1

2   MR. RICARDO DE PAOLI-ISEPPI (Orcid ID : 0000-0001-7724-9144)

3   DR. BRUCE E DEAGLE (Orcid ID : 0000-0001-7651-3687)

4   DR. SIMON N JARMAN (Orcid ID : 0000-0002-0792-9686)

5

6

7   Article type   : Resource Article

8

9

10   Age estimation in a long-lived seabird (*Ardenna tenuirostris*) using DNA

11   methylation-based biomarkers

12

13   R. De Paoli-Iseppi[1,2*], B. E. Deagle[2], A. M. Polanowski[2], C. R. McMahon[1,3], J. L.

14   Dickinson[4], M. A. Hindell[1,5] and S. N. Jarman[6,7]

15

16

17   [1] Institute for Marine and Antarctic Studies, University of Tasmania, Hobart,

18   Tasmania, Australia.

19   [2] Australian Antarctic Division, Hobart, Tasmania, Australia.

20   [3] Sydney Institute of Marine Science, Sydney, New South Wales, Australia.

21   [4] Cancer, Genetics and Immunology Group, Menzies Institute for Medical Research

22   Tasmania, Hobart, Tasmania, Australia.

23   [5] Antarctic Climate and Ecosystems CRC, Hobart, Tasmania, Australia

24   [6] Trace and Environmental DNA Laboratory, Department of Environment and

25   Agriculture, Curtin University, Perth, WA, Australia.

26   [7] CSIRO Indian Ocean Marine Research Centre, The University of Western Australia,

27   Perth, WA, Australia.

28

29

30   * Corresponding author

31   E-mail: ricardo.depaoliiseppi@utas.edu.au

32

Short title: DNA methylation age estimation in shearwater

Key words: age; birds; DNA methylation; epigenetics; DREAM;

Abstract

Age structure is a fundamental aspect of animal population biology. Age is strongly related to individual physiological condition, reproductive potential and mortality rate. Currently, there are no robust molecular methods for age estimation in birds. Instead, individuals must be ringed as chicks to establish known-age populations, which is a labour intensive and expensive process. The estimation of chronological age using DNA methylation is emerging as a robust approach in mammals including humans, mice and some non-model species. Here we quantified DNA methylation in whole blood samples from a total of 71 known-age Short-tailed shearwaters (*Ardenna tenuirostris*) using digital restriction enzyme analysis of methylation (DREAM). The DREAM method measures DNA methylation levels at thousands of CpG dinucleotides throughout the genome. We identified seven CpG sites with DNA methylation levels that correlated with age. A model based on these relationships estimated age with a mean difference of 2.8 years to known age, based on validation estimates from models created by repeated sampling of training and validation data subsets. Longitudinal observation of individuals re-sampled over 1 or 2 years generally showed an increase in estimated age (6/7 cases). For the first time, we have shown that epigenetic changes with age can be detected in a wild bird. This approach should be of broad interest to researchers studying age biomarkers in non-model species and will allow identification of markers that can be assessed using targeted techniques for accurate age estimation in large population studies.

1 | INTRODUCTION

Understanding the age structure of populations is a key aspect of animal ecology and conservation. Age estimate information can help to determine animal mortality, susceptibility to parasites, reproductive life history and the impact of anthropogenic activities (Froy *et al.* 2013; Gianuca *et al.* 2017; Musick 1999; Scott 1988). However,

65 measuring the chronological age of many wild animals is a difficult task due to the

66 lack of external changes that reflect age. Some animals have quantifiable physical

67 changes as they increase in age, for example, tooth length in deer (Pérez‑Barbería *et*

68 *al.* 2014) and growth rings in fish otoliths (Buckmeier *et al.* 2002; Campana 2001;

69 Gunn *et al.* 2008). However, few of these can be measured without capturing or even

70 killing the animal. The impact and ethics of these interventions on animals is often the

71 subject of debate (Festa-Bianchet *et al.* 2002; Nelson 2002). Other animals can show

72 general changes with life stage, for example, plumage variation in some seabirds

73 (Weimerskirch *et al.* 1989); or larval stage of arthropods and molluscs (Cobb &

74 Wahle 1994; Ernande *et al.* 2003), but these often only provide age information for

75 immature individuals. This lack of accessible chronological age information limits our

76 understanding of many wild animal species and it is only through long term,

77 expensive tracking or marking studies that age data can be collected and used

78 effectively.

79

80 Molecular biomarkers of age have recently been the focus of an increasing number of

81 studies (Ito *et al.* 2018; Maegawa *et al.* 2017; Wright *et al.* 2018). Neither telomere

82 length or DNA damage markers have been successfully used for chronological age

83 estimation in a wild animal population, so there is interest in developing alternative

84 molecular age biomarkers (Dunshea *et al.* 2011; Jarman *et al.* 2015). One promising

85 avenue is measuring epigenetic modification controlling changes in gene expression

86 that occur during animal ageing. Epigenetic regulation of gene expression can occur

87 at several different levels and can include histone modification, non-coding RNA

88 (ncRNA) and DNA methylation (DNAm). DNAm, the addition of a methyl group to a

89 cytosine followed by a guanine (CpG site), has been examined in the most detail and

90 recent evidence supports the use of this epigenetic modification for individual age

91 determination (Hannum *et al.* 2013; Horvath 2013; Vidal-Bralo *et al.* 2016).

92

93 Here, we refer to two types of changes in DNAm with age that could be used to

94 estimate age in wild animals. 'Epigenetic drift' generally refers to broad DNAm

95 signals at sites distributed across the genome, which in mammals, birds and fish has

96 been reported to decline with age (Gryzinska *et al.* 2013; Jakubczak *et al.* 2016;

97 Shimoda *et al.* 2014). Drift signals can also be enriched in CpG islands and enhancers

98   (Slieker *et al.* 2016). 'Clock-type' markers are specific CpG sites that show a strong

99   correlation with known chronological age. Correlations observed in this category can

100   be tissue specific and can involve an increase (hypermethylation), or decrease

101   (hypomethylation) with age (Horvath 2013; Slieker *et al.* 2018). Clock-type CpG age

102   markers have recently been referred to as "age-related DNA methylation positions"

103   (aDMPs) (Lowe *et al.* 2018; Slieker *et al.* 2018). aDMPs are generally located within

104   the promoter or first exon of a gene (Bekaert *et al.* 2015; Grönniger *et al.* 2010;

105   Horvath 2013; Sziráki *et al.* 2018; Zbieć-Piekarska *et al.* 2015). Epigenetic drift is

106   thought to occur due to a decline or imperfect replication of DNAm by an epigenetic

107   maintenance system with increasing age (Horvath 2013; Horvath & Raj 2018).

108   However, the mechanisms for specific 'clock-type' aDMP change have not yet been

109   characterised.

110

111   Very little is known about DNAm in most non-model species, especially birds.

112   Available studies have mostly focused on model species such as the Red junglefowl

113   (*Gallus gallus*) (Gryzinska *et al.* 2013; Hu *et al.* 2013; Li *et al.* 2011) and Japanese

114   quail (*Coturnix japonica*) (Andraszek *et al.* 2014). These studies show a distribution

115   of DNAm in the genome similar to that observed in mammals. Epigenetic drift is the

116   only age-related DNAm change that has been reported in birds. Gryzinska *at al.*

117   (2013) observed DNAm changes between chickens aged between 1 day and 32 weeks

118   using a colorimetric immunoenzymatic based protocol. We have previously reported

119   that the DNAm status of several mammalian clock-type age-related genes were not

120   conserved in homologous regions of a seabird (De Paoli-Iseppi *et al.* 2017).

121

122   Here, we used known age individuals from a long-term study of Short-tailed

123   shearwater (*Ardenna tenuirostris*) to investigate age related changes. The shearwater

124   has high breeding site and partner fidelity and is long-lived, making it an ideal species

125   in which to study population status and chronological ageing in a seabird population.

126   Fisher Island (Tasmania, Australia), is the site of a long-term banding study of this

127   species and as such can be used to collect known age blood and feather samples for

128   the investigation of DNAm and chronological age (Bradley *et al.* 1991). Epigenetic

129   age estimates of seabirds would be particularly valuable for use in population viability

130   analyses and could further our understanding of environmental effects on animal

131   performance or foraging (Velarde & Ezcurra 2018). For the first time, we have used

digital restriction enzyme analysis of methylation (DREAM) to assess DNAm in a

non-model vertebrate. We identified seven aDMPs in DNA extracted from 71 whole

blood samples. A model relating methylation at these aDMPs to age was made and

the precision evaluated using the mean absolute difference (MAD) between the

estimated and known chronological ages. Our study is the first to identify DNAm

changes with chronological age in a wild seabird and will provide a foundation for

further study of age-related DNAm in non-mammalian vertebrates.

2 | METHODS

2.1 | Samples and DNA extraction

In sampling trips between 2015 – 2018, blood samples were collected from adult

(November – December) and chick (March) *A. tenuirostris* from Fisher Island

(40°13'00.7"S 148°14'20.7"E) Tasmania, under Department of Primary Industries,

Parks, Water and Environment (DPIPWE) permit: FA15230 and University of

Tasmania (UTAS) Animal Ethics Committee permits: A14277 and A0016107. Blood

was collected onto Whatman FTA® Micro (WB120210) cards and stored as

previously described (De Paoli-Iseppi *et al.* 2017). DNA was extracted from a 3 mm

punch of immobilised blood using an Epicentre MasterPure™ (MCD85201) DNA

Purification Kit according to the manufacturer's instructions. We examined blood

DNA in two high-throughput sequencing runs of a total of $N = 71$ known-age

individuals. Age was determined by recording the band number of birds first marked

as chicks, and was rounded to whole years as all sampling occurred in a short time

window each year. Run 1 consisted of 35 known-age animals (5 – 21 years old, mean

= 12.14 years). Two individuals aged 8 and 14 years old were replicated within this

run. Run 2 consisted of DNA from 36 additional known-age samples (6 – 21 years

old, mean = 14.18 years). Run 2 contained three technical replicates from Run 1 (6,

12 and 21 years old) and three within-run replicates aged 8, 14 and 21 years old.

Several birds were recaptured in sampling trips in different years allowing us to

perform some limited longitudinal observations (Run 2: $N = 3$ x 2 samples and $N = 4$

x 2 samples at 1 and 2 year resights respectively). In total, $N = 63$ known-age

165 shearwater were used to calibrate the model following removal of replicates. Bird sex

166 was determined by *CHD-1* gene amplification in blood DNA using a previously

167 described method (Faux *et al.* 2014). Sample details for each age group and known

168 age distribution are shown in Table 1 and Supplementary Figure 1 respectively.

169

170 **2.2 | Analysis of genome-wide 'CCCGGG' methylation**

171 We examined DNAm at CpG sites throughout the genome using digital restriction

172 enzyme analysis of methylation (DREAM) of 71 Short-tailed shearwater whole blood

173 DNA samples (Jelinek & Madzo 2016). Briefly, genomic DNA (1 μg) extracted from

174 shearwater blood FTA samples was sequentially cut with two enzymes that recognise

175 the 'CCCGGG' sequence motif in DNA (Figure 1). Methyl-sensitive *SmaI* first cuts

176 only unmethylated sites leaving blunt 5'-GGG ends. Then, *XmaI* cleaves the

177 remaining methylated sites leaving 5'-CCGGG ends. Thus, unique sequences are

178 made for methylated or unmethylated CpG sites. Following this sequential digest,

179 DNA was used to create sequencing libraries using NEBNext Multiplex Oligos for

180 Illumina Index Primer Sets 1 – 3 and standard Illumina protocols. Blunt-end ligation

181 is done using NEBNext adaptor (10 μm) and T4 DNA ligase with hairpin loop

182 cleavage with USER enzyme. Dual size selection for 250 – 450 bp fragments was

183 done using AMPure XP beads. Unique barcodes were then added to DNA from

184 individual samples with 12x rounds of PCR using AmpliTaq Gold DNA Polymerase

185 (see Supplementary Table 1). Individual barcoded samples were analysed for correct

186 library size distribution (250 – 450 bp) using high sensitivity DNA 1000 kits on the

187 Bioanalyzer 2100. Two microliters of each sample was also quantified using a Qubit

188 2.0 to ensure equal volumes were pooled in the final library. Libraries were run at 2 –

189 4 ng/uL on the Illumina NextSeq 500 platform with a 15 – 25% PhiX control at the

190 Ramaciotti Centre for Genomics (UNSW, Sydney, AUS).

191

192

193 **2.3 | Statistical analysis and construction of an age prediction model**

194 *Sequencing data analysis pipeline*

195 Raw DNA sequence reads were run through an in-house data analysis pipeline in the

196 following steps.

197   1. *Quality filtering*. Demultiplexed Fastq sequences were filtered with a

198   maximum expected error (maxee) rate of 0.5 and converted to Fasta format

199   (Edgar & Flyvbjerg 2015).

200   2. *Dereplication*. A database of unique reads from all samples was generated

201   (dereplication) using trimmed sequences and the USEARCH10 command

202   'fastx_uniques' (Edgar 2010), with a min_unique size = 150.

203   3. *Methylated and non-methylated motif databases*. These dereplicated

204   sequences were duplicated to contain the unique sequence with either the 5'-

205   GGG or 5'-CCGGG motif, in separate databases (GG or CC databases).

206   4. *Motif database hits*. Each sample was then compared to each database using

207   the 'usearch-global' command with 97% identity and required an exact match

208   to the first 2 bp of the relevant motif (id_prefix = 2). Hits for each sequence

209   against both methylation databases were recorded.

210   5. *DNAm level calculation*. The methylation level for each sample was then

211   calculated as the count of the methylated signature divided by the total number

212   of hits for a specific CpG marker and the value was recorded between 0 and 1.

213   A value of 0 is unmethylated (i.e. all sequences from that site match the GG

214   sequence generated by methyl-sensitive *SmaI*) and 1 is methylated (i.e. all

215   sequences from that site match the CC sequence generated by *XmaI*).

216

217 Methylation scores were retained for read depths between 20 and 2000 reads. Scores

218 that were calculated outside of this range were converted to a 'NA'. To retain

219 potentially informative markers in the final analysis, markers with less than seven NA

220 values across all samples were imputed using the mean of the remaining non-NA

221 values for the marker. This method ensured that potential age-related markers would

222 not be omitted based on missing scores and that imputed values would have a

223 relatively small effect on any correlations observed. Since variation is required to find

224 correlations with age, we removed markers that had a DNAm standard deviation of

225 less than 5% across all samples. A small run effect was observed, so the mean DNAm

226 difference between run 1 and 2 replicates was used to adjust the score of each marker

227 in run 2.

228

229 *Predictor selection and age estimation model*

230   Markers that passed filtering were then used to fit penalised lasso regularisation paths

231   to each predictor using the R package 'glmnet' (Friedman *et al.* 2010). The penalty

232   value used to select coefficients, lambda 1 standard error (λ1se), was calculated after

233   repeated runs (100x) of the default k-fold cross validation function of glmnet

234   (cv.glmnet, 10-fold) with an alpha = 1 (lasso). This method randomly subsets the data

235   each cycle and assesses the linear relationship between age and DNAm. Following

236   repeated runs of this function a mean λ1se value was generated. The λ1se value

237   generally selects CpG sites for the simplest model with an error similar to the best

238   model (λ minimum), given the cross-validation uncertainty.

239

240   Individual markers that passed the λ1se cut-off were inspected visually using simple

241   linear regression and markers that had an $R^2 < 0.2$ or showed small changes in DNAm

242   range (< 15%) were removed from further analysis. Remaining age-related CpG sites

243   were then incorporated into a multiple linear regression model. To test the selected

244   markers, the original data set was randomly split into 75/25% training ($N = 47$) and

245   test ($N = 16$) data sets respectively. Training set DNAm values for each aDMP were

246   used to create a multiple linear regression model. The model was then tested with

247   remaining samples in the test set. This random sub-sampling method was run for 100

248   iterations. By substituting the calculated methylation values for each of the individual

249   shearwaters used in the training and test sets into the equation, we obtained the

250   predicted epigenetic age. Mean absolute difference (MAD), the uncertainty of age

251   estimates expressed in years, between the known and estimated age was then

252   calculated. The 77 bp sequence following the CG motif was analysed by BLASTn

253   searches of bird genomes available on the NCBI database to identify any regions

254   conserved between species (Altschul *et al.* 1990).

255

256   **2.4 | Global DNA methylation analyses**

257   *Global analysis of 2338 CpG sites using DREAM*

258   The mean DNAm of 2338 CpG sites identified using DREAM were analysed by age

259   group in years as follows: Chicks: 0.12 – 0.15 ($N = 2$), Young breeder: 5 – 9 ($N = 16$),

260   Middle: 10 – 18 ($N = 39$), Old: 19+ ($N = 6$). CpGs were analysed using a one-way

261   ANOVA followed by post-test for multiple comparisons (Tukey's HSD). Mean

262   DNAm differences were calculated in both the chick and young breeder context and

263   analysed as above. Significance was set at $P < 0.05$.

264

265 *Colorimetric DNA methylation analysis*

266 We also measured epigenetic drift in global DNAm using a commercially available

267 methylated DNA quantification assay for relative 5-mC content (Abcam,

268 Colorimetric, ab117128). Briefly, 42 shearwater blood DNA samples (chicks, 5 – 21

269 years old, mean = 10.9 years) were analysed in duplicate, alongside the supplied

270 positive (5 ng) and negative controls. Methylated DNA was captured and detected

271 using diluted (1:1000, 1:2000) 5-mC antibodies. Following the addition of a

272 developing solution, colour change was monitored and quantified at 450 nm (Tecan

273 Spark). Using the mean absorbance values of the duplicates, relative 5-mC for each

274 sample was calculated as follows: (((Sample OD – Negative control OD) / DNA input

275 (ng)) / (((Positive control OD – Negative control OD) x 2) / Positive control input (5

276 ng))) * 100. Analysis of duplicate colorimetric data was done using a one-way

277 ANOVA with Šidák correction for multiple comparisons for each age group in years

278 as above.

279

280

281

282 **3 | RESULTS**

283 **3.1 | Sequencing metrics**

284 Quality analysis of DREAM libraries showed bands in the expected post clean-up

285 range, (range = 194 – 974 bp, mean = 451 bp; Bioanalyzer gel and electropherogram

286 traces are shown in Supplementary Figure 2A – D). A total of 125 million sequences

287 (mean of 1761622 per sample) passed initial bioinformatic QC (maxee = 0.5 and

288 matched restriction site motif; Supplementary Table 2). The sum of reads from

289 sequences with a mean high read depth (> 2000x) represented approximately 6%

290 (mean = 84518 reads) of the total mean sequences per sample. Following filtering and

291 dereplication, we identified 93884 unique sequences that were used to create a

292 database of reference sequences (i.e. markers for specific CpG sites) for sample

293 matching (Supplementary Figure 3). Following the pipeline filtering described, a total

294 of 2338 unique CpGs were used for lasso analysis (glmnet).

295

296 **3.2 | Development and testing of an age prediction model in the Short-tailed**
297 **shearwater**
298 DNAm data from seven CpG sites obtained using DREAM were included in the age
299 prediction model based on our selection criteria (Figure 2A – G). Information on
300 removed CpG sites with weaker age correlations is provided in Supplementary Table
301 3 (e.g. just below our mean λ1se cut-off of 1.2; see Supplementary Figure 4). To
302 investigate potential sex-related DNAm effects in the seven aDMPs used in the age
303 prediction model, separate linear regressions were done for each sex (Supplementary
304 Figure 5A-G). Sex had a significant effect on DNAm age correlation in a single
305 aDMP in isolation (M1801, P = 0.0031, Bonferroni corrected), with males driving the
306 association (Supplementary Figure 5C). However, there was no sex-specific effect
307 when the methylation scores for all seven aDMPs were then used to create the age
308 estimation model (Figure 3, sex regression slopes and diagnostics are shown in
309 Supplementary Figures 5H and 6 respectively). Read depth had a mean of 51x for
310 these CpG sites (Supplementary Figure 7). The MAD between the known and
311 estimated age reports the uncertainty in age estimates expressed in years. Following
312 repeated cross-validation, the seven aDMP age assay provided epigenetic age
313 estimates in training subsamples with a MAD of 2.34 ± 1.73 (SD) years (mean R^2 =
314 .605, range: 0.46 – 0.72) (Supplementary Figure 8A). In the validation test
315 subsamples, the age estimates had an increased error; across all age estimates MAD =
316 2.81 ± 2.08 years (mean R^2 = .404, range: 0.03 – 0.80) (Supplementary Figure 8B).
317 The significant y-intercept of 5.13 indicated that the predicted ages were
318 overestimated for chicks and young birds and underestimated for older individuals,
319 and may indicate a non-linear relationship. The training set MAD ranged from 1.17 –
320 6.25 years, whilst in the test set MAD ranged from 1.58 – 7.86 years. The MADs for
321 each year and grouped age, as described in the methods, are shown in Figure 4.
322 Between run replicates for seven age-related CpG sites showed a mean DNAm score
323 difference of 11.29% (range: 3.79 – 12.80%) and 6.83% (range: 4.21 – 11.04%) pre-
324 and post-run adjustment respectively (Supplementary Table 4). Within run replicates
325 showed a mean absolute difference in DNAm of 8.65% (range: 5.18 – 11.91%) for the
326 age-related markers.
327
328 **3.3 | Biomarker sequence and gene conservation**

329　The seven aDMPs we identified were used to search for conserved regions in

330　available bird genomes and scaffolds using BLASTn. Of these seven markers, four

331　had low E values and > 50% query cover indicating a reasonable match with a known

332　sequence in the available avian databases (Table 2). Marker 1071 matched with the

333　*G3BP1* region in the Zebra finch (*Taeniopygia guttata*) genome, however the query

334　cover was only slightly above 50%. Marker 1934 had a 100% query cover match with

335　an uncharacterised locus in the Mallard (*Anas platyrhynchos*) genome. Marker 2083

336　matched against scaffold 4695 in the North Island brown kiwi (*Apteryx australis*

337　*mantelli*) genome. Finally, marker 3169 had a 100% query cover match to the *DHH*

338　gene in several species with the top hit to the Eurasian blue tit (*Cyanistes caeruleus*)

339　genome.

340

341　**3.4 | Longitudinal observations of DNA methylation in resighted individuals**

342　We observed that 6/7 (85%) age estimates for resighted individuals sampled 1 or 2

343　years apart showed the expected positive increase in predicted age relative to their

344　known age from leg bands (Figure 5). At many individual aDMPs the longitudinal

345　samples did not follow the expected DNAm trend (Supplementary Figure 9A-B).

346　However, when combined into the model, only one individual showed a negative

347　change in estimated age from two samples taken at 15 and 17 years of age. The mean

348　absolute difference between estimated and known age for 2-year resights was 0.74

349　years ($N = 8$) and 0.87 years ($N = 6$) for 1-year resights.

350

351　**3.5 | DNA methylation of 2338 CpGs using DREAM assay**

352　We show that a large proportion of the 2338 CpG sites that passed the filtering cut off

353　are highly methylated, with 50.2% of CpGs showing DNAm levels greater than 80%

354　across all ages (Figure 6A). We also observed a small, but non-significant linear

355　change in DNAm from young animals to old. The mean DNAm was .712, .724, .725

356　and .729, for chicks, young breeders, middle and old birds respectively (Figure 6B).

357　The difference in mean DNAm, relative to chick levels, for each individual CpG site

358　is shown in Figure 6C. This shows that relative to older birds, chicks are less

359　methylated at low DNAm levels (approximately < 10%) and more methylated at high

360　DNAm levels (approximately > 90%).

361

362　**3.6 | Global 5-mC using colorimetric assay**

363 Relative 5-mC was quantified against the supplied 5 ng positive control. Global blood

364 DNAm levels of the Short-tailed shearwater were combined into age groups as

365 described in the methods. Chicks and young breeders showed similar relative 5-mC

366 levels, (mean = 0.725, N = 4 and mean = 0.727, N = 15 respectively). Both of these

367 groups had slightly higher relative 5-mC than that observed in middle-aged birds

368 (mean = 0.614, N = 17) and old birds (mean = 0.498, N = 5). Following adjustment

369 for multiple comparisons, no significant differences were observed between the age

370 groups (Figure 7).

371

372

373 **4 | DISCUSSION**

374

375 Seabirds exhibit little or no external physical changes with age and there are currently

376 no reliable biomarkers of chronological age in most long-lived seabirds beyond

377 fledging. The identification of an accurate age biomarker would be a substantial

378 advance in our ability to understand seabird age-related demographics. Seabird age

379 estimation using molecular methods is currently not possible. DNAm changes with

380 age have been reported for both wild and model mammalian species in several tissues,

381 indicating that DNAm age biomarkers may be useful in birds. In this study, we

382 quantified the DNAm profile of known-age Short-tailed shearwaters using digital

383 restriction enzyme analysis of methylation (DREAM). We present evidence for

384 DNAm changes with chronological age in seven CpG sites.

385

386 **4.1 | Age related biomarkers in birds**

387

388 Previous bird ageing research has focused primarily on telomere length assays and

389 pentosidine accumulation in collagen. Studies of terminal telomere restriction

390 fragments (TRFs) have shown that telomere length can shorten with increasing age

391 and that the rate of change corresponds to lifespan in several species (Bize *et al.* 2009;

392 Juola *et al.* 2006; Tricola *et al.* 2018). However, this trend is not consistent amongst

393 all birds, with some species showing increases in TRF with age, as in the Leach's

394 storm-petrel (*Oceanodroma leucorhoa*) (Haussmann *et al.* 2003), and no decline in

395 length, or both as reported for the Magellanic penguin (*Spheniscus magellanicus*)

396 (Cerchiara *et al.* 2017). For individuals in some avian species, change in telomere

397 length can be tracked longitudinally and correlate with reproductive timing, however

398 the use of TRF for cross-sectional analysis of age has yet to be demonstrated (Bauer

399 *et al.* 2018).

400

401 Pentosidine is a less frequently studied age biomarker for birds. It forms cross-links

402 between amino acid residues in collagen and accumulates with age in birds (Fallon *et*

403 *al.* 2006; Iqbal *et al.* 1999). Pentosidine has been shown to accumulate in a linear

404 fashion in terrestrial birds and some seabirds including California gulls (*Larus*

405 *californicus*) (Chaney Jr *et al.* 2003) and Double-crested cormorants (*Phalocrocorax*

406 *auritus*) (Fallon *et al.* 2006). This technique has yielded age estimates with a precision

407 of 2 – 4 years in wild birds (Chaney Jr *et al.* 2003; Fallon *et al.* 2006; Rattiste *et al.*

408 2015). However, in a study of another long-lived seabird, the Bridled tern

409 (*Onychoprion anaethetus*), no correlation between pentosidine levels and age was

410 found (Labbé 2017). It is not known how pentosidine levels may respond to the

411 effects of changing biological age or environmental stressors. As a result of the

412 limited success in age estimation by these methods, our research aimed to build upon

413 recent successes in mammals by assessing DNAm estimates of age in the Short-tailed

414 shearwater.

415

416 We previously established that specific aDMPs from mammals were not conserved in

417 the shearwater (De Paoli-Iseppi *et al.* 2017). We therefore sought to identify bird-

418 specific aDMPs or a global DNAm signature associated with age using DREAM of

419 whole blood samples. This is the first epigenetic age assay developed for use in a

420 seabird, and one of the few used in a wild species. Using the DREAM method, we

421 identified seven novel aDMPs in shearwaters. Following repeated cross-validation of

422 our known-age samples to train and test the age-estimation model, we reported a test-

423 set MAD for all ages of 2.81 ± 2.08 years. The linear relationship with age in these

424 CpG sites is not as strong as those reported for whales (Polanowski *et al.* 2014) or

425 dogs (Thompson *et al.* 2017), but was similar to that reported for a bat species

426 (Wright *et al.* 2018). We also observed variation in MAD for different age classes,

427 with birds aged 5 – 9 years and 19+ providing less accurate age estimates compared to

428 other groups (Figure 4A). Additionally, the significant Y-intercept in our model

429 (Figure 3) causes an overestimation of age in younger individuals. A single marker

430 (M1801) showed evidence for male driven DNAm age correlation. Due to the reduced

431 sample size when comparing by sex only, more known-age samples would be

432 required to confirm the lack of association in females and ideally, whole genome

433 information could determine if this marker is located on a sex chromosome.

434

435 However, the biggest limitation in developing our model was the low number of

436 young non-breeding bird samples that we could capture in the field, which hinders our

437 understanding of the rate of DNAm change between chicks and early breeders (5 – 9

438 years), and with more samples, this may be correctable in future. The shearwaters

439 studied here typically do not return to their island of birth until their first year of

440 breeding at age five (Bradley *et al.* 1991; Bradley *et al.* 1989). However, for unknown

441 reasons we did not recover many individuals in the 5 – 9 early breeder age range. The

442 larger DNAm variability in these young animals could be due to the stressful effects

443 of the first year of breeding. Shearwaters lay one of the largest eggs relative to body

444 mass of all seabirds, and individuals face challenges including incubatory fasting and

445 intermittent foraging (Wooller *et al.* 1990). Additionally, both migration and

446 parenthood can reduce body condition, and evidence suggests that these birds may

447 undergo intermittent breeding if an individual determines its body condition is too

448 low (Bradley *et al.* 2000).

449

450 Despite some uncertainty in ages estimated with our model, this approach could

451 discriminate between relevant age classes (e.g. young and old adults). These

452 epigenetic age estimates, in combination with other parameters including sex and

453 weight, could be used to examine the effect of climate change on population viability

454 (Lee 2017). Recent studies also highlight other areas in which estimated age data

455 could be informative, including post-pest eradication monitoring of island-breeding

456 seabird populations (Brooke *et al.* 2018), parasite load in the Blue tit (Aguilar *et al.*

457 2016) and modelling the impacts of longline fisheries on effective population size

458 (Cortés *et al.* 2018; Mills & Ryan 2005).

459

460 Obtaining a broad age range of samples from long-lived, known-age birds is difficult

461 as extensive banding studies are rare. Whilst the Fisher Island shearwater population

462 has been followed for several decades, the youngest and oldest adult individuals we

463 recovered were 5 and 21 years old respectively. The oldest individual, at 21 years old,

464   represents a little over half of the maximum reported lifespan for this species of 39

465   years. However, research on age dependent survival on Fisher Island birds shows few

466   animals living beyond 25 years post first breeding, which would place our oldest

467   individual at closer to 70% of the expected lifespan of approximately 30 years (Baylis

468   *et al.* 2018; Bradley *et al.* 1989). The relationship we have observed with age should

469   be investigated further for older individuals, however previous studies in mammals

470   have primarily shown linear correlations with age (Maegawa *et al.* 2010; Polanowski

471   *et al.* 2014; Spiers *et al.* 2016). Although no recaptures were made within the 1 – 4

472   year age range, as these non-breeding birds are not at the nesting sites, the

473   relationship of adults to the DNAm level of the chicks suggest birds at these ages will

474   have a similar trend to the rest of the calibration range.

475

476   We quantified 'epigenetic drift' in DNAm levels observed across all 2338 CpG sites

477   included in our analysis. We did not identify a significant trend with chronological

478   age. However, we did observe some interesting differences between young and old

479   age groups at the lower and upper limits of DNAm. In contrast to mammalian and the

480   only other bird study, we found no clear trend of DNA hypomethylation in older

481   animals compared to that in younger individuals (Gaudet *et al.* 2003; Gryzinska *et al.*

482   2016; Portela & Esteller 2010). The lack of statistical significance could be due to the

483   analysis of this relatively small subset of total CpGs in the bird genome.

484   Immunoenzymatic analyses of chicken 5-mC levels have shown decreased global

485   methylation with age (Gryzinska *et al.* 2013). Using the same method, we found no

486   relationship between relative 5-mC levels and age in 42 known-age shearwater whole

487   bloods. However, we observed a non-significant trend towards decreasing

488   methylation across age groups. Our study of age-related global DNAm in shearwaters

489   is only the second of this phenomenon in birds and further work will be required to

490   determine if this approach could be suitable for age estimation in other bird species.

491

492   **4.2 | Measuring methylation in non-model organisms**

493

494   Despite the identification of several thousand unique CpG sites using the DREAM

495   method, the 20x read depth requirement for DNAm calculation resulted in the

496   exclusion of many sites from further analysis. A small percentage of the total reads

497   was also lost to repetitive elements. There is little doubt that as technologies improve

498   sequencing depths will increase, and direct analysis of CpG DNAm will be possible,

499   (Rand *et al.* 2017; Slatko *et al.* 2018). Improvements in bioinformatics will also help

500   to validate DNAm markers and predict age in large data sets (Vidaki *et al.* 2017). The

501   DREAM technique has been used previously to identify DNAm changes following

502   compound exposure in zebrafish embryos (Bouwmeester *et al.* 2016) and caloric

503   restriction in mice (Maegawa *et al.* 2017). A similar method, EpiRADSeq, also uses a

504   methylation sensitive restriction enzyme (*HpaII*) and NGS to quantify DNAm in CpG

505   sites (Schield *et al.* 2016). This technique differs from DREAM in that only a single

506   methylation sensitive enzyme is used in combination with a frequent cutter (*PstI*).

507   *HpaII* recognises a 'CCGG' motif, which is likely to lead to higher genomic coverage

508   of CpG sites due to increased cut frequency. However, DNAm scores generated using

509   this method are relative to the count of unmethylated EpiRADSeq reads only. This is

510   avoided when using a dual methylation sensitive digest as in DREAM, as reads are

511   generated for both methylated and unmethylated CpGs (Jelinek & Madzo 2016).

512   Reduced representation bisulphite sequencing (RRBS) can also be used to quantify

513   CpG DNAm, but does require a higher quantity of initial genomic DNA (Meissner *et*

514   *al.* 2005). The output of these various techniques depends upon several molecular,

515   platform and bioinformatic factors and choices, which is discussed in detail elsewhere

516   (O'Leary *et al.* 2018). Our results now show that the DREAM method can also be

517   used to quantify global DNAm and screen for aDMPs in non-model animals. The

518   primary limitation in applying this method is the high read depth required per CpG

519   site, particularly in organisms with relatively high quantities of repetitive DNA. This

520   makes it cost-prohibitive as a method for applying to population-wide samples, but

521   certainly effective as a screening method for identifying aDMPs. Once aDMPs are

522   identified by DREAM, targeted DNAm scoring assays could be developed to reduce

523   costs for high-throughput applications.

524

525   An additional limitation to the simple analysis of shearwater DREAM and indeed

526   most non-model NGS data, is the limited genomic resources available for further

527   analyses. Multiplex restriction site PCR (mRS-PCR) could be used to obtain both up

528   and downstream sequence around an aDMP of interest (Sarkar *et al.* 1993; Weber *et*

529   *al.* 1998). This method can generate larger reference sequences for use in targeted

530   bisulphite assays such as EpiTYPER, pyrosequencing or other NGS based techniques

531   (Ehrich *et al.* 2005). More sequence information may also result in more accurate

532   comparative genomic analyses against bird genomes that are currently undergoing

533   scaffold alignment. The genes *DHH* and *G3BP1* were identified as conserved age-

534   related sequences from our data and these could be used in future as part of a targeted

535   gene assay in shearwater (Table 2, M1071 and M3169). Whilst we cannot comment

536   on any potential functional effects of DNAm, *DHH* and *G3BP1* encode for signalling

537   molecules in cell morphogenesis and a DNA-unwinding enzyme, respectively. Two

538   other markers also showed high conservation with other bird species, however these

539   hits were either unassigned (M2083) or uncharacterised (M1934). These factors limit

540   our ability to identify biomarkers that have the potential to be used in closely related

541   species, and design a cost-effective, targeted age assay.

542

543

544   **5 | CONCLUSIONS**

545

546   This study demonstrates that seabird age estimates can be generated from a DNA

547   methylation age assay. This minimally invasive method could be used to produce age

548   estimates for Short-tailed shearwaters from chicks to 21 years old. This is the first

549   time an epigenetic assay has been applied to a wild seabird and could be used in

550   future to estimate population age structure. Further refinement of this method could

551   result in the identification, validation and use of target genes, similar to that in

552   mammals, for related seabird species and see wider use for monitoring and

553   conservation.

554

555   **Data and code availability**

556   DREAM count data, adjusted DNAm values for 2338 CpGs, fasta pipeline and

557   variable selection R scripts used in this publication have been deposited in the

558   Dryad Digital Repository at [doi: 10.5061/dryad.n4h3672]. Sample details and

559   raw Illumina sequence data (FASTQ) are available from NCBI/SRA using

560   accession:

561   PRJNA507458, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA507458.

562

563

564   **Acknowledgements**

Author contributions

All authors conceived the ideas and designed methodology; RDP, AMP, CRM and MAH collected samples; RDP and AMP did the genetics laboratory work; RDP, BED and SNJ analysed the data; RDP led manuscript writing. All authors contributed to drafts and gave final approval for publication.

References

Aguilar JRd, Westerdahl H, Puente JMdl, *et al.* (2016) MHC‑I provides both quantitative resistance and susceptibility to blood parasites in blue tits in the wild. *Journal of Avian Biology* **47**, 669-677.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Andraszek K, Gryzińska M, Wójcik E, Knaga S, Smalec E (2014) Age-dependent change in the morphology of nucleoli and methylation of genes of the nucleolar organizer region in Japanese quail (*Coturnix japonica*) model (Temminck and Schlegel, 1849)(Galliformes: Aves). *Folia Biologica* **62**, 293-300.

Bauer CM, Graham JL, Abolins-Abols M, *et al.* (2018) Chronological and Biological Age Predict Seasonal Reproductive Timing: An Investigation of Clutch Initiation and Telomeres in Birds of Known Age. *The American Naturalist* **191**, 777-782.

Baylis SM, Sunnucks P, Clarke R (2018) A model for first‑estimates of species‑specific, age‑specific mortality from centralized band‑recovery databases. *Ecosphere* **9**, e02136.

597   Bekaert B, Kamalandua A, Zapico SC, Van de Voorde W, Decorte R (2015)
598   Improved age determination of blood and teeth samples using a selected
599   set of DNA methylation markers. *Epigenetics* **10**, 922-930.

600   Bize P, Criscuolo F, Metcalfe NB, Nasir L, Monaghan P (2009) Telomere dynamics
601   rather than age predict life expectancy in the wild. *Proceedings of the*
602   *Royal Society of London B: Biological Sciences* **276**, 1679-1683.

603   Bouwmeester MC, Ruiter S, Lommelaars T, *et al.* (2016) Zebrafish embryos as a
604   screen for DNA methylation modifications after compound exposure.
605   *Toxicology and Applied Pharmacology* **291**, 84-96.

606   Bradley J, Skira I, Wooller R (1991) A long‐term study of Short‐tailed
607   Shearwaters (*Puffinus tenuirostris*) on Fisher Island, Australia. *Ibis* **133**,
608   55-61.

609   Bradley J, Wooller R, Skira I (2000) Intermittent breeding in the short‐tailed
610   shearwater Puffinus tenuirostris. *Journal of Animal Ecology* **69**, 639-650.

611   Bradley J, Wooller R, Skira I, Serventy D (1989) Age-dependent survival of
612   breeding Short-tailed Shearwaters (*Puffinus tenuirostris*). *The Journal of*
613   *Animal Ecology*, 175-188.

614   Brooke MdL, Bonnaud E, Dilley B, *et al.* (2018) Seabird population changes
615   following mammal eradications on islands. *Animal Conservation* **21**, 3-12.

616   Buckmeier DL, Irwin ER, Betsill RK, Prentice JA (2002) Validity of otoliths and
617   pectoral spines for estimating ages of channel catfish. *North American*
618   *Journal of Fisheries Management* **22**, 934-942.

619   Campana S (2001) Accuracy, precision and quality control in age determination,
620   including a review of the use and abuse of age validation methods. *Journal*
621   *of Fish Biology* **59**, 197-242.

622   Cerchiara JA, Risques RA, Prunkard D, *et al.* (2017) Telomeres shorten and then
623   lengthen before fledging in Magellanic penguins (*Spheniscus*
624   *magellanicus*). *Aging (Albany NY)* **9**, 487.

625   Chaney Jr RC, Blemings KP, Bonner J, Klandorf H (2003) Pentosidine as a
626   measure of chronological age in wild birds. *The Auk* **120**, 394-399.

627   Cobb JS, Wahle RA (1994) Early life history and recruitment processes of clawed
628   lobsters. *Crustaceana* **67**, 1-25.

629   Cortés V, García-Barcelona S, González-Solís J (2018) Sex-and age-biased
630   mortality of three shearwater species in longline fisheries of the
631   Mediterranean. *Marine Ecology Progress Series* **588**, 229-241.

632   De Paoli-Iseppi R, Polanowski AM, McMahon C, *et al.* (2017) DNA methylation
633   levels in candidate genes associated with chronological age in mammals
634   are not conserved in a long-lived seabird. *PLoS One* **12**, e0189181.

635   Dunshea G, Duffield D, Gales N, *et al.* (2011) Telomeres as age markers in
636   vertebrate molecular ecology. *Molecular Ecology Resources* **11**, 225-235.

637   Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST.
638   *Bioinformatics* **26**, 2460-2461.

639   Edgar RC, Flyvbjerg H (2015) Error filtering, pair assembly and error correction
640   for next-generation sequencing reads. *Bioinformatics* **31**, 3476-3482.

641   Ehrich M, Nelson MR, Stanssens P, *et al.* (2005) Quantitative high-throughput
642   analysis of DNA methylation patterns by base-specific cleavage and mass
643   spectrometry. *Proceedings of the National Academy of Sciences* **102**,
644   15785-15790.

645   Ernande B, Clobert J, McCombie H, Boudry P (2003) Genetic polymorphism and
646   trade‐offs in the early life‐history strategy of the Pacific oyster,
647   *Crassostrea gigas* (Thunberg, 1795): a quantitative genetic study. *Journal*
648   *of Evolutionary Biology* **16**, 399-414.

649   Fallon JA, Cochrane RL, Dorr B, Klandorf H (2006) Interspecies comparison of
650   pentosidine accumulation and its correlation with age in birds. *The Auk*
651   **123**, 870-876.

652   Faux CE, McInnes JC, Jarman SN (2014) High-throughput real-time PCR and melt
653   curve analysis for sexing Southern Ocean seabirds using fecal samples.
654   *Theriogenology* **81**, 870-874.

655   Festa-Bianchet M, Blanchard P, Gaillard J-M, Hewison AM (2002) Tooth
656   extraction is not an acceptable technique to age live ungulates. *Wildlife*
657   *Society Bulletin* **30**, 282-283.

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

660   Froy H, Phillips RA, Wood AG, Nussey DH, Lewis S (2013) Age‐related variation
661   in reproductive traits in the wandering albatross: evidence for terminal
662   improvement following senescence. *Ecology Letters* **16**, 642-649.

663   Gaudet F, Hodgson JG, Eden A*, et al.* (2003) Induction of tumors in mice by
664   genomic hypomethylation. *Science* **300**, 489-492.

665   Gianuca D, Phillips RA, Townley S, Votier SC (2017) Global patterns of sex-and
666   age-specific variation in seabird bycatch. *Biological Conservation* **205**, 60-
667   76.

668   Grönniger E, Weber B, Heil O*, et al.* (2010) Aging and chronic sun exposure cause
669   distinct epigenetic changes in human skin. *PLoS Genet* **6**, e1000971.

670   Gryzinska M, Blaszczak E, Strachecka A, Jezewska-Witkowska G (2013) Analysis
671   of age-related global DNA methylation in chicken. *Biochemical Genetics*
672   **51**, 554-563.

673   Gryzinska M, Jakubczak A, Listos P*, et al.* (2016) Association between body
674   weight and age of dogs and global DNA methylation. *Medycyna*
675   *Weterynaryjna* **72**, 64-67.

676   Gunn JS, Clear NP, Carter TI*, et al.* (2008) Age and growth in southern bluefin
677   tuna, *Thunnus maccoyii* (Castelnau): direct estimation from otoliths,
678   scales and vertebrae. *Fisheries Research* **92**, 207-220.

679   Hannum G, Guinney J, Zhao L*, et al.* (2013) Genome-wide methylation profiles
680   reveal quantitative views of human aging rates. *Molecular Cell* **49**, 359-
681   367.

682   Haussmann MF, Winkler DW, O'Reilly KM*, et al.* (2003) Telomeres shorten more
683   slowly in long-lived birds and mammals than in short–lived ones.
684   *Proceedings of the Royal Society of London B: Biological Sciences* **270**,
685   1387-1392.

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

688   Horvath S, Raj K (2018) DNA methylation-based biomarkers and the epigenetic
689   clock theory of ageing. *Nature Reviews Genetics*, 1.

690   Hu Y, Xu H, Li Z*, et al.* (2013) Comparison of the genome-wide DNA methylation
691   profiles between fast-growing and slow-growing broilers. *PLoS One* **8**,
692   e56411.

693   Iqbal M, Probert LL, Alhumadi NH, Klandorf H (1999) Protein glycosylation and
694   advanced glycosylated endproducts (AGEs) accumulation: an avian
695   solution? *Journals of Gerontology Series A: Biomedical Sciences and*
696   *Medical Sciences* **54**, B171-B176.

697   Ito H, Udono T, Hirata S, Inoue-Murayama M (2018) Estimation of chimpanzee
698   age based on DNA methylation. *Scientific Reports* **8**, 9998.

699   Jakubczak A, Listos P, Dudko P, Abramowicz K, Jeżewska-Witkowska G (2016)
700   Association between body weight and age of dogs and global DNA
701   methylation. *Medycyna Weterynaryjna* **72**, 64-67.

702   Jarman SN, Polanowski AM, Faux CE*, et al.* (2015) Molecular biomarkers for
703   chronological age in animal ecology. *Molecular Ecology* **24**, 4826-4847.

704   Jelinek J, Liang S, Lu Y*, et al.* (2012) Conserved DNA methylation patterns in
705   healthy blood cells and extensive changes in leukemia measured by a new
706   quantitative technique. *Epigenetics* **7**, 1368-1378.

707   Jelinek J, Madzo J (2016) DREAM: A Simple Method for DNA Methylation
708   Profiling by High-throughput Sequencing. *Chronic Myeloid Leukemia:*
709   *Methods and Protocols*, 111-127.

710   Juola FA, Haussmann MF, Dearborn DC, Vleck CM (2006) Telomere shortening in
711   a long-lived marine bird: cross-sectional analysis and test of an aging tool.
712   *The Auk* **123**, 775-783.

713   Labbé A (2017) *Effects of age on reproduction and chick rearing in bridled terns*
714   *(Onychoprion anaethetus) at Penguin Island, Western Australia*, Murdoch
715   University.

716   Lee CT (2017) Elasticity of population growth with respect to the intensity of
717   biotic or abiotic driving factors. *Ecology* **98**, 1016-1025.

718   Li Q, Li N, Hu X*, et al.* (2011) Genome-wide mapping of DNA methylation in
719   chicken. *PLoS One* **6**, e19428.

720   Lowe R, Barton C, Jenkins CA*, et al.* (2018) Ageing-associated DNA methylation
721   dynamics are a molecular readout of lifespan variation among
722   mammalian species. *Genome Biology* **19**, 22.

723   Maegawa S, Hinkal G, Kim HS*, et al.* (2010) Widespread and tissue specific age-
724   related DNA methylation changes in mice. *Genome Research* **20**, 332-340.

725   Maegawa S, Lu Y, Tahara T, *et al.* (2017) Caloric restriction delays age-related
726   methylation drift. *Nature Communications* **8**, 539.
727   Meissner A, Gnirke A, Bell GW, *et al.* (2005) Reduced representation bisulfite
728   sequencing for comparative high-resolution DNA methylation analysis.
729   *Nucleic Acids Research* **33**, 5868-5877.
730   Mills MS, Ryan PG (2005) Modelling impacts of long-line fishing: what are the
731   effects of pair-bond disruption and sex-biased mortality on albatross
732   fecundity? *Animal Conservation* **8**, 359-367.
733   Musick JA (1999) Ecology and conservation of long-lived marine animals **23**, 1-
734   10.
735   Nelson ME (2002) The science, ethics, and philosophy of tooth extractions from
736   live-captured white-tailed deer: a response to Festa-Bianchet et al.(2002).
737   *Wildlife Society Bulletin (1973-2006)* **30**, 284-288.
738   O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS (2018) These aren't
739   the loci you're looking for: Principles of effective SNP filtering for
740   molecular ecologists. *Molecular Ecology*.
741   Pérez‐Barbería F, Duff E, Brewer M, Guinness F (2014) Evaluation of methods
742   to age Scottish red deer: the balance between accuracy and practicality.
743   *Journal of Zoology* **294**, 180-189.
744   Polanowski AM, Robbins J, Chandler D, Jarman SN (2014) Epigenetic estimation
745   of age in humpback whales. *Molecular Ecology Resources* **14**, 976-987.
746   Portela A, Esteller M (2010) Epigenetic modifications and human disease. *Nature
747   Biotechnology* **28**, 1057.
748   Rand AC, Jain M, Eizenga JM, *et al.* (2017) Mapping DNA methylation with high-
749   throughput nanopore sequencing. *Nature Methods*.
750   Rattiste K, Klandorf H, Urvik J, *et al.* (2015) Skin pentosidine and telomere length
751   do not covary with age in a long-lived seabird. *Biogerontology* **16**, 435-
752   441.
753   Sarkar G, Turner RT, Bolander ME (1993) Restriction-site PCR: a direct method
754   of unknown sequence retrieval adjacent to a known locus by using
755   universal primers. *Genome Research* **2**, 318-322.

756   Schield DR, Walsh MR, Card DC*, et al.* (2016) EpiRADseq: scalable analysis of
757   genomewide patterns of methylation using next‐generation sequencing.
758   *Methods in Ecology and Evolution* **7**, 60-69.
759   Scott ME (1988) The impact of infection and disease on animal populations:
760   implications for conservation biology. *Conservation Biology* **2**, 40-56.
761   Shimoda N, Izawa T, Yoshizawa A*, et al.* (2014) Decrease in cytosine methylation
762   at CpG island shores and increase in DNA fragmentation during zebrafish
763   aging. *Age* **36**, 103-115.
764   Slatko BE, Gardner AF, Ausubel FM (2018) Overview of Next‐Generation
765   Sequencing Technologies. *Current Protocols in Molecular Biology* **122**,
766   e59.
767   Slieker RC, Relton CL, Gaunt TR, Slagboom PE, Heijmans BT (2018) Age-related
768   DNA methylation changes are tissue-specific with ELOVL2 promoter
769   methylation as exception. *Epigenetics & Chromatin* **11**, 25.
770   Slieker RC, van Iterson M, Luijk R*, et al.* (2016) Age-related accrual of
771   methylomic variability is linked to fundamental ageing mechanisms.
772   *Genome Biology* **17**, 191.
773   Spiers H, Hannon E, Wells S*, et al.* (2016) Age-associated changes in DNA
774   methylation across multiple tissues in an inbred mouse model.
775   *Mechanisms of Ageing and Development* **154**, 20-23.
776   Sziráki A, Tyshkovskiy A, Gladyshev VN (2018) Global remodeling of the mouse
777   DNA methylome during aging and in response to calorie restriction. *Aging*
778   *Cell* **17**, e12738.
779   Thompson MJ, vonHoldt B, Horvath S, Pellegrini M (2017) An epigenetic aging
780   clock for dogs and wolves. *Aging (Albany NY)* **9**, 1055.
781   Tricola GM, Simons MJ, Atema E*, et al.* (2018) The rate of telomere loss is related
782   to maximum lifespan in birds. *Phil. Trans. R. Soc. B* **373**, 20160445.
783   Velarde E, Ezcurra E (2018) Are seabirds' life history traits maladaptive under
784   present oceanographic variability? The case of Heermann's Gull (*Larus*
785   *heermanni*). *The Condor* **120**, 388-401.
786   Vidaki A, Ballard D, Aliferi A, Miller TH, Barron LP (2017) DNA methylation-
787   based forensic age prediction using artificial neural networks and next

788   generation sequencing. *Forensic Science International: Genetics* **28**, 225-
789   236.

790 Vidal-Bralo L, Lopez-Golan Y, Gonzalez A (2016) Simplified Assay for Epigenetic
791   Age Estimation in Whole Blood of Adults. *Frontiers in Genetics* **7**.

792 Weber KL, Bolander ME, Sarkab G (1998) Rapid acquisition of unknown DNA
793   sequence adjacent to a known segment by multiplex restriction site PCR.
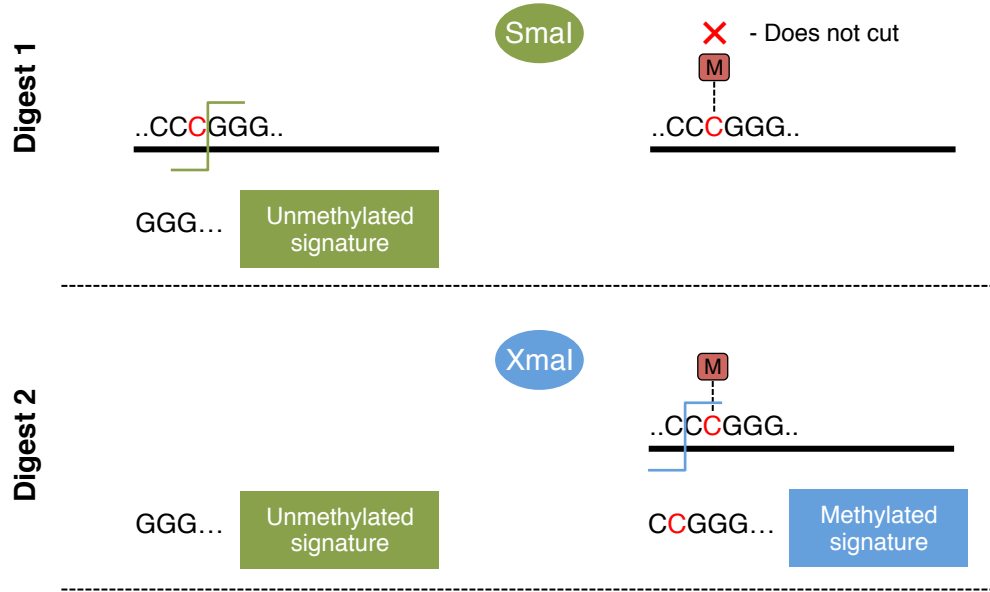794   *Biotechniques* **25**, 415-419.

795 Weimerskirch H, Lequette B, Jouventin P (1989) Development and maturation of
796   plumage in the wandering albatross (*Diomedea exulans*). *Journal of*
797   *Zoology* **219**, 411-421.

798 Wooller R, Bradley J, Skira I, Serventy D (1990) Reproductive success of short-
799   tailed shearwaters (*Puffinus tenuirostris*) in relation to their age and
800   breeding experience. *The Journal of Animal Ecology*, 161-170.

801 Wright PG, Mathews F, Schofield H*, et al.* (2018) Application of a novel molecular
802   method to age free‐living wild Bechstein's bats. *Molecular Ecology*
803   *Resources*.

804 Zbieć-Piekarska R, Spólnicka M, Kupiec T*, et al.* (2015) Examination of DNA
805   methylation status of the ELOVL2 marker may be useful for human age
806   prediction in forensic science. *Forensic Science International: Genetics* **14**,
807   161-167.

808

Digest 1

SmaI

✗ - Does not cut

M

..CCCGGG..

..CCCGGG..

GGG… Unmethylated signature

Digest 2

XmaI

M

..CCCGGG..

GGG… Unmethylated signature

CCGGG… Methylated signature

NextSeq Reads

Read 1	CCGGG	Read 6	GGG
Read 2	CCGGG	Read 7	GGG
Read 3	GGG	Read 8	CCGGG
Read 4	CCGGG	Read 9	CCGGG
Read 5	CCGGG	Read 10	GGG

DNAm score

$$\frac{6x\ CCGGG}{10x\ reads} = \mathbf{0.60}$$

M_001071_Seq1071
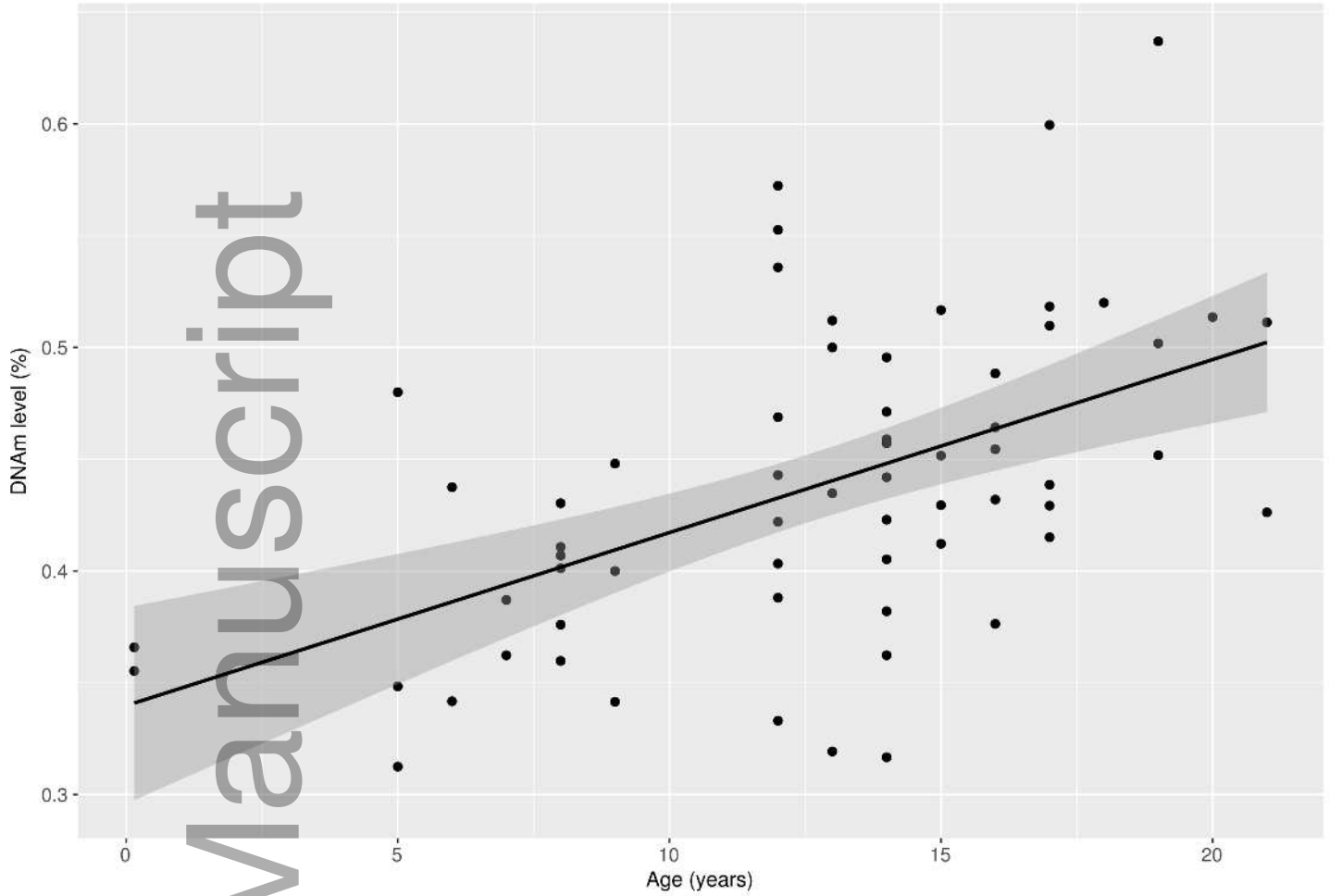
men_12981_f2a.tif

M_001158_Seq1158
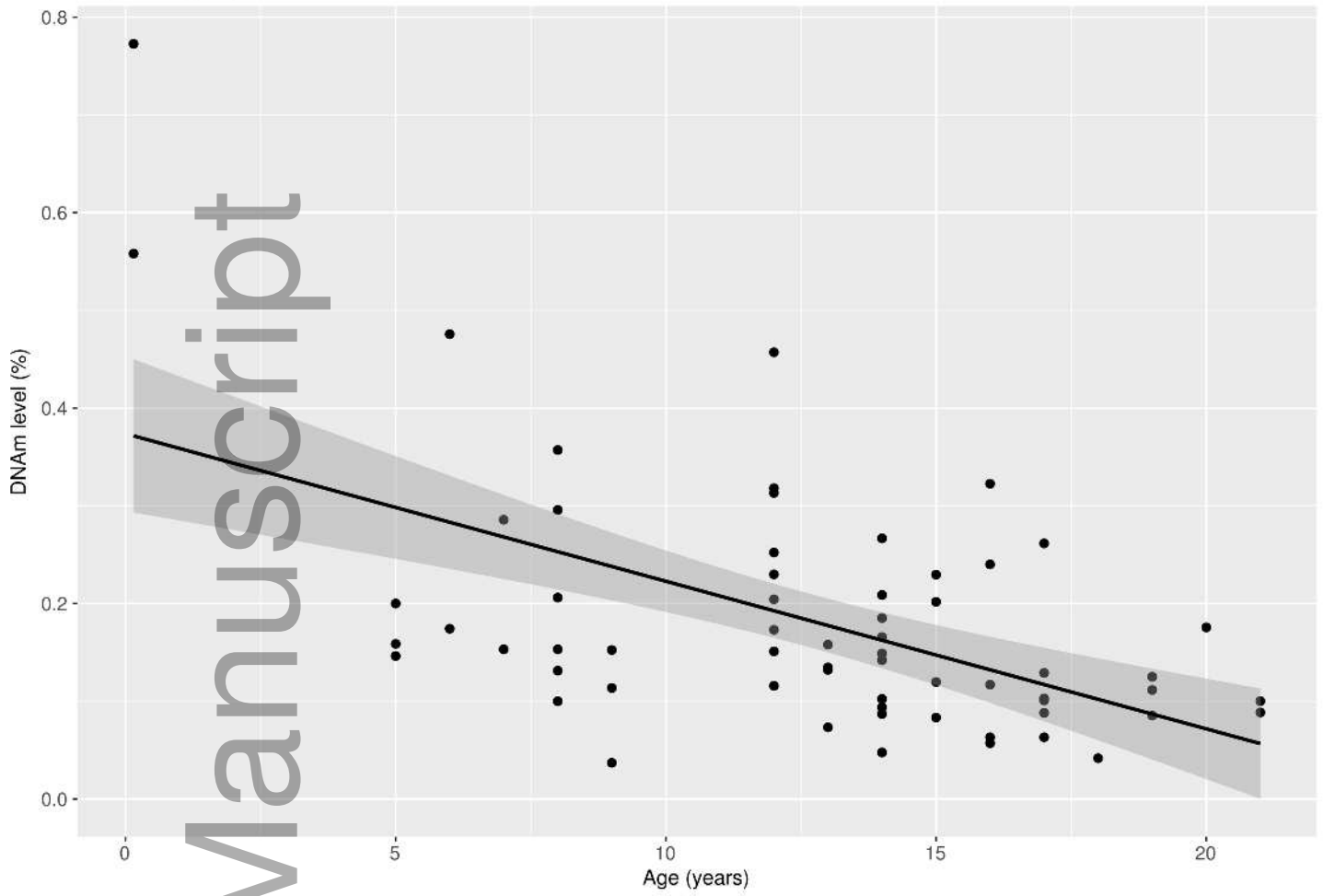
Adj R2 = 0.2896 Range = 0.736

men_12981_f2b.tif

M_001801_Seq1801

DNAm level (%)

Age (years)
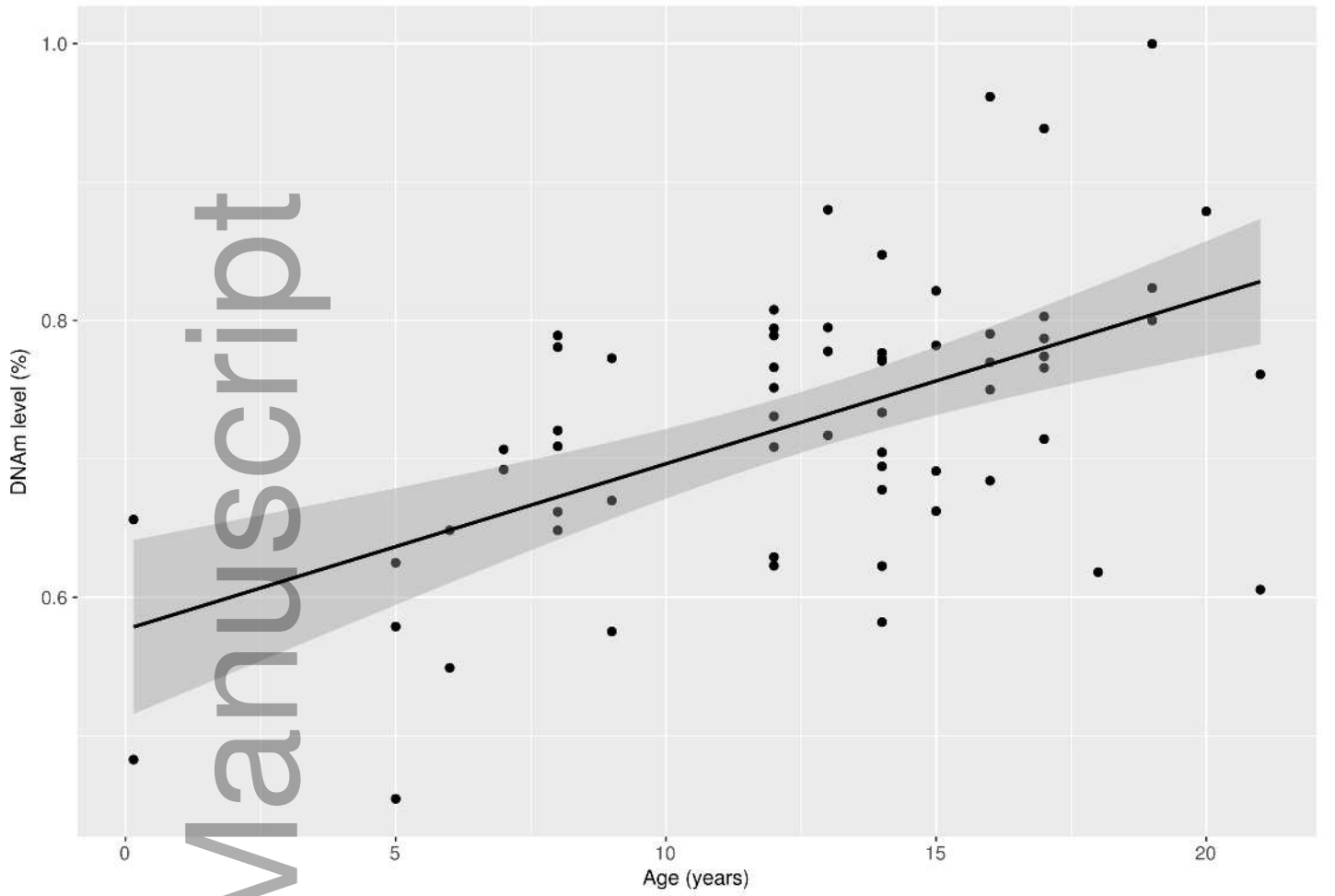
Adj R2 = 0.2849 Range = 0.545

men_12981_f2c.tif

men_12981_f2d.tif

men_12981_f2e.tif

M_003169_Seq3169

DNAm level (%)

Age (years)
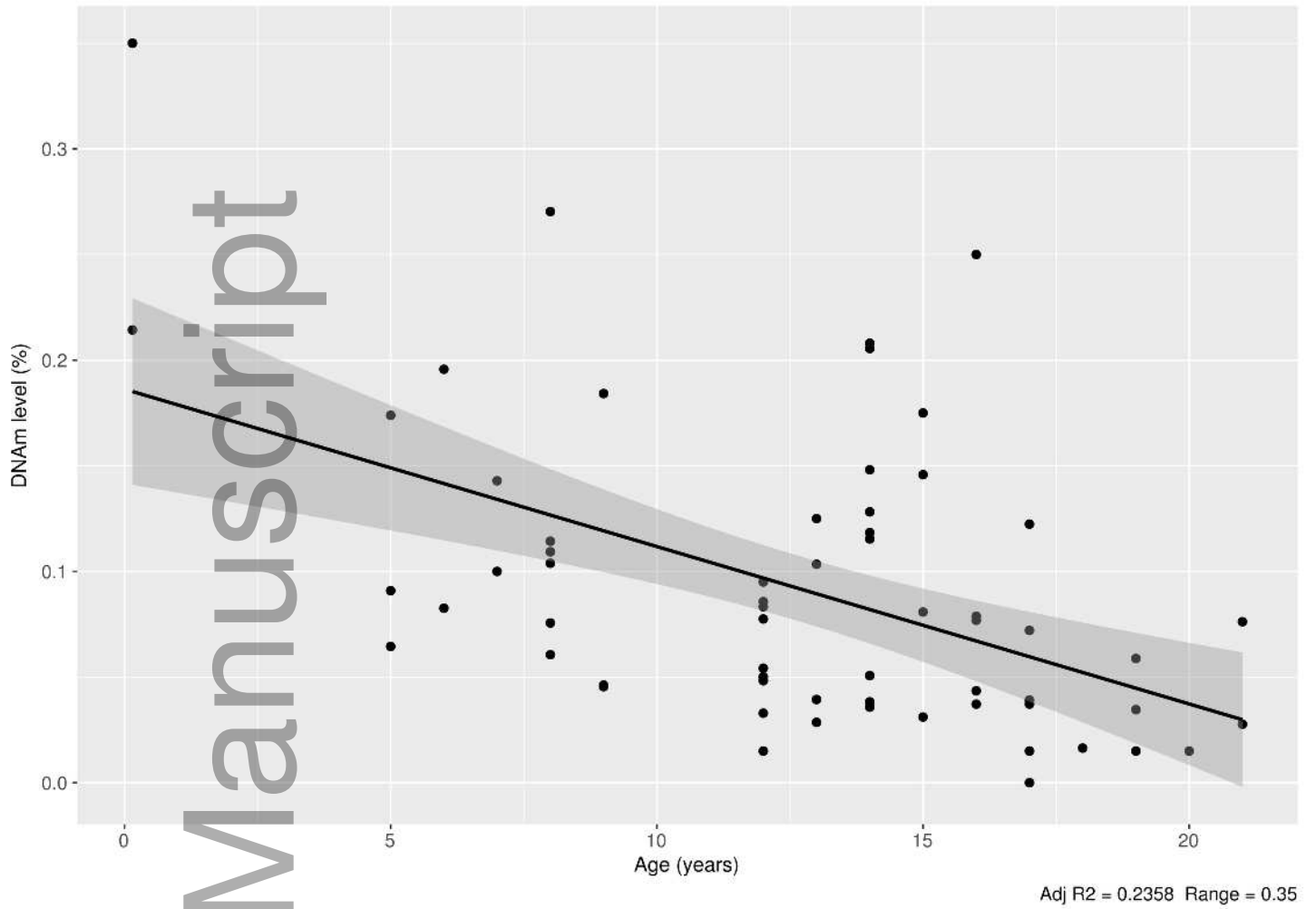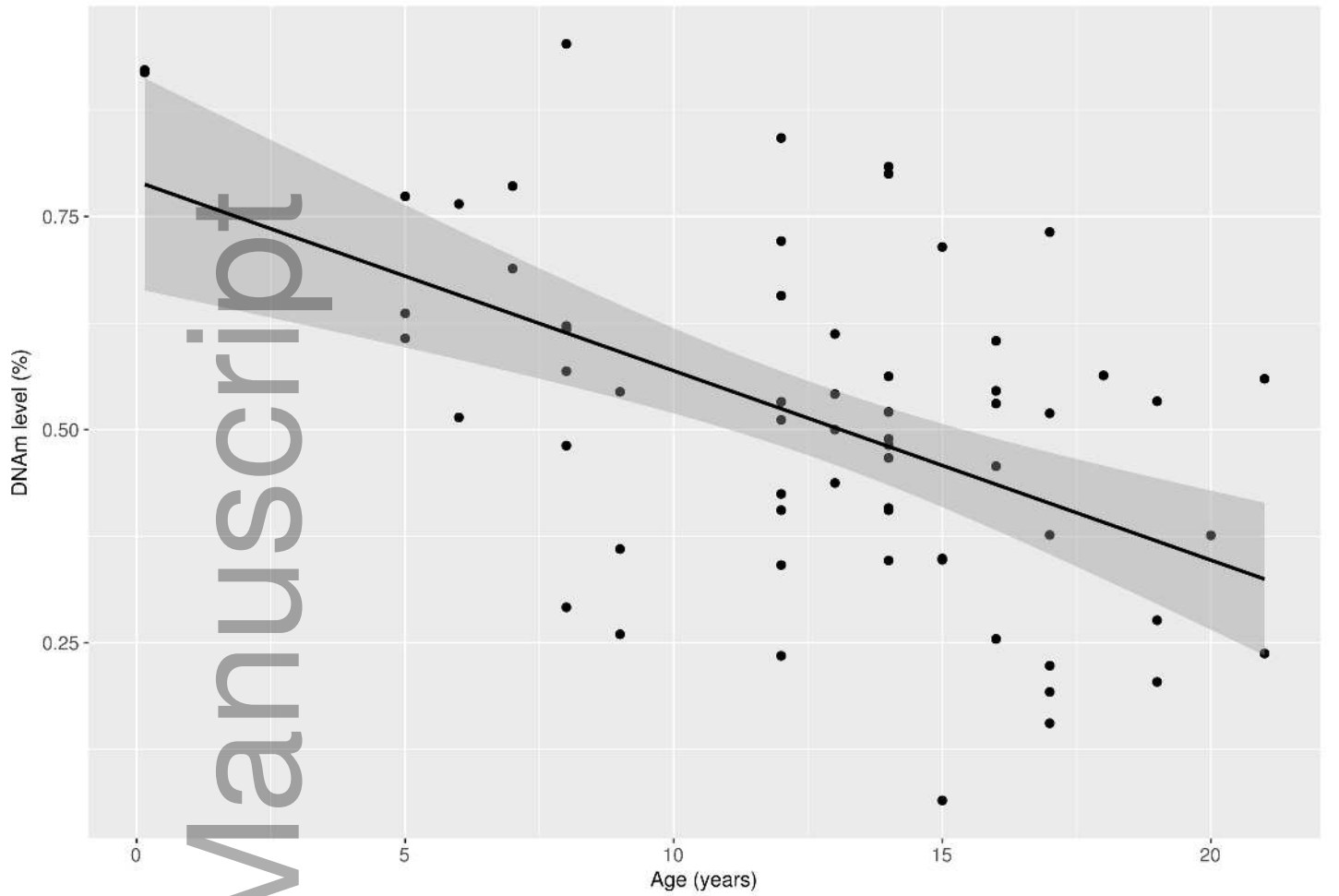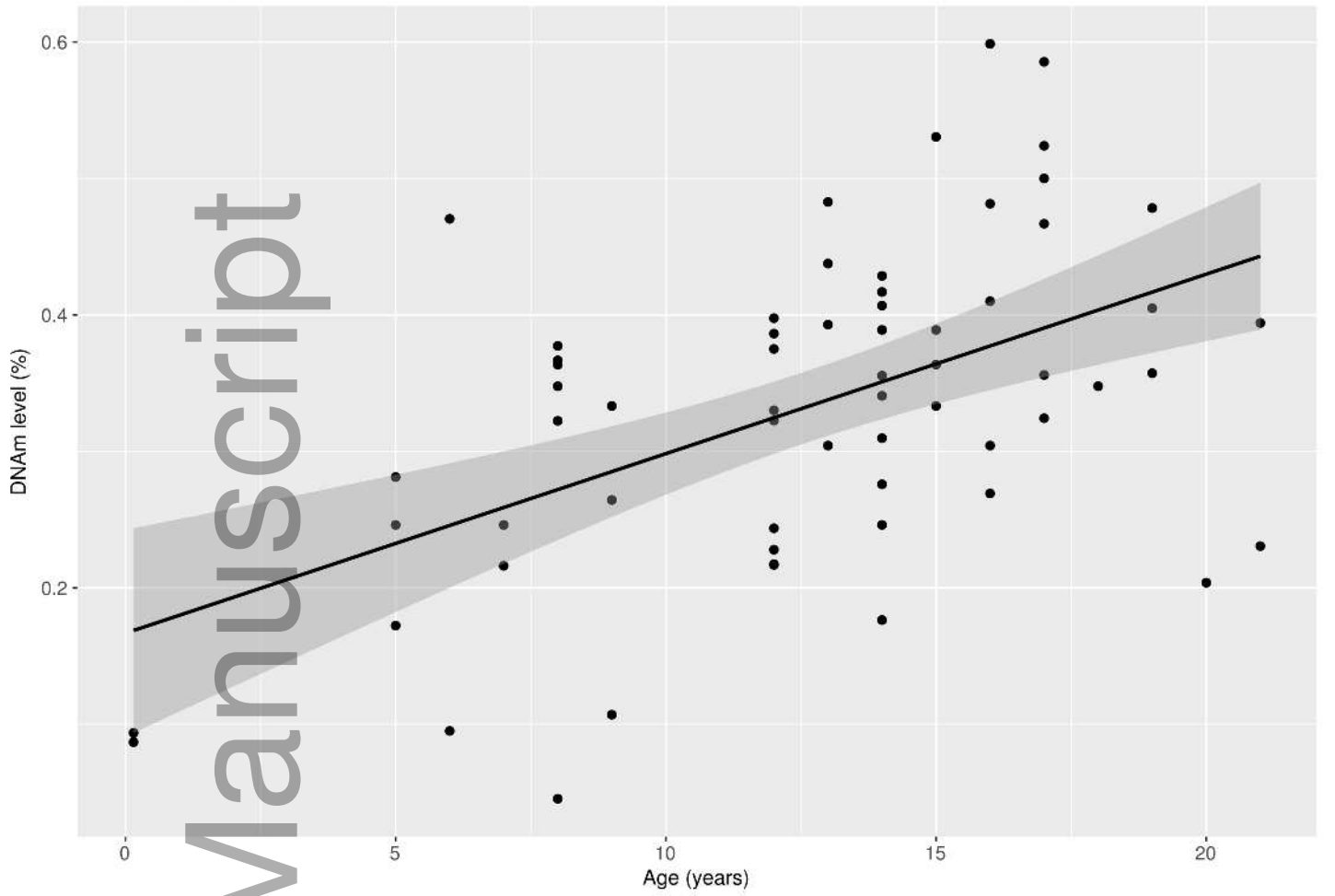
Adj R2 = 0.2513 Range = 0.553

men_12981_f2f.tif
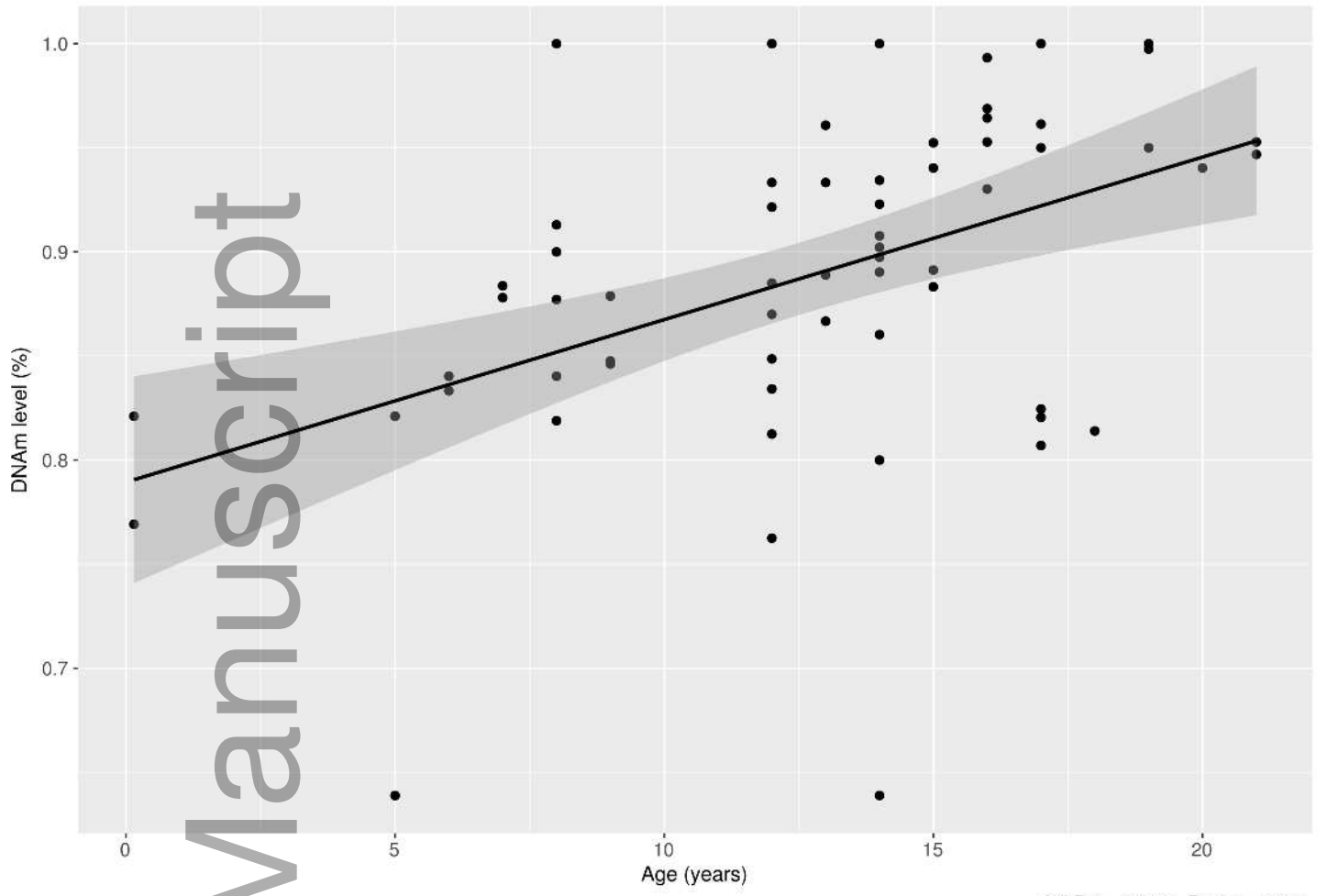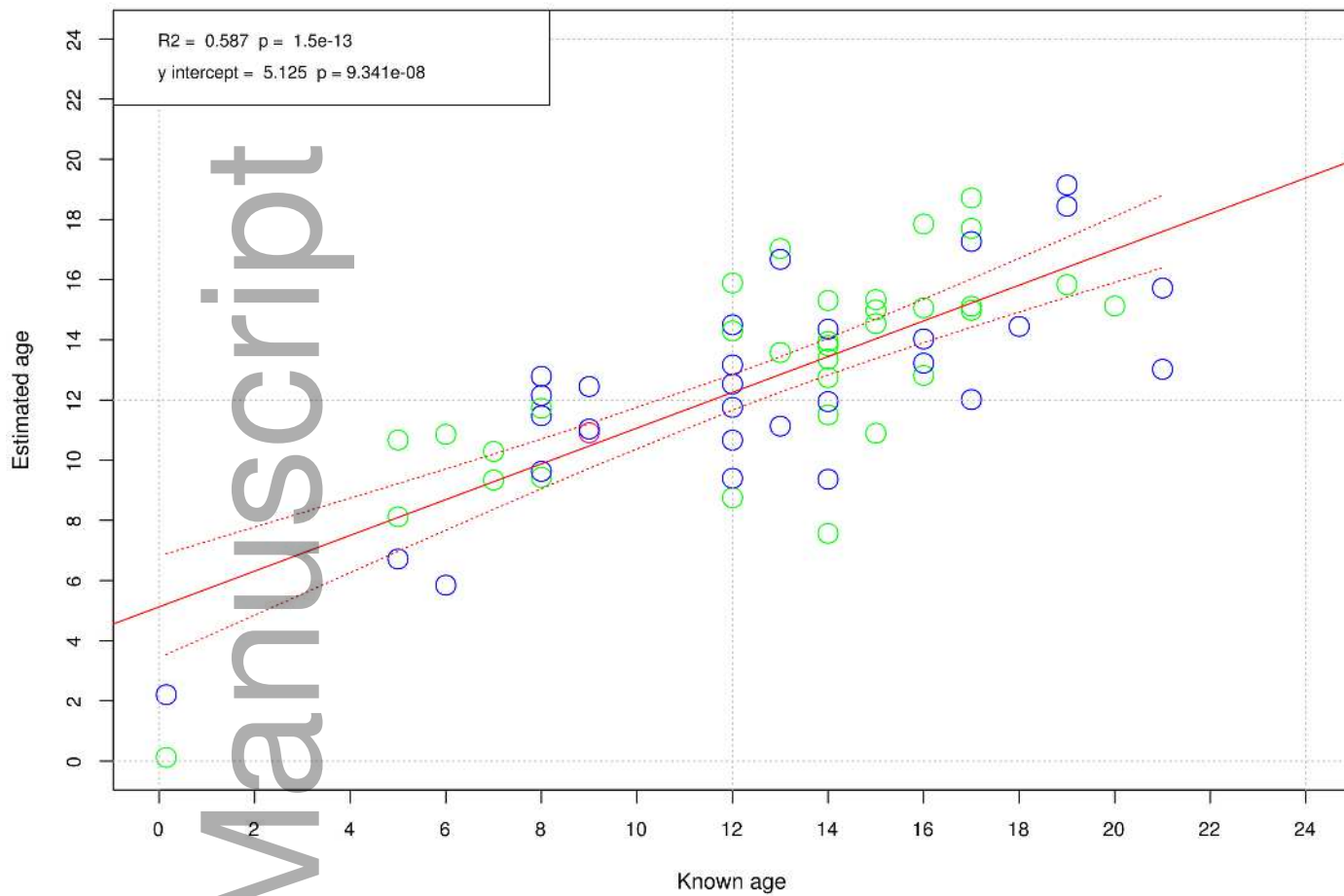
M_003784_Seq3784

men_12981_f2g.tif

Figure showing scatter plot with "Estimated age" (y-axis) versus "Known age" (x-axis). Text in plot: R2 = 0.587 p = 1.5e-13; y intercept = 5.125 p = 9.341e-08

men_12981_f3.tif

men_12981_f4a.eps

Known age (years)

men_12981_f4b.eps

Figure axes: Known age (years) on x-axis, Estimated age (years) on y-axis.

Data point labels: 39179, 39348, 39348, 39254, 39179, 38856, 52095, 39254, 38856, 52152, 52152, 52095, 52196, 52196

men_12981_f5.eps

Color Key and Histogram

Count

25000

15000

5000

0

0.0   0.2   0.4   0.6   0.8   1.0

Value
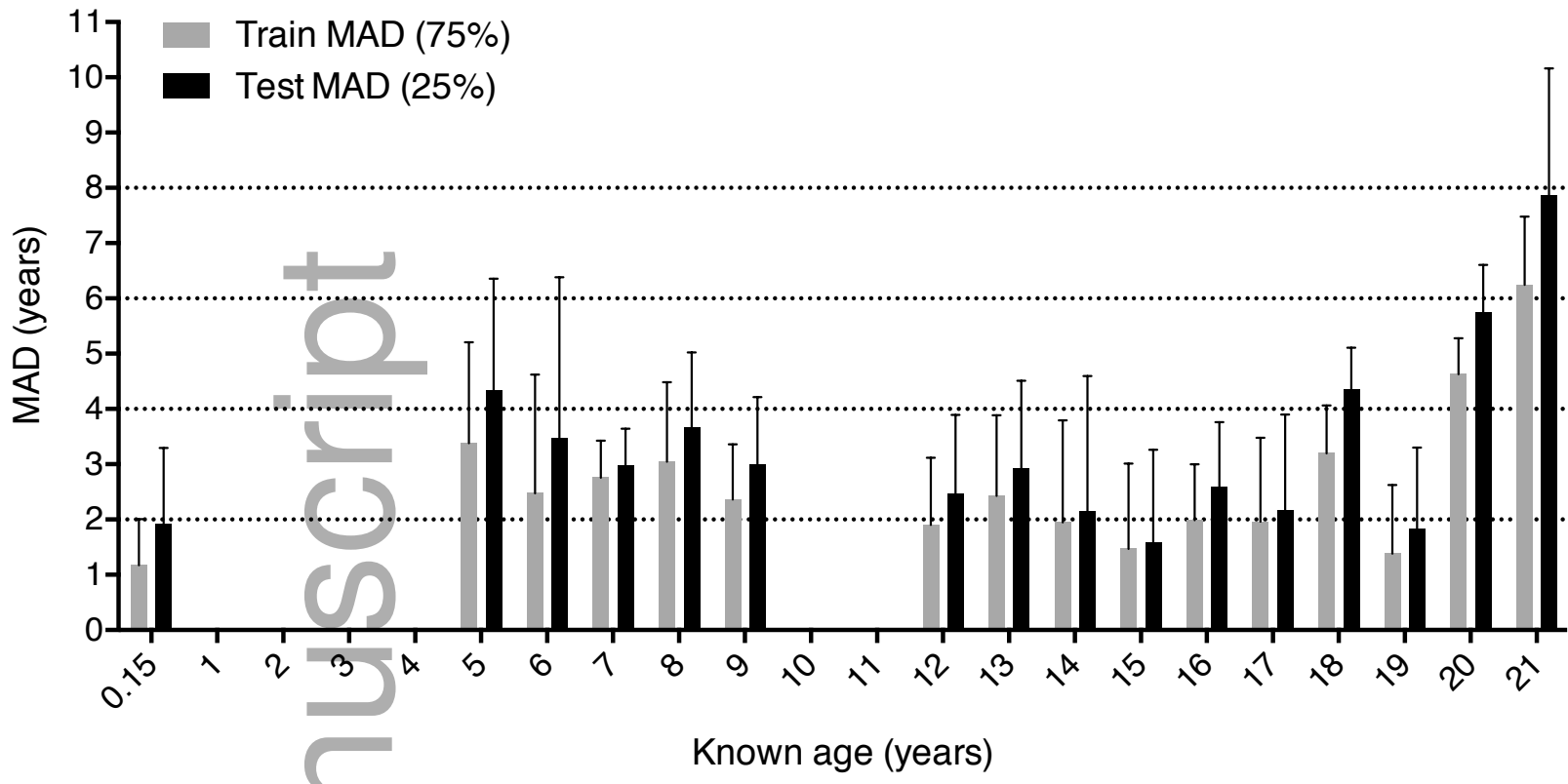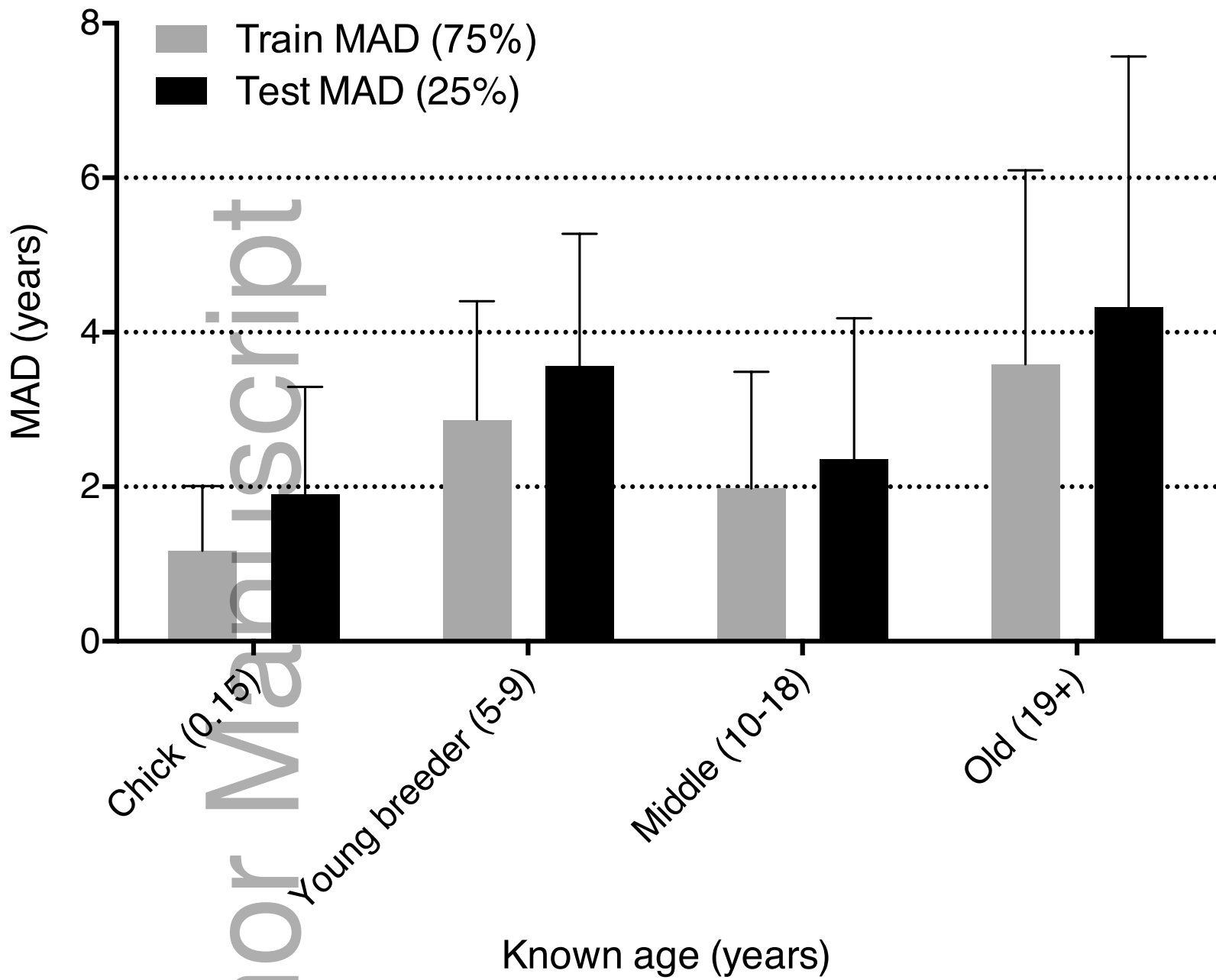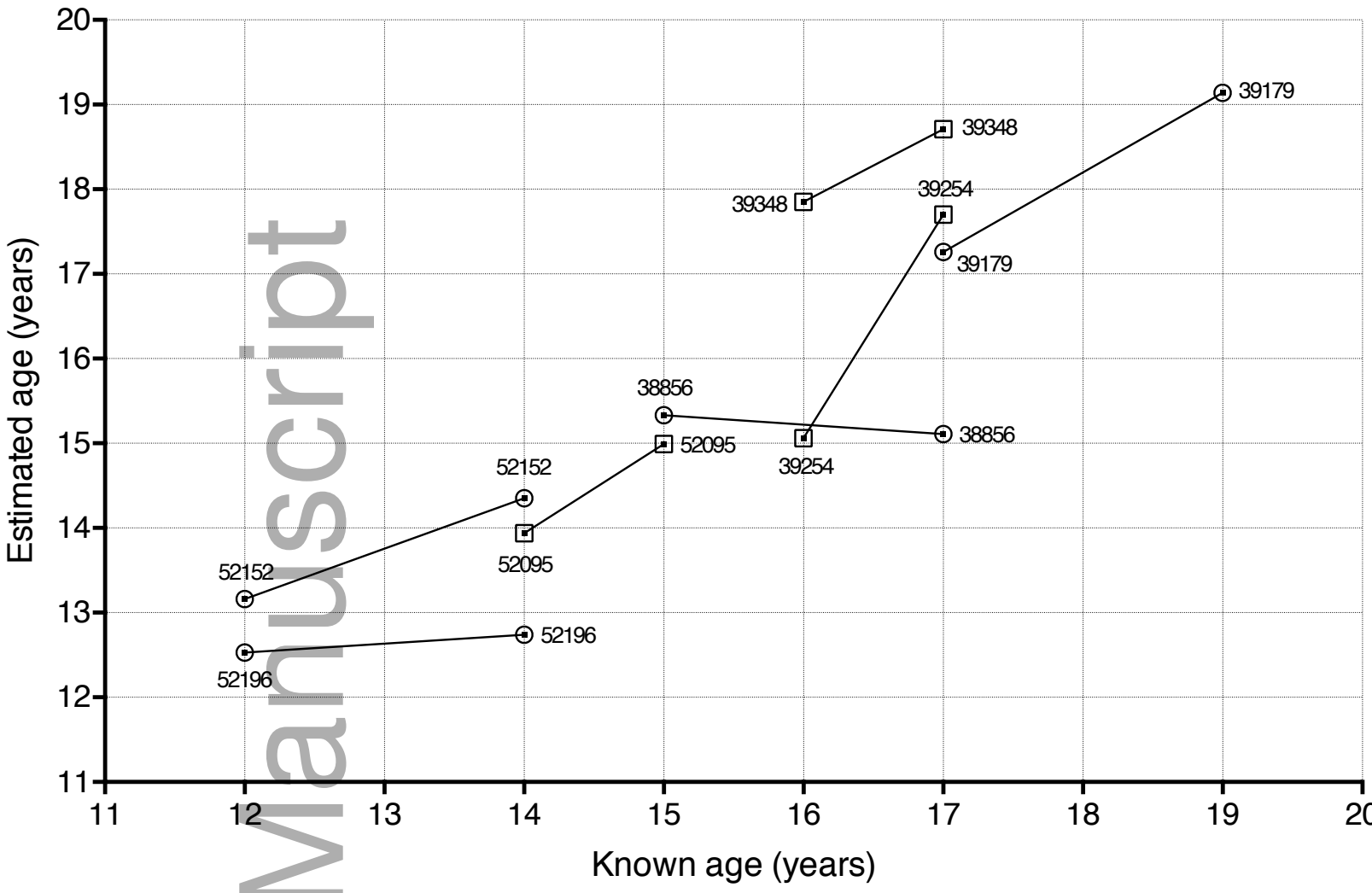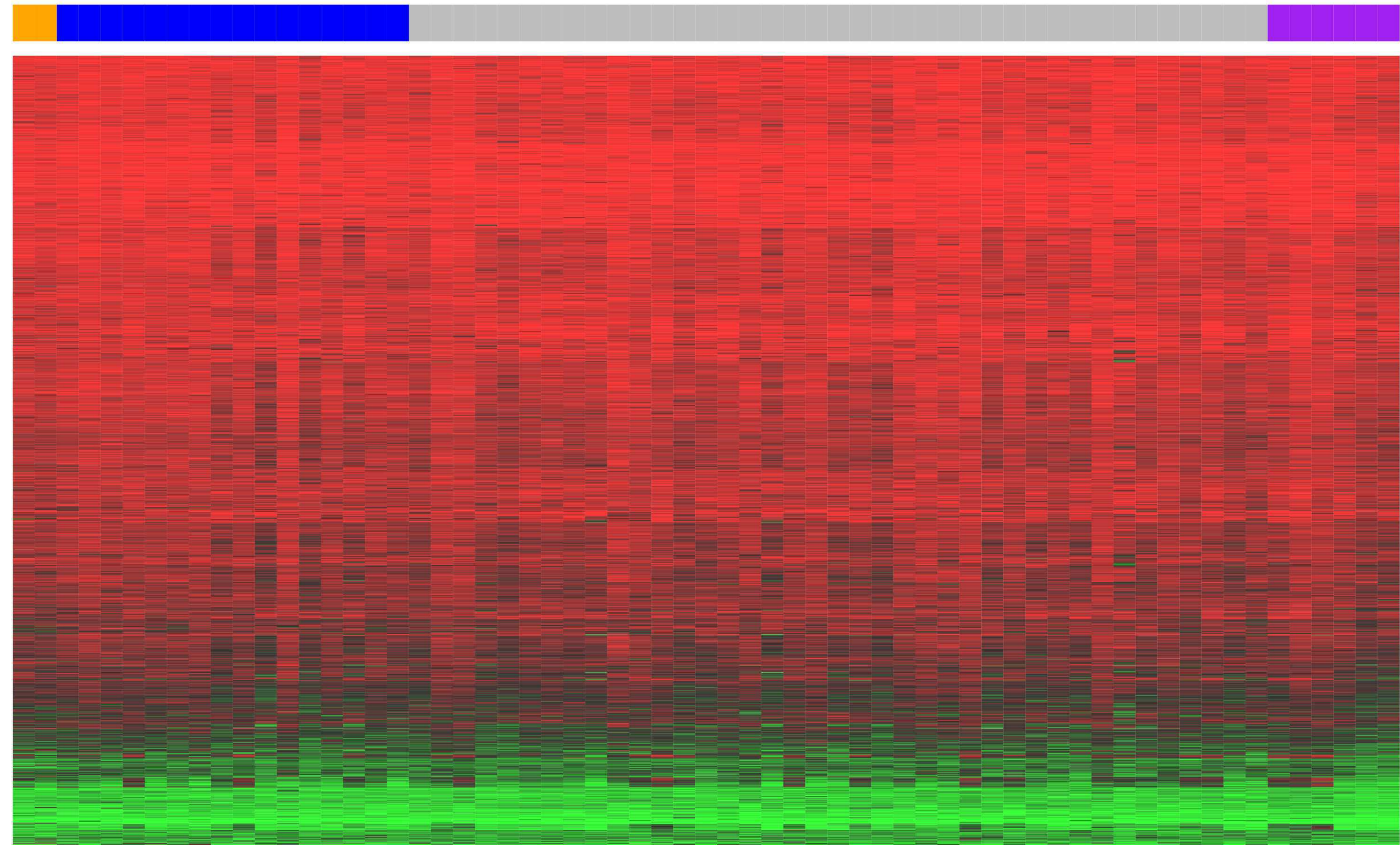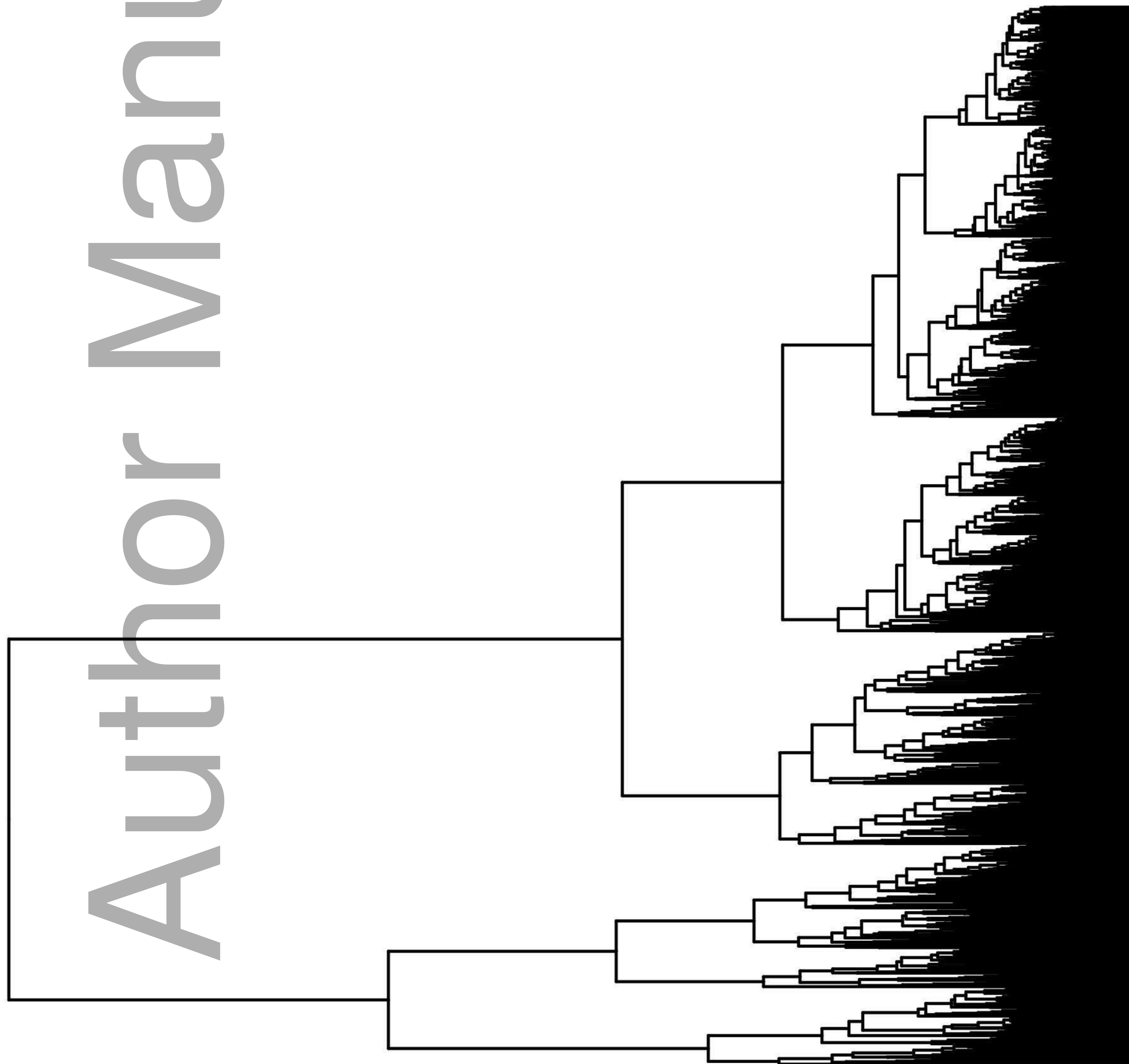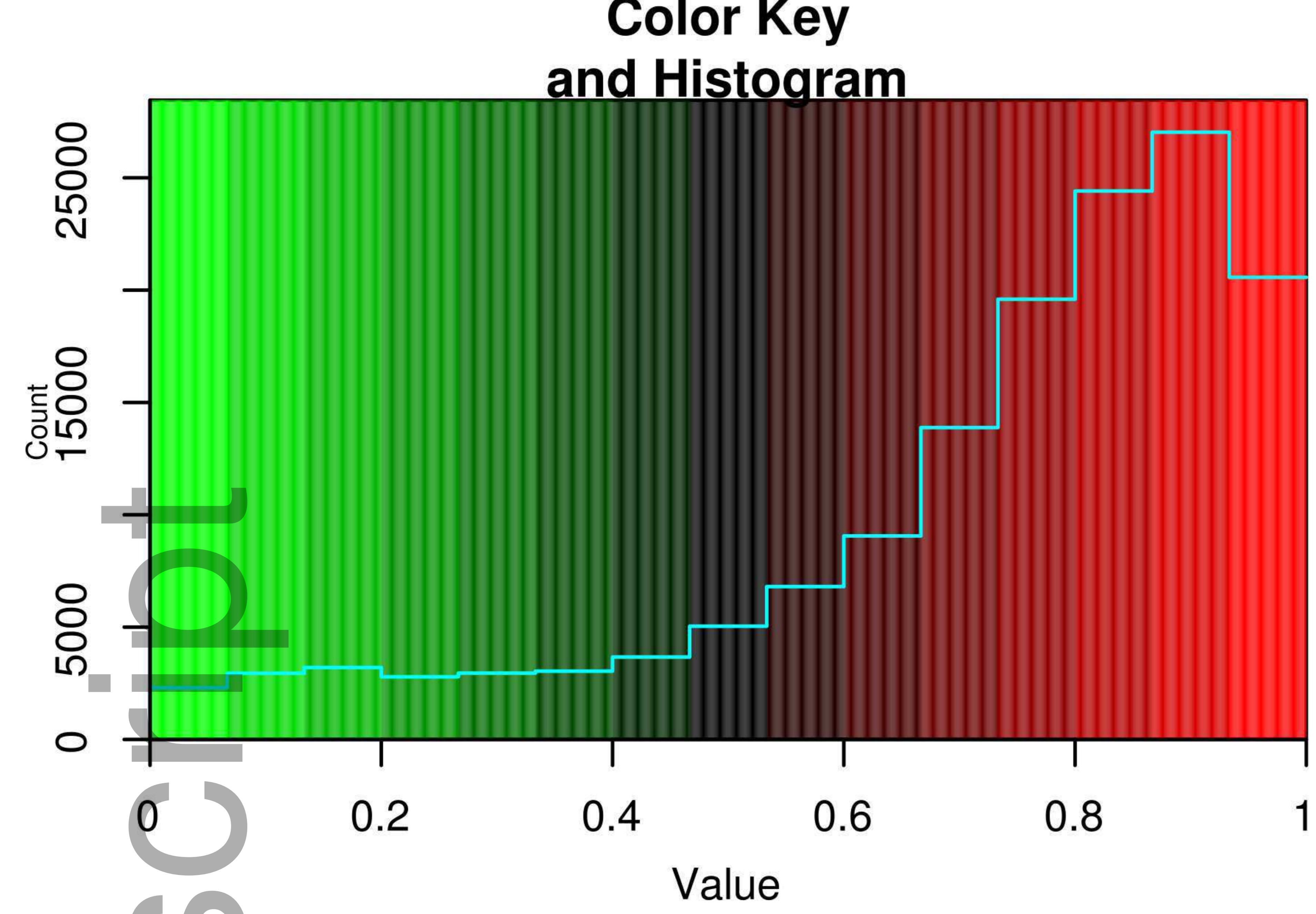
STS16020_S35
STS16011_S34
STS15007_S1_
STS15025_S2_
STS15103_S3_
STS15046_S1_
STS16074_S5_
STS16051_S7_
STS15105_S6_
STS16029_S4_
STS15050_S10
STS17060_S5_
STS15037_S9_
STS15091_S2_
STS16089_S11
STS17001_S6_
STS16080_S13
STS16046_S12
STS15057_S7_
STS15111_S16
STS15044_S14
STS15073_S9_
STS15065_S8_
STS15116_S13
STS15077_S11
STS15099_S12
STS15075_S10
STS16058_S17
STS16075_S18
STS16028_S14
STS16081_S19
STS17014_S19
STS16102_S18
STS16104_S23
STS17046_S22
STS16063_S22
STS16021_S21
STS17015_S20
STS15087_S15
STS17044_S21
STS16043_S17
STS15020_S24
STS17030_S24
STS16036_S25
STS15026_S23
STS15047_S27
STS16060_S26
STS16078_S28
STS15031_S26
STS15013_S25
STS16041_S28
STS15113_S27
STS16044_S30
STS17034_S29
STS17052_S30
STS16067_S31
STS15117_S31
STS17059_S32
STS16052_S32
STS17065_S33
STS15059_S34
STS17038_S35

Individual birds (samples 1-63)
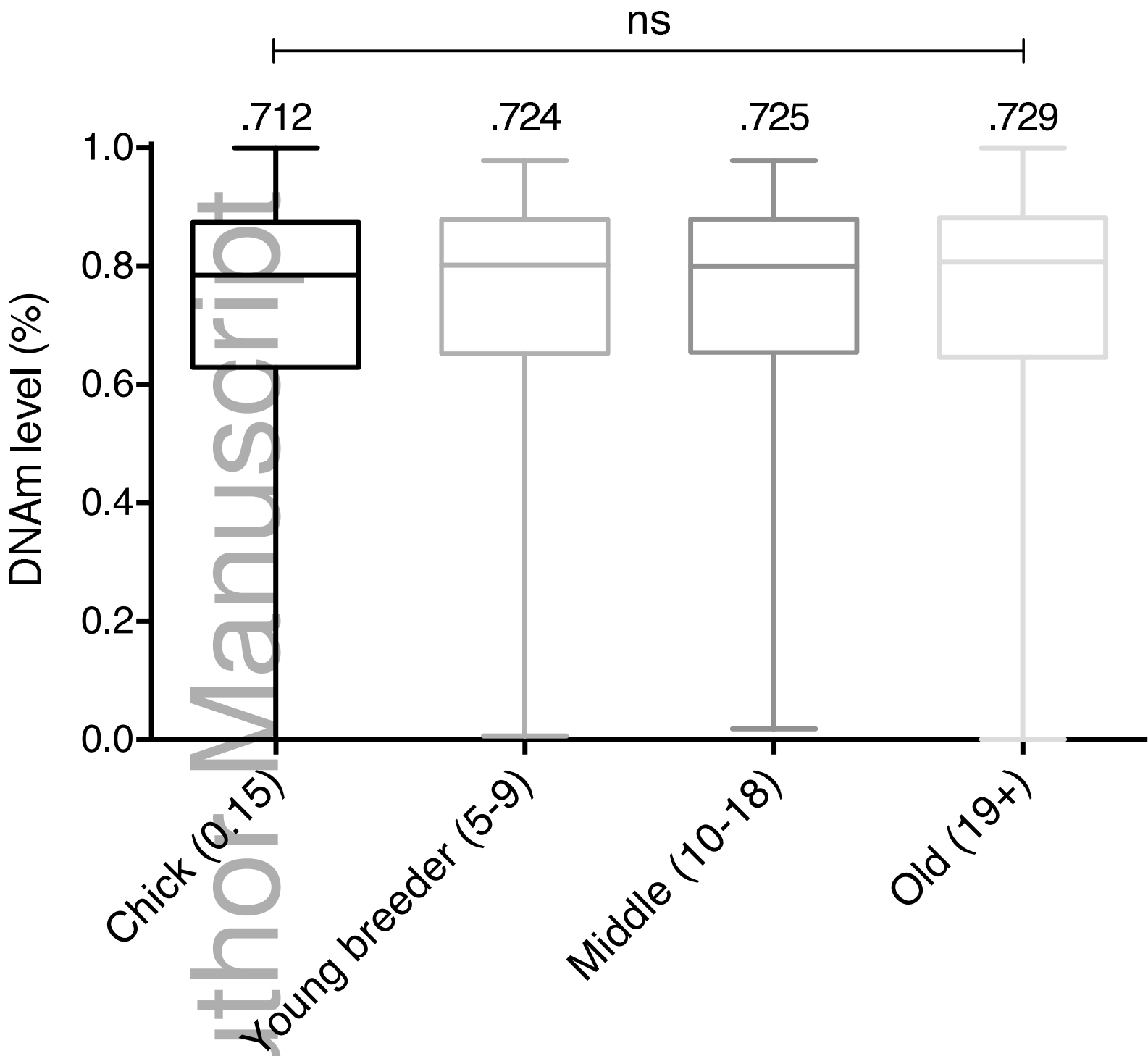
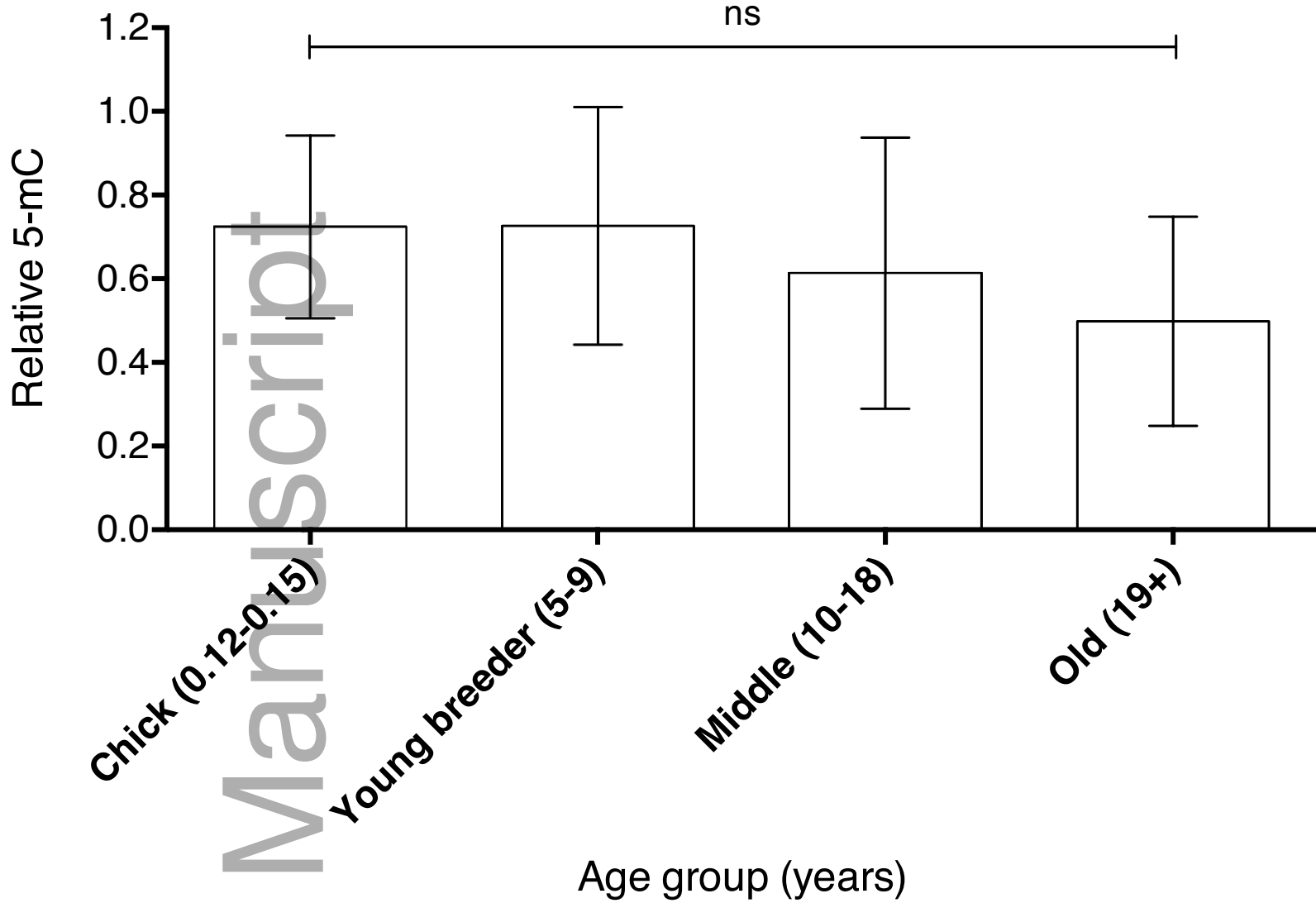men_12981_f6b.eps
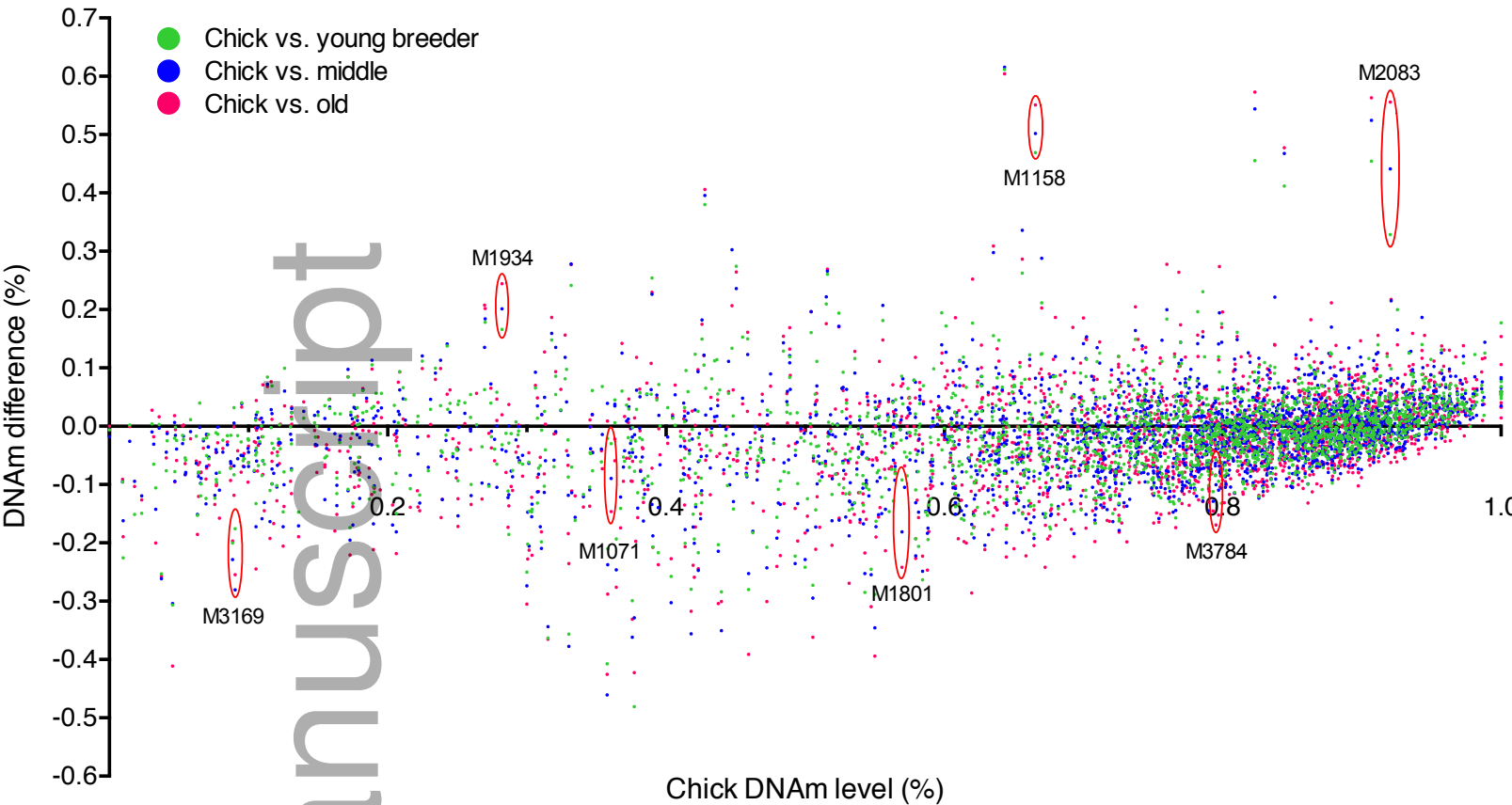
men_12981_f6c.eps

Age group (years)

men_12981_f7.eps

Author/s:
De Paoli-Iseppi, R; Deagle, BE; Polanowski, AM; McMahon, CR; Dickinson, JL; Hindell, MA; Jarman, SN

Title:
Age estimation in a long-lived seabird (Ardenna tenuirostris) using DNA methylation-based biomarkers

Date:
2019-03-01

Citation:
De Paoli-Iseppi, R., Deagle, B. E., Polanowski, A. M., McMahon, C. R., Dickinson, J. L., Hindell, M. A. & Jarman, S. N. (2019). Age estimation in a long-lived seabird (Ardenna tenuirostris) using DNA methylation-based biomarkers. MOLECULAR ECOLOGY RESOURCES, 19 (2), pp.411-425. https://doi.org/10.1111/1755-0998.12981.

Persistent Link:
http://hdl.handle.net/11343/285418

File Description:
Accepted version