Genetic Analyses of Epigenetic Predictors that Estimate Aging, Metabolic Traits, and Lifespan

Khyobeni Mozhui [1, 2*], Ake T. Lu[3], Caesar Z Li [3], Amin Haghani [4], Jose Vladimir Sandoval-Sierra [1], Robert W. Williams [2], Steve Horvath[3, 4]

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

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

[3] Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

[4] Department of Biostatistics, Fielding School of Public Health, University of California Los Angeles, Los Angeles, CA, USA

[*]Correspondence:
Khyobeni Mozhui
kmozhui@uthsc.edu

Abstract

DNA methylation (DNAm) remodeling is an important aspect of aging. Recently, a 'pan-mammalian' microarray was developed that provides a unifying platform to profile conserved CpGs across mammalian clades, and to develop universal biomarkers for aging. We used this array to measure DNAm in 339 predominantly female mice belonging to the BXD family. We computed epigenetic clocks and predicted maximum lifespan (predicted-maxLS), and tested associations with DNAm entropy, diet, weight, metabolic traits, and genetic variation. Epigenetic age acceleration (EAA) and predicted-maxLS were correlated with genotype-dependent lifespan of the BXD strains. Genetic analyses uncovered quantitative trait loci (QTLs) on chromosome (Chr) 11 that encompasses the *Erbb2/Her2* oncogenic region, and on Chr19 that contains a cytochrome P450 cluster. Both loci harbor candidate genes associated with EAA in humans (*STXBP4*, *NKX2*-3, *CUTC)*. Transcriptome and proteome analyses revealed enrichment in oxidation-reduction and metabolic genes. The results highlight loci concordant in humans and mice, and demonstrate intimate links between metabolism, body weight, and epigenetic aging.

Keywords: epigenetic clock, lifespan, aging, DNA methylation, QTL, weight, diet

Introduction

Epigenetic clocks are widely used molecular biomarkers of aging.[1] These biological clocks are based on the methylation status across an ensemble of "clock CpGs" that are collectively used to derive a DNA methylation (DNAm) based estimate of age (DNAmAge). This estimate tracks closely, but not perfectly, with an individual's chronological age. How much the DNAmAge deviates from the known chronological age is a measure of the rate of biological aging. Denoted as epigenetic age acceleration (EAA), a more accelerated measure (positive EAA) suggests an older biological age. While DNAmAge predicts age, its age-adjusted counterpart, EAA, is associated with health, fitness, exposure to stressors, body mass index (BMI), and even life expectancy.[2-6] These DNAm clocks were initially reported for humans.[7,8] Since then, the biomarker has been extended to model organisms,[9-11] and different variants of human clocks have also been developed. Some clocks are tissue specific, others are pan-tissue, and others perform well at predicting health and life expectancy.[5,8,12-14]

A new microarray platform was recently developed to profile CpGs that have high conservation across mammalian clades. This pan-mammalian DNAm array (HorvathMammalMethylChip40) provides a common platform to measure DNAm, and has been used to build universal epigenetic clocks that can estimate age across a variety of tissues and mammalian species.[15,16] Another remarkable development with this array is the novel lifespan predictor that can estimate the maximum lifespan of over 190 mammals at high accuracy.[17]

Here, we test the performance and applicability of these novel DNAm clocks and lifespan predictor in a genetically diverse cohort of mice belonging to the BXD family.[18,19] The BXDs are a well-established mouse genetic reference panel that were first created as a family of recombinant inbred (RI) strains by crossing two inbred progenitors: C57BL/6J (B6) and DBA/2J (D2). The family has been expanded to ~150 fully sequenced progeny strains.[20,21] The BXD population incorporates significant genetic variation, and likewise, show considerable phenotypic variability in their metabolic profiles, response to diet, aging rates, and natural life expectancy.[18-20,22-24] The availability of genetic data, and accompanying deep multi-omics and phenomics, make the BXDs a powerful experimental population to evaluate the DNAm biomarkers, and to dissect the genetic modulators of epigenetic aging. Previously, we explored the aging methylome in a small number of BXD cases and found that high fat diet (HFD) and higher body weight were associated with higher age-dependent changes in methylation.[25] In the present work, our goals were to (1) test the accuracy of the DNAm measures in predicting age and lifespan of the BXDs, (2) examine associations with diet and metabolic characteristics, and (3) apply quantitative trait locus (QTL) mapping and gene expression analyses to uncover loci and genes associated with these DNAm biomarkers.

Our results are consistent with a faster clock for cases on HFD, and with higher body weight. Both the DNAmAge and lifespan predictors were correlated with the genotype-dependent life expectancy of female BXDs. We report QTLs on chromosomes (Chrs) 11 and 19. A strong candidate gene in the chromosome (Chr) 11 interval (referred to as *Eaaq11*) is *Stxbp4*, a gene that has been consistently associated with EAA by human genome-wide association studies (GWAS).[26-28] The Chr19 QTL (*Eaaq19*) also harbors strong contenders including *Cyp26a1*, *Myof*, *Cutc*, and *Nkx2–3*, and the conserved genes in humans have been associated with longevity and

78 EAA.[28-30] *Eaaq19* may also have an effect on body weight change with age. We performed gene
79 expression analyses to clarify the molecular pathways associated with the DNAm traits, and
80 this, perhaps unsurprisingly, highlighted metabolic networks as strong expression correlates of
81 epigenetic aging.

82 Results

83 Description of samples

84 The present study used liver DNAm data from 339 predominantly female mice (18 males only)
85 belonging to 45 isogenic members of the BXD family, including F1 hybrids, and parental strains.
86 Age ranged from 5.6 to 33.4 months. Mice were all weaned onto a normal chow (control diet;
87 CD) and a balanced subset of cases were then randomly assigned to HFD (see Roy et al. for
88 details[19]). Tissues were collected at approximately six months intervals (see Williams et al.[18]).
89 Individual-level data of cases used in this study are in **Data S1**.

90 DNAm clocks and chronological age prediction

91 The first epigenetic clock we tested was the universal relative age clock described in Lu et al.[16]
92 We refer to the age estimates as univ.DNAmAge, and this had a correlation of r = 0.92 with the
93 chronological age of mice (**Table 1; Fig 1a**).

94 We also built three different mouse clock, and each was developed as a pair depending on
95 whether the training set used all tissues (pan-tissue) or only liver (liver specific). These are: (1) a
96 general DNAm clock (referred to simply as DNAmAge): clock trained without pre-selecting for
97 any specific CpG subsets; (2) developmental clock (dev.DNAmAge): built from CpGs that change
98 during development; and (3) interventional clock (int.DNAmAge): built from CpGs that change
99 in response to aging related interventions such as caloric restriction and growth hormone
100 knockout.[9,11,25] These clocks were trained in a large mouse dataset that excluded the BXDs
101 (previously reported in [16,17]), and are therefore unbiased to BXD characteristics. The specific
102 CpGs and coefficients for clock computation are in **Data S2**. All the mouse clocks performed
103 well in age estimation but the interventional clocks had higher deviation from chronological age

Table 1. Epigenetic predictors and correlation with chronological age and entropy

Clock type	DNAmAge name	Tissue	r with age (n=339)	r with entropy (n=339)
Universal clock	univ.DNAmAge	pan	0.92	0.59
Standard clocks	DNAmAge	pan	0.89	0.50
		liver	0.92	0.51
Developmental clocks	dev.DNAmAge	pan	0.87	0.46
		liver	0.91	0.45
Interventional clocks	int.DNAmAge	pan	0.85	0.38
		liver	0.86	0.49
Lifespan predictor	Predicted-maxLS	pan	-0.07 (ns)	-0.13 (p=0.02)

(**Table 1**). The age-adjusted EAA derived from these clocks showed wide individual variation (**Fig 1b**).

We also used the universal maximum lifespan predictor[17] to estimate the potential maximum lifespan (predicted-maxLS) of mice (**Data S1**). Predicted-maxLS was uncorrelated with chronological age (**Table 1**), and this is expected since the chronological age represents the time when the biospecimens were collected; not the time of natural demise. Instead, the predicted-maxLS showed an overall inverse correlation with EAA from the different clocks, and this suggests higher age-acceleration for mice with lower predicted-maxLS (**Data S3**).

Association with methylome entropy

The methylome-wide entropy provides a measure of randomness and information loss, and this increased with chronological age (**Fig 1c**).[7] As direct correlates of chronological age, all the DNAmAge estimates had significant positive correlations with entropy (**Table 1**).

We hypothesized that higher entropy levels will be associated with (a) higher EAA, and (b) lower predicted-maxLS. Indeed, the univ.EAA had a positive correlation with entropy that was significant regardless of diet (**Fig 1d**). However, the EAA from the mouse clocks showed only weak correlations with entropy (**Data S3**). Entropy had a modest negative correlation with predicted-maxLS primarily in the CD group (**Fig 1e**). Taken together, our results indicate that discordance in the methylome increases with age, and is higher with higher univ.EAA. Mice with shorter predicted-maxLS may also had slightly higher entropy.

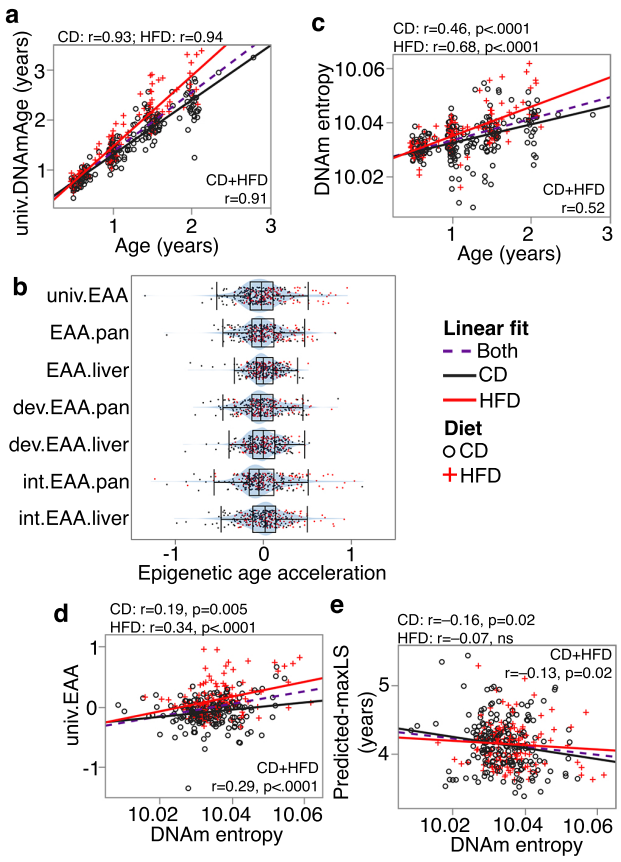


Fig 1. DNA methylation readouts and inter-correlations
(a) Correlation between predicted age (universal clock) and chronological age. Mice on control diet depicted by black circles (CD; normal lab chow; n = 210); mice on high fat diet (HFD; n = 129) by red crosses. **(b)** Violin plots of epigenetic age acceleration (EAA) derived from different DNAmAge clocks. **(c)** Shannon entropy, calculated from the full set of high quality CpGs, increases with age. **(d)** DNA methylation entropy has a direct correlation with EAA derived from the universal clock. **(e)** There is a slight inverse correlation between the entropy and predicted maximum lifespan (predicted-maxLS).

How the epigenetic readouts relate to diet, body weight, and sex

Diet. EAA from most of the clocks, including the universal clock, were significantly higher in the HFD group (**Table 2**). Entropy was also significantly higher in the HFD group. The maxLS did not differentiate between diets **(Table 2**).

Body weight. Body weight was first measured when mice were at an average age of 4.5 ± 2.7 months. We refer to this initial weight as baseline body weight (BW0). For mice on HFD, this was usually before introduction to the diet, with the exception of 48 cases that were first weighed 1 or 3 days after HFD (**Data S1**). In the CD group, only the EAA from the pan-tissue standard clock and liver developmental clock showed significant positive correlations with BW0 (**Table 2**). In the HFD group, the positive correlation with BW0 was more robust and consistent across all the clocks, and this may have been due to the inclusion of the 48 cases that had been on HFD for 1 or 3 days. Taking only these 48 cases, we found that higher weight even after 1 day of HFD had an age-accelerating effect (**Data S3**). This was particularly strong for the interventional clocks (r = 0.45, p = 0.001 for int.EAA, pan-tissue; r = 0.58, p < 0.0001 for int.EAA, liver), and for the universal clock (**Fig 2a**). Second weight was measured 7.4 ± 5.2 weeks after BW0 (mean age 6.3 ± 2.8 months). We refer to this as BW1 and we estimated the weight change as deltaBW = BW1 – BW0. DeltaBW was a positive correlate of EAA on both diets, albeit more pronounce in the HFD group (**Fig 2b**; **Data S3**). The final body weight (BWF) was measured

Table 2. Association with diet and weight, and heritability of the epigenetic readouts

Type	EAA	Diet	Mean (SD)	Diet (p)	r BW0 [a]	p BW0	r BWF [a]	p BWF	H^2	Strain r [b]
Universal clock	univ.EAA	CD	-0.08± 0.22	<.0001	0	ns	0.11	ns	0.37	0.67
		HFD	0.13 ± 0.27		0.35	<.0001	0.50	<.0001	0.43	
Mouse clocks	EAA, pan	CD	-0.05 ± 0.21	<.0001	0.19	0.006	0.29	<.0001	0.49	0.54
		HFD	0.07 ± 0.21		0.21	0.01	0.42	<.0001	0.50	
	EAA, liver	CD	0 ± 0.17	ns	0.09	ns	0.20	0.003	0.40	0.73
		HFD	0.03 ± 0.14		0.22	0.01	0.49	<.0001	0.52	
	dev.EAA, pan	CD	-0.04 ± 0.23	0.004	0.09	ns	0.22	0.001	0.53	0.76
		HFD	0.03 ± 0.22		0.27	0.002	0.45	<.0001	0.61	
	dev.EAA, liver	CD	0 ± 0.2	ns	0.19	0.002	0.29	<.0001	0.46	0.78
		HFD	0 ± 0.16		0.29	0.0007	0.47	<.0001	0.60	
	int.EAA, pan	CD	-0.05 ± 0.25	0.0003	0.03	ns	0.21	0.002	0.27	0.66
		HFD	0.06 ± 0.33		0.22	0.01	0.46	<.0001	0.45	
	int.EAA, liver	CD	-0.04 ± 0.22	<.0001	0.05	ns	0.18	0.01	0.59	0.80
		HFD	0.11 ± 0.25		0.27	0.002	0.58	<.0001	0.54	
Entropy	-	CD	10.034 ± 0.007	0.004	-0.15	0.03	-0.34	<.0001	0.39	0.20 (ns)
		HFD	10.036 ± 0.007		-0.11	0.21	0	ns	0.23	
Pred-maxLS	-	CD	4.14 ± 0.37	ns	0.03	ns	0.05	ns	0.66	0.89
		HFD	4.15 ± 0.32		-0.06	ns	-0.11	0.20	0.70	

[a] BW0 is body weight at about 4.5 months of age (n = 339; 210 CD and 129 HFD); BWF is final weight at tissue collection (1 HFD case missing data; n = 338; 210 CD and 128 HFD)

[b] Pearson correlation between strain means for n = 29 BXD genotypes kept on CD and HFD

at the time of tissue harvest, and EAA from all the mouse clocks were significant correlates of BWF on both diets **(Table 2**).

Somewhat unexpected, entropy had an inverse correlation with body weight. This effect was primarily in the CD mice **(Table 2)**. We found no association between predicted-maxLS and the body weight traits (**Table 2**).

Sex effect. Four BXD genotypes (B6D2F1, D2B6F1, BXD102, B6) had cases from both males and females. We used these to test for sex effects. All the mouse clocks showed significant age acceleration in male mice, and this effect was particularly strong for the pan-tissue int.EAA (**Fig 2c; Data S3**). The predicted-maxLS was significantly lower in males (**Fig 2d**). Entropy on the other hand, was significantly higher in females (**Fig 2e**).

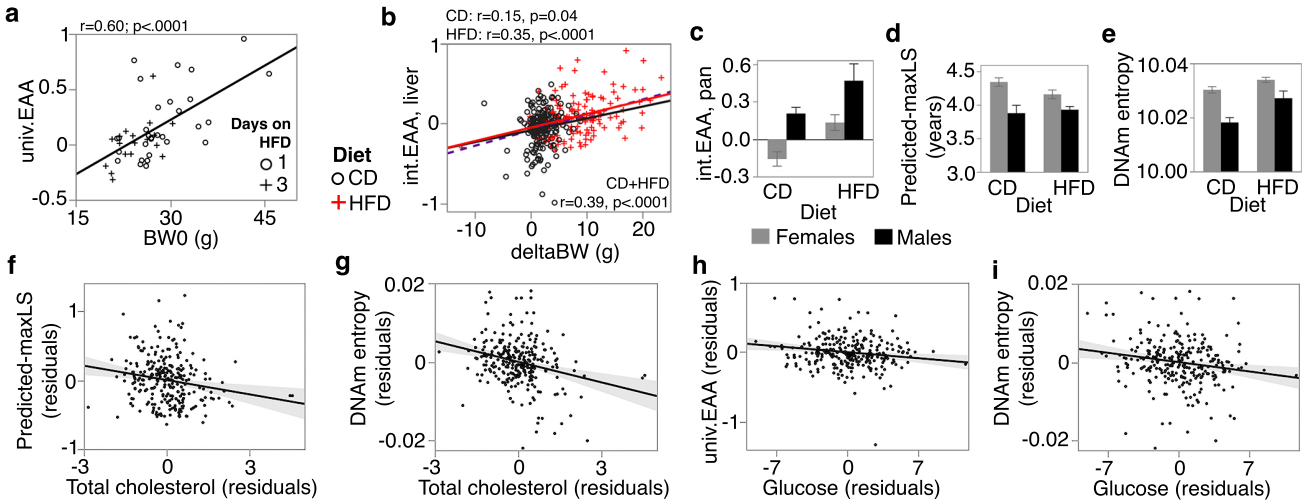


Fig 2. Correlates and modifiers of the epigenetic readouts
(a) For 48 mice, initial body weight (BW0) was measured 1 or 3 days after high fat diet (HFD), and these showed significant correlation between BW0 and epigenetic age acceleration (EAA shown for the universal clock). **(b)** Weight was first measured at mean age of 4.5 ± 2.7 months, and then at 6.3 ± 2.8 months (BW1). Weight gains during this interval (deltaBW = BW1 – BW0) is a direct correlate of EAA derived from the interventional clock in both normal chow (control diet or CD; n = 210) and HFD mice (n = 128). **(c)** For the BXD genotypes with samples from both males and females, males have higher age acceleration and this sex effect is highly significant for the pan-tissue interventional clock (int.EAA). Bars represent mean ± standard error; 40 females (26 CD, 14 HFD) and 18 males (10 CD, 8 HFD). Males had **(d)** lower predicted maximum-lifespan, and **(e)** lower average methylome entropy. **(f–i)** The residual plots display the direction of association between metabolic traits and DNAm readouts (n = 276 cases with metabolic data). After adjusting for chronological age, diet and body weight, serum cholesterol has inverse associations with **(f)** predicted maximum lifespan (p = 0.002), and **(g)** methylome entropy (p = 9.1E-06). Serum glucose level has inverse associations with **(h)** epigenetic age acceleration derived from the universal clock (p = 0.005), and **(i)** methylome entropy (p = 0.003).

Association with metabolic traits

276 cases with DNAm data also had fasted serum glucose and total cholesterol,[18,19] and we examined whether these metabolic traits are associated with three of the DNAm readouts: the univ.EAA, predicted-maxLS, and DNAm entropy. We applied linear regression with age, diet and final body weight as covariates, and this showed significant effects of cholesterol on predicted-maxLS ($p = 0.002$), and entropy ($p = 9E-06$) (**Table S1**). To visualize how cholesterol levels associate with these, we plotted the residual values after the respective predictor and outcome variables were adjusted for age, diet, and BWF. The residual plot shows an inverse association between cholesterol and predicted-maxLS (**Fig 2f**). For entropy, similar to how it related with weight, higher cholesterol predicted lower entropy. Cholesterol had no significant association with univ.EAA **(Table S1)**.

Glucose had an unexpected inverse association with the univ.EAA that indicates lower age acceleration with higher glucose ($p = 0.005$) (**Fig 2h; Table S1**). Higher glucose was also associated with lower entropy ($p = 0.003$) **(Fig 2i**). Glucose was not associated with predicted-maxLS **(Table S1)**.

Association with strain longevity

We next obtained longevity data from a parallel cohort of female BXD mice that were allowed to age on CD or HFD.[19] We evaluated whether the DNAm readouts were informative of strain-level lifespan. Since the strain lifespan was determined from female BXDs, we restricted this to only the female cases. For strains with natural death data from $n \geq 5$, we computed the minimum (minLS), 25th quartile (25Q-LS), mean, median lifespan, 75th quartile (75Q-LS), and maximum lifespan (maxLS) (**Data S1**). Specifically, we postulated (a) an accelerated clock for strains with shorter lifespan (i.e., inverse correlation), (b) a direct correlation between predicted-maxLS and observed lifespan, and (c) higher entropy with shorter lifespan.

Overall, the EAA measures showed the expected inverse correlation trend with the lifespan summaries, and this was highly significant for the universal clock (**Table S2**; **Fig 3a**,**b**). For the mouse clocks, this effect was significant for the liver int.EAA (**Table S2**). When separated by diet, these correlations became weaker, but the negative trend remained consistent.

The DNAm entropy had an inverse correlation trend with strain lifespan (**Table S2**). This was nominally significant only for the strain maxLS when CD and HFD groups were combined ($r = -0.13$, $p = 0.02$) but became non-significant when separated by diet.

The predicted-maxLS showed a positive correlation trend with the lifespan summaries, and this was significant for the observed strain maxLS (**Fig 3d**). When separated by diet, the predicted-maxLS remained a significant correlate of strain maxLS only in the CD group.

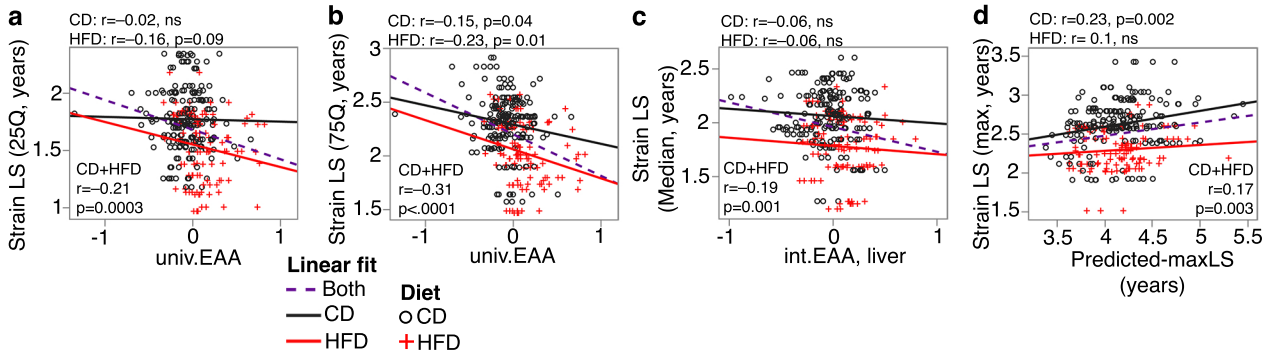


Fig 3. Predictors of strain longevity

BXD genotypes with shorter life expectancy tended to have a more accelerated universal clock (univ.EAA). This inverse relation is depicted for the **(a)** 25th and **(b)** 75th quartile age at natural death for female BXDs kept on either normal chow (CD) or high fat diet (HFD). **(c)** Age acceleration from the liver interventional mouse clock (int.EAA) also showed a similar inverse correlation with strain longevity, but effect was significant only when both diets were included (here illustrated with median lifespan). **(e)** The predicted maximum lifespan had a significant direct correlation with the observed strain maximum lifespan. Analysis in n = 302 female BXDs; 191 CD and 111 HFD.

Genetic analysis of epigenetic age acceleration and predicted-maxLS

The EAA traits had modest to high heritability, and averaged at 0.50 for mouse clocks (**Table 2**). The predicted-maxLS had heritability of 0.66 on CD, and 0.70 on HFD. Another way to gauge level of genetic correlation is to compare between members of strains maintained on different diets. All the EAA traits, and the predicted-maxLS had high strain-level correlations between diets, indicating an effect of background genotype that is robust to dietary differences (**Table 2**).

To uncover genetic loci, we applied QTL mapping using mixed linear modeling that corrects for the BXD kinship structure.[31] First, we performed the QTL mapping for the EAA derived from the different mouse and universal clocks, with adjustment for diet and body weight. EAA from the two interventional clocks had the strongest QTLs (**Data S4**). The pan-tissue int.EAA had a significant QTL on Chr11 (90–99 Mb) with the highest linkage at ~93 Mb (p = 3.5E-06; equivalent to a LOD score of 4.7) (**Fig 4a**). Taking a genotype marker at the peak interval (BXD variant ID DA0014408.4 at Chr11, 92.750 Mb)[20], we segregated the BXDs homozygous for either the D2 (*DD*) or the B6 (*BB*) alleles. Strains with *DD* genotype at this locus had significantly higher int.EAA (**Fig 4a** inset).

The liver int.EAA had the peak QTL on Chr19 (35–45 Mb) with the most significant linkage at markers between 38–42 Mb (p = 9E-07; LOD score of 5.2) (**Fig 4b**). We selected a marker at the peak interval (rs48062674 at Chr19, 38.650 Mb), and the *BB* genotype had significantly higher int.EAA compared to *DD* (**Fig 4b** inset). The QTL map for the univ.EAA did not reach genome-wide significance (**Fig 4c**). However, there were nominally significant peaks at the Chr19 (p = 0.0004), and Chr11 (p = 0.004) intervals.

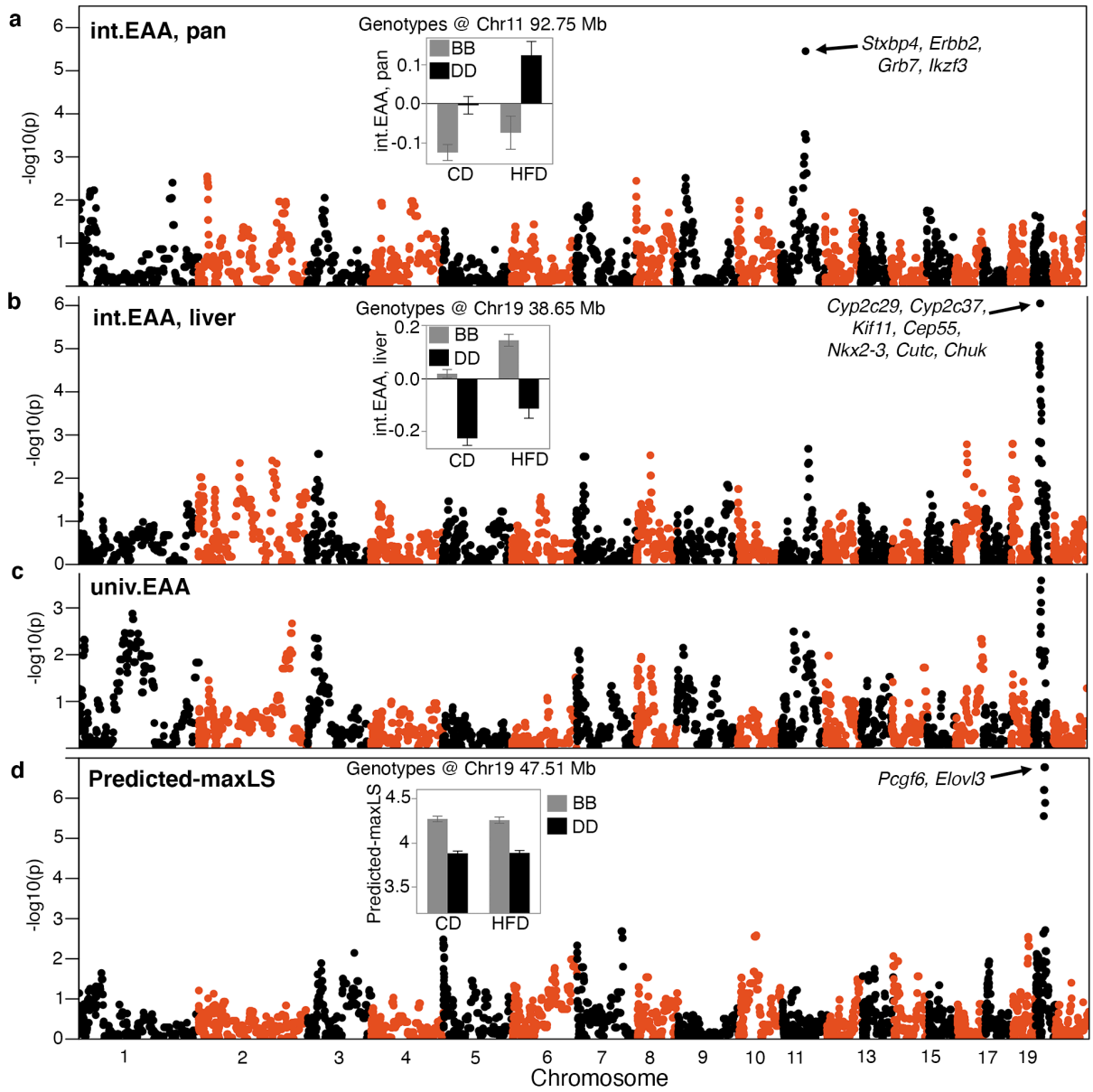


Fig 4. Genetic linkage analysis

The Manhattan plots represent the location of genotyped markers (x-axis), and linkage – $\log_{10}p$ (y-axis). **(a)** The peak quantitative trait locus (QTL) for age acceleration from the pan-tissue interventional clock (int.EAA) is on chromosome (Chr) 11 at ~93 Mb. The inset shows the mean (± standard error) trait values for BXDs the are homozygous for the C57BL/6J allele (*BB*; grey) versus BXDs homozygous for the DBA/2J allele (*DD*; black). **(b)** The liver-specific int.EAA has a peak QTL on Chr19 (~38 Mb). Trait means by genotype at this locus are shown in inset. **(c)** The linkage statistics are weaker for the EAA derived from the universal clock (univ.EAA). However, there are consistent nominally significant peaks on the Chr11 and Chr19 loci. **(d)** The DNA methylation based predicted maximum lifespan also maps to Chr19, but the peak markers are at ~47.5 Mb.

We next performed QTL mapping for DNAm entropy with adjustment for major covariates (diet,

232 chronological age, and body weight). No locus reached genome-wide significance (**DataS4**).
233 There were modest QTLs on Chrs11 and 19. However, the Chr11 region is slightly distal to the
234 markers linked to the EAA traits (minimum p = 0.009 at Chr11, ~103.7 Mb). The Chr19 locus
235 somewhat overlapped the QTL for EAA, but the peak marker (minimum p = 0.0009) is slightly
236 distal at ~48 Mb (**Data S4**).

237 The predicted-maxLS had a significant QTL on Chr19 (**Fig 4d; Data S4**) with the peak markers
238 between 44–48 Mb (p = 2E-07; LOD score of 5.9). This overlaps the EAA QTL, but the peak
239 markers are distal (rs30567369 at 47.510 Mb). At this locus, mice with the *BB* genotype had
240 significantly higher predicted-maxLS (**Fig 4d** inset).

241 To identify regulatory loci that are consistent across the different EAA measures, we applied a
242 multi-trait analysis and derived the linkage meta-p-value for the mouse and universal EAA
243 traits.[32] The peaks on Chrs 11 and 19 attained the highest consensus p-values (**Fig S1a; Data**
244 **S3**). Additional consensus peaks (at $-\log_{10}$meta-p > 6) were observed on Chrs 1 (~152 Mb), and
245 3 (~54 Mb).

246 We focus on the Chrs 11 and 19 QTLs and refer to these as *EAA QTL on Chr 11* (*Eaaq11*), and
247 *EAA QTL on Chr 19* (*Eaaq19*). *Eaaq11* extends from 90–99 Mb. For *Eaaq19*, we delineated a
248 broader interval from 35–48 Mb that also encompasses the peak markers for the predicted-
249 maxLS, albeit these may be separate loci related to EAA (~39 Mb of *Eaaq19*), and predicted-
250 maxLS (~47 Mb of *Eaaq19*).

251 We performed marker-specific linkage analyses for each of the mouse and universal clocks
252 using a regression model that adjusted for diet. With the exception of the liver int.EAA, all the
253 EAA traits had nominal to highly significant associations with the representative *Eaaq11* marker
254 (DA0014408.4), and the *DD* genotype had higher age acceleration (**Table 3**). Mean plots by
255 genotype and diet shows that this effect was primarily in the CD mice (**Fig S1b**). The effect of
256 this locus appeared to be higher for the pan-tissue clocks compared to the corresponding liver-
257 specific clocks. This marker in *Eaaq11* was not associated with either entropy or predicted-
258 maxLS.

259 For proximal *Eaaq19*, the representative marker (rs48062674) was associated with all the EAA
260 traits and the *BB* mice had higher age acceleration on both diets (**Fig S1c**). This marker was not
261 associated with entropy, and had only a weak effect on predicted-maxLS (**Table 3**). When we
262 performed the same analysis with the marker on distal *Eaaq19* (rs30567369), the association
263 with EAA became weaker, and the association with predicted-maxLS became much stronger
264 (**Table 3**). This suggests that the proximal part of *Eaaq19* is related to EAA while the distal part
265 is related to predicted-maxLS.

266 We also tested if these peak markers were associated with the recorded lifespan phenotype
267 and we found no significant association with the observed lifespan of the BXDs.

268

269

270

Association of EAA QTLs with body weight trajectory

Since gain in body weight with age was an accelerator of the clocks, we examined whether the selected markers in *Eaaq11* and *Eaaq19* were also related to body weight change. We retrieved

Table 3: Marker specific linkage analyses for epigenetic age acceleration, predicted maximum lifespan, and body weight trajectory

Predictor	Outcome	Estimate	Std Error	t Ratio	p
Eaaq11 DA0014408.4[DD] Chr11, 92.750 Mb (133 *BB* cases, and 173 *DD* cases)	EAA, pan	0.096	0.023	4.184	3.8E-05
	EAA, liver	0.067	0.017	3.880	0.0001
	dev.EAA, pan	0.077	0.025	3.041	0.003
	dev.EAA, liver	0.037	0.020	1.878	0.06
	int.EAA, pan	0.153	0.029	5.278	2.5E-07
	int.EAA, liver	-0.033	0.025	-1.284	0.20
	univ.EAA	0.101	0.025	4.057	6.3E-05
Eaaq19 rs48062674[DD] Chr19, 38.650 Mb (238 *BB* cases, and 67 *DD* cases)	EAA, pan	-0.083	0.028	-2.954	0.003
	EAA, liver	-0.137	0.020	-6.972	2.0E-11
	dev.EAA, pan	-0.206	0.029	-7.218	4.3E-12
	dev.EAA, liver	-0.124	0.023	-5.461	9.9E-08
	int.EAA, pan	-0.143	0.035	-4.028	7.1E-05
	int.EAA, liver	-0.250	0.027	-9.238	4.6E-18
	univ.EAA	-0.145	0.029	-4.932	1.3E-06
	Pred-maxLS	-0.100	0.048	-2.086	0.04
Distal Eaaq19 rs30567369[DD] Chr19, 47.510 Mb (198 *BB* cases, and 106 *DD* cases)	int.EAA, liver	-0.079	0.026	-2.995	0.003
	univ.EAA	-0.053	0.026	-2.012	0.05
	Pred-maxLS	-0.383	0.036	-10.781	3.9E-23
Mixed model for longitudinal change in body weight					
Predictor	**Outcome**	**Estimate**	**Std Error**	**t Ratio**	**p**
Eaaq11 DA0014408.4[DD] Number of observations = 6885; number of individuals = 2112	Body weight	0.619	0.345	1.794	0.07
Eaaq19 rs48062674[DD] Number of observations = 6132; number of individuals = 1852	Body weight	-1.847	0.374	-4.945	7.6E-07
Distal Eaaq19 rs30567369[DD] Number of observations = 6059; number of individuals = 1802	Body weight	-1.619	0.363	-4.458	8.3-06

longitudinal weight data from a larger cohort of the aging BXD mice that were weighed at regular intervals. After excluding heterozygotes, we tested the effect of genotype. Concordant with the higher EAA for the *DD* genotype at *Eaaq11* in the CD group, the *DD* genotype in the CD group also had slightly higher mean weight at older adulthood (12 and 18 months; **Fig 5a**). However, this marker had no significant association with body weight when tested using a mixed effects model (p = 0.07; **Table 3**). In proximal *Eaaq19*, it was the *BB* genotype that exhibited consistently accelerated clock on both diets, and the *BB* genotype also had higher average body weight by 6 months of age (**Fig 5b**), and this locus had a significant influence on the body weight trajectory (p = 7.6E-07; **Table 3**). The nearby marker on distal *Eaaq19* also showed a similar pattern of association with body weight (**Table 3**).

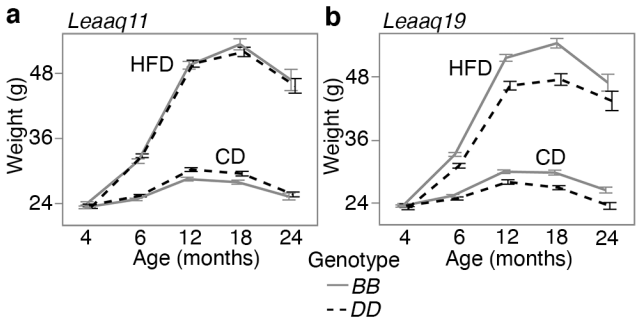


Fig 5. Body weight trajectory by diet and genotype
Body weight was measured at regular age intervals (x-axis) from **(a)** 2112 BXD mice that were homozygous at the *Leaaq11* marker (DA0014408.4; 842 *BB*, 1279 *DD*), and **(b)** 1852 BXD mice that were homozygous at the proximal *Leaaq19* marker (rs48062674; 1252 *BB*, 600 *DD*). Mice were maintained on either control diet (CD) or high fat diet (HFD). The graphs show the segregation of body weight over time by diet and genotype. Mean ± standard error; heterozygotes were excluded.

Candidate genes for epigenetic age acceleration

There are several positional candidate genes in *Eaaq11* and *Eaaq19*. To narrow the list, we applied two selection criteria: genes that (1) contain missense and/or stop variants, and/or (2) contain non-coding variants and regulated by a cis-acting expression QTL (eQTL). For the eQTL analysis, we utilized an existing liver transcriptome data from the same aging cohort.[18] We identified 24 positional candidates in *Eaaq11* that includes *Stxbp4*, *Erbb2* (*Her-2* oncogenic gene), and *Grb7* (growth factor receptor binding) (**Data S5**). *Eaaq19* has 81 such candidates that includes a cluster of cytochrome P450 genes, and *Chuk* (inhibitor of NF-kB) in the proximal region, and *Pcgf6* (epigenetic regulator) and *Elovl3* (lipid metabolic gene) in the distal region (**Data S5**).

For further prioritization, we converted the mouse QTL regions to the corresponding syntenic regions in the human genome, and retrieved GWAS annotations for these intervals.[33] We specifically searched for the traits: epigenetic aging, longevity, age of menarche/menopause/puberty, Alzheimer's disease, and age-related cognitive decline and dementia. This highlighted 5 genes in *Eaaq11*, and 3 genes in *Eaaq19* (**Table S4**). We also identified a GWAS study that found associations between variants near *Myof-Cyp26a1* and human longevity,[30] and a meta-GWAS that found gene-level associations between *Nkx2–3* and *Cutc*, and epigenetic aging (**Table S4**).[28]

Gene expression correlates of EAA and predicted max-LS

316 Liver RNA-seq data was available for 153 of the BXD cases that had DNAm data (94 CD, and 59
317 HFD).[18] We used this set to perform transcriptome-wide correlation analysis for the univ.EAA.
318 To gain insights into gene functions, we selected the top 2000 transcriptome correlates (|r| ≥
319 0.37, p ≤ 2.8E-06; **Data S6**) for functional enrichment analysis. These top correlates represented
320 transcript variants from 1052 unique genes and included a few positional candidates (e.g., *Ikzf3*,
321 *Kif11*, *Cep55*, *Cyp2c29*, *Cyp2c37*). Only 62 transcripts from 36 unique genes were negatively
322 correlated with univ.EAA, and this set was significantly enriched (Bonferroni correct p < 0.05) in
323 oxidation-reduction, and metabolic pathways (**Data S7**; **Fig 6a**). These functional categories
324 included the cytochrome genes, *Cyp2c29* and *Cyp2c37*, located in *Eaaq19*. This set was also
325 highly liver specific. The positive correlates were enriched in a variety of gene functions, and

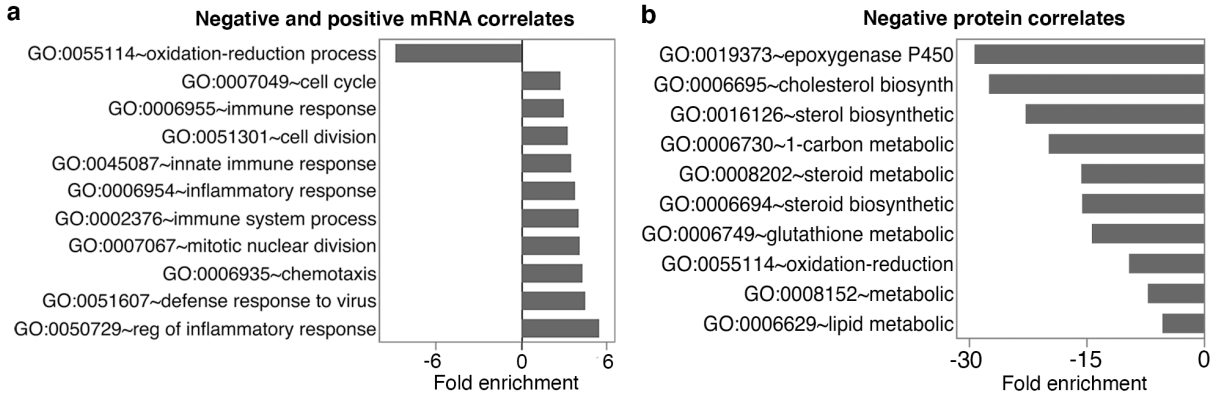


Fig 6. Gene expression correlates of the universal clock
(a) Graph displays the enriched gene ontologies (GO) for the top 2000 transcriptome correlates of epigenetic age acceleration (univ.EAA). The y-axis shows the fold enrichment; positive values are for the positive correlates, and negative values are for the negative correlates of univ.EAA. The smaller set of negative mRNA correlates is enriched in oxidation-reduction process. For positive mRNA correlates, the top 10 enriched GOs highlight immune response, cell cycle, and mitosis. **(b)** The top 10 negative proteome correlates of univ.EAA is also enriched in oxidation-reduction and related metabolic pathways, and these are populated by several cytochrome P450 genes.

326 was not a liver-specific gene set (**Data S7**). Taking the top 10 GO categories, we can broadly
327 discern two functional domains: immune and inflammatory response, and mitosis and cell cycle
328 (**Fig 6a**). To verify that these associations are robust to the effect of diet, we repeated the
329 correlation and enrichment analysis in the CD group only (n = 94). Again, taking the top 2000
330 correlates (|r| ≤ 0.30; p ≤ 0.003), we found the same enrichment profiles for the positive and
331 negative correlates (**Data S7**).

332 Next, we performed the correlational analysis using liver proteomic data that was available for
333 164 of the BXDs with DNAm data. The proteome data quantifies over 32000 protein variants
334 from 3940 unique genes.[18] We took the top 2000 protein correlates of univ.EAA (|r| ≥ 0.27, p ≤
335 6.0E-04) (**Data S8**). This represented protein levels from 563 unique genes. 1139 protein
336 variants (215 genes) had negative correlations, and similar to the mRNA correlates, there was
337 enrichment in oxidation-reduction and metabolic processes. This set was also enriched in liver

genes, and included pathways related to lipid and steroid metabolism, epoxygenase p450 pathway, and xenobiotics (**Fig 6b**; **Data S9**). These categories were populated by the cytochrome genes including candidates in *Eaaq19* (e.g., *Cyp2c29*, *Cyp2c37*). The positive proteome correlates showed a different functional profile than the transcriptomic set. These were enriched in genes related to transport (includes apolipoprotein such as APOE), cell adhesion, protein translation, protein folding, and metabolic pathways related to glycolysis and gluconeogenesis (**Data S9**).

We performed a similar transcriptome and proteome analysis for the predicted-maxLS. For mRNA, both the negative and positive correlates were enriched in metabolic pathways including glucose and lipid metabolism (**Data S10, S11**). Similarly, the positive and negative protein correlates of predicted-maxLS converged on oxidation-reduction processes (included cytochrome genes located in proximal *Eaaq19*) and metabolic pathways (**Data S12, S13**).

Discussion

The goals of this study were to examine the aging methylome at conserved CpGs, test the performance of the DNAm based biomarkers, and identify potential modifiers and genetic drivers. HFD had a strong age-accelerating effect that concurs with the association between EAA and obesity in humans.[25,34,35] Age-acceleration due to diet manifested within the first 1 to 3 days of transitioning from normal lab chow to HFD. Even among the CD mice, higher weight gain at a younger age was associated with an accelerated clock.

Somewhat surprising was how entropy related to the metabolic traits. Epigenetic entropy increases with age, and is likely an indicator of the increase in stochastic noise over time.[7,36] In biological systems, entropy is kept at bay by the uptake of energy, and investment in maintenance and repair.[37] As HFD increased entropy (possibly due to higher cellular heterogeneity and adiposity of liver tissue), we expected entropy to be higher with higher body weight. But instead, entropy had an inverse correlation with weight, an effect that was primarily in the CD mice. Higher levels of serum glucose and total cholesterol were also associated with lower entropy. The reason for this is unclear, and we can only speculate that the enhanced energy consumption in mice that had higher metabolic substrates may have kept the methylome in a more ordered state. Despite this, mice with higher entropy also tended to have higher EAA. Entropy had a modest negative correlations with both the the DNAm based predicted-maxLS, and the known strain-level maxLS. The predicted-maxLS on the other hand, showed no direct association with diet or body weight, but higher total cholesterol and higher EAA were associated with shorter predicted-maxLS.

For the BXDs, life expectancy is highly dependent on the background genotype, and mean lifespan varies from under 16 months for strains such as BXD8 and BXD13, to over 28 months in strains such as BXD91 and BXD175.[19,22,24] While the EAA from the universal was inversely correlated with strain lifespan data, the predicted-maxLS had a direct correlation with observed strain maxLS. We note that the predicted-maxLS overestimated the strain max-LS by 0.7 to 3 years (median error of +1.6 years). Nonetheless, the correlation between individual-level predicted-, and strain-level observed maxLS is remarkable considering that both the universal clock and max-LS predictor are pan-mammalian, and species- and tissue-agnostic.[16,17] Our

379 results suggest that these universal epigenetic predictors of biological aging and lifespan are
380 informative of the subtle and normative lifespan variation in a family of inbred mice. The
381 analysis between the epigenetic readouts and lifespan was also an indirect comparison. Unlike
382 the comparison with body weight and metabolic traits, which were traits measured from the
383 same individual, the lifespan data are strain characteristics computed from a parallel cohort of
384 mice that were allowed to survive till natural mortality. Nonetheless, this indirect comparison
385 demonstrates that these multi-species epigenetic predictors capture genotype-dependent
386 effects on aging and lifespan within a species.

387 We tested different versions of the mouse DNAmAge clocks, and these appeared to capture
388 slightly different aspects of epigenetic aging. For instance, the interventional clocks were
389 sensitive to diet and early weight change, but not related to BW0 in the CD mice. Instead, BW0
390 had a significant accelerating effect on the liver developmental clock (dev.EAA).

391 Our goal was to take these different clocks and identify regulatory loci that were the most
392 stable and robust to the slight algorithmic differences in building the clocks. A notable
393 candidate in *Eaaq11* is Syntaxin binding protein 4 (*Stxbp4*, aka, *Synip*), located at 90.5 Mb.
394 *Stxbp4* is a high-priority candidate due to the concordant evidence from human genetic studies.
395 The conserved gene in humans is a replicated GWAS hit for the intrinsic rate of epigenetic
396 aging.[26-28] In the BXDs, *Stxbp4* contains several non-coding variants, and a missense mutation
397 (rs3668623), and the expression of *Stxbp4* in liver is modulated by a *cis*-eQTL. *Stxbp4* plays a
398 key role in insulin signaling,[38] and has oncogenic activity and implicated in different cancers.[39,40]
399 Furthermore, GWAS have also associated *STXBP4* with age of menarche.[41,42] *Eaaq11*
400 corresponds to the 17q12-21 region in humans, and the location of additional oncogenic genes,
401 e.g., *ERBB2/HER2*, *GRB7*, and *BRCA1*.[43] The mouse *Brca1* gene is a little distal to the peak QTL
402 region and is not considered a candidate here, although it does segregate for two missense
403 variants in the BXDs. *Erbb2* and *Grb7* are in the QTL region, and *Erbb2* contains a missense
404 variant (rs29390172), and *Grb7* is modulated by a *cis*-eQTL. *Nr1d1* is another candidate in
405 *Eaaq11*, and the co-activation of *Erbb1*, *Grb7*, and *Nr1d1* has been linked to breast and other
406 cancers.[44,45]

407 *Eaaq19* was consistently associated with EAA from all the clocks we evaluated, and also with
408 body weight gains, irrespective of diet. The predicted-maxLS also maps to this region, and
409 DNAm entropy may also have a weak association with markers at this interval. The EAA traits
410 have peak markers in the proximal part of *Eaaq19* (around the cytochrome cluster), and the
411 predicted-maxLS peaks in the distal portion (over candidates like *Elovl3*, *Pcgf3*). Two candidates
412 in *Eaaq19* have been implicated in epigenetic aging in humans based on gene-level meta-
413 GWAS: NK homeobox 3 (*Nkx2-3*, a developmental gene), and CutC copper transporter (*Cutc*).[28]
414 *Eaaq19* is also the location of the *Cyp26a1-Myof* genes, and the human syntenic region is
415 associated with longevity, metabolic traits, and lipid profiles.[30,46,47] Another noteworthy
416 candidate in *Eaaq19* is *Chuk*, a regulator of mTORC2, that has been associated with age at
417 menopause.[41,48] Clearly, *Eaaq19* presents a complex and intriguing QTL related to the different
418 DNAm readouts, and potentially metabolic traits. Both *Eaaq19* and *Eaaq11* exemplify the major
419 challenge that follows when a genetic mapping approach leads to gene- and variant-dense
420 regions.[49,50] Both loci have several biologically relevant genes, and identifying the causal gene
421 (or genes) will require a more fine-scaled functional genomic dissection.

422 The gene expression analyses highlighted metabolic pathways related to lipids, glucose, and
423 proteins for both the univ.EAA and predicted-maxLS. Other enriched pathways were mitosis
424 and cell division, and immune processes, but this was specific to the positive transcriptomic
425 correlates. The more compelling evidence is for the cytochrome P450 genes, which are both
426 positional candidates, as well as expression correlates at the transcriptomic and proteomic
427 levels. These genes have high expression in liver, and have major downstream impact on
428 metabolism.[51-53] One caveat is that these CYP genes are part of a gene cluster in *Eaaq19* that
429 includes transcripts with *cis*-eQTLs (e.g., *Cyp2c66*, *Cyp2c39*, *Cyp2c68*), and the tight clustering of
430 the genes, and proximity of trait QTL and eQTLs may result in tight co-expression due to linkage
431 disequilibrium.[54] Nonetheless, the cytochrome genes in *Eaaq19* are strong candidate
432 modulators of EAA that calls for further investigation.

433 Aside from *Eaaq11* and *Eaaq19*, loci with evidence of consensus QTLs were also detected on
434 Chrs 1 and 3. We do not delve into these in the present work, but the Chr3 interval is near
435 genes associated with human epigenetic aging (*Ift80*, *Trim59*, *Kpna4*).[26,28] However, this QTL is
436 dispersed across a large interval, and the peak markers do not exactly overlap these human EAA
437 GWAS hits. While we have focused on *Eaaq11* and *Eaaq19*, these other loci also present
438 potentially important regions for EAA.

439 In summary, we have identified two main QTLs—*Eaaq11* and *Eaaq19*—that contribute to
440 variation in two DNAm readouts: EAA, and predicted-maxLS. *Eaaq11* contains several genes
441 with oncogenic properties (e.g., *Stxbp4*, *Erbb2*), while *Eaaq19* contains a dense cluster of
442 metabolic genes (e.g., *Elovl3*, *Chuk,* the cytochrome genes). We demonstrate that metabolic
443 profile and body weight are closely related to epigenetic aging. The convergence of evidence
444 from genetic and gene expression analyses indicates that genes involved in metabolism and
445 energy balance contribute to the age-dependent restructuring of the methylome, which in turn
446 forms the basis of the epigenetic predictors of aging and lifespan.

447 Materials and Methods

448 Biospecimen collection and processing

449 Samples for this study were selected from a larger colony of BXD mice that were housed in a
450 specific pathogen-free (SPF) facility at the University of Tennessee Health Science Center
451 (UTHSC). All animal procedures were in accordance with a protocol approved by the
452 Institutional Animal Care and Use Committee (IACUC) at the UTHSC. Detailed description of
453 housing conditions and diet can be found in.[18,19] Mice were given *ad libitum* access to water,
454 and either standard laboratory chow (Harlan Teklad; 2018, 18.6% protein, 6.2% fat, 75.2%
455 carbohydrates), or high-fat chow (Harlan Teklad 06414; 18.4% protein, 60.3% fat, 21.3%
456 carbohydrate). Animals were first weighed within the first few days of assignment to either
457 diets, and this was mostly but not always prior to introduction to HFD. Following this, animals
458 were weighed periodically, and a final time (BWF) when animals were humanely euthanized
459 (anesthetized with avertin at 0.02 ml per g of weight, followed by perfusion with phosphate-
460 buffered saline) at specific ages for tissue collection. The present work utilizes the biobanked
461 liver specimens that were pulverized and stored in -80 °C, and overlaps samples described in
462 Williams et al.[18] DNA was extracted using the DNeasy Blood & Tissue Kit from Qiagen. Nucleic

463 acid purity was inspected with a NanoDrop spectrophotometer, and quantified using a Qubit
464 fluorometer dsDNA BR Assay.

465 Methylation array, quality check, and entropy calculation

466 DNA samples from ~350 BXD mice were profiled on the Illumina HorvathHumanMethylChip40
467 array. Details of this array are described here Arneson et al.[15] The array contains probes that
468 target ~36K highly conserved CpGs in mammals. Over 33K probes map to homologous regions
469 in the mouse genome, and data from these were normalized using the SeSame method.[55]
470 Unsupervised hierarchical clustering was performed to identify outliers and failed arrays, and
471 those were excluded. We also performed strain verification as an additional quality check.
472 While majority of the probes were free of DNA sequence variants, we found 45 probes that
473 overlapped variants in the BXD family. We leveraged these as proxies for genotypes, and
474 performed a principal component analysis. The top principal component (PC1 and PC2)
475 segregated the samples by strain identity, and samples that did not cluster with the reported
476 strains were removed. After excluding outliers, failed arrays, and samples that failed strain
477 verification, the final liver DNAm data consisted of 339 samples.

478 For entropy calculation, we used 27966 probes that have been validated for the mouse genome
479 using calibration data generated from synthetic mouse DNA.[56] Shannon entropy was calculated
480 for each sample using the R package, "entropy" (v1.2.1) with method = "ML": maximum
481 likelihood.[57]

482 Clock estimation and maximum lifespan predictor

483 The development of the universal pan-tissue epigenetic clocks of age, and the universal
484 maximum lifespan predictor are described in Lu et al.[16], and Li et al.[17], respectively. For the
485 present work, we utilized the universal clock that predicts relative age, defined as individual age
486 relative to the maximum lifespan of its species, followed by inverse transformation to estimate
487 DNAmAge.[16]

488 The mouse specific clocks reported in this study were built with a similar strategy as described
489 in Lu et al.[16] using elastic net regression. First, a log-linear transformation was applied to the
490 chronological age using the function:

491 $$f(Age) = \begin{cases} \frac{Age}{1.2 + 0.06} + \log(1.2 + 0.06) - \frac{1.2}{1.2 + 0.06}, Age > 1.2 \\ \log(Age + 0.06), Age \leq 1.2 \end{cases}$$

492 This is similar to the age transformation described in the original Horvath pan-tissue human
493 clock, but with offset at 0.06, and adult mouse age at 1.2.[8] Following this, elastic net regression
494 was implemented to regress the transformed chronological age on the CpG beta-values in the
495 training set using the glmnet() function in R.[58] The training set included the mouse samples
496 reported in [16,17] but excluded the BXDs. The training was therefore BXD unbiased. For pan-
497 tissue clocks, all the samples were used, and for the liver specific clocks, the training was
498 limited to liver samples. The alpha was set at 0.5, and the optimal lamda parameter was
499 determined by 10-fold cross-validation (function cv.glmnet). This selected subsets of 'clock

CpGs' and coefficients (see **Table S2** for clock CpG IDs, intercept, and coefficients). DNAmAge was then calculated as:

$$DNAmAge = f^{-1}\left(\frac{b_0 + b_1 CpG_1 + b_2 CpG_2 + \cdots + b_i CpG_i}{b_0 + b_1 + b_2 + \cdots + b_i}\right)$$

where b_0 is the intercept, and b_1 to b_i are the coefficients, and CpG_1 to CpG_i denote the beta-values for the respective clock CpGs, and $f^{-1}()$ denotes the inverse function of $f()$.

Statistics

Statistical analyses between the epigenetic predictors and continuous variables (body weight, strain lifespan) were based on Pearson correlations, and t-test was used to evaluate the effect of categorical predictors (sex, diet).

Two metabolic traits were downloaded from the bioinformatics platform GeneNetwork 2 (GN2) [59]: (1) fasted serum glucose, and (2) fasted serum total cholesterol (more information on how to retrieve these data directly from GN2 are provided in **Data S14**). Association with metabolic traits was examined using multivariable linear regression (the R equations are provided in **Table S1**). For visualization, residuals for both the predictor and outcome variables were extracted after regressing on age, diet, and BWF using the R code: residuals(lm(~ age + diet + BWF)).

Longevity data (defined as age at natural death) was also downloaded from GN2 (**Data S13**).[19] Males were excluded and strain-by-diet lifespan summary statistics were derived. Only strain-by-diet groups with 5 or more observations were included in the correlational analyses with the epigenetic predictors.

Genetic analyses

The broad sense heritability within diet was estimated as the fraction of variability that was explained by background genotype.[20,60,61] For this, we applied a simple anova: aov(EAA ~ strain), and heritability was computed as: $H^2 = SSq_{strain}/(SSq_{strain} + SSq_{residual})$, where SSq_{strain} is the strain sum of squares, and $SSq_{residual}$ is the residual sum of squares.

All QTL mapping was done on the GN2 platform, and these traits can be accessed from this website [59] (trait accession IDs provided in **Data S14**). In the GN2 home page, the present set of BXD mice belongs to the **Group: BXD NIA Longevity Study**, and GN2 provides a direct interface to the genotype data. All QTL mapping was done for genotypes with minor allele frequency ≥ 0.05 using the genome-wide efficient mixed model association (GEMMA) algorithm,[31] which corrects for the BXD kinship matrix. For the EAA traits, diet, weight at 6 months, and final weight were fitted as cofactor. Chronological age had not correlation with EAA and this was not included as a cofactor (including age does not change the results). Genome-wide linkage statistics were downloaded for the full set of markers that were available from GN2 (3720 markers as of early 2021). For the combined p-values, QTL mapping was done separately using GEMMA for each EAA derived from all the unbiased mouse and universal clocks. Fisher's p-value combination was then applied to get the meta-p-value.[32] We used this method to simply highlight loci that had consistent linkage across the different EAA measures. QTL mapping for entropy, major covariates—age, diet, BW1, and BWF—were included as co-factors. QTL

538 mapping for predicted-maxLS was done without co-factors (including variables such as age,
539 weight, and diet did not change the results).

540 For marker specific linkage, we selected SNPs located at the peak QTL regions (DA0014408,
541 rs48062674, rs30567369), and grouped the BXDs by their genotypes (F1 hybrids and other
542 heterozygotes were excluded from this), and marker specific linkage was tested using ANOVA.
543 rs48062674 and rs30567369 are reference variants that is already catalogued in dbSNP,[62] and is
544 used as a marker in the QTL mapping. DA0014408.4 is an updated variant at a recombinant
545 region in the Chr11 interval and within the peak QTL interval.[20] Genotypes at these markers for
546 individual BXD samples are in **Data S1**.

547 For marker specific QTL analysis for EAA, we performed linear regression using the data in **Data**
548 **S1**. Heterozygotes at the respective markers were excluded, and we applied the following
549 regression model for each of the mouse and universal EAA traits separately: lm(EAA ~ genotype
550 + diet). To test the effect on body weight change, body weight data measured at approximately
551 4 (baseline), 6, 12, 18, and 24 months were downloaded from GN2 (**Data S14**). Detailed
552 description of these weight data are in Roy et al.[19] We then applied a mixed effects regression
553 model using the lme4 R package[63]: lmer(weight ~ age + diet + genotype + (1|ID)), where ID is
554 the identifier for individual mouse.

555 Bioinformatic tools for candidate gene selection

556 Sequence variation between B6 and D2 in the QTL intervals (Chr11:90–99 Mb, and Chr19:35–48
557 Mb) were retrieved from the Wellcome Sanger Institute Mouse Genomes Project database
558 (release 1505 for GRCm38/mm10).[64-66] Positional candidates were required to contain at least
559 one coding variant (missense and/or nonsense variants), or have non-coding variants with
560 evidence of *cis*-regulation in liver tissue of the BXDs. *Cis*-eQTLs for the candidate genes were
561 obtained from the liver RNA-seq data described in [18]. An interface to search and analyze this
562 transcriptome data is available from GN2, and is catalogued under *Group: BXD NIA Longevity*
563 *Study*; *Type: Liver mRNA*; and *Dataset: UTHSC BXD Liver RNA-seq (Oct 19) TMP Log2*. This data
564 was also used for the transcriptome-wide correlations analysis for univ.EAA in the 153 cases
565 that had both DNAm and RNA-seq data. We considered the top 2000 highest correlated
566 transcripts, and the list of transcripts were collapsed to a non-redundant list of gene symbols,
567 and this was uploaded to the DAVID Bioinformatics Database (version 6.8) for GO enrichment
568 analysis.[67,68] Similarly, proteome correlational analysis was carried out using the data: *Group:*
569 *BXD NIA Longevity Study*; *Type: Liver Proteome*; and *Dataset: EPFL/ETHZ BXD Liver Proteome*
570 *CD-HFD (Nov19)*.[18]

571 For human GWAS annotations, we navigated to the corresponding syntenic regions on the
572 human genome by using the coordinate conversion tool in the UCSC Genome Browser. The
573 Chr11 90–95 Mb interval on the mouse reference genome (GRCm38/mm10) corresponds to
574 human Chr17:50.14–55.75 Mb (GRCh38/hg38) (40.7% of bases; 100% span). The Chr11 95–99
575 Mb interval in the mouse corresponds to human Chr17:47.49–50.14 Mb (29.3% of bases, 57.9%
576 span), and Chr17:38.19–40.39 Mb (20.7% of bases, 44.1% span). Likewise, for the Chr19 QTL,
577 the mm10 35–40 Mb corresponds to hg38 Chr10:89.80–95.06 Mb (32.2% of bases, 89.2% span),
578 40–45 Mb corresponds to hg38 Chr10:95.23–100.98 Mb (46.6% of bases, 95.6% span), and 45–
579 48 Mb corresponds to hg38 Chr10:100.98–104.41 Mb (46.5% of bases, 100% span). We then

580 downloaded the GWAS data for these regions from the NHGRI-EBI GWAS catalogue,[33] and
581 retained the GWAS hits that were related to aging.

582 Data availability

583 The full microarray data will be released via NCBI's Gene Expression Omnibus upon official
584 publication. Genome annotations of the CpGs can be found on Github
585 https://github.com/shorvath/MammalianMethylationConsortium. Individual level BXD data are
586 available on www.genenetwork.org on FAIR+ compliant format; data identifiers, and way to
587 retrieve data are described in **Data S14**.

588 **Acknowledgement.** We thank the entire UTHSC BXD Aging Colony team, particularly
589 Casey J Chapman, Melinda S McCarty, Jesse Ingles, and everyone else who contributed to the
590 tissue harvest. We thank Evan G Williams for making the gene expression data readily available,
591 and to David Ashbrook for providing the BXD genotypes. We thank the GeneNetwork team,
592 especially Zach Sloan and Arthur Centeno, who have been extremely prompt and effective at
593 assisting with the GeneNetwork interface.

594 **Author contributions.** KM contributed to the data, conceived portion of the study, and
595 performed statistical analysis and drafted the article. ATL, CZL, AH contributed to the data
596 analysis and in computing the epigenetic clocks and predictor. JVS contributed to the lab work.
597 RWW conceived of the BXD Aging Colony, and provided access to the BXD biospecimen and
598 data. SH developed the array platform, and built the epigenetic clocks and predictor. All authors
599 contributed to, and approved the manuscript.

600 **Research funding.** This study was funded by the NIH NIA grants R21AG055841 and
601 R01AG043930

602 **Competing interests.** SH is a founder of the non-profit Epigenetic Clock Development
603 Foundation, which plans to license several of his patents from his employer UC Regents. The
604 other authors declare no conflicts of interest.

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

608 Reference:

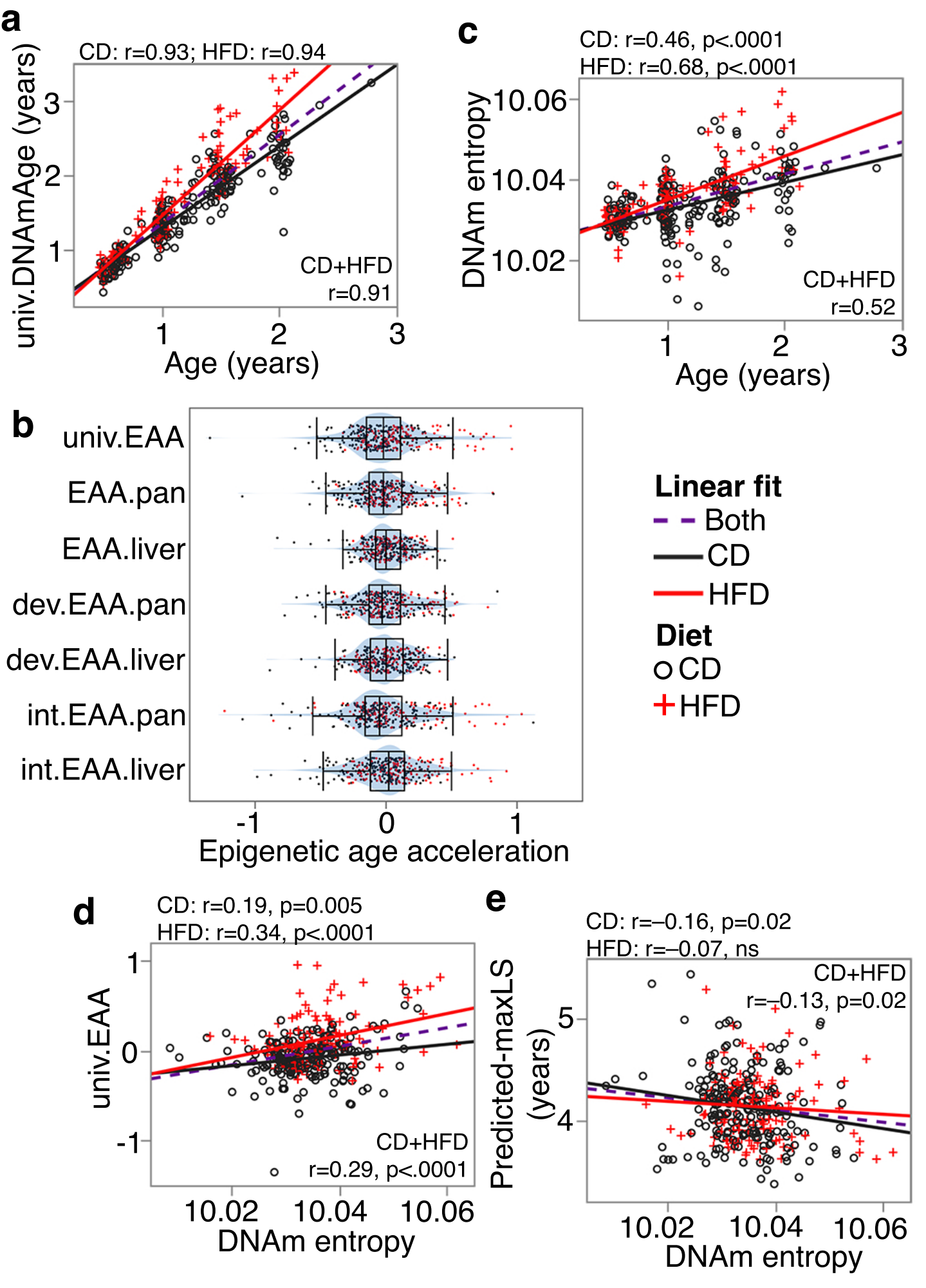
609 1 Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory
610 of ageing. *Nature reviews. Genetics* **19**, 371-384, doi:10.1038/s41576-018-0004-3
611 (2018).
612 2 Zannas, A. S. *et al.* Lifetime stress accelerates epigenetic aging in an urban, African
613 American cohort: relevance of glucocorticoid signaling. *Genome biology* **16**, 266,
614 doi:10.1186/s13059-015-0828-5 (2015).
615 3 Marioni, R. E. *et al.* The epigenetic clock is correlated with physical and cognitive fitness
616 in the Lothian Birth Cohort 1936. *International journal of epidemiology* **44**, 1388-1396,
617 doi:10.1093/ije/dyu277 (2015).

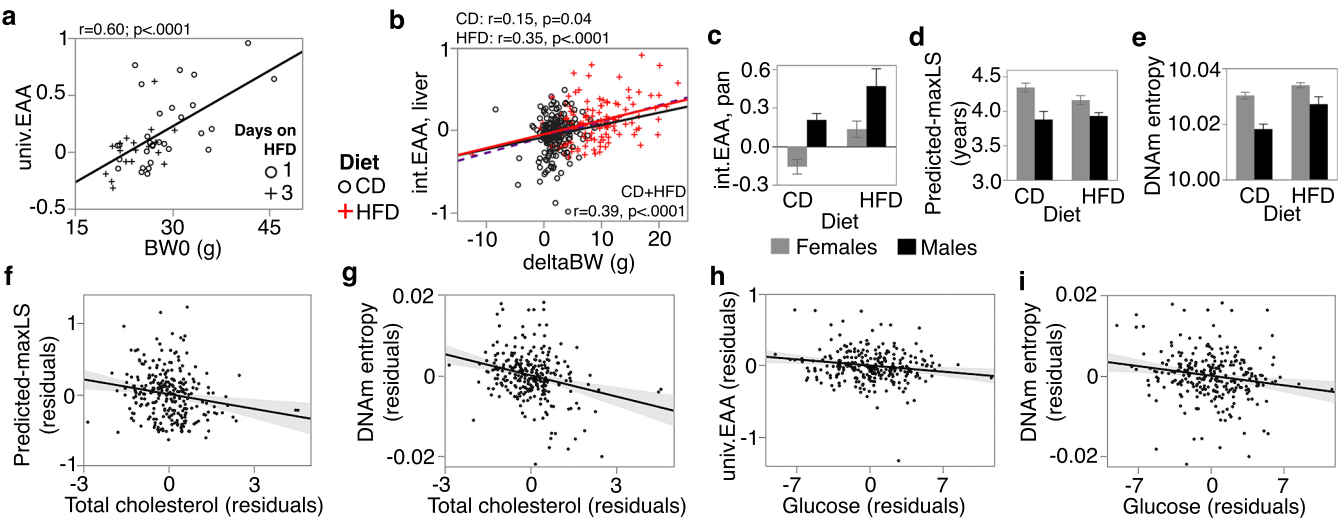
618 4 Dugue, P. A. *et al.* DNA methylation-based biological aging and cancer risk and survival:
619 Pooled analysis of seven prospective studies. *Int J Cancer* **142**, 1611-1619,
620 doi:10.1002/ijc.31189 (2018).
621 5 Lu, A. T. *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan.
622 *Aging* **11**, 303-327, doi:10.18632/aging.101684 (2019).
623 6 Ryan, J., Wrigglesworth, J., Loong, J., Fransquet, P. D. & Woods, R. L. A systematic review
624 and meta-analysis of environmental, lifestyle and health factors associated with DNA
625 methylation age. *The journals of gerontology. Series A, Biological sciences and medical*
626 *sciences*, doi:10.1093/gerona/glz099 (2019).
627 7 Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of
628 human aging rates. *Molecular cell* **49**, 359-367, doi:10.1016/j.molcel.2012.10.016
629 (2013).
630 8 Horvath, S. DNA methylation age of human tissues and cell types. *Genome biology* **14**,
631 R115, doi:10.1186/gb-2013-14-10-r115 (2013).
632 9 Petkovich, D. A. *et al.* Using DNA Methylation Profiling to Evaluate Biological Age and
633 Longevity Interventions. *Cell metabolism* **25**, 954-960 e956,
634 doi:10.1016/j.cmet.2017.03.016 (2017).
635 10 Stubbs, T. M. *et al.* Multi-tissue DNA methylation age predictor in mouse. *Genome*
636 *biology* **18**, 68, doi:10.1186/s13059-017-1203-5 (2017).
637 11 Wang, T. *et al.* Epigenetic aging signatures in mice livers are slowed by dwarfism, calorie
638 restriction and rapamycin treatment. *Genome biology* **18**, 57, doi:10.1186/s13059-017-
639 1186-2 (2017).
640 12 Horvath, S. *et al.* Epigenetic clock for skin and blood cells applied to Hutchinson Gilford
641 Progeria Syndrome and ex vivo studies. *Aging* **10**, 1758-1775,
642 doi:10.18632/aging.101508 (2018).
643 13 Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging*
644 **10**, 573-591, doi:10.18632/aging.101414 (2018).
645 14 Shireby, G. L. *et al.* Recalibrating the epigenetic clock: implications for assessing
646 biological age in the human cortex. *Brain* **143**, 3763-3775, doi:10.1093/brain/awaa334
647 (2020).
648 15 Arneson, A. *et al.* A mammalian methylation array for profiling methylation levels at
649 conserved sequences. *bioRxiv*, 2021.2001.2007.425637, doi:10.1101/2021.01.07.425637
650 (2021).
651 16 Lu, A. T. *et al.* Universal DNA methylation age across mammalian tissues. *bioRxiv*,
652 2021.2001.2018.426733, doi:10.1101/2021.01.18.426733 (2021).
653 17 Li, C. Z. *et al.* Epigenetic predictors of maximum lifespan and other life history traits in
654 mammals. *bioRxiv*, 2021.2005.2016.444078, doi:10.1101/2021.05.16.444078 (2021).
655 18 Williams, E. G. *et al.* Multi-Omic Profiling of the Liver Across Diets and Age in a Diverse
656 Mouse Population. *bioRxiv*, 2020.2008.2020.222968, doi:10.1101/2020.08.20.222968
657 (2020).
658 19 Roy, S. *et al.* Gene-by-environment modulation of lifespan and weight gain in the murine
659 BXD family. *Nat Metab* **3**, 1217-1227, doi:10.1038/s42255-021-00449-w (2021).
660 20 Ashbrook, D. G. *et al.* A Platform for Experimental Precision Medicine: The Extended
661 BXD Mouse Family. *Cell Syst*, doi:10.1016/j.cels.2020.12.002 (2021).

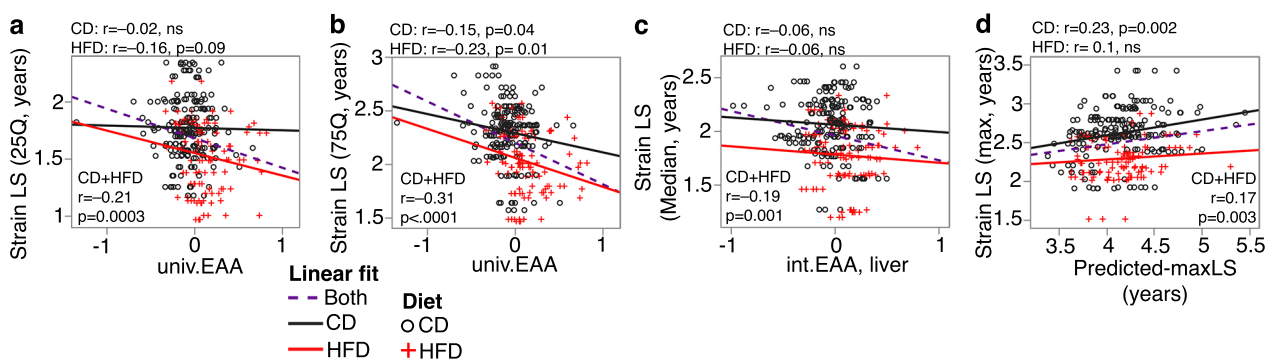
662 21 Peirce, J. L., Lu, L., Gu, J., Silver, L. M. & Williams, R. W. A new set of BXD recombinant
663 inbred lines from advanced intercross populations in mice. *BMC genetics* **5**, 7,
664 doi:10.1186/1471-2156-5-7 (2004).
665 22 de Haan, G. & Williams, R. W. A genetic and genomic approach to identify longevity
666 genes in mice. *Mechanisms of ageing and development* **126**, 133-138,
667 doi:10.1016/j.mad.2004.09.012 (2005).
668 23 Hsu, H. C. *et al.* Age-related thymic involution in C57BL/6J x DBA/2J recombinant-inbred
669 mice maps to mouse chromosomes 9 and 10. *Genes and immunity* **4**, 402-410,
670 doi:10.1038/sj.gene.6363982 (2003).
671 24 Lang, D. H. *et al.* Quantitative trait loci (QTL) analysis of longevity in C57BL/6J by DBA/2J
672 (BXD) recombinant inbred mice. *Aging clinical and experimental research* **22**, 8-19
673 (2010).
674 25 Sandoval-Sierra, J. V. *et al.* Body weight and high-fat diet are associated with epigenetic
675 aging in female members of the BXD murine family. *Aging cell*, e13207,
676 doi:10.1111/acel.13207 (2020).
677 26 Gibson, J. *et al.* A meta-analysis of genome-wide association studies of epigenetic age
678 acceleration. *PLoS genetics* **15**, e1008104, doi:10.1371/journal.pgen.1008104 (2019).
679 27 Lu, A. T. *et al.* GWAS of epigenetic aging rates in blood reveals a critical role for TERT.
680 *Nature communications* **9**, 387, doi:10.1038/s41467-017-02697-5 (2018).
681 28 McCartney, D. L. *et al.* Genome-wide association studies identify 137 genetic loci for
682 DNA methylation biomarkers of aging. *Genome biology* **22**, 194, doi:10.1186/s13059-
683 021-02398-9 (2021).
684 29 Benton, M. C. *et al.* Methylome-wide association study of whole blood DNA in the
685 Norfolk Island isolate identifies robust loci associated with age. *Aging* **9**, 753-768,
686 doi:10.18632/aging.101187 (2017).
687 30 Yashin, A. I. *et al.* Genetics of Human Longevity From Incomplete Data: New Findings
688 From the Long Life Family Study. *The journals of gerontology. Series A, Biological*
689 *sciences and medical sciences* **73**, 1472-1481, doi:10.1093/gerona/gly057 (2018).
690 31 Zhou, X. & Stephens, M. Efficient multivariate linear mixed model algorithms for
691 genome-wide association studies. *Nature methods* **11**, 407-409,
692 doi:10.1038/nmeth.2848 (2014).
693 32 Peirce, J. L., Broman, K. W., Lu, L. & Williams, R. W. A simple method for combining
694 genetic mapping data from multiple crosses and experimental designs. *PloS one* **2**,
695 e1036, doi:10.1371/journal.pone.0001036 (2007).
696 33 Catalog, G. <https://www.ebi.ac.uk/gwas/> (
697 34 Horvath, S. *et al.* Obesity accelerates epigenetic aging of human liver. *Proceedings of the*
698 *National Academy of Sciences of the United States of America* **111**, 15538-15543,
699 doi:10.1073/pnas.1412759111 (2014).
700 35 Nevalainen, T. *et al.* Obesity accelerates epigenetic aging in middle-aged but not in
701 elderly individuals. *Clinical epigenetics* **9**, 20, doi:10.1186/s13148-016-0301-7 (2017).
702 36 Xie, H. *et al.* Genome-wide quantitative assessment of variation in DNA methylation
703 patterns. *Nucleic acids research* **39**, 4099-4108, doi:10.1093/nar/gkr017 (2011).

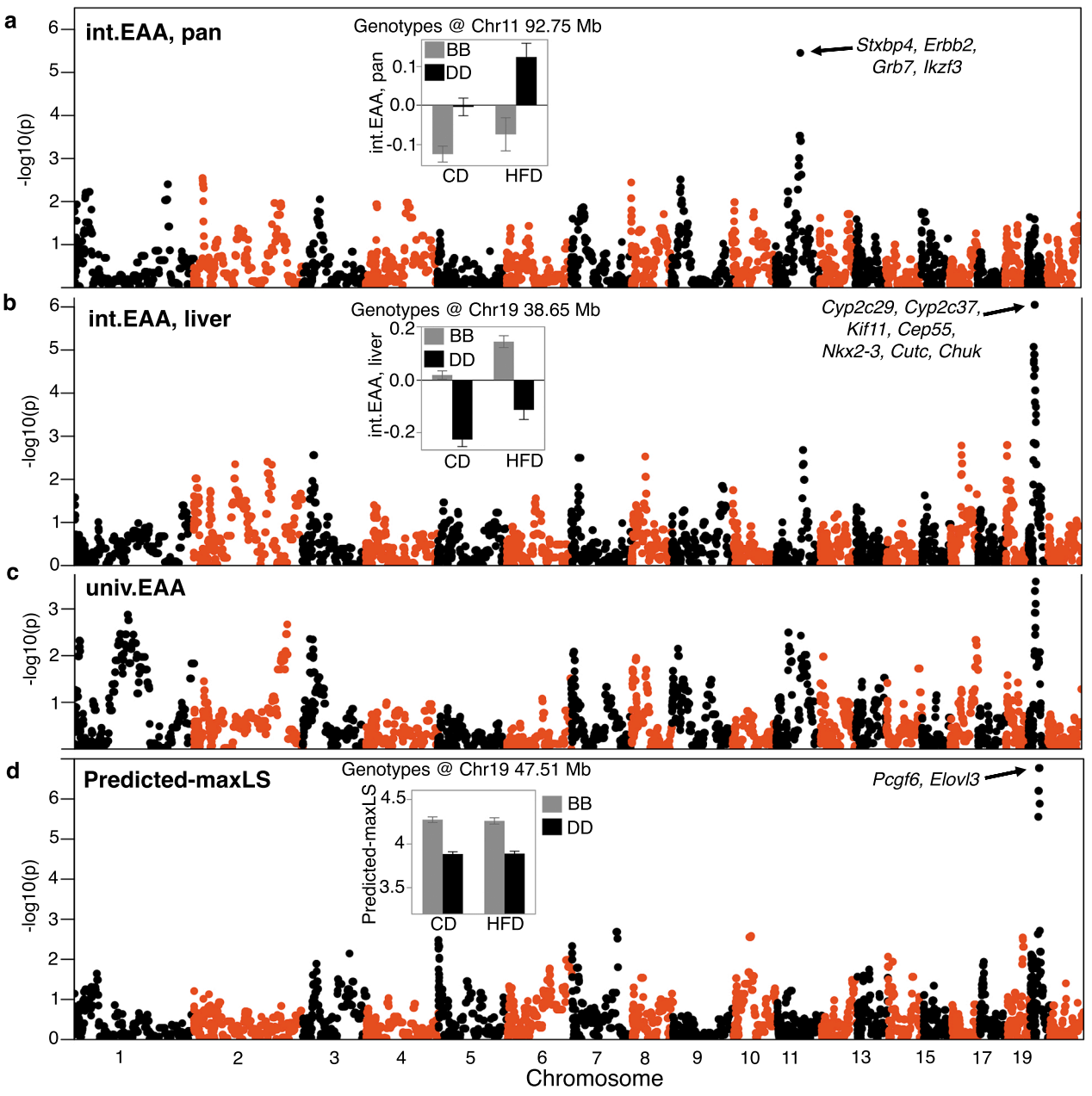
704 37 Hayflick, L. Entropy explains aging, genetic determinism explains longevity, and
705 undefined terminology explains misunderstanding both. *PLoS genetics* **3**, e220,
706 doi:10.1371/journal.pgen.0030220 (2007).
707 38 Holman, G. D. A new deadly Syn? *Curr Biol* **9**, R735-737, doi:10.1016/s0960-
708 9822(99)80471-0 (1999).
709 39 Michailidou, K. *et al.* Association analysis identifies 65 new breast cancer risk loci.
710 *Nature* **551**, 92-94, doi:10.1038/nature24284 (2017).
711 40 Rokudai, S. *et al.* STXBP4 regulates APC/C-mediated p63 turnover and drives squamous
712 cell carcinogenesis. *Proceedings of the National Academy of Sciences of the United*
713 *States of America* **115**, E4806-E4814, doi:10.1073/pnas.1718546115 (2018).
714 41 Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power.
715 *American journal of human genetics* **104**, 65-75, doi:10.1016/j.ajhg.2018.11.008 (2019).
716 42 Perry, J. R. *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for
717 age at menarche. *Nature* **514**, 92-97, doi:10.1038/nature13545 (2014).
718 43 Albertsen, H. M. *et al.* A physical map and candidate genes in the BRCA1 region on
719 chromosome 17q12-21. *Nature genetics* **7**, 472-479, doi:10.1038/ng0894-472 (1994).
720 44 Kauraniemi, P. & Kallioniemi, A. Activation of multiple cancer-associated genes at the
721 ERBB2 amplicon in breast cancer. *Endocr Relat Cancer* **13**, 39-49,
722 doi:10.1677/erc.1.01147 (2006).
723 45 Tanaka, S. *et al.* Coexpression of Grb7 with epidermal growth factor receptor or
724 Her2/erbB2 in human advanced esophageal carcinoma. *Cancer research* **57**, 28-31
725 (1997).
726 46 de Vries, P. S. *et al.* Multiancestry Genome-Wide Association Study of Lipid Levels
727 Incorporating Gene-Alcohol Interactions. *Am J Epidemiol* **188**, 1033-1054,
728 doi:10.1093/aje/kwz005 (2019).
729 47 Richardson, T. G. *et al.* Evaluating the relationship between circulating lipoprotein lipids
730 and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian
731 randomisation analysis. *PLoS Med* **17**, e1003062, doi:10.1371/journal.pmed.1003062
732 (2020).
733 48 Xu, Y. *et al.* IKK interacts with rictor and regulates mTORC2. *Cell Signal* **25**, 2239-2245,
734 doi:10.1016/j.cellsig.2013.07.008 (2013).
735 49 Gallagher, M. D. & Chen-Plotkin, A. S. The Post-GWAS Era: From Association to Function.
736 *American journal of human genetics* **102**, 717-730, doi:10.1016/j.ajhg.2018.04.002
737 (2018).
738 50 Lappalainen, T. Functional genomics bridges the gap between quantitative genetics and
739 molecular biology. *Genome research* **25**, 1427-1431, doi:10.1101/gr.190983.115 (2015).
740 51 Olona, A. *et al.* Epoxygenase inactivation exacerbates diet and aging-associated
741 metabolic dysfunction resulting from impaired adipogenesis. *Mol Metab* **11**, 18-32,
742 doi:10.1016/j.molmet.2018.03.003 (2018).
743 52 Schuck, R. N. *et al.* The cytochrome P450 epoxygenase pathway regulates the hepatic
744 inflammatory response in fatty liver disease. *PloS one* **9**, e110162,
745 doi:10.1371/journal.pone.0110162 (2014).

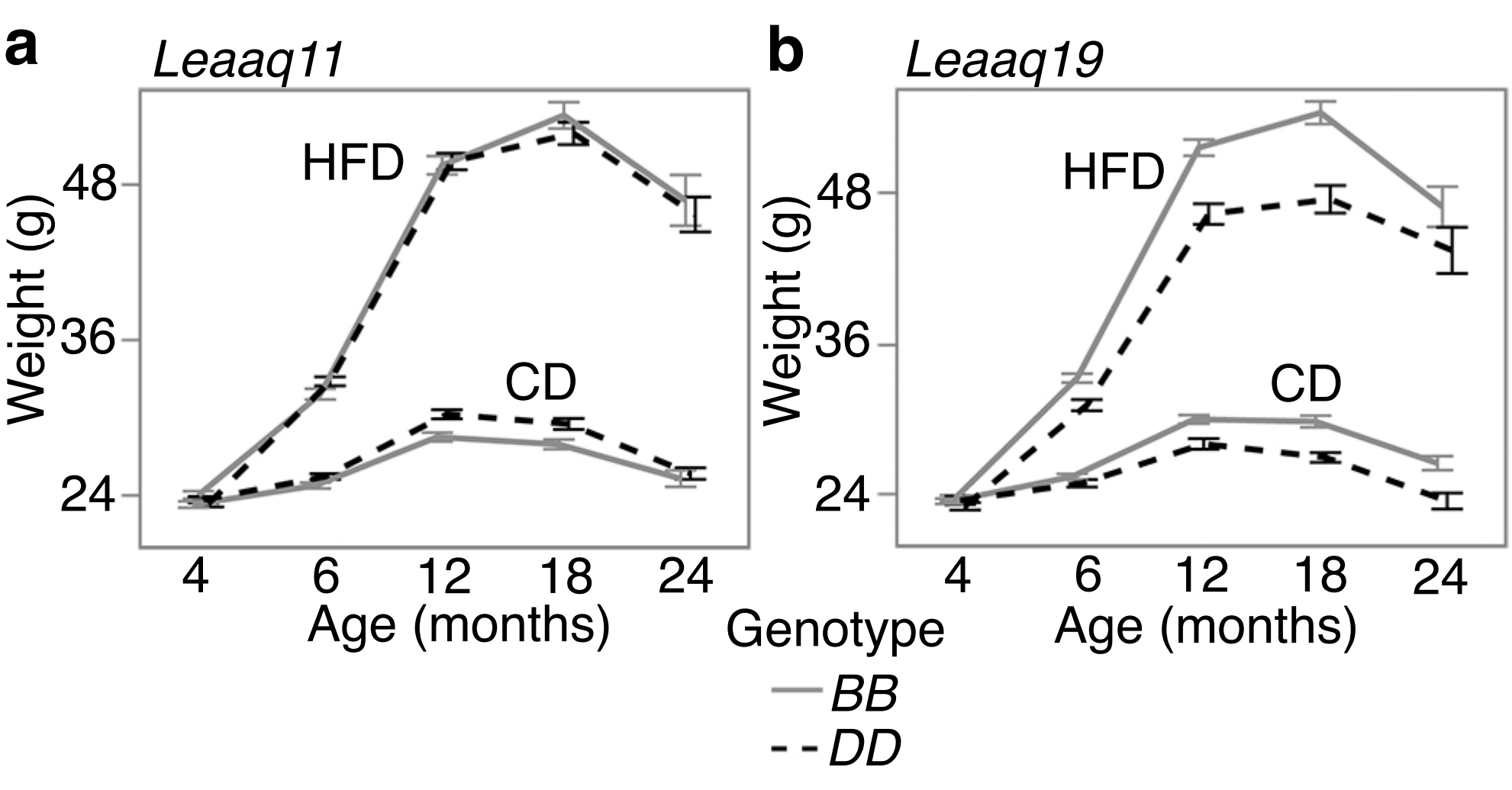
53 Wang, Q. *et al.* Time serial transcriptome reveals Cyp2c29 as a key gene in hepatocellular carcinoma development. *Cancer Biol Med* **17**, 401-417, doi:10.20892/j.issn.2095-3941.2019.0335 (2020).
54 Mozhui, K. *et al.* Dissection of a QTL hotspot on mouse distal chromosome 1 that modulates neurobehavioral phenotypes and gene expression. *PLoS genetics* **4**, e1000260, doi:10.1371/journal.pgen.1000260 (2008).
55 Zhou, W., Triche, T. J., Jr., Laird, P. W. & Shen, H. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic acids research* **46**, e123, doi:10.1093/nar/gky691 (2018).
56 Haghani, A. *et al.* DNA Methylation Networks Underlying Mammalian Traits. *bioRxiv*, 2021.2003.2016.435708, doi:10.1101/2021.03.16.435708 (2021).
57 Hausser, J. & Strimmer, K. Entropy Inference and the James-Stein Estimator, with Application to Nonlinear Gene Association Networks. *J. Mach. Learn. Res.* **10**, 1469–1484 (2009).
58 Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* **33**, 1-22 (2010).
59 *GeneNetwork 2*, <http://genenetwork.org> (
60 Ashbrook, D. G. *et al.* Born to Cry: A Genetic Dissection of Infant Vocalization. *Front Behav Neurosci* **12**, 250, doi:10.3389/fnbeh.2018.00250 (2018).
61 Belknap, J. K. Effect of within-strain sample size on QTL detection and mapping using recombinant inbred mouse strains. *Behavior genetics* **28**, 29-38 (1998).
62 Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic acids research* **29**, 308-311, doi:10.1093/nar/29.1.308 (2001).
63 Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: Linear mixed-effects models using Eigen and S4. . *R package version 1.1-7, http://CRAN.R-project.org/package=lme4*.
64 *Wellcome Sanger Institute Mouse Genome Project*, <https://www.sanger.ac.uk/sanger/Mouse_SnpViewer/rel-1505> (
65 Keane, T. M. *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289-294, doi:10.1038/nature10413 (2011).
66 Yalcin, B. *et al.* Sequence-based characterization of structural variation in the mouse genome. *Nature* **477**, 326-329, doi:10.1038/nature10432 (2011).
67 Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* **37**, 1-13, doi:10.1093/nar/gkn923 (2009).
68 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).

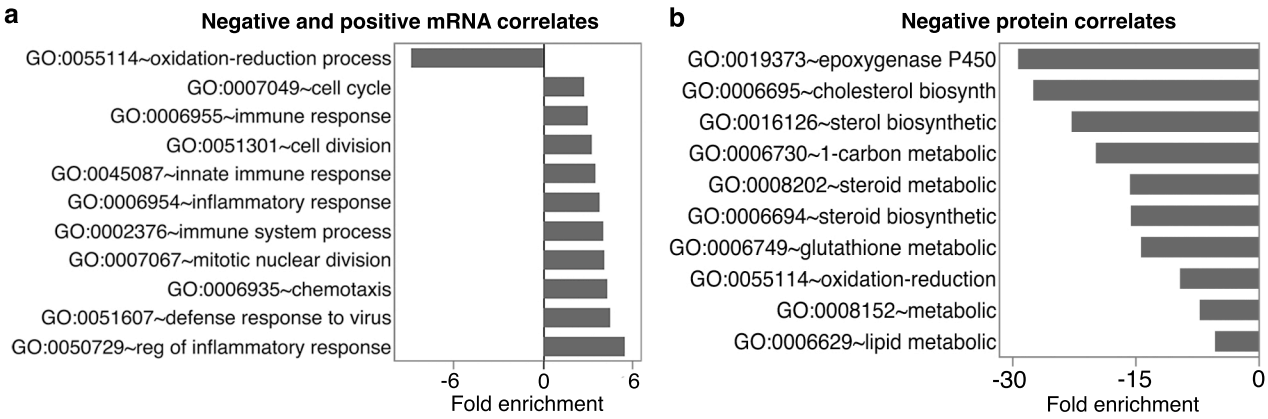

a
CD: r=0.93; HFD: r=0.94
univ.DNAmAge (years)
Age (years)
CD+HFD
r=0.91
c
CD: r=0.46, p<.0001
HFD: r=0.68, p<.0001
DNAm entropy
Age (years)
CD+HFD
r=0.52
b
univ.EAA
EAA.pan
EAA.liver
dev.EAA.pan
dev.EAA.liver
int.EAA.pan
int.EAA.liver
Epigenetic age acceleration
Linear fit
Both
CD
HFD
Diet
CD
HFD
d
CD: r=0.19, p=0.005
HFD: r=0.34, p<.0001
univ.EAA
DNAm entropy
CD+HFD
r=0.29, p<.0001
e
CD: r=–0.16, p=0.02
HFD: r=–0.07, ns
Predicted-maxLS (years)
DNAm entropy
CD+HFD
r=–0.13, p=0.02


a
r=0.60; p<.0001
univ.EAA
BW0 (g)
Days on HFD
1
3
b
CD: r=0.15, p=0.04
HFD: r=0.35, p<.0001
int.EAA, liver
deltaBW (g)
Diet
CD
HFD
CD+HFD
r=0.39, p<.0001
c
int.EAA, pan
CD
HFD
Diet
Females
Males
d
Predicted-maxLS (years)
CD
HFD
Diet
e
DNAm entropy
CD
HFD
Diet
f
Predicted-maxLS (residuals)
Total cholesterol (residuals)
g
DNAm entropy (residuals)
Total cholesterol (residuals)
h
univ.EAA (residuals)
Glucose (residuals)
i
DNAm entropy (residuals)
Glucose (residuals)


a
CD: r=−0.02, ns
HFD: r=−0.16, p=0.09
Strain LS (25Q, years)
CD+HFD
r=−0.21
p=0.0003
univ.EAA
b
CD: r=−0.15, p=0.04
HFD: r=−0.23, p= 0.01
Strain LS (75Q, years)
CD+HFD
r=−0.31
p<.0001
univ.EAA
c
CD: r=−0.06, ns
HFD: r=−0.06, ns
Strain LS (Median, years)
CD+HFD
r=−0.19
p=0.001
int.EAA, liver
d
CD: r=0.23, p=0.002
HFD: r= 0.1, ns
Strain LS (max, years)
CD+HFD
r=0.17
p=0.003
Predicted-maxLS (years)
Linear fit
Both
CD
HFD
Diet
CD
HFD


a
int.EAA, pan
Genotypes @ Chr11 92.75 Mb
BB
DD
int.EAA, pan
0.1
0.0
-0.1
CD
HFD
Stxbp4, Erbb2,
Grb7, Ikzf3
-log10(p)
b
int.EAA, liver
Genotypes @ Chr19 38.65 Mb
BB
DD
int.EAA, liver
0.2
0.0
-0.2
CD
HFD
Cyp2c29, Cyp2c37,
Kif11, Cep55,
Nkx2-3, Cutc, Chuk
-log10(p)
c
univ.EAA
-log10(p)
d
Predicted-maxLS
Genotypes @ Chr19 47.51 Mb
BB
DD
Predicted-maxLS
4.5
4
3.5
CD
HFD
Pcgf6, Elovl3
-log10(p)
1
2
3
4
5
6
7
8
9
10
11
13
15
17
19
Chromosome


a
Leaaq11
HFD
CD
Weight (g)
48
36
24
4
6
12
18
24
Age (months)
b
Leaaq19
HFD
CD
Weight (g)
48
36
24
4
6
12
18
24
Age (months)
Genotype
BB
DD


a
Negative and positive mRNA correlates
GO:0055114~oxidation-reduction process
GO:0007049~cell cycle
GO:0006955~immune response
GO:0051301~cell division
GO:0045087~innate immune response
GO:0006954~inflammatory response
GO:0002376~immune system process
GO:0007067~mitotic nuclear division
GO:0006935~chemotaxis
GO:0051607~defense response to virus
GO:0050729~reg of inflammatory response
-6
0
6
Fold enrichment
b
Negative protein correlates
GO:0019373~epoxygenase P450
GO:0006695~cholesterol biosynth
GO:0016126~sterol biosynthetic
GO:0006730~1-carbon metabolic
GO:0008202~steroid metabolic
GO:0006694~steroid biosynthetic
GO:0006749~glutathione metabolic
GO:0055114~oxidation-reduction
GO:0008152~metabolic
GO:0006629~lipid metabolic
-30
-15
0
Fold enrichment