Received: 13 November 2022 | Revised: 26 February 2023 | Accepted: 24 March 2023

DOI: 10.1111/1755-0998.13791

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

DNA methylation-based biomarkers for ageing long-lived cetaceans

Kim M. Parsons[1] | Amin Haghani[2] | Joseph A. Zoller[3] | Ake T. Lu[2] | Zhe Fei[3] | Steven H. Ferguson[4] | Eva Garde[5] | M. Bradley Hanson[1] | Candice K. Emmons[1] | Craig O. Matkin[6] | Brent G. Young[4] | William R. Koski[7] | Steve Horvath[2,3,8]

[1]Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, Washington, USA

[2]Altos Labs, San Diego, California, USA

[3]Department of Biostatistics, School of Public Health, University of California-Los Angeles, Los Angeles, California, USA

[4]Fisheries and Oceans Canada, Winnipeg, Manitoba, Canada

[5]Greenland Institute of Natural Resources, Nuuk, Greenland

[6]North Gulf Oceanic Society, Homer, Alaska, USA

[7]LGL Limited, King City, Ontario, Canada

[8]Department of Human Genetics, Gonda Research Center, David Geffen School of Medicine, Los Angeles, California, USA

Correspondence
Kim M. Parsons, Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, WA, USA.
Email: kim.parsons@noaa.gov

Steve Horvath, Altos Labs, San Diego, CA, USA.
Email: shorvath@altoslabs.com

Funding information
Paul G. Allen Family Foundation

Handling Editor: Paul A. Hohenlohe

Abstract

Epigenetic approaches for estimating the age of living organisms are revolutionizing studies of long-lived species. Molecular biomarkers that allow age estimates from small tissue biopsies promise to enhance studies of long-lived whales, addressing a fundamental and challenging parameter in wildlife management. DNA methylation (DNAm) can affect gene expression, and strong correlations between DNAm patterns and age have been documented in humans and nonhuman vertebrates and used to construct "epigenetic clocks". We present several epigenetic clocks for skin samples from two of the longest-lived cetaceans, killer whales and bowhead whales. Applying the mammalian methylation array to genomic DNA from skin samples we validate four different clocks with median errors of 2.3–3.7 years. These epigenetic clocks demonstrate the validity of using cytosine methylation data to estimate the age of long-lived cetaceans and have broad applications supporting the conservation and management of long-lived cetaceans using genomic DNA from remote tissue biopsies.

KEYWORDS
bowhead whale, DNA methylation, epigenetic clock, killer whale, skin biopsies

Kim M. Parsons and Amin Haghani are joint first authors

1 | INTRODUCTION

In wildlife ecology, longitudinal studies spanning the entire life cycle of an organism from birth through death are prized, yet exceedingly rare, and most often limited to those species with short generation times. The ability to recognize and follow individual animals through time provides unique insight into various life stages as well as accurate age estimates, but this degree of resolution is uncommon in most wildlife species due to practical limitations. In addition, the lifespan of many large mammals (both terrestrial and aquatic) often exceeds the duration of a typical field study, and these limitations are further amplified for wide-ranging and long-lived marine species such as cetaceans (whales, dolphins and porpoises) that spend most of their lives underwater. However, as most biological processes and life history parameters change through time, accurate age estimates are essential and represent key parameters for understanding population dynamics, identifying age-associated disease and physiological impacts (Hunt et al., 2013), and inferring individual and population-level effects of bio-accumulated toxicants (Ross et al., 2000). Demographic characteristics such as the age-structure of a population have been identified as useful indicator variables for monitoring marine mammal populations. Age-structured models can reveal contemporary shifts in survivorship and fecundity and detect early warning signals of population level effects resulting from sublethal disturbance to anthropogenic stimuli or from commercial or subsistence harvests (Booth et al., 2020; Doak & Morris, 1999; Holmes et al., 2007; Mosnier et al., 2014; Pirotta et al., 2018). Current best practices for ageing cetaceans most often rely on counting incremental growth layers (GLG; growth layer groups) in odontocete teeth (Hohn et al., 1989; Waugh et al., 2018) or ear plugs from mysticete whales (Gabriele et al., 2010; Lockyer, 1984). More recently, aspartic acid racemization of the eye lens has proved highly successful for estimating the age for phocoenids (Nielsen et al., 2013), odontocetes (Garde et al., 2007; Pleskach et al., 2016; Watt et al., 2020), and balaenopterids (George et al., 1999; Nielsen et al., 2013; Olsen & Sunde, 2002). However, all of these techniques rely on the collection of samples from carcasses or animals recently harvested for subsistence. Minimally invasive approaches that can be applied to small tissue biopsies such as ratios of endogenous fatty acids in blubber samples (Herman et al., 2008, 2009; Marcoux et al., 2015) show great promise and generated age estimates with a precision of ±3.8 years, but require frozen preserved samples of the subcutaneous blubber layer, potentially limiting applicability for some species and types of field studies.

Molecular biomarkers of age applied to epidermal tissues offer an alternative approach that could be applied to samples collected nonlethally via remote biopsy techniques, with a broad range of potential applications in wildlife ecology. Recent advances in technology have created new avenues for age estimation through molecular approaches. Attempts to age marine mammals from small skin samples using changes in telomere length over time were filled with both promise and pitfalls (Dennis, 2006; Izzo et al., 2011; Olsen et al., 2012) and the relationship between telomere length and age was found to be very weak in many species (Dunshea et al., 2011; Jarman et al., 2015). Most recently, attention has turned towards epigenetic markers and the high degree of precision with which validated "epigenetic clocks" can be used to estimate the chronological age of a mammal based on age related changes in DNA methylation (Beal et al., 2019; Bocklandt et al., 2011; Bors et al., 2020; Horvath, 2013; Polanowski et al., 2014; Tanabe et al., 2020).

Mounting evidence indicates that epigenetic age estimators based on age-associated changes in DNA methylation (DNAm) provide a powerful new opportunity for developing highly accurate estimators of chronological age in humans and other mammals (Barratclough et al., 2021; Horvath & Raj, 2018; Jarman et al., 2015; Jylhava et al., 2017; Tanabe et al., 2020). The term "epigenetics" encompasses a number of molecular processes including cytosine methylation. Cytosines are typically part of a cytosine-phosphate-guanine dinucleotide site, referred to as CpG sites. These CpG sites can be methylated when a methyl group binds to the cytosine to generate 5-methylcytosine. Such CpG sites are often clustered into "CG islands" around gene promoter regions and the chemical modification through methylation may affect gene expression despite the absence of modifications to the genomic sequence (Field et al., 2018; Jones et al., 2015; Razin & Cedar, 1991; Robeck, Fei, et al., 2021). The discovery that predictable changes in CpG methylation levels in specific gene regions are correlated with age in humans and other mammals created a unique opportunity for developing assays that can predict the chronological age of an organism based on empirical estimation of DNAm in these gene promoter regions.

The first DNA methylation-based age predictors, referred to as "epigenetic clocks", were developed for human saliva (Bocklandt et al., 2011) and later for all tissues (Horvath, 2013). Subsequent studies described epigenetic clocks for mice (Petkovich et al., 2017; Stubbs et al., 2017; Thompson et al., 2018; Wang et al., 2017) and many other mammalian species including bats (Wilkinson et al., 2021), primates (Horvath, Zoller, Haghani, Jasinska, et al., 2021; Horvath, Zoller, Haghani, Lu, et al., 2021; Jasinska et al., 2021), equids (Horvath, Haghani, et al., 2022; Horvath, Haghani, Zoller, et al., 2022; Larison et al., 2021), deer (Lemaître et al., 2022), dogs (Horvath, Lu, et al., 2022), cats (Raj et al., 2021), bottlenose dolphins (Beal et al., 2019; Robeck, Fei, Lu, et al., 2021), beluga whales (Bors et al., 2020). Methylation levels at highly conserved cytosines allows one to define pan mammalian ageing clocks that apply to all mammalian species (Lu et al., 2021). The accuracy of species-specific DNAm models, where fit is determined by the median absolute error (MAE) between DNAm age and chronological age for animals with known birthdates, highlights the value of these molecular biomarkers of age across a range of tissue types (e.g., $r = .91$, MAE 4.8 years) (Robeck, Fei, Haghani, et al., 2021).

The foundation of epigenetic clocks is based on the identification of age-related changes in DNA methylation patterns and both sample size and the range of ages in a calibration data set can influence the performance of age estimation models (Mayne

et al., 2021). Validating epigenetic clocks across the spectrum of ages for the group or species expected to be represented in future applications of the model will be key for evaluating the accuracy and precision of epigenetic age estimators. Meeting these requirements with data sets generated from natural populations is very challenging, particularly for long-lived species with an expected lifespan greater than the longevity of a typical study or research career. Here, we leverage samples from two unique data sets to develop an "epigenetic clock" for long-lived cetaceans using DNA methylation patterns generated from small epidermal biopsies. Bowhead whales (*Balaena mysticetus*, Linnaeus, 1758) are thought to be the longest lived extant mammal living up to 211 years (de Magalhaes et al., 2007; George et al., 1999; John & Bockstoce, 2008; Mayne et al., 2021). Incorporating skin samples from bowhead whales on the Eastern Canada-West Greenland (ECWG) population with associated age estimates based on either aspartic acid racemization or morphological characteristics provides an exceptional opportunity to examine changes in DNA methylation in individuals spanning ages ranging from 1 to 139 years. Killer whales (*Orcinus orca*, Linnaeus, 1758) are another long-lived cetacean, with an estimated longevity of 80–90 years (Olesiuk et al., 1990). A unique longitudinal data set focused on individual-based monitoring of killer whales in the eastern North Pacific spanning nearly 50 years (Centre for Whale Research, K. Balcomb) created the foundation for collecting remote biopsy skin samples from identified, known age killer whales that can be used to validate a species-specific epigenetic clock (Ford et al., 2011, 2018). Samples from individual killer whales of known chronological age, and bowhead whales with estimated ages based on established allometric growth relationships and/or aspartic acid racemization (AAR) were used to develop a method of epigenetic age estimation using a custom mammalian DNA methylation array. Validated epigenetic clocks for skin samples can be applied to remotely collected skin biopsies from living bowhead and killer whales, promising future potential as a tool supporting conservation and management efforts involving population viability analyses that depend on reliable age estimates for these (and other) endangered cetacean species.

2 | MATERIALS AND METHODS

2.1 | Animal use and ethics

For bowhead subsistence hunts, indigenous hunters had the authorization to conduct hunts and collected samples on behalf of Fisheries and Oceans Canada. Bowhead whale biopsy samples and associated drone photographs were collected in 2019 under Fisheries and Oceans Canada (DFO) licence to Fish for Scientific Purposes (LFSP) S-19/20-1007-NU and Animal Care approval (AUP) FWI-ACC-2019-14. Skin samples from eastern North Pacific killer whales were collected as previously described (Ford et al., 2018) under NMFS General Authorization no. 781–1725, and scientific research permits 781–1824-01, 16,163, 532–1822-00, 532–1822, 10,045, 18,786–03, 545–1488, 545–1761, and 15,616.

2.2 | Killer whale biopsy samples & DNA extraction

Killer whales in the eastern North Pacific are among the most intensively studied cetacean populations globally. The so-called "resident" killer whale populations inhabiting coastal waters from California to Alaska comprise individually identified whales that have been studied for over 40 years (Balcomb & Bigg, 1986; Bigg, 1982; Ford et al., 2000; Matkin et al., 1999, 2014). These longitudinal studies and reliable identification of individual whales within populations through annual photographic census data has provided unique insight into population dynamics and demographics. The resolution provided by annual documentation of births, ages at physical maturity (males) and age at first parturition (females) provides an unparalleled opportunity to validate epigenetic models of age from whale skin.

The killer whale validation data set is rare in both the representation across age classes and the number of known-age samples from wild populations based on direct observations (Bigg et al., 1990; Matkin et al., 1999, 2014; Olesiuk et al., 1990, 2005), providing an ideal training set for validating an epigenetic clock. Killer whale longevity is estimated to be 80 or 90 years (Olesiuk et al., 1990, 2005) and samples in the current data set represent individual whales ranging from age 0 (neonate) to 79 years (estimated). Reflecting killer whale age-related mortality patterns (Olesiuk et al., 2005), the number of individuals representing older age classes diminishes as expected with $n = 11$ whales estimated be >50 years old based on size, physical and reproductive maturity at the time of first observation.

Ages of individual killer whales were determined based on the sex and size of the animal during the year that it was first documented, following Olesiuk et al. (1990). Whales born during the study (post-1974) were aged in reference to the year in which they were born. Ages for whales that were juveniles or adults when field observations began in the early 1970s were aged based on the year they reached physical maturity or, for females, the year they gave birth to their first viable offspring (Olesiuk et al., 1990). Confidence estimates (0%–100%) were assigned to each individual sampled whale included in the data set reflecting the certainty around age estimates, frequency of encounters with the individual and age or state of physical maturity at the time of first identification (Matkin et al., 2014; Olesiuk et al., 1990, 2005). Genetic samples from 131 killer whales were included in the data set representing 118 different killer whales. Sampled killer whales include individuals from two different "resident" killer whale populations (Southern residents and Gulf of Alaska residents) as well as 11 individuals from an eastern North Pacific "transient" killer whale population. Killer whale samples represented individuals ranging from age 0 years (neonate or foetus; 100% certainty) to age 75+ years (50% certainty; Table S1).

Epidermal samples were collected from live killer whales using remote dart biopsy methods, and from dead stranded animals during routine post-mortem necropsy protocols (Barrett-Lennard et al., 1996; Parsons et al., 2003). Sampled individuals included in the methylation analyses include identified individual whales from the Southern Resident, Alaska Resident and Transient killer whale populations. Identities of individual whales were recorded photographically whenever possible at the time of sample collection. Total genomic DNA was extracted from skin biopsies either using a silica-membrane kit following manufacturer's protocols (DNeasy Blood and Tissue kit, Qiagen), or following a standard proteinase K phenol/chloroform/isoamyl alcohol extraction protocol (Sambrook et al., 1989).

2.3 | Bowhead whale biopsy samples and age estimates

Bowhead whales are very slow-growing and extremely long-lived baleen species that undergo periods with rapid growth (as a foetus), pauses in growth (ages 1 to about 6–7) and slow growth (age 8+) which gradually slows even more as they age (George & Thewissen, 2020). For this reason, accurate age estimates of individual bowhead whales are typically limited to early life stages when multiple sources of information are available.

Estimating age for individual bowhead whales requires a combination of several different types of data (summarized in George & Thewissen, 2020) including body length, length of the longest baleen plate, body condition measurements, other morphological measurements that include frequency of scars, colour of the peduncle and chin region, and aspartic acid racemization (AAR) analyses from eye lenses. Despite multiple data sources, bowhead age can only be approximated for whales more than approximately 7 years old because the variation in size at age is many times larger than the annual growth rate of individual whales. Whales 2–7 years old cannot be accurately aged without information on the longest baleen plate, which is only available for harvested whales, although body condition assessed from photogrammetry provides information to assist with age estimation.

Bowhead whale skin samples were collected from remote biopsy using a crossbow during photographic studies conducted in Cumberland Sound in 2017 and 2019 (Young et al., 2022), and samples collected during subsistence harvest (1996–2016). Data available for age estimates were limited to body length, body condition, frequency of scars, and colour of the peduncle region for biopsy samples. Skin samples collected during subsistence harvest included age estimates based on body length, length of the longest baleen plate, notes on scars or peduncle colour and AAR ages. Where multiple sources of data were available, ages obtained from AAR or age-at-length estimates were adjusted to take account of the additional data as below.

Age at length was estimated using data in Figure 7 of Lubetkin et al. (2012) with estimation for gradually reduced growth rates for whales older than 60 years old as suggested by their data and those of Koski et al. (1992). Yearlings were confirmed among the smallest whales by body condition using criteria in Koski et al. (2010). Body condition was also used to adjust the age of whales 2–7 years as determined by length. For example, the youngest of these whales (i.e., 2 years old) appeared to have the poorest body condition and body condition improved as young whales aged (see Figure 4 in Koski et al., 2010).

Age based on aspartic acid racemization was determined for 11 harvested bowhead whales (four females; seven males). Both eyeballs were stored at −20°C immediately after dissection from harvested whales, and one eyeball per whale was subsequently used for age estimation. Dissection of eye lenses and age estimation by the aspartic acid racemization (AAR) technique was performed using methods described in (Garde et al., 2007). Estimates of individual D/L ratios were converted to age estimates using the equation:

$$\text{Age (yrs)} = \frac{\ln\left(\frac{1+D/L}{1-D/L}\right) - \ln\left(\frac{1+D/L}{1-D/L}\right)}{2k_{\text{Asp}}}$$

where racemization rates (k_{Asp}) and D/L_0 values from bowhead whales in Alaska and Greenland were used (George et al., 1999; Heide-Jørgensen et al., 2012; Rosa et al., 2004, 2013).

In total, genetic samples from 79 bowhead whales were included in the data set representing whales aged 1–139 years (Table S2). Confidence in age estimates was assigned based on expert opinion associated with morphology age estimates and was generally lower for juveniles and the oldest whales. Total genomic DNA was extracted from skin biopsies using the automated sample preparation protocol on a QIAsymphony (Qiagen).

2.4 | Methylation array

Total genomic DNA was quantified on a Qubit (ThermoFisher) fluorometer and approximately 250 ng genomic DNA from each sample was bisulphite converted and carried forward for the custom methylation array. The mammalian DNA methylation arrays were profiled using a custom Infinium methylation array, HorvathMammalMethylChip40 (Arneson et al., 2022) based on 37,492 CpG sites, including 35,541 probes selected to assess cytosine DNA methylation levels in most mammalian species (Arneson et al., 2022). Oligonucleotide probes designed for each CpG site comprised 50 nucleotides terminating in a C-G dinucleotide to determine the methylation state of CpGs. The particular subset of species for each probe is provided in the chip manifest file found at Gene Expression Omnibus (GEO) at NCBI as platform GPL28271.

Raw data were normalized using the SeSaMe pipeline to derive beta values (and associated *p*-values) for each probe and every individual sample (Zhou et al., 2018). Beta (β) values ranged from 0 to 1, indicating the degree of methylation signal at each array probe and sample, where each array probe corresponds to a specific CpG site. A β value of 0 indicates that no gene copies were found to be

methylated at that site for a particular sample. Data were quality checked by comparing results for technical replicates (DNA aliquots for the same sample included >1 on the methylation array) and by performing unsupervised hierarchical clustering based on the interarray correlation coefficient (Pearson correlation) to cluster samples. Three killer whale samples and two bowhead whale samples were identified as technical outliers and removed from the data set.

2.5 | Biostatistical analysis

2.5.1 | Penalized regression models

Details on the clocks (CpGs, genome coordinates) and R software code are provided in the Supporting Information. Penalized regression models were implemented using the glmnet package for R (Friedman et al., 2010). We evaluated models produced by "elastic net" regression methods based on the two main parameters lambda and alpha. Alpha is the elastic net mixing parameter used to determine the blend between a ridge regression (alpha = 0.0) and a least absolute shrinkage and selection (LASSO) regression (alpha = 1.0). The optimal penalty parameter (lambda) was determined automatically using a 10-fold internal cross-validation (cv.glmnet) on the training set. The alpha value for the elastic net regression was set to 0.5 (midpoint between Ridge and LASSO type regressions) and was not optimized for model performance.

The epigenetic clock training data set included 131 killer whale samples and 79 bowhead whale samples, where morphological age was estimated with a high degree (≥90%) of confidence. We performed a cross-validation scheme to estimate the accuracy of the different DNA methylation-based age estimators. One type consisted of leaving out a single sample (LOOCV) from the regression, predicting an age for that sample, and iterating over all samples.

We also trained two sex specific clocks for skin samples from both whale species.

Species characteristics were chosen from an updated version of the anAge database (de Magalhaes et al., 2007).

DNA from seven killer whale samples and 11 bowhead samples were run in duplicate as independent technical replicates to evaluate variance across methylation estimates. Model fit was examined by calculating the median absolute error (MAE) based on the morphological age for each individual in the training data set.

2.5.2 | Epigenome wide association studies of age

EWAS was performed in each tissue separately using the R function "standardScreeningNumericTrait" from the "WGCNA" R package (Langfelder & Horvath, 2008). Next the results were combined across tissues using Stouffer's meta-analysis method. The analysis was performed using the genomic region of enrichment annotation tool (McLean et al., 2010), and the gene level enrichment was conducted using GREAT analysis and human Hg19 background (McLean et al., 2010).

2.5.3 | Gene ontology enrichment analysis

The analysis was done using the genomic region of enrichment annotation tool (McLean et al., 2010), and the gene level enrichment was performed as above using GREAT analysis (McLean et al., 2010) and human Hg19 background.

3 | RESULTS

High quality DNA methylation profiles were generated from $n = 131$ unique killer whale skin samples, from 3 different populations, and $n = 79$ unique skin samples from bowhead whales, plus technical replicates at a rate of 7% and 14% for killer whales and bowhead whales, respectively. The same (HorvathMammalMethylChip40) methylation array platform was used for both species. The mammalian methylation array profiles CpGs with neighbouring DNA sequences that are conserved between different species of the mammalian class. We find that 30,467 CpGs out of 37,492 CpGs on the mammalian array map to the genome of killer whales.

Hierarchical clustering of the skin samples revealed two technical outliers among the bowhead whale samples and three technical outliers among the killer whale samples (possibly due to low amounts of DNA) that were subsequently removed from the analysis.

3.1 | Epigenetic clocks

A confidence measure was associated with each chronological age estimate based on the best available data from repeated field observations or AAR data. Epigenetic clocks were developed from methylation data generated for individual whales with an assigned confidence of at least 90% in the chronological age estimate. In the elastic net regression model, we used a square root transformation: sqrt(Age+1) as dependent variable.

To arrive at unbiased estimates of the epigenetic clocks, we performed cross-validation analyses and obtained estimates of the age correlation r (defined as Pearson's correlation between the age estimate, DNAm age and chronological age), as well as the median absolute error (MAE).

From these we generated four epigenetic clocks (Figure 1) for cetacean skin samples based on varying numbers of CpGs (details can be found in Supporting Information and Table S3). The killer whale clock is based on 50 CpGs. The bowhead whale clock is based on 61 CpGs, and the bowhead whale clock for young animals (age <50 years) is based on 34 CpGs. The two-species clock for both killer whales and bowhead whales is based on 136 CpGs.

Technical replicates were generated from nine killer whale samples and 11 bowhead whale samples processed in duplicate. A

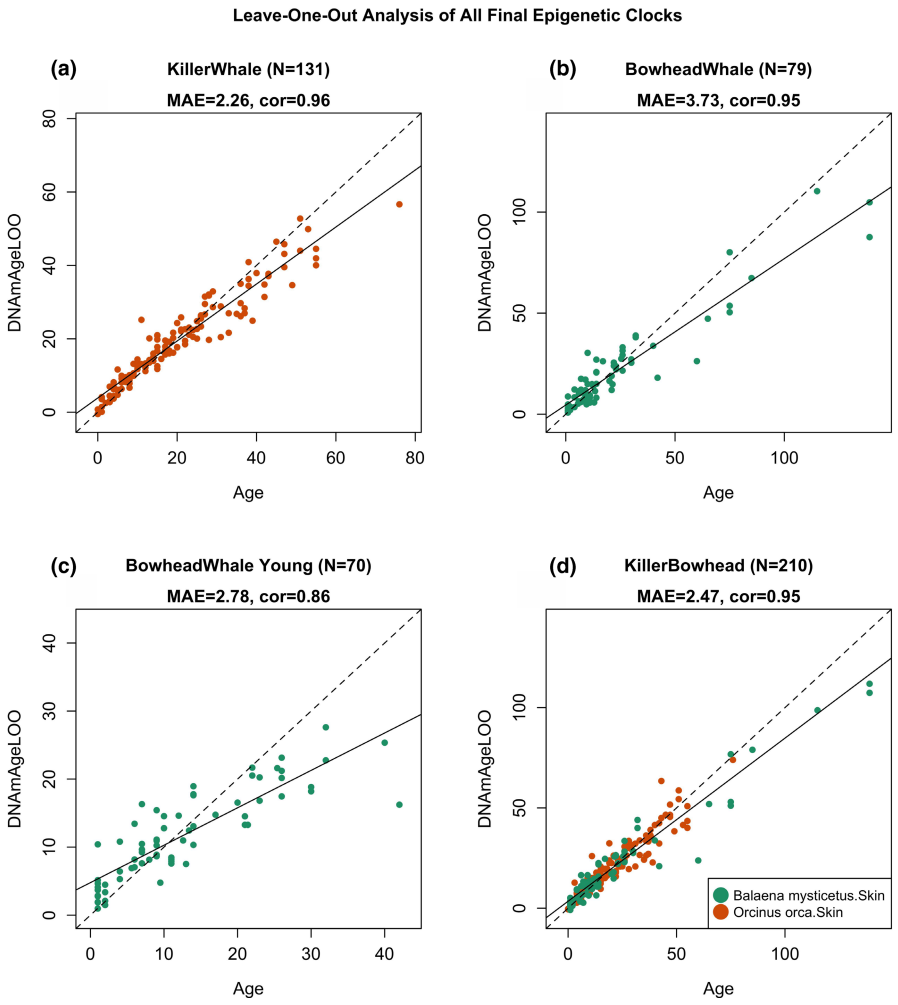

FIGURE 1 Accuracy of epigenetic clocks for skin samples from bowhead whales and killer whales. Each panel reports morphological or chronological age (x-axis) versus a leave one out (LOO) cross validation estimates of DNA methylation age (y-axis, in units of years). (a) Killer whale, (b) Bowhead whale clock trained on animals from the entire age range, (c) Bowhead whale clock trained on animals younger than 50. (d) Two-species clock for both killer whale and bowhead whales. Each panel reports the sample size, correlation coefficient, median absolute error (MAE) in units of years. Linear regression lines of predicted age ~ Age shown as a solid line. Dashed line is 1:1 reference lines if the predicted epigenetic age is identical to the chronological (observed) age.

single replicate each of two samples exhibited anomalous results for the killer whale model and were dropped from the data set (Figures S1 and S2). Of the remaining samples, seven killer whale and 11 bowhead whale samples were analysed in duplicate. The average difference in estimated age for replicates was 0.17 years (SD = 3.619) and 1.15 years (SD = 5.945) for the species-specific killer whale and bowhead whale models respectively. In addition to the technical replicates, 13 individual killer whales were represented by multiple skin biopsies collected at known time points providing an opportunity to examine DNAm estimates for the same individual whale over time periods ranging from 2 to 9 years. DNAm age estimates differed from the time elapsed between sampling events by an average of 2.375 years (SD = 2.212), less than the overall mean MAE, and DNAm estimates showed a slight trend towards underestimating (rather than overestimating) the age difference between consecutive biopsy samples by 1–7 years (Figure S3).

Sex specific clocks were trained using skin samples from both bowhead whales and killer whales (Figure 2). Both sex specific clocks are highly accurate in their respective sex (Figure 2a). However, the male clock applied to female samples (Figure 2c) leads to a systematic offset, whereby the male clock underestimates the ages of female samples, highlighting sex-specific differences in methylation rates with age.

3.2 | Comparison of age related CpGs in bowhead and killer whales

The custom mammalian methylation array used contains 30,467 probes that could be aligned to specific loci adjacent to 5730 unique genes in the killer whale (Orcinus_orca.GCF_000331955.2_Oorc_1.1) genome. The design of the mammalian methylation array targeting conserved DNA regions, extends this annotation to the bowhead whale data. Our epigenome wide association studies (EWAS) correlated each of these CpGs with chronological age in skin samples from bowhead whales ($n = 79$, age range = 1–39 years) and killer whales ($n = 131$, age range = 0–76 years). At a nominal p-value $<10^{-5}$, a total of 1122 and 1578 age-related CpGs were identified in bowhead and killer whales respectively (Figure 3a, Table S3). Some of the top age-related CpGs include bowhead whales, hypomethylation in *FOSB* exon, *CAMK2G* exon, *PEX14* intron; killer whales, hypomethylation in *CCSER2* exon, *LOC117200810* downstream, *LHX4* promoter, *LOC101288163* upstream, and *NOVA1* exon. In both

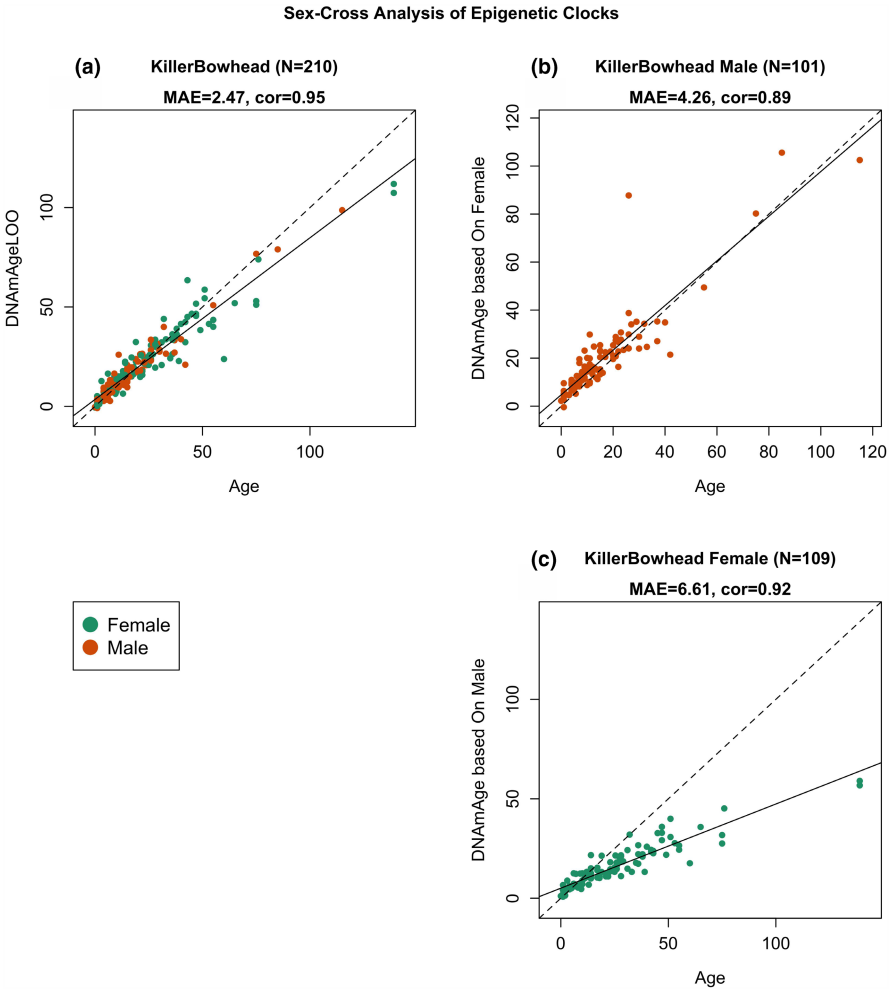

FIGURE 2 Sex specific clocks in killer whales and bowhead whales. Results for two separate epigenetic clocks for skin samples from males and females, respectively. Two-species clock for both killer whale and bowhead whales but trained on a specific sex. Each panel reports morphological or chronological age (x-axis) versus a leave one out (LOO) cross validation estimates of DNA methylation age (y-axis, in units of years). (a) Leave one out (LOO) cross validation results for (i) the female clock applied to female samples (green dots) and (ii) the male clock applied to male samples (orange samples). (b) Female clock (y-axis) applied to male samples. (c) Male clock (y-axis) applied to female samples. Each panel reports the sample size, correlation coefficient, median absolute error (MAE) in units of years. Linear regression lines of predicted age – age shown as a solid line. Dashed line is 1:1 reference lines if the predicted epigenetic age is identical to the chronological (observed) age.

of these species, there was a modest excess of positive association of age with sites located within CpG islands compared to CpG non-islands (Figure 3b). In general, age-related CpGs were found to be located in all genic and intergenic regions that can be defined relative to transcriptional start sites (Figure 3c). As expected, the proportion of hypermethylated CpGs was higher in regulatory regions (promoter and 5'UTR) than other genic regions.

Comparing the CpG level changes in these two species identified a subset of 318 CpGs with a similar ageing pattern across both bowhead and killer whales (Figure 3d). In general, the CpG level correlation of DNAm ageing in these two species was only 0.32 (Figure 3e). While this finding warrants validation with larger sample size, our result suggests that bowhead and killer whales each have unique DNAm ageing signatures. We categorized the CpGs based on the pattern of change in these two species: (1) Shared signatures, which are the 318 shared CpGs that related to age in the same direction in both. Some examples include hypermethylation in *EN1* 5'UTR and hypomethylation in *NOVA1* exon (Figure 4a). (2) Killer whale unique DNAm ageing signatures, for example, hypomethylation in *IKZF2* intron, *ABCB4 5'UTR* (Figure 4b). (3) Bowhead whale DNAm ageing signatures, for example, hypomethylation in *SNAP25* exon, *EHD4* exon (Figure 4c). (4) The CpGs with contrasting ageing pattern between these two whale species, for example, while FRMD4A exon is hypermethylated with age in killer whales, it is hypomethylated in bowhead whales (Figure 4d).

3.3 | Enrichment analysis of age related CpGs

We performed a gene level enrichment analysis of the identified age-related CpGs. Although there is a complex relationship between DNAm, gene and protein expression, an enrichment analysis can implicate biological processes that are associated with ageing effects that are (a) shared between killer whales and bowhead whales and (b) are specific to each of these species. Similar to all other mammals (Lu et al., 2021), age related CpGs that are shared between bowhead and killer whales are adjacent to genes that play a role in developmental pathways and are located near targets of polycomb repressive complex 2 (e.g., PRC2, EED, SUZ12, H3K27Me3, Figure 5, Table S4, Table S5, Table S6).

Bowhead whale specific age related CpGs were located near genes that play a role in Wnt, and Cadherin signalling, and several cancer related terms in MsigDB perturbation database (Figure 5a,b,e). The unique transcriptional factor motifs for age related changes in

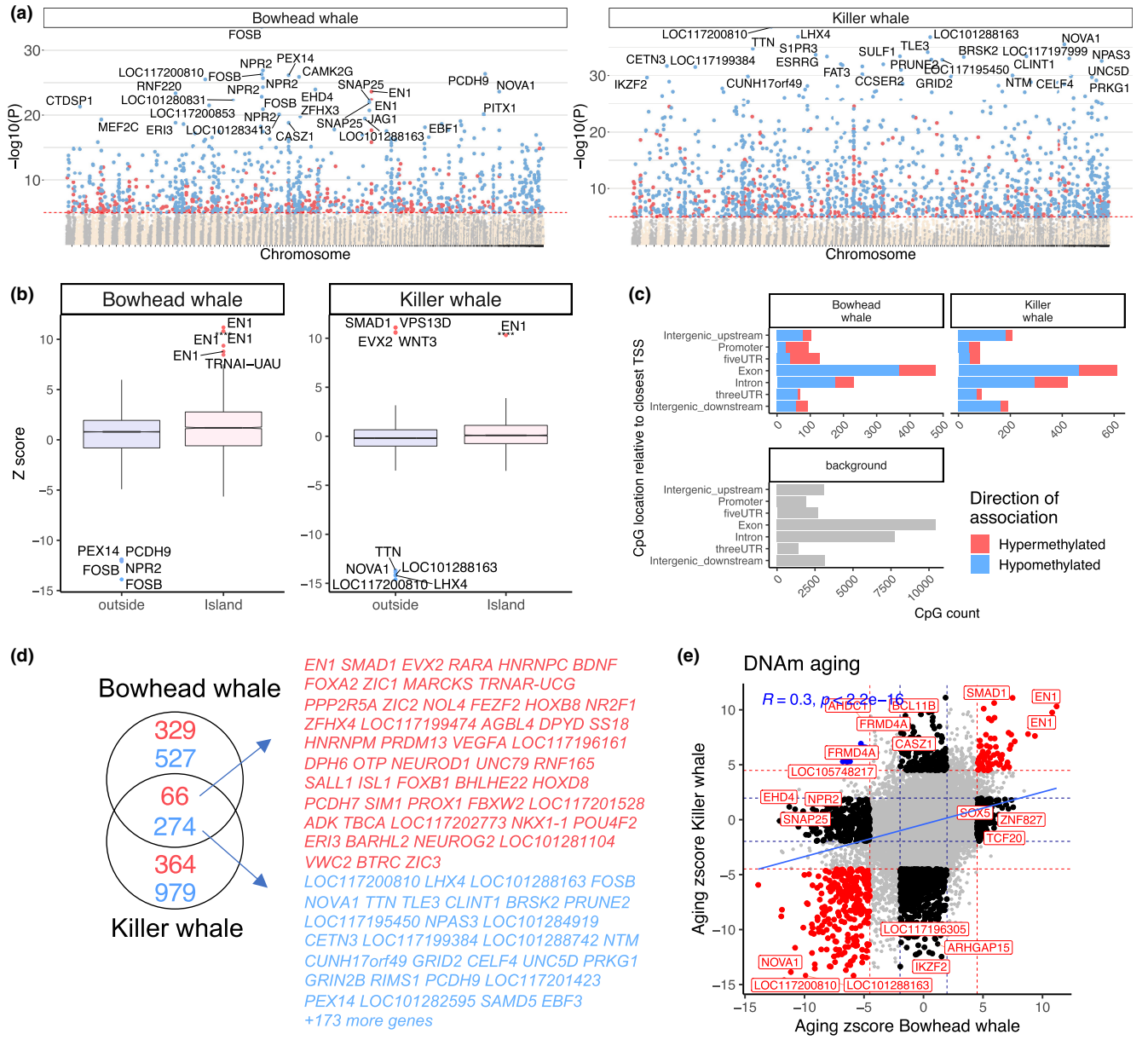

FIGURE 3 Comparative epigenome-wide association (EWAS) of age in bowhead and killer whales. (a) Manhattan plots of the EWAS of chronological age in skin of bowhead whales ($n = 96$) and Killer whales ($n = 133$). The coordinates are estimated based on the alignment of Mammalian array probes to Orcinus_orca.GCF_000331955.2_Oorc_1.1 genome assembly. The direction of associations with $p < 10^{-5}$ (red dotted line) is highlighted by red (hypermethylated) and blue (hypomethylated) colours. Top 30 CpGs was labelled by the neighbouring genes. Details can be found in Table S4. (b) Box plot analysis of DNAm ageing association by CpG island status. The top age related CpGs in each species are labelled by adjacent genes (**** $p < 10^{-4}$). (c) Location of top CpGs in each tissue relative to the adjacent transcriptional start site. The grey colour in the last panel represents the location of 30,467 mammalian methylation array probes mapped to the Killer whale genome. (d) Venn diagram of the overlap of significant CpGs for each species. Details can be found in Table S6. (e) Sector plot of DNA methylation ageing in skin of bowhead whales and Killer whales. Red dotted line: $p < 10^{-5}$; blue dotted line: $p > .05$; Red dots: shared CpGs; black dots: tissue specific changes; blue dots: CpGs whose age correlation differs between bowhead whales and killer whales. Details can be found in Table S7.

bowhead whale included targets of HMX1, SP1, ELK1, NRF2, and E12 (Figure 5d).

In contrast, killer whale specific CpGs implicated genes involved in membrane assembly, RXR signalling, nervous system development such as synapse formation (Figure 5a,c,d). The transcription factor AR was implicated by CpGs that are specific to killer whale (Figure 5d).

3.4 | Tissue atlas for the bowhead whale

It will be interesting to study how methylation levels of CpGs differ across different tissue types. As a first step, we present a small tissue atlas that profiles skin, skeletal muscle, liver, kidney, heart, and cerebral cortex from the bowhead whale (Figure 6). The tissue atlas

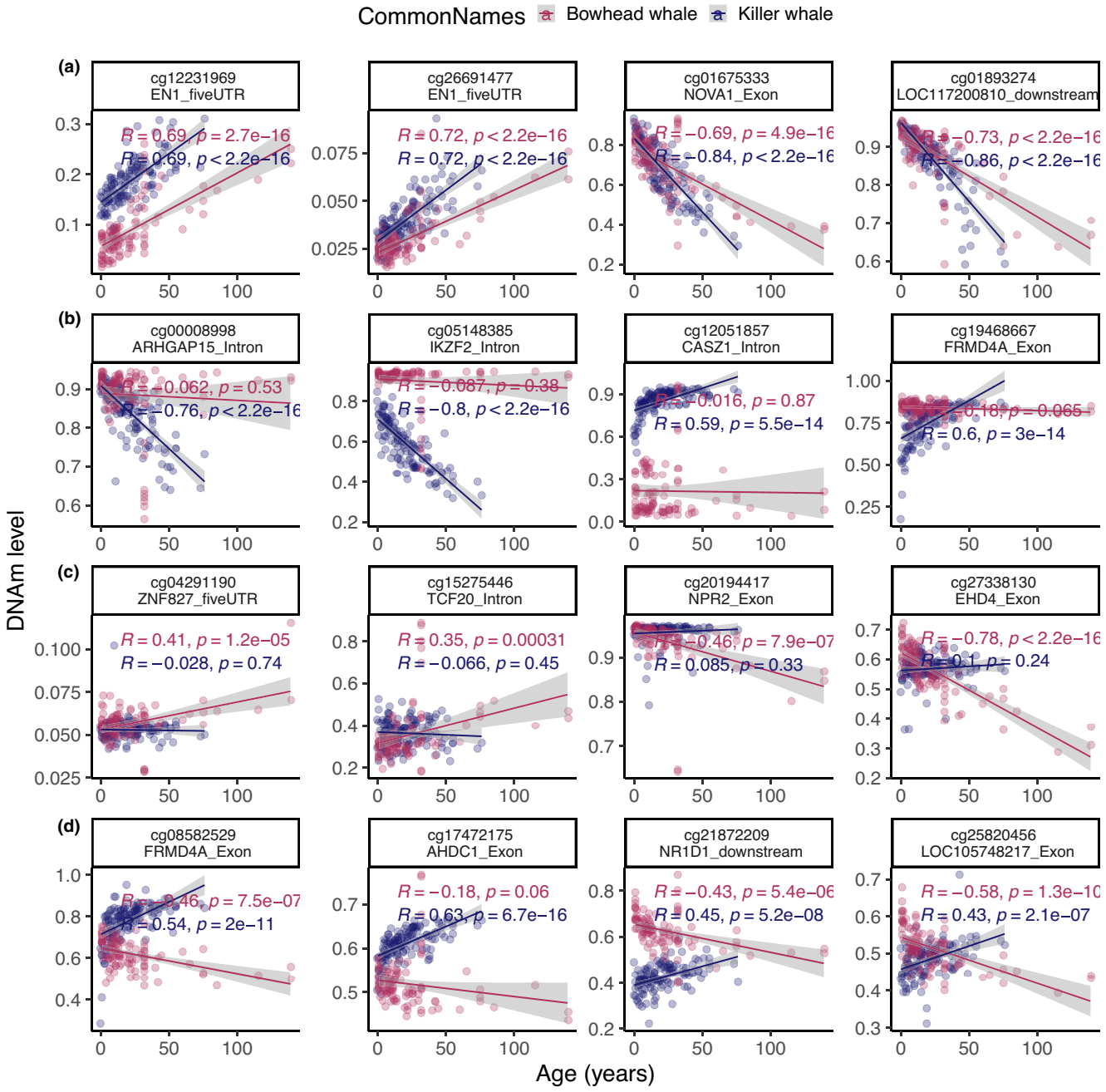

FIGURE 4 Scatter plots of age-related changes in selected CpGs in skin of bowhead and killer whales. (a) CpGs that change with age in both species. (b) Examples of killer whale specific changes. (c) Examples of bowhead whale specific changes. (d) Selected CpGs with divergent ageing pattern between bowhead and killer whales. Shaded regions indicate the 95% confidence intervals of the regression lines (solid lines). Details can be found in Table S6.

demonstrates that the top age related CpGs in bowhead whales and killer whales have consistent mean methylation levels across the different tissue types (Figure 6).

4 | DISCUSSION

Using data sets from two divergent long-lived cetacean species, we present a robust epigenetic clock for both killer whales and bowhead whales using genomic DNA isolated from small skin biopsies that can be collected using remote, minimally invasive techniques from live animals. The tight correlation between epigenetic and chronological age for the two species is based on CpGs that are highly conserved across mammalian species (Arneson et al., 2022). The fit of the species-specific DNAm models describing the relationship between epigenetic and chronological age over the majority of the lifespan of bowhead whales ($N = 96$, $r = .96$) and killer whales ($N = 133$, $r = .95$) is comparable to the two-species model ($R = .95$; Figure 1), as well as

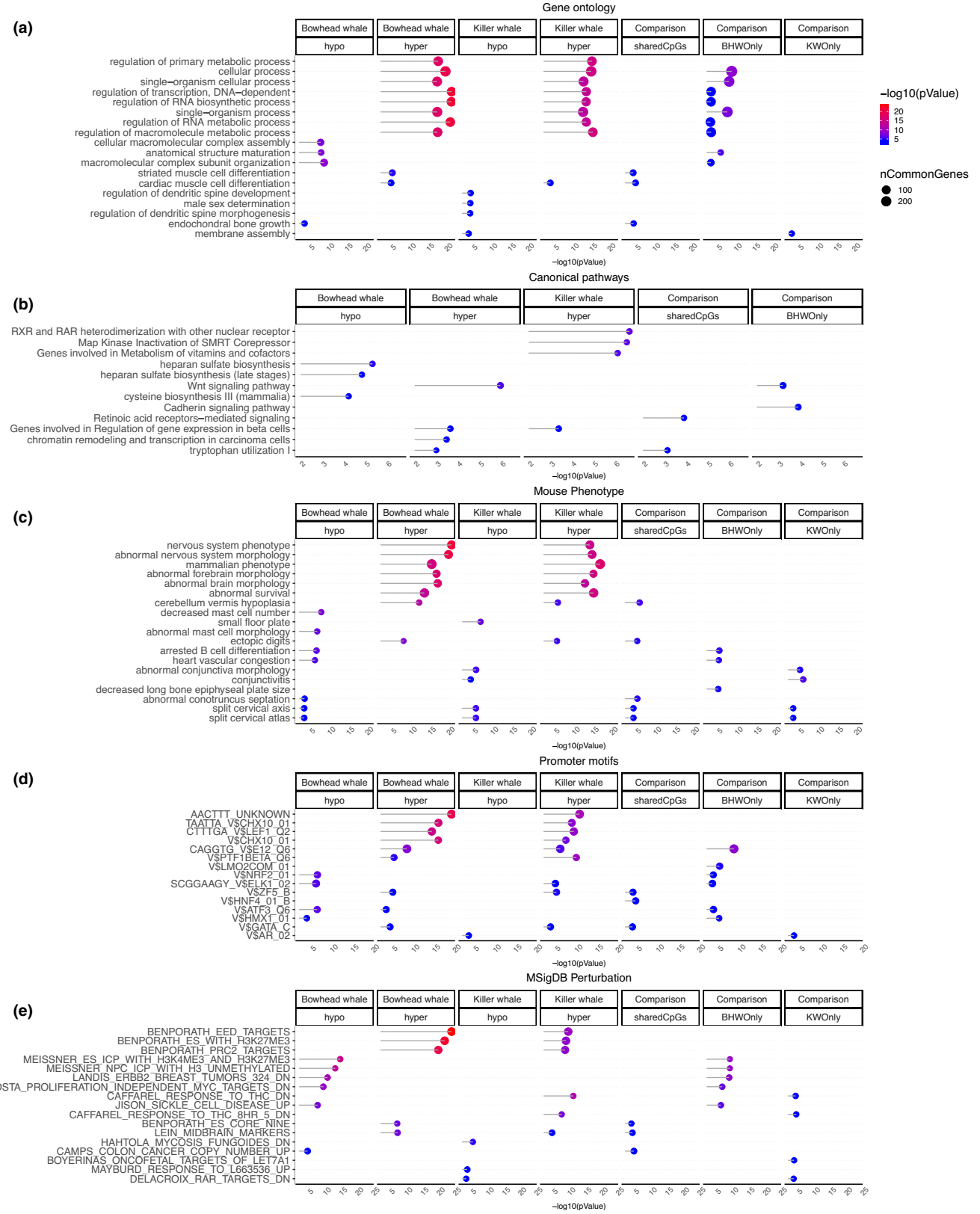

FIGURE 5 Gene set enrichment analysis of DNA methylation ageing in bowhead (BHW) and killer (KW) whales. The gene level enrichment was done using GREAT analysis and human Hg19 background. Data sets: (a) Gene ontology, (b) canonical pathways, (c) mouse phenotypes, (d) promoter motifs and (e) MSigDB perturbation. The results were filtered for significance at $p < 10^{-3}$. The comparison columns are based on the shared or unique CpGs that were identified in the sector plot analysis (Figure 3e) of DNAm ageing in two species. Details can be found in Table S5.

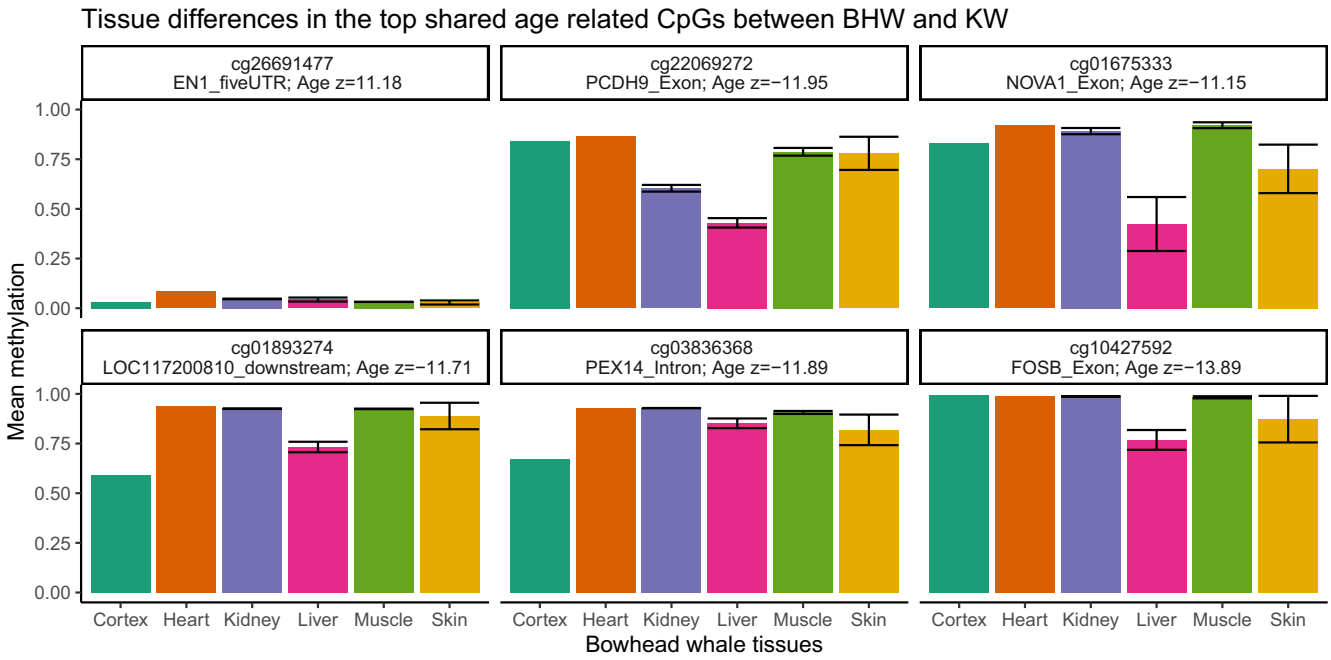

FIGURE 6 Tissue differences in the top age related CpGs in the skin of bowhead and killer whales. The bar plots indicate the mean methylation levels in different bowhead tissues. Sample size: Skin 79; Muscle 2; Liver 3; Kidney 2; Heart 1; Cortex 1. The error bars represent the standard deviation of methylation levels in tissues with more than one sample.

that described for other cetacean epigenetic age clocks, including those fit to data for beluga whales, humpback whales and bottlenose dolphins (Barratclough et al., 2021; Beal et al., 2019; Robeck, Fei, Lu, et al., 2021). The validated epigenetic clocks for skin described here provide a valuable tool for estimating the age of wild killer whales or bowhead whales from minimally invasive, remotely obtained skin samples, creating a new avenue for generating previously intractable life history parameters.

Recent technological advances and multiagency collaborations have supported the development and publication of a wide range of epigenetic clocks. A number of recent publications (including this one) have leveraged the use of an established, reproducible mammalian methylation array (HorvathMammalMethylChip40; Arneson et al., 2022) to validate epigenetic clocks for specific tissues and/or species, highlighting the value of a generalized array targeting a large taxonomic group. While the value and potential applications of a universal mammalian epigenetic clock (Lu et al., 2021) are unequivocal, the landscape of the epigenome varies considerably across both species and tissue types, and existing data suggest that the predictive power and accuracy of DNAm clocks are improved when validated across large data sets representing both the species and the specific tissues of interest (Field et al., 2018; Zhang et al., 2019).

Recently published epigenetic clocks for equids (Horvath, Haghani, et al., 2022; Horvath, Haghani, Zoller, et al., 2022; Larison et al., 2021), marsupials (Horvath, Haghani, et al., 2022; Horvath, Haghani, Zoller, et al., 2022), or odontocetes (Robeck, Fei, Lu, et al., 2021) have leveraged the same mammalian methylation array to support development of epigenetic clocks for targeted taxa. Epigenetic clocks targeting multiple species provide additional flexibility by leveraging data from related species that are better represented by available sample numbers from known age individuals. The recommended minimum sample size of 70 individuals distributed uniformly across the age range to calibrate an epigenetic clock is a goal that can often be challenging to meet for species of greatest conservation and management concern (Mayne et al., 2021). "Borrowing" information from related species to generate taxonomically broad epigenetic clocks can lessen the sampling burden on high priority species. The two-species clock presented here for both killer whales and bowhead whale performed similarly to the species-specific clocks, universal pan mammalian clocks have been developed using the same platform (Lu et al., 2021). However, both species and population specific effects on age-related methylation changes for comparable tissues are currently unknown, and, as demonstrated here, species specific clocks generally have greater accuracy than such universal clocks. Although speculated to be low, based on the conserved CpG sites targeted by methylation array employed here, additional studies examining intraspecific variance in age-related changes in DNAm would be of great valuable, particularly for species such as killer whales with sympatric, yet divergent, ecotypes (Baird et al., 1992; Foote et al., 2019; Ford et al., 1998; Riesch et al., 2012; Saulitis et al., 2000). Extrinsic factors including environmental stressors and dietary preferences are known to affect the epigenome (e.g. Feil & Fraga, 2012), and assessing whether or not dietary preferences (e.g., mammals vs. fish) affect the CpGs included in these epigenetic clocks, and the associated age-related methylation patterns, is important for understanding their broader applicability for globally distributed species such as killer whales.

The leave one out cross-validation (LOOCV) analysis indicates that the killer whale and bowhead epigenetic clocks estimated the age of sampled individuals with a median absolute error of 2.21 and 3.77 years, respectively. The two-species clock performed similarly with a median error of 2.57 years. Bowhead whale longevity has been estimated to be greater than 200 years, based on the recovery of fragments of historical weapons and aspartic acid racemization (AAR) of the eye lens nucleus (George et al., 1999, 2011; John & Bockstoce, 2008), highlighting that the epigenetic clock approximates age within +/− 2% of the bowhead lifespan. Here, we estimated the age of the longest-lived bowhead whale in our data set at 107 years for an individual with an AAR age 139 years. Among the longest lived vertebrates, and the longest lived mammal (George et al., 1999), the bowhead whale data set is valuable as a species of extreme longevity. Recent work on lifespan estimators from methylation data estimate the maximum longevity of the bowhead whale to be on the order of 268 years (Mayne et al., 2019).

Among the epigenetic clocks published recently for odontocetes based on skin samples, there is a general trend towards a best fit for mid-range ages with DNAm age estimates tending to diverge from the chronological age at both ends of the age continuum (Barratclough et al., 2021; Bors et al., 2020; Robeck, Fei, Lu, et al., 2021). This trend is particularly evident for very long-lived individuals where the epigenetic age tends to underestimate chronological age, based on the relatively limited sample of individuals representing older age classes. Improving the fit of the epigenetic clocks for these age classes is challenging due to the very few known age individuals approaching maximum longevity for a given species. In addition, studies of semi-supercentenarian humans suggest epigenetic changes associated with exceptional longevity resulting in epigenetic ages that are younger than their chronological age (Gutman et al., 2020; Horvath et al., 2015). If this pattern extends to nonhuman mammals, this may inherently limit the accuracy of epigenetic clocks for individuals with longevity in these extreme age classes.

Cross validation of sex specific clocks for the two whale species provides novel insight into age related differences in methylation rates between males and females. Applying the male clock to female skin samples underestimates the ages of females (Figure 2c), which could be attributed to two reasons. First, technical variation cannot be ruled out. Second, the observed differences could indicate that females age more slowly than males. Sex differences in longevity are well documented in several cetacean species, including killer whales (Olesiuk et al., 2005). However, the latter hypothesis should be more carefully tested with additional samples and other molecular biomarkers of ageing. Epigenetic clock studies in humans suggest that human females age more slowly than males (Horvath et al., 2016).

Age biomarkers, in general, often suffer from a paucity of samples of known age individuals for methods validation. Small sample sizes and intraspecific variability invariably contribute to uncertainty around age estimates for a range of biomarkers, however, the ability to assign individuals to age classes with a resolution greater than decades can significantly advance efforts to parameterize population demographic models. The accuracy of DNAm age estimates is comparable to, and a slight improvement on, that estimated from endogenous fatty acids (Herman et al., 2008, 2009) and eye-nucleus aspartic acid racemization (George et al., 1999; Nielsen et al., 2013; Olsen & Sunde, 2002; Rosa et al., 2013), with the distinct advantage of being able to apply epigenetic clocks to minimally invasive skin samples collected from living whales. Leveraging additional sources of morphological data through photographic identifications or aerial drone footage could further enhance the accuracy of age estimates and validation of age biomarkers such as epigenetic clocks (Cosens & Blouw, 2003; Frasier et al., 2020; Schweder et al., 2010; Young et al., 2022).

Information on population age structure is essential for identifying trends in key demographic parameters including age-specific survivorship and fecundity, providing valuable insights into potential sources of mortality and population drivers for populations of management and conservation concern (e.g., Caswell, 2001; Caughley, 1966; Fujiwara & Caswell, 2001; Holmes et al., 2007; Mosnier et al., 2014; Olesiuk et al., 2005). Beyond accurate age estimates for individuals, determining relative ages for pairs of individuals can further enhance abundance and survival estimates from capture-mark-recapture data, resolving ordinal age among parent-offspring pairs (Bravington et al., 2014; Jarman et al., 2015; Polanowski et al., 2014). The increasing accessibility of genomic sequencing services and age biomarkers, such as DNAm, that can be applied to tissues that can be collected with relative ease from living animals promises to increase the potential to address ecologically relevant questions for a wide range of cetacean species.

AUTHOR CONTRIBUTIONS

Kim M. Parsons, Steve Horvath and Steven H. Ferguson initially conceived and designed the study. Kim M. Parsons conducted DNA extraction and sample selection for all killer whale samples. Steven H. Ferguson and Brent G. Young collected all bowhead whale samples, and Candice K. Emmons, M. Bradley Hanson and Craig O. Matkin collected tissue biopsies from North Pacific killer whales. Eva Garde and William R. Koski estimated ages for bowhead whales based on morphometrics and aspartic acid racemization. Amin Haghani, Joseph A. Zoller, Ake T. Lu, Zhe Fei and Steve Horvath developed and validated the epigenetic clocks and conducted data analyses. Kim M. Parsons, Amin Haghani, Steven H. Ferguson and Steve Horvath wrote the manuscript with input for all other authors.

ACKNOWLEDGEMENTS

This project would not have been possible without the inspired and dedicated efforts of Michael Bigg and Ken Balcomb who initiated the individual based monitoring of the Southern Resident killer whale population. We are grateful to the Centre for Whale Research, WA, Peter Olesiuk and John Ford for early discussions about early age estimates for individual Southern Resident killer whales. We thank the community of Pangnirtung and the Pangnirtung Hunters and Trappers Organization for their support for the Cumberland Sound field research, in particular the

boat captains, field guides, and assistants who have contributed to this research over the years. We are grateful to the many field and laboratory technicians at the Northwest Fisheries Science Centre who support the ongoing genetic and genomic analyses of killer whale tissue samples. This work was supported by the Paul G. Allen Frontiers Group (SH). Funding for the bowhead collections was provided by the Nunavut Wildlife Management Board, Fisheries and Oceans Canada, LGL Limited, and World Wildlife Fund Canada.

CONFLICT OF INTEREST STATEMENT

Steve Horvath is a founder of the nonprofit Epigenetic Clock Development Foundation which plans to licence several patents from his employer UC Regents. These patents list Steve Horvath as inventor. The other authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The DNA methylation data have been posted on Gene Expression Omnibus and can be accessed through GSE225288 (bowhead whale) and GSE225289 (killer whale). The manifest for the Horvath Methylation Array is available at Gene Expression Omnibus (GPL28271: Illumina HorvathMammalianMethylChip40 BeadChip). The mammalian methylation array can be purchased from the nonprofit Epigenetic Clock Development Foundation (https://clockfoundation.org/).

ORCID

Joseph A. Zoller https://orcid.org/0000-0001-6309-0291
Steve Horvath https://orcid.org/0000-0002-4110-3589

REFERENCES

Arneson, A., Haghani, A., Thompson, M. J., Pellegrini, M., Kwon, S. B., Vu, H., Maciejewski, E., Yao, M., Li, C. Z., Lu, A. T., Morselli, M., Rubbi, L., Barnes, B., Hansen, K. D., Zhou, W., Breeze, C. E., Ernst, J., & Horvath, S. (2022). A mammalian methylation array for profiling methylation levels at conserved sequences. *Nature Communications*, *13*(1), 783.

Baird, R. W., Abrams, P. A., & Dill, L. M. (1992). Possible indirect interactions between transient and resident killer whales: Implications for the evolution of foraging specializations in the genus *Orcinus*. *Oecologia*, *89*, 125–132.

Balcomb, K. C., & Bigg, M. A. (1986). In B. C. Kerkevold & J. S. Lockard (Eds.), *Population biology of the three resident killer whale pods in Puget sound and off southern Vancouver Island. Behavioural biology of killer whales*. Alan R. Liss, Inc.

Barratclough, A., Smith, C. R., Gomez, F. M., Photopoulou, T., Takeshita, R., Pirotta, E., Thomas, L., McClain, A. M., Parry, C., & Zoller, J. A. (2021). Accurate epigenetic aging in bottlenose dolphins (*Tursiops truncatus*), an essential step in the conservation of at-risk dolphins. *Journal of Zoological and Botanical Gardens*, *2*(3), 416–420.

Barrett-Lennard, L. G., Smith, T. G., & Ellis, G. M. (1996). A cetacean biopsy system using lightweight pneumatic darts, and its effect on the behaviour of killer whales. *Marine Mammal Science*, *12*(1), 14–27.

Beal, A. P., Kiszka, J. J., Wells, R. S., & Eirin-Lopez, J. M. (2019). The bottlenose dolphin epigenetic aging tool (BEAT): A molecular age estimation tool for small cetaceans. *Frontiers in Marine Science*, *6*, 561.

Bigg, M. A. (1982). An assessment of killer whale (*Orcinus orca*) stocks off Vancouver Island, British Columbia. *Report of the International Whaling Commission*, *32*, 655–666.

Bigg, M. A., Olesiuk, P. F., Ellis, G. M., Ford, J. K. B., & Balcomb, K. C. (1990). Social organization and genealogy of resident killer whales (*Orcinus orca*) in the coastal waters of British Columbia and Washington state. Individual recognition of cetaceans: Use of photo-identification and other techniques to estimate population parameters. P. Hammond. Cambridge. *Rep Int Whal Comm. Special Issue*, *12*, 383–405.

Bocklandt, S., Lin, W., Sehl, M. E., Sanchez, F. J., Sinsheimer, J. S., Horvath, S., & Vilain, E. (2011). Epigenetic predictor of age. *PLoS One*, *6*(6), e14821.

Booth, C. G., Sinclair, R. R., & Harwood, J. (2020). Methods for monitoring for the population consequences of disturbance in marine mammals: A review. *Frontiers in Marine Science*, *7*(115), 00115.

Bors, E. K., Baker, C. S., Wade, P. R., O'Neill, K. B., Shelden, K. E., Thompson, M. J., Fei, Z., Jarman, S., & Horvath, S. (2020). An epigenetic clock to estimate the age of living beluga whales. *Evolutionary Applications*, *14*(5), 1263–1273.

Bravington, M. V., Jarman, S. N., & Skaug, H. J. (2014). *Antarctic blue whale surveys: Augmenting via genetics for close-kin and ordinal age*. I. W.

Caswell, H. (2001). *Matrix population models* (2nd ed.). Sunderland.

Caughley, G. (1966). Mortality patterns in mammals. *Ecology*, *47*(6), 906–918.

Cosens, S. E., & Blouw, A. (2003). Size-and age-class segregation of bowhead whales summering in northern Foxe Basin: A photogrammetric analysis. *Marine Mammal Science*, *19*(2), 284–296.

de Magalhaes, J. P., Costa, J., & Church, G. M. (2007). An analysis of the relationship between metabolism, developmental schedules, and longevity using phylogenetic independent contrasts. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *62*(2), 149–160.

Dennis, C. (2006). Conservation at a distance: A gentle way to age. *Nature*, *442*(7102), 507–508.

Doak, D. F., & Morris, W. (1999). Detecting population-level consequences of ongoing environmental change without long-term monitoring. *Ecology*, *80*(5), 1537–1551.

Dunshea, G., Duffield, D., Gales, N., Hindell, M., Wells, R. S., & Jarman, S. N. (2011). Telomeres as age markers in vertebrate molecular ecology. *Molecular Ecology Resources*, *11*(2), 225–235.

Feil, R., & Fraga, M. F. (2012). Epigenetics and the environment: Emerging patterns and implications. *Nature Reviews. Genetics*, *13*(2), 97–109.

Field, A. E., Robertson, N. A., Wang, T., Havas, A., Ideker, T., & Adams, P. D. (2018). DNA methylation clocks in aging: Categories, causes, and consequences. *Molecular Cell*, *71*(6), 882–895.

Foote, A. D., Martin, M. D., Louis, M., Pacheco, G., Robertson, K. M., Sinding, M.-H. S., Amaral, A. R., Baird, R. W., Baker, C. S., Ballance, L., Barlow, J., Brownlow, A., Collins, T., Constantine, R., Dabin, W., Dalla Rosa, L., Davison, N. J., Durban, J. W., Esteban, R., … Morin, P. A. (2019). Killer whale genomes reveal a complex history of recurrent admixture and vicariance. *Molecular Ecology*, *28*(14), 3427–3444.

Ford, J. K. B., Ellis, G. M., & Balcomb, K. C. (2000). *Killer whales: The natural history and genealogy of Orcinus orca in British Columbia and Washington state* (2nd ed.). B.C., UBC Press.

Ford, J. K. B., Ellis, G. M., Barrett-Lennard, L. G., Morton, A. B., Palm, R. S., & Balcomb, K. C. (1998). Dietary specialization in two sympatric populations of killer whales (*Orcinus orca*) in coastal British Columbia and adjacent waters. *Canadian Journal of Zoology*, *76*, 1456–1471.

Ford, M. J., Hanson, M. B., Hempelmann, J. A., Ayres, K. L., Emmons, C. K., Schorr, G. S., Baird, R. W., Balcomb, K. C., Wasser, S. K., & Parsons, K. M. (2011). Inferred paternity and male reproductive success in a killer whale (*Orcinus orca*) population. *Journal of Heredity*, *102*(5), 537–553.

Ford, M. J., Parsons, K., Ward, E., Hempelmann, J., Emmons, C. K., Bradley Hanson, M., Balcomb, K. C., & Park, L. K. (2018). Inbreeding in an endangered killer whale population. *Animal Conservation*, *21*(5), 423–432.

Frasier, T. R., Petersen, S. D., Postma, L., Johnson, L., Heide-Jørgensen, M. P., & Ferguson, S. H. (2020). Abundance estimation from genetic mark-recapture data when not all sites are sampled: An example with the bowhead whale. *Global Ecology and Conservation*, *22*, e00903.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22.

Fujiwara, M., & Caswell, H. (2001). Demography of the endangered North Atlantic right whale. *Nature*, *414*(6863), 537–541.

Gabriele, C. M., Lockyer, C., Straley, J. M., Jurasz, C. M., & Kato, H. (2010). Sighting history of a naturally marked humpback whale (Megaptera novaeangliae) suggests ear plug growth layer groups are deposited annually. *Marine Mammal Science*, *26*(2), 443–450.

Garde, E., Heide-Jørgensen, M. P., Hansen, S. H., Nachman, G., & Forchhammer, M. C. (2007). Age-specific growth and remarkable longevity in narwhals (Monodon monoceros) from West Greenland as estimated by aspartic acid racemization. *Journal of Mammalogy*, *88*(1), 49–58.

George, J., Bada, J., Zeh, J., Scott, L., Brown, S., O'Hara, T., & Suydam, R. (1999). Age and growth estimates of bowhead whales (*Balaena mysticetus*) via aspartic acid racemization. *Canadian Journal of Zoology*, *77*, 571–580.

George, J., Follmann, E., Zeh, J., Sousa, M., Tarpley, R., Suydam, R., & Horstmann-Dehn, L. (2011). A new way to estimate the age of bowhead whales (*Balaena mysticetus*) using ovarian corpora counts. *Canadian Journal of Zoology*, *89*(9), 840–852.

George, J. C., & Thewissen, J. (2020). *The bowhead whale: Balaena mysticetus: Biology and human interactions*. Academic Press.

Gutman, D., Rivkin, E., Fadida, A., Sharvit, L., Hermush, V., Rubin, E., Kirshner, D., Sabin, I., Dwolatzky, T., & Atzmon, G. (2020). Exceptionally long-lived individuals (ELLI) demonstrate slower aging rate calculated by DNA methylation clocks as possible modulators for healthy longevity. *International Journal of Molecular Sciences*, *21*(2), 615.

Heide-Jørgensen, M. P., Garde, E., Nielsen, N., & Andersen, O. (2012). A note on biological data from the hunt of bowhead whales in West Greenland 2009-2011. *Journal of Cetacean Research and Management*, *12*, 329–333.

Herman, D., Matkin, C., Ylitalo, G., Durban, J., Hanson, M., Dahlheim, M., Straley, J., Wade, P., Tilbury, K., Boyer, R., Pearce, R., & Krahn, M. (2008). Assessing age distributions of killer whale, *Orcinus orca*, populations from the composition of endogenous fatty acids in their outer blubber layers. *Marine Ecology Progress Series*, *372*, 289–302.

Herman, D., Ylitalo, G., Robbins, J., Straley, J., Gabriele, C., Clapham, P., Boyer, R., Tilbury, K., Pearce, R., & Krahn, M. (2009). Age determination of humpback whales Megaptera novaeangliae rough blubber fatty acid compositions of biopsy samples. *Marine Ecology Progress Series*, *392*, 277–293.

Hohn, A. A., Scott, M. D., Wells, R. S., Sweeney, J. C., & Irvine, A. B. (1989). Growth layers in teeth from known-age, free-ranging bottlenose dolphins. *Marine Mammal Science*, *5*(4), 315–342.

Holmes, E., Fritz, L., York, A., & Sweeney, K. (2007). Age-structured modeling reveals long-term declines in the natality of western Steller Sea lions. *Ecological Applications*, *17*(8), 2214-2232.A.

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), R115.

Horvath, S., Gurven, M., Levine, M. E., Trumble, B. C., Kaplan, H., Allayee, H., Ritz, B. R., Chen, B., Lu, A. T., Rickabaugh, T. M., Jamieson, B. D., Sun, D., Li, S., Chen, W., Quintana-Murci, L., Fagny, M., Kobor, M. S., Tsao, P. S., Reiner, A. P., … Assimes, T. L. (2016). An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biology*, *17*(1), 171.

Horvath, S., Haghani, A., Peng, S., Hales, E. N., Zoller, J. A., Raj, K., Larison, B., Robeck, T. R., Petersen, J. L., Bellone, R. R., & Finno, C. J. (2022). DNA methylation aging and transcriptomic studies in horses. *Nature Communications*, *13*(1), 40.

Horvath, S., Haghani, A., Zoller, J. A., Raj, K., Belov, K., Hogg, C. J., & Sears, K. E. (2022). Epigenetic clock and methylation studies in marsupials: Opossums, Tasmanian devils, kangaroos, and wallabies. *GeroScience*, *44*, 1825–1845.

Horvath, S., Lu, A. T., Haghani, A., Zoller, J. A., Li, C. Z., Lim, A. R., Brooke, R. T., Raj, K., Serres-Armero, A., Dreger, D. L., Hogan, A. N., Plassais, J., & Ostrander, E. A. (2022). DNA methylation clocks for dogs and humans. *Proceedings of the National Academy of Sciences*, *119*(21), e2120887119.

Horvath, S., Pirazzini, C., Bacalini, M. G., Gentilini, D., Blasio, A. M., & Delledonne, M. (2015). Decreased epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring. *Aging*, *7*, 1159–1170.

Horvath, S., & Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews. Genetics*, *19*, 371–384.

Horvath, S., Zoller, J. A., Haghani, A., Jasinska, A. J., Raj, K., Breeze, C. E., Ernst, J., Vaughan, K. L., & Mattison, J. A. (2021). Epigenetic clock and methylation studies in the rhesus macaque. *GeroScience*, *43*, 2441–2453.

Horvath, S., Zoller, J. A., Haghani, A., Lu, A. T., Raj, K., Jasinska, A. J., Mattison, J. A., & Salmon, A. B. (2021). DNA methylation age analysis of rapamycin in common marmosets. *GeroScience.*, *43*, 2413–2425.

Hunt, K. E., Moore, M. J., Rolland, R. M., Kellar, N. M., Hall, A. J., Kershaw, J., Raverty, S. A., Davis, C. E., Yeates, L. C., Fauquier, D. A., Rowles, T. K., & Kraus, S. D. (2013). Overcoming the challenges of studying conservation physiology in large whales: A review of available methods. *Conservation Physiology*, *1*(1), cot006.

Izzo, C., Hamer, D. J., Bertozzi, T., Donnellan, S. C., & Gillanders, B. M. (2011). Telomere length and age in pinnipeds: The endangered Australian sea lion as a case study. *Marine Mammal Science*, *27*(4), 841–851.

Jarman, S. N., Polanowski, A. M., Faux, C. E., Robbins, J., De Paoli-Iseppi, R., Bravington, M., & Deagle, B. E. (2015). Molecular biomarkers for chronological age in animal ecology. *Molecular Ecology*, *24*(19), 4826–4847.

Jasinska, A. J., Haghani, A., Zoller, J. A., Li, C. Z., Arneson, A., Ernst, J., Kavanagh, K., Jorgensen, M. J., Mattison, J. A., Wojta, K., Choi, O.-W., DeYoung, J., Li, X., Rao, A. W., Coppola, G., Freimer, N. B., Woods, R. P., & Horvath, S. (2021). Epigenetic clock and methylation studies in vervet monkeys. *GeroScience*, *44*(2), 699–717.

John, C., & Bockstoce, J. R. (2008). Two historical weapon fragments as an aid to estimating the longevity and movements of bowhead whales. *Polar Biology*, *31*(6), 751–754.

Jones, M. J., Goodman, S. J., & Kobor, M. S. (2015). DNA methylation and healthy human aging. *Aging Cell*, *14*, 924–932.

Jylhava, J., Pedersen, N. L., & Hagg, S. (2017). Biological age predictors. *eBioMedicine*, *21*, 29–36.

Koski, W. R., Davis, R. A., Miller, G. W., & Withrow, D. E. (1992). Growth rates of bowhead whales as determined from low-level aerial photogrammetry. *Report of the International Whaling Commission*, *42*, 491–499.

Koski, W. R., Zeh, J., Mocklin, J., Davis, A. R., Rugh, D. J., George, J. C., & Suydam, R. (2010). Abundance of Bering-Chukchi-Beaufort bowhead whales (Balaena mysticetus) in 2004 estimated from photo-identification data. *Journal of Cetacean Research and Management*, *11*, 89–99.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559.

Larison, B., Pinho, G. M., Haghani, A., Zoller, J. A., Li, C. Z., Finno, C. J., Farrell, C., Kaelin, C. B., Barsh, G. S., Wooding, B., Robeck, T. R., Maddox, D., Pellegrini, M., & Horvath, S. (2021). Epigenetic models developed for plains zebras predict age in domestic horses and endangered equids. *Communications Biology*, *4*(1), 1412.

Lemaître, J. F., Rey, B., Gaillard, J. M., Régis, C., Gilot-Fromont, E., Débias, F., Duhayer, J., Pardonnet, S., Pellerin, M., Haghani, A., Zoller, J. A., Li, C. Z., & Horvath, S. (2022). DNA methylation as a tool to explore ageing in wild roe deer populations. *Molecular Ecology Resources*, *22*(3), 1002–1015.

Lockyer, C. (1984). Age determination by means of the earplug in baleen whales. *Report of the International Whaling Commission*, *34*, 692–696.

Lu, A. T., Z. Fei, A. Haghani, T. R. Robeck, J. A. Zoller, C. Z. Li, J. Zhang, J. Ablaeva, D. M. Adams, J. Almunia, R. Ardehali, A. Arneson, C. S. Baker, K. Belov, P. Black, D. T. Blumstein, E. K. Bors, C. E. Breeze, R. T. Brooke, J. L. Brown, A. Caulton, J. M. Cavin, I. Chatzistamou, H. Chen, P. Chiavellini, O.-W. Choi, S. Clarke, J. DeYoung, C. Dold, C. K. Emmons, S. Emmrich, C. G. Faulkes, S. H. Ferguson, C. J. Finno, J.-M. Gaillard, E. Garde, V. N. Gladyshev, V. Gorbunova, R. G. Goya, M. J. Grant, E. N. Hales, M. B. Hanson, M. Haulena, A. N. Hogan, C. J. Hogg, T. A. Hore, A. J. Jasinska, G. Jones, E. Jourdain, O. Kashpur, H. Katcher, E. Katsumata, V. Kaza, H. Kiaris, M. S. Kobor, P. Kordowitzki, W. R. Koski, B. Larison, S.-G. Lee, Y. C. Lee, M. Lehmann, J.-F. Lemaitre, A. J. Levine, C. Li, X. Li, D. T. Lin, N. Macoretta, D. Maddox, C. O. Matkin, J. A. Mattison, J. Mergl, J. J. Meudt, K. Mozhui, A. Naderi, M. Nagy, P. Narayan, P. W. Nathanielsz, N. B. Nguyen, C. Niehrs, A. G. Ophir, E. A. Ostrander, P. O'Tierney Ginn, K. M. Parsons, K. C. Paul, M. Pellegrini, G. M. Pinho, J. Plassais, N. A. Prado, B. Rey, B. R. Ritz, J. Robbins, M. Rodriguez, J. Russell, E. Rydkina, L. L. Sailer, A. B. Salmon, A. Sanghavi, K. M. Schachtschneider, D. Schmitt, T. Schmitt, L. Schomacher, L. B. Schook, K. E. Sears, A. Seluanov, D. Shanmuganayagam, A. Shindyapina, K. Singh, I. Sinha, R. G. Snell, E. Soltanmaohammadi, M. L. Spangler, L. Staggs, K. J. Steinman, V. J. Sugrue, B. Szladovits, M. Takasugi, E. C. Teeling, M. J. Thompson, B. Van Bonn, S. C. Vernes, D. Villar, H. V. Vinters, M. C. Wallingford, N. Wang, R. K. Wayne, G. S. Wilkinson, C. K. Williams, R. W. Williams, X. W. Yang, B. G. Young, B. Zhang, Z. Zhang, P. Zhao, Y. Zhao, J. Zimmermann, W. Zhou, J. Ernst, K. Raj and S. Horvath (2021). "universal DNA methylation age across mammalian tissues." bioRxiv: 2021.2001.2018.426733.

Lubetkin, S., Zeh, J., & George, J. (2012). Statistical modeling of baleen and body length at age in bowhead whales (*Balaena mysticetus*). *Canadian Journal of Zoology*, *90*(8), 915–931.

Marcoux, M., Lesage, V., Thiemann, G. W., Iverson, S. J., & Ferguson, S. H. (2015). Age estimation of belugas, Delphinapterus leucas, using fatty acid composition: A promising method. *Marine Mammal Science*, *31*(3), 944–962.

Matkin, C. O., Ellis, G. M., Olesiuk, P. F., & Saulitis, E. L. (1999). Association patterns and inferred genealogies of resident killer whales, *Orcinus orca*, in Prince William sound, Alaska. *Fishery Bulletin*, *97*, 900–919.

Matkin, C. O., Ward Testa, J., Ellis, G. M., & Saulitis, E. L. (2014). Life history and population dynamics of southern Alaska resident killer whales (*Orcinus orca*). *Marine Mammal Science*, *30*(2), 460–479.

Mayne, B., Berry, O., Davies, C., Farley, J., & Jarman, S. (2019). A genomic predictor of lifespan in vertebrates. *Scientific Reports*, *9*(1), 17866.

Mayne, B., Berry, O., & Jarman, S. (2021). Optimal sample size for calibrating DNA methylation age estimators. *Molecular Ecology Resources*, *21*(7), 2316–2323.

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., & Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, *28*, 495–501.

Mosnier, A., Doniol-Valcroze, T., Gosselin, J.-F., Lesage, V., & Hammill, M. O. (2014). An age structured Bayesian populatio model for St. Lawrence Estuary beluga (*Delphinapterus leucas*) L. Measures and.

Nielsen, N. H., Garde, E., Heide-Jørgensen, M. P., Lockyer, C. H., Ditlevsen, S., Òlafsdóttir, D., & Hansen, S. H. (2013). Application of a novel method for age estimation of a baleen whale and a porpoise. *Marine Mammal Science*, *29*(2), E1–E23.

Olesiuk, P., Ellis, G., & Ford, J. (2005). *Life history and population dynamics of resident killer whales Orcinus orca in the coastal waters of British Columbia*. Nanaimo.

Olesiuk, P. F., Bigg, M. A., & Ellis, G. M. (1990). Life history and population dynamics of resident killer whales (*Orcinus orca*) in the coastal waters of British Columbia and Washington state. *Rep. Int. Whal. Commn. Special Issue*, *12*, 209–244.

Olsen, E., & Sunde, J. (2002). Age determination of minke whales (Balaenoptera acutorostrata) using the aspartic acid racemization technique. *Sarsia*, *87*(1), 1–8.

Olsen, M., Berube, M., Robbins, J., & Palsboll, P. (2012). Empirical evaluation of humpback whale telomere length estimates; quality control and factors causing variability in the singleplex and multiplex qPCR methods. *BMC Genetics*, *13*(77), 1471–2156.

Parsons, K. M., Durban, J. W., & Claridge, D. E. (2003). Comparing two alternative methods for genetic sampling of small cetaceans. *Marine Mammal Science*, *19*(1), 224–231.

Petkovich, D. A., Podolskiy, D. I., Lobanov, A. V., Lee, S.-G., Miller, R. A., & Gladyshev, V. N. (2017). Using DNA methylation profiling to evaluate biological age and longevity interventions. *Cell Metabolism*, *25*(4), 954–960 e956.

Pirotta, E., Booth, C. G., Costa, D. P., Fleishman, E., Kraus, S. D., Lusseau, D., Moretti, D., New, L. F., Schick, R. S., Schwarz, L. K., Simmons, S. E., Thomas, L., Tyack, P. L., Weise, M. J., Wells, R. S., & Harwood, J. (2018). Understanding the population consequences of disturbance. *Ecology and Evolution*, *8*(19), 9934–9946.

Pleskach, K., Hoang, W., Chu, M., Halldorson, T., Loseto, L., Ferguson, S. H., & Tomy, G. T. (2016). Use of mass spectrometry to measure aspartic acid racemization for ageing beluga whales. *Marine Mammal Science*, *32*(4), 1370–1380.

Polanowski, A. M., Robbins, J., Chandler, D., & Jarman, S. N. (2014). Epigenetic estimation of age in humpback whales. *Molecular Ecology Resources*, *14*(5), 976–987.

Raj, K., Szladovits, B., Haghani, A., Zoller, J. A., Li, C. Z., Black, P., Maddox, D., Robeck, T. R., & Horvath, S. (2021). Epigenetic clock and methylation studies in cats. *GeroScience.*, *43*, 2363–2378.

Razin, A., & Cedar, H. (1991). DNA methylation and gene expression. *Microbiology and Molecular Biology Reviews*, *55*(3), 451–458.

Riesch, R., Barrett-Lennard, L. G., Ellis, G. M., Ford, J. K. B., & Deecke, V. B. (2012). Cultural traditions and the evolution of reproductive isolation: Ecological speciation in killer whales? *Biological Journal of the Linnean Society*, *106*(1), 1–17.

Robeck, T. R., Fei, Z., Haghani, A., Zoller, J. A., Li, C. Z., Steinman, K. J., Dirocco, S., Staggs, L., Schmitt, T., Osborn, S., Montano, G., Rodriguez, M., & Horvath, S. (2021). "multi-tissue methylation clocks for age and sex estimation in the common bottlenose dolphin." Frontiers in marine. *Science*, *8*, 713373.

Robeck, T. R., Fei, Z., Lu, A. T., Haghani, A., Jourdain, E., Zoller, J. A., Li, C. Z., Steinman, K. J., DiRocco, S., Schmitt, T., Osborn, S., Van Bonn, B., Katsumata, E., Mergl, J., Almunia, J., Rodriguez, M., Haulena, M., Dold, C., & Horvath, S. (2021). Multi-species and multi-tissue methylation clocks for age estimation in toothed whales and dolphins. *Commun Biol*, *4*(1), 642.

Rosa, C., George, J., Zeh, J., Botta, O., Zauscher, M., Bada, J., & O'Hara, T. (2004). *Update on age estimation of bowhead whales (Balaena mysticetus) using aspartic acid racemization* (p. 15). IWC Scientific Committee.

Rosa, C., Zeh, J., Craig George, J., Botta, O., Zauscher, M., Bada, J., & O'Hara, T. M. (2013). Age estimates based on aspartic acid

racemization for bowhead whales (*Balaena mysticetus*) harvested in 1998–2000 and the relationship between racemization rate and body temperature. *Marine Mammal Science*, *29*(3), 424–445.

Ross, P., Ellis, G., Ikonomou, M., Barrett-Lennard, L., & Addison, R. (2000). High PCB concentrations in free-ranging pacific killer whales, *Orcinus orca*: Effects of age, sex and dietary preference. *Marine Pollution Bulletin*, *40*(6), 504–515.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: A laboratory manual* (2nd ed.). New York.

Saulitis, E., Matkin, C., Barrett-Lennard, L., Heise, K., & Ellis, G. (2000). Foraging strategies of sympatric killer whale (*Orcinus orca*) populations in Prince William sound, Alaska. *Marine Mammal Science*, *16*(1), 94–109.

Schweder, T., Sadykova, D., Rugh, D., & Koski, W. (2010). Population estimates from aerial photographic surveys of naturally and variably marked bowhead whales. *Journal of Agricultural, Biological, and Environmental Statistics*, *15*(1), 1–19.

Stubbs, T. M., Bonder, M. J., Stark, A.-K., Krueger, F., von Meyenn, F., Stegle, O., & Reik, W. (2017). Multi-tissue DNA methylation age predictor in mouse. *Genome Biology*, *18*(1), 1–14.

Tanabe, A., Shimizu, R., Osawa, Y., Suzuki, M., Ito, S., Goto, M., Pastene, L. A., Fujise, Y., & Sahara, H. (2020). Age estimation by DNA methylation in the Antarctic minke whale. *Fisheries Science*, *86*(1), 35–41.

Thompson, M. J., Chwiałkowska, K., Rubbi, L., Lusis, A. J., Davis, R. C., Srivastava, A., Korstanje, R., Churchill, G. A., Horvath, S., & Pellegrini, M. (2018). A multi-tissue full lifespan epigenetic clock for mice. *Aging (Albany NY)*, *10*(10), 2832–2854.

Wang, T., Tsui, B., Kreisberg, J. F., Robertson, N. A., Gross, A. M., Yu, M. K., Carter, H., Brown-Borg, H. M., Adams, P. D., & Ideker, T. (2017). Epigenetic aging signatures in mice livers are slowed by dwarfism, calorie restriction and rapamycin treatment. *Genome Biology*, *18*(1), 57.

Watt, C. A., Stewart, B. E., Loseto, L., Halldorson, T., & Ferguson, S. H. (2020). Estimating narwhal (Monodon monoceros) age using tooth layers and aspartic acid racemization of eye lens nuclei. *Marine Mammal Science*, *36*(1), 103–115.

Waugh, D. A., Suydam, R. S., Ortiz, J. D., & Thewissen, J. (2018). Validation of growth layer group (GLG) depositional rate using daily incremental growth lines in the dentin of beluga (Delphinapterus leucas (Pallas, 1776)) teeth. *PLoS One*, *13*(1), e0190498.

Wilkinson, G. S., Adams, D. M., Haghani, A., Lu, A. T., Zoller, J., Breeze, C. E., Arnold, B. D., Ball, H. C., Carter, G. G., & Cooper, L. N. (2021). DNA methylation predicts age and provides insight into exceptional longevity of bats. *Nature Communications*, *12*(1), 1–13.

Young, B. G., Koski, W. R., Kilabuk, R., Watt, C. A., Ryan, K. P., & Ferguson, S. H. (2022). Collaborative field research using drones for whale photo-identification studies in Cumberland sound, Nunavut. *Drone Systems and Applications*, *10*(1), 256–265.

Zhang, Q., Vallerga, C. L., Walker, R. M., Lin, T., Henders, A. K., Montgomery, G. W., He, J., Fan, D., Fowdar, J., & Kennedy, M. (2019). Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. *Genome Medicine*, *11*(1), 1–11.

Zhou, W., Triche, T. J., Jr., Laird, P. W., & Shen, H. (2018). SeSAMe: Reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Research*, *46*(20), e123.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Parsons, K. M., Haghani, A., Zoller, J. A., Lu, A. T., Fei, Z., Ferguson, S. H., Garde, E., Hanson, M. B., Emmons, C. K., Matkin, C. O., Young, B. G., Koski, W. R., & Horvath, S. (2023). DNA methylation-based biomarkers for ageing long-lived cetaceans. *Molecular Ecology Resources*, *23*, 1241–1256. https://doi.org/10.1111/1755-0998.13791