bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Influence of body weight at young adulthood on the epigenetic clock and lifespan in the BXD murine family

Jose Vladimir Sandoval-Sierra[1], Alexandra H. B. Helbing[1], Evan G. Williams[3], David G. Ashbrook[2], Suheeta Roy[2], Robert W. Williams[2], Khyobeni Mozhui[1,2*]

[1]Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, Tennessee, USA

[2]Department of Genetics, Genomics and Informatics, University of Tennessee Health Science Center, Memphis, Tennessee, USA

[3] Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Zurich CH-8093, Switzerland

[*]Corresponding author: Kmozhui@uthsc.edu

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Background. The DNA methylation landscape is shaped by genetic and environmental factors and is modulate by aging. Here we evaluate the "aging methylome" in 12 recombinant inbred mouse strains from the BXD family that have more than two-fold variation in longevity. We examine relations among body weight, diet, lifespan, and DNA methylation-based rate of biological aging. We used affinity capture with the methyl-CpG binding domain (MBD) protein, followed by deep sequencing (MBD-seq), to assay DNA methylation in 70 mostly female liver samples, ranging in age from 6 to 25 months from mice maintained either on low fat chow or high fat diet (HFD).

Results. Genetic background among strains is a major source of variation in genome-wide DNA methylation patterns. Surprisingly, body weight at young adulthood had a stronger association with the methylome than age itself. Nonetheless, age also had a strong effect on methylation at well-defined CpG regions largely located within genes. We used subsets of age-informative CpGs to build versions of the epigenetic clock and as expected, these were strongly correlated with chronological age. Both high initial body weight and the HFD were associated with accelerated epigenetic aging. A DNA methylation clock model built using CpGs associated with body weight correlated with longevity of strains rather than chronological age, implying an underlying lifespan clock. Complementary mRNA clocks were also informative of chronological age. **Conclusion.** Our results support the known association between body mass and lifespan, and indicate that the methylome provides a mechanistic link to accelerated aging.

Keywords: Epigenetic clock, DNA methylation, aging, lifespan, longevity, age acceleration

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Background

In the past few years, the "epigenetic clock" has emerged as a robust and widely used biomarker of aging that perhaps surpasses telomere length assays in its accuracy and utility [1-3]. Also referred to as the DNA methylation age (DNAmAge), the CpG based estimator of biological age comes in a few different versions for both humans and mice [4-11]. All these clocks share a common feature—they rely on the methylation status of preselected subsets of CpGs that are each assigned weights and are used collectively to estimate age. A critical question has been: are these DNAmAge clocks detecting changes that are purely a function of time and, therefore, correlates of chronological age? Or are they providing a measure of the intrinsic pace of biological aging that can be related to health, fitness, and life expectancy? Evidence from retrospective human epidemiological studies indicates that certain versions of the clock perform better at predicting life expectancy. In general, a younger DNAmAge relative to chronological age is associated with lower risk of disease and mortality [6, 12-18]. The age-dependent CpGs have also been studied in the context of lifespan variation among mammalian species [19], as well as variation within species, for example when lifespan is shifted by caloric restriction, treatment with rapamycin, or single gene mutations [7, 9, 10, 20-22].

As is the case with humans, aging trajectories vary considerably among mouse genotypes, and common DNA variants contribute to the pace of normal aging [23]. The BXD family has a long history in model organism aging and longevity research with median longevities of females ranging from at least 400 to 900 days [24-28]. Genomes of the BXDs are randomly recombinant versions of genomes of their two parents—strains C57BL/6J (B6) and DBA/2J (D2). On average,

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

D2 has a shorter lifespan than B6 [29, 30]. The more "accelerated aging" profile of D2 is consistent with other age-associated parameters such as more rapid thymic involution [31] and replicative senescence of hematopoietic progenitor cells [25], and increased tail tendon breakage in D2 compared to B6 [32, 33]. Variation in lifespan is much greater among the progeny BXD strains due to random assortment of independent gene variants [25].

Here, we have leveraged extensive experimental longevity data generated for the BXD family [28] to evaluate the associations between body weight, DNA methylation, aging, and lifespan. We used an affinity-capture enrichment followed by deep sequencing (MBD-seq) to profile the liver methylome in 12 members of the BXD family [34-36]. To evaluate the impact of a common metabolic stressor on aging, we also quantified the methylome of a subset of cases maintained on a high fat diet (HFD), which is known to decrease the longevity by as much as ~13% in the BXDs [28]. The main question we posed was: can we define DNA methylation clocks that are predictive of strain and diet dependent variation in lifespan? And how do these relate to strain characteristics such as body weight? Lastly, we tested whether age or lifespan predicted by methylome patterns can be carried over to the transcriptome, and whether transcripts can also be assembled into age-informative, and lifespan-predictive mRNA clocks.

Overall, the results reveal interdependence among body mass, the aging methylome, and lifespan. In particular, body weight of young adults (ca. 100 days) has a pronounced effect of DNA methylation. We were able to compute reasonably accurate DNAmAge estimates using subsets of age-informative CpG regions. A similar clock using CpG regions associated with body weight did not correlate with age, but was strongly correlated with strain lifespan. The clocks

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

77 highlight accelerated aging associated with higher initial body weight, and one version of the
78 DNAmAge estimator also showed accelerate aging on HFD. While mRNA-based clocks are a
79 feasible alternative, the methylome remains a more robust biomarker for aging and life
80 expectancy.

81 # Results

82 ## Characteristics of the study sample

83 The longevity data was collected from a cohort of females that were allowed to age until
84 mortality. (Details on this cohort in Roy et al., 2019 [28] and full data can be accessed from
85 GeneNetwork 2 [37].) A parallel cohort was used for biospecimen collection at different ages.
86 For the 12 selected strains and F1 hybrids, lifespans on *ad libitum* standard chow (control diet
87 or CD) ranged from an average of 417 ± 155 days (mean ± SD) to 933 ± 86 days (**Table 1**). HFD
88 generally reduces longevity, but with marked differences among strains [28]. Matched samples
89 from cases on HFD were included for five strains and sub-strains (**Table 1**). Each strain-by-diet
90 group was classified as short-lived (mean lifespan < 600 days), medium-lived (between 600 and
91 750 days), and long-lived (>800 days). A strain classified as long-lived on CD may also be
92 classified as short-lived on HFD, e.g., BXD65 (**Table 1**).

93 We performed methylome-wide assays in 70 liver samples collected from the strain and diet
94 groups at different ages. Since age and lifespan are the main variables of interest, liver
95 specimens were chosen so that distribution of age across the three lifespan groups are closely
96 matched **(Fig. 1a**; individual level sample information in **Table S1)**. We note that aside from

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

three male cases for BXD102, B6D2F1, and B6D2F1 (**Table 1**), all specimens were from females. While the samples were not chosen on the basis of body or organ weight, there was significant variation in body weight when the mice were initially weighed before introduction to HFD (**Fig. 1b)**. The F1s had higher body weights compared to BXDs on CD, and this hybrid vigor was apparent with or without the male cases. The final weight of mice (i.e., weight on day of sample collection) continued to show significant strain variation (**Fig. 1b**). The weight of the liver appeared fairly consistent across the strains (**Fig. 1b**). There was no group difference in body weight at baseline between the two diets (CD = 25 ± 7 g vs. HFD = 23 ± 5 g; $n = 70$). By final weighing, the group on HFD had become significantly heavier when compared to the full set of mice on CD (HFD = 41 ± 12 vs. CD = 29 ± 8 g, $p < 0.0001$; $n = 70$), or when compared only to the matched strains on CD (HFD = 41 ± 12 vs. CD = 26 ± 6 g, $p < 0.0001$; $n = 34$). The weight of the liver on HFD was slightly heavier but the effect was not statistically significant, likely due to the modest sampling of cases on HFD (HFD = 1.29 ± 0.23 g vs. CD = 1.22 ± 0.23 g; $p = 0.37$, $n = 34$). The baseline weight was measured at young adulthood (mean age of 134 ± 81 days) and at this point, age was a significant predictor of body weight with lower body weight for the younger mice (**Fig. 1c**). The age of mice at time of sample collection was, however, not correlated with final body weight ($r = 0.01$) or with weight of liver ($r = 0.12$). Instead, the baseline body weight remained a significant predictor of the final body weight **(Fig. 1d)** and liver weight **(Fig. 1e)**. When restricted to only the few HFD mice, baseline weight was still a significant correlate of final body weight ($r = 0.53$, $p = 0.05$, $n = 15$), but not liver weight ($r = 0.28$, $p = 0.30$).

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Characteristics of selected strains from the BXD panel

	Longevity trait				Biospecimen	
Strain/line	Mean lifespan (days)[1]	Median lifespan (days)[1]	Lifespan Group[2]	Diet[3]	N[3]	Age Range (days)[3]
B6D2F1	933±86	896	long	CD	5[4]	216–726
BXD102	861±222	891	long	CD	5[4]	183–714
BXD40	585±239	577	short	CD	8	284–719
BXD48	695±124	684	med	CD	3	188–731
BXD48	523±152	517	short	HFD	3	189–595
BXD48a	617±196	670	med	CD	3	233–604
BXD48a	635±113	650	med	HFD	3	233–543
BXD65	824±199	896	long	CD	6	181–711
BXD65	534±128	551	short	HFD	3[5]	230–541
BXD65b	726±91	751	med	CD	4	187–748
BXD73	702±116	687	med	CD	4	206–759
BXD73	699±112	715	med	HFD	3	206–694
BXD73b	820±129	807	long	CD	3	237–743
BXD73b	742±193	790	med	HFD	3	237–729
BXD79	417±155	330	short	CD	7	217–570
BXD9	507±135	462	short	CD	3	245–548
D2B6F1	771±143	791	long	CD	4[4]	210–744

[1] Average and median lifespans for strains under standard or high fat diet estimated from an aging cohort of mice co-housed with mice used for biospecimen collection and methylome assays

[2] Groups based on phenotypic lifespan: short = average lifespan < 600 days; med = average 640–750 days, and long = average lifespan > 800 days

[3] Diet, sample size, and age range of mice from MBD-seq data was generated. CD = control diet; HFD = high fat diet

[4] One male case; see Table S1 for individual level data

[5] One case excluded due to uncertain identity

Strain dependent patterns in global features of the methylome

Deep sequencing of the methylome was carried out after samples were enriched for methylated-CpGs (MBD-seq). After quality checks and filtering by read coverage, we retained a set of 368,300 regions, each 150 bp in length, with sufficient coverage in the 70 samples. The majority of the CpG regions (83%) contained no sequence variants (SNPs or small insertions/deletions) segregating in the BXDs. For the 17% with sequence variants, there was

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

an average of 2 ± 1.6 variants within the 150 bp bin. (Chromosomal coordinates, CpG density, variant counts, and normalized data will be made available from NCBI GEO upon official publication.) Consistent with the DNA enrichment and filtering protocols, the 368,300 CpG regions were enriched in annotated gene features and CpG islands and depleted in intergenic regions when compared to the genome-wide background (**Table S2**). Our first goal was to gain a global perspective, and for this, we performed two analyses: (1) dimension reduction with principal component analysis (PCA), and (2) analysis of genome-wide mean methylation and variance.

PC1 and PC2 captured 19% and 13% of the variance, respectively. A plot of the top two principal components (PCs) showed clustering of samples by strain identity, irrespective of diet (**Fig. 2a**). The one exception was a BXD65 on HFD; this case plotted away from the BXD65 cluster and, as this was of questionable identity, this sample was excluded from downstream analyses. Sub-strains (e.g., BXD73/BXD73b; BXD65/BXD65b) also clustered in close proximity with only slight separation. Unsupervised hierarchical clustering confirmed the clustering of samples by strain identity rather than age or diet groups (QC plots in **Fig. S1**). The top five PCs collectively explained 58% of the variance (PC1–PC5 in **Table S1**). We found no correlation between these five PCs and the age of mice. For strains with matched CD and HFD cases, the PCs did not differentiate between the two diets.

Next, we computed the within-individual genome-wide average methylation and variance. Global average methylation was inversely correlated with within-individual variance and showed extensive strain variability. When we divided the CpG regions into genic (i.e., CpG

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

regions that overlap annotated gene features) and intergenic sets, we found that the inverse correlation between global methylation and variance was significant for both sets but particularly pronounced for the intergenic set (**Fig. 2b, c;** and **Table S1** for individual level data on methylome means and variance). This reveals that for CpGs in intergenic regions, mice that have a hypermethylated profile have lower within-individual variance. There is extensive strain variation in this global methylation and variance patterns (**Fig. 2b, c**). For the intergenic set, the BXD73 sub-strains showed the lowest methylation scores and highest variance while the F1s had the highest methylation and lowest variance. Age was not a significant correlate of the genic and intergenic features. We evaluated if the top 5 PCs relate to the methylation averages and variance. The average methylation and variance at intergenic CpG regions were significantly correlated with PC1 (**Fig. 2d**), and PC3 (11% of variance) (**Fig. 2e**). PC4 (8% of variance) showed a significant negative correlation with mean methylation at genic CpG regions (**Fig. 2f**).

To summarize, the global analyses show that the clustering by strains in the PC plot, to a large extent, can be explained by the strain-dependent methylation averages. We did not detect any effect of age or diet on these large-scale methylome features.

Intercorrelations between body weight, methylome, and strain-level lifespan

We next examined whether the body weight measures could be associated with the large-scale methylome features. Since the F1s exhibit hybrid vigor both in body weight and mean methylation, we report the results after excluding the F1s (correlations with and without the F1s in **Table S3**). Of the top 5 PCs, only PC1 showed a weak but significant negative correlation

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with baseline weight (**Fig 3a**), and final weights of body and liver. Body weight at baseline, but not the final body weight, had a weak but significant negative correlation with mean methylation (**Fig 3b**), and significant positive correlation with variance (**Fig 3c**) at genic CpG regions that indicates a sustained inter-relatedness between body mass at younger age and the methylome. Unlike the baseline body weight, the final body weight was not correlated with methylation means or variance at either the genic or intergenic regions in the BXDs (**Table S3)**. We considered the possibility that the HFD-induced shifts in body weight may have reduced the correlation; however, when we restricted the analysis to only the CD BXDs (n = 46), the final body weight still showed no correlation with the intergenic and genic features. The liver weight had a significant positive correlation with mean methylation and negative correlation with variance at intergenic CpGs **(Table S3)**.

The lifespan data is from 17 strain-diet groups (**Table 1**), and while this is a small sampling of lifespan variability in the BXD panel, we used this longevity information to explore associations with the body weight measures and the global methylome features. This analysis was again performed without the F1s due to the vigor in both methylome features, body weight, and lifespan. Only PC4, an inverse correlate of genic CpG methylation means (**Fig. 2f**), showed a strong correlation with the strain longevity phenotype (**Fig 3d**). Lifespan showed no direct correlations with the methylation averages and variance at either the genic or intergenic CpG regions. For the weight measures, baseline body weight had a weak but significant negative correlation with strain mean lifespan (r = -0.27, p = 0.04; **Fig 3d**). This is in agreement with the stronger inverse correlation between body weight and lifespan that is seen in the larger BXD cohort [28].

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Taken together, while the individual correlations are weak (**Fig 3**), the results suggest that body weight at younger age is associated with multiple interdependent features of the global methylome that are also predictive of strain life expectancy.

Effect of age on site specific DNA methylation

Following the global analyses, we next evaluated the effect of age and diet on site-specific DNA methylation at each of the 368,300 CpG bins. Our goal was to identify age informative CpG regions that we can then use to estimate the epigenetic age of mice. We applied a multiple regression model for age and diet with adjustment for other major sources of variance (methylation ~ age + diet + PC1 + PC2 + PC3 + PC4 + PC5). Although we did not detect a generalized impact of aging on the global methylation, the distribution of *p*-values indicated a strong effect on few CpG regions (**Fig. 4a**). For the effect of diet, the *p*-value distribution was close to the null hypothesis (**Fig. 4b**), likely because with only 14 samples from the HFD group, the sampling is underpowered. For the remainder, we focused mainly on age, and considered diet only in the context of its effect on strain lifespan and potential age-accelerating effect.

At a Bonferroni threshold of 10% (unadjusted $p \leq 2.6 \times 10^{-7}$), there were 26 age-associated differentially methylation regions (age-DMRs). All of these were genic regions, including neighboring block of CpGs, and were located within 17 annotated genes **(Table S4)**. While only a small number of CpG bins were significant at the 10% Bonferroni threshold, given the non-independence of adjacent CpG regions, we used a lenient statistical threshold and selected the top 500 age-DMRs to define the general characteristics of the aging methylome (unadjusted $p \leq 3.1 \times 10^{-4}$, 23% FDR; Manhattan plot **Fig. 4c**; **Table S4**). Of these, 60% (299 of 500) were

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

218 associated with loss of methylation with age (age-hypomethylation), and 40% showed gains in
219 methylation with age (age-hypermethylation) [38, 39]. This set of 500 age-DMRs collectively
220 represented differentially methylated CpG regions located in or near 347 unique genes
221 (location within genes or distance to nearest promoter for intergenic CpG regions in **Table S4**).
222 Gene set enrichment analysis (GSEA) for the age-DMRs ranked by regression coefficient for age
223 showed no strong enrichment after FDR correction. However, we note that the most
224 overrepresented pathway was for signaling genes involved in stem cells pluripotency (KEGG ID
225 mmu04550; nominal enrichment p = 0.005, FDR = 0.10) and the five genes in this pathway
226 (*Fzd1*, *Fzd8*, *Wnt5a*, *Jak3*, *Meis1*) were associated with increase in DNA methylation with age.
227 Genes involved in mesenchyme development were also slightly overrepresented (GO ID
228 0060485; nominal enrichment p = 0.001, FDR = 0.24) and members of this GO (*Ptk7*, *Nrp2*,
229 *Sema5b*, *Zfp703*, and *Wnt5a*) were also associated with age-hypermethylation (**Fig. S2**).

230 Compared to the background set of 368,300 CpG bins, the age-DMRs were depleted in
231 intergenic regions and enriched in genic regions (enrichment and depletion p-values in **Table**
232 **S2**). The majority of the age-DMRs (86%) contained no sequence variants in the BXDs and the
233 age dependent methylation patterns are unlikely to be confounded by SNP or variant effects.
234 The age-hypermethylated set was highly enriched in CpG islands, promoters, and exons, and
235 also slightly enriched in 5'UTRs (**Fig. 4d**; **Table S2**). The age-hypomethylated set was enriched in
236 introns, and 3'UTRs and transcription termination sites (**Fig. 4d**; **Table S2**). For each CpG region,
237 we computed the average methylation and variance across the 69 samples, and compared
238 these to the age regression coefficients, which convey the change in methylation per unit
239 change in age. The most pronounced age-hypermethylation (positive regression coefficients for

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

age) occurred in bins with high CpG density and lower average methylation, and the magnitude of the age coefficient showed positive correlation with CpG density and negative correlation with average methylation (**Fig. 4e, f**). In contrast, age-hypomethylated regions (negative regression coefficient for age) featured low CpG density, and higher average methylation (**Fig. 4e, f**).

Age-DMR based epigenetic clocks and lifespan prediction

We next explored constructing age-informative and potentially lifespan-predictive clocks using three different approaches: (1) training-based age estimator by applying elastic net regression, (2) untrained age estimator with the set of 500 age-DMRs, and (3) untrained age estimator using only a subset of the age-DMRs that are also associated with cross-sectional variation in lifespan.

First, we implemented the standard clock-building protocol by deriving the clock in a training cohort, and then testing accuracy in a validation cohort [40, 41]. We are clearly limited in sample size; nevertheless, as assessment of feasibility, we randomly assigned 36 samples (52% of the 69 cases) to the training set to model an age-estimator (see **Table S1** for sample assignments). This constructed an age-estimator that was based on 60 "clock CpGs", i.e., CpG regions that are collectively used in determining the DNAmAge [5]. These clock CpGs included a few regions that individually showed no significant association with age; but most had modest to strong associations with age and included fifteen of the 500 age-DMRs (clock CpG regions and weighing coefficients in **Table S5**). In the training set, this age estimates had a near perfect correlation with chronological age at Pearson $r = 0.999$ (**Fig. 5a**; **Table 2**). In the test set of $n =$

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

33, the correlation between the estimated age and chronological age was 0.74 ($p < 0.0001$) with a median absolute error of 90 days (**Fig. 5b**; **Table 2**). Due to the small sample size of the test set, we did not use this clock to examine age acceleration.

For the non-training based estimator, we simply treated the 500 age-DMRs as possible clock CpGs and used the respective age regression coefficients as weighing factors to compute the weighted average for each sample (coefficients for each CpG region are in **Table S4**). The weighted averages were significantly correlated with chronological age of mice, and for a more direct comparison, the values were scaled to the age range for the 69 samples. This clock is therefore study-specific and calibrated to this cohort. The DNAmAge had a near linear correlation with chronological age at $r = 0.90$ ($p < 0.0001$), and with a median absolute error of 96 days (**Fig. 6a; Table 2**). We then estimated the age acceleration for each mouse (DNAmAge-acc) based on the residuals derived from the linear regression between DNAmAge and chronological age [9, 12] (individual-level DNAmAge data in **Table S1**). For this, positive residual values indicate an accelerated, and negative values a decelerated rate of biological aging. Unlike the DNAmAge, DNAmAge-acc was not correlated with the final age of mice, indicating that this measure of age acceleration is independent of chronological age [9]. The DNAmAge-acc derived from this clock did not correlate with the lifespan phenotype but showed a significant positive correlation with body weight measures that would suggest a more accelerated clock with increased body mass (**Fig. 6b, c**; **Table S6**). For the strains with matched samples, the DNAmAge-acc did not differentiate between the diet groups (**Fig. 6d**).

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We then examined if we can build a version of the DNAmAge estimator using only the age-DMRs that were also associated with strain differences in expected lifespan. We limited this analysis to the top 500 age-DMRs, and to test association with lifespan, we applied a mixed effects model with strain median lifespan and age as fixed variables and strain-diet as random variable. At a lenient statistical threshold of $p \leq 0.05$, there were 56 age-DMRs associated with lifespan and we refer to this set as putative lifespan-age-DMRs (**Table S4**). The majority of these (84% or 47 of 56) had lower methylation among the long-lived strains regardless of whether these were age-hypermethylated or age-hypomethylated (**Fig. 7a**). We illustrate this pattern using three of the top-ranking lifespan-age-DMRs: age-hypomethylated *Casz1* (**Fig. 7b**), and age-hypermethylated *Cyp46a1* (**Fig. 7c**) and *Abca7* (**Fig. 7d**). *Casz1* was associated with a small age-hypomethylation (age coefficient of only -0.0007, unadjusted $p = 1.6 \times 10^{-4}$) and most strain-diet groups had slight downward trajectory over time. *Cyp46a1* and *Abca7*, on the other hand, were associated with stronger age-hypermethylation, and almost all the strain-diet groups showed an upward trajectory with age. In addition to the change over time, cross-sectional comparison between lifespan groups showed lower average methylation in the long-lived strains relative to the medium and short-lived strain-diet groups.

While most of the CpG regions were only weakly associated with lifespan (for instance, *Abca7* in **Fig 7d**), we used the 56 lifespan-age-DMRs to calculate the DNAmAge by weighing each CpG by the respectively age regression coefficients (estimated ages in **Table S1**). Chronological age correlated strongly to this version of the DNAmAge ($r = 0.83$, $p < 0.0001$), although the trend appeared to fit a quadratic equation with a greater degree of plateau at older age, an observation that is consistent with the apparent underestimation or possible decline in aging

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

303 rate with increased chronological age [15, 42, 43] (**Fig. 6e**). The DNAmAge-acc derived from
304 this version of the clock showed a significant negative correlation with the longevity phenotype
305 (both mean and median lifespan) with more accelerated aging in mice from the short-lived
306 strain-diet groups (**Fig. 6f**). This correlation was robust and remained significant after excluding
307 the F1s (full correlation and probabilities in **Table S6**). The DNAmAge-acc was also positively
308 correlated with the body weight measures, again indicating more accelerated aging in mice
309 with higher body mass (**Fig. 6g, Table S6**). For the strains with matched samples from CD and
310 HFD, the group on HFD had a mean age acceleration of 21 ± 36 days compared to the mean
311 decelerated aging (-26 ± 57 days) for the group on CD (p = 0.01; n = 33; **Fig. 6h**). Therefore, this
312 DNAmAge estimator, built from fewer CpG regions, is therefore both age-informative and
313 lifespan-predictive.

314 **Table 2. Versions of DNA methylation and mRNA clocks**

Clock versions	Sample N	Clock CpGs/mRNAs	Age r[1]	median |error|[2] (days)	BW0 r[3]	Lifespan r[3]
DNA methylation elastic net	36 training	60	0.99 (<0.0001)	4.5	not tested	not tested
	33 testing		0.74 (<0.0001)	90	not tested	not tested
500 age-DMRs	69	500	0.9 (<0.0001)	96	0.38 (0.001)	0.004 (ns)
56 lifespan-ageDMRs	69	56	0.83 (<0.0001)	111	0.24 (0.05)	r = -0.58 (<0.0001)
500 BW0-DMRs	69	500	0.12 (ns)	*66	*r = -0.52 (<0.0001)	*0.47 (<0.0001)
RNA clock elastic net Ver1	150 training	65	0.89 (<0.0001)	52	not tested	not tested
	141 testing		0.65 (<0.0001)	85	-0.17 (0.05)	-0.15 (ns)
RNA clock elastic net	141 training	62	0.90 (<0.0001)	67	not tested	not tested

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ver2	150 testing	0.62 (<0.0001)	100	0.19 (0.02)	0.08 (ns)

[1] Pearson correlation between clock estimates and chronological age; significance within parenthesis
[2] Median of the absolute difference between estimated and chronological age as a measure of DNAmAge precision as defined in [40]. * For the 500 BW0-DMR clock, median of the absolute difference between estimates and mean lifespan was used as a measure of lifespan clock accuracy.
[3] Pearson correlation between age acceleration residuals and baseline body weight and strain mean lifespan. * For the 500 BW0-DMR clock, the estimated values instead of the residuals were used.

Baseline body weight based epigenetic clock and lifespan prediction

Given the strong influence on baseline body weight on the methylome, we then explored the possibility of deriving lifespan informative estimator from body weight associated CpG regions. To identify relevant CpG regions, we applied a regression analysis with baseline body weight (BW0) as predictor with adjustment for age and top 5 PCs as covariates. While age was associated with only a few strong age-DMRs, BW0 had a more widespread association with the methylome (**Fig. 8a**). At a Bonferroni threshold of 10% (unadjusted $p \leq 2.6 \times 10^{-7}$), there were 667 CpG regions associated with body weight. Unlike the age-DMRs, which were depleted in intergenic regions, 74% of the baseline body weight associated differentially methylated CpG regions (BW0-DMRs) were intergenic. The vast majority (581 of the 667) had positive regression estimates for body weight, i.e., higher levels of methylation in heavier mice.

As in the case of the age-DMR based clock, we then took the top 500 BW0-DMRs and weighed each CpG region by the respective age regression coefficient (weighing coefficients and the list of CpG regions in **Table S7**). We note that 27.6% (138 of the 500) of the BW0-DMRs contained sequence variants, and the differential methylation at these sites could also be capturing the effect of underlying sequence effects. Using the BW0-DMRs, the weighted averages did not

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

correlate with the chronological age of mice. Instead, the values had a significant positive correlation with strain lifespan. For direct comparison, we rescaled the weighted averages to the range of median lifespan for the study sample **(Fig. 8b**; **Table S1)**. The strong positive correlation with lifespan suggests that this clock is an indicator of the expected age at death rather than the chronological age of mice. The median absolute error for the predicted lifespan relative to the known lifespan for the strain-diet groups was at 66 days (**Table 2**), and we refer to this as a baseline body weight based lifespan clock (BW0.lifespan). For a few of the long-lived strains, particularly the robust F1 hybrids that have both higher body weight and longer lifespan (labelled as groups A and B **in Fig. 8b**), the BW0.lifespan was much lower than the expected lifespan. Since the CpGs regions were individually associated with body weight, as expected, the BW0.lifespan showed a strong correlation with the baseline weight (**Fig. 8c**) and with liver weight, but not strongly with final body weight (**Table S6**). For strains that had both CD and HFD cases, BW0.lifespan did not differentiate between the diet groups, suggesting that this clock is not sensitive to lifespan modification by HFD but is mainly dependent on strain differences in baseline body weight. The age acceleration residuals from the 56 lifespan-age-DMRs had a strong inverse correlation with the BW0.lifespan estimates indicating convergent information from both the clocks (**Fig. 8d**).

Impact on gene expression and mRNA based clocks

To examine age-dependent changes in the liver transcriptome, we made use of available liver RNA sequencing (RNA-seq) data from a larger cohort of aging BXDs. The data obtained from GeneNetwork was generated in two batches, and sample and batch information are in **Table**

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

S8. First, to examine whether the transcripts that correspond to the age-DMRs also show changes in expression levels with age, the 347 unique genes represented by the 500 age-DMR CpGs were matched to the corresponding transcripts by gene symbol. This resulted in 265 age-DMRs and transcript pairs. Of these, 110 of the transcripts (42%) were correlated with age at an unadjusted $p \leq 0.05$ ($|r| \geq 0.12$ in $n = 291$) (**Table S9**). The majority of the age-hypermethylated CpG regions were associated with increased gene expression with age with only nine age-hypermethylated DMRs (e.g., *Fzd8*, *Cyp46a1*, *Nfix*, *Rnf4*) associated with decreased gene expression with age (**Table S9**). For the age-hypomethylated DMRs, 37 transcripts showed a decrease in expression and 39 transcripts showed an increase in expression with age. Overall, this shows that many of the age-DMRs are also associated with gene expression changes with age.

Finally, we attempted to derive age-informative clocks using the training-based approach. We first performed the training in batch 1 and then tested in batch 2, and then vice versa using the liver transcriptome-wide data. The elastic net modeling trained on the 150 samples in batch 1 identified 65 transcripts for mRNA clock estimation (**Table S10**), and the age estimated from these transcripts (mRNAage) had a correlation of $r = 0.89$ with chronological age in the training set (**Fig. 5c**; **Table 2**). In the validation set (i.e., batch 2 with 141 samples), the correlation between the age estimates and chronological age was 0.65 with median error of 85 days (**Fig. 5d**; **Table 2**). The age acceleration, as estimated by the residuals from the regression between mRNAage and chronological age, showed a weak but significant ($p = 0.05$) negative correlation with baseline body weight, which is at odds with the expected age acceleration with higher body mass, and no significant correlation with strain lifespan (mRNA age estimates and age

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

acceleration for individual samples in **Table S8**). Training in batch 2 identified 62 transcripts for clock estimation **(Table S11**). Of these 62 transcripts, only 9 transcripts (*Rin3*, *Adcy3*, *Cpn1*, *Chuk*, *Rpl41*, *Rps27a*, *Serpina3m*, *Srm*, and *Ubb*) overlapped between the two versions of the mRNA clocks. Correlation between the mRNAage and chronological age was 0.90 in the training set, and 0.62 in the testing set (**Fig 5e**, **5f**; **Table 2**). The age acceleration for this version of the clock showed a significant positive correlation with baseline body weight and no correlation with lifespan (**Table S8**). For strains with samples from both CD and HFD, the age accelerations derived from the mRNA clocks did not differentiate between diets.

Discussion

The results we have presented convey deep associations between body mass, DNA methylation, aging, and lifespan. Somewhat to our surprise, we found that body weight at young adulthood, more so than age at time of tissue collection, had a stronger influence on the large-scale methylome. Age, on the other hand, had a strong effect on few discrete CpG regions. Age and weight appeared to exert effects on largely independent sets of CpGs, with the age-DMRs mostly located within gene bodies, while the BW0-DMRs included a large proportion of intergenic loci. Nonetheless, longevity information could be derived from both the age- and BW0-DMRs.

Consistency in age-dependent DNA methylation changes

In regard to whether genomic characteristics can explain whether a CpG region gains or loses methylation with time, we found a clear dependence on CpG density and methylation status

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that agreed with previous reports [44, 45]. Since DNA methylation levels were quantified over 150 bp non-overlapping bins, we could directly examine how the local CpG density relates to (1) average methylation levels and variance across the 69 samples, and (2) the change over time using the age regression coefficients. The results show that loss of methylation over time occurs at regions with low CpG density that also generally have higher average methylation. This is consistent with the general age-dependent hypomethylation at CpG spares regions of the genome where the few or isolated CpGs exist mostly in a methylated state [46-49]. For age-hypermethylated regions, the gains in methylation was directly correlated with CpG density, and inversely correlated with methylation, and occurred in regions with lower average methylation. This too, is consistent with reports that CpG dense regions on the genome, a feature of CpG islands, which typically remain unmethylated, are the sites that tend to gain methylation with age [50-53].

For the 347 unique genes that correspond to the 500 age-DMRs, we did not find a particularly strong functional theme. However, the list included notable members that have previously been linked to aging related diseases or with longevity and lifespan. For example, *Cyp46a1*, *Lrp1* and *Abca7* play roles in lipid shuttle and cholesterol metabolism, and have been associated with Alzheimer's disease [54, 55]. Other age-DMR genes that have been previously implicated in aging and human longevity include *Adarb2* [56], *Abcc4* [57], *Igf2r* [58], *Ucp3* [59], *Grb2* [60], *Il7r* [61], *Ikbkb* [62]. The current data replicated the CpG islands in *C1ql3* and *Ptk7*, which we previously reported as age-hypermethylated sites in the BXD parental strains, C57BL/6J and DBA/2J [63]. We also compared the list of age-DMRs from the present study to a list of 79 genes reported to harbor age-DMRs in a recent study of DNA methylation in aging human liver

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[42], and found only three genes in common: *Il4i1*, *Ildr2* and *Nfix*. The study by Bacalini et al. [42] reported an enrichment in mesenchymal and Wnt-signaling pathways, and similarly, we also find a slight over-representation of genes that are part of signaling pathways and mesenchyme development (e.g., *Fzd1*, *Fzd8*, *Wnt5a*, *Jak3*, *Ptk7*, *Nrp2*, etc.).

Building clocks from age dependent CpG regions

Currently, there are several different versions of the DNAmAge estimator available for both mice and humans. Some have multi-tissue application [5, 8, 9], and others are optimized for specific tissues [4, 7, 10, 11]. The standard protocol for developing DNAmAge clocks starts by applying a linear regression algorithm in a training dataset, followed by age estimation in validation cohorts to gauge the accuracy of the clock [5, 40, 41]. The training sessions identify age informative CpGs, referred to as clock CpGs, that are each assigned predetermined weight for estimating DNAmAge. This preselected ensemble of clock CpGs can then be used to estimate age in other independent data sets, under the condition that the same or at least the majority of the CpGs are also measured. To our knowledge, all these clocks have relied on bisulfite-based assays that provide single CpG resolution [4-11, 64, 65]. These existing clocks rarely share CpGs in common. For instance, a comparison of the three mouse clocks showed only two CpGs in common between the blood-based and multi-tissue clocks [7, 8, 41]. This is not particularly surprising given that these were constructed from reduced representation bisulfite sequencing (RRBS) data, a protocol for which CpG coverage can have low overlap between libraries. But even in the case of the human clocks that were built from microarray data with fixed and limited number of CpGs to select from, the overlap is still poor. The Horvath

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and Hannum clocks share only about six CpGs in common. This means that DNAmAge can be defined and calibrated from numerous different subsets of potential clock CpGs [9, 41].

While MBD-sequencing does not provide single CpG resolution, it is nonetheless known to deliver highly sensitive and reliable quantification of genome-wide methylation [34-36, 39]. Instead of measuring individual CpGs, MBD-seq quantifies methylation levels within circumscribed regions containing spatially proximal CpGs that are generally correlated in methylation patterns [66-68]. To test feasibility, we started by training to chronological age using linear regression. Despite the extremely humble sample number, the estimated age had a near perfect correlation with chronological age in the training set (r = 0.99). In the test set of 23 mice, the age correlation was at 0.74 with a fairly high median error of ± 90 days. While not precise, this DNAmAge estimator, consisting of 60 clock CpG regions, could still be applied to classify mice into broad age groups (young, middle, old), and may perform just as well as transcriptomic age-classifiers that have been defined from much larger sample numbers [69, 70]. The constituent clock CpG regions included sites in or near 29 genes that were represented among the age-DMRs (e.g., *Cyp46a1*, *Fzd8*, *Gata6*), although in terms of the exact CpG regions, only 15 age-DMRs were present in both lists (see **Table S5**). The RRBS-based clock published by Wang et al. [10] was for the liver tissue, and we compared the genes in their list of 149 clock CpGs and we found two common genes, *Cyp46a1* and *Sulf2*, represented in both, although again, the precise CpGs did not overlap.

We also applied a direct weighted averaging method that did not impose a training step. Both versions of the age-DMR based clocks (i.e., one built using the full set of 500 age-DMRs, and

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

another built from a restricted set of only 56 lifespan-age-DMRS) provided fairly close estimates of chronological ages. Of these two DNAmAge estimators, only the DNAmAge_56 appeared sensitive to lifespan variability, and the age acceleration derived from this clock was predictive of strain lifespan, and also accelerated in the HFD group. Using this clock, almost all cases from strain-diet groups classified as long-lived had negative DNAmAge-acc; the only exceptions were four F1 samples, including the male B6D2F1, that had positive DNAmAge-acc. For cases belonging to strain-diet groups classified as short-lived, all had positive DNAmAge-acc with the exception of one BXD65 kept on HFD that had negative DNAmAge-acc. The strain with the most decelerated clock, and presumably slowest rate of biological aging, was BXD102 on CD, which is also the longest lived BXD strain we had in the study (**Table 1**). For these mice, the DNAmAge-acc for the female were highly decelerated and ranged from –127 to –194 days. Interestingly, the one male BXD102 had a less decelerated clock with a DNAmAge-acc of –56 days.

The members of the DNAmAge_56 clock were selected because these CpGs were both correlates of age, and predictive of lifespan. The majority of these (31 of the 56) were age-hypermethylated and had lower methylation among the long-lived strains (lower right box of **Fig. 7a**). These CpGs also had the most pronounced association with both age and lifespan and were CpGs in the introns, exons and promoters of genes such as *Jak3*, *Wnt5a*, *Abca7* and *Cyp46a1*. The relation between the regression coefficients for age and lifespan would suggest that for these CpGs that gain methylation with age, the long-lived strains have a more "youthful" profile, i.e., lower average methylation. However, 16 of the age-lifespan-DMRs were age-hypomethylated, and yet had lower methylation among the long-lived strains (lower left box of **Fig. 7a**). A few epidemiological studies have also attempted to define predictors of all-

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cause mortality risk in humans, and these studies paint a very complex picture for humans. While two of the studies showed that hypomethylation at specific CpGs was prognostic of poor survival [71, 72], another study found a combination of both hypo- and hyper-methylated CpGs linked to mortality [73]. A complication with the human studies is that several of the mortality CpGs were also markers for tobacco use, indicating that a good proportion of the CpGs were likely capturing the effects of lifestyle and environmental exposures [72, 74, 75]. In the BXDs, given the controlled environment, the CpG methylation patterns are presumably providing a closer marker of genetic and epigenetic predisposition to life expectancy.

Body weight, DNA methylation, and lifespan prediction

We also explored a less conventional clock based on the methylation levels at the top 500 BW0-DMRs. This clock does not qualify as a DNAmAge clock since the estimates were uncorrelated with chronological age. Instead, the estimates were correlated with strain longevity data and we tentatively refer to this as lifespan clock (BW0.lifespan clock). The main mismatches between the lifespan predicted by the BW0.lifespan clock versus the recorded strain lifespans were for the long-lived F1 hybrids and BXD65, members of the BXD panel that, in this particular cohort, also exhibited higher body weight. For these F1s and strain, recorded median lifespans are over 790 days; the BW0.lifespan clock incorrectly predicted much shorter lifespans for these cases (all at under 566 days). We note that the simple averages for the top 500 BW0-DMRs (i.e., not weighed by any variable) were also correlated with lifespan but to a lesser degree ($r = –0.41$, $p = 0.0004$). Incorporating the age information by using the age regression

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

coefficients as weighing variables enhanced the predictive power and made it a directly comparable to the lifespan data.

A key point of distinction between the DNAmAge and BW0.lifespan clocks is that the DNAmAge is primarily an estimate of chronological age, and the age acceleration is a secondary derivative. The higher DNAmAge-acc in the HFD group also implied that the rate of the clock, as measured by the DNAmAge_56 clock, is a modifiable outcome. The BW0.lifespan clock on the other hand, provided a direct estimate of expected lifespan dependent on a phenotype from a younger age. One could speculate that the BW0.lifespan clock represents a genotype dependent "lifespan potential", while the DNAmAge-acc represents a health-related trait that can be accelerated or decelerated over the course of life, and is responsive to health promoting interventions. This distinction aside, the DNAmAge and the BW0.lifespan clocks had consistent associations with body weight. Specifically, DNAmAge-acc derived from both the 500 and 56 age-DMR clocks showed positive correlation with baseline body weight indicating more accelerated aging for mice with higher body mass. For the BW0.lifespan clock, the estimates were negatively correlated with body weight, also indicating shorter lifespan for mice with higher body mass. This conveys remarkably close interrelations among body weight/mass, the aging methylome and final lifespan. The influence of body mass on longevity and the more favorable health profile associated with smaller body mass has been known for a long time [76-81]. The results we present are further evidence that the epigenome plays a mechanistic intermediary role that links body weight to the rate of biological aging, and the DNA methylation clocks likely capture the effects of weight and other biological factors that collectively influence the longevity of the BXDs.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In terms of the direction of causality (i.e., whether the methylome patterns are due to body weight, or body weight differences due to DMRs), based on the timeline of data collection, a simple inference would be that the baseline body weight predicted the methylome patterns and the epigenetic age acceleration. In humans, differences in DNA methylation also appear to be more the effect of, rather that the cause for differences in body mass index [82]. However, both body weight and DNA methylation are genetically regulated phenotypes that are tightly interlinked. And in these mice, body weight from young adulthood had a sustained or even lifelong effect of weight measured later in life. This was also true for the few mice that were introduced to HFD as the mice that were heavier at young age, prior to HFD, continued to be heavier on HFD, indicating a genetic predisposition to heavier bodies. Given this, we are inclined to interpret that these three phenotypic domains (body weight, methylome, longevity) as inter-dependent and are ultimately modulated by underlying DNA sequence variation. A well-powered genetic dissection of these phenotypes will certainly shed more light on the nature of this inter-relatedness.

The analysis with the RNA-seq data shows that the aging methylome may have an impact on gene expression. However, the mRNA clocks, derived from much larger sample sizes, had lower correlations with chronological age. Furthermore, the age-accelerations from the two versions of the mRNA clocks had inconsistent associations with body weight, and were not predictive of lifespan.

Technical considerations and caveats

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

549 Before concluding, we should address a few caveats. The sequence alignment was done to the
550 mm10 B6 reference genome, which means that for regions with D2 haplotypes, the sequence
551 differences could compromise alignment. The strong population stratification in the PC plot is
552 likely the result of true quantitative variation in methylation, and also due to a portion of the
553 CpG regions serving as surrogates for underlying genotype [83]. The main methylome-wide
554 analysis we performed was for age, and then baseline body weight. Since age is independent of
555 strain and the different BXDs and lifespan groups were represented across the entire age range,
556 the age-DMRs are expected to be less susceptible to the confounding effect of DNA sequence
557 variants. For body weight, a phenotype that is closely linked to genotype, differential
558 quantification due to sequence effects could be more problematic. To partly control for this, we
559 included the top five PCs in the regression models as these PCs capture a portion of the
560 genotype effects and also possibly other unmeasured confounder variables (e.g., cryptic batch
561 effects, cell composition) [39, 84]. For the top 500 age-DMR, only 71 bins (14%) contained at
562 least one sequence variant and the remaining were monomorphic. This is lower than the 17%
563 of variant containing bins in the background set of 368,300 regions. For the top 500 BW0-
564 DMRs, 139 (27.6%) contained sequence variants, and this higher proportion shows that the
565 BW0-DMRs occur in genomic loci that are variable in the BXD, and some of the differential
566 methylation estimates may be conflated by sequence variants. Nonetheless, the majority of the
567 BW0.DMRs used to estimate the clock by weighing on the age coefficients were devoid of
568 sequence variants.

569 Another weakness to consider is that the present study made use of DNA from frozen liver
570 tissue and does not account for cellular composition. This is an issue that is inherent to

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

571 epigenomic and transcriptomic data from bulk tissue [85]. While the liver is relatively
572 homogeneous in cellular makeup, at least compared to blood or brain, cellular heterogeneity is
573 still a confounding factor, particularly since this is known to increase with age [53, 85]. For this
574 issue too, incorporating the top PCs in the statistics is an *in silico* method to partly mitigate the
575 confounding effect.

576 Conclusion

577 In conclusion, we have provided a comprehensive description of the aging methylome in a
578 subset of the BXD panel. Our results demonstrate that the DNA methylations based epigenetic
579 clock is sensitive to subtle differences in natural lifespan that arise from common genetic
580 variants. The effect of body mass and other biological and environmental variables on lifespan
581 may also be mechanistically mediated by the methylome.

582 Materials and Methods

583 Sample preparation and high throughput sequencing

584 Liver samples were obtained from the BXD colony maintained at the University of Tennessee
585 Health Science Center (see Roy et al. [28] for details). All animal procedures were in accordance
586 to the protocol approved by the Institutional Animal Care and Use Committee (IACUC) at the
587 University of Tennessee Health Science Center. DNA from liver tissue was extracted using the
588 DNeasy Blood & Tissue Kit from Qiagen. Nucleic acid purity was inspected using a NanoDrop
589 spectrophotometer, and quantified using Qubit fluorometer dsDNA BR Assay. Affinity based

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

590 enrichment was carried out using the MethylMiner DNA enrichment kit from ThermoFisher
591 Scientific according to the standard manufacturer's protocol. This kit relies on the methyl-CpG
592 binding domain (MBD) protein to capture methylated CpG fragments. In brief, 1 µg of DNA in
593 110 µl of low TE buffer was fragmented to ~150 bp using a Covaris S2 ultrasonicator. The
594 sonication settings were: cycle/burst of 1 for 10 cycles of 60 s, duty cycle of 10%, intensity of
595 5.0. DNA fragment size and quality were assessed using the Agilent Bioanalyzer 2100. Following
596 the MBD capture reactions, DNA was eluted in a single step using the high salt (2000 mM NaCl)
597 elution buffer, and re-concentrated by ethanol precipitation. The final concentration of
598 methylated-CpG enriched DNA ranged from 1.12 to 5.43 ηg per µl (2.46 ± 0.94). The library
599 construction and sequencing were carried out at Novogene Genomic Services facility.
600 Sequencing was done to a depth of approximately 50 million reads per sample (150 bp paired-
601 end) on the Illumina HiSeq 4000.

602 **Read alignment and data quality**

603 The FASTQ files were first inspected with the FastQC tool (v.0.11.8) [86], and all files passed the
604 initial quality checks. Alignment was then done to the mouse reference genome
605 (mm10/GRCm38) using the Bowtie2 aligner (v.2.3.4.3) [87], and alignment quality was
606 evaluated with SAMtools (v.1.9) [88], and SAMstat v.1.5.1[89]. Potential PCR duplicates and
607 reads with mapping quality less than 10 were filtered out, and indexed Bam files were created.
608 The Bam files were then loaded to the MEDIPS R package (v.1.36.0) [90], for additional quality
609 control and assessment of read coverage. Saturation analysis (MEDIPS.saturation) showed that
610 all 70 libraries had sufficient read coverage, and pair-wise correlations (MEDIPS.correlation)

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

showed high consistency between samples (Pearson $r > 0.90$ for all pairs). Given the MBD enrichment, all the samples were enriched for CpGs (mean CG enrichment score of 2.30 ± 0.19; using the MEDIPS MEDIPS.CpGenrich function), and on average, 51% of CpGs in the reference genome was covered by at least one mapped read, with 28% of CpGs covered by > 5 mapped reads. To quantify DNA methylation, the mouse genome was divided in 150 bp non-overlapping windows and reads were counted for each bin with normalization to the local CpG density (referred to as coupling factor or CF) using the function MEDIPS.meth and the following parameters: ws =150, extend = 150, uniq = 1, shift=0.

Global methylome analysis

Read counts generated in MEDIPS were filtered to retain only the 150 bp bins that had sufficient coverage for reliable quantification and statistical analyses. First, bins with no CpGs (CF = 0) and mean read counts ≤ 1 were excluded, resulting in 4,286,826 bins. The Y chromosome was also excluded. Following these filters, the data was loaded to the EdgeR R package (v3.24) [91] and further filtered on the basis of counts per million (CPM) such than only reads with more than 1 count per million in 2 or more libraries were retained. This resulted in 368,300 CpG regions with sufficient coverage across the libraries, these were normalized using the calcNormFactors in EdgeR, and RPKM values extracted using the parameters gene.length = 150, log = TRUE. These 368,300 bins had a mean CpG density of 5.4 ± 2.6 and in total cover 2,001,723 CpG sites. The compendium of SNPs and small insertions/deletions segregating in the BXDs have been catalogued for the BXDs [92, 93], and we used this information to count the number of variants in each of the 368,300 150 bp bins.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To evaluate sample clustering and to detect potential outliers, we performed PCA and hierarchical clustering (**Fig. S1a**). There were no outlier samples and the DNA methylation profile was highly consistent across the samples (**Fig. S1b**) and averaged at 3.8 ± 0.72 logRPKM. These CpG regions were then annotated for genomic features using the HOMER program (v4.10) [94] and evaluated for enrichment relative to the genome-wide set using a hypergeometric test (R codes in **Table S2**). Based on the annotations, the CpG regions were then divided into bins that occurred within annotated genes (genic set, 200531 bins), and those that were in intergenic regions (167769 bins). For each sample, the overall average methylation and variance for these genic and intergenic sets were computed. The intercorrelations between the large-scale methylome features, body weight measures, and strain-level lifespan phenotype were examined using Pearson correlations.

Statistics for differential methylation analyses

To detect age-DMRs, we applied a generalized multiple regression model that included diet and the top 5 PCs: glm(logRPKM ~ age + diet + PC1 + PC2 + PC3 + PC4 + PC5). The top 5 PCs were included so that the age-effect can be adjusted for other major sources of variance, including strain effect and other unmeasured potential confounder variables. The top age-DMRs were selected at a lenient statistical threshold of unadjusted $p \leq 3.1 \times 10^{-4}$ (23% FDR) since our primary goal was to examine the overall trends and to test if these sites can collectively be used to estimate chronological age and lifespan prediction. For these, GSEA was carried out using the WebGestalt platform (http://www.webgestalt.org) with genes ranked by the age regression coefficients [95, 96]. To test if the top 500 age-DMRs were associated with strain-level lifespan,

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

we applied a mixed effects model using the lme4 R package (v4_1.1-19) [97]. The lifespan data are at the level of strain within a diet group (CD or HFD), and to test association between the individual level DNA methylation and the strain-diet longevity data, we treated strain-diet as a random variable. The mixed effects model was: lmer(logRPKM ~ age + median.lifespan + (1 + age|strain.diet)), where age and median lifespan are numeric variables and strain.diet is a categorical identity. To identify CpG regions that are associated with baseline body weight variation, we used the regression model: glm(logRPKM ~ BW0 + age + PC1 + PC2 + PC3 + PC4 + PC5), where BW0 was baseline body weight as a numeric variable.

Epigenetic clock calculation

For the training-based clock construction, we applied elastic net regression (alpha=0.5 using the glmnet R package (v2.0-18) [98, 99]. For the DNA methylation data, we randomly selected 36 of the 69 samples as a training set, and training was done using the log RPKM values for the 368,300 CpG bins. The untransformed chronological age of mice in days was entered as the training variable. Model parameters were optimized using a 10-fold cross-validation with the following cv.glmnet parameter: nfolds = 50, alpha = 0,5, family ="gaussian". The performance of the age estimator were then tested in the remaining 33 MBD-seq samples. The accuracy of the predicted age was examined by Pearson correlation with chronological age. As a second measure of accuracy, we also used the median absolute error as described in Horvath & Raj [40]. As recommended in Thompson et al. 2018 [9], the "age acceleration" was computed as the residuals after fitting the predicted age to chronological ages: residuals(lm(DNAmAge ~ Age)).

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

For non-training DNA methylation clocks, each CpG region was weighed by the age regression coefficient derived from the equation: glm(logRPKM ~ age + diet + PC1 + PC2 + PC3 + PC4 + PC5), which means that each CpG region was weighed by the change in DNA methylation per unit change in age. For the for the 500 age-DMR version, the weighted average was computed for the 500 CpG regions; for the 56 lifespan-age-DMR version, the weighted averages were computed from the restricted set of 56 CpG regions (for each CpG region, the weighing factor is in the Estimate.Age column of **Table 54**). The DNAmAge_56 and DNAmAge_500 clocks were then derived by scaling the weighted average to the age scale in the 69 samples using the following formula: DNAmAge = (((weighted.average – min.weighted.average) x age.range) / weighted.average.range) + min.age, where min.weighted.average and weighted.average.range are the minimum value and range for the weighted averages in the 69 BXD samples, and age.range = 578 days is the range of chronological age in the 69 BXDs, and min.age = 181 days is the minimum age for the 69 BXDs. For the body weighted based clock, we again used the age coefficient from the same regression model as a weighing factor, but the weighted averages were computed for the top 500 CpG regions that were associated significantly with baseline body weight of mice. The same scaling formula was applied to bring the weighted averages to the same scale the median lifespan data.

Transcriptomes analyses

We used liver gene expression data from a larger cohort of 291 BXDs that is available from GeneNetwork 2 [100]. The RNA sequencing was carried out in two batches, with batch 1 generated in 2017 from 150 samples (77 CD and 73 HFD cases), and batch 2 generated in 2018

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

695 from another set of 141 samples (84 CD and 57 HFD). While the liver specimens all came from
696 the UTHSC Aging BXD Colony, the two batches did not share overlapping samples. Batch 2 had
697 lower sequencing depth than batch 1, and to make the two batches more comparable, we
698 excluded all transcripts with no read coverage (i.e., logRPKM values of 0) in 10% or more of the
699 291 samples, and this retained 25,676 Ensemble transcript IDs. The 347 age-DMR genes were
700 matched by gene symbol to the corresponding transcripts, and only 265 of the age-DMR genes
701 paired to one or more transcript variants in the liver transcriptome. We performed a simple
702 Pearson correlation between the expression of these transcripts and age using data from both
703 batches. In the case of age-DMR genes that matched to multiple transcript variants (specifically,
704 different Ensemble transcript variants) from the same gene, we retained only the transcript
705 with the most significant correlation with age that resulted in 265 unique age-DMR and mRNA
706 pairs (**Table S9**).

707 For mRNA based age estimation, we considered the 25,676 transcripts and applied the same
708 elastic net regression parameters as described for the DNA methylation clocks. In the case of
709 the RNA-seq data, the training to age was first carried out in the 150 cases from batch 1,
710 followed by validation in the 141 cases from batch 2. This was then complemented by
711 performing the training in the 141 cases from batch 2, followed by validate in batch 1.

712 ## Data availability

713 The normalized MBD-seq data for the 368,300 CpG bins that were considered for statistical
714 analyses and sample metadata will be available from the NCBI NIH Gene Expression

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

715 Omnibus upon official publication. The full raw alignment files for the MBD-seq data will be

716 made available from the NCBI NIH Sequence Repository Archive upon official publication.

717 Abbreviations

718 age-DMR: age-associated Differentially Methylation Region

719 bp: base pair

720 BW0-DMR: Baseline body Weight associated differentially methylated CpG region

721 BW0: Baseline body Weight

722 BW0.lifespan: baseline weight-based lifespan clock

723 CD: Control Diet

724 CF: Coupling Factor

725 CPM: Counts Per Million

726 DNAmAge-acc: DNA methylation age acceleration

727 DNAmAge: DNA methylation age

728 GO: Gene Ontology

729 GSEA: Gene Set Enrichment Analysis

730 HFD: High Fat Diet

731 KEGG: Kyoto Encyclopedia of Genes and Genomes

732 MBD-seq: Methyl-CpG binding domain sequencing

733 mRNAage: mRNA based age estimates

734 RRBS: Reduced Representation Bisulfite Sequencing

735 PC: Principal Component

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

PCA: Principal Component Analysis

QC: Quality Control

RPKM: Reads Per Kilo base per Million mapped reads

SNP: Single Nucleotide Polymorphism

Declarations

Ethics approval and consent to participate. All animal procedures were in accordance to protocol approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Tennessee Health Science Center.

Consent for publication. All author read and approved the current version of the manuscript

Availability of data and materials. The normalized data and the raw fastq files will be made available from NCBI NIH Gene Expression Omnibus and the Sequence Repository Archive upon official publication.

Competing interests. No competing interests

Funding: Funded by NIH NIA grants R21AG055841 and R01AG043930

Authors' contributions

KM supervised the study, contributed to conception of study design and analysis, and wrote the manuscript. RWW contributed to the conception of the BXD Aging Colony, and provided access to the biorepository resource. JVSS, AHBH, and EGW contributed to the lab work, and JVSS performed the primary bioinformatics and contributed to analysis and manuscript. DA and SR contributed to analysis and data acquisition. All authors contributed to and approved the final version of the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements

The present work relied on resources from the BXD Aging Colony, and we thank the entire team, particularly Dr. Lu Lu, Jesse Ingels, Casey J Chapman, Melinda S McCarty, Megan K Mulligan, and Arthur Centeno. We thank Dr. Saunak Sen for his invaluable comments of the statistics. We thank the UTHSC-Rhodes College Population Health Researcher Program and Dr. Teresa Waters for support and for bringing an excellent summer student to the PI's lab. We thank Dr. Karolina A. Aberg and her team at the Virginia Commonwealth University for their advice on the MBD-seq technique.

References.

1. Marioni RE, Harris SE, Shah S, McRae AF, von Zglinicki T, Martin-Ruiz C, et al. The epigenetic clock and telomere length are independently associated with chronological age and mortality. International journal of epidemiology. 2018;47(1):356. Epub 2017/12/01. doi: 10.1093/ije/dyx233. PubMed PMID: 29190382; PubMed Central PMCID: PMCPMC5837660.
2. Mitteldorf J. An epigenetic clock controls aging. Biogerontology. 2016;17(1):257-65. Epub 2015/11/27. doi: 10.1007/s10522-015-9617-5. PubMed PMID: 26608516.
3. Lowe D, Horvath S, Raj K. Epigenetic clock analyses of cellular senescence and ageing. Oncotarget. 2016;7(8):8524-31. PubMed PMID: WOS:000375618100007.
4. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Molecular cell. 2013;49(2):359-67. doi: 10.1016/j.molcel.2012.10.016. PubMed PMID: 23177740; PubMed Central PMCID: PMC3780611.
5. Horvath S. DNA methylation age of human tissues and cell types. Genome biology. 2013;14(10):R115. doi: 10.1186/gb-2013-14-10-r115. PubMed PMID: 24138928; PubMed Central PMCID: PMC4015143.
6. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging. 2018;10(4):573-91. Epub 2018/04/21. doi: 10.18632/aging.101414. PubMed PMID: 29676998; PubMed Central PMCID: PMCPMC5940111.
7. Petkovich DA, Podolskiy DI, Lobanov AV, Lee SG, Miller RA, Gladyshev VN. Using DNA Methylation Profiling to Evaluate Biological Age and Longevity Interventions. Cell metabolism. 2017;25(4):954-60 e6. Epub 2017/04/06. doi: 10.1016/j.cmet.2017.03.016. PubMed PMID: 28380383; PubMed Central PMCID: PMCPMC5578459.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

790 8. Stubbs TM, Bonder MJ, Stark AK, Krueger F, Team BIAC, von Meyenn F, et al. Multi-
791 tissue DNA methylation age predictor in mouse. Genome biology. 2017;18(1):68. doi:
792 10.1186/s13059-017-1203-5. PubMed PMID: 28399939; PubMed Central PMCID: PMC5389178.
793 9. Thompson MJ, Chwialkowska K, Rubbi L, Lusis AJ, Davis RC, Srivastava A, et al. A multi-
794 tissue full lifespan epigenetic clock for mice. Aging. 2018;10(10):2832-54. Epub 2018/10/24.
795 doi: 10.18632/aging.101590. PubMed PMID: 30348905; PubMed Central PMCID:
796 PMCPMC6224226.
797 10. Wang T, Tsui B, Kreisberg JF, Robertson NA, Gross AM, Yu MK, et al. Epigenetic aging
798 signatures in mice livers are slowed by dwarfism, calorie restriction and rapamycin treatment.
799 Genome biology. 2017;18(1):57. Epub 2017/03/30. doi: 10.1186/s13059-017-1186-2. PubMed
800 PMID: 28351423; PubMed Central PMCID: PMCPMC5371228.
801 11. Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, et al. Aging of blood can be
802 tracked by DNA methylation changes at just three CpG sites. Genome biology. 2014;15(2):R24.
803 doi: 10.1186/gb-2014-15-2-r24. PubMed PMID: 24490752; PubMed Central PMCID:
804 PMC4053864.
805 12. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, et al. DNA methylation age
806 of blood predicts all-cause mortality in later life. Genome biology. 2015;16(1):25. doi:
807 10.1186/s13059-015-0584-6. PubMed PMID: 25633388; PubMed Central PMCID: PMC4350614.
808 13. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA
809 methylation-based measures of biological age: meta-analysis predicting time to death. Aging.
810 2016;8(9):1844-65. Epub 2016/10/01. doi: 10.18632/aging.101020. PubMed PMID: 27690265;
811 PubMed Central PMCID: PMCPMC5076441.
812 14. Zhang Y, Hapala J, Brenner H, Wagner W. Individual CpG sites that are associated with
813 age and life expectancy become hypomethylated upon aging. Clinical epigenetics. 2017;9. doi:
814 ARTN 9
815 10.1186/s13148-017-0315-9. PubMed PMID: WOS:000393925000001.
816 15. Christiansen L, Lenart A, Tan Q, Vaupel JW, Aviv A, McGue M, et al. DNA methylation
817 age is associated with mortality in a longitudinal Danish twin study. Aging cell. 2016;15(1):149-
818 54. Epub 2015/11/26. doi: 10.1111/acel.12421. PubMed PMID: 26594032; PubMed Central
819 PMCID: PMCPMC4717264.
820 16. Lin Q, Weidner CI, Costa IG, Marioni RE, Ferreira MRP, Deary IJ, et al. DNA methylation
821 levels at individual age-associated CpG sites can be indicative for life expectancy. Aging-Us.
822 2016;8(2):394-401. doi: DOI 10.18632/aging.100908. PubMed PMID: WOS:000372086600017.
823 17. Marioni RE, Shah S, McRae AF, Ritchie SJ, Muniz-Terrera G, Harris SE, et al. The
824 epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort
825 1936. International journal of epidemiology. 2015;44(4):1388-96. doi: 10.1093/ije/dyu277.
826 PubMed PMID: 25617346; PubMed Central PMCID: PMC4588858.
827 18. Perna L, Zhang Y, Mons U, Holleczek B, Saum KU, Brenner H. Epigenetic age acceleration
828 predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. Clinical
829 epigenetics. 2016;8:64. Epub 2016/06/09. doi: 10.1186/s13148-016-0228-z. PubMed PMID:
830 27274774; PubMed Central PMCID: PMCPMC4891876.
831 19. Lowe R, Barton C, Jenkins CA, Ernst C, Forman O, Fernandez-Twinn DS, et al. Ageing-
832 associated DNA methylation dynamics are a molecular readout of lifespan variation among

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

833 mammalian species. Genome biology. 2018;19(1):22. Epub 2018/02/18. doi: 10.1186/s13059-
834 018-1397-1. PubMed PMID: 29452591; PubMed Central PMCID: PMCPMC5815211.
835 20. Cole JJ, Robertson NA, Rather MI, Thomson JP, McBryan T, Sproul D, et al. Diverse
836 interventions that extend mouse lifespan suppress shared age-associated epigenetic changes at
837 critical gene regulatory regions. Genome biology. 2017;18. doi: ARTN 58
838 10.1186/s13059-017-1185-3. PubMed PMID: WOS:000397557000003.
839 21. Hahn O, Gronke S, Stubbs TM, Ficz G, Hendrich O, Krueger F, et al. Dietary restriction
840 protects from age-associated DNA methylation and induces epigenetic reprogramming of lipid
841 metabolism. Genome biology. 2017;18(1):56. Epub 2017/03/30. doi: 10.1186/s13059-017-
842 1187-1. PubMed PMID: 28351387; PubMed Central PMCID: PMCPMC5370449.
843 22. Sziraki A, Tyshkovskiy A, Gladyshev VN. Global remodeling of the mouse DNA
844 methylome during aging and in response to calorie restriction. Aging cell. 2018;17(3):e12738.
845 Epub 2018/03/27. doi: 10.1111/acel.12738. PubMed PMID: 29575528; PubMed Central PMCID:
846 PMCPMC5946071.
847 23. Hook M, Roy S, Williams EG, Bou Sleiman M, Mozhui K, Nelson JF, et al. Genetic
848 cartography of longevity in humans and mice: Current landscape and horizons. Biochimica et
849 biophysica acta. 2018. Epub 2018/02/08. doi: 10.1016/j.bbadis.2018.01.026. PubMed PMID:
850 29410319.
851 24. de Haan G, Gelman R, Watson A, Yunis E, Van Zant G. A putative gene causes variability
852 in lifespan among genotypically identical mice. Nature genetics. 1998;19(2):114-6. doi:
853 10.1038/465. PubMed PMID: 9620762.
854 25. De Haan G, Van Zant G. Genetic analysis of hemopoietic cell cycling in mice suggests its
855 involvement in organismal life span. FASEB journal : official publication of the Federation of
856 American Societies for Experimental Biology. 1999;13(6):707-13. PubMed PMID: 10094931.
857 26. Gelman R, Watson A, Bronson R, Yunis E. Murine chromosomal regions correlated with
858 longevity. Genetics. 1988;118(4):693-704. PubMed PMID: 3163317; PubMed Central PMCID:
859 PMC1203324.
860 27. Lang DH, Gerhard GS, Griffith JW, Vogler GP, Vandenbergh DJ, Blizard DA, et al.
861 Quantitative trait loci (QTL) analysis of longevity in C57BL/6J by DBA/2J (BXD) recombinant
862 inbred mice. Aging clinical and experimental research. 2010;22(1):8-19. PubMed PMID:
863 20305363.
864 28. Roy S, Sleiman MB, Jha P, Williams EG, Ingels JF, Chapman CJ, et al. Modulation of
865 longevity by diet, and youthful body weight, but not by weight gain after maturity. bioRxiv.
866 2019:776559. doi: 10.1101/776559.
867 29. Yuan R, Peters LL, Paigen B. Mice as a mammalian model for research on the genetics of
868 aging. ILAR journal / National Research Council, Institute of Laboratory Animal Resources.
869 2011;52(1):4-15. PubMed PMID: 21411853; PubMed Central PMCID: PMC3074346.
870 30. Yuan R, Tsaih SW, Petkova SB, Marin de Evsikova C, Xing S, Marion MA, et al. Aging in
871 inbred strains of mice: study design and interim report on median lifespans and circulating IGF1
872 levels. Aging cell. 2009;8(3):277-87. doi: 10.1111/j.1474-9726.2009.00478.x. PubMed PMID:
873 19627267; PubMed Central PMCID: PMC2768517.
874 31. Hsu HC, Li L, Zhang HG, Mountz JD. Genetic regulation of thymic involution. Mechanisms
875 of ageing and development. 2005;126(1):87-97. doi: 10.1016/j.mad.2004.09.016. PubMed
876 PMID: 15610766.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

877 32. Sloane LB, Stout JT, Austad SN, McClearn GE. Tail tendon break time: a biomarker of
878 aging? The journals of gerontology Series A, Biological sciences and medical sciences.
879 2011;66(3):287-94. doi: 10.1093/gerona/glq196. PubMed PMID: 21059835; PubMed Central
880 PMCID: PMC3041475.
881 33. Sloane LB, Stout JT, Vandenbergh DJ, Vogler GP, Gerhard GS, McClearn GE. Quantitative
882 trait loci analysis of tail tendon break time in mice of C57BL/6J and DBA/2J lineage. The journals
883 of gerontology Series A, Biological sciences and medical sciences. 2011;66(2):170-8. doi:
884 10.1093/gerona/glq169. PubMed PMID: 21047976; PubMed Central PMCID: PMC3021371.
885 34. Aberg KA, Xie L, Chan RF, Zhao M, Pandey AK, Kumar G, et al. Evaluation of Methyl-
886 Binding Domain Based Enrichment Approaches Revisited. PloS one. 2015;10(7):e0132205. doi:
887 10.1371/journal.pone.0132205. PubMed PMID: 26177298; PubMed Central PMCID:
888 PMC4503759.
889 35. Aberg KA, McClay JL, Nerella S, Xie LY, Clark SL, Hudson AD, et al. MBD-seq as a cost-
890 effective approach for methylome-wide association studies: demonstration in 1500 case--
891 control samples. Epigenomics. 2012;4(6):605-21. doi: 10.2217/epi.12.59. PubMed PMID:
892 23244307; PubMed Central PMCID: PMC3923085.
893 36. Aberg KA, Chan RF, Xie L, Shabalin AA, van den Oord E. Methyl-CpG-Binding Domain
894 Sequencing: MBD-seq. Methods in molecular biology. 2018;1708:171-89. Epub 2017/12/11.
895 doi: 10.1007/978-1-4939-7481-8_10. PubMed PMID: 29224145.
896 37. GeneNetwork 2. Available from: http://gn2.genenetwork.org.
897 38. Johansson A, Enroth S, Gyllensten U. Continuous Aging of the Human DNA Methylome
898 Throughout the Human Lifespan. PloS one. 2013;8(6):e67378. doi:
899 10.1371/journal.pone.0067378. PubMed PMID: 23826282; PubMed Central PMCID:
900 PMC3695075.
901 39. McClay JL, Aberg KA, Clark SL, Nerella S, Kumar G, Xie LY, et al. A methylome-wide study
902 of aging using massively parallel sequencing of the methyl-CpG-enriched genomic fraction from
903 blood in over 700 subjects. Human molecular genetics. 2014;23(5):1175-85. doi:
904 10.1093/hmg/ddt511. PubMed PMID: 24135035; PubMed Central PMCID: PMC3919012.
905 40. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of
906 ageing. Nature reviews Genetics. 2018;19(6):371-84. Epub 2018/04/13. doi: 10.1038/s41576-
907 018-0004-3. PubMed PMID: 29643443.
908 41. Field AE, Robertson NA, Wang T, Havas A, Ideker T, Adams PD. DNA Methylation Clocks
909 in Aging: Categories, Causes, and Consequences. Molecular cell. 2018;71(6):882-95. Epub
910 2018/09/23. doi: 10.1016/j.molcel.2018.08.008. PubMed PMID: 30241605.
911 42. Bacalini MG, Franceschi C, Gentilini D, Ravaioli F, Zhou X, Remondini D, et al. Molecular
912 Aging of Human Liver: An Epigenetic/Transcriptomic Signature. The journals of gerontology
913 Series A, Biological sciences and medical sciences. 2019;74(1):1-8. Epub 2018/03/20. doi:
914 10.1093/gerona/gly048. PubMed PMID: 29554203.
915 43. Snir S, Farrell C, Pellegrini M. Human epigenetic ageing is logarithmic with time across
916 the entire lifespan. Epigenetics : official journal of the DNA Methylation Society.
917 2019;14(9):912-26. doi: 10.1080/15592294.2019.1623634. PubMed PMID:
918 WOS:000479080700005.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

44. Ciccarone F, Tagliatesta S, Caiafa P, Zampieri M. DNA methylation dynamics in aging: how far are we from understanding the mechanisms? Mechanisms of ageing and development. 2018;174:3-17. Epub 2017/12/23. doi: 10.1016/j.mad.2017.12.002. PubMed PMID: 29268958.
45. Zampieri M, Ciccarone F, Calabrese R, Franceschi C, Burkle A, Caiafa P. Reconfiguration of DNA methylation in aging. Mechanisms of ageing and development. 2015;151:60-70. doi: 10.1016/j.mad.2015.02.002. PubMed PMID: 25708826.
46. Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, et al. Distinct DNA methylomes of newborns and centenarians. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(26):10522-7. doi: 10.1073/pnas.1120658109. PubMed PMID: 22689993; PubMed Central PMCID: PMC3387108.
47. Maegawa S, Hinkal G, Kim HS, Shen L, Zhang L, Zhang J, et al. Widespread and tissue specific age-related DNA methylation changes in mice. Genome research. 2010;20(3):332-40. doi: 10.1101/gr.096826.109. PubMed PMID: 20107151; PubMed Central PMCID: PMC2840983.
48. Unnikrishnan A, Hadad N, Masser DR, Jackson J, Freeman WM, Richardson A. Revisiting the genomic hypomethylation hypothesis of aging. Annals of the New York Academy of Sciences. 2018;1418(1):69-79. Epub 2018/01/25. doi: 10.1111/nyas.13533. PubMed PMID: 29363785.
49. Singhal RP, Mays-Hoopes LL, Eichhorn GL. DNA methylation in aging of mice. Mechanisms of ageing and development. 1987;41(3):199-210. Epub 1987/12/01. doi: 10.1016/0047-6374(87)90040-6. PubMed PMID: 3431172.
50. Cruickshanks HA, McBryan T, Nelson DM, Vanderkraats ND, Shah PP, van Tuyn J, et al. Senescent cells harbour features of the cancer epigenome. Nat Cell Biol. 2013;15(12):1495-506. Epub 2013/11/26. doi: 10.1038/ncb2879. PubMed PMID: 24270890; PubMed Central PMCID: PMCPMC4106249.
51. Day K, Waite LL, Thalacker-Mercer A, West A, Bamman MM, Brooks JD, et al. Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. Genome biology. 2013;14(9):R102. Epub 2013/09/17. doi: 10.1186/gb-2013-14-9-r102. PubMed PMID: 24034465; PubMed Central PMCID: PMCPMC4053985.
52. Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome research. 2010;20(4):434-9. Epub 2010/03/12. doi: 10.1101/gr.103101.109. PubMed PMID: 20219945; PubMed Central PMCID: PMCPMC2847746.
53. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome research. 2010;20(4):440-6. doi: 10.1101/gr.103606.109. PubMed PMID: 20219944; PubMed Central PMCID: PMC2847747.
54. Carter CJ. Convergence of genes implicated in Alzheimer's disease on the cerebral cholesterol shuttle: APP, cholesterol, lipoproteins, and atherosclerosis. Neurochem Int. 2007;50(1):12-38. Epub 2006/09/16. doi: 10.1016/j.neuint.2006.07.007. PubMed PMID: 16973241.
55. Steinberg S, Stefansson H, Jonsson T, Johannsdottir H, Ingason A, Helgason H, et al. Loss-of-function variants in ABCA7 confer risk of Alzheimer's disease. Nature genetics. 2015;47(5):445-7. Epub 2015/03/26. doi: 10.1038/ng.3246. PubMed PMID: 25807283.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

963 56. Sebastiani P, Montano M, Puca A, Solovieff N, Kojima T, Wang MC, et al. RNA editing
964 genes associated with extreme old age in humans and with lifespan in C. elegans. PloS one.
965 2009;4(12):e8210. Epub 2009/12/17. doi: 10.1371/journal.pone.0008210. PubMed PMID:
966 20011587; PubMed Central PMCID: PMCPMC2788130.
967 57. Yashin AI, Wu D, Arbeev KG, Ukraintseva SV. Joint influence of small-effect genetic
968 variants on human longevity. Aging. 2010;2(9):612-20. Epub 2010/09/14. doi:
969 10.18632/aging.100191. PubMed PMID: 20834067; PubMed Central PMCID: PMCPMC2984609.
970 58. Soerensen M, Dato S, Tan Q, Thinggaard M, Kleindorp R, Beekman M, et al. Human
971 longevity and variation in GH/IGF-1/insulin signaling, DNA damage signaling and repair and
972 pro/antioxidant pathway genes: cross sectional and longitudinal studies. Experimental
973 gerontology. 2012;47(5):379-87. Epub 2012/03/13. doi: 10.1016/j.exger.2012.02.010. PubMed
974 PMID: 22406557; PubMed Central PMCID: PMCPMC3334439.
975 59. Dato S, Soerensen M, Montesanto A, Lagani V, Passarino G, Christensen K, et al. UCP3
976 polymorphisms, hand grip performance and survival at old age: association analysis in two
977 Danish middle aged and elderly cohorts. Mechanisms of ageing and development.
978 2012;133(8):530-7. Epub 2012/06/30. doi: 10.1016/j.mad.2012.06.004. PubMed PMID:
979 22743239; PubMed Central PMCID: PMCPMC3629378.
980 60. Slack C, Alic N, Foley A, Cabecinha M, Hoddinott MP, Partridge L. The Ras-Erk-ETS-
981 Signaling Pathway Is a Drug Target for Longevity. Cell. 2015;162(1):72-83. Epub 2015/06/30.
982 doi: 10.1016/j.cell.2015.06.023. PubMed PMID: 26119340; PubMed Central PMCID:
983 PMCPMC4518474.
984 61. Passtoors WM, van den Akker EB, Deelen J, Maier AB, van der Breggen R, Jansen R, et al.
985 IL7R gene expression network associates with human healthy ageing. Immunity & ageing : I & A.
986 2015;12:21. Epub 2015/11/14. doi: 10.1186/s12979-015-0048-6. PubMed PMID: 26566388;
987 PubMed Central PMCID: PMCPMC4642670.
988 62. Zhang G, Li J, Purkayastha S, Tang Y, Zhang H, Yin Y, et al. Hypothalamic programming of
989 systemic ageing involving IKK-beta, NF-kappaB and GnRH. Nature. 2013;497(7448):211-6. Epub
990 2013/05/03. doi: 10.1038/nature12143. PubMed PMID: 23636330; PubMed Central PMCID:
991 PMCPMC3756938.
992 63. Mozhui K, Pandey AK. Conserved effect of aging on DNA methylation and association
993 with EZH2 polycomb protein in mice and humans. Mechanisms of ageing and development.
994 2017;162:27-37. Epub 2017/03/03. doi: 10.1016/j.mad.2017.02.006. PubMed PMID: 28249716;
995 PubMed Central PMCID: PMCPMC5411177.
996 64. Huang Y, Yan J, Hou J, Fu X, Li L, Hou Y. Developing a DNA methylation assay for human
997 age prediction in blood and bloodstain. Forensic Sci Int Genet. 2015;17:129-36. Epub
998 2015/05/17. doi: 10.1016/j.fsigen.2015.05.007. PubMed PMID: 25979242.
999 65. Zbiec-Piekarska R, Spolnicka M, Kupiec T, Parys-Proszek A, Makowska Z, Paleczka A, et
1000 al. Development of a forensically useful age prediction method based on DNA methylation
1001 analysis. Forensic Sci Int Genet. 2015;17:173-9. Epub 2015/06/01. doi:
1002 10.1016/j.fsigen.2015.05.001. PubMed PMID: 26026729.
1003 66. Zhang W, Spector TD, Deloukas P, Bell JT, Engelhardt BE. Predicting genome-wide DNA
1004 methylation using methylation marks, genomic position, and DNA regulatory elements.
1005 Genome biology. 2015;16:14. doi: 10.1186/s13059-015-0581-9. PubMed PMID: 25616342;
1006 PubMed Central PMCID: PMC4389802.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

67. Lovkvist C, Dodd IB, Sneppen K, Haerter JO. DNA methylation in human epigenomes depends on local topology of CpG sites. Nucleic acids research. 2016;44(11):5123-32. doi: 10.1093/nar/gkw124. PubMed PMID: 26932361.
68. Li Y, Zhu J, Tian G, Li N, Li Q, Ye M, et al. The DNA methylome of human peripheral blood mononuclear cells. PLoS biology. 2010;8(11):e1000533. doi: 10.1371/journal.pbio.1000533. PubMed PMID: 21085693; PubMed Central PMCID: PMC2976721.
69. Janssens GE, Lin XX, Millan-Arino L, Kavsek A, Sen I, Seinstra RI, et al. Transcriptomics-Based Screening Identifies Pharmacological Inhibition of Hsp90 as a Means to Defer Aging. Cell Rep. 2019;27(2):467-80 e6. Epub 2019/04/11. doi: 10.1016/j.celrep.2019.03.044. PubMed PMID: 30970250; PubMed Central PMCID: PMCPMC6459000.
70. Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, et al. The transcriptional landscape of age in human peripheral blood. Nature communications. 2015;6:8570. Epub 2015/10/23. doi: 10.1038/ncomms9570. PubMed PMID: 26490707; PubMed Central PMCID: PMCPMC4639797.
71. Svane AM, Soerensen M, Lund J, Tan Q, Jylhava J, Wang Y, et al. DNA Methylation and All-Cause Mortality in Middle-Aged and Elderly Danish Twins. Genes (Basel). 2018;9(2). Epub 2018/02/09. doi: 10.3390/genes9020078. PubMed PMID: 29419728; PubMed Central PMCID: PMCPMC5852574.
72. Zhang Y, Wilson R, Heiss J, Breitling LP, Saum KU, Schottker B, et al. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. Nature communications. 2017;8:14617. Epub 2017/03/18. doi: 10.1038/ncomms14617. PubMed PMID: 28303888; PubMed Central PMCID: PMCPMC5357865.
73. Lund JB, Li S, Baumbach J, Svane AM, Hjelmborg J, Christiansen L, et al. DNA methylome profiling of all-cause mortality in comparison with age-associated methylation patterns. Clinical epigenetics. 2019;11(1):23. Epub 2019/02/10. doi: 10.1186/s13148-019-0622-4. PubMed PMID: 30736859; PubMed Central PMCID: PMCPMC6368749.
74. Bojesen SE, Timpson N, Relton C, Davey Smith G, Nordestgaard BG. AHRR (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. Thorax. 2017;72(7):646-53. Epub 2017/01/20. doi: 10.1136/thoraxjnl-2016-208789. PubMed PMID: 28100713; PubMed Central PMCID: PMCPMC5520281.
75. McCartney DL, Hillary RF, Stevenson AJ, Ritchie SJ, Walker RM, Zhang Q, et al. Epigenetic prediction of complex traits and death. Genome biology. 2018;19(1):136. Epub 2018/09/28. doi: 10.1186/s13059-018-1514-1. PubMed PMID: 30257690; PubMed Central PMCID: PMCPMC6158884.
76. Albanes D, Winick M. Are cell number and cell proliferation risk factors for cancer? J Natl Cancer Inst. 1988;80(10):772-4. Epub 1988/07/20. doi: 10.1093/jnci/80.10.772. PubMed PMID: 3385783.
77. Bartke A. Healthy aging: is smaller better? - a mini-review. Gerontology. 2012;58(4):337-43. Epub 2012/01/21. doi: 10.1159/000335166. PubMed PMID: 22261798; PubMed Central PMCID: PMCPMC3893695.
78. Global BMIMC, Di Angelantonio E, Bhupathiraju Sh N, Wormser D, Gao P, Kaptoge S, et al. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. Lancet. 2016;388(10046):776-86. Epub 2016/07/18. doi:

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

10.1016/S0140-6736(16)30175-1. PubMed PMID: 27423262; PubMed Central PMCID: PMCPMC4995441.

79. Lemez S, Wattie N, Baker J. Do "big guys" really die younger? An examination of height and lifespan in former professional basketball players. PloS one. 2017;12(10). doi: ARTN e0185617 10.1371/journal.pone.0185617. PubMed PMID: WOS:000412029600029.

80. McCay CM, Crowell MF, Maynard LA. The effect of retarded growth upon the length of life span and upon the ultimate body size. 1935. Nutrition. 1989;5(3):155-71; discussion 72. Epub 1989/05/01. PubMed PMID: 2520283.

81. Nunney L. Size matters: height, cell number and a person's risk of cancer. P Roy Soc B-Biol Sci. 2018;285(1889). doi: ARTN 20181743 10.1098/rspb.2018.1743. PubMed PMID: WOS:000448380800012.

82. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017;541(7635):81-6. Epub 2016/12/22. doi: 10.1038/nature20784. PubMed PMID: 28002404; PubMed Central PMCID: PMCPMC5570525.

83. Barfield RT, Almli LM, Kilaru V, Smith AK, Mercer KB, Duncan R, et al. Accounting for population stratification in DNA methylation studies. Genet Epidemiol. 2014;38(3):231-41. Epub 2014/01/31. doi: 10.1002/gepi.21789. PubMed PMID: 24478250; PubMed Central PMCID: PMCPMC4090102.

84. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nature reviews Genetics. 2011;12(8):529-41. Epub 2011/07/13. doi: 10.1038/nrg3000. PubMed PMID: 21747404; PubMed Central PMCID: PMCPMC3508712.

85. Lappalainen T, Greally JM. Associating cellular epigenetic models with human phenotypes. Nature reviews Genetics. 2017;18(7):441-51. Epub 2017/05/31. doi: 10.1038/nrg.2017.32. PubMed PMID: 28555657.

86. FastQC. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

87. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology. 2009;10(3):R25. doi: 10.1186/gb-2009-10-3-r25. PubMed PMID: 19261174; PubMed Central PMCID: PMC2690996.

88. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9. doi: 10.1093/bioinformatics/btp352. PubMed PMID: 19505943; PubMed Central PMCID: PMC2723002.

89. Lassmann T, Hayashizaki Y, Daub CO. SAMStat: monitoring biases in next generation sequencing data. Bioinformatics. 2011;27(1):130-1. Epub 2010/11/20. doi: 10.1093/bioinformatics/btq614. PubMed PMID: 21088025; PubMed Central PMCID: PMCPMC3008642.

90. Lienhard M, Grimm C, Morkel M, Herwig R, Chavez L. MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. Bioinformatics. 2014;30(2):284-6. doi: 10.1093/bioinformatics/btt650. PubMed PMID: 24227674; PubMed Central PMCID: PMC3892689.

91. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40. doi:

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

10.1093/bioinformatics/btp616. PubMed PMID: 19910308; PubMed Central PMCID: PMC2796818.

92. Ashbrook DG, Arends D, Mulligan MK, Williams EG, Lutz C, Valenzuela A, et al. Sequencing the BXD family, a cohort for experimental systems genetics and precision medicine. IBANGS MEETING: THE 20TH ANNUAL GENES, BRAIN & BEHAVIOR MEETING Rochester, MN, USA. 2018.

93. Ashbrook DG, Arends D, Prins P, Mulligan MK, Roy S, Williams EG, et al. The expanded BXD family of mice: A cohort for experimental systems genetics and precision medicine. bioRxiv. 2019:672097. doi: 10.1101/672097.

94. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell. 2010;38(4):576-89. Epub 2010/06/02. doi: 10.1016/j.molcel.2010.05.004. PubMed PMID: 20513432; PubMed Central PMCID: PMCPMC2898526.

95. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. Nucleic acids research. 2019;47(W1):W199-W205. Epub 2019/05/23. doi: 10.1093/nar/gkz401. PubMed PMID: 31114916; PubMed Central PMCID: PMCPMC6602449.

96. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic acids research. 2005;33(Web Server issue):W741-8. Epub 2005/06/28. doi: 10.1093/nar/gki475. PubMed PMID: 15980575; PubMed Central PMCID: PMCPMC1160236.

97. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. . R package version 11-7, http://CRANR-projectorg/package=lme4.

98. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33(1):1-22. Epub 2010/09/03. PubMed PMID: 20808728; PubMed Central PMCID: PMCPMC2929880.

99. Zou H, Hastie T. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2005;67(2):301-20. doi: 10.1111/j.1467-9868.2005.00503.x.

100. NIA Aging BXD CD+HFD Liver RNA-Seq (Aug18) Log2.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure legends.

Fig. 1. Age distribution and body weight characteristics.

(a) Each point depicts a mouse used for methylome assay. The purpose of this plot is to show a near uniform distribution of ages (y-axis) across the three lifespan groups (categorical classification on x-axis; short-lived with strain mean lifespan < 600 days, medium-lived: 640–750 days, and long-lived: lifespan > 800 days).

(b) The bar plots show significant strain variation in mean body weight at young adulthood (baseline weight) and at final weighing. By final weight, mice on high fat diet (HFD) had gained significant weight. The weight of the liver did not differ significantly between strains and showed only a slight gain for mice on HFD. These graphs were plotted using all 70 samples; excluding the three male cases in BXD102, B6D2F1, and D2B6F1, did not alter the strain distribution and the F1s still had robust body weights.

(c) At young adulthood, body weight was significantly correlated with age of mice. The baseline body weight was also correlated with final weight of body **(d)** and liver **(e)**.

Fig. 2. Global features of the methylome

(a) Scatter plot between the top 2 principal components—PC1 (19% of variance) and PC2 (13% of variance)—show a strong population structure with mice clustering by strain identity (color coded). Members of sub-strains also cluster in close proximity. For strains with animals from both standard chow (CD; solid circles) and high fat diet (HFD; squares), mice on HFD co-cluster with the CD mice.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

For each individual mouse, the overall mean methylation and within-individual variance was calculated for 200,531 genic CpG regions located with annotated genes **(b)**, and 167,769 CpG regions in intergenic sites **(c)**. The intergenic regions have wide variation between strains and the F1 hybrids have the highest mean methylation and lowest variance. The genic CpG regions are more consistent across strains. Scatter plots on the right show the correlation between methylation averages and variance in the 70 samples. Mean methylation is inversely correlated with variance, and this is particularly pronounced for the intergenic CpG regions. Average methylation at intergenic regions (x-axis) is correlated with PC1 **(d)**, and PC3 **(e)**, and average methylation at genic regions is correlated with PC4 **(f)**.

Fig. 3. Intercorrelation between body weight at young adulthood, the methylome, and strain longevity

Body weight at young adulthood has weak but significant correlations with **(a)** the methylome top principal component, PC1, and the **(b)** mean methylation and **(c)** variance at genic CpG regions. Given the timeline, the results indicate that the body weight at earlier time is predictive of DNA methylation (solid arrow). However, since both are genetically modulated phenotypes and the methylome may also have had a sustained effect on the body weight of mice (dashed arrow), the direction of causality cannot be clearly resolved, and we consider these as interdependent phenotypes.

(d) The methylome in turn may be predictive of lifespan (solid arrow), and PC4, a correlate of mean methylation at genic CpG regions, is strongly correlated with lifespan.

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1174 **(e)** Baseline body weight is also predictive of strain longevity (solid arrow), and the negative
1175 correlation conveys reduced lifespan for mice that had heavier body weight at young
1176 adulthood.
1177
1178 **Fig. 4. Features of age dependent differentially methylated CpGs regions (age-DMRs)**
1179 **(a)** Histogram of observed *p*-values for age dependent methylation shows that aging has a
1180 strong effect on few CpG regions. **(b)** The *p*-values for the effect of diet showed a null
1181 distribution, likely because of the small number of samples from mice kept on high fat diet
1182 **(c)** Each point in the Manhattan plot depicts the location of a CpG region (x-axis: autosomal
1183 chromosomes 1 to 19; and chromosome X as 20), and the $-\log_{10}p$ for age effect (y-axis). The
1184 genome-wide significant threshold was set at $-\log_{10}$(2.6e-7) (red line; 10% Bonferroni threshold
1185 for 368,300 tests) and the suggestive threshold at $-\log_{10}$(3.6e-4) (blue line). **(d)** The top 500 age-
1186 DMRs consisted of 206 regions that gained methylation with age (age-hypermethylated;
1187 positive regression coefficient for age), and 294 regions that were age-hypomethylated
1188 (negative regression coefficient for age). The bar-plots display the percent total of genomic
1189 features in the age-hypermethylated (burgundy) and age-hypomethylated (sandy brown) sets,
1190 relative to the set of 369,300 MBD-seq bins (grey), and the full genome-wide background set
1191 (black). Within the bins, the regression coefficients for age (i.e., change in DNA methylation per
1192 unit change in age in days, $\log_{10}$ scale) were dependent on the **(e)** CpG density, and **(f)** mean
1193 methylation.
1194
1195 **Fig. 5. Elastic net regression DNA methylation and mRNA clocks**

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The y-axis shows the chronological age of mice, and the x-axis shows the DNAmAges predicted by the 60 clock CpG regions in the **(a)** training set of 36 samples, and **(b)** the testing set of 33 samples.

For the transcriptomic data, the training in 150 RNA-seq samples (batch 1) identified 65 transcripts for age estimation. The plots compare the chronological ages to the predicted ages in the **(c)** training set of 150 samples, and **(d)** in the testing testing of 140 samples (batch 2). To complement this, the training was also carried out in the 140 samples (batch 2) and this identified 62 transcripts for age estimation. Correlations between chronological age and estimated age in the batch 2 training set (**e**), and batch 1 testing set (**f**) are shown.

Fig. 6. DNA methylation based epigenetic clocks derived from age-DMRs

The epigenetic age of mice was estimated from DNA methylation levels at the 500 age-DMRs. **(a)** The estimated DNAmAge (y-axis) was correlated with chronological age (x-axis) of mice. The age acceleration residuals (DNAmAge-acc) derived from this clock **(b)** did not correlate with strain lifespan, **(c)** but was positively correlated with body weight. **(d)** For strains with samples from both standard chow (control diet or CD) and high fat diet (HFD), the DNAmAge-acc did not differentiate between the diet groups.

A second version of the clock was made using 56 age-DMRs that were also associated with lifespan variation. **(e)** The DNAmAge derived for the 56 age-lifespan-DMRs correlated with chronological age on mice; however, the relation deviated from linearity and appeared to plateau for the older mice. The DNAmAge-acc from this clock was **(f)** inversely correlated with strain lifespan that indicates decelerated aging in long-lived strains, and **(g)** positively correlated

bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with baseline body weight that suggests a more accelerated clock for mice with higher body weight at younger age. **(h)** This clock also indicates greater age acceleration for the HFD group compared to the CD group (mean of 21 ± 36 for HFD, -26 ± 57 for CD; p = 0.01; n = 33).

Fig. 7. Lifespan predictive age-DMRs

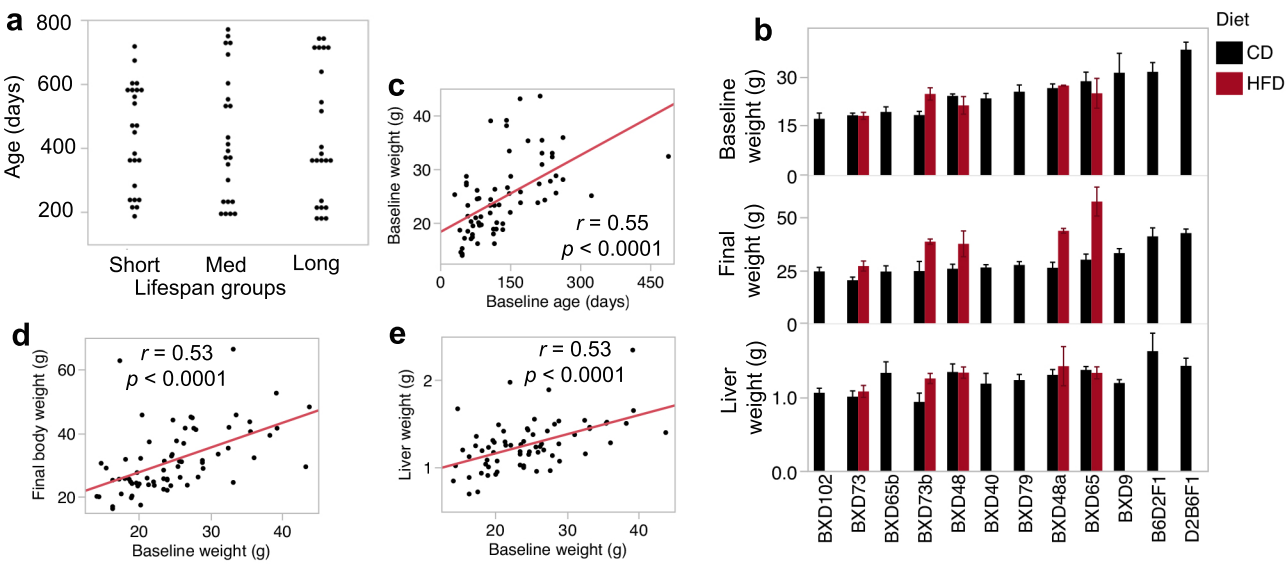
(a) Comparison between the regression coefficients for age (x-axis) vs. the regression coefficient for strain median lifespan (y-axis) shows that most of the age-DMRs that are also associated with lifespan have generally lower methylation levels in strains with longer median lifespan, as indicated by the negative regression coefficient for lifespan. As examples of age and lifespan associated CpG regions, the change in methylation over time along with cross-sectional variation is illustrated by the age-hypomethylated CpG region in **(b)** *Casz1*, and by the age-hypermethylated CpG regions in **(c)** *Cyp46a1* and **(d)** *Abca7*. Each point represents a mouse plotted by age (x-axis) and methylation level (y-axis, logRPKM). The lines represent linear regression lines for each strain-by-diet, classified by strain lifespan phenotype as short- (red), medium (yellow) or long-lived (blue). In all three CpG regions, cross-sectional comparisons show that mice belonging to short-lived groups have higher average methylation relative to mice belong to medium and long-lived groups (box plots). *Casz1* mean methylation log RPKM in short-lived = 3.96 ± 0.28, medium-lived = 3.62 ± 0.39, long-lived = 3.28 ± 0.42 (ANOVA p < 0.0001. *Cyp46a1* mean methylation log RPKM in short-lived = 3.34 ± 0.69, medium-lived = 3.08 ± 0.85, long-lived = 2.51 ± 0.83 (p = 0.0025. *Abca7* mean methylation log RPKM in short-lived = 3.85 ± 0.41, medium-lived = 3.75 ± 0.33, long-lived = 3.58 ± 0.32 (p = 0.04).

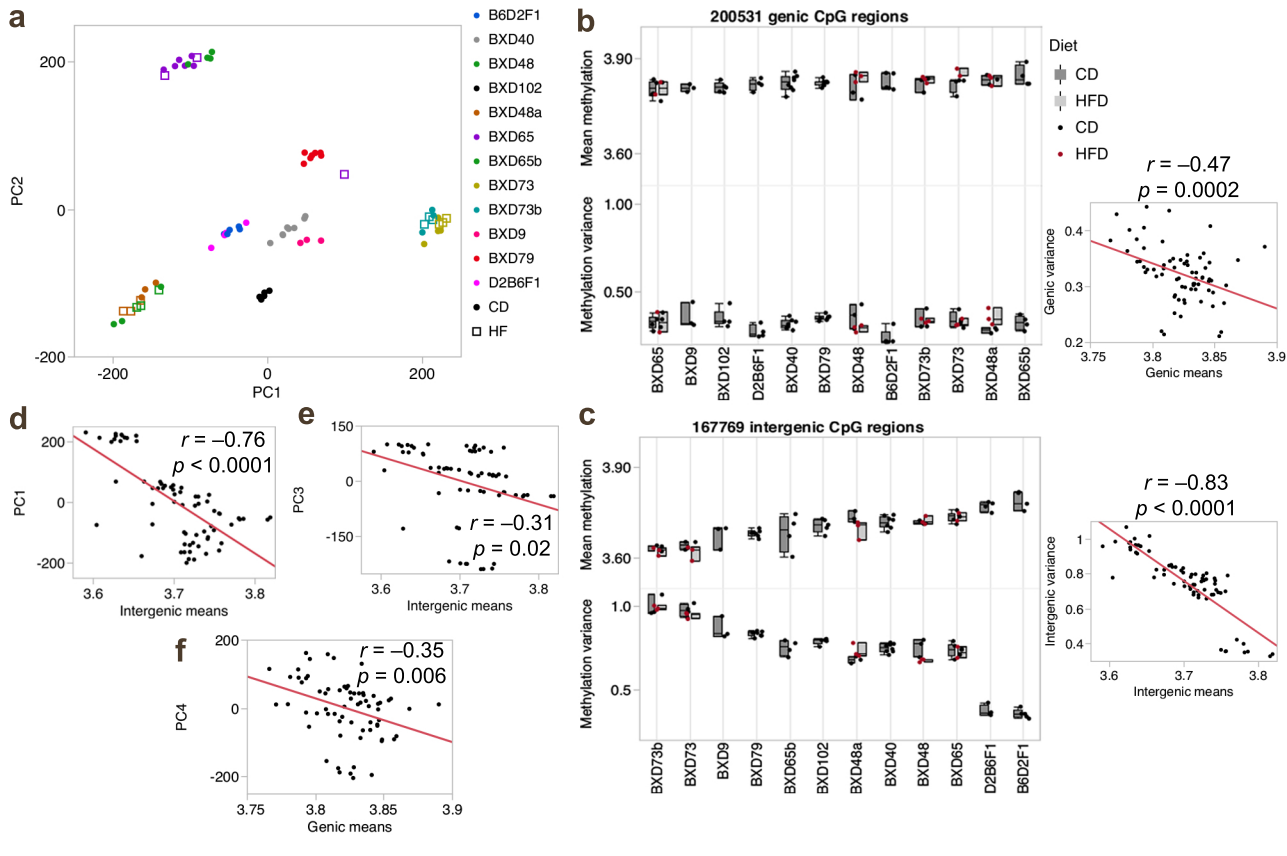
bioRxiv preprint doi: https://doi.org/10.1101/791582; this version posted October 3, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

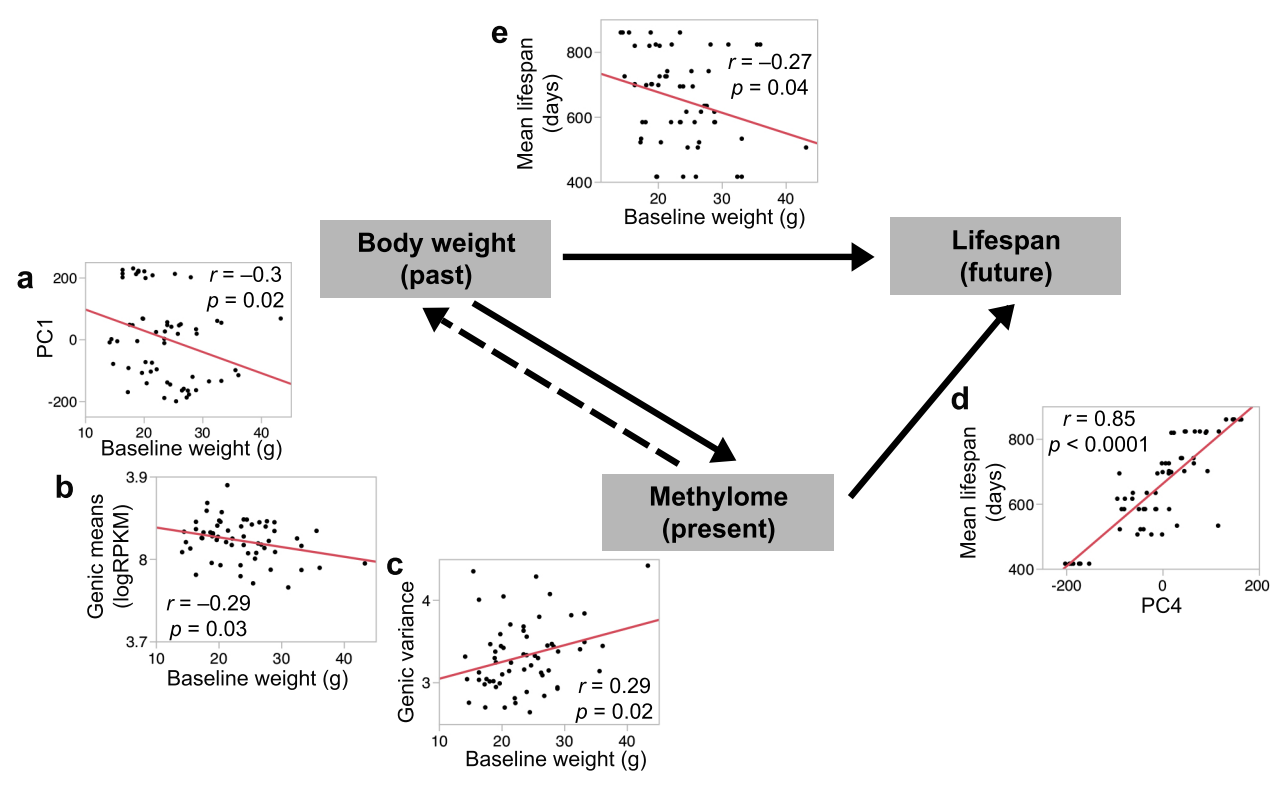
Fig. 8. Epigenetic clock derived from body weight associated CpG regions

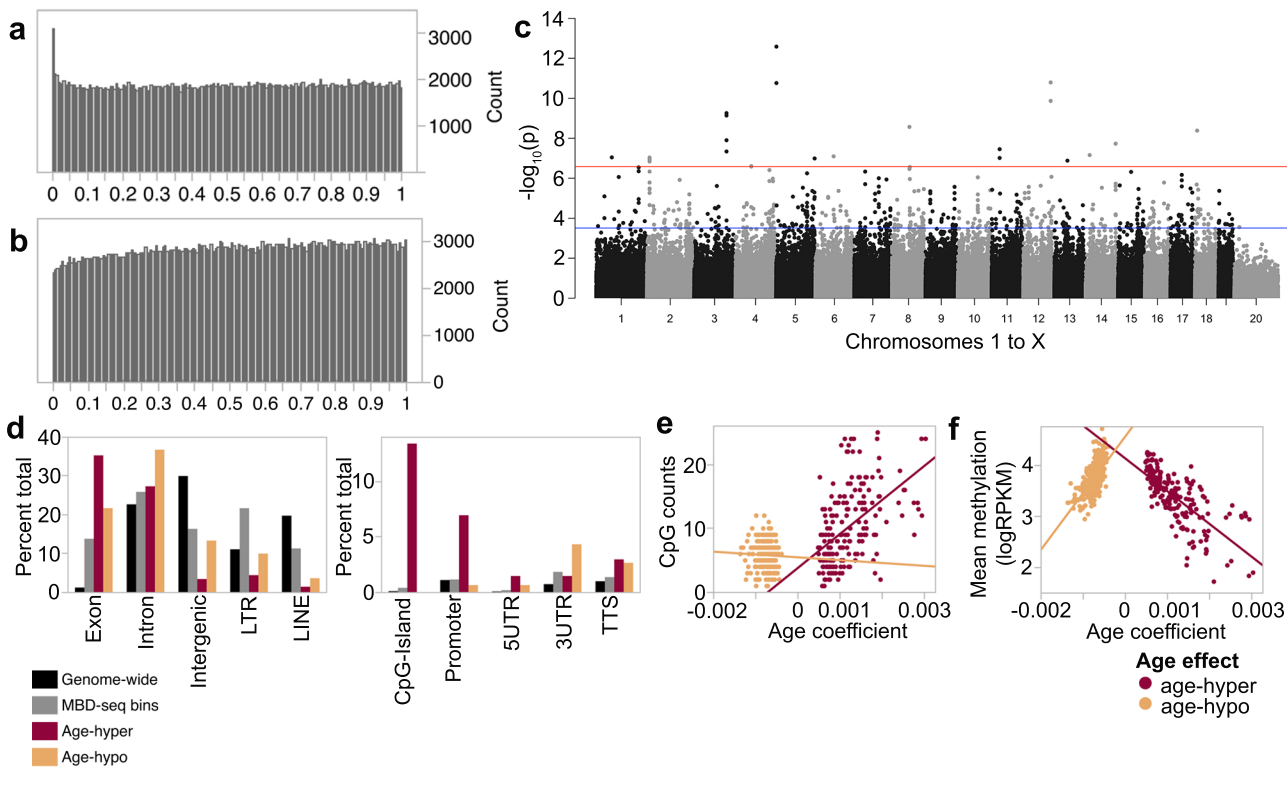
(a) The Manhattan plot for baseline body weight (BW0) shows widespread association with DNA methylation. The genome-wide significant threshold is set at -$\log_{10}$(2.6e-7) (red line; 10% Bonferroni threshold for 368,300 tests).

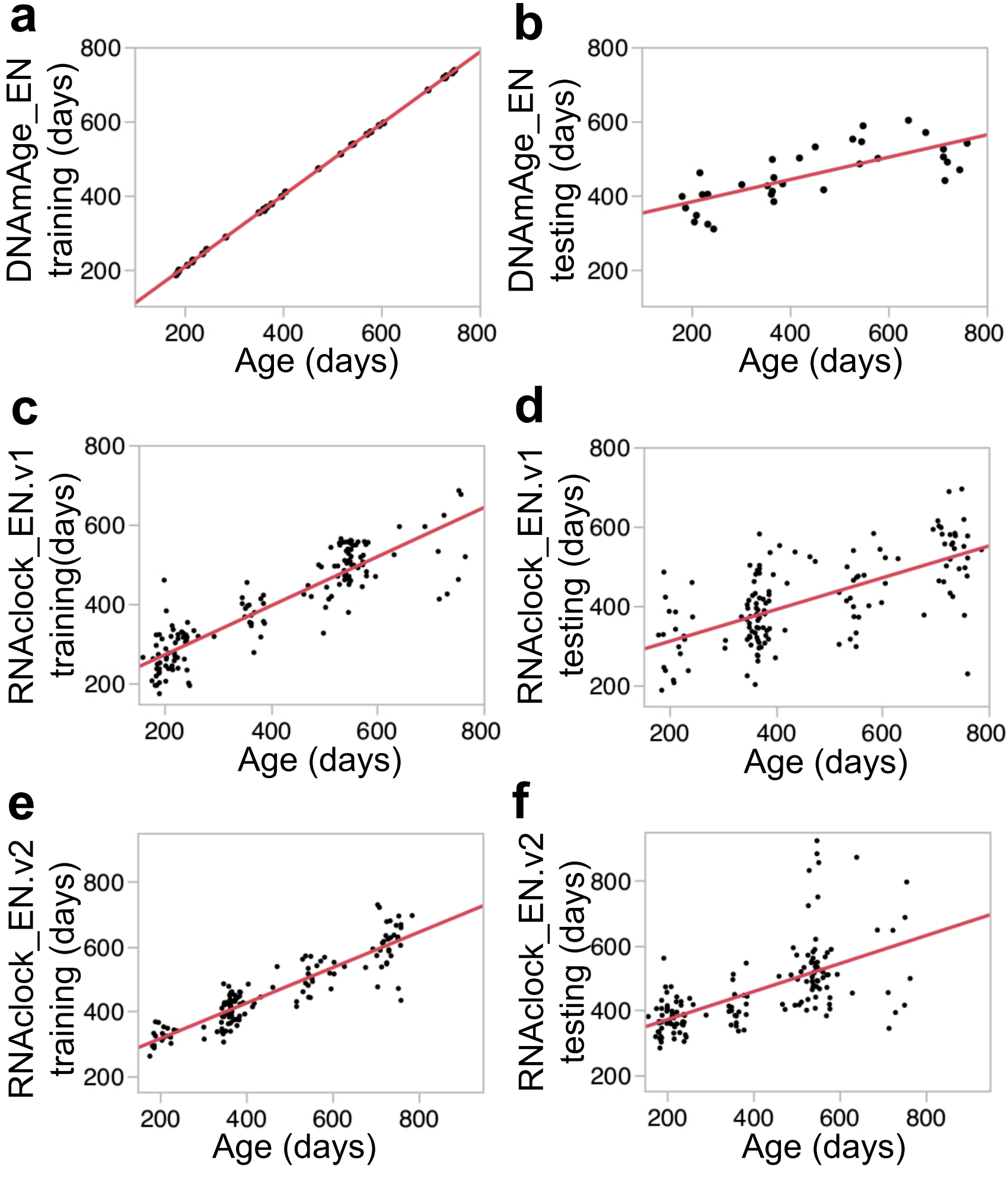
The epigenetic clock was estimated from DNA methylation levels at the top 500 BW0 associated CpG regions. **(b)** We refer to this version of the clock as BW0.lifespan as the estimates (y-axis) has a significant positive correlation with the expected lifespan of strains (x-axis). However, for the long-lived F1 hybrids and BXD65 cases that also had higher body weight, the estimates were shorter than the recorded lifespans (cluster A has B6D2F1 and BXD65, cluster B has D2B6F1). **(c)** Baseline body weight (x-axis) is negatively correlated with the BW0.lifespan estimates. **(d)** Age-acceleration from the 56 lifespan-age-DMR clock (x-axis) was negatively correlated with the BW0.lifespan estimated with more accelerated epigenetic aging for shorter-lived BXDs.

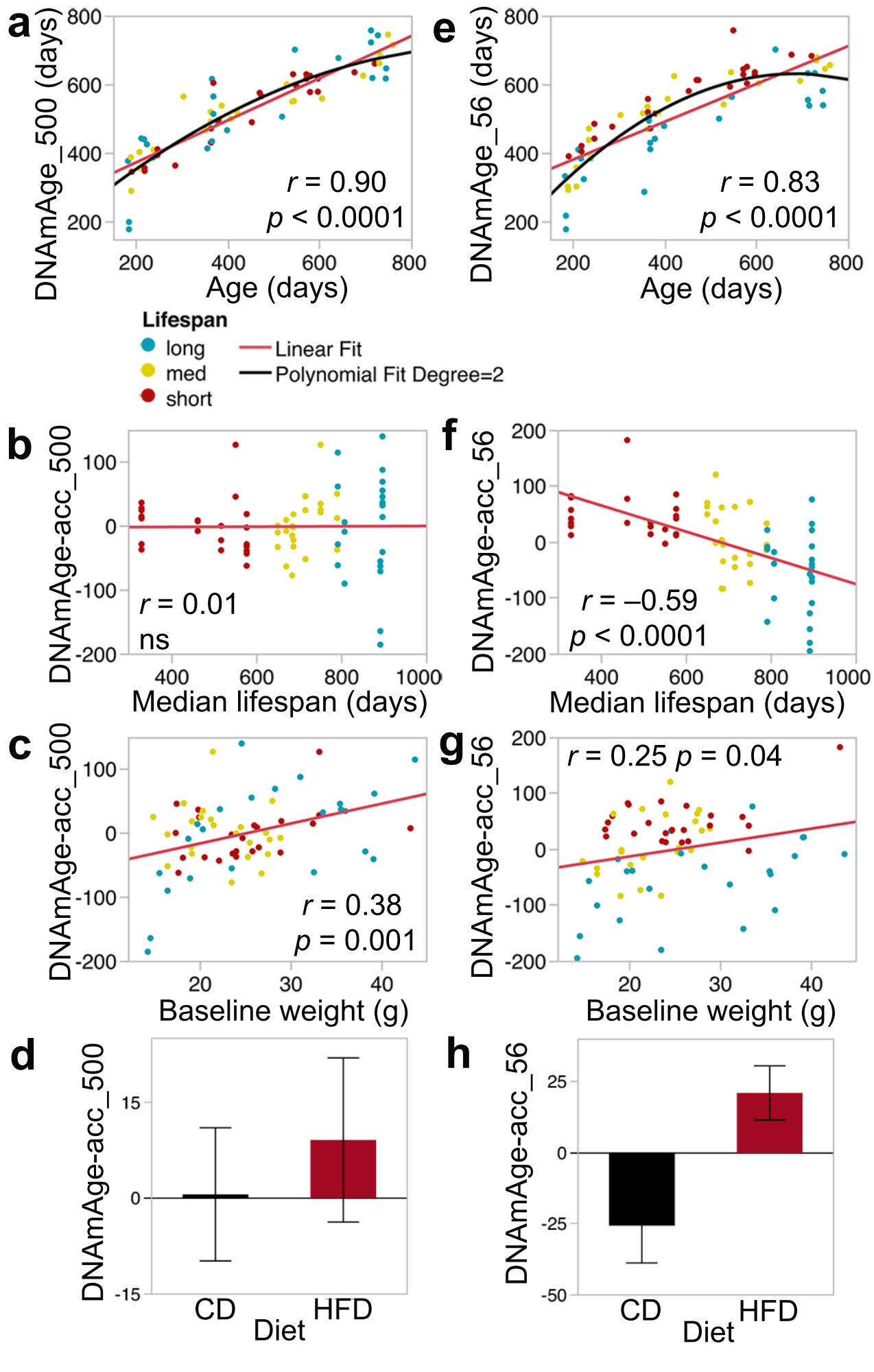

a
Age (days)
800
600
400
200
Short
Med
Long
Lifespan groups
c
Baseline weight (g)
r = 0.55
p < 0.0001
Baseline age (days)
b
Diet
CD
HFD
Baseline weight (g)
Final weight (g)
Liver weight (g)
BXD102
BXD73
BXD65b
BXD73b
BXD48
BXD40
BXD79
BXD48a
BXD65
BXD9
B6D2F1
D2B6F1
d
Final body weight (g)
r = 0.53
p < 0.0001
Baseline weight (g)
e
Liver weight (g)
r = 0.53
p < 0.0001
Baseline weight (g)

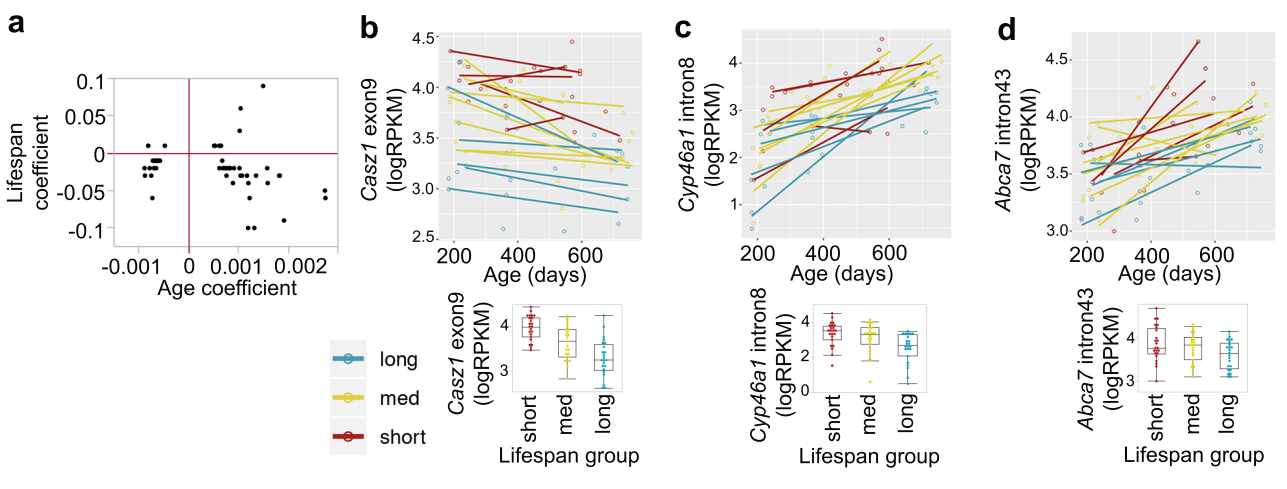

a
PC2
PC1
200
0
-200
-200
0
200
B6D2F1
BXD40
BXD48
BXD102
BXD48a
BXD65
BXD65b
BXD73
BXD73b
BXD9
BXD79
D2B6F1
CD
HF
b
200531 genic CpG regions
Mean methylation
Methylation variance
3.90
3.60
1.00
0.50
Diet
CD
HFD
CD
HFD
BXD65
BXD9
BXD102
D2B6F1
BXD40
BXD79
BXD48
B6D2F1
BXD73b
BXD73
BXD48a
BXD65b
r = –0.47
p = 0.0002
Genic variance
Genic means
0.4
0.3
0.2
3.75
3.8
3.85
3.9
c
167769 intergenic CpG regions
Mean methylation
Methylation variance
3.90
3.60
1.0
0.5
BXD73b
BXD73
BXD9
BXD79
BXD65b
BXD102
BXD48a
BXD40
BXD48
BXD65
D2B6F1
B6D2F1
r = –0.83
p < 0.0001
Intergenic variance
Intergenic means
1
0.8
0.6
0.4
3.6
3.7
3.8
d
PC1
200
0
-200
r = –0.76
p < 0.0001
3.6
3.7
3.8
Intergenic means
e
PC3
150
0
-150
r = –0.31
p = 0.02
3.6
3.7
3.8
Intergenic means
f
PC4
200
0
-200
r = –0.35
p = 0.006
3.75
3.8
3.85
3.9
Genic means

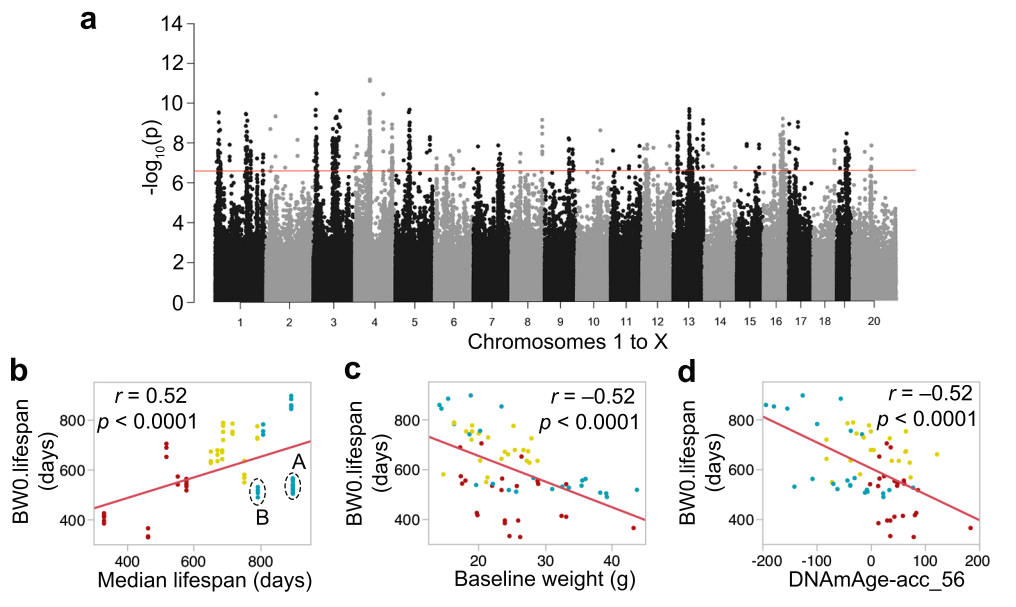

e
Mean lifespan (days)
800
600
400
20
30
40
Baseline weight (g)
r = –0.27
p = 0.04
Body weight (past)
Lifespan (future)
a
PC1
200
0
-200
10
20
30
40
Baseline weight (g)
r = –0.3
p = 0.02
b
Genic means (logRPKM)
3.9
8
3.7
10
20
30
40
Baseline weight (g)
r = –0.29
p = 0.03
Methylome (present)
c
Genic variance
4
3
10
20
30
40
Baseline weight (g)
r = 0.29
p = 0.02
d
Mean lifespan (days)
800
600
400
-200
0
200
PC4
r = 0.85
p < 0.0001

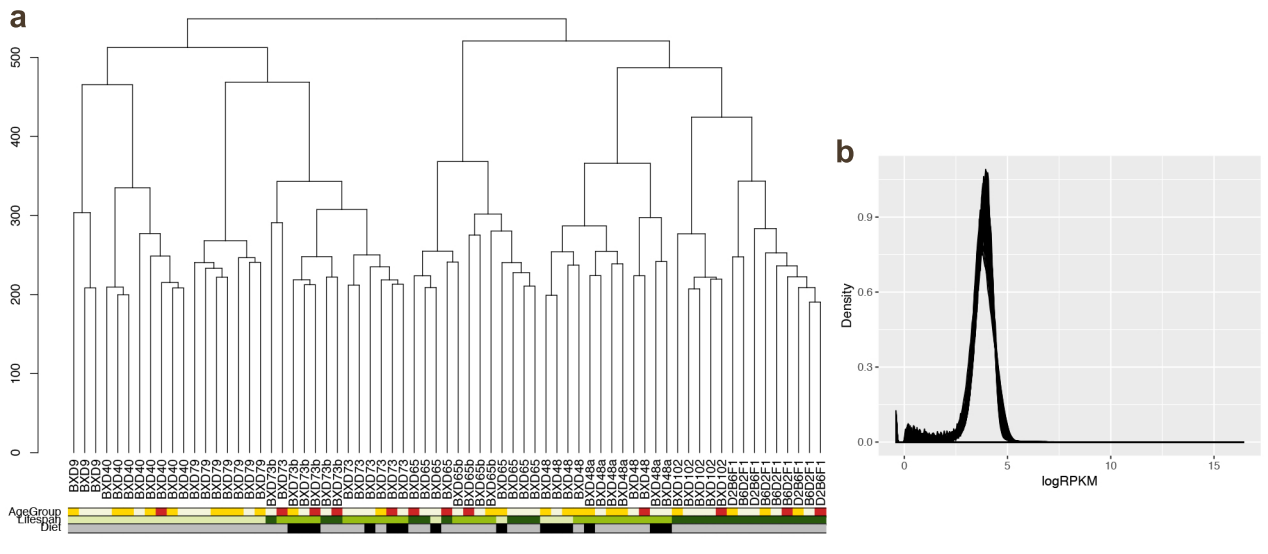

a
3000
2000
1000
Count
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
b
3000
2000
1000
0
Count
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
c
14
12
10
8
6
4
2
0
-log10(p)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 20
Chromosomes 1 to X
d
40
30
20
10
0
Percent total
Exon
Intron
Intergenic
LTR
LINE
10
5
0
Percent total
CpG-Island
Promoter
5UTR
3UTR
TTS
Genome-wide
MBD-seq bins
Age-hyper
Age-hypo
e
20
10
0
CpG counts
-0.002
0
0.001
0.003
Age coefficient
f
4
3
2
Mean methylation (logRPKM)
-0.002
0
0.001
0.003
Age coefficient
Age effect
age-hyper
age-hypo

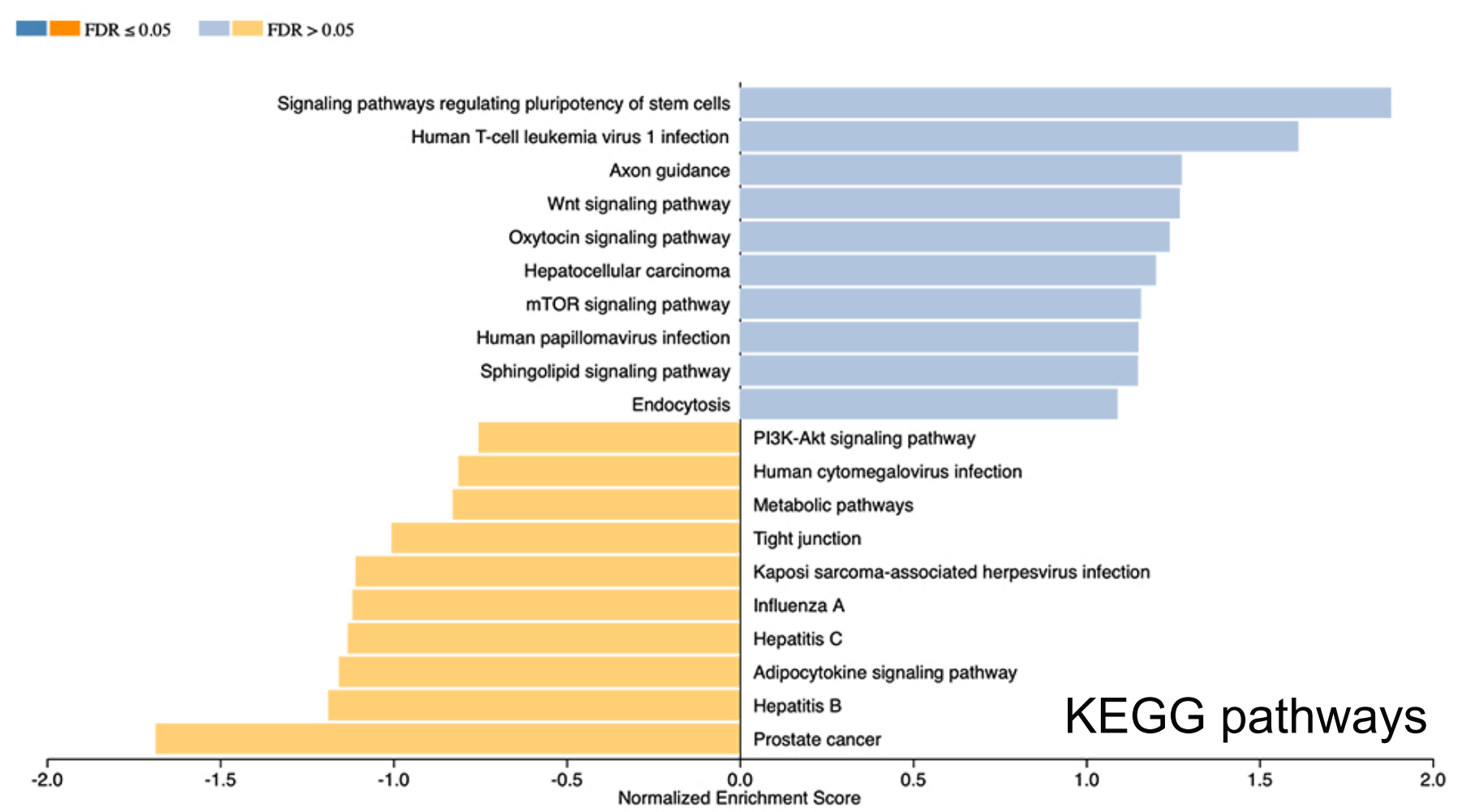

a
DNAmAge_EN training (days)
800
600
400
200
200
400
600
800
Age (days)
b
DNAmAge_EN testing (days)
800
600
400
200
200
400
600
800
Age (days)
c
RNAclock_EN.v1 training(days)
800
600
400
200
200
400
600
800
Age (days)
d
RNAclock_EN.v1 testing (days)
800
600
400
200
200
400
600
800
Age (days)
e
RNAclock_EN.v2 training (days)
800
600
400
200
200
400
600
800
Age (days)
f
RNAclock_EN.v2 testing (days)
800
600
400
200
200
400
600
800
Age (days)

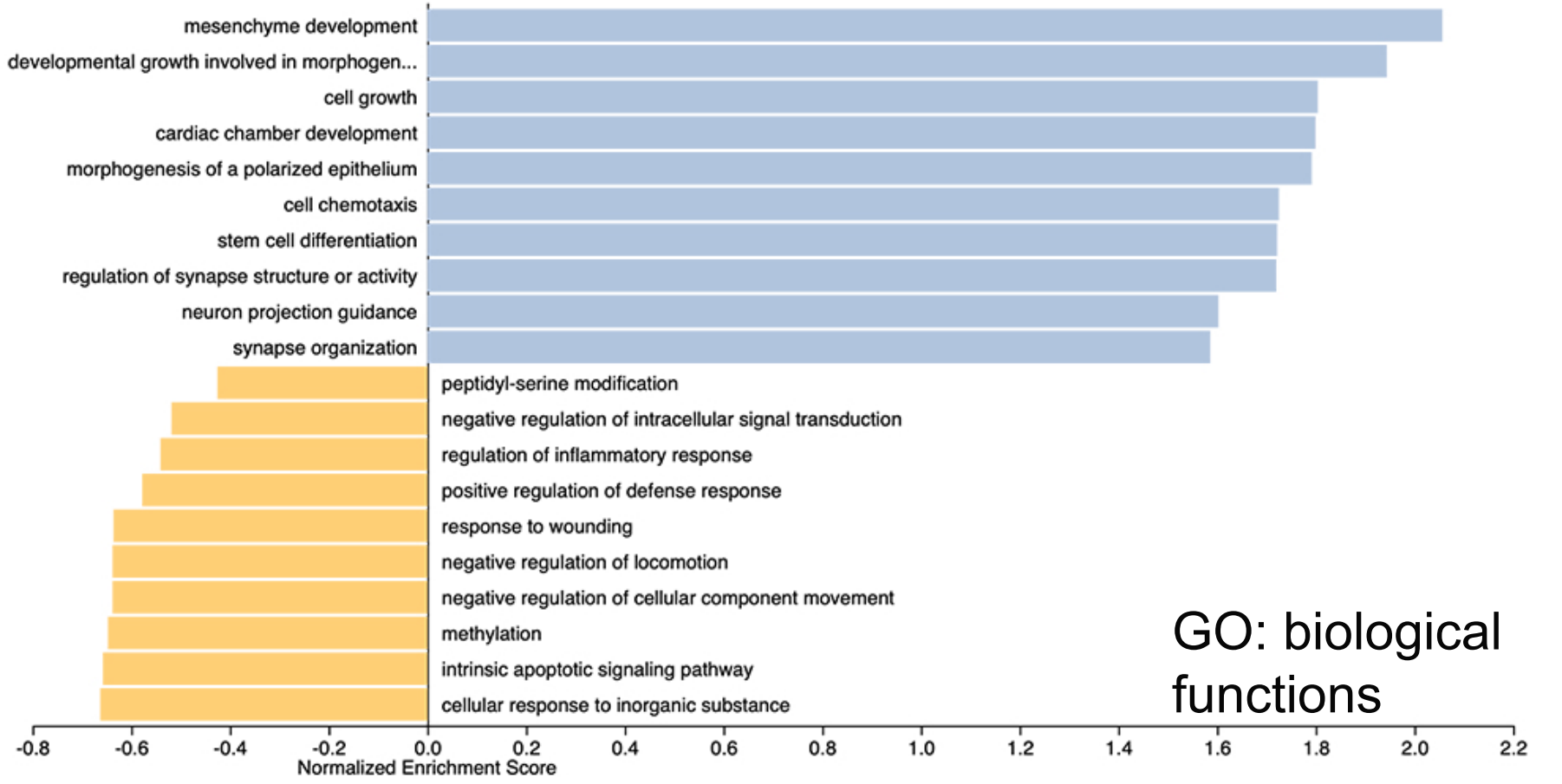

a
DNAmAge_500 (days)
800
600
400
200
200
400
600
800
Age (days)
r = 0.90
p < 0.0001
e
DNAmAge_56 (days)
800
600
400
200
200
400
600
800
Age (days)
r = 0.83
p < 0.0001
Lifespan
long
med
short
Linear Fit
Polynomial Fit Degree=2
b
DNAmAge-acc_500
100
0
-100
-200
400
600
800
1000
Median lifespan (days)
r = 0.01
ns
f
DNAmAge-acc_56
200
100
0
-100
-200
400
600
800
1000
Median lifespan (days)
r = –0.59
p < 0.0001
c
DNAmAge-acc_500
100
0
-100
-200
20
30
40
Baseline weight (g)
r = 0.38
p = 0.001
g
DNAmAge-acc_56
200
100
0
-100
-200
20
30
40
Baseline weight (g)
r = 0.25 p = 0.04
d
DNAmAge-acc_500
15
0
-15
CD
HFD
Diet
h
DNAmAge-acc_56
25
0
-25
-50
CD
HFD
Diet


a
Lifespan coefficient
0.1
0.05
0
-0.05
-0.1
-0.001
0
0.001
0.002
Age coefficient
b
Casz1 exon9 (logRPKM)
4.5
4.0
3.5
3.0
2.5
200
400
600
Age (days)
long
med
short
Casz1 exon9 (logRPKM)
4
3
short
med
long
Lifespan group
c
Cyp46a1 intron8 (logRPKM)
4
3
2
1
200
400
600
Age (days)
Cyp46a1 intron8 (logRPKM)
4
2
0
short
med
long
Lifespan group
d
Abca7 intron43 (logRPKM)
4.5
4.0
3.5
3.0
200
400
600
Age (days)
Abca7 intron43 (logRPKM)
4
3
short
med
long
Lifespan group


a
-log10(p)
Chromosomes 1 to X
b
r = 0.52
p < 0.0001
BW0.lifespan (days)
Median lifespan (days)
A
B
c
r = –0.52
p < 0.0001
BW0.lifespan (days)
Baseline weight (g)
d
r = –0.52
p < 0.0001
BW0.lifespan (days)
DNAmAge-acc_56


a
b
0
100
200
300
400
500
BXD9
BXD9
BXD9
BXD40
BXD40
BXD40
BXD40
BXD40
BXD40
BXD40
BXD40
BXD79
BXD79
BXD79
BXD79
BXD79
BXD79
BXD79
BXD73b
BXD73
BXD73b
BXD73b
BXD73b
BXD73b
BXD73b
BXD73b
BXD73
BXD73
BXD73
BXD73
BXD73
BXD73
BXD73
BXD65
BXD65
BXD65
BXD65
BXD65
BXD65
BXD65
BXD65b
BXD65b
BXD65b
BXD65b
BXD65b
BXD65b
BXD65b
BXD65
BXD65
BXD65
BXD65
BXD65
BXD65
BXD65
BXD48
BXD48
BXD48
BXD48
BXD48
BXD48
BXD48
BXD48a
BXD48a
BXD48a
BXD48a
BXD48a
BXD48a
BXD48a
BXD48
BXD48
BXD48
BXD48
BXD48
BXD48
BXD48
BXD102
BXD102
BXD102
BXD102
BXD102
BXD102
BXD102
D2B6F1
D2B6F1
D2B6F1
D2B6F1
D2B6F1
D2B6F1
D2B6F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
B6D2F1
D2B6F1
D2B6F1
D2B6F1
D2B6F1
D2B6F1
D2B6F1
D2B6F1
AgeGroup
Lifespan
Diet
Density
0.0
0.3
0.6
0.9
0
5
10
15
logRPKM


FDR ≤ 0.05
FDR > 0.05
Signaling pathways regulating pluripotency of stem cells
Human T-cell leukemia virus 1 infection
Axon guidance
Wnt signaling pathway
Oxytocin signaling pathway
Hepatocellular carcinoma
mTOR signaling pathway
Human papillomavirus infection
Sphingolipid signaling pathway
Endocytosis
PI3K-Akt signaling pathway
Human cytomegalovirus infection
Metabolic pathways
Tight junction
Kaposi sarcoma-associated herpesvirus infection
Influenza A
Hepatitis C
Adipocytokine signaling pathway
Hepatitis B
Prostate cancer
KEGG pathways
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Normalized Enrichment Score


mesenchyme development
developmental growth involved in morphogen...
cell growth
cardiac chamber development
morphogenesis of a polarized epithelium
cell chemotaxis
stem cell differentiation
regulation of synapse structure or activity
neuron projection guidance
synapse organization
peptidyl-serine modification
negative regulation of intracellular signal transduction
regulation of inflammatory response
positive regulation of defense response
response to wounding
negative regulation of locomotion
negative regulation of cellular component movement
methylation
intrinsic apoptotic signaling pathway
cellular response to inorganic substance
GO: biological functions
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
Normalized Enrichment Score