

RESOURCE ARTICLE OPEN ACCESS

# An Epigenetic Clock for Accurate Age Prediction in Atlantic Cod Populations for Improved Fisheries Management

Dafni Anastasiadi<sup>1</sup>  | Yassine Kasmi<sup>2</sup>  | Christoph Stransky<sup>3</sup>  | Laura Casas<sup>4</sup>  | Erik Eschbach<sup>5</sup>  | Francesc Piferrer<sup>6</sup> 

<sup>1</sup>Plant and Food Research Group, New Zealand Institute for Bioeconomy Science Limited, Nelson, New Zealand | <sup>2</sup>Thünen Institute of Fisheries Ecology, Bremerhaven, Germany | <sup>3</sup>Thünen Institute of Sea Fisheries, Bremerhaven, Germany | <sup>4</sup>Institute of Marine Research (IIM), Spanish National Research Council (CSIC), Vigo, Spain | <sup>5</sup>Thünen Institute, Public Relations, Braunschweig, Germany | <sup>6</sup>Institute of Marine Sciences (ICM), Spanish National Research Council (CSIC), Barcelona, Spain

**Correspondence:** Dafni Anastasiadi ([dafanast@gmail.com](mailto:dafanast@gmail.com)) | Francesc Piferrer ([piferrer@icm.csic.es](mailto:piferrer@icm.csic.es))

**Received:** 19 October 2024 | **Revised:** 27 January 2026 | **Accepted:** 30 January 2026

**Keywords:** age determination | biomarker | conservation | DNA methylation | fish | stock assessment

## ABSTRACT

Fisheries management relies on accurate stock assessments, which in turn depend on precise age information. Recent molecular tools called ‘epigenetic clocks’ harness age-related DNA methylation changes to build accurate and precise age-prediction models. However, the influences of intrinsic and extrinsic factors on clock performance remain uncertain. In this study, we examined Atlantic cod aged 0 to 7 years, sampled from various locations across the North Sea, and developed an epigenetic clock using DNA methylation data of 73 CpG sites from fin clips obtained by bisulfite restriction-site associated DNA sequencing (bis-RAD-seq). This clock predicted age with 97.5% accuracy and a precision of 2.8 months and generalised well in unseen data. Further, we addressed critical variables such as sex and maturity status, which are often overlooked, and we showed that clock performance was unaffected by sex-specific differences in growth, and it was lower in advanced sexually mature individuals, reflecting a slight bias towards younger fish. A key finding of our study is the discovery of a latitudinal cline in global DNA methylation patterns. We found that DNA methylation varied with latitude, despite the absence of genetic differences, while our clock maintained consistent performance across geographic locations. This resolves a major question regarding how generalizable epigenetic clocks are within the distribution of a species. Our clock demonstrates extensive applicability and enhanced practicality for real-world fisheries management. It provides accurate and precise age prediction for Atlantic cod irrespective of intrinsic differences or environmental influences associated with geographic locations.

## 1 | Introduction

The iconic Atlantic cod (*Gadus morhua*) has been one of the most historically targeted and heavily exploited fish species. Atlantic cod stocks declined dramatically, or even collapsed, in the 20th century owing to overfishing and environmental changes, prompting the implementation of management plans to help stock recovery. Fishery stock assessments utilise information on both the fisheries and fish populations. Key population parameters essential for stock assessment include

abundance, fecundity, recruitment, age and mortality (Casas and Saborido-Rey 2023). Age is used to estimate life history traits such as age at first maturity, growth and mortality rates, and the demographic structure of the populations (Campana 2001; Gebremedhin et al. 2021). Atlantic cod fishery management plans rely on key factors such as survey coverage, number of subsamples and age information. The latter is based on otolith readings, which are time-consuming, lethal, laborious and require highly skilled personnel (Gebremedhin et al. 2021; Piferrer and Anastasiadi 2023), while they show an accuracy of <80% and

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *Molecular Ecology Resources* published by John Wiley & Sons Ltd.

uncertainties due to different interpretations of growth zones by age readers (Moen et al. 2023).

Modern molecular methods based on high throughput sequencing of epigenetic marks are revolutionising age estimation in non-model species (De Paoli-Iseppi et al. 2017; Heydenrych et al. 2021; Tangili et al. 2023). Since the development of the first epigenetic clocks in humans (Bocklandt et al. 2011; Horvath 2013), several highly accurate and precise epigenetic clocks have been developed across a wide range of species, including other primates (Horvath et al. 2023; Jasinska et al. 2022), marine and terrestrial mammals (Hernandez et al. 2023; Polanowski et al. 2014; Wilkinson et al. 2021), ectotherms like fish (Anastasiadi and Piferrer 2020; Piferrer and Anastasiadi 2023; Weber et al. 2022, 2025), as well as marine invertebrates (Fairfield et al. 2021) and even a tree (Gardner et al. 2023). Epigenetic clocks are based on the identification of ‘clock-like’ changes in specific cytosine-guanine (CpG) loci that increase or decrease DNA methylation in a time-dependent manner among the general trend of genomic hypomethylation with age, a phenomenon well described in human genomes (Heyn et al. 2012; Jones et al. 2015). Epigenetic clocks often outperform traditional age estimation methods in fishes, as they consistently show extremely high accuracy and precision. When assessing the performance of epigenetic clocks, accuracy is measured by the correlation coefficient between chronological age and estimated age, with values  $>0.95$  being nowadays standard in fish species and commonly achieved under ideal circumstances (Piferrer and Anastasiadi 2023; Weber, Fields, et al. 2024; Weber et al. 2025). Precision is typically measured in years, with values usually below 1 year and sometimes around half a year being common. This is in sharp contrast to otolith age estimation, where, aside from the accuracy challenges in some species, precision is typically above 1 year, and sometimes even greater (Helser et al. 2019). In addition, epigenetic clocks can be non-lethal depending on the target tissue (for instance, fin-clips or blood), can be cost-effective if a ready-to-use kit is developed, show higher consistency between labs, and require only training for standard molecular biology techniques (Anastasiadi and Piferrer 2023).

For piscine epigenetic clocks to be effective in fisheries stock assessments, they must perform consistently across the entire target population. DNA methylation at the same loci varies depending on intrinsic factors such as sex, health and physiological status, as well as genetics. DNA methylation can also be influenced by extrinsic factors such as sea water temperature, salinity and other environmental cues, while is also subject to stochasticity (Jung and Pfeifer 2015). Piscine epigenetic clocks have now been developed in several fish species (Piferrer and Anastasiadi 2023), including important commercially exploited species, such as the Northern red snapper (*Lutjanus campechanus*) and the red grouper (*Epinephelus morio*) (Weber et al. 2022), the blackbelly rosefish (*Helicolenus dactylopterus*) (Weber, Fields, et al. 2024), the Florida bass (*Micropterus salmoides*) (Weber et al. 2025) and the haddock (*Melanogrammus aeglefinus*) (Strand et al. 2025).

Since the development of epigenetic clocks in fishes is  $<6$  years old, significant knowledge gaps remain. Consequently, the influence of key intrinsic and extrinsic factors on their

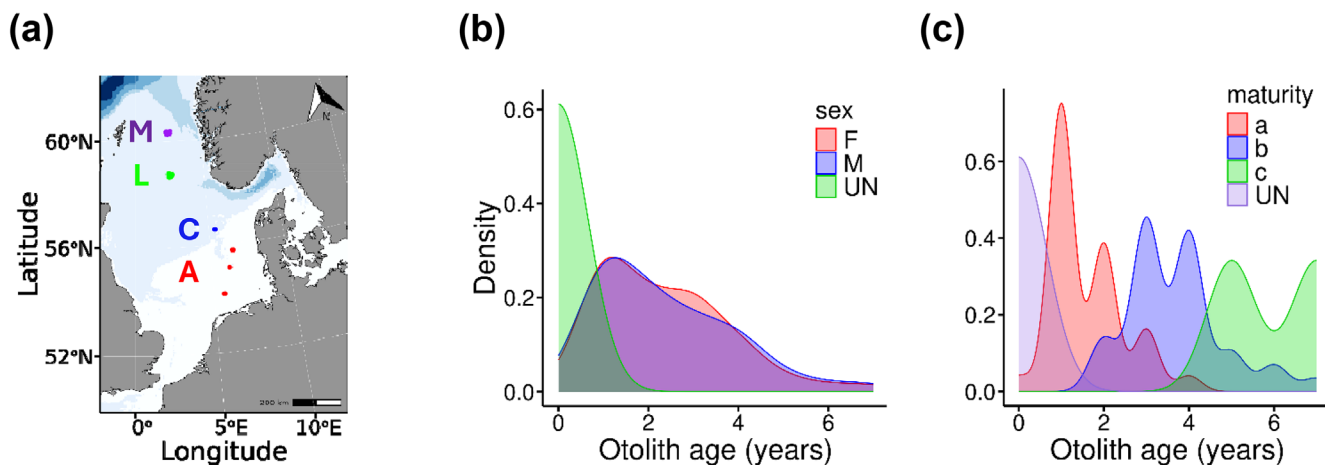
performance is seldom investigated and remains fragmentary (Piferrer and Anastasiadi 2023). For example, in fisheries stock assessments, sex and maturity status are key variables routinely assessed, as they are crucial, along with age, to estimate the reproductive capacity of populations. Only recently, studies have reported the influence of sex (Strand et al. 2025; Weber, Fields, et al. 2024; Weber, Wyffels, et al. 2024), and none has yet examined the influence of maturity status on the performance of epigenetic clocks. Furthermore, another important aspect, especially when surveying species widely distributed across large geographic areas, is whether spatial factors influence DNA methylation patterns. This may be manifested as a ‘latitudinal epigenetic cline’, that is, a gradual change in DNA methylation patterns with latitude, similar to the latitudinal clines observed in other molecular traits, like the clock genes of *Drosophila melanogaster* (Costa et al. 1992). In this regard, it has been observed that extensive latitudinal differences in gene expression in *Drosophila* correlate with cis-regulatory differences between populations, probably associated with epigenetic differences (Juneja et al. 2016). However, to the best of our knowledge, no latitudinal clines in DNA methylation patterns have ever been reported in animals. Spatial distribution, nevertheless, influenced global DNA methylation in panmictic populations of European eel (*Anguilla anguilla*) (Liu et al. 2022). Thus, clines associated with spatial and environmental factors are likely to influence DNA methylation patterns as well. These are all relevant aspects that must be addressed for epigenetic clocks to become practically applicable, as they need to perform consistently well regardless of intrinsic or extrinsic factors.

Here, we used the Atlantic cod to develop an epigenetic clock that accurately predicts age across individuals, regardless of intrinsic and extrinsic factors such as sex, maturity status and geographic origin. To achieve this, we collected individuals from different geographic areas of the North Sea, representing both sexes and various sexual maturity statuses, and aged them through otolith readings. We extracted DNA from fin-clips, routinely sampled for genetic analyses without compromising the commercial aspect of the fish, and processed it using bisulfite restriction-site DNA associated sequencing (bis-RAD-seq) (Trucchi et al. 2016) to obtain DNA methylation values, and with RAD-seq to obtain genetic sequencing data subsequently analysed using bioinformatics and machine learning procedures. Thus, the aims of this study were: (1) to develop an epigenetic clock in the Atlantic cod and (2) to test its performance across sexes, maturity statuses and geographic origins.

## 2 | Materials and Methods

### 2.1 | Sample Collection and Otolith Reading

The samples were collected during research cruise no. 428 of the German fisheries research vessel ‘Walther Herwig III’ during July/August 2019 as part of the International Bottom Trawl Survey (IBTS) in the third quarter. Demersal trawling was carried out according to the IBTS standards (ICES 2020). Individual cod were sampled throughout the survey area in the North Sea (see map in Figure 1a). For each fish, total length, weight, sex



**FIGURE 1** | Extrinsic and intrinsic influences on Atlantic cod samples. (a) Map of areas in the North Sea where individuals were caught. Letters (A, C, L, M) indicate the four areas of origin used in this study (Southern area: A and C; Viking area: L and M). (b) Smoothed (kernel) density estimates for number of individuals in age classes estimated by otoliths separated by sex (F, female; M, male; UN, undetermined sex) and (c) sexual maturity status (a, immature/juvenile; b, maturing; c, spent; UN, undetermined maturity). Year class 0 individuals are sexually undifferentiated and immature and hence were classified as unknowns (UN).

and sexual maturity stage were recorded, otolith samples were collected for age determination, and fin clips were taken for molecular analyses.

The maturity stage was determined by visual inspection in accordance with the scaling system used in the IBTS (ICES 2020), where 'UN' = undetermined maturity, 'a' = immature/juvenile, 'b' = maturing and 'c' = spent.

For manual age readings, both otoliths were embedded in polyester resin and, after hardening, thin-sectioned using diamond-coated high-speed sawing blades with 0.5 mm distance. The growth increments (annuli) visible in the thin sections were counted under a binocular microscope with transmitted light. Age determination was typically based on one otolith per fish, with readings performed by two age readers to reach consensus on the interpretation of growth zones. For challenging growth structures, several alternative reading axes from the nucleus (core zone) to the edge of the otolith section were used. Individuals were assigned to an age class according to the consensus otolith reading. The individuals analysed in this study were selected from a larger sample based on their well-defined growth patterns and the reliability of their determined ages.

Information on the length, weight, sex, maturity status, geographic location, latitude, longitude and age of each of the 110 Atlantic cod samples analysed in this study can be found in Table S1.

## 2.2 | DNA Extraction

Genomic DNA was extracted from fin clips using the DNeasy Blood & Tissue Kit (Qiagen, Germany) according to the manufacturer's protocol, including treatment with RNase. DNA was eluted in 100  $\mu$ L of elution buffer. The quality and quantity of the extracted DNA were assessed using a NanoDrop spectrophotometer and samples were normalised to a final concentration of 20 ng/ $\mu$ L.

## 2.3 | RAD-Sequencing Library Preparation and Sequencing

DNA was used to prepare a single-digested RAD-seq library for genetic analyses. Approximately 1  $\mu$ g of purified genomic DNA per individual was digested at 37°C for 15 min with the restriction endonuclease *SbfI* (CCATG/C recognition site) (NEB). Modified Illumina adapters containing five nucleotides of barcode sequence unique (P1 adapters) to an individual in the library were ligated with T4 DNA ligase (NEB) to enable sample multiplexing. The pooled samples were then sheared using a Q800R3 ultrasonicator (Qsonica) and size-selected to isolate DNA fragments spanning 300–500 bp by agarose gel electrophoresis. The MinElute Gel Purification Kit (Qiagen) was used to purify and recover the DNA from the gel following the manufacturer's instructions. The purified fragments were subsequently blunt-end repaired and A-tailed with the Quick Blunting Kit (NEB) and Klenow Fragment (NEB), respectively. A second adapter was ligated (P2 adapter) to enable selective PCR amplification. The reactions were then purified with a QIAquick column (Qiagen) and amplified via 12 cycles of PCR using the Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB). The amplified library was purified with a QIAquick column (Qiagen) and quantified with a Qubit fluorometer. The final RAD-seq library was sequenced on a HiSeqX platform (Illumina) as 150-bp paired reads, by the company Macrogen (Seoul, South Korea).

## 2.4 | Bisulfite RAD-Sequencing (Bis-RAD-Seq) Libraries Preparation and Sequencing

Bis-RAD-seq libraries for DNA methylation analyses were prepared and sequenced in two batches. In Batch 1, 60 samples of Atlantic cod were processed to develop the bis-RAD-seq protocol and evaluate its suitability for epigenetic clock construction. Batch 2 involved the processing of 60 additional cod samples. Bis-RAD-seq libraries were prepared by Floragenex (Eugene, USA) using an adaptation of the original protocol by Trucchi et al. (2016). Briefly, 200 ng of DNA were digested with the

restriction enzyme *SbfI*-HF (New England Biolabs; R3642L) and ligated to individually barcoded, restriction-site overhang-specific P1 adapters with methylated cytosines, matching the sticky ends created during digestion. The P1 adapters contain forward amplification and Illumina sequencing primer sites, as well as a 10bp nucleotide barcode for sample identification. To minimise erroneous sample assignment due to sequencing errors, barcodes differed by at least two nucleotides. Five individually barcoded samples with unique P1 barcoded adapters were pooled together. The pooled DNA was sheared to an average size of 500bp using a Bioruptor NGS (Diagenode) and cleaned up using MinElute Reaction Cleanup Kit (Qiagen). Fragments between 300 and 500bp were selected on a 1.5% agarose gel and further purified using the MinElute Gel Purification Kit (Qiagen). End repair and dA-tailing were performed using the NEBNext Ultra II End Repair/dA-Tailing Module (NEB). The cytosine methylated P2 adapter, a ‘Y’ adapter with divergent ends that contains a 3’ dT overhang, was ligated to the dA-tailed DNA fragments. The DNA was bisulfite-converted using the EZ DNA Methylation Direct Kit (Zymo Research). High-fidelity PCR amplification was performed on P1 and P2 adapter-ligated, bisulfite-treated ssDNA fragments using Phusion U High-Fidelity (Fisher Scientific). This step added Illumina flow cell annealing sequences, Illumina multiplexing indices and sequencing primer annealing regions to fragments containing both P1 and P2 adapters. Library fragments between 400 and 600bp were size selected on a 1.5% agarose gel and cleaned up. Final libraries were quantified using a fluorometer, and their molarity and fragment size was quantified on an Agilent Bioanalyzer. The batch 1 library was sequenced as paired-end 150bp reads to maximise information from the pilot run, while the batch 2 library was sequenced as single-end 150bp reads, typically sufficient for RAD-seq applications.

## 2.5 | Bioinformatic Analysis of Genetic Data

Raw sequence reads quality was assessed with FastQC (Andrews 2010). Stacks 2.4 (Catchen et al. 2013; Rochette et al. 2019) was used for RAD-seq data processing, SNP (Single Nucleotide Polymorphism) discovery and the calculation of standard population genetic statistics. We used the `process_radtags` program for demultiplexing and filtering with the following parameters: `-e sbfI` (enzyme specification), `-r` (rescue barcodes and RAD-Tags), `-c` (clean: clean data, remove any read with an uncalled base), `-q` (quality: discard reads with low quality scores). Retained reads from all samples were aligned to the *Gadus morhua* reference genome (assembly version gadMor3.0, GCA\_902167405.1, Jul 2019) (Martin et al. 2023) using Bowtie2 and analysed for building loci and calling SNPs using the ‘reference-based’ pipeline. Finally, data were filtered using the ‘populations’ module. We removed PCR duplicates using the ‘`--rm-pcr-duplicates`’ function. Loci were retained only if they were present in all geographical locations, and in at least 80% of samples within each one of them. Only loci with a maximum observed heterozygosity of 0.7 and a minimum allele frequency of 0.05 were kept, to minimise the impact of sequencing errors. In addition, loci with uncorrected Hardy–Weinberg equilibrium (HWE)  $p$ -values  $\leq 0.05$  (global test) were excluded, to avoid erroneous genotypes. To minimise physical linkage, only one SNP within each RAD locus was randomly selected, and SNPs with

more than two alleles were discarded. The ‘populations’ module was also used to calculate population-level summary statistics, including observed and expected heterozygosities, genetic diversity ( $\pi$ ) and  $F$  statistics. Discriminant analysis of principal components (DAPC), implemented in R v.4.4 (R Core Team 2024) using the *adegenet* v.2.1.10 package (Jombart et al. 2010; Jombart and Ahmed 2011), was used to visualise relationships among individuals. The number of principal components (PCs) retained in the analysis was determined using the ‘`find.clusters`’ function. Individual ancestry coefficients were explored using the Sparse Non-Negative Matrix Factorization (sNMF) algorithm implemented in the *lea* v.3.18 package for Landscape and Ecological Associations studies (Frichot and François 2015).

## 2.6 | Bioinformatic Analysis of DNA Methylation Data

Bioinformatics and machine learning analyses were performed as described in Anastasiadi and Piferrer (2023). The quality of demultiplexed raw sequencing reads was assessed using FastQC (Andrews 2010) and MultiQC for summarising reports across multiple samples (Ewels et al. 2016). Quality filtering was performed using `process_radtags` from Stacks (v.2.2) (Catchen et al. 2013; Rochette et al. 2019) with the following parameters: `-e sbfI` (enzyme specification), `-r` (rescue barcodes and RAD-Tags), `-c` (clean data, remove any read with an uncalled base), `-q` (discard reads with low quality scores), `--disable_rad_check` (disable RAD tag structure checks). PCR duplicates, a common artefact in high throughput approaches like bis-RAD-seq, were removed using `clone_filter` from Stacks (v.2.2). Quality control of the filtered data was reassessed using FastQC and MultiQC.

A bisulfite-converted version of the genome version gadMor3.0 was created using Bismark (v.20.0) (Krueger and Andrews 2011). Filtered reads were aligned to the bisulfite converted genome with Bismark, allowing alignments to all possible strands and relaxing the stringency with the setting `--score-min L,0,-0.6`. Methylation values were extracted from aligned reads using the function `bismark_methylation_extractor`, while ignoring the first four bases, since they correspond to the P1 adapter overhang. DNA methylation values in non-CpG context were merged, and a genome-wide report of DNA methylation at all cytosines was produced using the option `-scaffolds`. Custom AWK scripts were used to filter CpG sites with a coverage greater than four and to convert files into a format compatible with downstream analysis. Bisulfite conversion efficiency was calculated as 100 minus the percentage of Cs methylated in a CHH context, where H represents any nucleotide except G.

## 2.7 | DNA Methylation Data Pre-Processing

DNA methylation values for each sample were processed using methylKit (v.1.14.2) (Akalin et al. 2012) in R (v.3.5). Reading files with methylKit requires assignment of samples to one of two groups (control vs. treated), and samples were randomly assigned to one or the other. CpGs covered by fewer than 10 reads or exceeding the 99.9% percentile of coverage distribution were filtered out, and coverage was normalised across samples. Only CpGs present in at least 48 out of 60 samples (80%) within each

group were kept. Individuals with fewer than 100,000 CpGs were identified as outliers and eliminated, resulting in a dataset of 110 individuals. Filtering for coverage, normalisation and retention of common CpGs was repeated with the reduced number of individuals using methylKit. Percentage DNA methylation values were obtained for all retained CpGs.

Downstream analyses were performed in R (v.4.2). Missing values were imputed using the overall mean DNA methylation via the `na.aggregate` function from the `zoo` package (v.1.8-12) (Zeileis and Grothendieck 2005). Overall clustering of samples was evaluated using a Principal Component Analysis (PCA). A batch effect detected in the PCA was confirmed as significant by the function `PCAtest` of the `PCAtest` package (Camargo 2022) (v.0.0.1, Figure S1a). CpGs associated with the batch effect were identified using the `BEclear` package (v.2.18.0) from Bioconductor (Akulenko et al. 2016) and eliminated from further analysis, ensuring no clear clustering after batch correction (Figure S1b).

## 2.8 | Global DNA Methylation Patterns

Global DNA methylation per age class was calculated by averaging the methylation values of all CpGs within each sample, followed by calculating the mean methylation across all samples within each age class.

The exact GPS coordinates of each sample were used to calculate pairwise distances in meters between sampling points using the Geographic Distance Matrix Generator v.1.2.3 (Ersts, n.d.). From this matrix, distance-based Moran's eigenvector maps (dbMEM spatial eigenfunctions) were computed via the function `dbmem` of the `adespatial` package (Dray et al. 2018). These dbMEM eigenfunctions served as explanatory variables in a Redundancy Analysis (RDA), with global DNA methylation values as the response matrix. The significance of the RDA was evaluated using ANOVA with 999 permutations.

## 2.9 | Machine Learning (ML) Model Building

We followed the detailed outline of machine learning (ML) model building described in the Epigenetic Biomarker Discovery Procedure (EDBP) for aquatic organisms, adapted for age prediction, which is a regression problem since age is a continuous numeric variable with values always equal or above zero (Anastasiadi and Beemelmanns 2023). A tutorial with all necessary R code on ML model building in EBDP for age prediction can be found on Codeberg: [https://codeberg.org/dafanast/Epigenetic\\_clock](https://codeberg.org/dafanast/Epigenetic_clock). The Atlantic cod dataset consisted of 110 individuals ( $n_s$ ; number of samples) and 47,464 CpGs with DNA methylation values ( $n_f$ ; number of features; Figure S2). Taking into account age structure, data were split into 75% training and 25% test sets, with the test set reserved for final evaluation. Five pre-processing steps were applied to the training dataset. (1) CpGs with zero or near-zero variance were excluded (remaining  $n_f = 46,359$ ). (2) CpGs correlated between them above a threshold of 0.7 Pearson correlation coefficient were excluded (remaining  $n_f = 35,328$ ). (3) CpGs with DNA methylation above 80% or below 20% in more than 30% of the samples were excluded (remaining  $n_f = 12,874$ ). (4) CpGs correlated with age with an absolute

Pearson correlation coefficient higher than 0.5 and  $p < 0.01$  were maintained (remaining  $n_f = 12,610$ ). (5) The final dataset was centred and scaled using the `preProcess` and `predict` functions of the `caret` package (v.6) (Kuhn 2008), resulting in 86 samples and 12,610 CpG loci present in all samples. The response variable was transformed to the log scale as  $\log(\text{age} + 0.1)$  to deal with samples from age class 0.

Penalised regressions are typically used for epigenetic clock development, since they can handle large multivariate datasets where  $n_f$  is orders of magnitude higher than  $n_s$  and where most variables contribute little to prediction of the response variable. Here, we trialled two types of penalised regression: (1) Least Absolute Shrinkage and Selection Operator (LASSO) regression, where variables are forced to zero and (2) elastic net regression, which allows some variables to be shrunk instead of forcing them to zero. The third type of penalised regression, Ridge regression, shrinks the coefficients of the least contributing variables towards zero but keeps all variables; this was considered unsuitable for our large dataset of  $n_f = 12,610$ .

Models were tuned using the `train` function of `caret` (v.6) over a grid of 100 lambda values, which controls the amount of shrinkage. A 10-fold cross validation (CV) approach was used to evaluate the model performance tuned with the training set. The training set was iteratively split between training and validation sets, models were iteratively fit in the training sets, and prediction errors were averaged across each round. The optimal LASSO and elastic net models with minimised errors were evaluated on the 25% unseen test dataset. The process of model tuning was reiterated until the model performed well in the test dataset, to avoid overfitting and improve generalisation. Model performance was evaluated using loss function metrics: Root Squared Mean Error (RMSE) and Mean Absolute Error (MAE), together with  $R^2$ . The number of features or CpGs selected in the final model was considered an important factor for choosing between LASSO and elastic net, since fewer CpGs facilitate the development of a ready-to-use kit for age prediction for fisheries and conservation programmes. Pearson correlations between age estimated by otoliths versus age estimated by the epigenetic clock were also taken into account. The final epigenetic clock was applied for age prediction across the entire dataset.

Clock CpGs were identified based on non-zero coefficients in the LASSO model. Pearson correlations between otolith age and DNA methylation values at these CpGs were calculated. Their genomic coordinates were queried into the Ensembl database to identify overlapping genes and their associated Gene Ontology (GO) terms. GO-terms enrichment analysis was performed using PANTHER with the *Danio rerio* database. GO-terms enrichment reduction and visualisation were mediated by REVIGO (Supek et al. 2011) and finalised in R.

## 2.10 | Effect of Biological Variables and Geographic Origin on Age Prediction

We evaluated the effects of two biological variables, sex and maturity status, as well as geographic origin, on age prediction, using four approaches. Firstly, we used PCA on global DNA methylation data to visually assess clustering due to the factor of

interest. Secondly, we used PCA on the DNA methylation values of the CpGs of the final epigenetic clock. Thirdly, we compared the MAE of the final epigenetic clock between groups using the Wilcoxon rank test for sex (male/female) and the Kruskal–Wallis test for maturity status (stages a, b, c) and geographic origin (locations A, C, L, M). Finally, we evaluated the correlation between otolith-based age versus epigenetic age within each group separately.

### 3 | Results

#### 3.1 | Development of the Epigenetic Clock for Age Prediction

A selection of 110 Atlantic cod samples was used in this study. Fish ranged from 8.5 to 98.5 cm in total length and from 4 to 8140 g in body weight in accordance with the age distribution. Individuals were captured from four geographic locations across the North Sea named after different countries' fisheries territories: A=Germany; C=Denmark; L and M=Norway. Sites L and M were located to the northeast of the North Sea, with M positioned farther north, towards the northern edge of the North Sea and closer to the northeastern edge of the survey area (Figure 1a). Sampling points at location A were more widely distributed than at other areas. Age distribution, estimated by otolith readings, ranged from 0+ to 7 years old and was similar between males and females, with a skewed distribution towards younger individuals from age classes 1, 2 and 3 (Figure 1b). Sexual maturity status was linked to age, with more individuals classified as maturity status 'a' in age classes 1 and 2 and more individuals classified as maturity status 'c' in age classes 5, 6 and 7 (Figure 1c). Individuals from year class 0+ were undifferentiated (UN), with maturity status undetermined (UN).

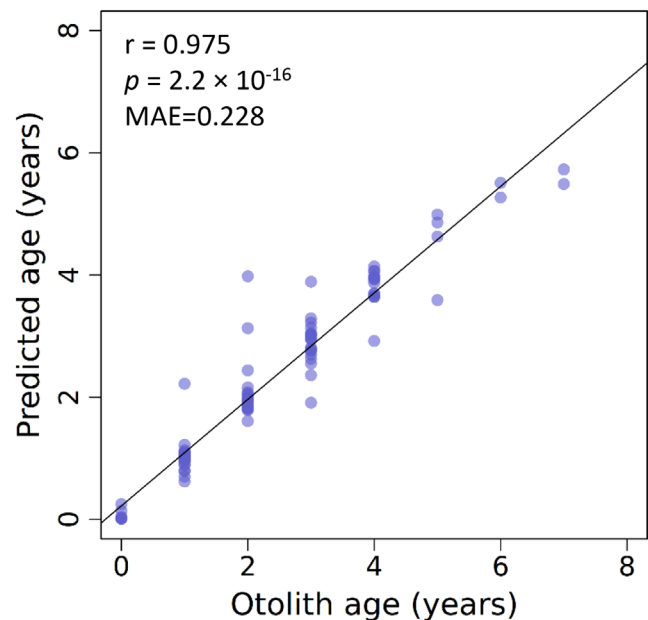
BisRAD-seq libraries were prepared for all 110 samples, yielding a mean of 3,388,756 raw sequences, with a mean alignment rate of 68.59% and a mean bisulfite conversion efficiency of 98.4%. On average, 254,534 raw CpGs per sample were obtained, which was reduced to 185,690 CpGs after filtering (Table S2). A total of 47,464 CpGs were retained for downstream analysis after identifying common CpGs across samples and correcting for batch effects. The mean DNA methylation across this set of CpGs was 77.26% (Table S3). Global DNA methylation distributions across this set of CpGs differed significantly between all age classes, with only one exception (Pairwise comparisons using Wilcoxon rank sum test with continuity correction; Figure S3, Table S3).

A RAD-seq library was processed in parallel using the same samples. We obtained a total of 309,851,627 raw paired-end sequences, of which 81.5% passed our quality filters and were built into stacks. The mean number of retained reads per sample was 3,481,479 (min: 865,510; max: 8,083,086), with a mean depth of coverage per sample of 19.5× (min: 5.5×; max: 51.0×). The final catalogue consisted of 512,457 loci. We excluded 21 samples that had a depth of coverage <10× or that had more than 5% missing data. From the 89 individual samples remaining in the dataset, 38,103 loci composed of 18,640,333 sites were variable biallelic SNPs that passed population (present in all sites), sample (present in at least 80% of samples) and remaining filtering constraints. These SNPs constitute the final dataset used in

subsequent analyses (see within-population summary statistics considering variant sites only in Table S4, and all variants in Table S5).

The ML model building in Epigenetic Biomarker Discovery Procedure in aquatic organisms (Anastasiadi and Beemelmanns 2023) was used in our dataset derived from the BisRAD-seq libraries. All individuals were used for model construction independently of their internal and external variables and split between training ( $n=86$ ) and test ( $n=24$ ) datasets while accounting for age structure. We evaluated two penalised regression models, LASSO and elastic net, which performed similarly on the training and test datasets, but elastic net produced lower median MAE and RMSE and higher median  $R^2$  calculated from 10-fold CV with the training dataset (Figure S4 and Table S6). However, LASSO was selected in the final model since it used a lower number of CpGs ( $n=73$ ) than elastic net ( $n=207$ ). This parameter is crucial for future implementation of a ready-to-use kit in fisheries and conservation applications. Log-transforming age improved epigenetic clock performance and was, thus, used in model creation.

The final epigenetic clock for age prediction achieved an accuracy of 97.5% and a precision of 0.228 years, using DNA methylation values of 73 CpGs (Figure 2). The correlation between otolith-based age and age predicted via the epigenetic clock was significant (Pearson correlation coefficient  $r=0.975$ ,  $p<2.2 \times 10^{-16}$ ). The clock not only performed well on the training dataset ( $r=0.998$ ,  $p<2.2 \times 10^{-16}$ , MAE=0.117; Figure S5a) but also performed well on the test dataset, which was held out since the beginning of the ML building and never seen by the model ( $r=0.914$ ,  $p<2.2 \times 10^{-16}$ , MAE=0.635; Figure S5b), showing its ability to generalise.



**FIGURE 2** | Epigenetic clock for age prediction in Atlantic cod populations. Predicted age based on the DNA methylation of 73 CpGs by Least Absolute Shrinkage and Selection Operator (LASSO) correlates to otolith-based age. Accuracy is measured by Pearson's correlation coefficient ( $r$ ) and its associated significance level ( $p$ ). Precision is measured by the Mean Absolute Error (MAE, in years) between predicted and true (otolith) age.

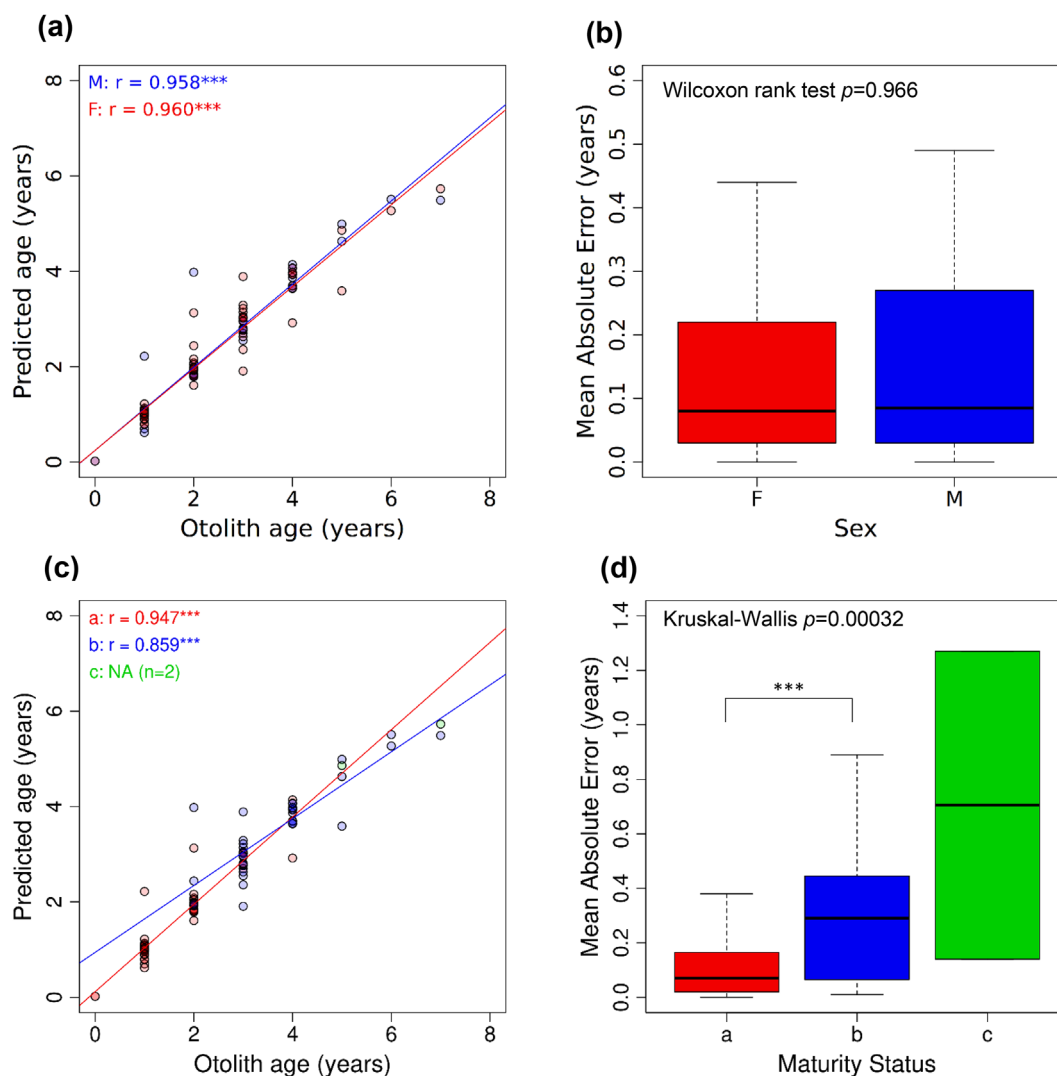
### 3.2 | Internal and External Influences on Age Prediction

DNA methylation and the performance of age prediction were similar between males and females, both for global DNA methylation (Figure S6a) and with the DNA methylation values of the clock CpGs only (Figure S6b). When the epigenetic clock was applied for age prediction separately by sex, accuracy was equal for females ( $r=0.960$ ,  $p < 2.2 \times 10^{-16}$ ; Figure 3a) and males ( $r=0.958$ ,  $p < 2.2 \times 10^{-16}$ ). Precision of age prediction between sexes (mean MAE in females=0.24 and males=0.228, median MAE in females=0.080 and males=0.085) was also equal (Wilcoxon rank sum test  $p=0.966$ ; Figure 3b).

Sexual maturity status was associated with differences in DNA methylation and age prediction. As visually assessed by

PCAs, both DNA methylation across CpG sites of the genome (Figure S7a) and at the clock CpGs (Figure S7b) showed clustering trends based on maturity status. When the epigenetic clock was applied for age prediction separately by maturity status, accuracy was high ( $r=0.947$ ,  $p < 2.2 \times 10^{-16}$ ) for individuals with maturity status 'a' and lower ( $r=0.859$ ) but still strong and statistically significant ( $p=3.9 \times 10^{-11}$ ) for individuals with maturity status 'b' (Figure 3c). Precision differed significantly between maturity status 'a' and 'b' (Dunn's test of multiple comparisons, Bonferroni  $p$ -adjusted=0.000482 following Kruskal–Wallis rank sum test  $p=0.00032$ ; Figure 3d). The low sample size for individuals with maturity status 'c' ( $n=2$ ) precluded meaningful statistical analysis.

Geographic origin of the Atlantic cod individuals influenced DNA methylation patterns, and the epigenetic clock performed



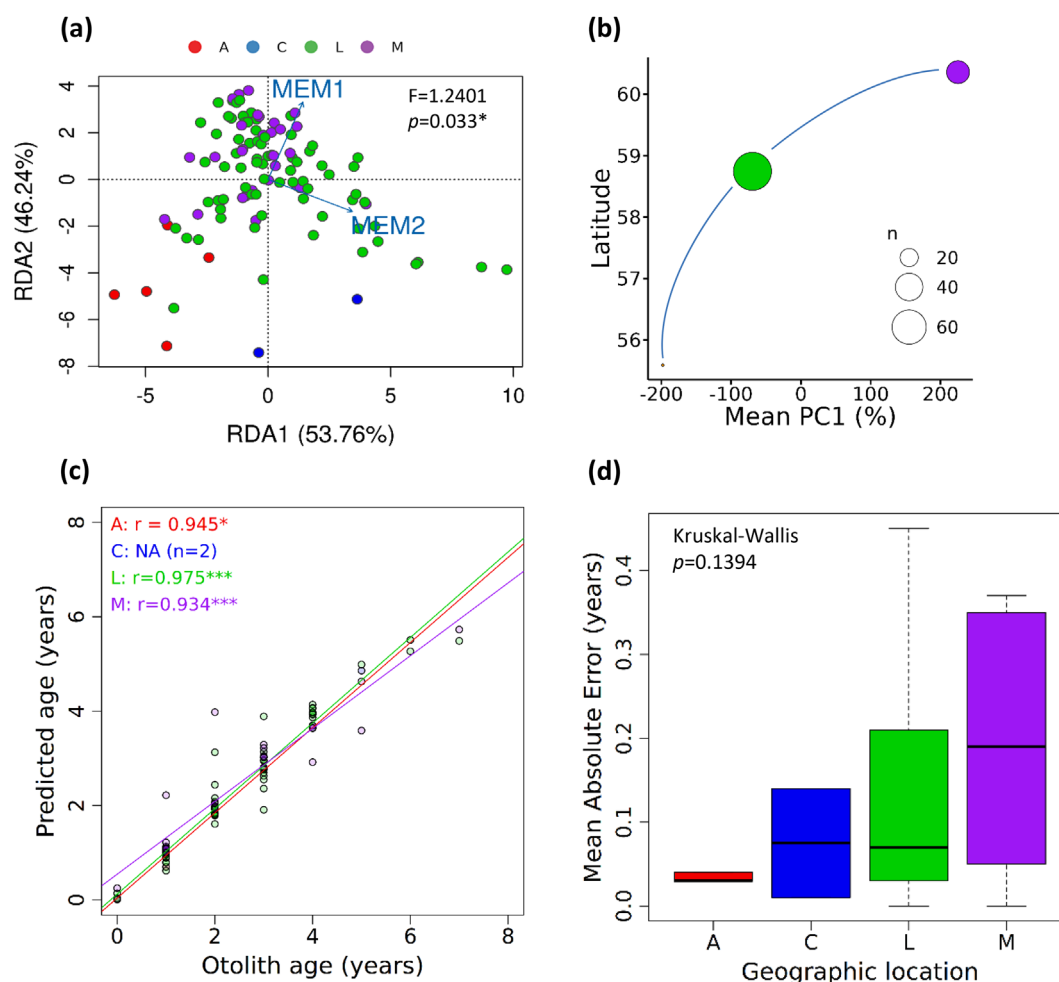
**FIGURE 3** | Influence of sex and sexual maturity status on the performance of the epigenetic clock. Predicted age based on the DNA methylation of 73 clock CpGs by Least Absolute Shrinkage and Selection Operator (LASSO) correlates to the otolith-based age separately per sex (a) and sexual maturity status (c). Pearson's correlation coefficient ( $r$ ) and associated significance level ( $p$ , with  $p < 0.001$  denoted by  $^{***}$ ) are calculated separately per sex (a) and sexual maturity status (c). Differences in precision between sexes were evaluated by Wilcoxon rank test of the Mean Absolute Error (MAE) per sex (b). Points, lines and boxes are coloured according to sex either as females (red), males (blue) or undifferentiated (UN) individuals (green). Differences in precision between sexual maturity statuses were evaluated by Kruskal–Wallis rank test of the MAE per sexual maturity status (d). Points, lines and boxes are coloured according to sexual maturity status (a, immature/juvenile, red; b, maturing, blue; c, spent, green). Maturity status 'c' contained only two individuals and statistical analyses did not include this group.

well across all locations. DNA methylation was generally similar across all four geographic locations, as shown by PCA of both global DNA methylation (Figure S8a) and DNA methylation values of the clock CpGs (Figure S8b). However, spatial factors explained 2.27% of the total variance of global DNA methylation, and this effect was significant ( $F=1.2401$ ,  $p=0.033$ , Figure 4a). A latitudinal trend in DNA methylation was observed when we focused on the mean PC1 axis per geographic location of the global DNA methylation (Figure 4b). Accuracy of the epigenetic clock was high (above 93%; Figure 4c) in geographic location A ( $r=0.945$ ,  $p=0.01552$ ), L ( $r=0.975$ ,  $p<2.2 \times 10^{-16}$ ) and M ( $r=0.934$ ,  $p=1.5 \times 10^{-13}$ ). Only two samples were available from location C, preventing meaningful statistical testing. No significant differences in clock precision were observed across geographic locations (Kruskal–Wallis rank sum test  $p=0.1394$ ; Figure 4d).

Regarding population genetic structure derived from the RAD-seq data, we found that discriminant analysis of principal components (DAPC) analysis with all loci revealed a unique genetic cluster for all individuals independent of their geographic location (Figure S9). Clustering of samples by Hardy–Weinberg equilibrium (HWE) with sNMF revealed that a value of  $K=1$  was the most likely configuration, according to the cross-entropy criterion. Exploration of the individual ancestry coefficients for  $K=3$  revealed no support for differentiation among geographic locations, in agreement with DAPC results (Figure S10).

### 3.3 | Clock CpGs and Associated Genes

DNA methylation at the 73 CpGs included in the final epigenetic clock was associated with age. Of these 73 CpGs, 33 showed



**FIGURE 4** | Influence of geographic origin on the performance of the epigenetic clock. Redundancy analysis of the effect of spatial factors (MEM1 and MEM2) on the DNA methylation values across the genome (a;  $n=47,464$ ) tested for significance using ANOVA with 999 permutations. Mean coordinates of Principal Component 1 (PC1) of the DNA methylation values across the genome were calculated per geographic location and are shown in relationship to the latitude (b). Size of bubbles is proportional to the number of individuals found in each geographic area. Predicted age based on the DNA methylation of 73 clock CpGs by Least Absolute Shrinkage and Selection Operator (LASSO) is regressed to the otolith-based age separately per geographic origin (c). Pearson's correlation coefficient ( $r$ ) and associated significance level ( $p$ , with  $p<0.001$  denoted by  $^{***}$  and  $p<0.05$  denoted by  $^*$ ) are calculated separately per geographic origin. Geographic area 'C' contained only two individuals and could not be included in the statistical analyses. Differences in precision between individuals of different geographic origin were evaluated by Kruskal–Wallis rank test of the Mean Absolute Error (MAE) per geographic origin (d). Points, lines and boxes are coloured according to geographic origin as 'A' (red), 'C' (blue), 'L' (green) or 'M' (purple).

a positive (Figure S11) and 40 a negative association between their DNA methylation and age (Figure S12). These associations were significant (Pearson's correlation  $p < 0.05$ ) for the majority of the clock CpGs (67/73; Table S7). The clock CpGs overlapped with 54 unique genes (Table S8). Among these, 21 genes had annotated names and were assigned GO terms, enabling GO-terms enrichment analysis. Genes associated with the clock CpGs showed significant enrichment for biological processes, including multicellular organism development, developmental processes, rhythmic processes, cell population differentiation and N-terminal protein amino acid methylation (Figure S13).

#### 4 | Discussion

In this study, we constructed a highly accurate and precise epigenetic clock for age prediction in the Atlantic cod, a species of significant importance to global fisheries. The epigenetic clock predicted accurately the age of samples unseen during model development, confirming its potential applicability for real-world usage. Importantly, the epigenetic clock performed consistently across sexes and showed only slightly lower accuracy in individuals of advanced sexual maturity, reflecting a bias towards younger age classes. Additionally, the epigenetic clock performance was similar across individuals of different geographic origins, despite a significant influence of spatial distribution on global DNA methylation patterns. These results underscore the robustness of our epigenetic clock under a variety of intrinsic and extrinsic conditions known to be associated with global DNA methylation patterns. Worldwide, every year tens of thousands of Atlantic cod individuals are aged via otoliths, and age is used in stock assessment models underlying management decisions across the world. The precise age determination achieved in this study represents a significant milestone that greatly enhances essential scientific data for stock assessments, which are crucial for effective fisheries management.

Traditional age determination in Atlantic cod, as in other fisheries, relies heavily on otolith readings. The process is time-consuming, especially given the large and relatively opaque otoliths of this demersal species. Even advanced approaches, such as automatic image processing and machine learning, achieve a maximum accuracy of 78.6% (Moen et al. 2023). Discrepancies in otolith readings are not uncommon, with a calibration experiment for North Sea cod reporting only 74% agreement among readers and a coefficient of variation (CV) of 39.8% being reached (ICES 2008). Using deep learning, Cayetano et al. (2024) were able to improve the CV down to 7.4%, but some uncertainty in age estimates remained, with a maximum agreement between traditional and artificial intelligence (AI) age readings reaching only 55%. In absolute values, the error varied by age group, with lower variation ( $\pm 1$  year) in younger individuals and higher variation ( $\pm 2$ – $3$  years) in older individuals. Our epigenetic clock relied on otolith age estimates for ground-truthing with inherent perpetuation of otolith age estimation inaccuracies, which could only be circumvented with the use of validated age based on strong cohorts, tagging or microchemistry. However, the individuals analysed in this study were chosen based on well-defined growth patterns and a high reliability of

their otolith-based determined ages. Moreover, after ground-truthing, the epigenetic clock showed an accuracy of 97.5%, significantly outperforming the best otolith-based accuracy of 78.6%, and precision of 2.8 months, which is less than a quarter of the minimum variation of 1 year observed for younger individuals in otolith-based methods. Thus, age prediction based on the epigenetic clock clearly outperforms traditional otolith age estimates in terms of both accuracy and precision.

In recent years, piscine epigenetic clocks have been developed for various species, including commercially exploited species (Strand et al. 2025; Weber et al. 2022; Weber, Fields, et al. 2024; Weber et al. 2025), species of conservation interest (Weber, Wyffels, et al. 2024), as well as model species (Anastasiadi and Piferrer 2020; Bertucci et al. 2021). The error of age estimation in the Atlantic cod epigenetic clock (2.8 months) corresponds to only 0.93% of the maximum species lifespan, which is estimated to be 25 years in the wild (Froese and Pauly 2010). Here, following suggested guidelines (Anastasiadi and Piferrer 2023; Piferrer and Anastasiadi 2023), we were able to build our model with a sample size of  $n = 110$ , that is, midway between the minimum ( $n = 72$ ) and the optimal ( $n = 134$ ) number of samples necessary to build a reliable epigenetic clock, as assessed by simulations (Mayne et al. 2021). In addition, the 73 CpGs of the Atlantic cod epigenetic clock fall within the typical range (26–199) of CpGs numbers used in piscine epigenetic clocks developed to date (Piferrer and Anastasiadi 2023). Interestingly, functional enrichment analysis of the genes associated with these 73 clock CpGs showed significant involvement in developmental and rhythmic processes, a finding consistent with mammalian studies involving 128 species (Lu et al. 2023). This design and methodology allowed us to construct a robust epigenetic clock for our study system. In addition, we provide a detailed tutorial in R with all necessary code, designed not only to replicate our analyses but also to facilitate the development of new epigenetic clocks.

Intrinsic factors are known to influence DNA methylation patterns. For instance, sex differences in DNA methylation are well known in vertebrates (Grant et al. 2022; Piferrer 2021; Piferrer et al. 2024). Also, sexual dimorphism, where one sex grows faster than the other, occurs in many important fisheries species (Parker 1992; Piferrer and Anastasiadi 2023; Beato et al. 2026). Thus, accounting for sex differences is important when developing epigenetic clocks. So far, it has been explicitly considered in the case of European sea bass (*Dicentrarchus labrax*) using reproductive tissues (Anastasiadi and Piferrer 2020) and, more recently, in the blackbelly rosefish (Weber, Fields, et al. 2024) and the cownose ray (Weber, Wyffels, et al. 2024), where clock performance improved with sex-specific models, and the haddock (Strand et al. 2025). These findings suggest that the same panel of CpGs biomarkers can simultaneously serve as tools for predicting both sex and age. However, from a practical perspective, sex-independent epigenetic clocks may be easier to implement, particularly in fisheries management, as their performance remains consistent regardless of the sex of the captured specimens. This is exemplified by our epigenetic clock for Atlantic cod, which achieves high accuracy and precision in age prediction despite large sex-specific differences in growth rate and maximum size (Beacham 1983).

Early developmental processes and time-related events occurring later in life are associated with DNA methylation patterning (Moore et al. 2020). These include key life history events, such as reproductive maturity and reproductive senescence, which have been suggested to explain the discordance between epigenetic age and chronological age (Parrott and Bertucci 2019), well known in mammals. Reproductive maturity in fishes is routinely determined during sampling and provides essential information used by stock assessment models. Our epigenetic clock performed better when applied to sexually immature than to maturing individuals. Sexual maturity status is closely aligned with chronological ageing, as sexually immature individuals are typically younger than sexually mature individuals, although overlapping between both groups occurs frequently. Thus, this difference in performance based on sexual maturity status probably reflects a bias towards younger individuals. This in turn probably results from a combination of the lower error of otolith-based age estimates in younger individuals and the well-documented bias in epigenetic clocks towards younger individuals (El Khoury et al. 2019; Simpson and Chandra 2021). This bias may be related to differences in the tick rate of the clock. Examples include the exponential rate of change until around 20 years of age, transitioning to a linear rate thereafter, of the original human clock (Horvath 2013) and the faster rate of change in female humans undergoing puberty (Binder et al. 2018), which can be explained by stage-dependent differences in growth and cell division. In addition to these biological factors, technical properties of model building, including the choice of machine learning models (penalised regressions) and the under-representation of older individuals in calibration datasets, may also be responsible or at least contributing to the bias (Simpson and Chandra 2021).

Extrinsic influences on DNA methylation, mainly environmental factors because of their ability to induce epigenetic changes, have been extensively studied in a laboratory setting. However, studies of DNA methylation patterns in natural animal populations are still limited, especially in fishes, despite their importance for understanding the contribution of epigenetics to adaptation (Anastasiadi et al. 2021). In the panmictic European eel (*Anguilla anguilla*), not only environmental factors but also spatial distribution had a significant effect on global DNA methylation variation (Liu et al. 2022). In the Atlantic salmon, epigenetic divergence between rivers was apparent in the absence of detectable genetic differentiation based on a low resolution technique (Morán and Pérez-Figueroa 2011). Similarly, in Atlantic cod, our study found a significant influence of spatial distribution on global DNA methylation for populations distributed across the North Sea. Current population structure information of cod in the North Sea identifies two genetic clusters (Viking and Dogger) and uses three sub-stocks (Viking, Northwestern and Southern) within the stock assessment framework (ICES 2023). In our dataset, we probably sampled individuals from both the Viking (areas L and M) and the Southern (areas A and C) regions. However, no genetic clustering was detected, possibly because of the lower number of individuals from the Southern areas, or mixing of individuals, which is supported by the proximity of areas C and L to the boundaries between sub-stocks. Despite the absence of genetic differentiation, geographic origin explained 2.27% of the variance in DNA methylation, a value

comparable to the 2.98% of variance attributed to geographic location in European eel populations (Liu et al. 2022). This spatial effect may reflect a latitudinal cline in DNA methylation. If confirmed, this would be the first such evidence in any animal species. Of note, despite the significant effect of spatial distribution, our epigenetic clock was able to predict age with similar accuracy and precision across all geographic areas, irrespective of underlying genetic clustering. This finding underscores the generalizability of our epigenetic clock, making it a valuable tool for age prediction across Atlantic cod populations. Importantly, it demonstrates that the development of population- or area-specific clocks is not necessary, which is a major contribution towards the use of piscine epigenetic clocks in the real world.

We addressed a crucial and previously unresolved question regarding the generalizability of epigenetic clocks across a species' distribution, using Atlantic cod as a model. Our findings demonstrate that the epigenetic clock developed in this study can be widely applied across populations from different geographic areas across a species' distribution, with accurate and precise performance. Consequently, our tool constitutes a major resource for practical applications, with the potential to aid the sustainable management of Atlantic cod populations. Furthermore, the methodologies and findings presented here pave the way for the development of epigenetic clocks in other fish species. Such clocks could perform well irrespective of intrinsic and extrinsic factors, across the whole geographic distribution, supporting effective management and conservation efforts on a broader scale.

#### Author Contributions

D.A. contributed to implement the bis-RAD-seq protocol, performed all bioinformatic and statistical analyses of the bis-RAD-seq data, produced all figures and wrote the paper. Y.K. extracted DNA prior to normalisation. C.S. planned fish sampling, collected the fish and aged them by reading otoliths. L.C. provided and analysed RAD-seq data. E.E. organised and contributed to fish sampling, including tissue and otoliths, and supervised and contributed to DNA extraction. F.P. contributed to research design, obtained funding, contributed to paper writing and supervised the whole project. All authors contributed to and approved the final version of the paper.

#### Acknowledgements

We thank Tina Blancke and Benita Möckel for their support at fish sampling during the WH428 cruise and Sílvia Joly for DNA sample equalisation. Research was funded through EU Contract—EASME/EMFF/2017/1.3.2.10/SI2.790889 to F.P. and C.S. The sampling for this study on the research surveys at sea was conducted under the Data Collection Framework (DCF) and financially supported by the European Maritime, Fisheries and Aquaculture Fund of the European Union (EMFAF). Funding was also received from the Spanish government through the 'Severo Ochoa Centre of Excellence' accreditation (CEX2019-000928-S). Open access publishing facilitated by Bioeconomy Science Institute, as part of the Wiley - Bioeconomy Science Institute agreement via the Council of Australasian University Librarians.

#### Funding

This work was supported by Agencia Estatal de Investigación (CEX2019-000928-S) and the European Union through the Service Contract (EASME/EMFF/2017/1.3.2.10/SI2.790889), signed with the European Climate, Infrastructure and Environment Executive Agency (CINEA).

## Disclosure

Benefits Generated: A research collaboration was developed with scientists from the countries that provided genetic samples, and all collaborators are included as coauthors. Results of this research have been shared with the provider communities and the broader scientific community, and the research addresses a priority concern of conservation of Atlantic cod populations.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

All code is available on Codeberg (EU-hosted version control service for open-source projects): [https://codeberg.org/dafanast/Epigenetic\\_clock](https://codeberg.org/dafanast/Epigenetic_clock). DNA methylation data from all samples used in this study are available on Dryad: <https://doi.org/10.5061/dryad.tmpg4f565>. Raw RADseq data are publicly available in Genbank under the accession number PRJNA1422464.

## References

- Akalin, A., M. Kormaksson, S. Li, et al. 2012. "methylKit: A Comprehensive R Package for the Analysis of Genome-Wide DNA Methylation Profiles." *Genome Biology* 13, no. 10: R87. <https://doi.org/10.1186/gb-2012-13-10-r87>.
- Akulenko, R., M. Merl, and V. Helms. 2016. "BEclear: Batch Effect Detection and Adjustment in DNA Methylation Data." *PLoS One* 11, no. 8: e0159921. <https://doi.org/10.1371/journal.pone.0159921>.
- Anastasiadi, D., and A. Beemelmans. 2023. "Development of Epigenetic Biomarkers in Aquatic Organisms." In *Epigenetics in Aquaculture*, edited by F. Piferrer and H. Wang, First ed. Wiley-Blackwell.
- Anastasiadi, D., and F. Piferrer. 2020. "A Clockwork Fish: Age Prediction Using DNA Methylation-Based Biomarkers in the European Seabass." *Molecular Ecology Resources* 20, no. 2: 387–397. <https://doi.org/10.1111/1755-0998.13111>.
- Anastasiadi, D., and F. Piferrer. 2023. "Bioinformatic Analysis for Age Prediction Using Epigenetic Clocks: Application to Fisheries Management and Conservation Biology." *Frontiers in Marine Science* 10: 1096909. <https://doi.org/10.3389/fmars.2023.1096909>.
- Anastasiadi, D., C. J. Venney, L. Bernatchez, and M. Wellenreuther. 2021. "Epigenetic Inheritance and Reproductive Mode in Plants and Animals." *Trends in Ecology & Evolution* 36, no. 12: 1124–1140. <https://doi.org/10.1016/j.tree.2021.08.006>.
- Andrews, S. 2010. "FastQC: A Quality Control Tool for High Throughput Sequence Data." Java.
- Beacham, T. 1983. "Growth and Maturity of Atlantic Cod (*Gadus morhua*) in the Southern Gulf of St. Lawrence." <https://doi.org/10.13140/RG.2.1.4344.7528>.
- Beato, S., G. Ecker-Eckhofen, C. Shao, and F. Piferrer. 2026. "Genetic, Endocrine and Epigenetic Mechanisms Underlying Sexual Size Dimorphism in Fish." *Sexual Development*. <https://doi.org/10.1159/000550574>.
- Bertucci, E. M., M. W. Mason, O. E. Rhodes, and B. B. Parrott. 2021. "Exposure to Ionizing Radiation Disrupts Normal Epigenetic Aging in Japanese Medaka." *Aging* 13, no. 19: 22752–22771. <https://doi.org/10.18632/aging.203624>.
- Binder, A. M., C. Corvalan, V. Mericq, et al. 2018. "Faster Ticking Rate of the Epigenetic Clock Is Associated With Faster Pubertal Development in Girls." *Epigenetics* 13, no. 1: 85–94. <https://doi.org/10.1080/15592294.2017.1414127>.
- Bocklandt, S., W. Lin, M. E. Sehl, et al. 2011. "Epigenetic Predictor of Age." *PLoS One* 6, no. 6: e14821. <https://doi.org/10.1371/journal.pone.0014821>.
- Camargo, A. 2022. "PCAtest: Testing the Statistical Significance of Principal Component Analysis in R." *PeerJ* 10: e12967. <https://doi.org/10.7717/peerj.12967>.
- Campana, S. E. 2001. "Accuracy, Precision and Quality Control in Age Determination, Including a Review of the Use and Abuse of Age Validation Methods." *Journal of Fish Biology* 59, no. 2: 197–242. <https://doi.org/10.1111/j.1095-8649.2001.tb00127.x>.
- Casas, L., and F. Saborido-Rey. 2023. "A Review of an Emerging Tool to Estimate Population Parameters: The Close-Kin Mark-Recapture Method." *Frontiers in Marine Science* 10: 1087027. <https://doi.org/10.3389/fmars.2023.1087027>.
- Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. "Stacks: An Analysis Tool Set for Population Genomics." *Molecular Ecology* 22, no. 11: 3124–3140. <https://doi.org/10.1111/mec.12354>.
- Cayetano, A., C. Stransky, A. Birk, and T. Brey. 2024. "Fish Age Reading Using Deep Learning Methods for Object-Detection and Segmentation." *ICES Journal of Marine Science* 81, no. 4: 687–700. <https://doi.org/10.1093/icesjms/fsae020>.
- Costa, R., A. A. Peixoto, G. Barbujani, and C. P. Kyriacou. 1992. "A Latitudinal Cline in a Drosophila Clock Gene." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 250, no. 1327: 43–49. <https://doi.org/10.1098/rspb.1992.0128>.
- De Paoli-Iseppi, R., B. E. Deagle, C. R. McMahon, M. A. Hindell, J. L. Dickinson, and S. N. Jarman. 2017. "Measuring Animal Age With DNA Methylation: From Humans to Wild Animals." *Frontiers in Genetics* 8: 106. <https://doi.org/10.3389/fgene.2017.00106>.
- Dray, S., G. Blanchet, D. Borcard, et al. 2018. "Package 'adespatial.'" R Package, 2018, 3–8.
- El Khoury, L. Y., T. Gorrie-Stone, M. Smart, et al. 2019. "Systematic Underestimation of the Epigenetic Clock and Age Acceleration in Older Subjects." *Genome Biology* 20: 1–10. <https://doi.org/10.1186/s13059-019-1810-4>.
- Ersts, P. J. n.d. "Geographic Distance Matrix Generator (Version 1.2.3) [Computer Software]." American Museum of Natural History, Center for Biodiversity and Conservation. Accessed June 11, 2024. [http://biodiversityinformatics.amnh.org/open\\_source/gdmg](http://biodiversityinformatics.amnh.org/open_source/gdmg).
- Ewels, P., M. Magnusson, S. Lundin, and M. Källér. 2016. "MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report." *Bioinformatics* 32, no. 19: 3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>.
- Fairfield, E. A., D. S. Richardson, C. L. Daniels, C. L. Butler, E. Bell, and M. I. Taylor. 2021. "Ageing European Lobsters (*Homarus gammarus*) Using DNA Methylation of Evolutionarily Conserved Ribosomal DNA." *Evolutionary Applications* 14, no. 9: 2305–2318. <https://doi.org/10.1111/eva.13296>.
- Frichot, E., and O. François. 2015. "LEA: An R Package for Landscape and Ecological Association Studies." *Methods in Ecology and Evolution* 6, no. 8: 925–929. <https://doi.org/10.1111/2041-210X.12382>.
- Froese, R., and D. Pauly. 2010. "FishBase." World Wide Web Electronic Publication.
- Gardner, S. T., E. M. Bertucci, R. Sutton, A. Horcher, D. Aubrey, and B. B. Parrott. 2023. "Development of DNA Methylation-Based Epigenetic Age Predictors in Loblolly Pine (*Pinus taeda*)." *Molecular Ecology Resources* 23, no. 1: 131–144. <https://doi.org/10.1111/1755-0998.13698>.
- Gebremedhin, S., S. Bruneel, A. Getahun, W. Anteneh, and P. Goethals. 2021. "Scientific Methods to Understand Fish Population Dynamics and Support Sustainable Fisheries Management." *Water* 13, no. 4: 4. <https://doi.org/10.3390/w13040574>.
- Grant, O. A., Y. Wang, M. Kumari, N. R. Zabet, and L. Schalkwyk. 2022. "Characterising Sex Differences of Autosomal DNA Methylation in Whole Blood Using the Illumina EPIC Array." *Clinical Epigenetics* 14, no. 1: 62. <https://doi.org/10.1186/s13148-022-01279-7>.

- Helser, T. E., I. M. Benson, and B. K. Barnett. 2019. *Proceedings of the Research Workshop on the Rapid Estimation of Fish Age Using Fourier Transform Near Infrared Spectroscopy (FT-NIRS)*. AFSC Processed Rmep. Alaska Fisheries Science Center, NOAA's National Marine Fisheries Service, Seattle WA. 2019-06, 195.
- Hernandez, K. M., K. B. O'Neill, E. K. Bors, et al. 2023. "Using Epigenetic Clocks to Investigate Changes in the Age Structure of Critically Endangered Māui Dolphins." *Ecology and Evolution* 13, no. 10: e10562. <https://doi.org/10.1002/ece3.10562>.
- Heydenrych, M. J., B. J. Saunders, M. Bunce, and S. N. Jarman. 2021. "Epigenetic Measurement of Key Vertebrate Population Biology Parameters." *Frontiers in Ecology and Evolution* 9: 617376. <https://doi.org/10.3389/fevo.2021.617376>.
- Heyn, H., N. Li, H. J. Ferreira, et al. 2012. "Distinct DNA Methylomes of Newborns and Centenarians." *Proceedings of the National Academy of Sciences of the United States of America* 109, no. 26: 10522–10527. <https://doi.org/10.1073/pnas.1120658109>.
- Horvath, S. 2013. "DNA Methylation Age of Human Tissues and Cell Types." *Genome Biology* 14, no. 10: R115. <https://doi.org/10.1186/gb-2013-14-10-r115>.
- Horvath, S., A. Haghani, J. A. Zoller, et al. 2023. "Pan-Primate Studies of Age and Sex." *GeroScience* 45, no. 6: 3187–3209. <https://doi.org/10.1007/s11357-023-00878-3>.
- ICES. 2008. "Report of the Workshop on Age Reading of North Sea Cod (WKARNSC) [Report]." ICES Expert Group Reports (Until 2018). <https://doi.org/10.17895/ices.pub.19280396.v1>.
- ICES. 2020. "Manual for the North Sea International Bottom Trawl Surveys." Series of ICES Survey Protocols SISP 10-IBTS 10, SISP 10-IB, 106.
- ICES. 2023. "Benchmark Workshop on Northern Shelf Cod Stocks (WKBCOD) [Report]." ICES Scientific Reports. <https://doi.org/10.17895/ices.pub.22591423.v1>.
- Jasinska, A. J., A. Haghani, J. A. Zoller, et al. 2022. "Epigenetic Clock and Methylation Studies in Vervet Monkeys." *GeroScience* 44, no. 2: 699–717. <https://doi.org/10.1007/s11357-021-00466-3>.
- Jombart, T., and I. Ahmed. 2011. "Adegenet 1.3-1: New Tools for the Analysis of Genome-Wide SNP Data." *Bioinformatics* 27, no. 21: 3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>.
- Jombart, T., S. Devillard, and F. Balloux. 2010. "Discriminant Analysis of Principal Components: A New Method for the Analysis of Genetically Structured Populations." *BMC Genetics* 11, no. 1: 94. <https://doi.org/10.1186/1471-2156-11-94>.
- Jones, M. J., S. J. Goodman, and M. S. Kobor. 2015. "DNA Methylation and Healthy Human Aging." *Aging Cell* 14, no. 6: 924–932. <https://doi.org/10.1111/accel.12349>.
- Juneja, P., A. Quinn, and F. M. Jiggins. 2016. "Latitudinal Clines in Gene Expression and Cis-Regulatory Element Variation in *Drosophila melanogaster*." *BMC Genomics* 17, no. 1: 981. <https://doi.org/10.1186/s12864-016-3333-7>.
- Jung, M., and G. P. Pfeifer. 2015. "Aging and DNA Methylation." *BMC Biology* 13: 7. <https://doi.org/10.1186/s12915-015-0118-4>.
- Krueger, F., and S. R. Andrews. 2011. "Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications." *Bioinformatics* 27, no. 11: 1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>.
- Kuhn, M. 2008. "Building Predictive Models in R Using the Caret Package." *Journal of Statistical Software* 28, no. 5: 5. <https://doi.org/10.18637/jss.v028.i05>.
- Liu, S., A. N. B. Tengstedt, M. W. Jacobsen, et al. 2022. "Genome-Wide Methylation in the Panmictic European Eel (*Anguilla anguilla*)." *Molecular Ecology* 31, no. 16: 4286–4306. <https://doi.org/10.1111/mec.16586>.
- Lu, A. T., Z. Fei, A. Haghani, et al. 2023. "Universal DNA Methylation Age Across Mammalian Tissues." *Nature Aging* 3, no. 9: 1144–1166. <https://doi.org/10.1038/s43587-023-00462-6>.
- Martin, F. J., M. R. Amode, A. Aneja, et al. 2023. "Ensembl 2023." *Nucleic Acids Research* 51, no. D1: D933–D941. <https://doi.org/10.1093/nar/gkac958>.
- Mayne, B., O. Berry, and S. Jarman. 2021. "Optimal Sample Size for Calibrating DNA Methylation Age Estimators." *Molecular Ecology Resources* 21, no. 7: 2316–2323. <https://doi.org/10.1111/1755-0998.13437>.
- Moen, E., R. Vabø, S. Smoliński, C. Denechaud, N. O. Handegard, and K. Malde. 2023. "Age Interpretation of Cod Otoliths Using Deep Learning." *Ecological Informatics* 78: 102325. <https://doi.org/10.1016/j.ecoinf.2023.102325>.
- Moore, S. R., K. L. Humphreys, N. L. Colich, et al. 2020. "Distinctions Between Sex and Time in Patterns of DNA Methylation Across Puberty." *BMC Genomics* 21, no. 1: 389. <https://doi.org/10.1186/s12864-020-06789-3>.
- Morán, P., and A. Pérez-Figueroa. 2011. "Methylation Changes Associated With Early Maturation Stages in the Atlantic Salmon." *BMC Genetics* 12, no. 1: 86.
- Parker, G. A. 1992. "The Evolution of Sexual Size Dimorphism in Fish." *Journal of Fish Biology* 41: 1–20. <https://doi.org/10.1111/j.1095-8649.1992.tb03864.x>.
- Parrott, B. B., and E. M. Bertucci. 2019. "Epigenetic Aging Clocks in Ecology and Evolution." *Trends in Ecology & Evolution* 34, no. 9: 767–770. <https://doi.org/10.1016/j.tree.2019.06.008>.
- Piferrer, F. 2021. "Epigenetic Mechanisms in Sex Determination and in the Evolutionary Transitions Between Sexual Systems." *Philosophical Transactions of the Royal Society B* 376, no. 1832: 20200110. <https://doi.org/10.1098/rstb.2020.0110>.
- Piferrer, F., and D. Anastasiadi. 2023. "Age Estimation in Fishes Using Epigenetic Clocks: Applications to Fisheries Management and Conservation Biology." *Frontiers in Marine Science* 10: 1062151. <https://doi.org/10.3389/fmars.2023.1062151>.
- Piferrer, F., E. A. Miska, and D. Anastasiadi. 2024. "Epigenetics in Fish Evolution." In *On Epigenetics and Evolution*, edited by C. Guerrero-Bosagna, 283–306. Elsevier. <https://www.sciencedirect.com/science/article/pii/B9780443190513000103>.
- Polanowski, A. M., J. Robbins, D. Chandler, and S. N. Jarman. 2014. "Epigenetic Estimation of Age in Humpback Whales." *Molecular Ecology Resources* 14, no. 5: 976–987. <https://doi.org/10.1111/1755-0998.12247>.
- R Core Team. 2024. "R: A Language and Environment for Statistical Computing [Computer Software]." <https://www.R-project.org/>.
- Rochette, N. C., A. G. Rivera-Colón, and J. M. Catchen. 2019. "Stacks 2: Analytical Methods for Paired-End Sequencing Improve RADseq-Based Population Genomics." *Molecular Ecology* 28, no. 21: 4737–4754. <https://doi.org/10.1111/mec.15253>.
- Simpson, D. J., and T. Chandra. 2021. "Epigenetic Age Prediction." *Aging Cell* 20, no. 9: e13452. <https://doi.org/10.1111/accel.13452>.
- Strand, E. L., R. McBride, E. Robillard, A. G. Bodnar, S. A. Wanamaker, and T. P. O'Donnell. 2025. "Building an 'Epigenetic Clock': Utilizing Whole Genome DNA Methylation Patterns to Predict Age in Haddock, *Melanogrammus aeglefinus*." *bioRxiv*. <https://doi.org/10.1101/2025.10.16.682892>.
- Supek, F., M. Bošnjak, N. Škunca, and T. Šmuc. 2011. "REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms." *PLoS One* 6, no. 7: e21800. <https://doi.org/10.1371/journal.pone.0021800>.
- Tangili, M., A. J. Slettenhaar, J. Sudyka, et al. 2023. "DNA Methylation Markers of Age(ing) in Non-Model Animals." *Molecular Ecology* 32, no. 17: 4725–4741. <https://doi.org/10.1111/mec.17065>.

Trucchi, E., A. B. Mazzeella, G. D. Gilfillan, M. T. Lorenzo, P. Schönswetter, and O. Paun. 2016. "BsRADseq: Screening DNA Methylation in Natural Populations of Non-Model Species." *Molecular Ecology* 25, no. 8: 1697–1713. <https://doi.org/10.1111/mec.13550>.

Weber, D., J. T. Wyffels, C. Buckner, et al. 2024. "Noninvasive, Epigenetic Age Estimation in an Elasmobranch, the Cownose Ray (*Rhinoptera bonasus*)." *Scientific Reports* 14, no. 1: 26261. <https://doi.org/10.1038/s41598-024-78004-2>.

Weber, D. N., A. T. Fields, D. W. Chamberlin, W. F. Patterson, and D. S. Portnoy. 2024. "Epigenetic Age Estimation in a Long-Lived, Deepwater Scorpionfish: Insights Into Epigenetic Clock Development." *Canadian Journal of Fisheries and Aquatic Sciences* 81: 620–631. <https://doi.org/10.1139/cjfas-2023-0296>.

Weber, D. N., A. T. Fields, W. F. Patterson, B. K. Barnett, C. M. Hollenbeck, and D. S. Portnoy. 2022. "Novel Epigenetic Age Estimation in Wild-Caught Gulf of Mexico Reef Fishes." *Canadian Journal of Fisheries and Aquatic Sciences* 79: 1–5. <https://doi.org/10.1139/cjfas-2021-0240>.

Weber, D. N., S. Lindelien, A. C. Dutterer, et al. 2025. "Nonlethal, Epigenetic Age Estimation in a Freshwater Sportfish, Florida Bass (*Micropterus salmoides*)." *Ecology and Evolution* 15, no. 11: e72495. <https://doi.org/10.1002/ece3.72495>.

Wilkinson, G. S., D. M. Adams, A. Haghani, et al. 2021. "DNA Methylation Predicts Age and Provides Insight Into Exceptional Longevity of Bats." *Nature Communications* 12, no. 1: 1615. <https://doi.org/10.1038/s41467-021-21900-2>.

Zeileis, A., and G. Grothendieck. 2005. "Zoo: S3 Infrastructure for Regular and Irregular Time Series." *Journal of Statistical Software* 14: 1–27. <https://doi.org/10.18637/jss.v014.i06>.

### Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** men70109-sup-0001-DataS1.pdf.