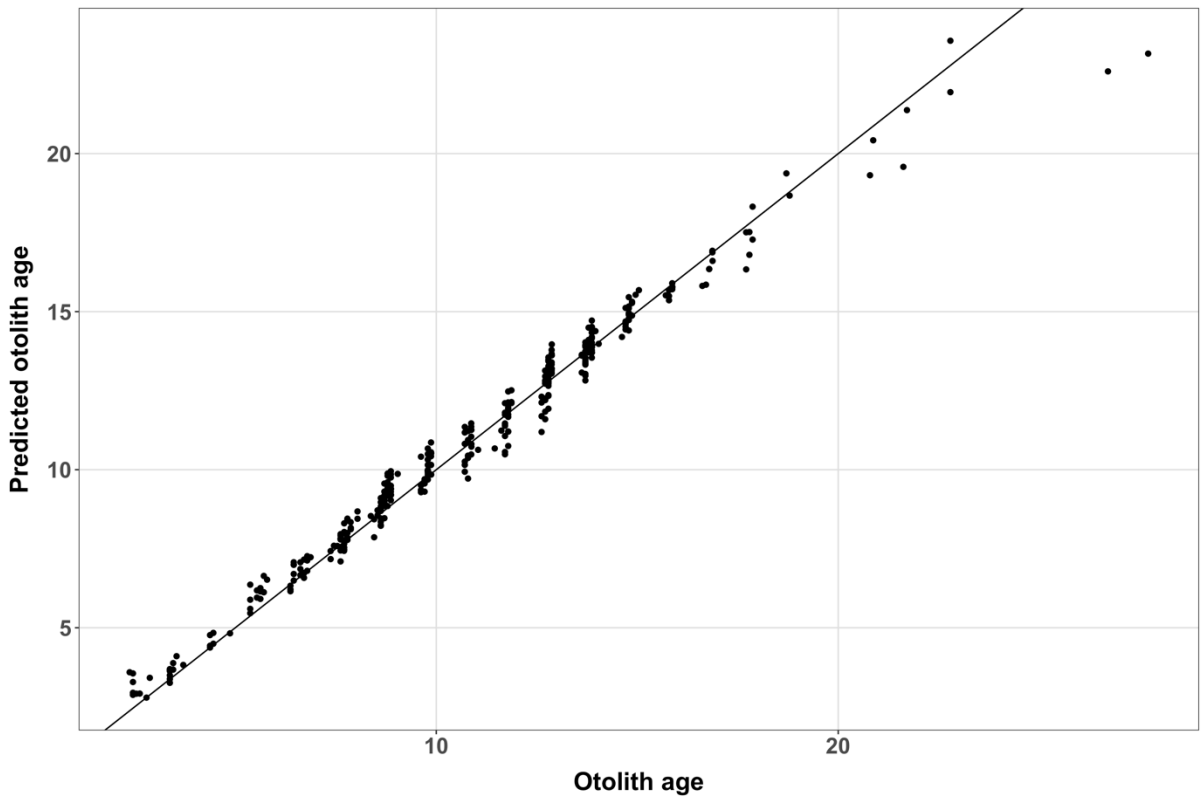

Figure 6: *Predicted otolith ages calculated using the NL-INV with improved accounting for heteroscedasticity and probe selection. The predictions from the five cross-validation test sets are presented as an aggregated plot for each of the species. The solid black line is the y = x line.*

Appendix 1: Abridged Terms of Reference for ICCAT GBYP 02/2023 – Pilot study on epigenetic ageing technique for age estimation for Atlantic bluefin tuna Atlantic-wide research program for bluefin tuna (ICCAT GBYP – Phase 12).

The contractor will perform a pilot study on epigenetic ageing technique (DNA methylation-based analysis) for age estimation of Atlantic bluefin tuna, according with the following terms of reference:

1. The aim of the pilot study will be to evaluate the accuracy of the epigenetic clock in comparison with direct age readings. Precision vs costs must also be considered in this comparison.
2. At least 500 samples of Atlantic bluefin tuna specimens (250 from each management areas, West vs East Atlantic plus Mediterranean Sea) will be analyzed. In these 250 samples for each management area, the entire age range should be represented, with at least approximately 10 specimens by age class. Samples used for testing the epigenetic clock should include individuals of both sexes. Specimens captured in West or East Atlantic management areas should cover all possible geographical locations. Most of the samples caught in the East area should come from the Mediterranean.
3. The genetic origin of each sample must be accredited for those samples for which genetic origin has not been previously assigned.
4. Muscle samples from specimens will be provided to the analytical team from the GBYP and other laboratories tissue banks. Sample metadata to include sex, location and date of sampling, gear and sampling lab will be provided to the analytical team. Direct ageing by means of schlerochronological methods will also be provided based on standard and validated protocols.
5. In order to reduce costs and meet the deadlines for the pilot study, samples of specimens with genetic origin already assigned and/or whose otoliths have been previously read (with high reading confidence) should be favoured.

Clarification on "accuracy of epigenetic clock" in ToRs

The application of epigenetic ageing to a particular species/taxon involves a number of steps, including the identification of appropriate epigenetic sites (CpG sites) and optimisation of the multiplex PCR for sequencing; (see more detailed explanation in 2 Methods, below); followed by calibration of the analysis model with samples from "known age" animals (see more detailed explanation in 2 Methods, below). In the case of tuna, the "known age" samples are sourced from individuals with high confidence, otolith-based, age estimates. As a result, the accuracy of the calibrated epigenetic age estimates are influenced, to a degree, by the accuracy of the otolith-based age estimates used in the calibration, and by the age signal from the degree of methylation at multiple CpG sites. Given the lack of true "known age" fish (e.g., animals sampled from breeding programs) that could be used as an independent source of "true age" for BFT, comparing the accuracy of the epigenetic age independently of the otolith ages is not possible. We can, however, estimate the precision and the level of uncertainty associated with each source as part of the development of the calibration model. This is how we propose to assess the performance of the epigenetic age model.