CpG methylation frequency of TET2, GRIA2 and CDKN2A genes in the North Atlantic fin whale varies with age and between populations

AUTHORS

Raquel García-Vernet[1,2,*], Berta Martín[3], Miguel A. Peinado[3], Gísli Víkingsson[4], Marta Riutort[2], Alex Aguilar[1]

[1]Institute of Biodiversity Research (IRBio) and Department of Evolutionary Biology, Ecology and Environmental Sciences. Faculty of Biology. University of Barcelona. Barcelona. Spain.

[2]Institute of Biodiversity Research (IRBio) and Department of Genetics, Microbiology and Statistics. Faculty of Biology. University of Barcelona. Barcelona. Spain.

[3]Germans Trias i Pujol Research Institute (IGTP), Program of Predictive and Personalized Medicine of Cancer, Badalona, Spain.

[4]Marine and Freshwater Research Institute, Fornubúðum 5, 220 Hafnarfjörður, Reykjavík, Iceland

***Correspondence:** Raquel García-Vernet (Telephone number: + 34 93 402 1453; Barcelona, Spain).

Email: rgarciaver@ub.edu

ABSTRACT

Recovery rates for baleen whales that were decimated by exploitation vary between species and populations. Age determination is critical for the understanding of recovery trends and population structure, but determining age in free-ranging individuals remains challenging. Recent research has suggested that the methylation level of some genes in skin samples may provide age determinations with accuracy. We selected nine CpG sites from three genes (*TET2*, *CDKN2A* and *GRIA2*) and analyzed them in 40 skin samples from known-age individuals pertaining to two different populations of fin whales from the North Atlantic. We observed significant correlations with age in five CpG sites. We used three of these CpG sites to perform an epigenetic age estimation. Predictions had a standard deviation of 2.94, but regression between observed and predicted ages showed a clear underestimation for older fin whales. For further development, we suggest: i) screening for new CpG sites associated with age that exhibit higher variability between individuals, and ii) including older animals whenever the sampling allows it. We also observed subtle, but significant differences between the two populations studied in one of the CpG sites (TET2_CpG+21). We attributed these differences to genetic differences or to the dissimilar environments that affect both populations.

KEYWORDS: cetaceans, epigenetics, marine mammals, molecular biology, population biology

58 **INTRODUCTION**

59 Most baleen whale species have been subject to intense commercial exploitation
60 for hundreds of years and particularly during the 20th century, when catches
61 dramatically intensified (Clapham & Baker, 2018; Rocha et al., 2014). Whaling
62 led to the depletion of many large whale populations worldwide until the
63 International Whaling Commission (IWC) implemented management regulations
64 that eventually led to the establishment of a moratorium on commercial whaling
65 that started in 1985-86 (Clapham & Baker, 2018). Since then, recovery of
66 harvested populations has been unequal among regions and species and this
67 highlighted the need for case-by-case monitoring and for a better understanding
68 of population biology. In this context, substantial efforts have been invested in
69 determining population structure and abundance, and in developing techniques
70 that, in the absence of catches, might allow the determination of the main
71 biological traits of individuals, such as sex, reproductive status, and age.

72 Age is critical to assess some basic population parameters such as somatic
73 growth rate, physical maturation, age at onset of reproductive activity, longevity,
74 or survival. In baleen whales, diverse techniques have been used to determine
75 age. The one most widely applied, and considered to be the most reliable, has
76 been the determination of the number of laminations or growth layer groups in
77 the earplug, a structure located in the ear canal that is composed of successive
78 depositions of wax and keratin that build up in layers over time. Early research
79 demonstrated that one growth layer group is formed every year in most species
80 (Lockyer, 1984). Counting of growth layer groups has been extensively used to
81 obtain highly accurate age estimates in fin (*Balaenoptera physalus*) and sei
82 whales (*Balaenoptera borealis*), but the technique has not worked well in other

mysticete species such as blue (*Balaenoptera musculus*) or minke whales (*Balaenoptera acutorostrata*) (Lockyer, 1984). Other techniques, such as the degree of racemization of aspartic acid in eye globes (George et al., 1999; Olsen & Sunde, 2002; Yasunaga et al., 2017), the deposition of growth layers in the tympanic bullae (Christensen, 1995), or the occurrence of ridges in baleen plates (Chittleborough, 1959; Lubetkin et al., 2008) have also been attempted to age baleen whales but results have been uneven. In all cases the results obtained had little precision or, as it occurs with baleen plate ridges, the age record was limited to a short period of the whale's life-span.

However, all these methods require samples from internal tissues that are impossible to obtain from living animals, so in recent years efforts have focused on developing techniques that permit age-determination through noninvasive means. Photo-identification studies have proven useful in species with easily recognizable morphological traits, such as humpback whales (*Megaptera novaeangliae*) (Barlow & Clapham, 1997), but age determination of an individual is only possible when it is identified at birth and re-sighted subsequently. As a consequence, application of this technique at the population level is only possible in studies with intensive research effort deployed for protracted periods of time. Moreover, it cannot be applied to species with poor fidelity to summer or winter grounds because re-sightings are sparse, or in those in which individual morphological identification is difficult, conditions that apply to most species of the *Balaenoptera* genus.

Biopsies collected from free-ranging individuals have since long become a ubiquitous technique used to monitor populations of cetaceans because they provide fresh skin samples with low impact on the sampled individuals (Aguilar &

108 Borrell, 1994; Aguilar & Nadal, 1984; Clapham & Mattila, 1993). A number of
109 studies have attempted to use these samples to assess age. Telomere length in
110 skin cells correlates with age in a number of species but does not appear accurate
111 as an age predictor because telomere length at birth is highly variable between
112 individuals and, afterwards, it is affected by environmental and physiological
113 features (Dunshea et al., 2011; Olsen et al., 2014). The ratio between certain fatty
114 acids present in the lipids of the blubber has also been found to be correlated
115 with age in humpback whales (Herman et al., 2009), but the precision found in
116 the age determination was low and, because lipids are to some degree
117 dependent on the diet of whales, its application would require calibration for each
118 population and may vary with time. More promising was the analysis of the
119 methylation levels of selected genes, which in humpback whales appeared to
120 provide high accuracy in age determinations (Polanowski et al., 2014; Riekkola
121 et al., 2018a). However, two of the three genes used in this study were later
122 investigated in Antarctic minke whales and provided dissimilar results (Goto et
123 al., 2020; Tanabe et al., 2020), thus suggesting differences between closely
124 related species.

125 In vertebrates, methylation mainly occurs in the cytosines of cytosine-phosphate-
126 guanine (CpG dinucleotides) and is associated with transcriptional repression of
127 nearby genes (Bird & Wolffe, 1999; Deaton & Bird, 2011; Klose & Bird, 2006).
128 Methylation is not static over time, but its rate changes during development
129 (Smith & Meissner, 2013) and aging (D'Aquila et al., 2013). However, aging is not
130 the only factor that exerts modifications in methylation patterns over time.
131 Epigenetic modifications induced by the environment have been reported in many
132 studies and, in some of them, these modifications have been associated with

133 differences in gene expression (Feil & Fraga, 2012). Such environmentally-
134 induced changes might play a relevant role in adaptation processes, providing
135 heterogeneity between and within populations, even if they are genetically
136 homogenous (Flores et al., 2013). Therefore, it is not unreasonable to expect
137 epigenetic differences driven by the environment between populations that have
138 adapted to distinct conditions. Thus, the rate between methylation levels and age
139 may not only be species-specific, but population-specific or even cohort-specific.

140 In this research we amplified and analyzed nine CpG sites of the three genes
141 previously studied in humpback whales (Polanowski et al., 2014). *TET2* (*ten*
142 *eleven translocation 2*) is an evolutionarily conserved dioxygenase from the *TET*
143 family. In humans, acquired *TET2* disruption and mutations have been related to
144 hematopoietic malignancies (Mullinghan, 2009; Solary et al., 2014). *CDKN2A*
145 (*cyclin dependent kinase inhibitor 2A*) encodes for several tumor suppressor
146 proteins. Mutation and silencing of *CDKN2A* has been linked to several types of
147 human cancers (Foulkes et al., 1997; Zhao et al., 2016). Finally, *GRIA2*
148 (*glutamate receptor Ia2/AMPA2*) encodes for one of the four subunits that
149 compose AMPA receptors, which are glutamate receptors. Hypermethylation of
150 these three genes have been associated with aging in humans (Grönninger et
151 al., 2010; Koch & Wagner, 2011).

152 The analyses were conducted on skin samples from fin whales from two separate
153 populations (off West Iceland and North West Spain). Age of all individuals was
154 determined through counts of growth layer groups in ear plugs that had been
155 previously conducted for population biology assessments. The study had a
156 double aim: a) establishing whether the correlation between age and methylation
157 levels observed in the three genes of the humpback whales previously studied

158 were maintained in fin whales, and b) examining inter-population differences
159 suggestive of environmental impact on the methylation rate of these CpG sites.

160

161 <u>**MATERIALS AND METHODS**</u>

162 **1. Sample collection and age determination**

163 The skin samples analyzed for the study were collected from fin whales caught
164 by commercial whaling operations in two feeding grounds in the North Atlantic,
165 which are managed as separate units by the International Whaling Commission
166 (IWC, 2009). In the first one, located off the northwestern coast of Spain (from
167 now on, referred to as the "Spanish population"), the samples were collected at
168 the Caneliñas station from 20 individuals caught during the 1985 summer season.
169 In the second, located off the western coast of Iceland (from now on, referred to
170 as the "Icelandic population"), the samples were collected at the Hvalur H/F
171 whaling station from 20 individuals caught during the summer of 1986. The skin
172 samples were collected by AA and GV always from the central part of the dorsal
173 region of the body to avoid epigenetic variation due to different sampling positions
174 (Goto et al., 2020). Immediately after collection, samples were frozen at -20ºC
175 and preserved in this condition until analysis.

176 The age of individuals was determined by counting the growth layers present on
177 a longitudinal section of their ear-plug core and following the methods described
178 by Aguilar and Lockyer (1987). Each plug was examined by at least two
179 researchers (AA and GV among others) and, in the case of obtaining different
180 values in the multiple readings, the average of both estimates was used. The
181 researchers that performed the ear plug readings have more than 30 years of

182 experience. Individuals from the Icelandic population ranged between 7 to 27
183 years old, while individuals from the Spanish population ranged from 0 (two
184 fetuses) to 49 years old.

185

186 **2. DNA extraction, bisulfite conversion, amplification and**
187 **pyrosequencing**

188 DNA was isolated using a Speedtools Tissue DNA extraction kit (Biotools) and
189 bisulfite converted using an EZ DNA methylation gold kit (Zymo Research, Irvine,
190 CA, USA). Both procedures were carried out following the corresponding
191 manufacturer's instructions.

192 The genes *TET2*, *CDKN2A,* and *GRIA2* were amplified as in Polanowski et al.
193 (2014). Since fin and humpback whales are closely-related species (Árnason et
194 al., 2018; Nikaido et al., 2006; Nishida et al., 2007; Sasaki et al., 2005), the
195 primers designed by Polanowski et al. (2014) were used for the amplification of
196 the three selected genes. PCR reactions were performed in a final volume of 30
197 µl, with 0.4 µl of Immolase DNA polymerase (Bioline 5U/µl), 3µl of 10x
198 ImmoBuffer, 1.2 µl of 50mM $MgCl_2$ solution, 0.4 µl of each primer at 0.25 µM and
199 6 µl of 0.5 mM dNTPs. When amplifying GRIA2, 6 µl of Betaine solution (Sigma)
200 were added to the reaction to enhance the amplification. Thermocycling
201 conditions for all three genes consisted of an initial step of 10 min at 95ºC,
202 followed by 30 cycles of 30 s at 95ºC, 30 s at 58ºC and 30 s at 72ºC, with a final
203 step of 5 min at 72ºC.

204 All PCR products were visualized using gel electrophoresis. When weak
205 amplification was detected in the agarose gel, samples were amplified again but

206 the annealing temperature was set at 56ºC instead of 58ºC. For 31 samples, PCR
207 reactions were performed twice. Of these, in 7 samples PCR reactions were
208 performed three times. Obtaining replicates of the PCR products allowed us to
209 assess the potential error of the posterior pyrosequencing analysis.

210 PCR products were analyzed using a Pyromark Q96MD pyrosequencer (Qiagen
211 GmbH Hilden, Germany) at the PEBC platform (Bellvitge Biomedical Research
212 Institute, IDIBELL) and the percentage of cytosine methylation at each CpG site
213 was extracted using Pyro Q-CpG Software (Qiagen). To avoid confusion, the
214 same notation as Polanowski et al. (2014) has been followed from now on and
215 the position of the cytosine is indicated relative to the start codon of each gene.
216 Negative or positive values indicate distance (5' or 3', respectively) in base pairs
217 from the start codon.

218 **3. Statistical analysis**

219 **3.1. Preliminary analyses**

220 For each of the cytosine sites, presence of outliers was tested graphically using
221 boxplots. To avoid a possible population effect, for each cytosine site the samples
222 from Spain and Iceland were plotted separately. All the percentage values that
223 fell outside the range of the whiskers were removed (that is, values located below
224 or above the Interquartile Range * 1.5), but one exception was made when all
225 PCR replicates of a same sample appeared as an outlier. In this case, we
226 considered that it was a biological but not a methodological outlier and all values
227 were maintained (Supplementary Information 1). After removal of outliers, the
228 means and standard deviation for each sample and CpG site were calculated.

229 Data were tested for normality (Shapiro-Wilk test) and homoscedasticity (Bartlett
230 test).

231 3.2. Age effect

232 Relation between age and methylation was examined in depth for each individual
233 position through a linear regression model (lm). In this case, models were fit
234 considering the methylation percentage of each CpG position as the dependent
235 variable and age as the covariate. Linear models were fitted for each population
236 as well as for the whole set of samples. When model p-value was significant (less
237 than 0.05), the linear regressions were plotted and, in case of having PCR
238 replicates, standard deviations were plotted as well at each individual point.

239 Samples with missing data were excluded from the following analyses. All CpG
240 sites showing significant relationship with age were considered for being included
241 into a multiple regression model. Multiple linear regressions were fitted only for
242 combinations of CpG sites from different genes, to avoid multicollinearity issues
243 between the independent variables (Polanowski et al., 2014).

244 Precision of multiple regression models was assessed through LOOCV (Leave-
245 One-Out Cross-Validation). The LOOCV approach trains the model using all
246 samples but one, which is excluded to test the model obtained. This procedure is
247 repeated for every observation, and the final result is calculated by taking the
248 mean of all individual calculations. RMSE (Root Mean Square Error) was used to
249 determine the average model prediction error and to decide which of the ten
250 combinations of two or three CpG sites was the best model.

251 3.3. Population and sex effect

252 To assess the effect of other biological factors (sex and population) on
253 methylation percentages, generalized linear models were fitted to the data. Each
254 model was fitted considering each CpG site as the dependent variable, and sex
255 and population were included as fixed factors with interaction. In addition, for
256 *TET2_CpG+21*, we reanalyzed the data through a Student's *t*-test to determine
257 if differences between populations were still significant after extracting the
258 samples of the Spanish population with ages that were not overlapping those of
259 the Icelandic population.

260 Finally, to examine in depth the population effect on the methylation percentages
261 on the CpG sites included in the best multiple regression model, we performed a
262 Principal Component Analysis (PCA). PCA was fitted only for samples without
263 missing data.

264 Statistical analyses were carried out with the IBM SPSS 23 software package or
265 in R 3.5.2. (R Core Team. http:// www.R-project.org).

266 **RESULTS**

267 **Determination of percentage of methylation**

268 In total, nine CpG sites were successfully assayed through pyrosequencing: three
269 for *TET2* (*TET2_CpG+16*, *TET2_CpG+21*, *TET2_CpG+31*), three for *CDKN2A*
270 (*CDKN2A_CpG+297*, *CDKN2A_CpG+303*, *CDKN2A_CpG+309*), and three for
271 *GRIA2* (*GRIA2_CpG+202*, *GRIA2_CpG+188*, *GRIA2_CpG+183*). We were not
272 able to obtain results for the fourth site of *TET2* (*TET2_CpG+58*) or *CDKN2A*
273 (*CDKN2A_CpG+327*), but we obtained two additional sites in the *GRIA2* gene
274 (*GRIA2_CpG+188* and *GRIA2_CpG+183*) that have been previously analyzed in
275 minke whales (Tanabe et al., 2020).

276 We did not obtain results at each CpG site for all samples, getting results for all
277 40 samples only for the two last CpG sites of the gene *CDKN2A*. In the other
278 cases, we obtained results from 36-39 samples, depending on the CpG site
279 analyzed. Total range of variation between samples differed between genes and
280 CpG sites. Differences between the highest and the lowest values of percentage
281 of methylation were larger in *TET2* sites, followed by *GRIA2* and *CDKN2A* (Table
282 1).

283 For all CpG sites, we did not obtain results in all the replicates, so the number of
284 replicates is variable depending on the gene and the CpG position analyzed
285 (Table 1). All replicates were analyzed to calculate standard deviation (SD) and
286 assess the reproducibility of our results. Average SD were variable between
287 genes and CpG sites (Table 1). All CpG sites met the assumptions of normality
288 and homoscedasticity, except *GRIA2_CpG+183* site.

289

290 **Table 1**: Total number of samples analyzed for each CpG site. Range of
291 methylation indicates the lowest and the highest values obtained for different
292 samples in a given CpG site. Diff. between samples indicates the difference
293 between the highest and the lowest value of percentage of methylation for
294 different samples in a given CpG site. Replicated samples column indicates the
295 number of samples that were amplified and pyrosequenced for at least one
296 additional time. Finally, SD between replicas indicates the mean standard
297 deviation between replicates for a given CpG site.

298

299 **<u>Sources of variation in the percentage of methylation: Age</u>**

300 Linear regression models were fit for the whole data set and for each population
301 separately at all CpG sites (Figure 1). When the two populations were analyzed

302 together, linear models showed a significant effect of age on the methylation
303 percentages in five CpG sites: three sites of *CDKN2A* ($p < 0.05$ for
304 *CDKN2A_CpG+297, CDKN2A_CpG+303,* and *CDKN2A_CpG+309*), the second
305 CpG site in *TET2* ($p < 0.05$ for *TET2_CpG+21*), and the second CpG site in
306 *GRIA2* ($p < 0.01$ for *GRIA2_CpG+188*). However, when each location was
307 treated as a distinct data set, significant correlations were only found for the
308 Spanish population in four out of five cases and for the Icelandic population in
309 two out of five cases (Figure 1).

310 **Figure 1**. Linear regressions between age and % of cytosine methylation for each
311 CpG site that showed significant correlation with age. Linear regressions were
312 calibrated for each population and the whole data set (only those with correlations
313 with $p < 0.05$ are shown). When data were available, standard deviations were
314 plotted in their corresponding points. Adjusted R^2 is indicated for each regression
315 line.

316 The 5 CpG sites with significant relationship with age were included in a multiple
317 linear regression model. All possible combinations of two or three CpG sites were
318 assessed with Leave One Out Cross Validation (LOOCV) and compared through
319 their RMSE (Root Mean Square Error). For this comparison we chose RMSE as
320 the indicator (and not MAE, the Mean Absolute Error) because although both
321 MAE and RMSE express average model prediction, RMSE gives higher weight to
322 large errors, which are highly undesirable for age predictions. Five samples had
323 missing data for some of the CpG sites and were excluded from the analyses, so
324 multiple regressions were initially calibrated using 35 samples.

325 The multiple regression model with smallest RMSE had the combination of three
326 CpG sites: *TET2_CpG+21, CDKN2A_CpG+303,* and *GRIA2_CpG+188.* After
327 checking the normality and homoscedasticity of the residuals, we removed an

328 additional sample from the Spanish population that was placed as outlier, so the
329 model was run using the 34 remaining samples (Supplementary information 2).
330 The RMSE of the final model was 5.91 and the MAE was 4.87.

331 The precision of the multiple regression model was additionally assessed by
332 checking the differences between the observed and predicted ages (residuals),
333 that presented a standard deviation of 2.941 and a mean absolute difference of
334 4.264. The 95% prediction interval for the model was 11.84 (Figure 2).

335 **Figure 2**: Regression between estimated and observed ages of the 34 samples
336 included in the multiple regression model. Estimated ages are the results of the
337 LOOCV analysis, which included *TET2_CpG+21, CDKN2A_CpG+303,* and
338 *GRIA2_CpG+188* sites. 95% prediction intervals are shown. The R^2 of the linear
339 regression was 0.488, and the prediction intervals 11.84.

340 Linear regression between known and predicted age is shown in Figure 2. The
341 linear regression had a significant y-intercept of 6.603, indicating that our model
342 initially overestimates young whale's age, but the slope of the regression showed
343 that older whales' ages are highly underestimated. R^2 of the regression was
344 0.488, pointing that an important part of model's variation is due to other unknown
345 factors.

346 **<u>Sources of variation in the percentage of methylation: Populations</u>**

347 Generalized linear models showed no significant effect of sex in the methylation
348 levels of any CpG site ($p > 0.05$), in accordance with previous results. However,
349 a significant effect of population was detected in one CpG site (*TET2_CpG+21*,
350 $p < 0.001$). It should be noted that the age ranges of each population are not
351 identical (7 – 27 years for Iceland population; 0 – 49 years for Spanish
352 population), and in this CpG site a significant effect of age was detected (p <

353 0.05). Hence, all samples with ages that were outside the range of the Icelandic
354 (7 – 27) were removed before reanalyzing possible differences between both
355 populations through a Student's *t*-test. Significant differences were detected
356 again between the Spanish and the Icelandic samples ($p < 0.001$; Figure 1),
357 suggesting that both populations were exhibiting different methylation patterns in
358 this CpG site.

359 **Figure 3.** Boxplot distributions of the percentage of methylation in *TET2_CpG+21*
360 after the removal of outliers, using all individuals ("Full Data set") or only those
361 with overlapping ages in both populations ("With overlapping ages"). The top and
362 bottom boundaries of each box indicate the 75th and 25th quartile values,
363 respectively, and lines within each box represent the 50th quartile values.

364 Finally, results from the PCA showed that the first two components explain most
365 of the variation in the data, with a cumulative proportion of variance of 82.41%.
366 For the first principal component (PC1), *CDKN2A_CpG+303* contributed with
367 49.18% and *GRIA2_CpG+188* contributed with 40.20%. For the second principal
368 component (PC2), the most important variable was *TET2_CpG+21,* which
369 explained the 81.03% of the contribution. PC2 showed differences between the
370 two populations, which were barely separated by PC1 (Figure 4).

371 **Figure 4**: Results from the Principal Component Analysis (PCA). Each population
372 is shown in different colors. For the first component (PC1), shown as X axis,
373 *CDKN2A_CpG+303* and *GRIA2_CpG+188* are the most important variables. For
374 the second component (PC2), shown as Y axis, *TET2_CpG+21* is the most
375 important variable.

376

377 DISCUSSION

378 Although the exact mechanisms that drive the association between gene
379 methylation and age are not totally understood, this correlation is well established
380 and several researchers have attempted to create protocols to infer human age
381 through DNA methylation markers. Most of the efforts have focused on finding
382 epigenetic markers that correlate with age in the forensically relevant tissues,
383 such as blood (Bekaert et al., 2015; Hannum et al., 2013; Naue et al., 2017; Vidal-
384 Bralo et al., 2016; Weidner et al., 2014), saliva (Bocklandt et al., 2011; Eipel et
385 al., 2016), semen (Lee et al., 2015), and skin (Koch & Wagner, 2011). Outcomes
386 of these studies were promising and provided high accuracy when predicting age,
387 and this opened thrilling perspectives for the study and monitoring of other
388 mammals, particularly those which live in the wild and are of restricted access.

389 In recent years, driven by the emergence of new cost-effective techniques, the
390 use of epigenetic analysis in studies of ecology and evolution has experienced a
391 substantial increase (De Paoli-Iseppi et al., 2017; Jarman et al., 2015; Verhoeven
392 et al., 2016). Here we have analyzed CpG sites from three genes, previously
393 applied with dissimilar results to infer age from skin samples of humpback and
394 minke whales (Polanowski et al., 2014; Riekkola et al., 2018; Tanabe et al., 2020;
395 Goto et al., 2020), to study their applicability to a different whale species.

396 Ranges of variation observed in our study between individuals' methylation levels
397 for each CpG site were similar to those observed in previous studies (Polanowski
398 et al., 2014; Tanabe et al., 2020; Goto et al., 2020). For the nine sites analyzed,
399 we detected significant correlations with age in five, a finding that reflects the
400 conserved relation between methylation patterns and aging in mammal species
401 (Booth & Brunet, 2016; Sen et al., 2016). It is well known that methylation and
402 aging are tightly related, since global levels of 5-methylcytosine tend to decrease

403 as aging occurs (Barbot et al., 2002; Bollati et al., 2009; Fuke et al., 2004; Nilsen
404 et al., 2016; Singhal et al.,1987). However, this process of hypomethylation does
405 not occur homogenously in the entire genome, and hypermethylation in
406 promoters of several genes and other specific regions has been reported (Fraga
407 & Esteller, 2007; Issa, 2003). In our study, all five sites with significant results
408 showed positive trends with aging (hypermethylation), including the CpG sites
409 located in the gene *TET2*, thus producing different results than in humpbacks and
410 minke whales but in concordance with those obtained in humans (Grönniger et
411 al., 2010). Although our results indicated significant correlations with age for
412 several CpG sites, trends found in fin whales were less consistent and with
413 smaller R^2 values than those found in humpbacks (Polanowski et al., 2014).

414 Even though significant correlations were found in several CpG sites, there is one
415 remarkable limitation in our study that deserves discussion. Standard deviations
416 between PCR replicates were relatively high (mean around 2%) in some CpG
417 sites, while the differences among individual samples hardly ever exceed 5%.
418 These differences are in accordance with the technical accuracy and
419 reproducibility of pyrosequencing when analyzing samples from different PCR
420 reactions (Kurdyukov & Bullock, 2016; Tost & Gut, 2007). Unfortunately, total
421 variation between individuals was small (Table 1), especially for those sites
422 located in *CDKN2A* and *GRIA2*. This brings forward the necessity to screen for
423 new markers associated with age that exhibit higher variability between
424 individuals in order to reduce the effect of the analytical error. It is likely that the
425 use of this kind of assay in fin whales will be only practicable with reasonable
426 accuracy through the use of Next Generation Sequencing techniques to perform
427 a wider screening, as suggested by Goto et al. (2020).

428 The standard deviation of the mean difference between known and predicted
429 ages was 2.94, similar to the results obtained in bats (Wright et al., 2018) and
430 humpbacks (Polanowski et al., 2014), and lower than in minke whales (Goto et
431 al., 2020). However, the y-intercept and the slope of the regression indicated that
432 our model overestimates the age of young whale's and, importantly, highly
433 underestimates the age of older fin whales, similar to the results obtained by Goto
434 et al. (2020). In addition, when we investigated the correlation between age and
435 each individual CpG site, we found that this relation was not totally maintained
436 when data were split into populations. In four out of five CpG sites, correlation
437 with age was significant in the Spanish population, but only in two out of five CpG
438 sites was there a significant correlation in the Icelandic population.

439 We hypothesize that both problems, the low accuracy for estimating the age of
440 older fin whales and the dissimilar results between populations, may be due to
441 the lack of very old individuals, especially in the Icelandic population. In fin
442 whales, longevity has not been properly estimated, but individuals around 80 - 90
443 years old have been reported (Aguilar & García-Vernet, 2018). In order to obtain
444 robust calibrations it would be highly desirable to analyze a higher number of
445 samples from old individuals, since in our data set most of the oldest animals
446 were 25 – 30 years. Indeed, most of the animals caught both in Spain and Iceland
447 during the 1980's were younger than 30 years old and the most frequent age
448 classes were 5-7 years old in Spain and 4-9 years old in Iceland (Aguilar &
449 Lockyer, 1987; IWC, 2009).

450 While some of the differences observed between the results of the two
451 populations may be explained by the dissimilar age range in the two data sets, in
452 the second CpG site of *TET2* (*TET2_CpG+21*) we detected subtle but significant

453 differences regardless of the age of the animals. Changes in patterns of
454 methylation along individuals' lifetime may not only be affected by intrinsic factors,
455 such as genetic variants, but also by biological and extrinsic factors, such as
456 aging and environmental effects, respectively (Feil & Fraga, 2012; Fraga et al.,
457 2005). In humans, for example, a number of studies have reported correlation
458 between methylation levels in different markers and environmental factors such
459 as sun exposure (Grönniger et al., 2020), short-term changes in diet (Jacobsen
460 et al., 2012), exercise (Barrès et al., 2012), early-life stress factors (Naumova et
461 al., 2012), and socioeconomic status (Lam et al., 2012; Mcguiness et al., 2012).

462 With regard to ecology, epigenetics can provide an answer to rapid adaptations
463 to the environment, promoting a high phenotypic plasticity in front of a constantly
464 changing environment (Verhoeven et al., 2016). Therefore, methylation patterns
465 may play an important role by linking environmental cues and phenotypes. In
466 addition, a potential transgenerational transmission of methylation patterns may
467 have an important role in evolutionary processes (Jablonka & Raz, 2009;
468 Verhoeven et al., 2016).

469 Several researchers have observed correlations between methylation patterns
470 and different environmental exposures. For example, resource availability may
471 shape differences in the epigenome of two groups of wild baboons with different
472 foraging strategies (Lea et al., 2016). Similarly, epigenetic differences have been
473 detected between urban and rural populations of Darwin's finches (McNew et al.,
474 2017). Methylation patterns have been related not only to different foraging
475 strategies, but also to different migratory strategies (Baerwald et al., 2016). Other
476 studies in different species have established associations between methylation
477 patterns and other environmental factors, such as contamination (Nilsen et al.,

478 2016), early life stress (Moghadam et al., 2017) or domestication processes
479 (Koch et al., 2016). Therefore, we suggest that differences detected in
480 *TET2_CpG+21* between Icelandic and Spanish populations could be a result of
481 the distinct environmental cues affecting both populations.

482 In the North Atlantic Ocean, the International Whaling Commission (IWC, 2009)
483 currently considers seven fin whale stocks that have been mainly identified
484 according to the location of the feeding grounds occupied during the summer.
485 Since whales reproduce during the winter in different areas, it is of the highest
486 relevance to establish the degree of mixing between animals that belong to each
487 stock. In general, low levels of genetic divergence have been detected between
488 fin whales sampled in different feeding areas (Pampoulie & Daníelsdóttir, 2013)
489 maybe reflecting some degree of mixing in the breeding grounds. However, fin
490 whales inhabiting waters off W. Iceland and N. W. Spain seem to belong to
491 different breeding grounds, implying a low degree of mixing between both
492 populations. Such segregation appears to be supported by previous studies using
493 different approaches, such as internal tagging (Gunnlaugsson & Sigurjónsson,
494 1989), internal and external morphology (Jover, 1992; Lockyer, 1982;
495 Víkingsson, 1992), stable isotope analysis (Vighi et al., 2016), trace element
496 concentrations (Sanpera et al., 1996), and other miscellaneous sources of
497 information (Víkingsson & Gunnlaugsson, 2006). The stock of fin whales feeding
498 off northwestern Spain is assumed to be composed of animals reproducing in a
499 single breeding ground (IWC, 2009), which although of unidentified location, it
500 may be located off eastern Africa (Vighi et al., 2016). Conversely, the stock
501 summering off western Icelandic appears to be composed of animals wintering in
502 either the central Atlantic (around the Azores) or off eastern Canada (IWC, 2009;

503 Silva et al., 2013). Whatever the case, more studies are needed to determine if
504 differences associated with breeding grounds and migration patterns of the two
505 groups of fin whales are associated not only to a differential genetic pool but also
506 to different environmental pressures that in turn may cause differences in their
507 epigenomic profiling. Again, we suggest that a wider screening of new CpG sites,
508 based on Next Generation Sequencing techniques, would shed some light on this
509 topic.

510 **Acknowledgements**

511 The samples were collected at the whaling stations of IBSA (Spain) and Hvalur
512 H/F (Iceland). The authors are grateful to Christina H. Lockyer for her contribution
513 to the earplug readings, to the personnel of the PEBC platform (Bellvitge
514 Biomedical Research Institute, IDIBELL) where the PCR products were analyzed,
515 and to Dr. Concepción Arenas for comments that improved the statistical
516 analysis. The study was part of project CGL2015-70468-R funded by the
517 Ministerio de Economía y Competitividad of Spain and the Fondo Europeo de
518 Desarrollo Regional (FEDER). R.G-V's work was funded through a Ph.D. APIF
519 scholarship granted by the University of Barcelona. R.G-V wants to thank the
520 Spanish State Public Employment Service (SEPE) for its financial support during
521 the review process.

522

523 **Disclosure statement**

524 MAP is cofounder and equity holder of Aniling, a biotech company with no
525 interests in this paper. MAP lab has received research funding from Celgene. The
526 rest of the authors declare no conflict of interest.

527

528 **Supplementary Information 1**: Methylation values for each CpG site were
529 plotted for each population. Values that fell outside the range of the whiskers
530 were considered outliers. However, when all PCR replicates of a same sample
531 appeared as an outlier, we considered that it was a biological but not a
532 methodological outlier and all values were maintained.

533

534 **Supplementary Information 2**: Regression diagnostic plots for the multiple
535 regression model that included *TET2_CpG+21, CDKN2A_CpG+303,* and
536 *GRIA2_CpG+188*. A panel includes all data, while in B panel we eliminated one
537 outlier (sample 35). Briefly, Residuals *vs* Fitted plot is used to check linear
538 relationship. Normal Q-Q plot is used to check if residuals are normally
539 distributed. Scale-Location plot checks the homocedasticity of the residuals.
540 Finally, Residuals *vs* Leverage plot checks if there are influential outliers (this is,
541 if outliers have an effect in our model). In all plots, sample 35 is placed as outlier.

542

543 **References**

544 Aguilar, A., & Borrell, A. (1994). Assessment of organochlorine pollutants in
545 cetaceans by means of skin and hypodermic biopsies. In M. C. Fossi & C.
546 Leonzio (Eds.), Nondestructive Biomarkers in Vertebrates (pp. 245–267). Boca
547 Ratón (USA): Lewis Publishers.

548 Aguilar, A., & García-Vernet, R. (2018). Fin Whale. In W. F. Perrin, B. Würsig, &
549 J. G. M. Thewissen (Eds.), Encyclopedia of Marine Mammals (Third Edition) (pp.
550 368–3714). https://doi.org/10.1016/B978-0-12-373553-9.00102-4

551 Aguilar, A., & Nadal, J. (1984). Obtención de biopsias hipodérmicas de cetáceos
552 en libertad. Investigación Pesquera, 48(1), 23–29.

553 Árnason, Ú., Lammers, F., Kumar, V., Nilsson, M. A., & Janke, A. (2018). Whole-
554 genome sequencing of the blue whale and other rorquals finds signatures for
555 introgressive gene flow. Science Advances, 4, eaap9873.
556 https://doi.org/10.1126/sciadv.aap9873

557 Baerwald, M. R., Meek, M. H., Stephens, M. R., Nagarajan, R. P., Goodbla, A.
558 M., Tomalty, K. M. H., Thorgaard, G. H., May, B., & Nichols, K. M. (2016).
559 Migration-related phenotypic divergence is associated with epigenetic
560 modifications in rainbow trout. Molecular Ecology, 25(8), 1785–1800.
561 https://doi.org/10.1111/mec.13231

562 Barbot, W., Dupressoir, A., Lazar, V., & Heidmann, T. (2002). Epigenetic
563 regulation of an IAP retrotransposon in the aging mouse: progressive
564 demethylation and de-silencing of the element by its repetitive induction. Nucleic
565 Acids Research, 30(11), 2365–2373. https://doi.org/10.1093/nar/30.11.2365

566 Barlow, J., & Clapham, P. J. (1997). A new birth-interval approach to estimating
567 demographic parameters of humpback whales. Ecology, 78(2), 535–546.
568 https://doi.org/10.1890/0012-9658(1997)078[0535:ANBIAT]2.0.CO;2

569 Barrès, R., Yan, J., Egan, B., Treebak, J. T., Rasmussen, M., Fritz, T., … Zierath,
570 J. R. (2012). Acute exercise remodels promoter methylation in human skeletal
571 muscle. Cell Metabolism, 15(3), 405–411.
572 https://doi.org/10.1016/j.cmet.2012.01.001

573 Bekaert, B., Kamalandua, A., Zapico, S. C., Van De Voorde, W., & Decorte, R.
574 (2015). Improved age determination of blood and teeth samples using a selected

set of DNA methylation markers. Epigenetics, 10(10), 922–930. https://doi.org/10.1080/15592294.2015.1080413

Bird, A. P., & Wolffe, A. P. (1999). Methylation-induced repression— belts, braces, and chromatin. Cell, 99(5), 451–454. https://doi.org/10.1016/S0092-8674(00)81532-9

Bocklandt, S., Lin, W., Sehl, M. E., Sánchez, F. J., Sinsheimer, J. S., Horvath, S., & Vilain, E. (2011). Epigenetic predictor of age. PLoS ONE, 6(6), e14821. https://doi.org/10.1371/Citation

Bollati, V., Schwartz, J., Wright, R., Litonjua, A., Tarantini, L., Suh, H., Sparrow, D., Vokonas, P., & Baccarelli, A. (2009). Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mechanisms of Ageing and Development, 130(4), 234–239. https://doi.org/10.1016/j.mad.2008.12.003

Booth, L. N., & Brunet, A. (2016). The aging epigenome. Molecular Cell, 62(5), 728–744. https://doi.org/10.1016/j.molcel.2016.05.013

Chittleborough, R. G. (1959). Determination of age in the humpback whale, *Megaptera nodosa* (Bonnaterre). Australian Journal of Marine and Freshwater Research, 10, 125–143.

Christensen, I. (1995). Interpretation of growth layers in the periosteal zone of tympanic bulla from minke whales *Balaenoptera acutorostrata*. Developments in Marine Biology, 4(C), 413–423. https://doi.org/10.1016/S0163-6995(06)80043-0

Clapham, P. J., & Baker, C. S. (2018). Modern Whaling. In W. F. Perrin, B. Würsig, & J. G. M. Thewissen (Eds.), Encyclopedia of Marine Mammals (Third Edition) (pp. 1070–1074). https://doi.org/10.1016/b978-0-12-804327-1.00272-7

Clapham, P. J., & Mattila, D. K. (1993). Reactions of Humpback whales to skin biopsy sampling on a West Indies breeding ground. Marine Mammal Science, 9, 382–391.

D'Aquila, P., Rose, G., Bellizzi, D., & Passarino, G. (2013). Epigenetics and aging. Maturitas, 74(2), 130–136. https://doi.org/10.1016/j.maturitas.2012.11.005

De Paoli-Iseppi, R., Deagle, B. E., McMahon, C. R., Hindell, M. A., Dickinson, J. L., & Jarman, S. N. (2017). Measuring animal age with DNA methylation: From humans to wild animals. Frontiers in Genetics, 8, 2010–2017. https://doi.org/10.3389/fgene.2017.00106

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. Genes & Development, 25(10), 1010–1022. https://doi.org/10.1101/gad.2037511.1010

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. https://doi.org/10.1111/j.1755-0998.2010.02976.x

Eipel, M., Mayer, F., Arent, T., Ferreira, M. R. P., Birkhofer, C., Gerstenmaier, U., … Wagner, W. (2016). Epigenetic age predictions based on buccal swabs are

616 more precise in combination with cell type-specific DNA methylation signatures.
617 Aging, 8(5), 1034–1048. https://doi.org/10.18632/aging.100972

618 Feil, R., & Fraga, M. F. (2012). Epigenetics and the environment: emerging
619 patterns and implications. Nature Reviews Genetics, 13(2), 97–109.
620 https://doi.org/10.1038/nrg3142

621 Flores, K. B., Wolschin, F., & Amdam, G. V. (2013). The role of methylation of
622 DNA in environmental adaptation. Integrative and Comparative Biology, 53(2),
623 359–372. https://doi.org/10.1093/icb/ict019

624 Foulkes, W. D., Flanders, T. Y., Pollock, P. M., & Haywardt, N. K. (1997). The
625 CDKN2A (p16) gene and human cancer. Molecular Medicine, 3(1), 5–20.

626 Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L.,
627 Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M.,
628 Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z.,
629 Spector, T. D., Wu, Y., (…), Esteller, M. (2005). Epigenetic differences arise
630 during the lifetime of monozygotic twins. Proceedings of the National Academy
631 of Sciences of the United States of America, 102(30), 10604–10609.
632 https://doi.org/10.1073/pnas.0500398102

633 Fraga, M. F., & Esteller, M. (2007). Epigenetics and aging: the targets and the
634 marks. Trends in Genetics, 23(8), 413–418.
635 https://doi.org/10.1016/j.tig.2007.05.008

636 Fuke, C., Shimabukuro, M., Petronis, A., Sugimoto, J., Oda, T., Miura, K.,
637 Miyazaki, T., Ogura, C., Okazaki Y., & Jinno, Y. (2004). Age related changes in
638 5-methylcytosine content in human peripheral leukocytes and placentas: An
639 HPLC-based study. Annals of Human Genetics, 68(3), 196–204.
640 https://doi.org/10.1046/j.1529-8817.2004.00081.x

641 George, J. C., Bada, J., Zeh, J., Scott, L., Brown, S. E., O'Hara, T., & Suydam,
642 R. (1999). Age and growth estimates of bowhead whales (*Balaena mysticetus*)
643 via aspartic acid racemization. Canadian Journal of Zoology, 77(4), 571–580.
644 https://doi.org/10.1139/z99-015

645 Goto, M., Kitakado, T., & Pastene, L. A. (2020). A preliminary study of epigenetic
646 estimation of age of the Antarctic minke whale *Balaenoptera bonarensis*.
647 Cetacean Population Studies, 2, 5–14.

648 Grönniger, E., Weber, B., Heil, O., Peters, N., Stäb, F., Wenck, H., Korn, B.,
649 Winnefeld, M., Lyko, F. (2010). Aging and chronic sun exposure cause distinct
650 epigenetic changes in human skin. PLoS Genetics, 6(5), 6.
651 https://doi.org/10.1371/journal.pgen.1000971

652 Gunnlaugsson, T., & Sigurjónsson, J. (1989). Analysis of the North-Atlantic fin
653 whale marking data from 1979-1988 with special reference of Iceland. Reports of
654 the International Whaling Commission, 39, 383–388.

655 Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B.,
656 Bibikova, M., Fan, J., Gao, Y., Deconde, R., Chen, M., Rajapakse, I., Friend, S.,
657 Ideker, T., & Zhang, K. (2013). Genome-wide methylation profiles reveal

658 quantitative views of human aging rates. Molecular Cell, 49(2), 359–367.
659 https://doi.org/10.1016/j.molcel.2012.10.016

660 Issa, J. P. (2003). Age-related epigenetic changes and the immune system.
661 Clinical Immunology, 109(1), 103–108. https://doi.org/10.1016/S1521-
662 6616(03)00203-1

663 IWC. (2009). Report of the first intersessional RMP workshop on North Atlantic
664 fin whales. Journal of Cetacean Research and Management, 11(SUPPL.), 425–
665 452.

666 Jablonka, E., & Raz, G. (2009). Transgenerational epigenetic inheritance:
667 prevalence, mechanisms, and implications for the study of heredity and evolution.
668 The Quarterly Review of Biology, 84(2), 131–176. https://doi.org/10.1086/598822

669 Jacobsen, S. C., Brøns, C., Bork-Jensen, J., Ribel-Madsen, R., Yang, B., Lara,
670 E., Hall, E., Calvanese, V., Nilsson, E., Jørgensen, S. W., Mandrup, S., Ling, C.,
671 Fernandez, A. F., Fraga, M.F., & Vaag, A. (2012). Effects of short-term high-fat
672 overfeeding on genome-wide DNA methylation in the skeletal muscle of healthy
673 young men. Diabetologia, 55(12), 3341–3349. https://doi.org/10.1007/s00125-
674 012-2717-8

675 Koch, I. J., Clark, M. M., Thompson, M. J., Deere-Machemer, K. A., Wang, J.,
676 Duarte, L., Gnanadesikan, G. E., Mccoy, E. L., Rubbi, L., Stahler, D. R., Pellegrini,
677 M., Ostrander, E. A., Wayne, R. K., Sinsheimer, J. S., & VonHoldt, B. M. (2016).
678 The concerted impact of domestication and transposon insertions on methylation
679 patterns between dogs and grey wolves. Molecular Ecology, 25(8), 1838–1855.
680 https://doi.org/10.1111/mec.13480

681 Jarman, S. N., Polanowski, A. M., Faux, C. E., Robbins, J., De Paoli-Iseppi, R.,
682 Bravington, M., & Deagle, B. E. (2015). Molecular biomarkers for chronological
683 age in animal ecology. Molecular Ecology, 24(19), 4826–4847.
684 https://doi.org/10.1111/mec.13357

685 Jover, L. (1992). Morphometric differences between Icelandic and Spanish fin
686 whales (*Balaenoptera physalus*). Reports of the International Whaling
687 Commission, 42, 747–750.

688 Klose, R. J., & Bird, A. P. (2006). Genomic DNA methylation: The mark and its
689 mediators. Trends in Biochemical Sciences, 31(2), 89–97.
690 https://doi.org/10.1016/j.tibs.2005.12.008

691 Koch, C. M., & Wagner, W. (2011). Epigenetic aging signature to determine age
692 in different tissues. Aging, 3(1), 1–10. https://doi.org/100395 [pii]

693 Kurdyukov, S., & Bullock, M. (2016). DNA methylation analysis: choosing the right
694 method. Biology, 5(1), 3. https://doi.org/10.3390/biology5010003

695 Lam, L. L., Emberly, E., Fraser, H. B., Neumann, S. M., Chen, E., Miller, G. E., &
696 Kobor, M. S. (2012). Factors underlying variable DNA methylation in a human
697 community cohort. Proceedings of the National Academy of Sciences,
698 109(Supplement 2), 17253–17260. https://doi.org/10.1073/pnas.1121249109

699 Lea, A. J., Altmann, J., Alberts, S. C., & Tung, J. (2016). Resource base
700 influences genome-wide DNA methylation levels in wild baboons (*Papio*
701 *cynocephalus*). Molecular Ecology, 25(8), 1681–1696.
702 https://doi.org/10.1111/mec.13436

703 Lee, H. Y., Jung, S. E., Oh, Y. N., Choi, A., Yang, W. I., & Shin, K. J. (2015).
704 Epigenetic age signatures in the forensically relevant body fluid of semen: A
705 preliminary study. Forensic Science International: Genetics, 19, 28–34.
706 https://doi.org/10.1016/j.fsigen.2015.05.014

707 Lockyer, C. H. (1982). Preliminary investigation of some anatomical characters
708 of fin whale ear plugs collected from different regions of the N.E. Atlantic. Reports
709 of the International Whaling Commission, 32, 101–103.

710 Lockyer, C. H. (1984). Age determination by means of the earplug in baleen
711 whales. Rep. Int. Whaling Comm, 34, 692–696.

712 Lubetkin, S. C., Zeh, J. E., Rosa, C., & George, J. C. (2008). Age estimation for
713 young bowhead whales (*Balaena mysticetus*) using annual baleen growth
714 increments. Canadian Journal of Zoology, 86(6), 525–538.
715 https://doi.org/10.1139/Z08-028

716 Mcguinness, D., Mcglynn, L. M., Johnson, P. C. D., Macintyre, A., Batty, G. D.,
717 Burns, H., Cavanagh, J., Deans, K. A., Ford, A., McConnachie, A., McGinty, A.,
718 McLean, J. S., Millar, K., Packard, C. J., Sattar, N. A., Tannahill, C., Velupillai, Y.
719 N., & Shiels, P. G. (2012). Socio-economic status is associated with epigenetic
720 differences in the pSoBid cohort. International Journal of Epidemiology, 41(1),
721 151–160. https://doi.org/10.1093/ije/dyr215

722 McNew, S. M., Beck, D., Sadler-Riggleman, I., Knutie, S. A., Koop, J. A. H.,
723 Clayton, D. H., & Skinner, M. K. (2017). Epigenetic variation between urban and
724 rural populations of Darwin's finches. BMC Evolutionary Biology, 17(1), 1–14.
725 https://doi.org/10.1186/s12862-017-1025-9

726 Moghadam, H. K., Johnsen, H., Robinson, N., Andersen, Ø., H. Jørgensen, E.,
727 Johnsen, H. K., Bæhr, V. J., & Tveiten, H. (2017). Impacts of early life stress on
728 the methylome and transcriptome of Atlantic Salmon. Scientific Reports, 7(1),
729 5023. https://doi.org/10.1038/s41598-017-05222-2

730 Mullighan, C. G. (2009). TET2 mutations in myelodysplasia and myeloid
731 malignancies. Nature Genetics, 41(7), 766–767. https://doi.org/10.1038/ng0709-
732 766

733 Naue, J., Hoefsloot, H. C. J., Mook, O. R. F., Rijlaarsdam-Hoekstra, L., van der
734 Zwalm, M. C. H., Henneman, P., Kloosterman, A. D., & Verschure, P. J. (2017).
735 Chronological age prediction based on DNA methylation: Massive parallel
736 sequencing and random forest regression. Forensic Science International:
737 Genetics, 31, 19–28. https://doi.org/10.1016/j.fsigen.2017.07.015

738 Naumova, O. Y., Lee, M., Koposov, R., Szyf, M., Dozier, M., & Grigorenko, E. L.
739 (2012). Differential patterns of whole-genome DNA methylation in
740 institutionalized children and children raised by their biological parents.

741 Development and Psychopathology, 24(1), 143–155.
742 https://doi.org/10.1017/S0954579411000605

743 Nikaido, M., Hamilton, H., Makino, H., Sasaki, T., Takahashi, K., Goto, M., Kanda,
744 N., Pastene, L. A., Okada, N. (2006). Baleen whale phylogeny and a past
745 extensive radiation event revealed by SINE insertion analysis. Molecular Biology
746 and Evolution, 23(5), 866–873. https://doi.org/10.1093/molbev/msj071

747 Nilsen, F. M., Parrott, B. B., Bowden, J. A., Kassim, B. L., Somerville, S. E., Bryan,
748 T. A., Bryan, C. E., Lange, T. R., Delaney, J. P., Brunell, A. M., Long, S. E., &
749 Guillette, L. J. (2016). Global DNA methylation loss associated with mercury
750 contamination and aging in the American alligator (*Alligator mississippiensis*).
751 Science of the Total Environment, 545–546, 389–397.
752 https://doi.org/10.1016/j.scitotenv.2015.12.059

753 Nishida, S., Goto, M., Pastene, L. A., Kanda, N., & Koike, H. (2007). Phylogenetic
754 relationships among cetaceans revealed by Y-chromosome sequences.
755 Zoological Science, 24(7), 723–732. https://doi.org/10.2108/zsj.24.723

756 Olsen, E., & Sunde, J. (2002). Age determination of minke whales (*Balaenoptera*
757 *acutorostrata*) using the aspartic acid racemization technique. Sarsia, 87(1), 1–
758 8. https://doi.org/10.1080/003648202753631686

759 Olsen, M. T., Robbins, J., Bérubé, M., Rew, M. B., & Palsbøll, P. J. (2014). Utility
760 of telomere length measurements for age determination of humpback whales.
761 NAMMCO Scientific Publications, 10(2014).
762 https://doi.org/http://dx.doi.org/10.7557/3.3194

763 Pampoulie, C., & Daníelsdóttir, A. K. (2013). Review on the genetic stock
764 structure of North Atlantic fin whales (*Balaenoptera physalus*): Past, present and
765 future. International Whaling Commission Document, SC/65a/RMP, 1–8.

766 Polanowski, A. M., Robbins, J., Chandler, D., & Jarman, S. N. (2014). Epigenetic
767 estimation of age in humpback whales. Molecular Ecology Resources, 14(5),
768 976–987. https://doi.org/10.1111/1755-0998.12247

769 Riekkola, L., Zerbini, A. N., Andrews, O., Andrews-Goff, V., Baker, C. S.,
770 Chandler, D., Childerhouse, S., Clapham, P., Dodémont, R., Donnelly, D.,
771 Friedlaender, A., Gallego, R., Garrigue, C., Ivashchenko, Y., Jarman, S., Lindsay,
772 R., Pallin, L., Robbins, J., Steel, D., Tremlett, J., Vindenes, S., & Constantine, R.
773 (2018a). Application of a multi-disciplinary approach to reveal population
774 structure and Southern Ocean feeding grounds of humpback whales. Ecological
775 Indicators, 89, 455–465. https://doi.org/10.1016/j.ecolind.2018.02.030

776 Rocha, R. C., Clapham, P. J., & Ivashchenko, Y. V. (2014). Emptying the oceans:
777 A summary of industrial whaling catches in the 20th century. Marine Fisheries
778 Review, 76(4), 37–48. https://doi.org/10.7755/MFR.76.4.3

779 Sanpera, C., González, M., & Jover, L. (1996). Heavy metals in two populations
780 of North Atlantic fin whales (*Balaenoptera physalus*). Environmental Pollution,
781 91(3), 299–307.

782 Sasaki, T., Nikaido, M., Hamilton, H., Goto, M., Kato, H., Kanda, N., Pastene, L.
783 A., Cao, Y., Fordyce, R. E., Hasegawa, M., & Okada, N. (2005). Mitochondrial
784 phylogenetics and evolution of mysticete whales. Systematic Biology, 54(1), 77–
785 90. https://doi.org/10.1080/10635150590905939

786 Sen, P., Shah, P. P., Nativio, R., & Berger, S. L. (2016). Epigenetic mechanisms
787 of longevity and aging. Cell, 166(4), 822–839.
788 https://doi.org/10.1016/j.cell.2016.07.050

789 Silva, M. A., Prieto, R., Jonsen, I., Baumgartner, M. F., & Santos, R. S. (2013).
790 North Atlantic blue and fin whales suspend their spring migration to forage in
791 middle latitudes: building up energy reserves for the journey? PLoS ONE, 8(10).
792 https://doi.org/10.1371/journal.pone.0076507

793 Singhal, R. P., Mays-Hoopes, L. L., & Eichhorn, G. L. (1987). DNA methylation in
794 aging of mice. Mechanisms of Ageing and Development, 41(3), 199–210.
795 https://doi.org/10.1016/0047-6374(87)90040-6

796 Smith, Z. D., & Meissner, A. (2013). DNA methylation: roles in mammalian
797 development. Nature Reviews Genetics, 14(3), 204–220.
798 https://doi.org/10.1038/nrg3354

799 Solary, E., Bernard, O. A., Tefferi, A., Fuks, F., & Vainchenker, W. (2014). The
800 Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic
801 diseases. Leukemia, 28(3), 485–496. https://doi.org/10.1038/leu.2013.337

802 Tanabe, A., Shimizu, R., Osawa, Y., Suzuki, M., Ito, S., Goto, M., Pastene, L. A.,
803 Yoshihiro, F., & Sahara, H. (2020). Age estimation by DNA methylation in the
804 Antarctic minke whale. Fisheries Science, 86(1), 35–41.
805 https://doi.org/10.1007/s12562-019-01371-7

806 Tost, J., & Gut, I. G. (2007). DNA methylation analysis by pyrosequencing. Nature
807 Protocols, 2(9), 2265–2275. https://doi.org/10.1038/nprot.2007.314

808 Verhoeven, K. J. F., VonHoldt, B. M., & Sork, V. L. (2016). Epigenetics in ecology
809 and evolution: What we know and what we need to know. Molecular Ecology,
810 25(8), 1631–1638. https://doi.org/10.1111/mec.13617

811 Vidal-Bralo, L., Lopez-Golan, Y., & Gonzalez, A. (2016). Simplified assay for
812 epigenetic age estimation in whole blood of adults. Frontiers in Genetics, 7(JUL),
813 1–7. https://doi.org/10.3389/fgene.2016.00126

814 Vighi, M., Borrell, A., & Aguilar, A. (2016). Stable isotope analysis and fin whale
815 subpopulation structure in the eastern North Atlantic. Marine Mammal Science,
816 32(2), 535–551. https://doi.org/10.1111/mms.12283

817 Víkingsson, G. A. (1992). Morphometrics of fin whales off Iceland and Spain.
818 Report of the Comprehensive Assessment Special Meeting on North Atlantic Fin
819 Whales. Anex D. Reports of the International Whaling Commission, 42, 611.

820 Víkingsson, G., & Gunnlaugsson, T. (2006). Stock structure of fin whales
821 (*Balaenoptera physalus*) in the North Atlantic – Indications from non-genetic data.
822 Doc. IWC/SC/57/PFI3. Available from International Whaling Commission (2005)
823 http://www.iwcoffice.Org/ at https://iwc.int/home.

824 Weidner, C. I., Lin, Q., Koch, C. M., Eisele, L., Beier, F., Ziegler, P., Bauerschlag,
825 D. O., Jöckel, K., Erbel, R., Mühleisen, T. W., Zenke, M., Brümmendorf, T. H., &
826 Wagner, W. (2014). Aging of blood can be tracked by DNA methylation changes
827 at just three CpG sites. Genome Biology, 15(2), R24. https://doi.org/10.1186/gb-
828 2014-15-2-r24

829 Wright, P. G. R., Mathews, F., Schofield, H., Morris, C., Burrage, J., Smith, A.,
830 Dempster, E. L., & Hamilton, P. B. (2018). Application of a novel molecular
831 method to age free-living wild Bechstein's bats. Molecular Ecology Resources,
832 18(6), 1374–1380. https://doi.org/10.1111/1755-0998.12925

833 Yasunaga, G., Pastene, L. A., Bando, T., Hakamada, T., & Fujise, Y. (2017). Age
834 estimation of Antarctic minke whales *Balaenoptera bonaerensis* based on
835 aspartic acid racemization technique. Fisheries Science, 83(6), 947–954.
836 https://doi.org/10.1007/s12562-017-1122-0

837 Zhao, R., Choi, B. Y., Lee, M. H., Bode, A. M., & Dong, Z. (2016). Implications of
838 genetic and epigenetic alterations of CDKN2A (p16INK4a) in cancer.
839 EBioMedicine, 8(127), 30–39. https://doi.org/10.1016/j.ebiom.2016.04.017

840

841