Research Paper

Epigenetic drift underlies epigenetic clock signals, but displays distinct responses to lifespan interventions, development, and cellular dedifferentiation

Emily M. Bertucci-Richter[1,2], Ethan P. Shealy[1,2,3], Benjamin B. Parrott[1,2,3]

[1]Savannah River Ecology Laboratory, University of Georgia, Aiken, SC 29802, USA
[2]Eugene P. Odum School of Ecology, University of Georgia, Athens, GA 30602, USA
[3]Interdisciplinary Toxicology Program, University of Georgia, Athens, GA 30602, USA

Correspondence to: Emily M. Bertucci-Richter; **email:** embertucci@gmail.com, https://orcid.org/0000-0002-8348-0024
Keywords: epigenetic aging, epigenetic drift, epigenetic rejuvenation, lifespan, DNA methylation
Received: August 3, 2023 **Accepted:** December 1, 2023 **Published:**

Copyright: © 2024 Bertucci-Richter et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Changes in DNA methylation with age are observed across the tree of life. The stereotypical nature of these changes can be modeled to produce epigenetic clocks capable of predicting chronological age with unprecedented accuracy. Despite the predictive ability of epigenetic clocks and their utility as biomarkers in clinical applications, the underlying processes that produce clock signals are not fully resolved, which limits their interpretability. Here, we develop a computational approach to spatially resolve the within read variability or "disorder" in DNA methylation patterns and test if age-associated changes in DNA methylation disorder underlie signals comprising epigenetic clocks. We find that epigenetic clock loci are enriched in regions that both accumulate and lose disorder with age, suggesting a link between DNA methylation disorder and epigenetic clocks. We then develop epigenetic clocks that are based on regional disorder of DNA methylation patterns and compare their performance to other epigenetic clocks by investigating the influences of development, lifespan interventions, and cellular dedifferentiation. We identify common responses as well as critical differences between canonical epigenetic clocks and those based on regional disorder, demonstrating a fundamental decoupling of epigenetic aging processes. Collectively, we identify key linkages between epigenetic disorder and epigenetic clocks and demonstrate the multifaceted nature of epigenetic aging in which stochastic processes occurring at non-random loci produce predictable outcomes.

INTRODUCTION

Changes in DNA methylation with age, a component of "epigenetic aging", are widely observed across the tree of life. Age-associated DNA methylation patterns manifest as two general phenomena; one leading to stereotypical shifts in mean methylation levels at individual cytosines that can be modeled to predict individual age with high accuracy [1], and the other leading to increased variability or "disorder" in DNA methylation states due to the erosion of the epigenetic landscape [2–4]. These phenomena are hypothesized to be linked as average methylation values of individual cytosines are reported to drift from hyper- or hypo-methylated (e.g., ≥80%, ≤20%) states to more intermediate levels (e.g., 20–80%) with age [5]. However, the extent to which age-associated changes to the DNA methylome reflect distinct or similar underlying processes remains unresolved.

Over the last decade, dozens of epigenetic clocks have been developed for a range of taxonomic groups including humans [6], rodents [7], fish [8], birds [9], and trees [10]. Epigenetic clocks are typically

constructed as linear models that predict chronological age or age-related phenotypes using mean methylation levels from a relatively small number of individual cytosines. The rate of epigenetic aging, measured as the discrepancy between chronological age and epigenetic age estimates, is associated with environmental conditions [6], life history traits (e.g., age at first menarche [11] and menopause [12]), and has become a widely used indicator of biological age and attendant disease risk [1, 13, 14]. More recently, epigenetic clocks have been applied to understanding epigenetic rejuvenation events occurring either naturally during early embryonic development or as a consequence of cellular reprogramming. Whereas epigenetic age estimates of induced pluripotent stem cells (iPSCs) are typically reset to zero [6], transient treatments with Yamanaka factors that do not fully induce dedifferentiation also reduce epigenetic age estimates and have been recognized as a promising anti-aging therapeutic avenue [15]. Kerepesi, et al. (2021) have also reported a period of epigenetic rejuvenation occurring during early development in which epigenetic age estimates decrease after conception until reaching a "ground zero" state coinciding with gastrulation [16]. Yet, age estimates derived from epigenetic clocks may not fully capture other facets of epigenetic aging, and here, we integrate multiple measures of age-associated DNA methylation patterns to examine these phenomena more broadly.

The mechanistic underpinnings of epigenetic clock signals are still unclear, but with millions of CpG dinucleotides in the genome [17], and minimal overlap of individual CpGs included across different epigenetic clocks [18], the "ticking" of epigenetic clocks is suggested to be the product of a more general epigenetic maintenance system than can be reflected by the clock sites alone [1]. Commonly referred to as epigenetic "drift", the failure of this maintenance system has many references in the recent literature [2, 5, 19–22]. Yet, despite an abundance of reports examining age-related epigenetic drift [5, 21, 23], a consensus definition is lacking, with studies often defining drift to mirror the analytical approach employed [24]. For example, "epiallele frequency" [25], "discordance" [26], "disorder" [3], "entropy" [27], and "heterogeneity" [3] have all been used to assess epigenetic drift and reflect different analytical approaches. Perhaps the most inclusive definition of epigenetic drift is a change in the status of DNA methylation over time [21, 23]. Yet, according to this definition, even programmed changes which guide developmental processes could be considered epigenetic drift, and it is likely more useful to define epigenetic drift as a stochastic, rather than a deterministic change in methylation states. One popular approach for assessing stochastic changes in methylation is using Shannon's Entropy [28]. Originating in information theory, this metric measures the amount of uncertainty in an occurrence or event. However, when applied to DNA methylation, Shannon's Entropy simply reflects average methylation values (whether genome wide or at a specific CpG) and is also likely influenced by heterogeneity among cells. Heterogeneity of epigenetic patterning *within* cells requires analyzing single cells or in the case of bisulfite sequencing experiments, can be inferred from linked CpGs occurring on individual reads [3, 29].

Herein, we apply novel read-based strategies to resolve age-associated epigenetic disorder across the mouse genome. By considering methylation states between individual CpGs and their immediate neighbors, we directly assess epigenetic disorder and investigate its relationship to epigenetic clock signals, embryonic development, lifespan interventions, and cellular reprogramming. Borrowing from the conceptual framework of Waddington's epigenetic landscape, we hypothesize that low levels of epigenetic disorder characterize robust epigenomic states and that gains in disorder occurring with age lead to "erosion" of this landscape [2, 22, 30–32]. We find that approximately 30% of the genome is disproportionately affected by age-related epigenetic disorder. Loci which act as predictors in conventional epigenetic clocks based on mean methylation levels appear to be enriched in regions that both accumulate and lose disorder with age, suggesting a direct link between epigenetic disorder dynamics and clock signals. We subsequently develop epigenetic clocks based on our regional disorder (RD) metric and compare age predictions with those produced using conventional epigenetic clocks and those based on entropy. Upon exploring the influences of development, lifespan interventions, and cellular dedifferentiation, we identify similarities as well as clear divergence between epigenetic clock signals based on either mean DNA methylation or regional DNA methylation disorder. Contrary to predictions based on prior studies, we find that disorder increases during early development and global levels of disorder are unaffected after cellular rejuvenation. Collectively, our findings suggest that DNA methylation disorder dynamics are a key contributor to epigenetic clock signals, yet also highlight a fundamental decoupling of disorder dynamics from canonical epigenetic aging that is likely to inform the potential of lifespan intervention strategies.

MATERIALS AND METHODS

Data acquisition

Reduced representation bisulfite sequencing (RRBS) data from 255 mouse samples were acquired from

NCBI's Sequence Read Archive (Accession: PRJNA319643). Individuals ranged in age from 0.67 to 35 months, and represented both sexes, four strains (DW/J × C3H/HEJ)/F2, (C57BL/6J × BALB/cByJ)/F2, B6D2F1 and C57BL/6), and two diets (standard and caloric restriction). This dataset included methylomes from whole blood samples, induced pluripotent stem cells (iPSCs) derived from kidney (n = 3) and lung (n = 3), as well as the fibroblasts they were derived from (n = 3 lung, n = 3 kidney). Sample collection and library preparation methods are detailed in [33].

Data processing

Raw sequence reads were trimmed of low-quality sequences using Trim Galore! (v0.6.5, options: --paired –rrbs –quality 25 –illumina). Trimmed reads were then aligned in paired end mode to a bisulfite index of the latest version of the mouse genome (GRCm39) using Bismark (v0.22.3), with mapping efficiency ranging from 54–70% among samples. Following alignment, reads were sorted by genomic coordinate, and converted to human readable SAM files using the Samtools (v1.10) functions 'sort' and 'view', respectively. The methylation call strings from each read were extracted in R (v3.6.1) using a custom R script. Reads with less than 2 CpGs were removed from the analysis. Each CpG within a methylation call string was then scored based on whether its methylation status matched the methylation status of its nearest neighbors. Because the first and last CpG on a string has only one nearest neighbor, the maximum disorder score is one (1), while each CpG in the internal part of a string has two nearest neighbors (one upstream and one downstream), giving a maximum disorder score of two (2).

Calculation of disorder

The proportion of disordered neighbor pairs (PDN) was calculated on a per read basis by taking the proportion of neighbor pairs within the read that were disordered (i.e. methylation state differed) over the total number of neighbor pairs within the read. Practically, this was calculated as follows:

$$\text{PDN} = \frac{sum\ of\ disordered\ neighabor\ pairs}{sum\ of\ total\ neighabor\ pairs}$$

Calculation of regional disorder and methylation

Due to differences in coverage across individuals, we normalized our metric of disorder across 200 bp windows of the genome, subsequently referred to as regional disorder (RD; Figure 1A). To measure RD, we binned the genome into 200 bp windows using the Bedtools (v2.26.0) function 'makewindows' and used the Bedtools 'map' function to average the per-read PDN, methylation, and CpG density for all reads for which >51% of the read mapped to a specific window, preventing reads from being represented in more than one region. Regional methylation (RM) was calculated using the mean proportion of methylated cytosines within each region. Regions with less than five reads per sequencing run were excluded from analysis, and data from separate sequencing runs were merged together on a per individual basis using a weighted average based on the number of reads from each run. We then removed regions which were not present in at least 80% of all 255 samples.

Calculation of regional entropy

Regional entropy (RE) was calculated for each 200 bp window as follows:

$$\text{RE} = -RM \times log2(RM) - (1 - RM) \times log2(1 - RM)$$

Age-associated disorder, methylation, and entropy

To test if disorder increased with age, we selected a subset of whole blood methylomes from 153 male, C57BL/6 mice fed a standard diet, with individuals ranging from 0.67 to 35 months of age. Using this subset, we performed individual Spearman correlations between age and both RD and RM with a false discovery rate (FDR) correction for multiple comparisons using the corr.test function from the package psych in R [34]. Regions with a correlation coefficient ≥0.5 and an FDR corrected p-value < 0.05 were considered to gain disorder or methylation with age, and those with a correlation coefficient ≤−0.5 and p-value < 0.05 were considered to lose disorder or methylation with age.

Calculation of global disorder

For each sample we calculated global disorder using the mean RD values of those regions passing the filtering approach outlined above, which allowed us to directly compare disorder between individuals despite differences in coverage or depth of coverage across the genome. We also calculated global disorder using only regions which displayed any modest gain disorder with age (correlation coefficient ≥0.25; n = 45,668) and only regions which lost disorder with age (correlation coefficient ≤−0.25; n = 3,789). We then modeled the relationships between all three global disorder metrics and age using the lme package in R, and age-adjusted global disorder was calculated using the residuals from the quadratic relationship between global disorder and age.

Genomic enrichment of age-associated disorder

We then classified each 200 bp region by its genomic localization according to annotations of genes, introns, exons, CpG density, promoters, enhancers, transcription factor (TF) binding sites, CTCF binding sites, polycomb repressive complex 2 targets (PRC2), and Petkovich epigenetic clock sites [33]. Coordinates for genes, introns, and exons were used as listed by the most recent Refseq annotation of the mouse genome (GRCm39) with genes considered as the entire interval between transcription start and end coordinates. Coordinates of promoters, enhancers, TF binding sites and CTCF binding sites were determined using the Expression and Regulation annotation track from UCSC Genome Browser (GRCm39), CpG density was calculated using the average number of CpGs per informative read over the 100 possible CpGs per region, and PRC2 target genes were determined by previously published ChIP-seq data of PRC2 subunit binding in mouse ESCs [35], with any gene binding at least one PRC2 subunit being considered a PRC2 target. Coordinates from the Petkovich epigenetic clock [33] were translated to the current mouse genome annotation using NCBI's coordinate remapping service. Overlap between the 200 bp regions and each genomic category (at least 1 bp) was determined using a custom R script. Genomic enrichment was determined using binomial tests using all other covered loci as a background.

Gene ontology

Genes in regions determined to gain or lose RD with age were split into lists and compared against the background (all represented genes) for gene ontology enrichment using gProfiler. Genes spanning multiple age-associated regions were only counted once per gene list.

CpG methylation

Merged alignment files for each sample were also used to produce CpG methylation matrices using Bioconductor's MethylKit. Individual cytosines from opposite strands were merged into single CpGs (destrand = TRUE). Only CpGs which were covered at a depth of 10× reads across all 153 male, C57BL/6, standard diet samples were retained for further analysis.

Clock optimization

To compare our measures of disorder with epigenetic aging, we developed four different epigenetic clocks based on RD, RM, RE, and CpG methylation as predictors of chronological age. We used the glmnet package in R to select predictors using elastic net regularized regression and a leave-one-out cross validation (LOOCV) approach to assess model performance. Alpha values for each model were set to 0.5 (true elastic net) and lambda was cross validated across all samples in the training set for each individual model. Age estimates from test samples (i.e., remaining individuals not used to train the model) were used to assess the error of the clocks. To assess robustness of individual predictor sites, we extracted predictors from each model and determined the proportion of the 153 data-type-specific clocks each was included in. The robustness of CpG clocks was assessed by assigning individual CpGs to their respective genomic region, with each region being counted only once per clock iteration (i.e. multiple clock sites per region were not multiply counted.) We then determined the overlap between selected clock regions between RM, RD, RE, and CpG clocks.

Representative clock building

While LOOCV approaches provide a more inclusive estimate of predictive power, they do not provide a singular model appropriate for downstream applications. Thus, we constructed an additional set of clocks by randomly splitting samples into a training set (n = 14) and a test set (n = 39), which consisted of 2 or 3 individuals from each age class. We refer to these models as the "representative" clocks for each data type (Supplementary Figure 1), and the same training and test set were used for every data type.

Testing the effects of lifespan interventions

We then tested the effects of three lifespan interventions: caloric restriction beginning at 14-weeks of age, knock out of growth hormone receptor (GHR), and dwarfism using the representative clocks. The dataset consisted of 22 male and female individuals from mouse strain Snell Dwarf (DW/J × C3H/HEJ)/F2, split between Snell Dwarf (mutation in Pit-1 gene; n = 10) and their respective controls ("Snell Dwarf Control"; n = 12), 26 male and female individuals from strain (C57BL/6J × BALB/cByJ)/F2, split between GHR knock out (GHRKO, n = 11) and GHR wild type (GHR WT, n = 15), 22 male B6D2F1 mice, split between standard diet (n = 10) and caloric restriction (n = 12), and 20 male individuals from line C57BL/6 on a calorie restricted diet. Specific details of lifespan extending treatments can be found in Petkovich, et al. (2017) [33]. We calculated age adjusted global disorder, RD, RM, and RE as described above, and extracted CpG methylation information for each individual and then applied our representative epigenetic clocks from each data type to acquire epigenetic age estimates. Data from individuals experiencing lifespan interventions

was handled exactly as described above and any missing predictor was assigned a zero value so as to be dropped from the model. Differences between treatments were determined using a one-way ANOVA.

Disorder during de-differentiation and development

We analyzed DNA methylomes from mouse iPSCs and their respective lung (n = 6) or kidney (n = 6) fibroblast precursors. Specific details regarding the de-differentiation of fibroblasts can be found in Petkovich, et al. (2017) [33]. Datasets for the analysis of methylation dynamics across embryonic development were acquired from SRA accessions PRJNA150129 and PRJNA221793. Methods for sample preparation and sequencing in these datasets are detailed in Kerepesi, et al. (2021) [16] and Smith, et al. (2012) [36], respectively. Sample selection and filtering for loci comprising the Stubbs epigenetic clock [7] was modeled after the epigenetic clock methods in Kerepesi, et al. (2021) [16] to reproduce reported results with the traditional CpG-based approach. This included removal of samples retaining the polar bodies, as well as those derived from pre-fertilization gametes and ESCs. A similar sample-selection strategy was utilized for the region-based metrics, but the filtering strategy instead followed that outlined earlier in this paper for RD, RM, and RE. Overall, 36 samples were included in the window-based analyses, and 38 in the CpG-based Stubbs clock (due to differences in filtering requirements between the two approaches). Developmental stages represented in the 'early' developmental group ranged from zygote to ICM (approximated to 0.5–3.5 days after [16]), with the 'late' group consisting of embryonic and extraembryonic tissue from E6.5 and E7.5 embryos.

Age adjusted global disorder, and epigenetic age estimates for embryonic samples were calculated using representative RD, RM, RE clocks and the Stubbs CpG clock [7] as described for the lifespan intervention experiments. As data originated from two different datasets and consisted of different tissues than those used to train representative clocks, age adjusted global disorder and epigenetic age predictions were normalized within their respective datasets. Differences between the epigenetic ages of iPSCs were determined using a two-way ANOVA with cell type (iPSC or fibroblast) and tissue (kidney or lung) as predictors. To further investigate the role of disorder during development and de-differentiation, we performed two-tailed t-tests to determine differences in RD occurring after de-differentiation (fibroblast vs iPSC), or across development. For this analysis, we grouped both fibroblast types (kidney and lung) to compare against the iPSCs, as well as grouping the developmental datasets into early (E0.5–3.5) and late (E6.5–7.5) development. Given the especially low sample size for the iPSC dataset (n = 12), we also removed any regions with missing values. P-values from t-tests were corrected using FDR, via the function p.adjust in R. Significant differences between groups were determined by an adjusted p-value ≤ 0.05 and a mean difference in disorder between groups of at least |0.1|. Significant differences in disorder were then further characterized into regions which gained disorder during development or de-differentiation, and regions which lost disorder during development or de-differentiation. To determine the effect size of any given region on epigenetic age prediction, we took the mean difference between groups (either de-differentiation or development) at that region and multiplied it by the beta value for that region used in the RD epigenetic clock model. The effect size for each region was then normalized to the percent of the total effect size for the clock.

Data availability

The data that support the findings of this study are openly available in NCBI's Sequence Read Archive (Accession: PRJNA319643). Examples of custom R scripts used to calculate regional disorder are available on GitHub (https://github.com/embertucci/epigenetic-disorder).

RESULTS

Disorder in DNA methylation patterns are strongly correlated with age on a regional and global scale. Of the 249,015 regions assessed, RD was significantly correlated with age in 76,353 regions (30.7%), with RD increasing with age in 70,094 genomic regions (91.8%; Figure 1B) and decreasing with age in 6,259 genomic regions (8.2%; Figure 1B). The average RD across all regions, or global disorder, increases with chronological age according to a quadratic relationship (R^2 = 0.51, p < 2.2e-16; Figure 1C). Consistent with increases and decreases in RD being driven by distinct processes, regions experiencing increases in RD (cor ≥0.25) display a quadratic relationship to age (R^2 = 0.74, p < 2.2e-16; Figure 1D), whereas regions experiencing decreasing RD (cor ≤0.25) display a linear relation with age (R^2 = 0.77, p < 2.2e-16; Figure 1E).

With the exception of the Y chromosome, every chromosome incurs significant age-related accumulation of RD (Figure 1F). Given that more than 30% of the genome experiences age-associated RD, only sites with a p-value ≤ 0.05 and a correlation coefficient greater ≥0.5 (n = 4149) or ≤−0.5 (n = 286) were considered as age-associated for enrichment tests, with all other regions considered background

(n = 244,580). Regions accumulating RD with age were significantly enriched in genes (p < 2.2e-16; Figure 1G) and promoters (p < 2.2e-16; Figure 1H), depleted in enhancers (p < 2.2e-16; Figure 1I), enriched in PRC2 target genes (p < 2.2e-16; Figure 1J), and depleted in both transcription factor (p = 0.00015; Figure 1K) and CTCF (p = 1.35e-12; Figure 1L) binding sites. Although enrichment scores were less robust, regions losing RD with age were significantly enriched in genes (p = 8.725e-05; Figure 1G), promoters (p = 7.073e-11; Figure 1H), and PRC2 target genes (p = 0.020; Figure 1J) and were depleted in enhancers (p = 0.030; Figure 1I). The mean CpG density did not differ in age-associated regions when compared to background (Figure 1M). Genes which accumulate disorder with age (n = 1,635) were significantly enriched in 552 different biological processes (GO:BP) with the most significant terms relating to nervous system development and differentiation (Supplementary Table 1; Figure 1N). Genes losing disorder with age (n = 197) were enriched in 14 different biological processes, with the most significant terms relating to multicellular organismal development (Supplementary Table 2; Figure 1N).

We next examined the relationship between RD and regional averages of Shannon's Entropy, a commonly used measure of epigenetic drift. Regional entropy (RE) is calculated directly from mean methylation values, and thus has a strong relationship to regional methylation (RM), even when regional values are averaged across all individuals (Figure 2A). However, in loci where RE reaches its maximum (RE = 1, mean

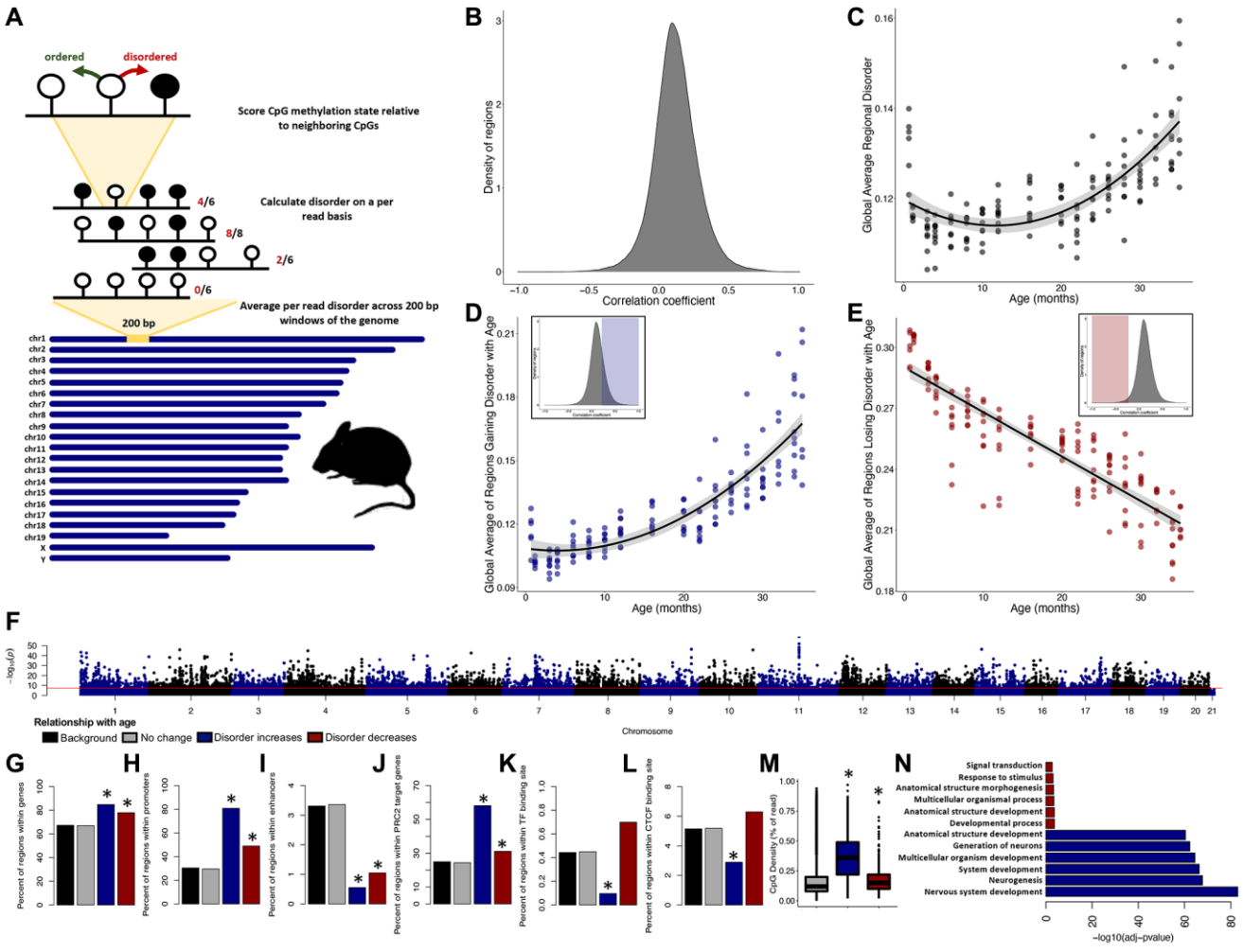


Figure 1. Epigenetic disorder increases across the murine lifespan. (**A**) Diagram of the approach for measuring regional disorder (RD). (**B**) Density of all genomic regions assessed with respect to their Spearman correlation coefficients between RD and age. (**C**) The relationship between global disorder and age in mice. (**D**) Average RD across all regions that gain disorder with age (correlation coefficient ≥0.25), or (**E**) lose disorder with age (correlation coefficient ≤−0.25. (**F**) Manhattan plot of the distribution of FDR corrected p-values of the relationship between RD and age. Red line marks a commonly used genome wide significance value of $p = 5 \times 10^{-8}$. Enrichment of age associated RD in genes (**G**), promoters (**H**), enhancers (**I**), PRC2 target genes (**J**), transcription factor binding sites (**K**), CTCF binding sites (**L**), and average CpG density (**M**). (**N**) The six most significant gene ontology biological processes (GO:BP) for regions gaining or losing disorder with age. Regions which gain disorder with age are shown in blue and regions which lose disorder with age are shown in red.

methylation = 50%), RD spans from fully ordered to fully disordered (RD = 0–1). We found that the relationship between average RD and average RE is best explained by a quadratic relationship (R^2 = 0.83, p < 2.2e-16; Figure 2B) with increasing RE generally indicating increases in RD. While the relationship between global RD and global RE does not change with age (Supplementary Figure 2), the relationship becomes increasingly variable at greater values of RE (Supplementary Figure 3). Age dependent changes to RM and RD are linked as 78.7% (n = 38,926) of regions experiencing modest age-associated RD (cor ≥ |0.25|) also incur modest age-associated RM (cor ≥ |0.25|); however, the remaining 21.3% (n = 10,531) of age-associated changes in RD do not correspond with RM, and 34.1% (n = 20,152) of age-associated RM occur independently of changes in RD (Figure 2C).

We also aimed to understand how signals underlying epigenetic clocks relate to epigenetic disorder. Interestingly, there is a clear enrichment of Petkovich epigenetic clock loci in regions which increase and decrease in RD with age (Figure 3A), with the absolute correlation coefficient of RD and age being significantly higher in Petkovich epigenetic clock regions when compared to those not included in the clock (p < 2e-16; Figure 3B). However, 37 of the 90 total clock CpGs fall into the same 200 bp genomic region. To more thoroughly resolve the relationship between epigenetic clock signals and epigenetic disorder, we built a series of epigenetic clocks based on CpG, RM, RE, and RD states. Over the 153 LOO iterations for each clock type, there was no difference in absolute error across clocks, suggesting that each methylation metric is capable of predicting chronological age with equivalent accuracy (Figure 3C). Similarly, there was no difference in the mean absolute error produced by the representative clocks (Supplementary Figure 1).

To further compare the influence of methylation context on clock composition, we assessed the overlap of loci incorporated into each clock type as well as the frequency in which they were selected (referred to as robustness). Of the LOO iterations, the CpG clocks selected 312 different regions with an average robustness of 0.11 (Figure 3D), RM clocks selected 106 different regions with an average robustness of 0.05 (Figure 3E), RE clocks selected 330 different regions with an average robustness of 0.11 (Figure 3F), and RD clocks selected 483 different regions with an average robustness of 0.13 (Figure 3G). Interestingly, the mean absolute RD correlation coefficients for age were significantly higher for CpG, RM, RE and RD clock regions when compared to non-clock regions (CpG p < 2e-16, RM p < 2e-16, RE p < 2e-16, RD p < 2e-16; Figure 3H–3K). The majority of clock sites (86.5%) were specific to each clock type; however, seven regions were selected across all clock types. Pan-clock regions are all associated with genes (Map10, Nlrp5-ps, Rasef, Rnf220, Evx2, Gm21297, and Apba1), with five (71%) regions located within promoters, and all increase in disorder with age (Figure 3L). While all four datatypes produce low errors in age prediction (Figure 3C), the discordance of chronological age with the age prediction (or "delta epigenetic age") is highly correlated with age adjusted global disorder in all datatypes (CpG R^2 = 0.09, p = 7.53e-05; RM R^2 = 0.07, p = 0.0007, RE R^2 = 0.12, p = 6.89e-06, RD R^2 = 0.18; p = 3.00e-08; Figure 3M).

We then tested the influence of common lifespan manipulations on epigenetic age estimates across different clock types. Caloric restriction led to a

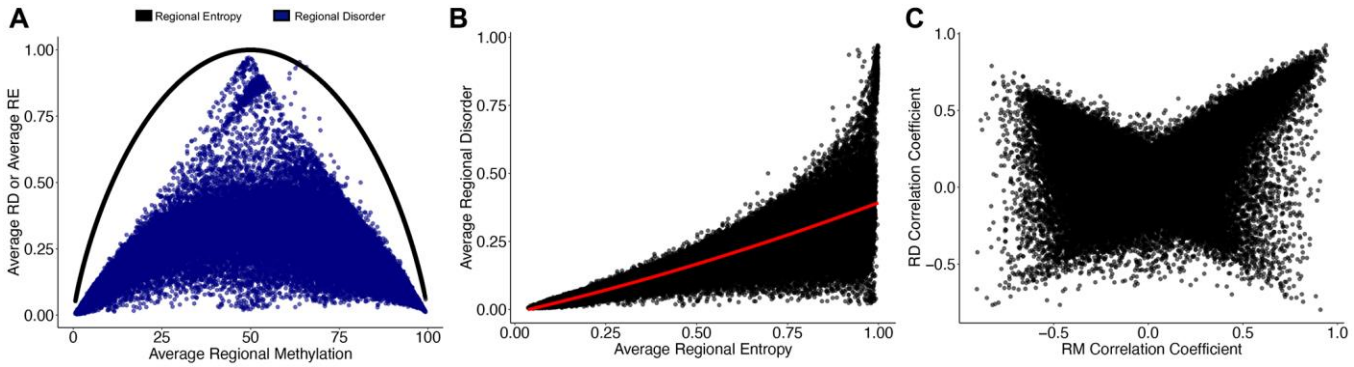


Figure 2. Regional disorder is distinct from Shannon's entropy and age-associated changes in mean methylation. (**A**) The relationship of regional entropy (RE; black line) with regional methylation and regional disorder (RD; blue dots) with regional methylation (RM). Data points show a single region averaged across all samples. (**B**) Relationship between RD and RE averaged across all samples. (**C**) Correlation coefficients of RD with age and RM with age across the 153 samples used to build the epigenetic clock. Regions which increase in RM or RD with age have positive correlation coefficients, regions which decrease in RM or RD with age have negative correlation coefficients.

reduction in age-associated RD, but the effect varied across strains. Male C57BL/6 mice fed a calorie restricted diet had significantly younger epigenetic ages when compared to controls as determined by all clock types (CpG p = 0.00043, RM p = 2.16e-07, RE p = 4.62e-06, RD p = 5.75e-08; Figure 4A). However, mean age adjusted global disorder appeared unaffected (p = 0.62; Figure 4B). Conversely, male B6D2F1 mice fed a calorie restricted diet only had significantly younger epigenetic ages as determined by the RM epigenetic clock (CpG p = 0.46, RM p = 0.018, RE p = 0.073, RD p = 0.39; Figure 4C). However, there was a slight trend

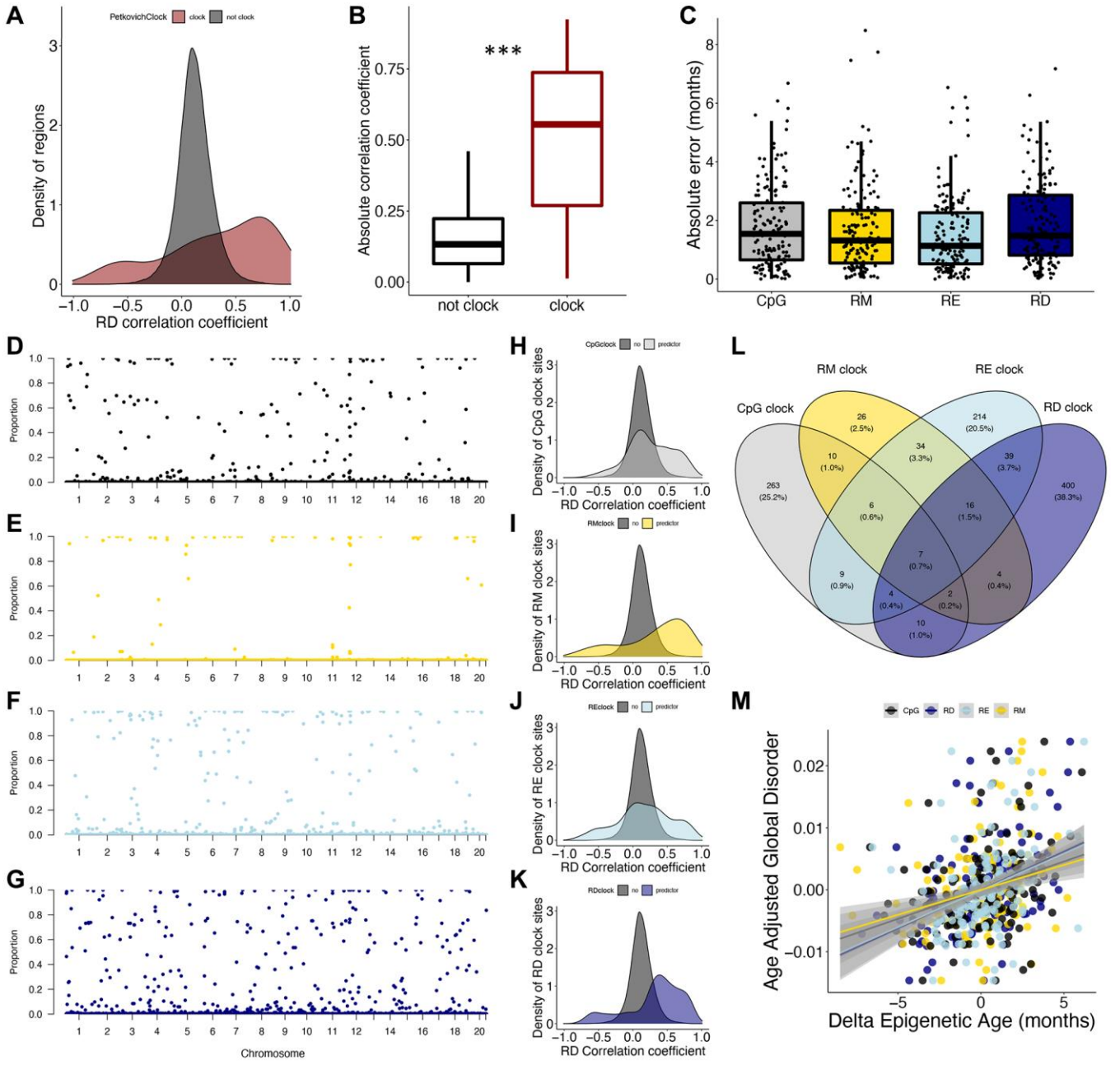


Figure 3. Epigenetic disorder underlies epigenetic clock signals. (**A**) Distribution of Petkovich epigenetic clock sites (red) across correlation coefficients between regional disorder (RD) and age. (**B**) Average absolute correlation coefficient between RD and age of regions which are included in the Petkovich epigenetic clock (red) compared to those which are not included. (**C**) Error of epigenetic age estimates produced by leave-one-out cross validation (LOOCV) for each data type. (**D**–**G**) Manhattan plots showing the robustness for each region (i.e., the proportion of clocks each region was selected in) across (**D**) CpG methylation (black), (**E**) regional methylation (RM; yellow), (**F**) regional entropy (RE; light blue), and (**G**) RD (dark blue) contexts. (**H**–**K**) Density plots showing the distribution of clock sites for each data type across correlation coefficients between RD and age. (**L**) Overlap between regions included in epigenetic clocks produced from each data type. (**M**) Relationship between delta epigenetic age (chronological age – predicted age) and age-adjusted global disorder for each data type.

for calorie restricted individuals to have greater mean age adjusted global disorder when compared to mice on a standard diet (p = 0.097; Figure 4D). Genetic interventions which extend lifespan resulted in a general decrease in epigenetic age. Snell Dwarf mice had significantly younger epigenetic ages when compared to controls according to the RD and RM clocks, but not the CpG or RE clocks (CpG p = 0.078, RM p = 0.0059, RE p = 0.23, RD p = 0.011; Figure 4E). Snell Dwarf mice also showed reduced mean age adjusted global disorder compared to control mice (p = 0.027; Figure 4F). GHR knock out also resulted in significantly younger epigenetic ages according to the RM and RE clocks, but not the CpG or RD clocks (CpG p = 0.088, RM p = 0.0080, RE p = 0.031, RD p = 0.64; Figure 4G), and no difference in age adjusted regional disorder was observed between GHRKO and control mice (p = 0.39; Figure 4H).

We next investigated the impacts of cellular dedifferentiation on epigenetic disorder by comparing DNA methylomes of iPSC cells and their differentiated precursors. A significant reduction in epigenetic age predictions after dedifferentiation was observed across all clock types except for RE (CpG p = 7.11e-08, RM p = 3.41e-06, RE p = 0.20, RD p = 3.93e-08; Figure 5A). Interestingly, there was no difference in the global disorder between kidney or lung fibroblasts when compared to their respective iPSCs (p = 0.28; Figure 5B). Given that dedifferentiation led to a reduction in epigenetic age estimates but did not affect global disorder, we sought to identify those regions in which RD is affected by dedifferentiation. Upon comparing RD across all fibroblasts and iPSCs, 26,512 regions significantly increase in RD after differentiation and 19,419 regions significantly decrease in disorder after dedifferentiation, but these regions do not disproportionately acquire age-associated RD relative to background (Figure 5C). Interestingly, the influence of dedifferentiation on RD epigenetic clock estimates are driven by differences in RD at just several clock sites (Figure 5D), with four regions contributing 35.7% of the overall effect.

Consistent with a previous report identifying an epigenetic rejuvenation event occurring during early development [16], we observed a significant decrease in epigenetic age predictions occurring between embryonic days 4 and 6 using the Stubbs CpG methylation epigenetic clock [7] (CpG p = 1.32e-08; Figure 5E) and the RM clock (RM p = 0.0019; Figure 5F). Conversely, we observe a significant increase in

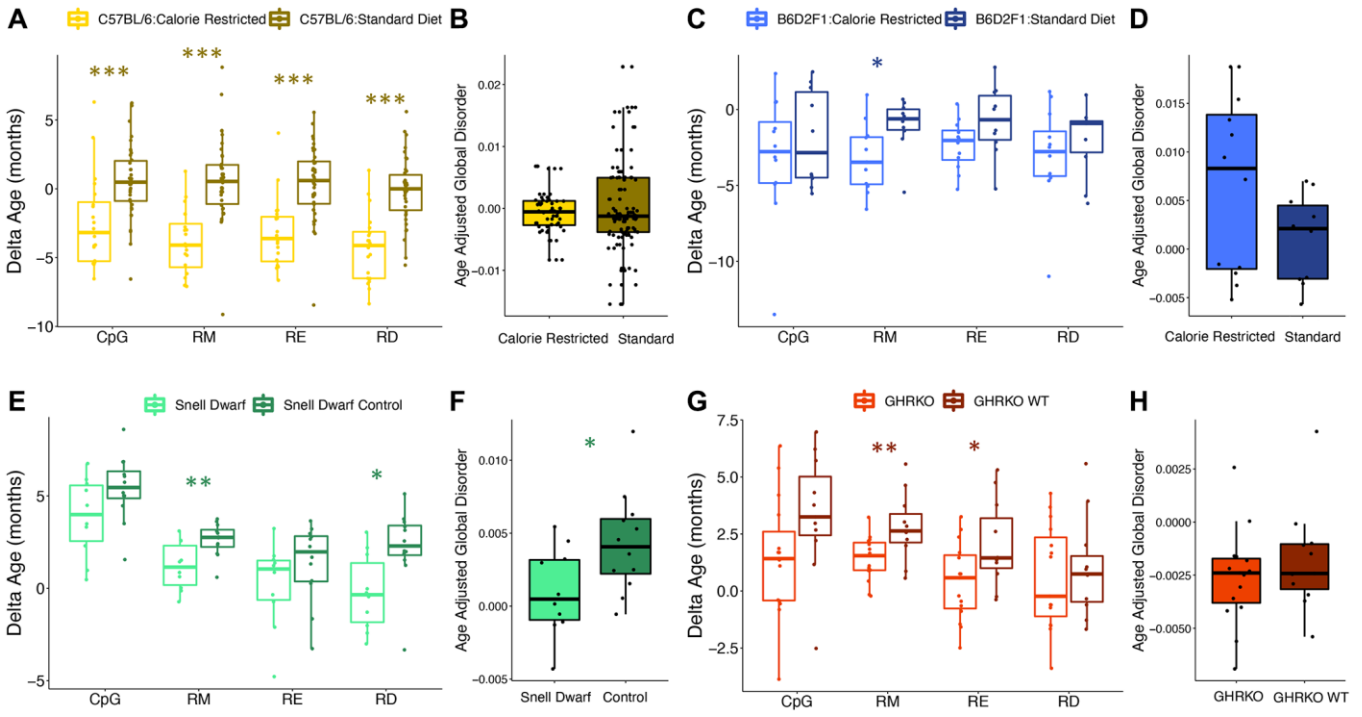


Figure 4. Epigenetic disorder is influenced by lifespan extending manipulations. The effect of caloric restriction in C57BL/6 mice on (**A**) epigenetic age predictions from each data type and (**B**) age-adjusted global disorder. The effect of caloric restriction in B6D2F1 mice on (**C**) epigenetic age predictions from each data type and (**D**) age-adjusted global disorder. Comparison of Snell dwarf and control mice on (**E**) epigenetic age predictions from each data type and (**F**) age-adjusted global disorder. The effect of growth hormone receptor knock-out (GHRKO) on (**G**) epigenetic age predictions from each data type and (**H**) age-adjusted global disorder. All plots show median, upper, and lower quartiles, and maximum and minimum. Outliers beyond 1.5 interquartile range are plotted.

epigenetic age predictions during early embryonic development when using the RD and RE clocks (RD p = 3.81e-05; RE p = 1.90e-08; Figure 5G, 5H). In addition, global disorder is strongly increased during development (Figure 5I). Upon comparing RD across all samples from embryonic days 0.5–3.5 (n = 24) and embryonic days 6.5–7.5 (n = 12), 31,687 regions significantly increase in RD during development and 368 regions significantly decrease in RD during development. These regions were not significantly enriched in regions with age-associated RD (Figure 5J). Similar to our findings examining the influence of de-differentiation, the predictions of the epigenetic clock appear to be driven strongly by differences in RD at just several clock sites, with one region having a total effect size of 2.51 months, contributing 14.6% of the difference in ages between groups (Figure 5K). Given that the clock has 86 of the regions included as predictors represented, we would expect each region to contribute just 1.16% to the overall effect size.

DISCUSSION

Epigenetic drift is broadly hypothesized to be a primary contributor to epigenetic aging. However, drift is a multifaceted phenomenon encompassing both stochastic and deterministic processes and is unlikely to be fully captured by a single metric. In this study, we report an approach for spatially resolving genomic patterns of DNA methylation disorder, which is distinct from

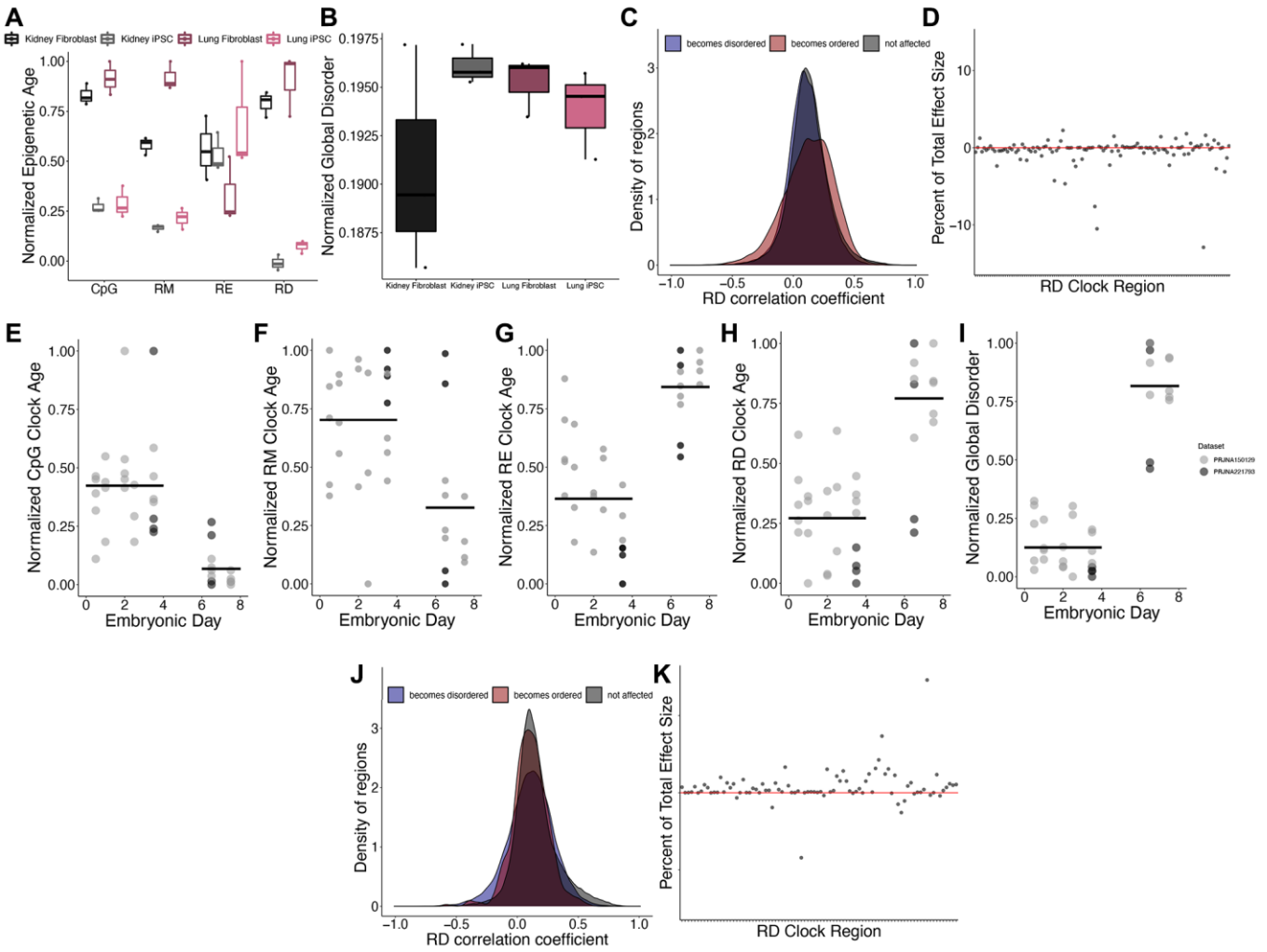


Figure 5. Epigenetic disorder during de-differentiation and development. (**A**) Epigenetic age predictions using each of the representative epigenetic clocks and (**B**) global disorder of kidney fibroblasts (black), kidney derived iPSCs (grey), lung fibroblasts (dark purple), and lung derived iPSCs (pink). Plot shows median, upper and lower quartiles, maximum, and minimum. Outliers beyond 1.5 interquartile range are plotted. (**C**) Distribution of regions which gain (blue) or lose (red) disorder after de-differentiation across correlation coefficients between regional disorder (RD) and age. (**D**) Effect sizes of de-differentiation on the RD epigenetic clock. Stubbs CpG methylation (**E**), RM (**F**), RE (**G**), and RD (**H**) epigenetic clock predictions of samples during embryonic development. (**I**) Global disorder of samples during embryonic development. (**J**) distribution of regions which gain (blue) or lose (red) disorder during early development across correlation coefficients between regional disorder (RD) and age. (**K**) Effect sizes of development on the RD epigenetic clock.

measures of both average methylation and entropy. Age-associated changes in regional disorder (RD) are found in approximately one third of the genome and generally reflect the accumulation of disordered methylation; however, the opposite pattern is observed in a subset of regions in which DNA methylation patterns become more ordered with age. Given that epigenetic drift or disorder is thought to be driven in part by stochastic processes, we hypothesize that the directionality of changes in RD represent different functional pathways. Yet, the specific biological mechanisms that mediate losses of RD with age remain unclear. Age-associated gains and losses of RD were disproportionately observed in coding regions, promoters, and regions harboring PRC2 target genes, and age-associated increases in RD were strongly enriched in developmental genes, especially those functioning in neural development. These gains in disorder with age support the deleteriome model of aging [37], wherein small deleterious errors accumulate in the epigenome without effect until later in life, when epigenomic stability is compromised [2]. We suggest that disorder accumulates across the genome until it reaches a critical threshold – this may explain why the majority of regions across the genome are characterized by relatively low disorder (RD <0.5; data not shown). The value of this hypothetical threshold and the factors which contribute to the accumulation of disorder have the potential to explain the rate of aging and possibly maximum lifespan across species [32].

Consistent with the hypothesis that disorder in DNA methylation patterns underlies signals in conventional CpG clocks (i.e., those based on mean CpG methylation levels), we find that loci comprising clocks constructed using RD, RM, RE, and CpG contexts are all enriched for regions in which disorder changes with age, and loci which are shared across clocks are involved biological processes such as cell cycle, tumor suppression, and development. Interestingly, Evx2 is shared across all clock contexts, and is frequently reported as being age-associated across a variety of vertebrate species [38–41]. However, we identify notable distinctions and minimal overlap across clock contexts. For example, while a subset of CpGs selected as predictors in an epigenetic clock were enriched in regions with age-associated disorder, many CpG clock sites also fell into regions lacking age-associated changes in disorder. Thus, while disorder underlies some components of traditional CpG epigenetic clocks, other components may be attributed to other processes like coordinated changes in methylation or cell type composition. By contrasting the effects of caloric restriction, genetic manipulations, cellular reprogramming, and development across different clock types, we further identify both similarities and clear distinctions according to DNA methylation context and genomic scale. For example, while traditional lifespan extending treatments, such as caloric restriction, broadly affect RD epigenetic clocks, there is no observable effects on global disorder. This may be because the magnitude of the effect of caloric restriction is too small to be seen when averaged across the entire genome. Similarly, while CpG and RM clocks demonstrate a "ground zero" occurring during mid-development [16], we see an increase in RE and RD clock predicted ages during the same period, suggesting that average methylation states may not fully reflect how the DNA methylome changes throughout development. Collectively, these findings demonstrate the connections between epigenetic drift and other aspects of epigenetic aging, while also highlighting a complexity that should be considered when assessing read outs from epigenetic clocks alone.

In mice, global disorder changes with age according to a quadratic function, with decreases in disorder occurring rapidly earlier in life prior to a steady increase with age. This pattern is consistent with previous findings of a quadratic relationship between global DNA methylation entropy and age in the naked-mole rat [42]. The initial high level of global disorder suggests that development, as well as aging, may be characterized by a disorganized epigenetic landscape – possibly due to a transitionary period between methylation states. Given the dynamic nature of the DNA methylome during development [6, 43, 44], it is likely that RD metrics, like other measures of DNA methylation that provide temporal snapshots, capture this transition as high disorder. While data from embryonic samples suggest that disorder increases during early development, the trajectory of global disorder throughout development, and whether it corresponds with previous findings of an epigenetic "ground zero" during development [16], will require a more complete developmental series to fully resolve.

Age estimates derived from epigenetic clocks are ideal for predicting chronological age (i.e., forensics, conservation and management applications [45, 46]) as well as identifying the consequences of accelerated epigenetic aging (i.e., biomarkers in biomedical approaches [1]). Yet, collapsing mean methylation levels into a single value presents challenges for understanding the drivers and biological pathways responsible for epigenetic aging. Given the push towards targeted, high-throughput approaches (e.g., bead-based assays) for acquiring data on age-associated methylation [38, 47], critical biological information is missed. While CpG level resolution has been integral in developing our understanding of epigenetic aging, clocks built using regionally averaged methylation perform with similar accuracy to those trained on

individual CpGs. We further demonstrate that the effect size of individual clock sites varies widely, and thus, changes in methylation states of just one or several clock loci can be misinterpreted as wholesale changes in epigenetic age. This is important especially when age estimates are compared across studies and different datasets. For example, we report that a single RD clock region accounted for nearly 10% of the difference in age estimation between fibroblasts and iPSCs. While the age predictions generated corroborate previous findings [6, 33], the inclusion (or exclusion) of this single region vastly changes our interpretation of the effects of de-differentiation on epigenetic age. Thus, using epigenetic clocks of any kind gives us a narrow, and potentially easily skewed, understanding of epigenetic aging at the genomic scale.

Overall, this study provides robust empirical evidence that epigenetic drift, as measured by epigenetic disorder, accumulates with age in non-random places of the mouse genome. Our analyses suggest that epigenetic disorder underlies aspects of traditional epigenetic clocks and highlights critical gaps in our interpretation of epigenetic aging. Although more work needs to be done to better resolve the drivers of epigenetic disorder – we provide an empirical basis for testing assumptions about this emerging phenomenon.

AUTHOR CONTRIBUTIONS

E.M.B and B.B.P. conceived and designed the study. E.M.B. and E.P.S. performed experimental analyses on previously published datasets. E.M.B, E.P.S., and B.B.P. wrote the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

FUNDING

This material is based upon work supported by the Department of Energy Office of Environmental Management under Award Number DE-EM0005228 to the University of Georgia Research Foundation. In addition, this work was supported by awards from the National Science Foundation (Award No. 2026210 to BBP) and the National Institutes of Health (Award No. 1R56AG078336-01 to BBP).

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness. Or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States.

REFERENCES

1. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet. 2018; 19:371–84. https://doi.org/10.1038/s41576-018-0004-3 PMID:29643443

2. Yang JH, Hayano M, Griffin PT, Amorim JA, Bonkowski MS, Apostolides JK, Salfati EL, Blanchette M, Munding EM, Bhakta M, Chew YC, Guo W, Yang X, et al. Loss of epigenetic information as a cause of mammalian aging. Cell. 2023; 186:305–26.e27. https://doi.org/10.1016/j.cell.2022.12.027 PMID:36638792

3. Landau DA, Clement K, Ziller MJ, Boyle P, Fan J, Gu H, Stevenson K, Sougnez C, Wang L, Li S, Kotliar D, Zhang W, Ghandi M, et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. Cancer Cell. 2014; 26:813–25. https://doi.org/10.1016/j.ccell.2014.10.012 PMID:25490447

4. Kane AE, Sinclair DA. Epigenetic changes during aging and their reprogramming potential. Crit Rev Biochem Mol Biol. 2019; 54:61–83. https://doi.org/10.1080/10409238.2019.1570075 PMID:30822165

5. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R, Chen M, Rajapakse I, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013; 49:359–67. https://doi.org/10.1016/j.molcel.2012.10.016 PMID:23177740

6. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013; 14:R115. https://doi.org/10.1186/gb-2013-14-10-r115 PMID:24138928

7. Stubbs TM, Bonder MJ, Stark AK, Krueger F, von Meyenn F, Stegle O, Reik W, and BI Ageing Clock Team. Multi-tissue DNA methylation age predictor in mouse. Genome Biol. 2017; 18:68. https://doi.org/10.1186/s13059-017-1203-5 PMID:28399939

8. Bertucci EM, Mason MW, Rhodes OE, Parrott BB. Exposure to ionizing radiation disrupts normal epigenetic aging in Japanese medaka. Aging (Albany NY). 2021; 13:22752–71. https://doi.org/10.18632/aging.203624 PMID:34644261

9. De Paoli-Iseppi R, Deagle BE, Polanowski AM, McMahon CR, Dickinson JL, Hindell MA, Jarman SN. Age estimation in a long-lived seabird (Ardenna tenuirostris) using DNA methylation-based biomarkers. Mol Ecol Resour. 2019; 19:411–25. https://doi.org/10.1111/1755-0998.12981 PMID:30576072

10. Gardner ST, Bertucci EM, Sutton R, Horcher A, Aubrey D, Parrott BB. Development of DNA methylation-based epigenetic age predictors in loblolly pine (Pinus taeda). Mol Ecol Resour. 2023; 23:131–44. https://doi.org/10.1111/1755-0998.13698 PMID:35957540

11. Binder AM, Corvalan C, Mericq V, Pereira A, Santos JL, Horvath S, Shepherd J, Michels KB. Faster ticking rate of the epigenetic clock is associated with faster pubertal development in girls. Epigenetics. 2018; 13:85–94. https://doi.org/10.1080/15592294.2017.1414127 PMID:29235933

12. Levine ME, Lu AT, Chen BH, Hernandez DG, Singleton AB, Ferrucci L, Bandinelli S, Salfati E, Manson JE, Quach A, Kusters CD, Kuh D, Wong A, et al. Menopause accelerates biological aging. Proc Natl Acad Sci U S A. 2016; 113:9327–32. https://doi.org/10.1073/pnas.1604558113 PMID:27457926

13. Xiao FH, Wang HT, Kong QP. Dynamic DNA Methylation During Aging: A "Prophet" of Age-Related Outcomes. Front Genet. 2019; 10:107. https://doi.org/10.3389/fgene.2019.00107 PMID:30833961

14. Oblak L, van der Zaag J, Higgins-Chen AT, Levine ME, Boks MP. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. Ageing Res Rev. 2021; 69:101348. https://doi.org/10.1016/j.arr.2021.101348 PMID:33930583

15. Sarkar TJ, Quarta M, Mukherjee S, Colville A, Paine P, Doan L, Tran CM, Chu CR, Horvath S, Qi LS, Bhutani N, Rando TA, Sebastiano V. Transient non-integrative expression of nuclear reprogramming factors promotes multifaceted amelioration of aging in human cells. Nat Commun. 2020; 11:1545. https://doi.org/10.1038/s41467-020-15174-3 PMID:32210226

16. Kerepesi C, Zhang B, Lee SG, Trapp A, Gladyshev VN. Epigenetic clocks reveal a rejuvenation event during embryogenesis followed by aging. Sci Adv. 2021; 7:eabg6082. https://doi.org/10.1126/sciadv.abg6082 PMID:34172448

17. Babenko VN, Chadaeva IV, Orlov YL. Genomic landscape of CpG rich elements in human. BMC Evol Biol. 2017; 17:19. https://doi.org/10.1186/s12862-016-0864-0 PMID:28251877

18. Bergsma T, Rogaeva E. DNA Methylation Clocks and Their Predictive Capacity for Aging Phenotypes and Healthspan. Neurosci Insights. 2020; 15:2633105520942221. https://doi.org/10.1177/2633105520942221 PMID:32743556

19. Martin GM. Epigenetic gambling and epigenetic drift as an antagonistic pleiotropic mechanism of aging. Aging Cell. 2009; 8:761–4. https://doi.org/10.1111/j.1474-9726.2009.00515.x PMID:19732045

20. Maegawa S, Lu Y, Tahara T, Lee JT, Madzo J, Liang S, Jelinek J, Colman RJ, Issa JJ. Caloric restriction delays age-related methylation drift. Nat Commun. 2017; 8:539. https://doi.org/10.1038/s41467-017-00607-3 PMID:28912502

21. Martin GM. Epigenetic drift in aging identical twins. Proc Natl Acad Sci U S A. 2005; 102:10413–4. https://doi.org/10.1073/pnas.0504743102 PMID:16027353

22. Hayano M, Yang JH, Bonkowski MS, Amorim JA, Ross JM, Coppotelli G, Griffin P, Chew YC, Guo W, Yang X, Vera DL, Salfati EL, Das A, et al. DNA Break-Induced Epigenetic Drift as a Cause of Mammalian Aging. bioRxiv. 2019. https://doi.org/10.2139/ssrn.3466338

23. Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, Diez J, Sanchez-Mut JV, Setien F, Carmona FJ, Puca AA, Sayols S, Pujana MA, et al. Distinct DNA methylomes of newborns and centenarians. Proc Natl Acad Sci U S A. 2012; 109:10522–7. https://doi.org/10.1073/pnas.1120658109 PMID:22689993

24. Scherer M, Nebel A, Franke A, Walter J, Lengauer T, Bock C, Müller F, List M. Quantitative comparison of within-sample heterogeneity scores for DNA methylation data. Nucleic Acids Res. 2020; 48:e46. https://doi.org/10.1093/nar/gkaa120 PMID:32103242

25. Rakyan VK, Beck S. Epigenetic variation and inheritance in mammals. Curr Opin Genet Dev. 2006; 16:573–7.
https://doi.org/10.1016/j.gde.2006.09.002
PMID:17005390

26. Farré P, Jones MJ, Meaney MJ, Emberly E, Turecki G, Kobor MS. Concordant and discordant DNA methylation signatures of aging in human blood and brain. Epigenetics Chromatin. 2015; 8:19.
https://doi.org/10.1186/s13072-015-0011-y
PMID:25977707

27. Mozhui K, Lu AT, Li CZ, Haghani A, Sandoval-Sierra JV, Wu Y, Williams RW, Horvath S. Genetic loci and metabolic states associated with murine epigenetic aging. Elife. 2022; 11:e75244.
https://doi.org/10.7554/eLife.75244
PMID:35389339

28. Jenkinson G, Pujadas E, Goutsias J, Feinberg AP. Potential energy landscapes identify the information-theoretic nature of the epigenome. Nat Genet. 2017; 49:719–29.
https://doi.org/10.1038/ng.3811
PMID:28346445

29. Xie H, Wang M, de Andrade A, Bonaldo Mde F, Galat V, Arndt K, Rajaram V, Goldman S, Tomita T, Soares MB. Genome-wide quantitative assessment of variation in DNA methylation patterns. Nucleic Acids Res. 2011; 39:4099–108.
https://doi.org/10.1093/nar/gkr017
PMID:21278160

30. Noble D. Conrad Waddington and the origin of epigenetics. J Exp Biol. 2015; 218:816–8.
https://doi.org/10.1242/jeb.120071
PMID:25788723

31. Waddington CH. The strategy of the genes. J Chem Inf Model. 1957; 1–259.

32. Bertucci EM, Parrott BB. Is CpG Density the Link between Epigenetic Aging and Lifespan? Trends Genet. 2020; 36:725–7.
https://doi.org/10.1016/j.tig.2020.06.003
PMID:32624337

33. Petkovich DA, Podolskiy DI, Lobanov AV, Lee SG, Miller RA, Gladyshev VN. Using DNA Methylation Profiling to Evaluate Biological Age and Longevity Interventions. Cell Metab. 2017; 25:954–60.e6.
https://doi.org/10.1016/j.cmet.2017.03.016
PMID:28380383

34. Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 2.2.9. Northwestern University, Evanston, Illinois. 2022.

35. Ngan CY, Wong CH, Tjong H, Wang W, Goldfeder RL, Choi C, He H, Gong L, Lin J, Urban B, Chow J, Li M, Lim J, et al. Chromatin interaction analyses elucidate the roles of PRC2-bound silencers in mouse development. Nat Genet. 2020; 52:264–72.
https://doi.org/10.1038/s41588-020-0581-x
PMID:32094912

36. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature. 2012; 484:339–44.
https://doi.org/10.1038/nature10960
PMID:22456710

37. Gladyshev VN. Aging: progressive decline in fitness due to the rising deleteriome adjusted by genetic, environmental, and stochastic processes. Aging Cell. 2016; 15:594–602.
https://doi.org/10.1111/acel.12480
PMID:27060562

38. Lu AT, Fei Z, Haghani A, Robeck TR, Zoller JA, Li CZ, Lowe R, Yan Q, Zhang J, Vu H, Ablaeva J, Acosta-Rodriguez VA, Adams DM, et al. Universal DNA methylation age across mammalian tissues. Nat Aging. 2023; 3:1144–66.
https://doi.org/10.1038/s43587-023-00462-6
PMID:37563227

39. Robeck TR, Fei Z, Lu AT, Haghani A, Jourdain E, Zoller JA, Li CZ, Steinman KJ, DiRocco S, Schmitt T, Osborn S, Van Bonn B, Katsumata E, et al. Multi-species and multi-tissue methylation clocks for age estimation in toothed whales and dolphins. Commun Biol. 2021; 4:642.
https://doi.org/10.1038/s42003-021-02179-x
PMID:34059764

40. Pinho GM, Martin JGA, Farrell C, Haghani A, Zoller JA, Zhang J, Snir S, Pellegrini M, Wayne RK, Blumstein DT, Horvath S. Hibernation slows epigenetic ageing in yellow-bellied marmots. Nat Ecol Evol. 2022; 6:418–26.
https://doi.org/10.1038/s41559-022-01679-1
PMID:35256811

41. Zoller JA, Parasyraki E, Lu AT, Haghani A, Niehrs C, Horvath S. DNA methylation clocks for clawed frogs reveal evolutionary conservation of epigenetic aging. Geroscience. 2023. [Epub ahead of print].
https://doi.org/10.1007/s11357-023-00840-3
PMID:37270437

42. Kerepesi C, Meer MV, Ablaeva J, Amoroso VG, Lee SG, Zhang B, Gerashchenko MV, Trapp A, Yim SH, Lu AT, Levine ME, Seluanov A, Horvath S, et al. Epigenetic aging of the demographically non-aging naked mole-rat. Nat Commun. 2022; 13:355.
https://doi.org/10.1038/s41467-022-27959-9
PMID:35039495

43. Wang T, Ma J, Hogan AN, Fong S, Licon K, Tsui B, Kreisberg JF, Adams PD, Carvunis AR, Bannasch DL, Ostrander EA, Ideker T. Quantitative Translation of Dog-to-Human Aging by Conserved Remodeling of the DNA Methylome. Cell Syst. 2020; 11:176–85.e6. https://doi.org/10.1016/j.cels.2020.06.006 PMID:32619550

44. Zhou FC. DNA methylation program during development. Front Biol (Beijing). 2012; 7:485–94. https://doi.org/10.1007/s11515-012-9246-1 PMID:23687512

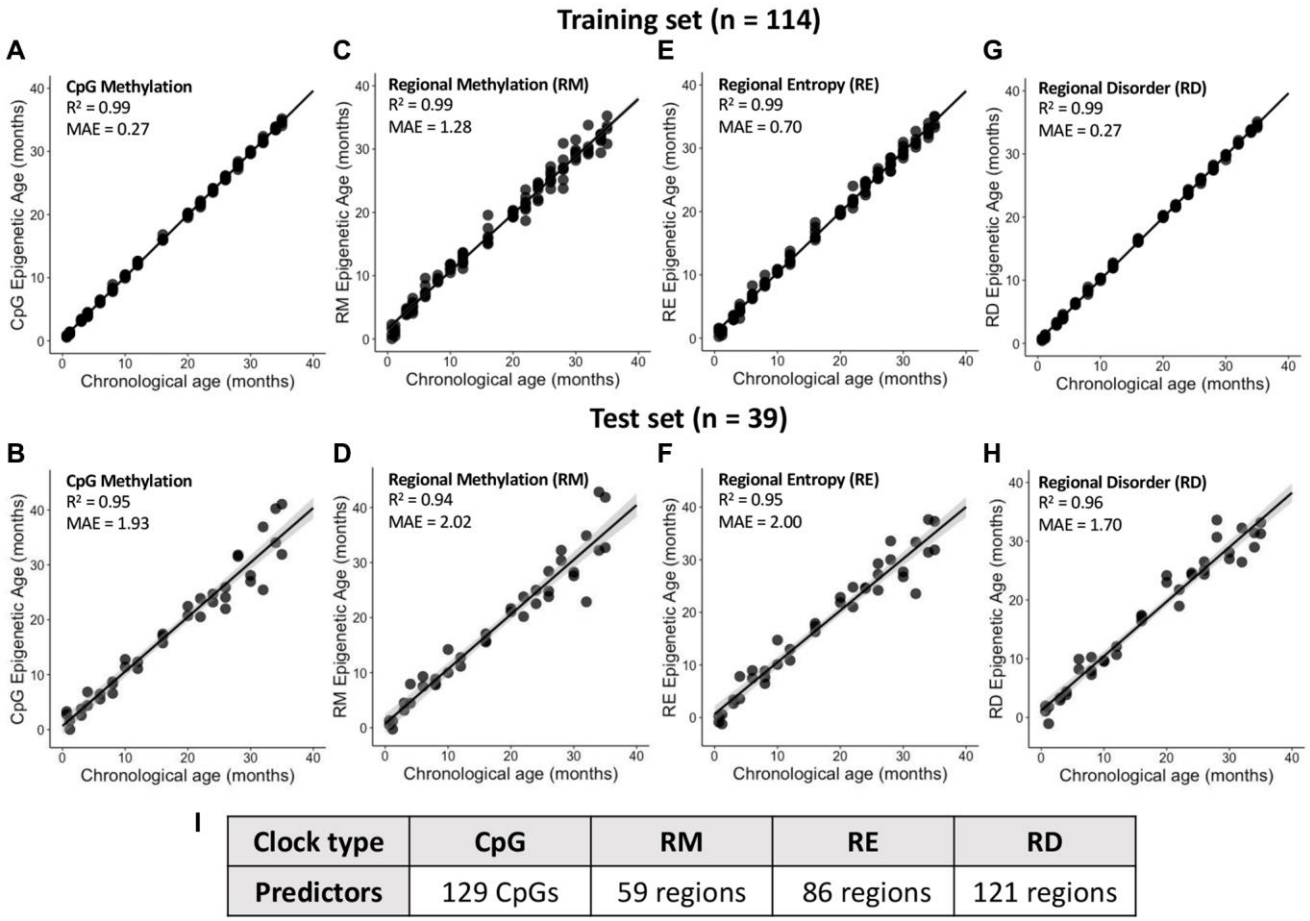
45. Parrott BB, Bertucci EM. Epigenetic Aging Clocks in Ecology and Evolution. Trends Ecol Evol. 2019; 34:767–70. https://doi.org/10.1016/j.tree.2019.06.008 PMID:31296344

46. Spólnicka M, Pośpiech E, Pepłońska B, Zbieć-Piekarska R, Makowska Ż, Pięta A, Karłowska-Pik J, Ziemkiewicz B, Wężyk M, Gasperowicz P, Bednarczuk T, Barcikowska M, Żekanowski C, et al. DNA methylation in ELOVL2 and C1orf132 correctly predicted chronological age of individuals from three disease groups. Int J Legal Med. 2018; 132:1–11. https://doi.org/10.1007/s00414-017-1636-0 PMID:28725932

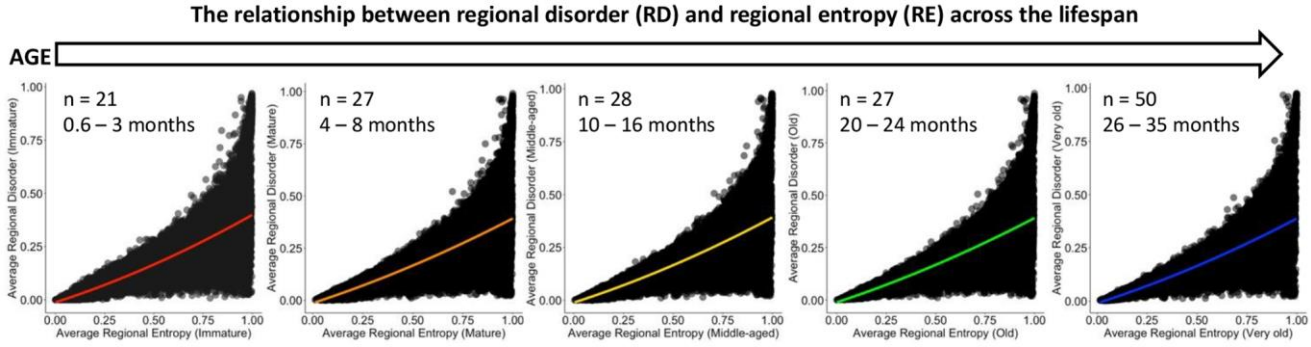
47. Han Y, Franzen J, Stiehl T, Gobs M, Kuo CC, Nikolić M, Hapala J, Koop BE, Strathmann K, Ritz-Timme S, Wagner W. New targeted approaches for epigenetic age predictions. BMC Biol. 2020; 18:71. https://doi.org/10.1186/s12915-020-00807-2 PMID:32580727

SUPPLEMENTARY MATERIALS

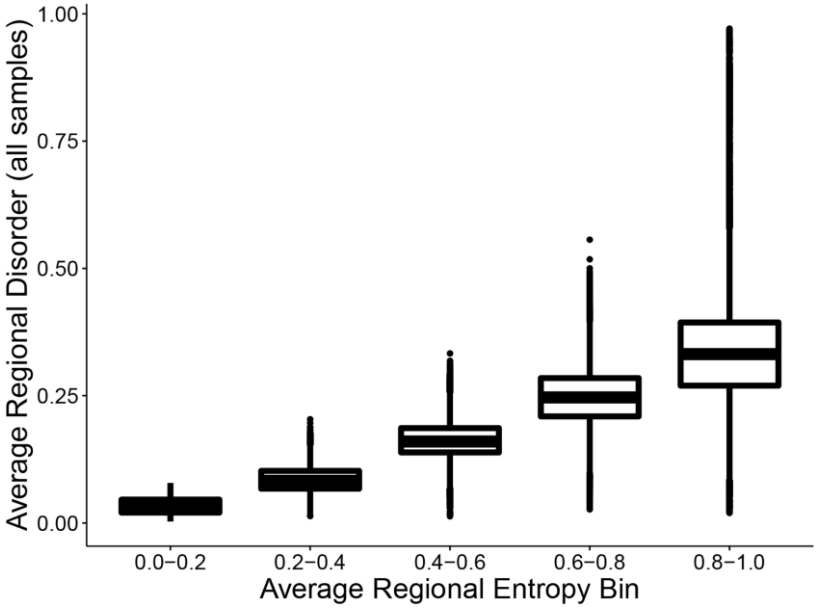
Supplementary Figures



Supplementary Figure 1. Representative epigenetic clocks from each datatype. (**A**, **B**) show results for the epigenetic clock constructed using CpG methylation data, (**C**, **D**) show results for the epigenetic clock constructed using regional methylation (RM) data, (**E**, **F**) show results from an epigenetic clock based on regional entropy (RE) data, and (**G**, **H**) show results from the epigenetic clock based on regional disorder (RD) data. Panel (**I**) indicates the number of predictors comprising each clock.



Supplementary Figure 2. The relationship between average regional disorder and average regional entropy across the lifespan of mice. Samples are broken up into distinct life stages which are denoted by different colors (immature – red, mature – orange, middle-aged – yellow, old – green, and very old – blue).



Supplementary Figure 3. Variation in average regional disorder increases with increasing values of average regional entropy.

Supplementary Tables

Supplementary Table 1. Gene ontology results from regions with age-associated gains in disorder.

Source	GO term name	GO ID	Adjusted *p*-value
GO:MF	DNA-binding transcription factor activity, RNA polymerase II-specific	GO:0000981	1.76E-28
GO:MF	DNA-binding transcription factor activity	GO:0003700	7.58E-28
GO:MF	sequence-specific double-stranded DNA binding	GO:1990837	5.72E-23
GO:MF	double-stranded DNA binding	GO:0003690	4.64E-21
GO:MF	sequence-specific DNA binding	GO:0043565	1.37E-20
GO:MF	RNA polymerase II transcription regulatory region sequence-specific DNA binding	GO:0000977	1.79E-20
GO:MF	transcription regulator activity	GO:0140110	5.41E-20
GO:MF	transcription cis-regulatory region binding	GO:0000976	7.85E-19
GO:MF	transcription regulatory region nucleic acid binding	GO:0001067	1.17E-18
GO:MF	RNA polymerase II cis-regulatory region sequence-specific DNA binding	GO:0000978	1.70E-17
GO:MF	cis-regulatory region sequence-specific DNA binding	GO:0000987	2.57E-17
GO:MF	binding	GO:0005488	8.47E-17
GO:MF	protein binding	GO:0005515	1.12E-15
GO:MF	DNA-binding transcription activator activity	GO:0001216	1.01E-11
GO:MF	DNA-binding transcription activator activity, RNA polymerase II-specific	GO:0001228	1.31E-11
GO:MF	gated channel activity	GO:0022836	1.80E-11
GO:MF	DNA binding	GO:0003677	8.54E-09
GO:MF	ion channel activity	GO:0005216	8.80E-09
GO:MF	voltage-gated cation channel activity	GO:0022843	1.21E-08
GO:MF	channel activity	GO:0015267	2.18E-08
GO:BP	nervous system development	GO:0007399	9.36E-84
GO:BP	neurogenesis	GO:0022008	1.71E-68
GO:BP	system development	GO:0048731	4.56E-67
GO:BP	multicellular organism development	GO:0007275	3.14E-65
GO:BP	generation of neurons	GO:0048699	4.99E-63
GO:BP	anatomical structure development	GO:0048856	4.01E-61
GO:BP	neuron differentiation	GO:0030182	3.05E-60
GO:BP	multicellular organismal process	GO:0032501	5.25E-58
GO:BP	developmental process	GO:0032502	5.17E-57
GO:BP	anatomical structure morphogenesis	GO:0009653	1.15E-54
GO:BP	cell-cell signaling	GO:0007267	1.09E-50
GO:BP	central nervous system development	GO:0007417	4.10E-48
GO:BP	neuron development	GO:0048666	2.56E-47
GO:BP	cell differentiation	GO:0030154	5.75E-45
GO:BP	cell development	GO:0048468	9.25E-45
GO:BP	cellular developmental process	GO:0048869	1.71E-44
GO:BP	neuron projection development	GO:0031175	2.60E-41
GO:BP	animal organ development	GO:0048513	3.67E-41
GO:BP	brain development	GO:0007420	2.06E-40
GO:BP	head development	GO:0060322	4.07E-40

Top 20 terms from molecular function (MF) and biological process (BP) based on significance values are shown.

Supplementary Table 2. Gene ontology results from regions with age-associated losses in disorder.

Source	GO term name	GO ID	Adjusted *p*-value
GO:MF	protein binding	GO:0005515	0.00063605
GO:MF	binding	GO:0005488	0.00159718
GO:MF	protein kinase activity	GO:0004672	0.00357349
GO:MF	phosphotransferase activity, alcohol group as acceptor	GO:0016773	0.01376325
GO:MF	transferase activity	GO:0016740	0.04052106
GO:BP	developmental process	GO:0032502	0.00019855
GO:BP	anatomical structure development	GO:0048856	0.00021537
GO:BP	multicellular organismal process	GO:0032501	0.00032362
GO:BP	anatomical structure morphogenesis	GO:0009653	0.00042435
GO:BP	response to stimulus	GO:0050896	0.00074356
GO:BP	signal transduction	GO:0007165	0.0017411
GO:BP	regulation of cellular process	GO:0050794	0.0030705
GO:BP	cellular response to stimulus	GO:0051716	0.0044849
GO:BP	signaling	GO:0023052	0.00807516
GO:BP	cell communication	GO:0007154	0.01493627
GO:BP	regulation of biological process	GO:0050789	0.01644778
GO:BP	biological regulation	GO:0065007	0.02094827
GO:BP	intracellular signal transduction	GO:0035556	0.03178697
GO:BP	multicellular organism development	GO:0007275	0.04073998
GO:MF	protein binding	GO:0005515	0.00063605
GO:MF	binding	GO:0005488	0.00159718
GO:MF	protein kinase activity	GO:0004672	0.00357349
GO:MF	phosphotransferase activity, alcohol group as acceptor	GO:0016773	0.01376325
GO:MF	transferase activity	GO:0016740	0.04052106
GO:BP	developmental process	GO:0032502	0.00019855
GO:BP	anatomical structure development	GO:0048856	0.00021537
GO:BP	multicellular organismal process	GO:0032501	0.00032362
GO:BP	anatomical structure morphogenesis	GO:0009653	0.00042435
GO:BP	response to stimulus	GO:0050896	0.00074356
GO:BP	signal transduction	GO:0007165	0.0017411
GO:BP	regulation of cellular process	GO:0050794	0.0030705
GO:BP	cellular response to stimulus	GO:0051716	0.0044849
GO:BP	signaling	GO:0023052	0.00807516
GO:BP	cell communication	GO:0007154	0.01493627
GO:BP	regulation of biological process	GO:0050789	0.01644778
GO:BP	biological regulation	GO:0065007	0.02094827
GO:BP	intracellular signal transduction	GO:0035556	0.03178697
GO:BP	multicellular organism development	GO:0007275	0.04073998

Terms from molecular function (MF) and biological process (BP) are shown.