



OPEN

DNA methylation profile in beef cattle is influenced by additive genetics and age

André Mauric F. Ribeiro¹, Leticia P. Sanglard¹, Hiruni R. Wijesena¹, Daniel C. Ciobanu¹, Steve Horvath^{2,3}✉ & Matthew L. Spangler¹✉

DNA methylation (DNAm) has been considered a promising indicator of biological age in mammals and could be useful to increase the accuracy of phenotypic prediction in livestock. The objectives of this study were to estimate the heritability and age effects of site-specific DNAm (DNAm level) and cumulative DNAm across all sites (DNAm load) in beef cattle. Blood samples were collected from cows ranging from 217 to 3,192 days (0.6 to 8.7 years) of age ($n = 136$). All animals were genotyped, and DNAm was obtained using the Infinium array HorvathMammalMethylChip40. Genetic parameters for DNAm were obtained from an animal model based on the genomic relationship matrix, including the fixed effects of age and breed composition. Heritability estimates of DNAm levels ranged from 0.18 to 0.72, with a similar average across all regions and chromosomes. Heritability estimate of DNAm load was 0.45. The average age effect on DNAm level varied among genomic regions. The DNAm level across the genome increased with age in the promoter and 5' UTR and decreased in the exonic, intronic, 3' UTR, and intergenic regions. In addition, DNAm level increased with age in regions enriched in CpG and decreased in regions deficient in CpG. Results suggest DNAm profiles are influenced by both genetics and the environmental effect of age in beef cattle.

DNA methylation (DNAm) is an epigenetic mark characterized by a modification of the chromatin structure in mammals¹. This modification exists throughout the genome with remarkably less abundance in cytosine-guanine (CpG) islands². In mammalian genomes, CpG islands are greatly under-represented due to spontaneous evolutionary deamination of 5-methylcytosine to thymine and are usually located within or near promoters and first exons of housekeeping genes³. In general, it has been recognized that DNAm in the promoter region of genes indicates transcriptional suppression while hypomethylation is associated with transcriptional activation and increased gene expression².

DNA methylation profiles are dynamic and susceptible to factors such as age^{4–6}, diet^{7,8}, sex⁹, and exposure to diseases¹⁰. It has been established that the DNAm level of mammalian genomes increases with age, although this relationship seems to differ across the regions of the genome. Marttila et al.¹¹ observed that age-associated DNA hypermethylation seems to originate from programmed developmental changes while DNA hypomethylation could result from environmental influences¹¹. Furthermore, DNAm profiles in several species have been demonstrated to predict individuals' biological age, known as epigenetic clocks^{12–23}. Recently, Caulton et al.¹³ reported a prediction accuracy of 0.95 for epigenetic clocks based on DNAm in a dairy herd. In general, DNAm has been considered a promising indicator of biological age in mammals.

In addition to influences of environmental factors, genetic control of DNAm levels has been reported^{9,24,25}. For example, DNAm of human peripheral blood lymphocytes at sites with deficient CpG density tended to have higher heritability than regions with high-CpG density²⁵. Integrating genomic data with DNAm information could be a potential strategy to increase the accuracy of phenotypic prediction in livestock. However, the expected improvement depends on the heritability of DNAm, or in other words, the additional information beyond additive genetics to be gained with DNAm information. Therefore, the objectives of this study were to estimate the heritability and age effects of site-specific and cumulative DNAm across the cattle genome.

¹Department of Animal Science, University of Nebraska, Lincoln, NE 68583, USA. ²Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA. ³Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, CA 90095, USA. ✉email: SHorvath@mednet.ucla.edu; mspangler2@unl.edu

Parameters	Promoter	5' UTR	Exon	Intron	3' UTR	Intergenic	All Regions	
Number of sites	2,445	1,456	11,500	7,852	879	10,192	34,324	
M-value	Mean \pm SD	-1.28 \pm 1.97	-1.31 \pm 2.36	1.24 \pm 2.20	1.22 \pm 2.08	1.63 \pm 1.93	0.83 \pm 2.01	0.84 \pm 2.24
	Maximum	5.67	5.27	5.57	5.49	4.90	5.08	5.67
	Minimum	-5.25	-5.27	-5.51	-5.18	-4.83	-5.38	-5.51
σ_p^2	Mean \pm SD	0.05 \pm 0.16	0.04 \pm 0.05	0.04 \pm 0.11	0.05 \pm 0.15	0.04 \pm 0.05	0.06 \pm 0.23	0.05 \pm 0.17
	Maximum	6.24	0.75	4.93	5.18	2.73	6.51	6.51
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00
σ_a^2	Mean \pm SD	0.02 \pm 0.09	0.01 \pm 0.02	0.02 \pm 0.06	0.02 \pm 0.08	0.02 \pm 0.07	0.03 \pm 0.14	0.02 \pm 0.10
	Maximum	3.33	0.33	2.97	2.85	1.62	4.08	4.08
	Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00
h^2	Mean \pm SD	0.36 \pm 0.07	0.35 \pm 0.07	0.34 \pm 0.07	0.34 \pm 0.07	0.34 \pm 0.07	0.36 \pm 0.08	0.35 \pm 0.07
	Maximum	0.67	0.63	0.69	0.71	0.72	0.73	0.73
	Minimum	0.19	0.20	0.18	0.20	0.20	0.19	0.18
Age effect	Mean \pm SD	9.2E-05 \pm 1.6E-04	4.8E-05 \pm 1.2E-04	-3.8E-05 \pm 1.3E-04	-5.4E-05 \pm 1.4E-04	-5.7E-05 \pm 1.4E-04	-6.8E-05 \pm 1.6E-04	-3.8E-05 \pm 1.6E-04
	Maximum	9.9E-04	1.0E-03	1.1E-03	7.0E-04	7.7E-04	8.9E-04	1.1E-03
	Minimum	-7.6E-04	-5.2E-04	-7.1E-04	-5.6E-04	-4.8E-04	-8.6E-04	-8.6E-04

Table 1. Summary of estimated parameters for DNA methylation (DNAm) levels (M-values) for all DNAm sites in different functional regions of the bovine genome. σ_p^2 , phenotypic variance (i.e., additive genetic variance + residual variance); σ_a^2 , additive genetic variance; h^2 , heritability.

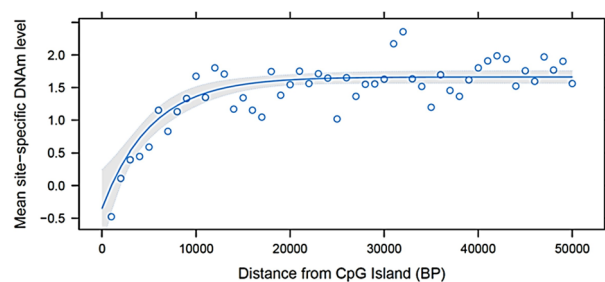


Figure 1. Exponential regression of mean DNAm level on distance of the methylated cytosines from CpG islands into bins of 1000 bp.

Results

Description of DNAm level and load. DNA methylation levels, defined as logit transformed standardized intensity (M-value), varied with functional regions of the genome (Table 1). On average, promoter (M-value = -1.28) and 5' UTR (M-value = -1.31) were hypomethylated while exonic, intronic, 3' UTR, and intergenic regions were hypermethylated, with M-values ranging from 0.83 to 1.24. CpG islands were hypomethylated (M-value = -0.68), while regions deficient in CpG were hypermethylated (M-value = 1.18). The average DNAm load, defined as the sum of M-values across all DNAm sites, was $12,336 \pm 3,584$. The animals with the highest and lowest DNAm load had 20,808 and 6,399 sites, respectively. Additionally, as the distance of methylated cytosines from the CpG islands increased, there was an exponential increase in the mean site-specific DNAm level, reaching a plateau when the distance was approximately 1 kb (Fig. 1).

Genetic parameters of DNAm level and load. Estimates of heritability (h^2) of DNAm sites ranged from 0.18 to 0.72 (Fig. 2a). A similar average estimate was obtained across the functional regions, ranging from 0.34 ± 0.07 to 0.36 ± 0.08 (Table 1). The distribution of h^2 estimates was uniform across the genome (Fig. 2b). The distribution of genetic and residual variance estimates across the genome is depicted in Supplemental Fig. S1. Chromosomes 20 and 21 showed the smallest and largest average heritabilities (0.35 ± 0.08 and 0.37 ± 0.08), respectively. The h^2 estimate of DNAm load was moderate to high (0.45 ± 0.10). The average h^2 estimates of DNAm were similar in regions enriched or deficient in CpG islands (0.36 ± 0.07 and 0.36 ± 0.08 , respectively).

Age effect on DNAm level and load. Overall, the average age effect on DNAm level across all sites was negative (-0.013 ± 0.06) (Fig. 3); however, the effect of age varied by functional region. The DNAm level increased with age in the promoter and 5' UTR and decreased in the exonic, intronic, 3' UTR, and intergenic regions (Table 1). In addition, DNAm level increased with age in regions enriched in CpG islands and decreased in regions deficient in CpG islands (Results not shown). DNAm sites with significant age effects were observed when h^2 estimates were intermediate rather than extreme values (Supplemental Fig. S2).

In addition, DNAm level increased with age in regions enriched in CpG islands and decreased in regions deficient in CpG islands (Results not shown). DNAm sites with significant age effects were observed when h^2 estimates were intermediate rather than extreme values (Supplemental Fig. S2).

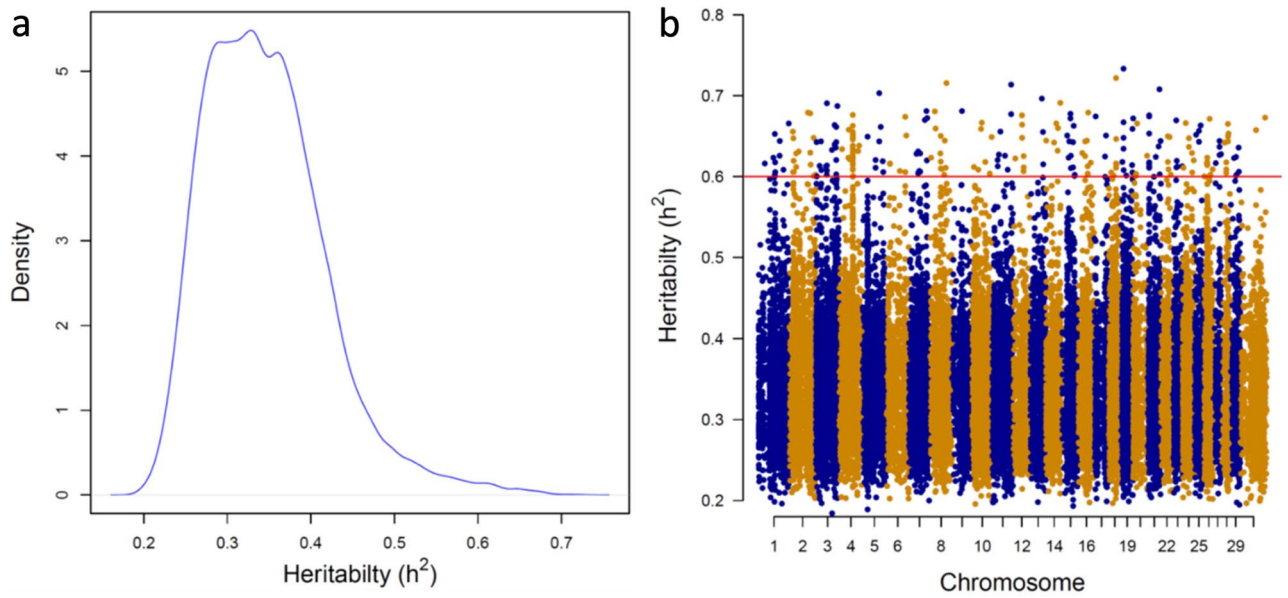


Figure 2. Density of heritability estimates of DNAm levels for each DNAm site (a) and heritability estimates of DNAm levels across all chromosomes (b).

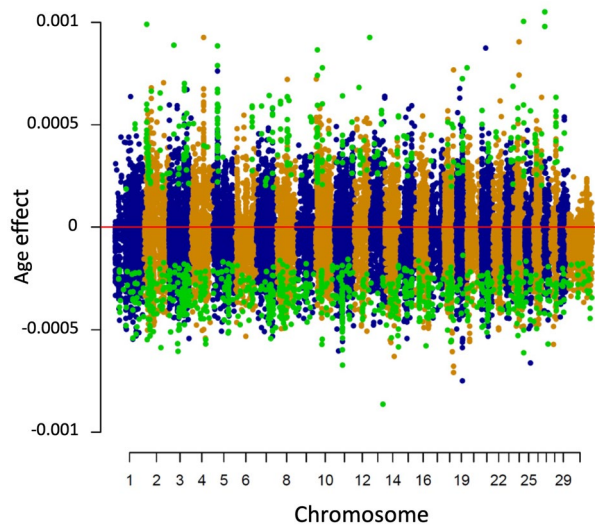


Figure 3. Distribution of age effects on DNAm levels across all chromosomes. DNAm sites with significant age effect ($P \leq 10^{-25}$) are highlighted in green.

There was a non-linear relationship between DNAm load and animal age (years) ($R^2 = 0.61$, Fig. 4), with younger animals having a greater DNAm load than older animals. As the distance from CpG islands increased, there was an exponential decrease in age effect on DNAm, reaching a plateau when the distance was approximately 1 kb (Fig. 5).

Predicted values for age were generated from five Bayesian regression models: Bayesian Ridge Regression (BRR), BayesA, BayesB, Bayes π , and Bayesian LASSO (BLASSO). Results can be seen in Table 2. In general, the mean correlation (r) between the predicted and true ages was high, ranging from 0.97 to 0.99, and the mean slope of the regression of predicted on true age ranged from 0.88 to 0.92, close to the expected slope of 1. A BayesA model performed best ($r = 0.99$, $MSE = 0.11$ and slope = 0.93), while BLASSO was the least accurate, albeit similar, and yielded predictions that were under-dispersed to a greater degree ($r = 0.97$, $MSE = 0.26$ and slope = 0.88).

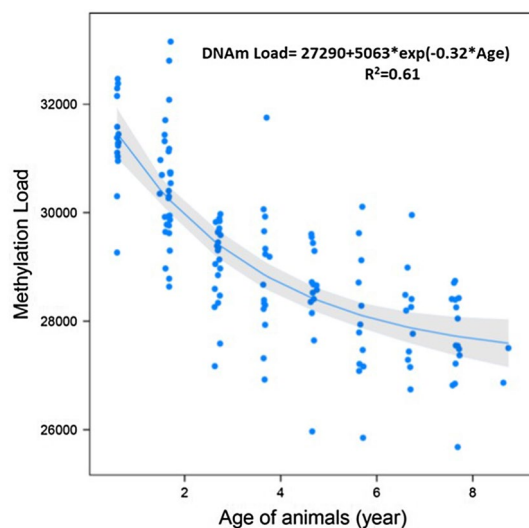


Figure 4. Exponential decay regression of Methylation load (DNAm load) on the age of animals (years).

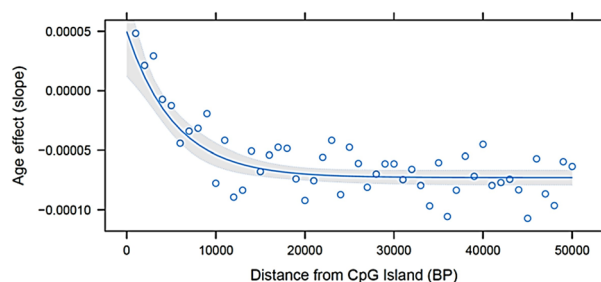


Figure 5. Exponential regression of age effect on distance of the methylated cytosines from CpG islands into bins of 1000 bp.

Parameter	BRR	BayesA	BayesB	BayesCπ	BLASSO
Mean correlation	0.98 (0.003)	0.99 (0.003)	0.98 (0.003)	0.98 (0.003)	0.97 (0.006)
Mean of MSE	0.15 (0.036)	0.11 (0.031)	0.14 (0.033)	0.15 (0.035)	0.26 (0.071)
Mean of slope	0.92 (0.023)	0.93 (0.021)	0.92 (0.022)	0.92 (0.023)	0.88 (0.031)

Table 2. Mean correlation, MSE, and slope of regression of predicted age based on DNAm status on true age. MSE Mean Square Error; BRR Bayesian Ridge Regression; BLASSO Bayesian LASSO.

Discussion

Changes in DNA methylation profiles can alter gene expression leading to phenotypic variation. DNA methylation is subjected to changes due to age and potentially by underlying genetic variation. Furthermore, methylation levels can vary across the genome depending on the functional role of the region and the location and enrichment of CpG islands. This study estimated heritabilities and the effect of age on site-specific and cumulative DNAm in a beef cattle population.

DNAm level was lower in the promoter and 5' UTR than 3' UTR, intergenic, exonic, and intronic regions. An exonic region was defined as a region of the genome that forms an mRNA molecule and is translated into a protein. Correspondingly, a lower DNAm level was observed in CpG islands than non-CpG islands. An increase in DNAm levels was observed as their distance from the CpG islands increased. This was expected given many promoters in the mammalian genome are characterized by unmethylated CpG islands, and promoter hypomethylation is associated with transcriptional activation¹. In addition, hypomethylation in the 5' UTR could be because CpG site-associated genes usually extend from the 5' UTR end into the gene body¹. Hypermethylation in intronic and exonic regions was also observed by Chodavarapu et al.²⁶ and was associated with an enrichment of nucleosomes, which are the basic structural unit of DNA packaging in eukaryotes and correspond to segments of DNA coiled around a core of histone proteins. However, different from this study, the authors observed a

greater DNAm level in exons than in introns²⁶. The hypermethylation in coding exons and 3'UTR is consistent with the presence of DNAm in transcribed regions.

Heritability estimates of DNAm levels were similar and moderate among functional regions, CpG islands vs. non-CpG islands, and chromosomes. These results indicate potentially no relationship between h^2 of DNAm levels and CpG density. These results contrast with McRae et al. and Gordon et al.^{24,25}, who reported higher h^2 of DNAm levels in low-CpG density sites than regions with high-CpG density in the genome of human peripheral blood lymphocytes. Estimates of heritability ranged from moderate to high across DNAm sites and were high for DNAm load, in agreement with previous reports. Boks et al.⁹ estimated h^2 ranging from 0.03 to 0.90 in specific CpG islands in a twin study. Another twin study reported h^2 estimates ranging from 0.48 to 0.94 for the top 5% of sites investigated in cord blood mononuclear cells²⁴. These results indicate that a substantial part of the variation in DNAm is due to genetic variance. However, there is still extra information beyond additive genetics to be gained with DNAm patterns that could help improve phenotypic prediction, especially for lowly heritable traits.

The relationship between DNAm level and aging has been well established. In the current study, the effect of age was dependent on the genomic region. DNAm level increased with age in the promoter and 5' UTR but decreased in the other regions. Similarly, DNAm level increased with age in CpG regions but decreased in non-CpG regions. These results are in agreement with Marttila et al.¹¹, who observed enrichment of aging-associated hypermethylation at CpG islands and hypomethylation at non-CpG islands. The majority of CpG islands are not initially methylated, and the change observed during aging is hypermethylation. The opposite is true for regions with few CpG islands that initially are heavily methylated, and the non-CpG are associated with hypomethylation. Marttila et al.¹¹ observed a slight majority (54%) of the identified sites were hypomethylated in the elderly. Similarly, Johansson et al.²⁷ also reported an excess of hypomethylation over hypermethylation with aging. On the other hand, Korkmaz and Kerr²⁸ observed hypomethylation of 67% of the CpG islands analyzed in 5 m vs. 16 m old dermal fibroblasts in dairy cattle. Age-associated DNA hypermethylation seems to be associated with genes belonging to a common pathway of sequence-specific DNA binding and transcription factor binding regulating genes associated with development and morphogenesis; and metabolic processes, gene expression and nucleotide metabolism, while hypomethylation was not associated with specific pathways¹¹.

The current study also showed that age could be accurately ($r \geq 0.97$) predicted based on DNAm status. In general, all models slightly under-predicted age. The slightly better performance of BayesA indicated that fitting all DNAm in the model simultaneously improves age prediction, and there is no advantage in performing variable selection²⁹. Additionally, BayesA assumes a heavy-tailed distribution of effects that may be more appropriate for modeling the effects of DNAm on age²⁹. Previous studies have shown DNAm can be used as features for constructing epigenetic clocks in several species, including cattle. Caulton et al.¹³ reported a high prediction accuracy (> 0.94) for different age clocks based on DNAm of CpG islands in deer, goat, sheep, and dairy cattle. More recently, based on DNAm of liver and blood samples of horses, Horvath et al.¹⁶ reported accuracy greater than 0.95 for epigenetic clocks. Kordowitzki et al.¹⁸ developed epigenetic clocks of reproductive aging in oocytes and blood from cattle with an accuracy greater than 0.86. Hayes et al.¹⁷ achieved a prediction accuracy for birthdate of 0.71 in beef cattle from a diverse genetic background. These results are encouraging for using epigenetic clocks to estimate age in cattle reared in extensive production systems, which usually have unknown or inaccurate birthdate records. However, accurately predicting differences in birthdates among cohorts requires substantial reductions in the average error reported by Hayes et al. that could perhaps be achieved with much larger training sets. Overcoming the challenges of industry birthdate recording with epigenetic clocks could contribute to herd management and estimation of genomic breeding values for economically important traits that depend on accurate age records, such as growth rate and calving interval¹⁷. Given epigenetic clocks have proven to be accurate predictors of biological age, epigenetic indicators could provide a mechanism to infer age-related pathologies and identify genetic, environmental, or lifestyle factors that accelerate or slow aging in humans and livestock³⁰.

Material and methods

Population resource and data collection. All animal procedures implemented in this study were approved by the University of Nebraska-Lincoln Animal Care and Use Committee. All methods were performed in accordance with relevant guidelines and regulations. Methods are reported in the manuscript following the recommendations in the ARRIVE guidelines.

The blood obtained from 136 cows was centrifuged at $2500 \times g$ for 10 min at room temperature. The buffy coat was collected and stored at -80°C . All samples were collected from October 2 to November 27 of 2018 at the Eastern Nebraska Research and Extension Center at the University of Nebraska-Lincoln. Cows were chosen to provide variation in age and ranged from 217 to 3,192 days (0.6 to 8.7 years) of age at the time of sample collection. Animals were from an admixed population comprised of purebred Angus and composites of differing proportions of Angus, Simmental, and Red Angus.

All animals were genotyped with the medium-density Illumina BovineSNP50 (~ 50 K SNPs) BeadChip (Illumina, San Diego, CA, USA). Genotype filtering included removing non-autosomal SNPs, SNPs with minor allele frequency < 0.02 , and SNPs with Hardy-Weinberg equilibrium P -value $> 10^{-5}$.

DNA methylation. DNA was extracted from buffy coats using the DNeasy Blood and Tissue Kit (Qiagen, Cat No. 69506) and, subsequently, bisulfite-converted using the EZ DNA Methylation Kit (ZymoResearch, Irvine, CA, USA). DNAm data was obtained from bisulfite-treated samples using the mammalian array (HorvathMammalMethylChip40)³¹. DNAm level for each site was calculated as methylation β -value (β -value = intensity of the methylated allele/intensity of the unmethylated allele + intensity of the methylated allele + 100). The addition of 100 was used to stabilize β -values when both intensities of the methylated and unmethylated alleles were small³². The SeSAMe pipeline³³ was used to generate normalized β -values and for quality control. The

β -value has severe heteroscedasticity outside the intermediate methylation range; thus, a logit transformation of the β -values (M -values) was used to approximate homoscedasticity. M -values of 0 correspond to 50% of methylation, and positive and negative values correspond to greater and lesser than 50% methylation level, respectively. M -values were used to quantify DNAm level by region (i.e., promoter, 5' and 3' UTR, exonic, intronic, and intergenic) and location related to CpG islands (i.e., within or outside). DNA methylation load was calculated as the sum of all DNAm levels (M -value).

DNAm status of each site was determined by the distribution of the methylation β -values. For example, β -values below, within, and above 2 standard deviations were classified as unmethylated (0), intermediately methylated (1), and methylated (2), respectively. DNAm status was used for prediction purposes.

Statistical analyses. *Genetic parameters for DNAm level and DNAm load.* Genetic parameters (i.e., additive genetic and residual variances) for DNAm level (M -values) and DNAm load were estimated using the following animal model fitted in a Bayesian genomic best linear unbiased prediction (GBLUP) framework.

$$y = Xb + Zu + e$$

where y corresponds to the vector of phenotypes (DNAm level or DNAm load); X corresponds to the design matrix linking the fixed effects to the phenotypes; b corresponds to the vector of fixed effects including the intercept, the linear and quadratic (DNAm only) covariates of age, and covariates for proportion of each breed; Z corresponds to the incidence matrix linking the random animal effect to the phenotypes; u corresponds to the vector of random animal effects, where $u \sim N(0, G\sigma_u^2)$, where G corresponds to the genomic relationship matrix constructed following the first method of VanRaden³⁴ and σ_u^2 corresponds to the additive genetic variance; and e corresponds to the vector of random residual effects associated with the phenotype, where $e \sim N(0, I\sigma_e^2)$, where I corresponds to the identity matrix and σ_e^2 corresponds to the residual variance. The h^2 was obtained as the ratio of additive genetic variance divided by phenotypic variance (additive genetic variance + residual variance).

Gibbs sampling was used to sample a posterior parameter distribution with a chain length of 20,000 iterations, burn-in of 2,000 samples, and a thinning interval of 100. Analyses were performed using the *BGLR* package³⁵ in R software.

Age effect on DNAm load and age prediction. The effect of age on DNAm load was estimated by fitting an exponential regression of DNAm load on the age of animals (years). DNAm status was included as a variable to predict the age of animals using five Bayesian regression models: BRR²⁹, BayesA²⁹, BayesB²⁹, BayesC π ³⁶, and Bayesian LASSO (BLASSO)³⁷, as follow:

$$y = Xb + \sum_{i=1}^k m_i \alpha_i \delta_i + e$$

where X and e have been previously described; y corresponds to the vector of ages in years; b corresponds to the vector of fixed effects, including the intercept and the linear covariates for each breed; m_i is the vector of DNAm status for site i (coded as 0, 1, and 2); α_i is the effect of DNAm status for site i for each of k sites; δ_i is an indicator whether DNAm status site i was included ($\delta_i = 1$) or excluded ($\delta_i = 0$) from the model for a given iteration of gibbs sampling (BayesRR and BayesA, $\delta_i = 1$). In BayesRR and BayesC π , the effect of DNAm status is assumed to follow a normal distribution. In BayesA and BayesB, the effect of DNAm status is assumed to follow a t -distribution with site-specific variances. In Bayes Lasso, the effect of DNAm status is assumed to follow a double-exponential distribution.

Gibbs sampling was used to sample a posterior parameter distribution with a chain length of 20,000 iterations, burn-in of 2,000 samples, and a thinning interval of 100. Bootstrapping ($n = 400$) was used to evaluate the performance of the models with 102 and 34 individuals as training and validation populations, respectively. The performance of the models was evaluated based on the correlation between true and predicted age, mean square error, and slope of the regression of predicted age on true age. Analyses were performed in *BGLR* package³⁵ in R software.

Conclusion

Estimates of heritability suggest generational inheritance of DNAm load in the genome of cattle. The heritability estimates of each DNA site were moderate and well distributed across different functional regions of the genome, locations related to CpG islands, and chromosomes. The variation of heritability across the genome indicates the possibility of using DNAm as an additional source of information to improve phenotypic prediction. However, the effect of age on DNAm differed across the regions of the genome. The promoter and 5' UTR were initially hypomethylated, and DNAm levels increased with age, while there was initial hypermethylation in the exonic, intronic, 3' UTR, and intergenic regions followed by a decrease in DNAm levels with age. The high prediction accuracy of true age based on DNAm indicates a potential use of DNAm to predict chronological age in beef cattle. Overall, these findings indicate DNAm profiles are determined by both genetics and environmental influences such as age in beef cattle.

Data availability

The methylation data will be released to Gene Expression Omnibus as part of the data release from the Mammalian Methylation Array Consortium. The mammalian methylation array (HorvathMammalMethylChip40)

is registered at the NCBI Gene Expression Omnibus as platform GPL28271. The mammalian methylation array is distributed by the non profit Epigenetic Clock Development Foundation <https://clockfoundation.org/>.

Received: 27 May 2022; Accepted: 8 July 2022

Published online: 14 July 2022

References

1. Suzuki, M. M. & Bird, A. DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–476 (2008).
2. Halušková, J., Holečková, B. & Staničová, J. DNA methylation studies in cattle. *J. Appl. Genet.* **62**, 121–136 (2021).
3. Jammes, H., Junien, C. & Chavatte-Palmer, P. Epigenetic control of development and expression of quantitative traits. *Reprod. Fertil. Dev.* **23**, 64–74 (2011).
4. Gensous, N. *et al.* Age-related DNA methylation changes: potential impact on skeletal muscle aging in humans. *Front. Physiol.* **10**, 996 (2019).
5. Kane, A. E. & Sinclair, D. A. Epigenetic changes during aging and their reprogramming potential. *Crit. Rev. Biochem. Mol. Biol.* **54**, 61–83 (2019).
6. Lee, J. H., Kim, E. W., Croteau, D. L. & Bohr, V. A. Heterochromatin: an epigenetic point of view in aging. *Exp. Mol. Med.* **52**, 1466–1474 (2020).
7. Zhang, N. Role of methionine on epigenetic modification of DNA methylation and gene expression in animals. *Anim. Nutr.* **4**, 11–16 (2018).
8. Dabek, A., Wojtala, M., Pirola, L. & Balcerczyk, A. Metabolomics by Ketone Bodies: implications of the pathological states. *Nutrients* **12**, 788 (2020).
9. Boks, M. P. *et al.* The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PLoS ONE* **4**, 21–23 (2009).
10. Wang, M. & Ibeagha-Awemu, E. M. Impacts of epigenetic processes on the health and productivity of livestock. *Front. Genet.* **11**, 1–31 (2021).
11. Marttila, S. *et al.* Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. *BMC Genomics* **16**, 1–17 (2015).
12. Jung, S. E., Shin, K. J. & Lee, H. Y. DNA methylation-based age prediction from various tissues and body fluids. *BMB Rep.* **50**, 546–553 (2017).
13. Caulton, A. *et al.* Development of epigenetic clocks for key ruminant species. *Genes (Basel)*. **13**, 96 (2022).
14. Larison, B. *et al.* Epigenetic models developed for plains zebras predict age in domestic horses and endangered equids. *Commun. Biol.* **4**, 1412 (2021).
15. Horvath, S. *et al.* DNA methylation age analysis of rapamycin in common marmosets. *GeroScience* **43**, 2413–2425 (2021).
16. Horvath, S. *et al.* DNA methylation aging and transcriptomic studies in horses. *Nat. Commun.* **13**, 1–13 (2022).
17. Hayes, B. J. *et al.* An Epigenetic aging clock for cattle using portable sequencing technology. *Front. Genet.* **12**, 1–8 (2021).
18. Kordowitzki, P. *et al.* Epigenetic clock and methylation study of oocytes from a bovine model of reproductive aging. *Aging Cell* **20**, 1–11 (2021).
19. Wilkinson, G. S. *et al.* DNA methylation predicts age and provides insight into exceptional longevity of bats. *Nat. Commun.* **12**, 1615 (2021).
20. Prado, N. A. *et al.* Epigenetic clock and methylation studies in elephants. *Aging Cell* **20**, 1–11 (2021).
21. Sugrue, V. J. *et al.* Castration delays epigenetic aging and feminizes dna methylation at androgen-regulated loci. *Elife* **10**, 1–20 (2021).
22. Schachtschneider, K. M. *et al.* Epigenetic clock and DNA methylation analysis of porcine models of aging and obesity. *GeroScience* **43**, 2467–2483 (2021).
23. Horvath, S. *et al.* DNA methylation clocks tick in naked mole rats but queens age more slowly than nonbreeders. *Nat. Aging* **2**, 46–59 (2022).
24. Gordon, L. *et al.* Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence. *Genome Res.* **22**, 1395–1406 (2012).
25. McRae, A. F. *et al.* Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol.* **15**, R73 (2014).
26. Chodavarapu, R. K. *et al.* Relationship between nucleosome positioning and DNA methylation. *Nature* **466**, 388–392 (2010).
27. Johansson, Å., Enroth, S. & Gyllenstein, U. Continuous aging of the human DNA methylome throughout the human lifespan. *PLoS ONE* **8**, e67378 (2013).
28. Korkmaz, F. T. & Kerr, D. E. Genome-wide methylation analysis reveals differentially methylated loci that are associated with an age-dependent increase in bovine fibroblast response to LPS. *BMC Genomics* **18**, 1–18 (2017).
29. Meuwissen, T. H. E., Hayes, B. J. & Goddard, M. E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829 (2001).
30. Horvath, S. *et al.* Decreased epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring. *Aging (Albany NY)*. **7**, 1159–1170 (2015).
31. Arneson, A. *et al.* A mammalian methylation array for profiling methylation levels at conserved sequences. *Nat. Commun.* **13**, 1–13 (2022).
32. Weinhold, L., Wahl, S., Pechlivanis, S., Hoffmann, P. & Schmid, M. A statistical model for the analysis of beta values in DNA methylation studies. *BMC Bioinform.* **17**, 1–11 (2016).
33. Zhou, W., Triche, T. J., Laird, P. W. & Shen, H. SeSAME: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res.* **46**, 1–15 (2018).
34. VanRaden, P. M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).
35. Pérez, P. & De Los Campos, G. Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **198**, 483–495 (2014).
36. Habier, D., Fernando, R. L., Kizilkaya, K. & Garrick, D. J. Extension of the bayesian alphabet for genomic selection. *BMC Bioinform.* **12**, 186 (2011).
37. Gianola, D. & Fernando, R. L. A multiple-trait Bayesian lasso for genome-enabled analysis and prediction of complex traits. *Genetics* **214**, 305–331 (2020).

Acknowledgements

The authors would like to thank the staff at the University of Nebraska-Lincoln (UNL) Eastern Research, Extension, and Education Center's beef cattle teaching unit their assistance in collecting tissue samples and Bonnie Long for her assistance in collecting and curating the samples. The authors also acknowledge the UNL Holland Computing Center for the use of their computational resources.

Author contributions

A.F.R.: performed data analyses and interpreted results; L.P.S.: helped in the interpretation of the results and drafted the manuscript; H.W.: performed DNA extraction; D.C.C.: performed DNA extraction, assisted with planning of research, and helped interpret results; S.H.: generated the methylation data; M.L.S.: designed and planned the study, interpreted results, and drafted the manuscript. All authors contributed to and approved the final manuscript.

Competing interests

SH is a founder of the non-profit Epigenetic Clock Development Foundation which plans to license several patents from his employer UC Regents. The other authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-16350-9>.

Correspondence and requests for materials should be addressed to S.H. or M.L.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022