Received: 20 November 2023 | Revised: 5 February 2024 | Accepted: 9 February 2024

DOI: 10.1111/1755-0998.13940

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Methylation-based markers for the estimation of age in African cheetah, *Acinonyx jubatus*

Louis-Stéphane Le Clercq[1,2] | Antoinette Kotzé[1,2] | J. Paul Grobler[2] | Desiré L. Dalton[3]

[1]South African National Biodiversity Institute, Pretoria, South Africa

[2]Department of Genetics, University of the Free State, Bloemfontein, South Africa

[3]School of Health and Life Sciences, Teesside University, Middlesbrough, UK

Correspondence
Louis-Stéphane Le Clercq, South African National Biodiversity Institute, Pretoria 0001, South Africa.
Email: leclercq.l.s@gmail.com

Funding information
National Research Foundation, Grant/Award Number: 112062

Handling Editor: Alana Alexander

Abstract

Age is a key demographic in conservation where age classes show differences in important population metrics such as morbidity and mortality. Several traits, including reproductive potential, also show senescence with ageing. Thus, the ability to estimate age of individuals in a population is critical in understanding the current structure as well as their future fitness. Many methods exist to determine age in wildlife, with most using morphological features that show inherent variability with age. These methods require significant expertise and become less accurate in adult age classes, often the most critical groups to model. Molecular methods have been applied to measuring key population attributes, and more recently epigenetic attributes such as methylation have been explored as biomarkers for age. There are, however, several factors such as permits, sample sovereignty, and costs that may preclude the use of extant methods in a conservation context. This study explored the utility of measuring age-related changes in methylation in candidate genes using mass array technology. Novel methods are described for using gene orthologues to identify and assay regions for differential methylation. To illustrate the potential application, African cheetah was used as a case study. Correlation analyses identified six methylation sites with an age relationship, used to develop a model with sufficient predictive power for most conservation contexts. This model was more accurate than previous attempts using PCR and performed similarly to candidate gene studies in other mammal species. Mass array presents an accurate and cost-effective method for age estimation in wildlife of conservation concern.

KEYWORDS

Acinonyx jubatus, African, age, age estimation, biological clocks, biomarker, cheetah, genes, mass array, methylation

1 | INTRODUCTION

Being able to accurately identify the age of an animal is often helpful to wildlife conservation. In understanding population dynamics—including effective population size—the number of individuals that are of young, reproductive, or old age serve as indicators for the present and future fitness of that species (Cooper et al., 2021). Thus, knowledge of these classes are crucial in monitoring populations and measuring the efficacy of current conservation efforts, as most population models and algorithms assume a fixed population distribution of known ages (Durant et al., 2017). In conservation genetics, many methods have been developed to use DNA to study pertinent individual as well as population characteristics, including biological sex (Morin et al., 2005), parentage (Double et al., 1997), population assignment (Kim & Sappington, 2013), hybridization (Brelsford et al., 2011), speciation (Linck et al., 2019), and migration (Merlin & Liedvogel, 2019). As ageing is a biological process under genetic control (Horvath, 2013) or the result of unrepaired damage to DNA (Kujoth et al., 2005), the cellular and genetic hallmarks of ageing can be used to determine age. One genetic method for age estimation is the epigenetic regulatory elements including methylation (Johansson et al., 2013) which has become a well-established method in humans (Gopalan et al., 2017). Recently, these methods have been extended to many species (Le Clercq, Kotzé, et al., 2023), including many of the more than 280 carnivore species (order: Carnivora) such as canids (Horvath et al., 2022; Ito et al., 2017; Thompson et al., 2017), pinnipeds (Robeck et al., 2023; Sareisian, 2014), felines (Cantrell et al., 2020; Qi et al., 2021; Raj et al., 2021), and bears (Nakamura et al., 2023).

The cheetah, *Acinonyx jubatus* [von Schreber, 1775], is a carnivore and the last remaining species in the genus *Acinonyx* (O'Brien et al., 1985). There are currently four recognized subspecies based on geographic location and morphology. These include the most abundant *A. j. jubatus* [von Schreber, 1775] in Southern Africa, *A. j. venaticus* [Griffith, 1821] in Iran, *A. j. hecki* [Hilzheimer, 1913] in Northern Africa, and *A. j. soemmeringii* [Fitzinger, 1855] in Central Africa (Krausman & Morales, 2005). *Felidae* species diverged from other carnivores approximately 100 million years ago (Nyakatura & Bininda-Emonds, 2012). The *Acinonyx* genus diverged from other *Felidae* species approximately 4.02 to 17.30 million years ago, while a divergence time between 32 and 67 thousand years ago is estimated between Asiatic and African cheetahs and African subspecies, *A. j. soemmeringii* and *A. j. jubatus*, diverged in parallel, approximately 16 to 72 thousand years ago (Charruau et al., 2011).

The International Union for Conservation of Nature (IUCN) uses a particular 'Red List' algorithm which considers several key ecological parameters as criteria to categorize individual species for their extinction risk to facilitate the prioritization of conservation efforts to those with the highest risk (Mace et al., 2008). Cheetah, along with many other *Felidae* species including the fishing cat (*Prionailurus viverrinus*), clouded leopard (*Neofelis nebulosa*), snow leopard (*Panthera uncia*), leopard (*Panthera pardus*), lion (*Panthera leo*), and tiger (*Panthera tigris*), are listed as either vulnerable (Durant et al., 2015) or endangered by the IUCN. Serious concerns exist about the potential extinction of these species considering the American lion, *Panthera atrox*, three species of American cheetah, genus *Miracinonyx*, and three species of Sabre-tooth tiger, genus *Smilodon*, went extinct over the past 15 thousand years, while several subspecies of lion (Black et al., 2013), cougar (Cardoza & Langlois, 2002), and tiger (Yamaguchi et al., 2013) were declared extinct between 1960 and 2018.

Cheetah once ranged across large stretches of Asia, India, the Middle East, and Africa; however, the population trend is estimated to be decreasing (Durant et al., 2015). As of 2016, approximately 7100 cheetahs remain in the wild; dispersed into 33 populations with fluctuating effective population sizes of around one hundred (Durant et al., 2015). At present, they only occupy about 6%–9% or less of their historical range, mostly in Eastern and Southern Africa, with an estimated 1326 in South Africa; and the majority (77%) of that range falls outside of protected areas (Durant et al., 2017). The cheetah has a maximum lifespan of up to 15 years in captivity; however, in the wild there is a less than 5% survival rate for cubs and most individuals only live between 6 and 8 years, with high mortality rates around the age of 2 years old when they first mate (Caro, 1994).

In the cheetah, individuals are classified into eight age groups, ranging from young cubs (0–6 months) to very old adults (>144 months) (Table 1). Current methods to age wild cheetah largely consist of physical examination of phenotypical traits such as teeth, fur, and body condition (Marker et al., 2003), which become progressively less accurate with older age groups (summarized in Table 1). The morphological method of age determination in cheetah was developed through observations made from pet cheetah (Adamson, 1970; Burney, 1980) along with published data (Broom, 1949) as well as fieldwork from the Serengeti Cheetah Project (Caro, 1994; Kelly et al., 1998) in Tanzania and Cheetah Conservation Fund (Marker et al., 2003) in Namibia. The accuracy of physical features was cross-validated using the more established method of tooth cementum annulation (TCA) on teeth extracted from dead individuals (Marker & Dickman, 2003). The TCA method uses a biological clock that follows predictable annual changes, the deposition of a dark and light layer each year following the eruption of the tooth, that enable histological age estimation (Naji et al., 2016; Sharma et al., 2021). The accuracy of TCA itself for age estimation in cheetah, however, remains dubious as they are known to suffer from frontal palatine erosion, particularly in captive individuals, due to mandible malocclusion (Fitch & Fagan, 1982) which affects tooth wear patterns and could obscure annuli or lines of annulated cementum (Nakanishi et al., 2009). Furthermore, the utility of morphometric observations for age determination is reliant upon specialized training and considerable experience, while an exploratory or validating study remains to be published in the main scientific literature. This has already been done in, for example, many

TABLE 1 Summary of morphological features used to determine age in cheetah (Marker & Dickman, 2003).

Class	Teeth	Coat	Body
1. Young cubs 0–6 months	Deciduous canines and incisors erupt at 28–30 days; molars erupt at 45–50 days	Spots on legs and yellow hair colouring develop at 6–7 weeks, mantle present at 4 weeks and lost at 3–4 months	Eyes open at 7–10 days, cubs emerge from den at about 6 weeks
2. Large cubs 6–12 months	Lower incisors fall at about 7 months, adult teeth erupt at about 8 months	Long hair on back of neck remains, although it is no longer a defined mantle	Lanky appearance until about 9 months, then body begins to fill out; body mass about two-thirds by 12 months
3. Adolescents 12–18 months	No tartar or yellowing of teeth	Some long fur on back of neck; fur on face and body fuzzy and scruffy rather than smooth	Attain full height but not adult weight, leggy, with dam
4. Independent 18–30 months	No tartar or yellowing of teeth	Some long fur on back of neck, smooth, sleek coat	Develop muscle tone, usually not with dam but may be with littermates
5. Young adults 30–48 months	Slight tartar and yellowing of teeth	Slight mane still	Fully grown but not fully muscled, in prime physical condition. Males have scars, females usually pregnant or with cubs
6. Prime adults 48–96 months	Tartar and yellowing of teeth, slight gum recession, some gingivitis	Mane on back of neck is gone	Fully muscled, prime physical condition but starting to show signs of ageing
7. Old adults 96–144 months	Tartar and yellowing of teeth, gum recession, gingivitis, canines tipped, loss of teeth, especially incisors	Coat beginning to look ragged, poorly groomed, scarred	Pads becoming smooth and elongated, sunken face, thinner, loss of muscle tone
8. Very old adults >144 months	Tartar and yellowing of teeth, gum recession, gingivitis, canines tipped, loss of teeth, especially incisors and canines, broken teeth	Ragged, poorly groomed, scarred coat	Pads quite smooth and elongated, sunken face, body delicate, and frail

other species in the family *Felidae* such as leopards (Stander, 1997), tigers (Fàbregas & Garcés-Narro, 2014), and African lions (White & Belant, 2016). Consequently, this makes the morphological method difficult to apply to cheetahs across their full range with serious questions remaining regarding the reliability of the traits used to assess age.

While the utility of methylation for age determination has been previously illustrated in the cheetah (Raj et al., 2021), there are difficulties with applying the Horvath Mammalian Array to cheetah, due to low-throughput sample processing that increases the per-sample cost when applied in a conservation context. Although the Clock Foundation does facilitate research collaborations that could reduce the cost, this requires submitting samples internationally, due to CITES regulations for species of conservation concern and potential concerns over sample/data sovereignty (Knight et al., 2022; Zainol et al., 2011) mean this is not currently feasible. Therefore, the aim of the present study is to investigate the potential of age-related changes in methylation, using a candidate gene approach for genes under epigenetic clock control assayed in previous studies, to be used to establish an accurate age estimation model using *Acinonyx jubatus* as a case study. Beyond cheetahs, the process we outline here for developing a candidate gene approach should be broadly useful to other species, where resource limitation and/or difficulty in internationally shipping samples precludes the use of existing array technology.

2 | METHODS

2.1 | Samples and ethics

Approval for the present study was obtained from the protocol committee of the Department of Genetics, University of the Free State (approval number: Res18/2020). Ethics approvals were obtained from the University of the Free State (approval number: UFS-AED2020/0015/1709) as well as the South African National Biodiversity Institute (approval number: SANBI/RES/P2020/30). Appropriate research permits were also obtained from South African regulatory authorities including the Department of Agriculture, Land Reform, and Rural Development (Section 20 permit: 12/11/1/1/18(1824JD)) and the Department of Environmental Affairs (Threatened Or Protected Species (TOPS) permit: O-52903). An appropriate minimum sample size required was calculated using G*Power 3.1 (Faul et al., 2009) given an anticipated effect size large enough to provide a correlation coefficient of 90% and power of at least 0.8; close to previous approximations with similar data (Mayne et al., 2021). Fifty cheetah samples (25 male, 25 female) were subsequently received from the SANBI Biobank, where they were stored at −80°C. Blood samples were previously collected during routine screening and health checks of captive cheetah from several localities in South Africa (Figure 1). Venous blood was collected from either the leg or tail

in an EDTA collection tube using an 18-gage needle. The samples were selected in such a way that each year of the expected life span of 15 years old was represented. Samples included young cubs ($N = 12$), adolescents ($N = 2$), independents ($N = 8$), young adults ($N = 7$), prime adults ($N = 6$), old adults ($N = 12$), and very old adults ($N = 3$). This project was registered in the National Centre for Biotechnology Information (NCBI) BioProject database as PRJNA737185 and sample details along with NCBI BioSample numbers are listed in Table S1. Mapping was done in QGIS 3.22.6 based on distribution maps for terrestrial mammals from the IUCN (International Union for Conservation of Nature, 2022) used in the assessment of cheetah (Durant et al., 2015).

2.2 | DNA extraction and quantitation

DNA was extracted from 100 µL of whole blood with the commercially available E.Z.N.A.® blood DNA mini kit (Omega Bio-Tek Inc., Norcross, Georgia, USA), according to manufacturer's instructions (Le Clercq, Dalton, et al., 2023b). This kit makes use of the basic digest-bind-wash-elute spin-column technology where the blood sample was digested with a protease and applied to a column with a binding buffer. These columns were centrifuged to remove debris using a wash buffer. The final step was eluting the bound and purified DNA in 50 µL elution buffer. The DNA concentration and purity were determined by spectrophotometric measurement of 1 µL extract applied to the Nanodrop™ 1000 (Thermo Fisher Scientific, Waltham, Massachusetts, USA) instrument. This is achieved by measuring the absorbance at wavelengths A_{260} and A_{260}/A_{280} to report concentration in ng/µl as well as indicate the purity. Additional screening was done by determining concentration and purity using the Qubit® 4 fluorometer (Invitrogen, Thermo Fisher Scientific, Waltham, Massachusetts, USA) with the Qubit® dsDNA Broad Range Assay Kits.

2.3 | Study design and target genes

Methylation status was assayed using methods adapted from human studies (Le Clercq, Dalton, et al., 2023a) for CpG sites within six promising gene regions (Table 2). Genes were selected based on several criteria, including (1) multiple studies (>2) providing evidentiary support of an epigenetic clock, (2) high correlation with age (>70%) in previous studies, (3) presence of multiple, potentially assayable, CpG's within the target region, and (4) availability of coordinate data in the form of Illumina 'cg' numbers (e.g. cg09809672) to facilitate sequence retrieval. Cheetah orthologues were retrieved using human sequences from the University of California Santa Cruz (UCSC) Genome Browser (Haeussler et al., 2019) in basic local alignment search tool (BLAST) searches (Agarwala et al., 2018) against the cheetah reference genome (NCBI RefSeq: GCF027475565.1). The final set of genes were: *EDARADD* (NCBI Gene ID: 106982554), *ELOVL2* (NCBI Gene ID: 106972032), *FHL2* (NCBI Gene ID: 106977955), *GRIA2* (NCBI Gene ID: 106986586), *ITGA2B* (NCBI Gene ID: 106974321), and *PENK* (NCBI Gene ID: 106982523). The degree to which the target regions are conserved among species

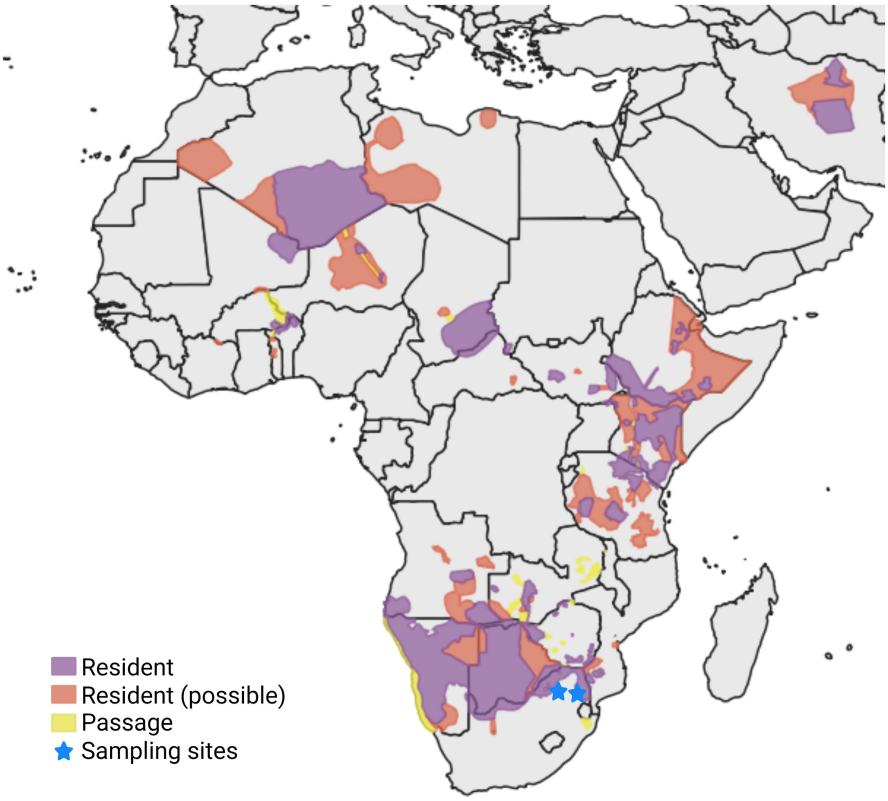

FIGURE 1 Map of distribution for *Acinonyx jubatus* species. Areas indicated in purple represent known resident ranges of cheetah, while pink areas indicate possible additional resident ranges. Areas indicated in yellow represent vagrant ranges where cheetah may occur sporadically as part of passage between established ranges. Sampling sites are indicated by blue stars. Image created in BioRender.com.

TABLE 2 Results from individual linear regressions for CpGs screened in six cheetah genes found to have a near-significant or significant correlation.

Gene	Evidence	CpG	Correlation (R^2)	*p*-Value
EDARADD	(Bocklandt et al., 2011)	**CpG 3**	**.16**	**.007*****
CpGs: 10		CpG 6	.02	.18
		CpG 8	.01	.28
ELOVL2	(Bekaert et al., 2015)	CpG 17	.06	.12
CpGs: 39		**CpG 18/19/20**	**.13**	**.03****
		CpG 21	.01	.26
FHL2	(Giuliani et al., 2016)	CpG 1	.02	.16
CpGs: 16		**CpG 9**	**.12**	**.08***
		CpG 14	.01	.24
GRIA2	(Koch & Wagner, 2011)	CpG 2	.02	.39
CpGs: 6		CpG 3	.02	.38
		CpG 4	**.09**	**.04****
ITGA2B	(Weidner et al., 2014)	CpG 4	.03	.15
CpGs: 6				
PENK	(Giuliani et al., 2016)	**CpG 2**	**.12**	**.05****
CpGs: 23		**CpG 3/4**	**.13**	**.05****
		CpG 6	.05	.11
		CpG 11/12	.02	.22
		CpG 16/17/18	.01	.24
		CpG 19/20	.02	.22

Note: Gene abbreviations are indicated along with the relevant reference study that reported on the use of these genes for age determination. Individual CpG's are listed according to their numbering on the EpiGrams generated with EpiTYPER® with significant CpG's indicated in bold.
*$p < .10$; **$p \leq .05$; ***$p < .01$.

within the family *Felidae* was determined by comparing amplicon sequences for *A. jubatus* to reference genomes using BLAST. Primers were designed using the EpiDesigner website (Agena Bioscience Inc., 2017) and are listed in Table S2. Success of the designed assay was determined in R 4.0.2 (R Core Team, 2020) with the RSeqMeth 1.0.2 (Statham & Csárdi, 2008) and MassArray 1.46.0 (Thompson & Greally, 2020) packages that analyses the predicted amplicon for fragmentation pattern in terms of size and resolution. This gave a direct indication of the number of assayable CpG's per gene. This platform was registered to the NCBI Genome Expression Omnibus (GEO) database as GPL32927.

2.4 | Bisulphite conversion and EpiTYPER® MASS array sequencing

Genomic DNA from all samples was treated with EpiMark® Bisulfite Conversion Kit (New England Biolabs, Ipswich, Massachusetts, USA). Briefly, 10 μL of DNA was mixed with prepared bisulphite mix (sodium metabisulphite) followed by brief vortexing and thermal cycled using alternating denaturation (95°C, 5 min) and incubation (65°C, 30–60–90 s) phases for three cycles. This treatment converted any non-methylated cytosine residues into uracil, while methylated cytosine residues remained unaffected. Samples were purified with spin columns, as for extractions, using desulphonation and wash buffers. Finally, the converted DNA was eluted and quantitated. Complete bisulphite conversion of DNA was monitored with 'TpC' internal conversion controls (Suchiman et al., 2015) that have been identified within each gene amplicon. At a 'TpC', the 'C' is unmethylated and is expected to produce 100% 'TpA' sequence reads in the final data for those sites when conversion was complete.

The EpiTYPER® Assay methylation analysis was run in three phases, starting with (1) PCR using T7-promoter tagged reverse primers to amplify the target regions while preserving the bisulphite-induced sequence changes and treatment of PCR product with Shrimp Alkaline Phosphatase, (2) in vitro transcription was performed, and (3) the resulting RNA transcripts were specifically cleaved at uracil residues. The resulting fragments differed in size and mass, depending on the sequence changes generated through bisulphite treatment, and through enzymatic cleavage of uracils. The resulting fragments were loaded onto a SpectroCHIP® Array (Chip) that was placed in the MASSarray (MALDI-TOF) mass spectrometer for data acquisition. Results showed individual peaks for the generated fragments. Fragmentation pattern was analysed for mass for each fragment (based on size) and the peak height, which corresponds to the percentage of methylation compared to an internal control. The output was automatically loaded into a database for data analysis with EpiTYPER® software.

2.5 | Statistical analysis

The percentage of methylation per gene site was compared to known age of each sample. Methylation data were first analysed for normality and 'goodness of fit' to a Gaussian distribution using the Anderson-Darling test for normality to eliminate possible errors due to skewedness or kurtosis. Hereafter, a standard linear regression was performed individually for each gene and CpG to identify those with the strongest correlation in RStudio 1.4.1106 (Rstudio Team, 2021), running R version 4.0.5 (R Core Team, 2020). Subset analyses to identify the best targets for regression were explored with *leaps* version 3.1 (Lumley, 2020). Those sites with a significant correlation coefficient and probability (at least $p < .10$), explaining the larger part of the Sum of Standard Error (SSE), were used to construct and test multiple-regression models.

2.6 | Model training and validation

Models were trained using machine learning approaches implemented through elastic net regression with the *glmnet* version 4.1–7 (Friedman et al., 2010) and *caret* version 6.0–94 (Kuhn, 2008) packages in R. Models were optimized based on the alpha (α) and lambda (λ) parameters. Furthermore, model performance was accessed based on the values obtained for the Root Mean Standard Error (RMSE) of the model, as per Equation 1, as well as the correlation coefficient (R^2) for the model. Here, the RMSE is derived from the square root of the sum for the quotient of the squared difference between the sample mean ($\hat{y}_i$) and the inferred value (y_i) and the total sample size (n).

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{\left(\hat{y}_i - y_i\right)^2}{n}} \quad (1)$$

These models were validated using Leave-Group-Out Cross-Validation (LGOCV), also known as Monte Carlo Cross-Validation (MCCV), best suited to hierarchical data such as age groups and smaller sample sizes (Shan, 2022; Shao, 1993). LGOCV has the added benefit of avoiding over-fitting while being able to subsample test sets of variable size (between 50% and 75% of the data). This method was applied using fivefold repeats ($k = 5$) with subsample sets of 39 samples (75%) from the test data ($N = 50$). The accuracy of the model predictions was expressed as the Mean Absolute Error (MAE) in months, according to Equation 2. Here, the MAE is derived from the sum of the absolute difference between the inferred (y_1) and actual (x_1) values, divided by the sample size (n).

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} \mid y_1 - x_1 \mid}{n} \quad (2)$$

The final model predictions were assessed for both validity and reliability (Cohen et al., 2017; Hammersley, 1987) on a subset of 12 samples, in lieu of samples of unknown ages, that had an equal distribution of males and females and included at least one sample per cheetah age group. Validity, in this context, refers to the ability of the metric to accurately measure the intended attribute. This was verified by comparing the known chronological age of individuals to their predicted age from the model. Reliability of the measure was assessed by comparing the age class assignment correctness for the predicted ages. Although the ultimate purpose of designing an age estimation model such as that presented in this paper is to apply the panel to wild animals of unknown age, due to travel restrictions related to reducing human-to-human spread of COVID and concern about potential transmission of COVID to cheetahs (Heeney et al., 1990; Kennedy et al., 2002; Tewari et al., 2023), the application to wild animals was beyond the scope of this study at the time the study was carried out.

3 | RESULTS

3.1 | CpG screening and selection

Orthologous gene regions were identified for cheetah in all six of the selected target genes. Initial estimates based on sequence revealed a total of 100 CpG's (Table 2) within the target regions. Fragmentation patterns for the selected amplicons in R showed that once fragmentation is considered a total of 77 CpG clusters that could potentially be screened using EpiTYPER® mass assay analyses. EpiTYPER® analyses (see Figure S1) were run on 50 samples. Good amplification was observed; however, eight of the screened CpG fragments fell outside the emission spectra and were not quantifiable: two for *EDARADD*, four for *ELOVL2*, one for *FHL2*, and one for *ITGA2B*. Linear models constructed for the relationship between age and differential methylation detected some correspondence in 19 of the remaining CpG's, of which six showed a statistically significant relationship (Table 2, Figure 2). These CpG's were present in five of the six candidate genes, excluding *ITGA2B*, with four CpG's showing hypomethylation with age and two showing hypermethylation. For *EDARADD*, CpG 3 showed progressive hypomethylation with methylation levels ranging from $\beta = .01$ to $\beta = .29$ and a significant ($F = 8.11$, $R^2 = .16$, $p < .01$) relationship with age. *FHL2* showed similar hypomethylation for CpG 9, with values ranging between $\beta = .01$ and $\beta = .28$ and a significant ($F = 3.28$, $N = 25$, $R^2 = .12$, $p < .10$) age correspondence. *PENK* was the only gene with several significant CpG's, both CpG 2 ($F = 4.1$, $R^2 = .12$, $p \leq .05$) and the cluster for CpG's 3 and 4 ($F = 4.32$, $R^2 = .13$, $p \leq .05$), with beta values of 0.21–0.56 and 0.50–0.84, respectively. *ELOVL2* (CpG cluster of 18, 19, and 20) and *GRIA2* (CpG 4) both showed hypermethylation with age. For *ELOVL2*, methylation levels ranged from $\beta = .01$ to $\beta = .14$ and reached statistical significance ($F = 3.99$, $R^2 = .14$, $p < .05$). Similar to *EDARADD* and *FHL2*, *GRIA2* showed methylation levels ranging between $\beta = .04$ and $\beta = .29$ but was significantly hypermethylated ($F = 4.35$, $R^2 = .09$, $p < .05$) with age. Average methylation levels for all CpG's per gene were not significant with only *PENK* showing a

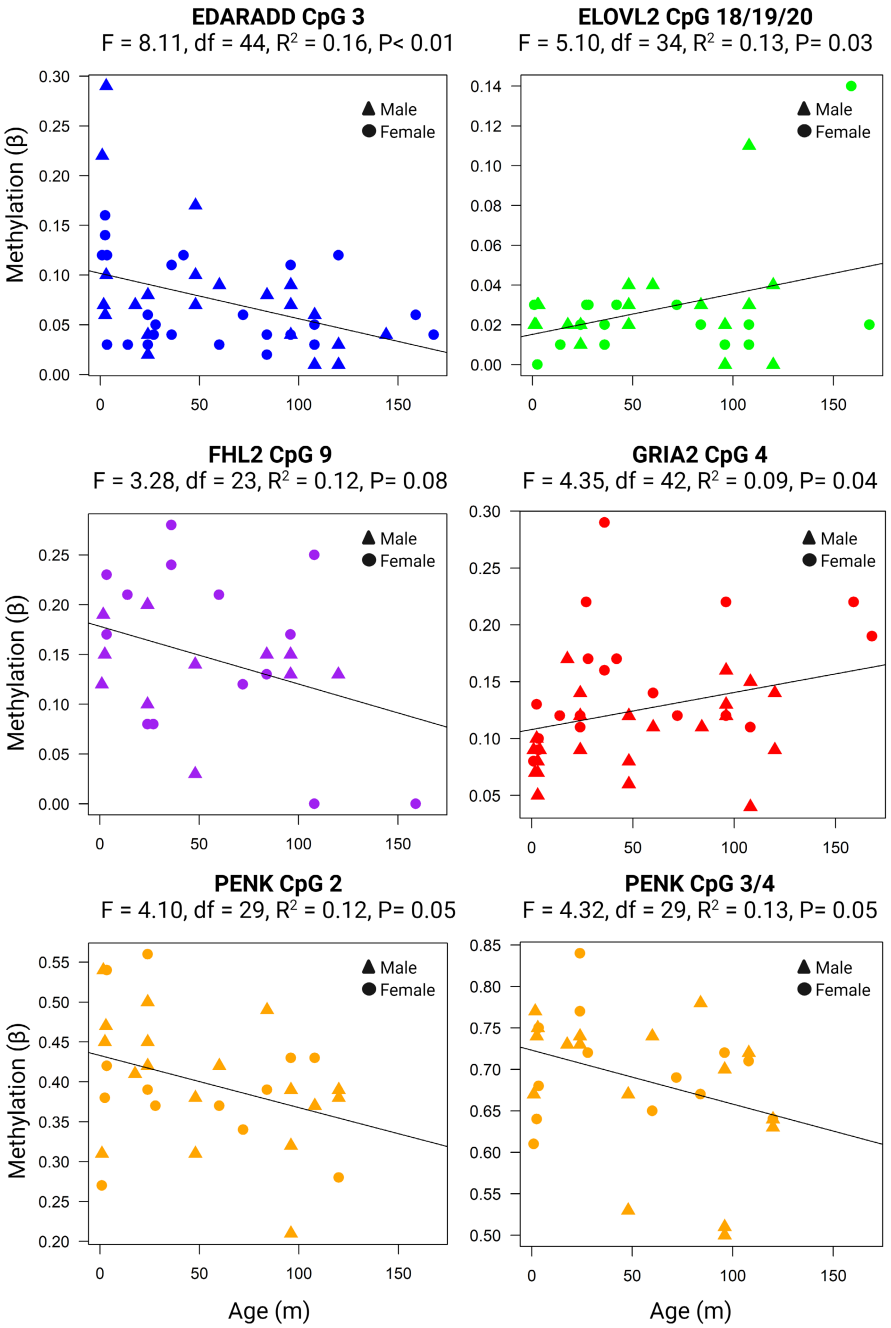

FIGURE 2 Scatter plots for the six CpGs identified as having a significant relationship with age. The y-axis indicates the relative level of methylation as expressed by the β-values, while the *x*-axis indicates the known ages expressed as months. Four CpG's, in the EDARADD (blue), FHL2 (purple), and PENK (orange) genes, showed a linear trend on hypomethylation with age and another two CpG's, in the ELOVL2 (green) and GRIA2 (red) genes, showed a linear trend of hypermethylation with age. Males are indicated by triangles, while females are indicated by circles. No statistically significant trend for methylation was detected for sex. Image created in BioRender.com.

near-significant correlation ($F = 2.64$, $R^2 = .08$, $p = .11$). A CpG subset analysis prior to model building identified four CpG's in *FHL2*, *GRIA2*, and *PENK*, as the most promising targets.

3.2 | Creating an age estimation model

The six significant CpG's were used to train and test several elastic net regression models using between two and six CpG's (Table 3). Models generated with fewer CpG's tended to have higher RMSE and MAE values and explained a lower percentage of the variance based on the correlation coefficient. For example, both the model based on two CpG's and the model based on three CpG's had an RMSE in the forties and explained less than 50% of the variance. The two best models were based on six CpG's (RMSE = 31.30, MAE = 25, $R^2 = .70$) and four CpG's (RMSE = 30.29, MAE = 25, $R^2 = .59$). While both models performed similarly based on RMSE and MAE, the model based on six CpG's performed better than the four CpGs by explaining a higher percentage of the variance as assessed by the correlation coefficient.

3.3 | Characteristics of the age estimation model

The best model was based on six CpG's and had an alpha value of .05 and lambda of 0.01. Subsets sampled from the dataset used to

TABLE 3 Results from machine learning experiments to create a model for age prediction in cheetah.

Model type	RMSE	MAE	R^2
6 CpG's ($\alpha = .05$, $\lambda = 0.01$)	31.30	25	.70
EDARADD CpG 3 + ELOVL2 CpG 18/19/20 + FHL2 CpG 9 + GRIA2 CpG 4 + PENK CpG 2 + PENK CpG 3/4			
5 CpG's ($\alpha = .97$, $\lambda = 5.17$)	35.31	29	.54
EDARADD CpG 3 + ELOVL2 CpG 18/19/20 + GRIA2 CpG 4 + PENK CpG 2 + PENK CpG 3/4			
4 CpG's ($\alpha = .05$, $\lambda = 0.001$)	30.29	25	.59
EDARADD CpG 3 + FHL2 CpG 9 + GRIA2 CpG 4 + PENK CpG 3/4			
3 CpG's ($\alpha = .07$, $\lambda = 5.90$)	40.09	35	.36
EDARADD CpG 3 + ELOVL2 CpG 18/19/20 + GRIA2 CpG 4			
2 CpG's ($\alpha = .73$, $\lambda = 6.28$)	48.56	43	.14
EDARADD CpG 3 + ELOVL2 CpG 18/19/20			

Note: Five models were tested using between two and six CpG's, respectively. In each step, the CpG with the lowest significance was omitted with the exception of the four CpG models which was based on the best targets from the subset analyses. Models were compared in terms of Root Mean Standard Error (RMSE), Mean Absolute Error (MAE), and correlation (R^2). The model with the lowest RMSE and MAE and highest correlation was based on six CpG's and was selected for validation.

train the model for validation testing using LGOCV performed well (Figure 3, $R^2 = .70$), with most samples having a near-perfect correlation between predicted and chronological ages with an accuracy of approximately 25 months or 2 years and 1 month of their actual age. When comparing predicted age versus chronological age, using only 12 samples, a high degree of correlation was observed ($R^2 = .65$) with most samples falling within the 95% confidence interval. There was, however, some evidence that the accuracy of the model is dependent on the age class as estimates for young individuals and very old adults showed the greatest deviance from their chronological age. Using predicted age for age class assignments, more than 80% of samples were assigned to the correct age class based on the predicted age from the model, showing a reasonably high level of correspondence.

3.4 | Degree of gene conservation within the family *Felidae*

Species within the family *Felidae* diverged from approximately 12 million years ago (MYA). Using gene sequences from *Acinonyx jubatus*, BLAST matches were retrieved for 14 additional species from seven genera within the family *Felidae* including the most closely related *Puma*, *Lynx*, and *Felis* species as well as the more distant *Panthera* species (Figure 4). BLAST matches ranged from 75% to 97% for gene similarity depending upon the gene and species with *EDARADD* and *GRIA2* returning the highest number of matches. As expected, species most closely related to cheetah, including the *Puma*, *Lynx*, *Felis*, and *Panthera* genera, showed the highest degree of conservation. Most species within this family also have similar lifespans of between 10 and 20 years in the wild and captivity, respectively (Le Clercq, Kotzé, et al., 2023), and several species from the matches have previously been used as model species in age and methylation studies. This includes studies in the bobcat (Cantrell et al., 2020; Lachance et al., 2015), *Lynx rufus*, domestic cat (Qi et al., 2021), *Felis catus*, lion (Raj et al., 2021), snow leopard (Qi et al., 2021), and tiger (Raj et al., 2021).

4 | DISCUSSION

This study investigated the potential use of age-related changes in methylation within candidate genes, previously identified in human studies, to establish an age estimation model in wildlife using mass array technology. For this purpose, orthologous gene regions from numerous human studies were retrieved for six target genes with a hundred CpG's within the gene regions of interest. Mass array technology was chosen as a sensitive but cost-effective alternative to existing methods, to assay differential methylation. The predicted fragmentation patterns resulted in a total of 77 CpG clusters that could be analysed for the study species. Screening, performed on 50 samples, provided data for 69 of the fragments: the remainder falling outside the weight spectra for the analyses. Correlation analyses between CpG methylation and known chronological age identified six CpG's with a significant relationship of which four showed progressive hypomethylation with age while two showed hypermethylation. Regression models were fitted using different combinations of CpG's and the optimal model, based on performance metrics and validation experiments, indicated that age models using either four or six CpG's were most accurate. The model based on six CpG's did, however, outperform the four CpG models in terms of correlation and predictive power.

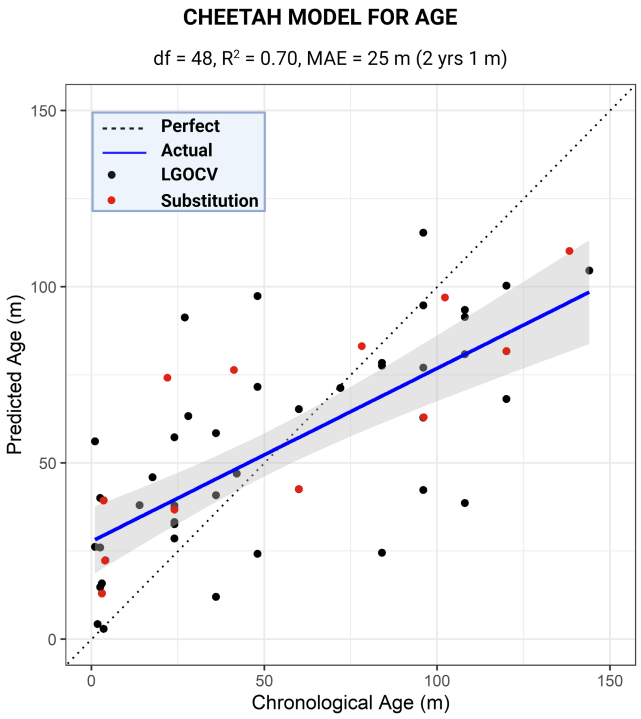

FIGURE 3 Scatter plot for the validation testing of the cheetah model for predicting age based on methylation at six CpG's. The model was validated using Leave-Group-Out Cross-Validation (LGOCV) on a test set of randomly selected samples from the original set used to train the model (black) as well as through the resubstitution of 12 samples (red). The age, as predicted by the methylation model, is plotted against the known chronological age of the individual in months. The dotted line indicated a theoretically perfect correspondence. The blue line shows the regression for predicted versus chronological age and the 95% confidence interval (grey). The model showed a high degree of accuracy (MAE = 25 months or 2 years and 1 month, $R^2 = .70$, $p < .05$) with many samples having a linear correlation between their predicted and actual ages and >55% of samples falling within or close to the 95% confidence interval. Image created in BioRender.com.

Previous studies aimed at creating age estimation models based on differential CpG methylation in candidate genes have used as little as one to four (Nakamura et al., 2023; Weidner et al., 2014) CpG's to subsets of seven (Freire-Aradas et al., 2016; Polanowski et al., 2014; Wright et al., 2018) or eight (Fleckhaus & Schneider, 2020). While most of these studies also used similar methods, such as pyromark bisulphite sequencing, methods such as methylation-sensitive PCR (Mawlood et al., 2016) and mass array technology (Freire-Aradas et al., 2016; Giuliani et al., 2016) have also proven useful. Of the six genes assayed in this study, six CpG's in five genes showed a correlation with age for cheetah. For most genes, apart from several CpG's in PENK, methylation levels were below 50% and differed between age classes by less than 10%. This was consistent with previous findings for *EDARADD* (Bekaert et al., 2015; De Paoli-Iseppi et al., 2017; Fleckhaus & Schneider, 2020; Pan et al., 2020), *ELOVL2* (Giuliani et al., 2016; Márquez-Ruiz et al., 2020; Spiers et al., 2016), *FHL2* (Dias et al., 2020; Fleckhaus & Schneider, 2020; Giuliani et al., 2016), *GRIA2* (Koch & Wagner, 2011; Mawlood et al., 2016; Tanabe et al., 2020), and *PENK* (Garagnani et al., 2012; Giuliani et al., 2016). Four CpG's showed hypomethylation which is characteristic of global hypomethylation associated with age (Le Clercq, Kotzé, et al., 2023) or epigenetic drift (Zampieri et al., 2015); however, the strong correlation with age may indicate that methylations at these sites are not random. Furthermore, two CpG's showed evidence of hypermethylation. As such, these CpG's likely represent highly regulated and site-directed epigenetic clocks (Horvath & Raj, 2018).

The best model for cheetah age was generated using six CpG's and was able to accurately predict age in validation experiments, with a correlation of 0.70 and mean absolute error of 2 years and 1 month, accounting for a large percent of the observed variance. A systematic review and meta-analysis comparing methylation-based age estimation models across several vertebrate classes found that the average effect for included studies was approximately 0.92 with a confidence interval ranging between 0.89 and 0.94 (Le Clercq, Kotzé, et al., 2023). While the model of this study falls below this range, it is comparable to models from several previous studies including 0.74 for beluga whales (Bors et al., 2021), chimpanzees (Ito et al., 2018), grey wolves (Thompson et al., 2017), and sheep (Sugrue et al., 2021); and between 0.76 and 0.79 for baboons (Anderson et al., 2021), bottlenose dolphins (Beal et al., 2019), domestic dog (Thompson et al., 2017), and humpback whales (Polanowski et al., 2014). In comparison to previous model attempts in cheetah, mass array proved more useful than methylation-sensitive PCR (Le Clercq et al., 2018) but was slightly less accurate compared to cheetah samples tested using a feline-specific model from probe-based array bisulphite sequencing (Raj et al., 2021). Furthermore, while the error of the estimate was approximately 2 years, the functional age classes of cheetah typically span more than 2 years of age for adults (Marker & Dickman, 2003), meaning the model could still classify individuals to a close enough approximation of chronological age.

In comparison to previous studies for *Felidae* species, mass array provided more informative data than studies using global methylation detection methods such as ELISA as applied to domestic cats (Tamazian et al., 2014), snow leopards (Jabbari et al., 1997), and bobcats (Cantrell et al., 2020). The study model also performed better than previous models for the domestic cat ($R^2 = .41$, MAE = 2 years and 11 months) and snow leopard ($R^2 = .63$, MAE = 2 years and 10 months), which used methylation-sensitive PCR and included 13 CpG's in the *ELOVL2* gene (Qi et al., 2021). It was, however, less accurate than models tested in lions and tigers (Raj et al., 2021). BLAST searches revealed that many of the genes used to design the assay for cheetah are conserved in other *Felidae*, and possibly carnivore, species with similar lifespans. As such future research could extend the method described here to other species either directly, for those with a near-perfect sequence match, or through only minor modifications to the primers.

It should be noted that there are several limitations that were identified. First, the model predictions were most accurate in adult age classes and performed poorly on young individuals such

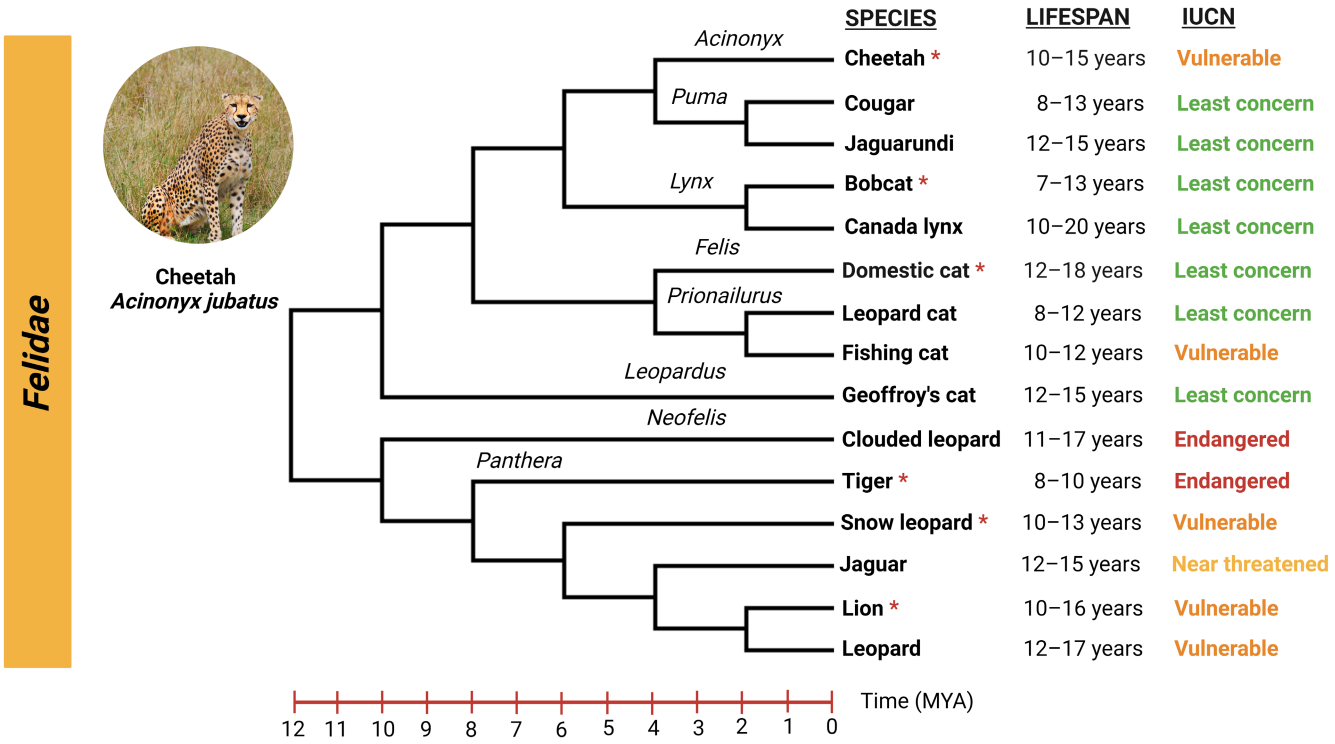

FIGURE 4 Time tree for *Felidae* species for which BLAST matches were identified from study genes. Species within this family diverged over the past 12 million years. Matches were retrieved for 14 additional species from 7 genera within the family *Felidae* including 5 species used in previous methylation studies (red asterisks). BLAST matches ranged between 75% and 97% gene conservation depending upon the gene, with the *Puma*, *Lynx*, and *Felis* genera showing the highest degree of gene similarity compared to *Acinonyx*. Most species within this family also have similar lifespans of between 10 and 20 years in the wild and captivity, respectively. In terms of species vulnerability as assessed by the IUCN, six species are of least concern while one species is classified as near threatened, five species are classified as vulnerable, and two species are classified as endangered. Image created in BioRender.com.

as cubs and adolescents. This does not decrease the overall utility of the method, considering other age estimation methods are highly accurate in the younger classes and lose resolution in adult classes (Marker & Dickman, 2003) but should be considered when using methylation-based models. More specifically, future studies should consider if the inclusion of equal numbers for each functional age class is desirable or if a focus should be placed on adult age classes. Second, a large amount of variance in methylation levels was observed between individuals of the same age, which increased the mean absolute error of the predictions. Variance can often be reduced through either including larger sample sizes for a given measure or using a method with greater sensitivity (James, 1985). While sample size was considered in the study design and based on previously recommended sample sizes (Mayne et al., 2021), some samples did fail for difficult-to-amplify targets such as *FHL2* and *PENK*, resulting in fewer data being available for these regions than is ideal with many of the available data being for younger individuals such as cubs that were less informative. Lastly, the model presented here is population-specific and was calibrated from samples of individuals in captivity and, considering the effects of stress (Poganik et al., 2022), environment (Viñuela et al., 2018), and damage (Kujoth et al., 2005) on epigenetics, may not be suitable to age determination in wild cheetah without further testing.

In conclusion, the present study was able to use age-correlated CpG's from human studies to identify orthologous regions in an unrelated mammal for which the correlations were conserved between species. This enabled the development of an age estimation model that performed similarly to previous studies using a candidate gene approach and was able to assign adult individuals to their relevant functional age classes. Therefore, methylation measured by mass array is a useful molecular tool for the development of age estimation models and could provide useful age information for individuals and populations. Furthermore, the flexibility of the assay design procedure described here could assist in the development of methylation-based age determination methods in most vertebrates. Thus, mass array is a cost-effective alternative to whole-genome bisulphite sequencing approaches and may find widespread applications in conservation—particularly for threatened, vulnerable, and endangered species subject to strict local and global regulations for sampling.

ACKNOWLEDGEMENTS

Images were created in BioRender.com. We would like to thank Professor David Burney for providing a scanned copy of his M.Sc. dissertation as the original was lost during student unrest and political activity on the University of Nairobi campus. We would also like to thank our Post Graduate Research Librarian (UFS), Carmen Nel,

for assisting in the retrieval of several difficult-to-find resources. Parts of this research were presented at the Student Symposium of the South African Society for Genetics. This work is based on the research supported wholly/in part by the National Research Foundation of South Africa (Grant Number: 112062).

CONFLICT OF INTEREST STATEMENT

The authors have no competing interests to declare.

DATA AVAILABILITY STATEMENT

Sample details were deposited in the National Centre for Biotechnology Information (NCBI) BioSample database (SAMN31850912-60), while CpG methylation data generated in this study were deposited to the NBCI Genome Expression Omnibus (GEO) database as dataset GSE252541, associated with GEO platform GPL32927. Accession numbers linked to NCBI BioProject PRJNA737185 are listed in Table S1. The method used to design the assay is available on Protocols.io (Le Clercq, Dalton, et al., 2023a). Additional data including R code used for data analyses are available from the Zenodo depository at the following link: https://doi.org/10.5281/zenodo.10159400. The PYTHON scripted algorithm based on the study model, useful to calculate the age estimates of unknown specimens, is under development and will be available on GitHub at the following link: https://github.com/LSLeClercq/ChEATAE.

BENEFIT-SHARING STATEMENT

Benefits from this research accrue from the sharing of our data and results on public databases as described above. This is particularly important considering that cheetah, used as a case study, are classified as vulnerable and access to sufficient samples may be restricted to specific parts of Africa. As we are aware of historical inequality and underrepresentation of specific demographics within the scientific community and publication, efforts were made to ensure a diverse and balanced author team. This paper includes an LGBT scientist, as well as two authors who are women. The authors are from four institutions and two countries including the UK and South Africa.

ORCID

Louis-Stéphane Le Clercq https://orcid.org/0000-0002-8713-8920
Antoinette Kotzé https://orcid.org/0000-0003-2367-1483
J. Paul Grobler https://orcid.org/0000-0002-5913-7031
Desiré L. Dalton https://orcid.org/0000-0001-5975-6425

REFERENCES

Adamson, J. (1970). *Pippa, the cheetah, and her cubs*. Harcourt.

Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C., Bolton, E., Bourexis, D., Brister, J. R., Bryant, S. H., Canese, K., Cavanaugh, M., Charowhas, C., Clark, K., Dondoshansky, I., Feolo, M., Fitzpatrick, L., Funk, K., Geer, L. Y., Gorelenkov, V., … Zbicz, K. (2018). Database resources of the National Center for biotechnology information. *Nucleic Acids Research*, *46*(D1), D8–D13. https://doi.org/10.1093/nar/gkx1095

Agena Bioscience Inc. (2017). *EpiDesigner*. http://www.epidesigner.com/

Anderson, J. A., Johnston, R. A., Lea, A. J., Campos, F. A., Voyles, T. N., Akinyi, M. Y., Alberts, S. C., Archie, E. A., & Tung, J. (2021). High social status males experience accelerated epigenetic aging in wild baboons. *eLife*, *10*, e66128. https://doi.org/10.7554/ELIFE.66128

Beal, A. P., Kiszka, J. J., Wells, R. S., & Eirin-Lopez, J. M. (2019). The bottlenose dolphin epigenetic aging tool (BEAT): A molecular age estimation tool for small cetaceans. *Frontiers in Marine Science*, *6*, 561. https://doi.org/10.3389/fmars.2019.00561

Bekaert, B., Kamalandua, A., Zapico, S. C., Van De Voorde, W., & Decorte, R. (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

Black, S. A., Fellous, A., Yamaguchi, N., & Roberts, D. L. (2013). Examining the extinction of the barbary lion and its implications for felid conservation. *PLoS One*, *8*(4), e60174. https://doi.org/10.1371/journal.pone.0060174

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/JOURNAL.PONE.0014821

Bors, E. K., Baker, C. S., Wade, P. R., O'Neill, K. B., Shelden, K. E. W., Thompson, M. J., Fei, Z., Jarman, S., & Horvath, S. (2021). An epigenetic clock to estimate the age of living beluga whales. *Evolutionary Applications*, *14*(5), 1263–1273. https://doi.org/10.1111/eva.13195

Brelsford, A., Milá, B., & Irwin, D. E. (2011). Hybrid origin of Audubon's warbler. *Molecular Ecology*, *20*(11), 2380–2389. https://doi.org/10.1111/j.1365-294X.2011.05055.x

Broom, R. (1949). Notes on the milk dentition of the lion, leopard and cheetah. *Annals of the Transvaal Museum*, *21*(1), 183–185.

Burney, D. A. (1980). *The effects of human activities on cheetah (Acinonyx jubatus) in the Mara region of Kenya*. Masters dissertation. University of Nairobi, Kenya.

Cantrell, B., Friedman, S., Lachance, H., Bernier, C., Murdoch, B., Frattini, S., Talenti, A., Crepaldi, P., & McKay, S. (2020). A novel understanding of global DNA methylation in bobcat (*Lynx rufus*). *Genome*, *63*, 125–130.

Cardoza, J. E., & Langlois, S. A. (2002). The eastern cougar: A management failure? *Wildlife Society Bulletin*, *30*(1), 265–273.

Caro, T. (1994). In G. B. Schaller (Ed.), *Cheetahs of the Serengeti Plains: Group living in an asocial species* (1st ed.). University of Chicago Press.

Charruau, P., Fernandes, C., Orozco-Terwengel, P., Peters, J., Hunter, L., Ziaie, H., Jourabchian, A., Jowkar, H., Schaller, G., Ostrowski, S., Vercammen, P., Grange, T., Schlötterer, C., Kotze, A., Geigl, E. M., Walzer, C., & Burger, P. A. (2011). Phylogeography, genetic structure and population divergence time of cheetahs in Africa and Asia: Evidence for long-term geographic isolates. *Molecular Ecology*, *20*(4), 706–724. https://doi.org/10.1111/j.1365-294X.2010.04986.x

Cohen, L., Manion, L., & Morrison, K. (2017). Validity and reliability. In *Research methods in education* (Vol. *44*, 8th ed., pp. 245–284). Routledge. https://doi.org/10.4324/9781315456539-14

Cooper, E. B., Bonnet, T., Osmond, H. L., Cockburn, A., & Kruuk, L. E. B. (2021). Aging and senescence across reproductive traits and survival in superb fairy-wrens (*Malurus cyaneus*). *American Naturalist*, *197*(1), 111–127. https://doi.org/10.1086/711755

De Paoli-Iseppi, R., Polanowski, A. M. A. M., McMahon, C., Deagle, B. E. B. E., Dickinson, J. L. J. L., Hindell, M. A. M. A., & Jarman, S. N. S. N. (2017). DNA methylation levels in candidate genes associated with chronological age in mammals are not conserved in a long-lived seabird. *PLoS One*, *12*(12), 1–20. https://doi.org/10.1371/journal.pone.0189181

Dias, H. C., Cordeiro, C., Pereira, J., Pinto, C., Real, F. C., Cunha, E., & Manco, L. (2020). DNA methylation age estimation in blood samples

of living and deceased individuals using a multiplex SNaPshot assay. *Forensic Science International*, *311*, 110267. https://doi.org/10.1016/j.forsciint.2020.110267

Double, M. C., Dawson, D., Burke, T., & Cockburn, A. (1997). Finding the fathers in the least faithful bird: A microsatellite-based genotyping system for the superb fairy-wren *Malurus cyaneus*. *Molecular Ecology*, *6*(7), 691–693. https://doi.org/10.1046/J.1365-294X.1997.00228.X

Durant, S. M., Mitchell, N., Groom, R., Pettorelli, N., Ipavec, A., Jacobson, A. P., Woodroffe, R., Böhm, M., Hunter, L. T. B., Becker, M. S., Broekhuis, F., Bashir, S., Andresen, L., Aschenborn, O., Beddiaf, M., Belbachir, F., Belbachir-Bazi, A., Berbash, A., de Matos Machado, I. B., … Young-Overton, K. (2017). The global decline of cheetah *Acinonyx jubatus* and what it means for conservation. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(3), 528–533. https://doi.org/10.1073/pnas.1611122114

Durant, S. M., Mitchell, N., Ipavec, A., & Groom, R. (2015). *Acinonyx jubatus*, cheetah. The IUCN Red List of Threathened Species, 8235.

Fàbregas, M. C., & Garcés-Narro, C. (2014). Validation of gum-line recession as a reliable technique to age tigers. *European Journal of Wildlife Research*, *60*(6), 947–950. https://doi.org/10.1007/s10344-014-0869-1

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fitch, H. M., & Fagan, D. A. (1982). Focal palatine erosion associated with dental malocclusion in captive cheetahs. *Zoo Biology*, *1*(4), 295–310. https://doi.org/10.1002/zoo.1430010403

Fleckhaus, J., & Schneider, P. M. (2020). Novel multiplex strategy for DNA methylation-based age prediction from small amounts of DNA via pyrosequencing. *Forensic Science International: Genetics*, *44*, 102189. https://doi.org/10.1016/j.fsigen.2019.102189

Freire-Aradas, A., Phillips, C., Mosquera-Miguel, A., Girón-Santamaría, L., Gómez-Tato, A., de Casares Cal, M., Álvarez-Dios, J., Ansede-Bermejo, J., Torres-Español, M., Schneider, P. M., Pośpiech, E., Branicki, W., Carracedo, Á., & Lareu, M. V. (2016). Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system. *Forensic Science International: Genetics*, *24*, 65–74. https://doi.org/10.1016/j.fsigen.2016.06.005

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22. https://doi.org/10.18637/jss.v033.i01

Garagnani, P., Bacalini, M. G., Pirazzini, C., Gori, D., Giuliani, C., Mari, D., Di Blasio, A. M., Gentilini, D., Vitale, G., Collino, S., Rezzi, S., Castellani, G., Capri, M., Salvioli, S., & Franceschi, C. (2012). Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell*, *11*(6), 1132–1134. https://doi.org/10.1111/acel.12005

Giuliani, C., Cilli, E., Bacalini, M. G., Pirazzini, C., Sazzini, M., Gruppioni, G., Franceschi, C., Garagnani, P., & Luiselli, D. (2016). Inferring chronological age from DNA methylation patterns of human teeth. *American Journal of Physical Anthropology*, *159*(4), 585–595. https://doi.org/10.1002/ajpa.22921

Gopalan, S., Carja, O., Fagny, M., Patin, E., Myrick, J. W., McEwen, L. M., Mah, S. M., Kobor, M. S., Froment, A., Feldman, M. W., Quintana-Murci, L., & Henn, B. M. (2017). Trends in DNA methylation with age replicate across diverse human populations. *Genetics*, *206*(3), 1659–1674. https://doi.org/10.1534/genetics.116.195594

Haeussler, M., Zweig, A. S., Tyner, C., Speir, M. L., Rosenbloom, K. R., Raney, B. J., Lee, C. M., Lee, B. T., Hinrichs, A. S., Gonzalez, J. N., Gibson, D., Diekhans, M., Clawson, H., Casper, J., Barber, G. P., Haussler, D., Kuhn, R. M., & Kent, W. J. (2019). The UCSC Genome Browser database: 2019 update. *Nucleic Acids Research*, *47*(D1), D853–D858. https://doi.org/10.1093/nar/gky1095

Hammersley, M. (1987). Some notes on the terms 'Validity' and 'Reliability'. *British Educational Research Journal*, *13*(1), 73–82. https://doi.org/10.1080/0141192870130107

Heeney, J. L., Evermann, J. F., McKeirnan, A. J., Marker-Kraus, L., Roelke, M. E., Bush, M., Wildt, D. E., Meltzer, D. G., Colly, L., & Lukas, J. (1990). Prevalence and implications of feline coronavirus infections of captive and free-ranging cheetahs (*Acinonyx jubatus*). *Journal of Virology*, *64*(5), 1964–1972. https://doi.org/10.1128/JVI.64.5.1964-1972.1990

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), 3156. https://doi.org/10.1186/gb-2013-14-10-r115

Horvath, S., Lu, A. T., Haghani, A., Zoller, J. A., Li, C. Z., Lim, A. R., Brooke, R. T., Raj, K., Serres-Armero, A., Dreger, D. L., Hogan, A. N., Plassais, J., & Ostrander, E. A. (2022). DNA methylation clocks for dogs and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(21), e2120887119. https://doi.org/10.1073/pnas.2120887119

Horvath, S., & Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics*, *19*(6), 371–384. https://doi.org/10.1038/s41576-018-0004-3

International Union for Conservation of Nature (IUCN). (2022). *Digital Distribution Maps on The IUCN Red List of Threatened Species* (p. Version 6.2). p. Version 6.2. https://www.iucnredlist.org/resources/spatial-data-download

Ito, G., Yoshimura, K., & Momoi, Y. (2017). Analysis of DNA methylation of potential age-related methylation sites in canine peripheral blood leukocytes. *Journal of Veterinary Medical Science*, *79*(4), 745–750. https://doi.org/10.1292/jvms.16-0341

Ito, H., Udono, T., Hirata, S., & Inoue-Murayama, M. (2018). Estimation of chimpanzee age based on DNA methylation. *Scientific Reports*, *8*(1), 1–5. https://doi.org/10.1038/s41598-018-28318-9

Jabbari, K., Cacciò, S., Païs De Barros, J. P., Desgrès, J., & Bernardi, G. (1997). Evolutionary changes in CpG and methylation levels in the genome of vertebrates. *Gene*, *205*(1–2), 109–118. https://doi.org/10.1016/S0378-1119(97)00475-7

James, B. A. P. (1985). Variance reduction techniques. *Journal of the Operational Research Society*, *36*(6), 525–530. https://doi.org/10.1057/JORS.1985.88/METRICS

Johansson, Å., Enroth, S., & Gyllensten, U. (2013). Continuous aging of the human DNA methylome throughout the human lifespan. *PLoS One*, *8*(6), e67378. https://doi.org/10.1371/journal.pone.0067378

Kelly, M. J., Laurenson, M. K., FitzGibbon, C. D., Collins, D. A., Durant, S. M., Frame, G. W., Bertram, B. C., & Caro, T. M. (1998). Demography of the Serengeti cheetah (*Acinonyx jubatus*) population: The first 25 years. *Journal of Zoology*, *244*(4), 473–488. https://doi.org/10.1111/J.1469-7998.1998.TB00053.X

Kennedy, M., Citino, S., Hillis McNabb, A., Serino Moffatt, A., Gertz, K., & Kania, S. (2002). Detection of feline coronavirus in captive Felidae in the USA. *Journal of Veterinary Diagnostic Investigation*, *14*(6), 520–522. https://doi.org/10.1177/104063870201400615

Kim, K. S., & Sappington, T. W. (2013). Microsatellite data analysis for population genetics. *Methods in Molecular Biology*, *1006*, 271–295. https://doi.org/10.1007/978-1-62703-389-3_19

Knight, J., Flack-Davison, E., Engelbrecht, S., Visagie, R. G., Beukes, W., Coetzee, T., Mwale, M., & Ralefala, D. (2022). A literature review analysis of engagement with the Nagoya Protocol, with specific application to Africa. *South African Journal of Bioethics and Law*, *15*(2), 69–74. https://doi.org/10.7196/SAJBL.2022.v15i2.811

Koch, C. M., & Wagner, W. (2011). Epigenetic-aging-signature to determine age in different tissues. *Aging*, *3*(10), 1018–1027.

Krausman, P. R., & Morales, S. M. (2005). Acinonyx jubatus. *Mammalian Species*, *771*(1), 1. https://doi.org/10.1644/1545-1410(2005)771[0001:aj]2.0.co;2

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Kujoth, C. C., Hiona, A., Pugh, T. D., Someya, S., Panzer, K., Wohlgemuth, S. E., Hofer, T., Seo, A. Y., Sullivan, R., Jobling, W. A., Morrow, J. D., Van Remmen, H., Sedivy, J. M., Yamasoba, T., Tanokura, M., Weindruch, R., Leeuwenburgh, C., & Prolla, T. A. (2005). Medicine: Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science*, *309*(5733), 481–484. https://doi.org/10.1126/science.1112125

Lachance, H., Cantrell, B., Friedman, S., Mckay, S., & Yee, R. (2015). Genome-wide epigenetic survey of Vermont bobcat (*Lynx rufus*) and fisher (*Martes pennanti*). In: 74th Annual Conference of the Northeast Association of Fish and Wildlife Agencies.

Le Clercq, L., Dalton, D. L., Kotzé, A., & Grobler, J. P. (2023a). Designing an EpiTYPER bisulfite sequencing assay for age estimation in *Acinonyx jubatus* based on human orthologues. *Protocols.Io*. https://doi.org/10.17504/protocols.io.j8nlk4yk1g5r/v1

Le Clercq, L., Dalton, D. L., Kotzé, A., & Grobler, J. P. (2023b). DNA extraction protocol for animal blood samples using the E.Z.N.A blood mini kit. *Protocols.Io*. https://doi.org/10.17504/protocols.io.ewov141xpvr2/v1

Le Clercq, L., Kotzé, A., Grobler, J. P., & Dalton, D. L. (2023). Biological clocks as age estimation markers in animals: A systematic review and meta-analysis. *Biological Reviews*, *98*(6), 1972–2011. https://doi.org/10.1111/brv.12992

Le Clercq, L. S. I., Dalton, D. L., & Kotze, A. (2018). Molecular age estimation based on promotor CpG methylation using methylation sensitive PCR. 9th Annual Research Symposium of the National Zoological Gardens. Pretoria: ScienceOpen https://doi.org/10.14293/S2199-1006.1.SOR-.PPROAEF.V1

Linck, E., Freeman, B. G., & Dumbacher, J. P. (2019). Speciation and gene flow across an elevational gradient in New Guinea kingfishers. *BioRxiv*. https://doi.org/10.1101/589044

Lumley, T. (2020). *leaps: Regression Subset Selection. R package*. https://cran.r-project.org/package=leaps

Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J., & Stuart, S. N. (2008). Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology*, *22*(6), 1424–1442. https://doi.org/10.1111/j.1523-1739.2008.01044.x

Marker, L. L., & Dickman, A. J. (2003). Morphology, physical condition, and growth of the cheetah (*Acinonyx jubatus jubatus*). *Journal of Mammalogy*, *84*(3), 840–850. https://doi.org/10.1644/BRB-036

Marker, L. L., Dickman, A. J., Jeo, R. M., Mills, M. G. L., & Macdonald, D. W. (2003). Demography of the Namibian cheetah, *Acinonyx jubatus jubatus*. *Biological Conservation*, *114*(3), 413–425. https://doi.org/10.1016/S0006-3207(03)00069-7

Márquez-Ruiz, A. B., González-Herrera, L., de Luna, J. D., & Valenzuela, A. (2020). DNA methylation levels and telomere length in human teeth: Usefulness for age estimation. *International Journal of Legal Medicine*, *134*(2), 451–459. https://doi.org/10.1007/s00414-019-02242-7

Mawlood, S. K., Dennany, L., Watson, N., & Pickard, B. S. (2016). The EpiTect methyl qPCR assay as novel age estimation method in forensic biology. *Forensic Science International*, *264*(April), 132–138. https://doi.org/10.1016/j.forsciint.2016.03.047

Mayne, B., Berry, O., & Jarman, S. (2021). Optimal sample size for calibrating DNA methylation age estimators. *Molecular Ecology Resources*, *21*(7), 2316–2323. https://doi.org/10.1111/1755-0998.13437

Merlin, C., & Liedvogel, M. (2019). The genetics and epigenetics of animal migration and orientation: Birds, butterflies and beyond. *The Journal of Experimental Biology*, *222*(Pt Suppl 1), jeb191890. https://doi.org/10.1242/jeb.191890

Morin, P. A., Nestler, A., Rubio-Cisneros, N. T., Robertson, K. M., & Mesnick, S. L. (2005). Interfamilial characterization of a region of the ZFX and ZFY genes facilitates sex determination in cetaceans and other mammals. *Molecular Ecology*, *14*(10), 3275–3286. https://doi.org/10.1111/j.1365-294X.2005.02651.x

Naji, S., Colard, T., Blondiaux, J., Bertrand, B., D'Incau, E., & Bocquet-Appel, J. P. (2016). Cementochronology, to cut or not to cut? *International Journal of Paleopathology*, *15*, 113–119. https://doi.org/10.1016/j.ijpp.2014.05.003

Nakamura, S., Yamazaki, J., Matsumoto, N., Inoue-Murayama, M., Qi, H., Yamanaka, M., Nakanishi, M., Yanagawa, Y., Sashika, M., Tsubota, T., Ito, H., & Shimozuru, M. (2023). Age estimation based on blood DNA methylation levels in brown bears. *Molecular Ecology Resources*, *23*(6), 1211–1225. https://doi.org/10.1111/1755-0998.13788

Nakanishi, N., Ichinose, F., Higa, G., & Izawa, M. (2009). Age determination of the iriomote cat by using cementum annuli. *Journal of Zoology*, *279*(4), 338–348. https://doi.org/10.1111/j.1469-7998.2009.00624.x

Nyakatura, K., & Bininda-Emonds, O. R. P. (2012). Updating the evolutionary history of Carnivora (Mammalia): A new species-level supertree complete with divergence time estimates. *BMC Biology*, *10*, 1–31. https://doi.org/10.1186/1741-7007-10-12

O'Brien, S. J., Roelke, M. E., Marker, L., Newman, A., Winkler, C. A., Meltzer, D., Colly, L., Evermann, J. F., & Wildt, D. E. (1985). Genetic basis for species vulnerability in the cheetah. *Science*, *227*, 1428–1434. https://doi.org/10.1126/science.2983425

Pan, C., Yi, S., Xiao, C., Huang, Y., Chen, X., & Huang, D. (2020). The evaluation of seven age-related CpGs for forensic purpose in blood from Chinese Han population. *Forensic Science International: Genetics*, *46*, 102251. https://doi.org/10.1016/j.fsigen.2020.102251

Poganik, J. R., Zhang, B., Baht, G. S., Kerepesi, C., Yim, S. H., Lu, A. T., Haghani, A., Gong, T., Hedman, A. M., Andolf, E., Pershagen, G., Almqvist, C., Clish, C. B., Horvath, S., White, J. P., Horvath, S., & Gladyshev, V. N. (2022). Biological age is increased by stress and restored upon recovery. *BioRxiv*. https://doi.org/10.1101/2022.05.04.490686

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

Qi, H., Kinoshita, K., Mori, T., Matsumoto, K., Matsui, Y., & Inoue-Murayama, M. (2021). Age estimation using methylation-sensitive high-resolution melting (MS-HRM) in both healthy felines and those with chronic kidney disease. *Scientific Reports*, *11*(1), 1–10. https://doi.org/10.1038/s41598-021-99424-4

R Core Team. (2020). *R: A language and environment for statisitical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Raj, K., Szladovits, B., Haghani, A., Zoller, J. A., Li, C. Z., Black, P., Maddox, D., Robeck, T. R., & Horvath, S. (2021). Epigenetic clock and methylation studies in cats. *GeroScience*, *43*(5), 2363–2378. https://doi.org/10.1007/S11357-021-00445-8/FIGURES/5

Robeck, T. R., Haghani, A., Fei, Z., Lindemann, D. M., Russell, J., Herrick, K. E. S., Montano, G., Steinman, K. J., Katsumata, E., Zoller, J. A., & Horvath, S. (2023). Multi-tissue DNA methylation aging clocks for sea lions, walruses and seals. *Communications Biology*, *6*(1), 359. https://doi.org/10.1038/s42003-023-04734-0

RStudio Team. (2021). RStudio: Integrated Development for R. *RStudio, PBC, Boston*, MA URL http://www.rstudio.com/.

Sareisian, L. (2014). *Global DNA methylation and persistent organic pollutants in ringed seals (Phoca hispida) from Svalbard and the Baltic Sea*. University of Oslo.

Shan, G. (2022). Monte Carlo cross-validation for a study with binary outcome and limited sample size. *BMC Medical Informatics and*

Decision Making, *22*(1), 1–15. https://doi.org/10.1186/s12911-022-02016-z

Shao, J. (1993). Linear model selection by cross-validation. *Journal of the American Statistical Association*, *88*(422), 486–494. https://doi.org/10.1080/01621459.1993.10476299

Sharma, V., Sharma, C. P., Sharma, V., Goyal, S. P., & Gupta, S. K. (2021). Fast and cost-effective age estimation in Bengal tiger and Asiatic lion: Applicability of cementum analysis method. *BioRxiv*. https://doi.org/10.1101/2021.09.27.461978

Spiers, H., Hannon, E., Wells, S., Williams, B., Fernandes, C., & Mill, J. (2016). Age-associated changes in DNA methylation across multiple tissues in an inbred mouse model. *Mechanisms of Ageing and Development*, *154*(21), 20–23. https://doi.org/10.1016/j.mad.2016.02.001

Stander, P. E. (1997). Field age determination of leopards by tooth wear. *African Journal of Ecology*, *35*(2), 156–161. https://doi.org/10.1111/j.1365-2028.1997.068-89068.x

Statham, A., & Csárdi, G. (2008). *RSeqMeth: Package for analysis of Sequenom EpiTYPER data*. Sydney. https://github.com/cran/RSeqMeth

Suchiman, H. E. D., Slieker, R. C., Kremer, D., Slagboom, P. E., Heijmans, B. T., & Tobi, E. W. (2015). Design, measurement and processing of region-specific DNA methylation assays: The mass spectrometry-based method EpiTYPER. *Frontiers in Genetics*, *6*, 287. https://doi.org/10.3389/fgene.2015.00287

Sugrue, V. J., Zoller, J. A., Narayan, P., Lu, A. T., Ortega-Recalde, O. J., Grant, M. J., Bawden, C. S., Rudiger, S. R., Haghani, A., Bond, D. M., Hore, R. R., Garratt, M., Sears, K. E., Wang, N., Yang, X. W., Snell, R. G., Hore, T. A., & Horvath, S. (2021). Castration delays epigenetic aging and feminizes dna methylation at androgen-regulated loci. *eLife*, *10*, e64932. https://doi.org/10.7554/eLife.64932

Tamazian, G., Simonov, S., Dobrynin, P., Makunin, A., Logachev, A., Komissarov, A., Shevchenko, A., Brukhin, V., Cherkasov, N., Svitin, A., Koepfli, K. P., Pontius, J., Driscoll, C. A., Blackistone, K., Barr, C., Goldman, D., Antunes, A., Quilez, J., Lorente-Galdos, B., … O'Brien, S. J. (2014). Annotated features of domestic cat – Felis catus genome. *GigaScience*, *3*(1), 1–3. https://doi.org/10.1186/2047-217X-3-13/TABLES/1

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

Tewari, D., Miller, R., Livengood, J., Wang, L., Killian, M. L., Bustamante, F., Kessler, C., Thirumalapura, N., Terio, K., Torchetti, M., Lantz, K., & Rosenberg, J. (2023). SARS-CoV-2 infection dynamics in the Pittsburgh zoo wild felids with two viral variants (Delta and alpha) during the 2021–2022 pandemic in the United States. *Animals*, *13*(19), 3094. https://doi.org/10.3390/ani13193094

Thompson, M. J., VonHoldt, B., Horvath, S., & Pellegrini, M. (2017). An epigenetic aging clock for dogs and wolves. *Aging*, *9*(3), 1055–1068. https://doi.org/10.18632/aging.101211

Thompson, R. F., & Greally, J. M. (2020). *MassArray: Analytical Tools for MassArray Data*. https://doi.org/10.18129/B9.bioc.MassArray

Viñuela, A., Brown, A. A., Buil, A., Tsai, P. C., Davies, M. N., Bell, J. T., Dermitzakis, E. T., Spector, T. D., & Small, K. S. (2018). Age-dependent changes in mean and variance of gene expression across tissues in a twin cohort. *Human Molecular Genetics*, *27*(4), 732–741. https://doi.org/10.1093/hmg/ddx424

von Schreber, J. C. D. (1775). Der Gepard. In *Die Säugthiere in Abbildungen nach der Natur mit Beschreibungen*. Walther'sche Kunst- und Buchhandlung. https://doi.org/10.11588/diglit.3543

Weidner, C. I., Lin, Q., Koch, C. M., Eisele, L., Beier, F., Ziegler, P., Bauerschlag, D. O., Jöckel, K. H., Erbel, R., Mühleisen, T. W., Zenke, M., Brümmendorf, T. H., & Wagner, W. (2014). Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biology*, *15*(2), 1–12. https://doi.org/10.1186/gb-2014-15-2-r24

White, P. A., & Belant, J. L. (2016). Individual variation in dental characteristics for estimating age of African lions. *Wildlife Biology*, *22*(3), 71–77. https://doi.org/10.2981/wlb.00180

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

Yamaguchi, N., Driscoll, C. A., Werdelin, L., Abramov, A. V., Csorba, G., Cuisin, J., Fernholm, B., Hiermeier, M., Hills, D., Hunter, L., Itakura, H., Johansson, U. S., Kascheev, V., Krohmann, K., Martin, T., Nowak-Kemp, M., Pavlinov, I. Y., Renoud, F., Tomsett, L., … Macdonald, D. W. (2013). Locating specimens of Extinct Tiger (*Panthera tigris*) subspecies: Javan Tiger (P. T. sondaica), Balinese Tiger (P. T. balica), and Caspian Tiger (P. T. virgata), including previously unpublished specimens. *Mammal Study*, *38*(3), 187–198. https://doi.org/10.3106/041.038.0307

Zainol, Z. A., Amin, L., Akpoviri, F., & Ramli, R. (2011). Biopiracy and states' sovereignty over their biological resources. *African Journal of Biotechnology*, *10*(58), 12395–12408.

Zampieri, M., Ciccarone, F., Calabrese, R., Franceschi, C., Bürkle, A., & Caiafa, P. (2015). Reconfiguration of DNA methylation in aging. *Mechanisms of Ageing and Development*, *151*, 60–70. https://doi.org/10.1016/j.mad.2015.02.002

SUPPORTING INFORMATION

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

How to cite this article: Le Clercq, L.-S., Kotzé, A., Grobler, J. P., & Dalton, D. L. (2024). Methylation-based markers for the estimation of age in African cheetah, *Acinonyx jubatus*. *Molecular Ecology Resources*, *00*, e13940. https://doi.org/10.1111/1755-0998.13940