



# Many chronological aging clocks can be found throughout the epigenome: Implications for quantifying biological aging

Hunter L. Porter<sup>1,2,3</sup> | Chase A. Brown<sup>1,2</sup> | Xiavan Roopnarinesingh<sup>1,2</sup> |  
Cory B. Giles<sup>1,2,3</sup> | Constantin Georgescu<sup>2</sup> | Willard M. Freeman<sup>1,2</sup> |  
Jonathan D. Wren<sup>1,2,3</sup>

<sup>1</sup>Oklahoma Medical Research Foundation, Oklahoma, OK, USA

<sup>2</sup>University of Oklahoma Health Sciences Center, Oklahoma, OK, USA

<sup>3</sup>Oklahoma Center for Geroscience and Healthy Brain Aging, Oklahoma, OK, USA

## Correspondence

Jonathan D. Wren, Oklahoma Medical Research Foundation, Oklahoma, OK, USA.

Email: jonathan-wren@omrf.org

## Funding information

The authors would like to thank NIH grants #5P30AG050911 (to J.D.W. and W.M.F.), NIH grant #2P20GM103636 (to J.D.W.), #1T32AG052363-01 (To William Sonntag), #1F31AG063493 (to H.L.P.), R01AG059430 (to W.M.F.), and VA grant #I01BX003906 (to W.M.F.)

## Abstract

Epigenetic alterations are a hallmark of aging and age-related diseases. Computational models using DNA methylation data can create “epigenetic clocks” which are proposed to reflect “biological” aging. Thus, it is important to understand the relationship between predictive clock sites and aging biology. To do this, we examined over 450,000 methylation sites from 9,699 samples. We found ~20% of the measured genomic cytosines can be used to make many different epigenetic clocks whose age prediction performance surpasses that of telomere length. Of these predictive sites, the average methylation change over a lifetime was small (~1.5%) and these sites were under-represented in canonical regions of epigenetic regulation. There was only a weak association between “accelerated” epigenetic aging and disease. We also compare tissue-specific and pan-tissue clock performance. This is critical to applying clocks both to new sample sets in basic research, as well as understanding if clinically available tissues will be feasible samples to evaluate “epigenetic aging” in unavailable tissues (e.g., brain). Despite the reproducible and accurate age predictions from DNA methylation data, these findings suggest they may have limited utility as currently designed in understanding the molecular biology of aging and may not be suitable as surrogate endpoints in studies of anti-aging interventions. Purpose-built clocks for specific tissues age ranges or phenotypes may perform better for their specific purpose. However, if purpose-built clocks are necessary for meaningful predictions, then the utility of clocks and their application in the field needs to be considered in that context.

## KEYWORDS

Aging, bioinformatics, epigenetic clocks, epigenetics

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Aging Cell* published by Anatomical Society and John Wiley & Sons Ltd.



## 1 | INTRODUCTION

Predicting age from molecular data has been a long-standing interest in aging research because it implies that we can identify the correlative/causal factors behind aging. By extension, if molecular changes associated with age-related diseases can be identified, hypotheses about potentially effective interventions that extend biological life and healthspan can be made. The first molecular predictors of chronological age included telomere length and p16INK4A levels (Tsygankov et al., 2009) but, recently, “epigenetic clocks” have supplanted them in accuracy and precision (Horvath, 2013), as well as their ability to predict all-cause mortality risk (Perna et al., 2016). Since these original reports on mortality risk, studies have increasingly used epigenetic clocks to evaluate interventions that may extend lifespan in both mice (Wang et al., 2017) and humans (Fahy et al., 2019). Epigenetic clocks that can quantify a “biological age” that is distinct from chronological age, could have a profound impact on aging research. However, such clocks depend upon understanding the degree to which epigenetic changes at clock sites reflect altered physiological states, versus solely the passage of time which, outside of forensic applications, is already known.

Epigenetic clocks are multivariate machine learning models that predict age using methylation levels from a set of genomic cytosine sites. A downside of the machine learning approach inherent to epigenetic clocks is that the behavior of individual sites across time, as well as the biological regulation and impact, is often obscured. In addition, the ability to build equivalent performing clocks from the same data through small algorithmic differences does not provide a prioritization for biological relevance of the clock loci. Indeed, any machine learning method exhibits a “black box” nature which can obscure the biological relationship between methylation and variables of interest. Based upon the training data, which includes chronological age and CpG methylation levels at specific loci in a set of samples, algorithmic “learning” weighs each data point and combination of data points to choose sites whose methylation pattern best linearly correlates with chronological age. Unlike telomere shortening, which has a fairly straightforward biological impact to understand, there is no obvious biological interpretation of the relevance of sites used by epigenetic clocks to predict age. Prior studies demonstrate enrichment for some functionally relevant features, such as polycomb repressor targets (Horvath, 2013) and even that clocks can be trained on data from just these targets (Yang et al., 2016). This, however, leaves open the question of whether epigenetic marks in these regions are causative of aging, or if “reversing” epigenetic aging is sufficient to ameliorate age-associated dysfunctions. Nonetheless, the field has extrapolated that epigenetic clock predicted ages that are older or younger than the individual's chronological age may represent “accelerated” and “decelerated” biological aging (Fahy et al., 2019; Fransquet et al., 2019; Horvath & Raj, 2018). Thus, it is becoming increasingly important to better understand how epigenetic clocks work and the biological relevance of the sites used.

Epigenetic clocks have been built using a wide variety of different genomic loci to predict age. Clocks can vary from three sites

(Weidner et al., 2014) to any number, with the most popular epigenetic clock built by Horvath using 353 sites (Horvath, 2013). These clocks do have some overlap in sites chosen, but they are predominantly composed of unique genomic sites (Horvath & Raj, 2018). The fact that different clocks can arrive at similar accuracy using many different loci in the genome raises at least two questions. First, just how much of one's epigenome changes with age and to what extent? To this end, prior work has identified many potentially age-related loci in blood (Slieker et al., 2016), but the extent of changes in pan-tissue models and outside of array data is unknown. Second, if multiple regions are equally predictive, then do they have anything in common biologically?

A key biological insight from many of the epigenetic clocks is found in their ability to identify loci that molecularly “age” similarly across tissues, that is, pan-tissue clocks. Epigenetic clocks select loci that predict age regardless of their tissue source, implying some pan-tissue mechanism is driving epigenetic aging. This is despite the established role of methylation in determining cell types and large methylome differences between tissues. In order to be used as an effective biomarker of clinical aging, epigenetic clocks must be able to predict aging using tissues that can be sampled pre-mortem (e.g., blood) and reflect changes in tissues affected by aging (e.g., brain). If these age-predictive epigenetic changes are not common across tissues, we must find a way to translate within individuals if clocks are to be used to evaluate interventions that prevent age-related diseases. Since epigenetic age “acceleration” is predictive of all-cause mortality using just blood samples (Perna et al., 2016), we hypothesize that the age-predictive methylation changes will be similar across tissues. However, prior work studying the relationship between pan-tissue clocks and neurodegenerative disease has found the potential for “false positives” coming from these pan-tissue signals that warrant even further scrutiny (Shireby et al., 2020).

While methylation differences between tissues are large, we found epigenetic clock sites change by only 1.5% on average between young (<35 years of age) and aged (>65 years of age) samples. Applications of epigenetic clocks in the field often use pan-tissue clocks to evaluate tissue-specific “biological aging” (Levine et al., 2016; McKinney et al., 2018; Ward-Caviness et al., 2016). Tissue-specific clocks may be necessary for clock models to predict age with loci that will also correlate with age-related dysfunction in that tissue. Because the process of aging comes with predictable phenotypes and increased risk of age-related diseases, when methylation changes are hypothesized to reflect “biological aging,” they should have a relationship to the cellular aging of the examined tissues. This leads us to two hypotheses. First, that samples from individuals with age-related diseases would be epigenetically “older” than control counterparts. Second, age-predictive loci should be consistent when training clocks across the lifespan and tissue types.

Elastic net regression, used to create epigenetic clocks, is predicated on selecting the best set of unique predictors, not necessarily all age-predictive sites in the genome. Epigenetic clocks using other genomic sites have been developed for cellular senescence (Lowe et al., 2016), obesity (Sargent, 2015), cancer (Zheng et al., 2016), and

even spawned their own theories of aging (Horvath & Raj, 2018). Taken together, all of these clocks suggest methylation at a large number of sites in the genome can be indicative of health and disease and that there are commonalities across tissues. However, the actual age-related variation in methylation at clock loci is very small (<5%). Thus, the detected methylation changes with age could be attributable to changes in cellular abundance of a relatively rare cell subtype in examined tissues (e.g., infiltrating macrophages) rather than an epigenetic change in the resident cells of a tissue. Alternatively, there could be a common form of age-related cellular change across cell types and tissues (e.g., cellular senescence). Or worse, the correlation could be spurious. Quoting Calude and Longo (Calude, 2017), "One of the main ideas supporting data analytics is that a series of correlations will continue or iterate similarly along the chosen parameter (recurrence). If, for example, time is the main parameter [...], then the correlation will extend into the future by iterating a similar "distance", typically, between the chosen observables." Algorithmic identification of potentially interesting patterns within large datasets holds great potential for advancing scientific understanding. It is, however, not a substitute for it. Mathematically, the larger a dataset, the more arbitrary correlations it will contain (Calude, 2017). Thus, once identified, it is incumbent upon us to identify how (and if) such patterns inform our current understanding. To assay the breadth of sites in the genome that are associated with age, potential mechanisms that could explain the ability to generate a clock model, potential interacting partners of DNA methylation-modifying enzymes that influence changing methylation with age, and understand what pathways may be affected by age-related methylation changes, we conducted a large-scale analysis of human methylation data using the Illumina 450k methylation array platform, for which we collected 9,699 samples of adults (aged 25+) with age and tissue descriptions using our automated sample annotation approach (Giles et al., 2017).

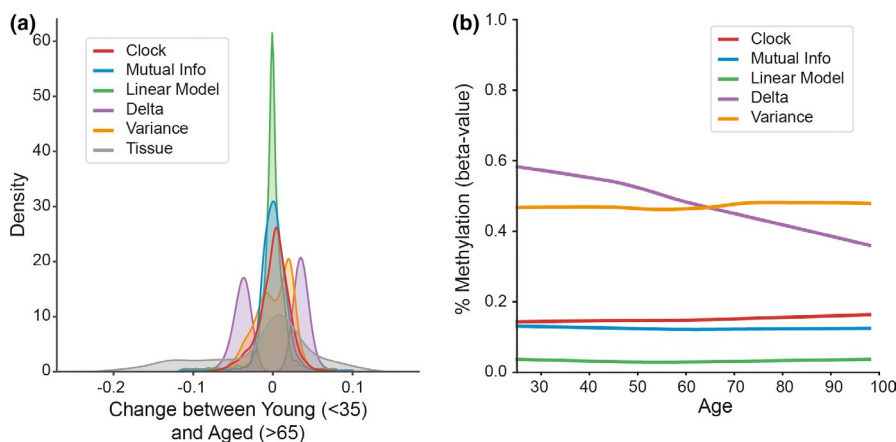
## 2 | RESULTS

### 2.1 | Epigenetic changes with age are small in magnitude across the lifespan

The magnitude of epigenetic clock site (ECS) methylation changes across the lifespan is significantly smaller than the largest age-related changes within non-clock sites, and smaller still than known differences in DNA methylation between tissues (Figure 1). Looking at age-related sites by a variety of metrics, we see that age-predictive sites by linear regression and non-linear mutual information regression follow the same pattern of small magnitude changes found in clock sites. We then analyze the highest weighted locus by each metric (e.g., epigenetic clock site in red) (Figure 1b). Even the most "age-related" locus, as identified by multiple methods, shows a small magnitude difference over the lifespan. This finding leads to many questions. How do such small age-related changes occur consistently across individuals and tissues composed of multiple cell types? Are these small differences occurring in a restricted set of loci, or are there large numbers of these sites from which a variety of clocks can be constructed? Perhaps methylation values reach an asymptote and regress toward the young values?

### 2.2 | Age predictions from epigenetic clocks are replicable, but use different loci

A logical starting point in determining the number of possible ECSs is to recreate Horvath's original clock. We independently collected as much of the original data that was publicly available from NCBI Gene Expression Omnibus (GEO), obtaining ~75% of the training data. We selected a set of loci whose methylation state can predict age using



**FIGURE 1** Average lifespan change of age-predictive loci is small. Age-related loci were selected by different methods: epigenetic clocks (red), linear models (green), top 10% mutual information (blue), top 10% greatest change in mean ("delta," purple) or variance (orange). (a) Histogram of age-related changes in methylation over the lifespan (beta values from 0 to 1.0), averaged across samples. An example tissue difference of blood versus brain is also presented (gray). Aside from sites with the greatest mean difference with age (purple) and between tissues (gray), age-related sites exhibit small magnitude changes over the lifespan. (b) The most informative individual locus by each method is displayed as a lowess-smoothed fit over age



the same preprocessing, including imputation, normalization, and elastic net modeling.

We replicated the results of the original Horvath model (Figure 2a, 2b). We also trained a model on 9699 samples of Illumina 450k methylation array data deposited in NCBI GEO (Figure 2C), which allowed us to use data from more sites (~450,000) than the 21,369 sites used in the Horvath report. We only included samples from experiments with over 100 samples, as we found that including smaller experiments quickly diminishes clock performance. Our full 450k model outperformed the original Horvath model, but this is expected as it had access to more features in both the training and test sets. While some data from Horvath's original paper were unavailable and/or lacked age data that were publicly available, we demonstrated roughly equivalent performance across the three models with mainly different sets of sites in each model (Figure 2d).

Because clock models trained on almost identical data can select different loci as the most predictive set and perform equivalently, we examined how many loci could be used as clock sites via an iterative “knockout” approach. After using elastic net regression to identify the most informative loci, these sites were then removed and new clock models trained on the remaining sites. We fit 5 new clock models after every removal and removed any locus used by any of the 5 clocks. By iteratively removing the most predictive sites every round, subsequent models should become increasingly hindered in their performance. Interestingly, rather than a rapid depletion of predicted performance, we saw a gradual linear decrease in prediction accuracy (as defined by Pearson correlation coefficient between predicted and actual age, Figure 2e) and rise in model error (Figure 2e) until about 20% of sites have been removed from the training data. These data illustrate the breadth of CpGs that are strong age predictors from the sites measured by the Illumina 450k microarray.

### 2.3 | Pan-tissue epigenetic clocks fail to identify tissue-specific epigenetic aging

DNA methylation has known roles in specifying tissue identity (Koh & Rao, 2013; Macaluso & Giordano, 2004) via differentially suppressing and activating specific regions of the genome in a cell type-specific manner. Our own data and previous reports (Horvath, 2013; Lowe et al., 2016; Thompson et al., 2018) on pan-tissue epigenetic clocks demonstrate that chronological age-predictive models work equally well across most tissues when trained on many tissues. However, clocks trained on blood alone have poor performance at predicting tissue-specific impairments, such as cognitive decline (Starnawska et al., 2017).

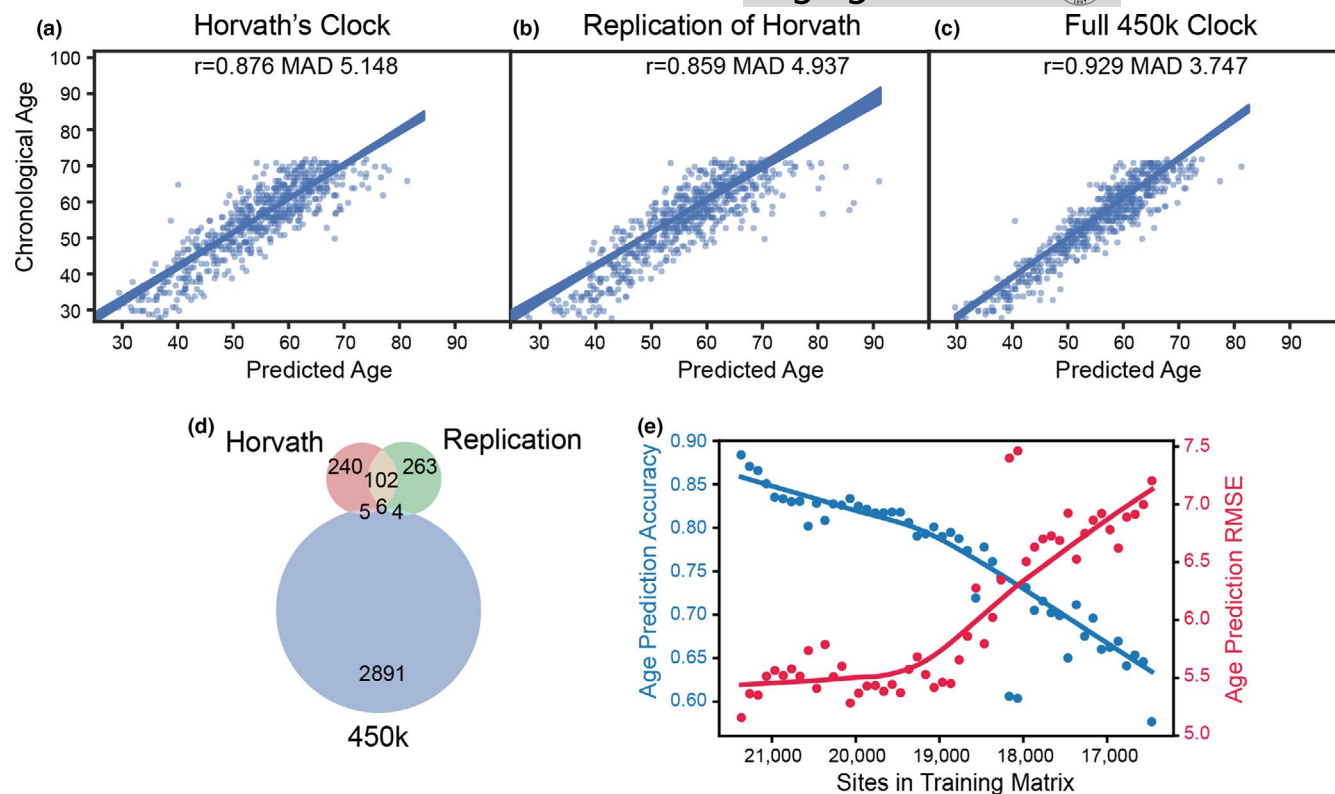
This raises the question of whether pan-tissue clocks either still capture enough tissue-specific/relevant changes by inclusion of multiple tissues in the training set, or if pan-tissue clocks are missing tissue-relevant changes, and potentially sacrificing insight into tissue-specific aging and disease.

To look for tissue-specific age association, we used an ordinary least squares regression to model the effects of age, tissue, and their interactions. 39823 loci were significantly ( $q < 0.05$ ) associated with tissue, while 9587 were associated with age. Of the 6226 sites with significant main effects of both tissue and age, 3939 had a significant interaction effect, suggesting a tissue-specific aging change. The rest of the age-associated sites had tissue-independent effects (as defined by non-significant tissue:age interactions), suggesting common age-related changes across tissues (Figure 3a, c). Looking closer at the loci with significant age main effects, the tissue-specific aging interactions are weak (Figure 3b). This suggests that there is some common process across tissues that causes these changes regardless of whether it is a regulated program or a systemic form of epigenomic entropy. Dysregulations of epigenetic machinery have been implicated as the driver of the relationship between DNA methylation and aging (Bell et al., 2019), but our prior work in mouse found no changes in the direct methylation machinery with aging (Hadad et al., 2016).

Observing that most loci with a main effect of age were similar across tissues (except for saliva) was surprising in light of the strong relationship between methylation and tissue identity. This leads us to question whether elastic net would identify ideal pan-tissue predictors as the best sites to predict age even in a single tissue, or if they would be too weak of a signal compared to tissue-specific signals. When we trained clocks on data from only one tissue at a time, they perform markedly worse on predicting age in other tissues (Figure 3d and e). This indicates that the tissue-specific signals are in fact strong enough to drown out the non-tissue-specific signals, which are more abundant. Notably, saliva sample ages were uniquely well-predicted by all other tissue models, but poor predictors of other tissue ages. When comparing the coefficients of loci identified by elastic net for multiple tissues, we do see moderate associations ( $r > 0.5$ ) for all tissues—indicating that some loci are being chosen that covary even by tissue-specific models (Figure 3f). The loci chosen by these tissue-specific clocks are largely uniformly distributed across genomic features, but are notably enriched near GTEx (Lonsdale et al., 2013) eQTLs. Perhaps the enrichment of these loci near regions where point mutations are sufficient to affect disease-related gene expression is indicative of a potential biological link, but determining whether these changes are causal, compensatory, or silent in the aging process will require further studies in epigenome editing/manipulation.

### 2.4 | Age-predictive loci depend upon which ages are used to train and test the clock

We have made two key observations regarding methylation aging, namely 1) non-linear site selection outperforms linear for training epigenetic clocks (Figure S2) and 2) age-predictive loci have very small changes over the lifespan (Figure 1). This raised the question of how sites that would need to change by a fraction of a percent per year can consistently predict age. These small changes are observed



**FIGURE 2** Different epigenetic clocks select different sites but perform similarly. (a) Horvath's original model sites (353) showing predicted vs actual age on a single test set of blood samples (GSE42861) (b) The model based on replication of Horvath's method produced predicted vs actual age on GSE42861 using all available training data from Horvath's original paper (Horvath, 2013). The model performed slightly worse with the smaller training set, and also selected fewer sites (252). (c) A model trained using all age-annotated 450k data with sample labels from ALE(Giles et al., 2017), was tested on the same set of blood samples (GSE42861). This model performed better than the original Horvath clock, but also had access to a much larger set of loci for training and selected a larger set of loci (2906) to predict age. (d) Venn diagram of clock loci used in each model. The models selected different sites despite similar prediction quality with the only variables being different training samples and a different random seed. (e) Pearson correlation between predicted and actual age (blue) and root mean squared error (RMSE, red) by number of loci remaining in training set. Each point represents a new model trained with all loci used by previous models removed

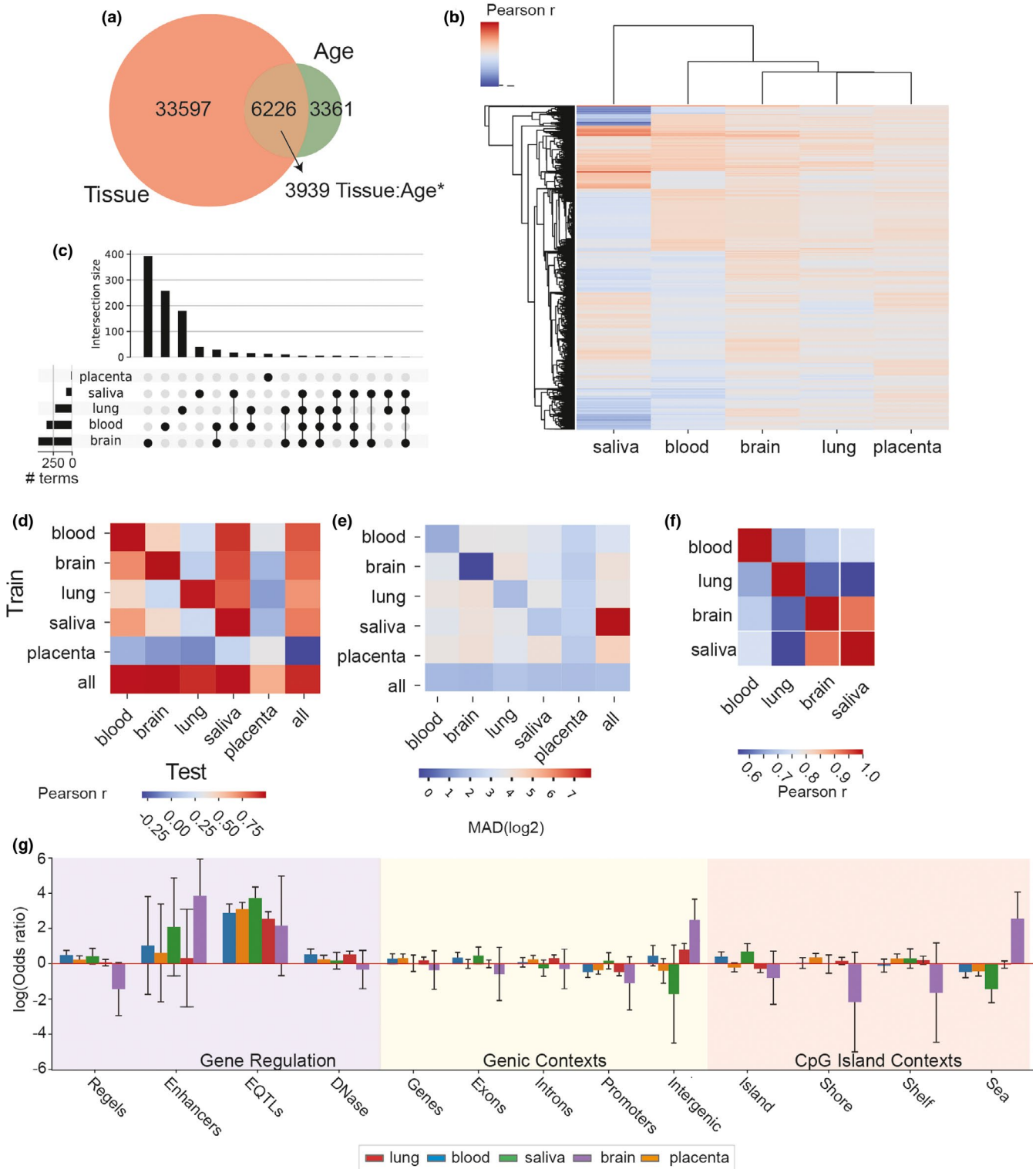
even in a pan-tissue context. The changes are much larger than we would expect from senescent cells, which are relatively rare and appear at different rates between tissues (Tuttle et al., 2020), and also are unlikely explained by some change in blood constituent cells (Chen et al., 2016) as they would vary wildly in their abundance with vascularity of the tissue, such as between whole blood and saliva. To answer this question, we performed an experiment to see if the most predictive loci from full age range models are identifiable in windows within the lifespan (Figure 4a). We trained three epigenetic clocks separately using data from three groups: young (25–50), middle (50–75), and aged (75–100) samples, then tested them on each age group. We found sites most predictive of aging in younger ages (25–50) were poor predictors later in life (75–100). Based on our observations of the linear and non-linear relationships, we then looked at how clock sites from both Horvath and our own clocks change throughout the lifespan. We found that 99.5% of loci exhibit a distribution with at least one inflection point with aging that could be regressed into an overall linear trend (Figure 4d, Figure S4), alongside loci where locally fitted regression matched the sites' linear fits (Figure 4e). While other reports (e.g., - (Marioni et al., 2019)) have

noted a non-linear change with age, the parabolic distribution we observe appears to be novel. These changes may be attributable to changes in variance with aging (Slieker et al., 2016), but they describe an overall trend to increasing variance that would not necessarily lead to a parabolic distribution. Since these parabolic sites are also selected by epigenetic clocks on the basis of an extrapolated linear trend, further exploration of the trajectory of methylation aging may yield understanding of how clock predictions would respond to aging interventions and diseases. Similar to tissue-specific clocks, clocks trained on discrete sections of the lifespan are largely uniformly distributed with an enrichment near eQTLs (Figure 4f).

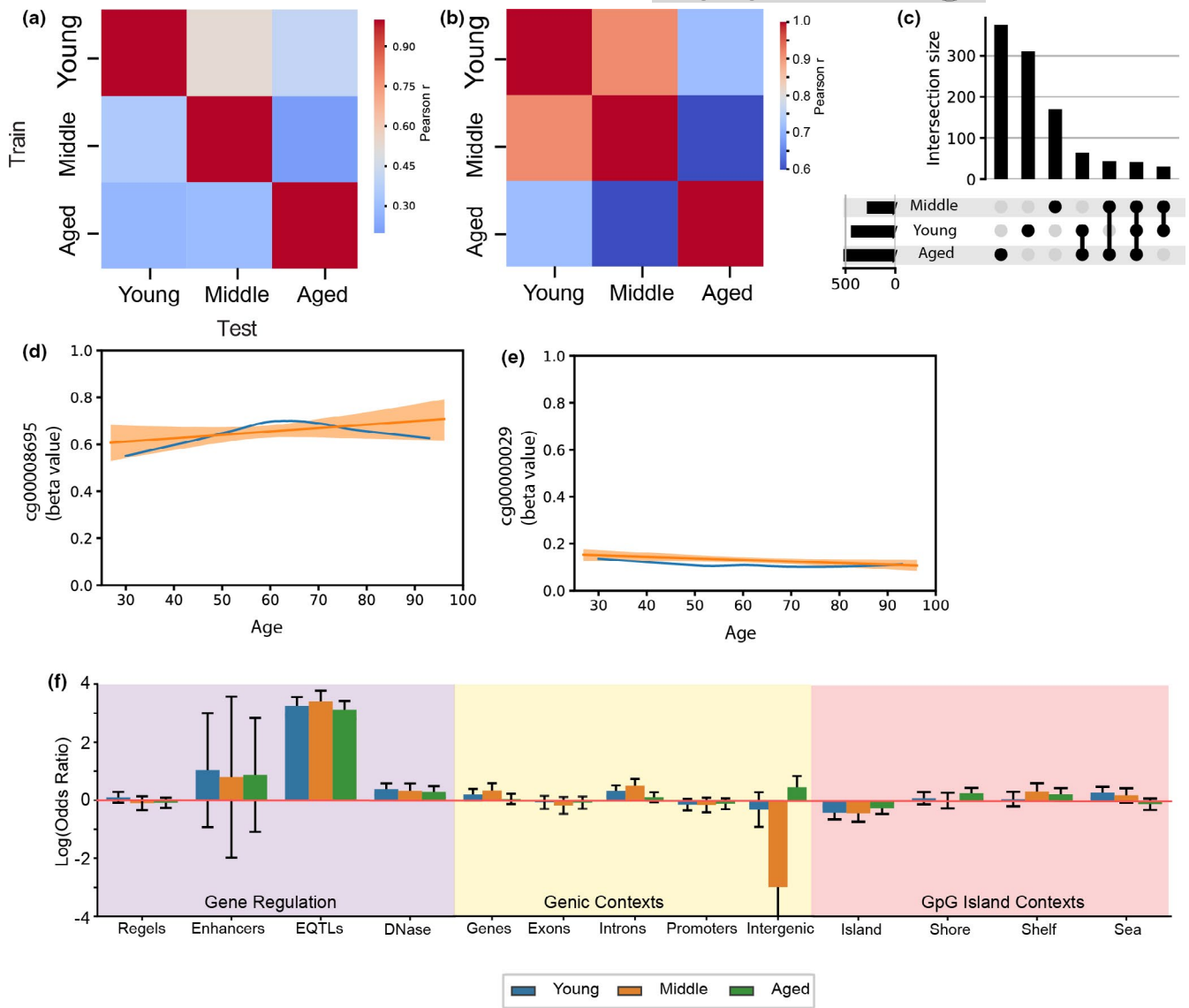
## 2.5 | Age-predictive methylation loci are depleted in biologically informative regions

To understand the biological relevance of different loci that are predictive of or related to aging, we performed genomic enrichment analyses on three sets of methylation loci. These sets were as follows: 1) 6666 loci we determined that change with age by linear regression,





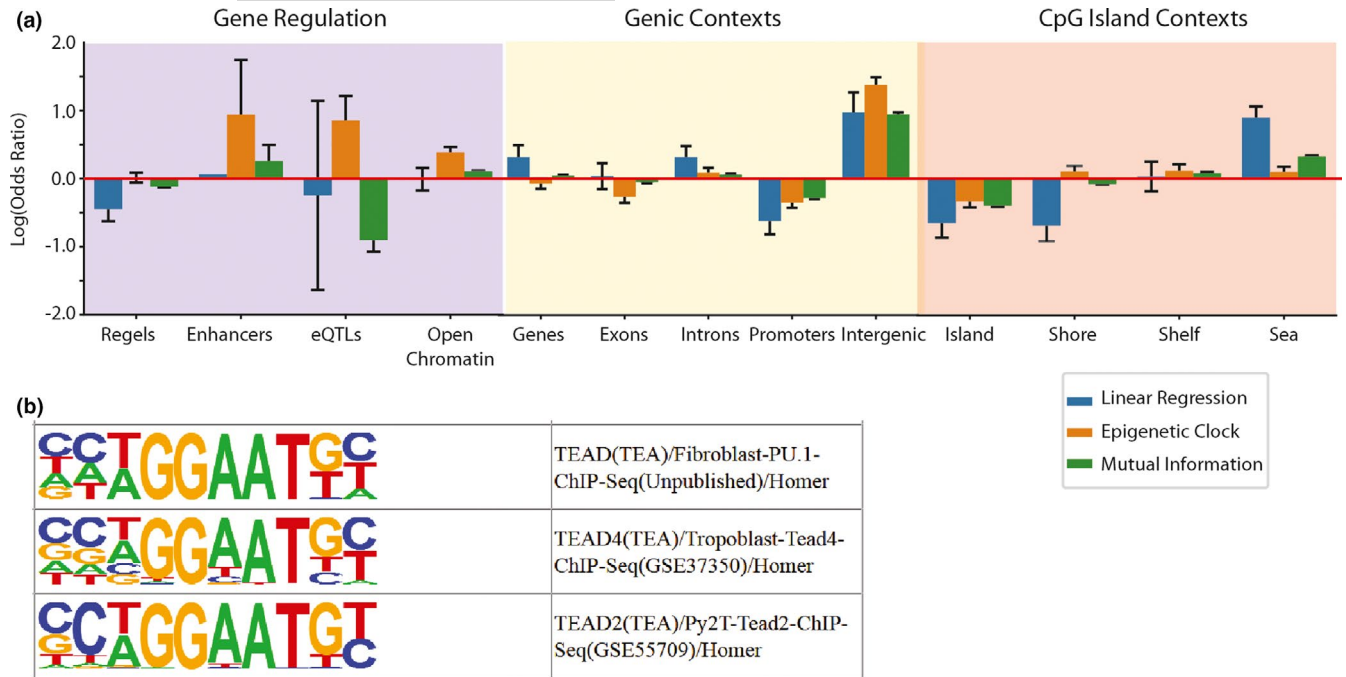
**FIGURE 3** Modeling tissue-specific aging methylation changes. (a) Venn diagram of the overlap between methylated loci that significantly differ between tissues and those that change with age in the 450k data. Multiple testing correction was done using a Benjamini-Hochberg 10% FDR correction. (b) Clustering of all sites with a significant age association by their per-tissue aging correlation coefficients. Coloration indicates methylation-age correlation coefficients for that specific tissue. (c) UpSet plot showing the number of loci in each trained clock and their combinatorial overlaps. (d and e) Matrix of correlation between predicted and actual age from epigenetic clocks trained and tested on a per-tissue basis. Epigenetic clocks were trained on the tissues in the Y-axis and tested on the tissues in the X-axis. Model correlations between predicted and actual age shown in (d) while median absolute deviation is shown in (e). (f) Matrix of correlations between clock site coefficients common between at least 3 of 5 clocks trained on specific tissues. The placenta clock did not have any sites common to at least 3 other models. (g) Bar plots of log odds ratios showing enrichment/depletion of loci from clocks with different training tissues with respect to regulatory features, genic features, and CpG islands, corrected for the coverage of 450k methylation array



**FIGURE 4** Clocks trained on discrete age groups show markedly poor performance at predicting higher/lower age groups. (a) Matrix of prediction accuracy from epigenetic clocks trained and tested on age three age bins. Each age bin represents a third of the ages from 25 to 100 (e.g., - Young = ages 25–50). (b) Matrix of correlations between clock site coefficients common between clocks trained on Young, Middle, and Aged groups. (c) Upset plot showing the number of loci in each trained clock and their combinatorial overlaps. (d and e) Line plots of two example loci comparing loess regression (blue) to linear regression (orange) in the GSE60185 dataset. (d) a clock site with a parabolic distribution that reflects into a linear one. (e) a clock site where a linear model is a good fit for age-related changes in methylation. (f) Bar plots of log odds ratios showing enrichment/depletion of loci from clocks with different training age range with respect to regulatory features, genic features, and CpG islands, corrected for genomic distribution of sites measured on the 450k array

2) 2624 that serve as predictive clock sites, and 3) 79921 sites related to aging as measured by mutual information. When exploring the 6666 age-related sites identified by traditional statistical analysis (linear regression) we found strong enrichments for genomic regions with limited known biological function for DNA methylation such as intergenic regions and sites outside of CpG islands, with underrepresentation in regions such as promoters and CpG islands where DNA methylation is understood to regulate gene expression/genomic accessibility (Figure 5). The results were similar but weaker for age-related sites selected using mutual information. In contrast, epigenetic clock sites were previously reported as enriched in VISTA (Visel et al.,

2006) enhancers and near GTEX (Lonsdale et al., 2013) eQTLs. We attempted to further interrogate the activity state of promoters near age-predictive cytosines, but it is impossible to generate a meaningful multi-tissue prediction of activity with current data. Performing the enrichment against a target tissue of interest did not yield any significant results (Figure S8). They were also enriched in open chromatin regions as defined by aggregated DNase hypersensitivity data. However, the regulatory element enrichment appears to be driven by enhancers as aggregate regulatory elements, because subsets of enhancers, super enhancers, and repressors show neither enrichment nor depletion. All methods of identifying age-related methylation sites found



**FIGURE 5** Genomic feature enrichments for age-related loci. Epigenetic clock loci are enriched in information-poor regions of the genome. The top quartile of mutual information loci (green), sites with significant age effects by OLS regression (blue), and sites chosen in an epigenetic clock built on the full 450k dataset (orange) was compared to known genomic features. (a) Bar plots of log odds ratios showing enrichment/depletion of loci from multiple models with respect to regulatory features, genic features, and CpG islands. Age-related methylation loci are enriched in intergenic regions and depleted in gene promoters. For CpG islands, shores were defined as 2kb up and downstream of the CpG island body, and shelves were defined as 2kb up and downstream of the shores. Age-related methylation loci are enriched in the open sea (outside of CpG islands) and depleted in the island bodies. Age-related methylation loci selected by mutual information and regression are depleted in enhancers, eQTLs, and other regulatory elements (TFBS, repressors, etc.), and distributed evenly throughout euchromatin and heterochromatin. Epigenetic clock sites are however enriched in open chromatin, near eQTLs and gene enhancers. (b) Table with logos of most enriched motifs from HOMER analysis of the probe sequence from epigenetic clock sites

they were depleted in CpG island bodies and gene promoters while being enriched in intergenic regions.

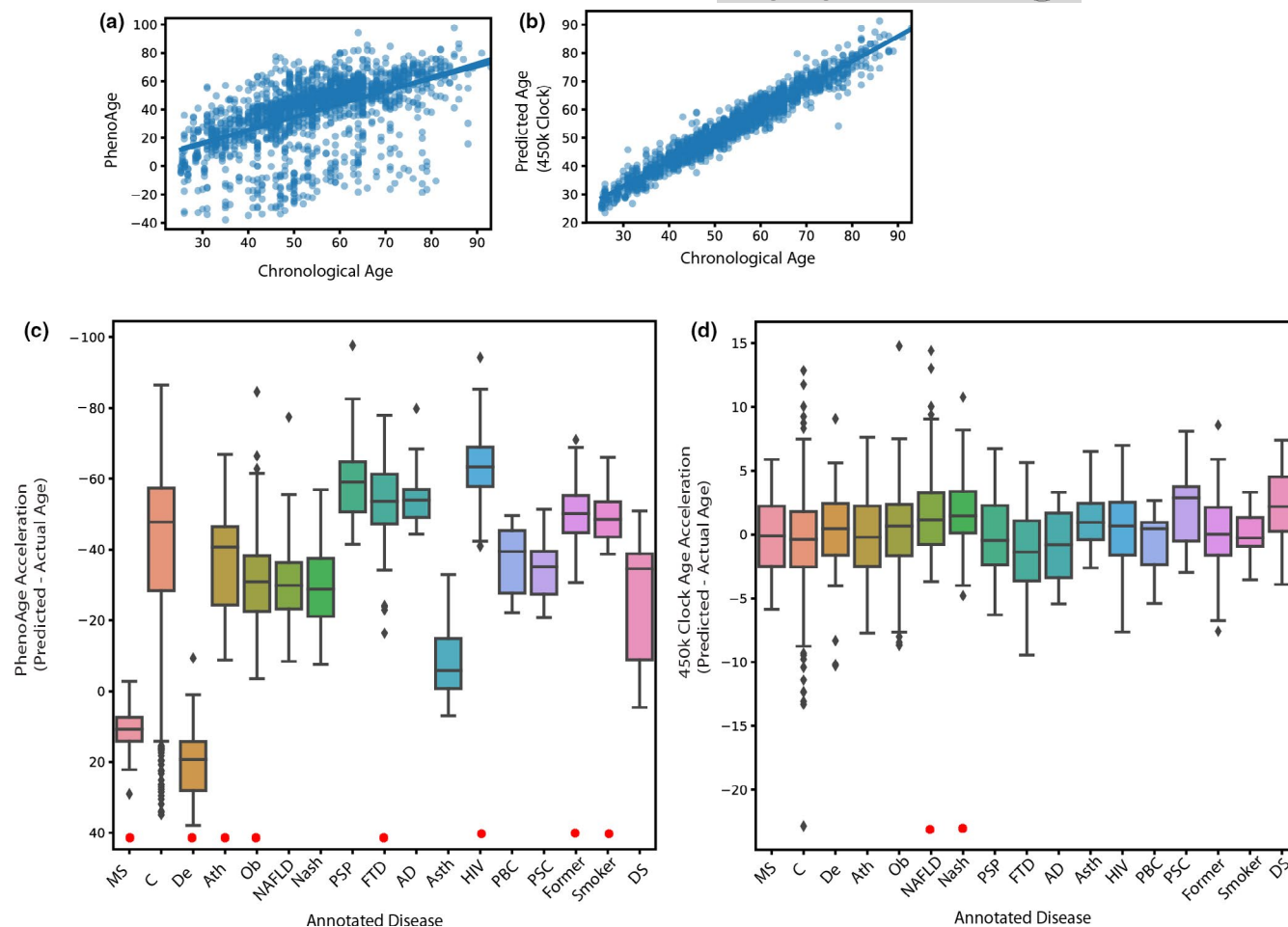
We then analyzed the clock sites' relationship to DNA-binding proteins by comparing the probe sequence used to detect methylation loci to DNA binding motifs using HOMER. We attempted to interrogate the enrichments with respect to the "targeted" cytosines but found no significant results after multiple testing (Supplement S3). The top 3 most significant sequence enrichments for known DNA-binding protein motifs were all TEAD proteins, which are enhancer binding proteins that aid in the initiation of transcription and have been linked to cellular senescence (Xie et al., 2013) and age-related disease (Tsika et al., 2010). These enhancer enrichments were previously reported (Bell et al., 2019), but in their analysis were under the threshold for significance after adjusting for the array background. Nonetheless, these provide a potential mechanism to further explore the biological impact of altered clock site methylation.

## 2.6 | Epigenetic Age Acceleration, Aging Diseases, and the Trajectory of Aging

If epigenetic age acceleration is truly a biomarker of aging, we would expect samples from patients with age-related diseases and

conditions tend to have higher predicted ages than healthy samples. To this end, we annotated 1767 samples for a variety of conditions including multiple sclerosis, obesity, Alzheimer's disease, and smoking status (Figure 6). We compared measures of "age-acceleration" in these groups using both our own chronological 450k clock and the published PhenoAge (Levine et al., 2018) "biological" clock. PhenoAge predicted most samples were age-decelerated, with multiple sclerosis, obesity, and depression being significantly "accelerated" compared to controls (Figure 6a). Meanwhile, smoking status and HIV decreasing age acceleration seem contradictory to prior reports (Esteban-Cantos et al., 2021; Levine et al., 2018). Since the PhenoAge clock is regressed against 10-year mortality risk coerced into units of years, perhaps it is not surprising that even "healthy" control samples are predicted to be age-accelerated. In contrast, our chronological 450k clock had much closer to zero average age acceleration. Notable exceptions are the significant increase in "age acceleration" of patients with NAFLD or NASH (Figure 6b). In both cases, these analyses have an additional uncertainty term in their age predictions that make interpreting age acceleration values difficult, with most samples falling within the range of model error. A potential explanation for age acceleration predictions going the opposite direction one might expect is found in looking at trajectory of aging methylation values over the lifespan. Namely, some show





**FIGURE 6** Epigenetic age acceleration associations from chronological and biological clocks across many age-related states. We annotated 1767 samples for their control or potentially age-related states. (a) PhenoAge predicted age vs chronological age in the disease-annotated samples. (b), Our 450k clock predicted age vs chronological age. (b) and (c), Age acceleration distributions were compared using a one-way linear model with holm-adjusted t test post hoc tests. Significant differences ( $p < .05$ ) from pooled healthy controls marked with a red asterisk (\*). (c) Age acceleration as computed with PhenoAge (Levine et al., 2018). (d), Age acceleration computed with our 450k clock model. A table of the included experiments and their annotations can be found in the supplement (Supplemental S2). Abbreviations: MS—multiple sclerosis, C—aggregate controls from all included studies, De, depression; Ath, atherosclerosis; Ob, obese; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; PSP, progressive supranuclear palsy; FTD, fronto-temporal dementia; AD, Alzheimer's disease; Asth, Asthma; HIV, human immunodeficiency virus; PBC, primary biliary cholangitis; PSC, primary sclerosing cholangitis; Former, former smoker; Smoker, current smoker; DS, Down's Syndrome

a non-linear fit that is “hidden” to the linear elastic net regression (Figures S4 and S5). While not all epigenetic clock sites exhibit this non-linear relationship, we noted in our primary feature selection that selecting sites with the best linear relationship reduced performance much more than selecting sites by non-linear feature selection (Figure S2). This may explain why some age-related diseases appear to be “decelerated” due to the same methylation values being present earlier and later in life.

### 3 | DISCUSSION

Our meta-analysis of the largest available age-annotated methylation dataset to date found: 1) as much as one fifth of the measured cytosines contains age-predictive methylation patterns; 2) tissues

show largely similar aging patterns despite having methylated regions that define their identity; 3) epigenetic clock sites are enriched in intergenic regions, gene enhancers and sites near eQTLs and 4) are depleted in the regions generally thought to have the largest direct impact upon gene expression (e.g., CpG Islands and gene promoters); 5) patients with age-correlated diseases did not appear significantly age-accelerated according to the chronological epigenetic clock.

The fact that many different sites can be used to create an epigenetic clock with minimal impact on predictive performance argues against the idea that methylation changes are either programmed or individually important. Yet, because the clock is robustly predictive and age-related methylation changes are mostly similar between tissues, this argues against entropy as a driving force. This could be reconciled by hypothesizing some genomic regions and/or features receive less methylation maintenance than others. Perhaps the



changes occur in regions of the genome where they have no consequence, and instead, vary with absolute time such as in determining speciation time using pseudogene mutation rates. This “pseudo-methylation” would be problematic for modeling aging biology, as they would likely not respond to aging intervention. Methylation maintenance mechanisms (e.g., DNMT1) serve as a counterbalance against entropy. However, if some genomic regions are less maintained than others, then we would expect the probability of a methylation state change with age to be correlated with the degree to which it is subject to methylation surveillance and maintenance. Because maintenance costs energy, it is reasonable to hypothesize the degree of maintenance correlates with the adverse impact an unregulated change in methylation would cause. If so, the probability a site's methylation will vary with age would inversely correlate with its impact on an organism's survival.

It is interesting that in spite of tissue aging interactions being rare in the age-related differentially methylated loci (Figure 3A and B), training clocks on specific tissues selects loci that poorly predict other tissues (3C). This is further seen in the case study on age-related neurological diseases, where the disease-associated methylation sites are also rarely clock sites (Figure 6b). The tissue-independent aging loci that are selected by clocks trained on multiple tissues are depleted in regions canonically associated with DNA methylation, namely promoters and CpG islands (Figure 5a), and simultaneously defining age acceleration using these clocks poorly predicts age-related disease in our case study. Although combining the knowledge that these disease-afflicted samples are predicted to be “younger” and the identification of non-linear methylation changes in clock sites, perhaps these values of age acceleration are rather measuring proper compensation by the system. This would be consistent with the observation that clock sites are significantly enriched in open chromatin and TEAD-binding regions (TEAD requires co-factors to act, plus literature-mining analysis (Marioni et al., 2019) of TEAD-binding regions near genes merely suggest TATA-binding proteins as a commonality, which is also a general motif). Under this assumption, being epigenetically “younger” could be a mixture of those failing to compensate (and thus died and had tissue collected to be measured). This could also simultaneously make aging interventions show “decompensation” as they are no longer needing to respond to the pressures of aging and thus the samples would be predicted to be younger.

Given that methylation changes with age are robust across tissues, yet small in magnitude, leads the field to question whether the “ticking” that drives them is due to changes in cell population composition, such as a reduction of pluripotent stem cells or an increase in senescent cells within every tissue, or possibly high magnitude effects in rare cell populations (e.g., immune cells in the CNS compared to astrocytes/neurons). In either case, it is not clear whether the phenomenon driving ticking clock sites is due to healthy compensatory changes or deleterious drift toward age-related fragility. To address the whole tissue versus individual cell type hypothesis, we are currently working on an aging study using mice with cell type-specific markers that will allow whole-genome sequencing

from specific cell types. By comparing the mouse clocks' predicted ages between different cell types, we hope to identify if the clock is indeed a pan-tissue phenomenon or is affecting some subset of cells “contaminating” all tissues. We are also working on analyzing paired senescent and non-senescent cells from the same patients to determine if senescent cells are driving the clock's predictive accuracy. These analyses will use BS-Seq instead of methylation arrays, allowing us to simultaneously explore how robust our biological enrichments are when determined using whole-genome sequencing data.

Our finding that the relationship between age-related disease and age acceleration seems contradictory to other research (Fransquet et al., 2019; Levine et al., 2015). However, these prior publications show average age acceleration values for patients that are within the error of predictive accuracy (i.e., disease samples are “age accelerated” by less than the ~4 year error range of epigenetic clocks). While we cannot provide a resolution to this dilemma with the currently available data, it should act as a caution when evaluating “age-acceleration” as a researcher using smaller subsets of samples, especially when the used clock is not trained on similar sample types with similar distributions of methylation values.

In summary, the predictive power of the epigenetic clock is robust, but such a large fraction of the genome can be used to predict, the magnitude of the changes is small, and these regions tend to be depleted near genes. This leads us to hypothesize that the pan-tissue predictive loci are more likely to be molecularly “silent” methylation changes that accrue outside of strong regulatory regions due to entropy in methylation maintenance, which must be explored in the future studies. Furthermore, if current models inconsistently annotate patients with age-related diseases as “age-accelerated” and the confidence by which one can declare a sample age-accelerated is small, this argues against the idea that epigenetic clocks can disentangle biological age from chronological age.

## 4 | EXPERIMENTAL PROCEDURES

### 4.1 | Data and Label Collection

Raw data were collected from NCBI's Gene Expression Omnibus (GEO). For replicating Horvath's experiments, individual datasets were extracted and manually curated from metadata using `geoquery` (Davis & Meltzer, 2007). Of these, only 20 datasets were available and matched the given age distributions and sample numbers as reported in Horvath's original report (Horvath, 2013). These data make up dataset 1, used for the direct Horvath replication (Figure 2b, d) and the iterative trimming models (Figure 2e). Our second dataset consists of all publically available data from the Illumina 450k Human Methylation BeadChip Array that could be annotated by our label extraction program ALE (Giles et al., 2017), with accompanying metadata as presented in `geometadb` (Zhu et al., 2008). Sample labels for all samples' sex, age, and tissue were extracted from text using our previously published tool ALE (Giles et al., 2017). We then exclude any samples with annotated ages under 25 for two reasons.



First, we see aging as a process that begins post-development, with human development ending around 25 years of age. Second, our annotation model is often incorrect about units, and as a result, our confidence in ages <25 is much lower than the rest of the lifespan due to common age labels with units of weeks/months falling in this range. Disease annotations were hand curated by reading their metadata and the metadata of the associated GSEs. A table of annotations can be found in Supplemental Data S2.

## 4.2 | Data Preprocessing

The replication dataset for Horvath's clock model was preprocessed using code from Horvath 2013 to ensure identical normalization and imputation. We also used the same age transformation as reported in Horvath 2013 (Horvath, 2013). The full 450k data were imputed in a similar pipeline, using KNN imputation in sets of 120,000 probes with subsequent normalization. For our linear models, batch effect correction was performed using ComBat (Leek et al., 2012) to control for experiment ID after removing samples which did not fall in the beta distribution.

## 4.3 | Feature Selection

Due to the large number of samples and probes included in the full 450k dataset, modeling could not be performed on the full dataset simultaneously. As such, we tested many feature selection pipelines in the smaller Horvath subset to determine their ability to preserve useful sites for age prediction (Figure S2). This led us to using mutual information as the primary feature selection method, from which we selected the top ~20% of sites and using those 79931 loci for all downstream analyses.

F-regression and mutual information were performed using sklearn's (Fabian Pedregosa et al., 2011) implementation on one quarter of all loci at a time due to memory constraints. Top 10% delta mean and variance were computed by comparing the mean and variance differences for each locus between the young (25–35) and aged (65–100) sets of samples.

## 4.4 | Statistical Analysis

Elastic net regression was utilized to generate clock models, as in Horvath 2013. Horvath's replication was performed using R's glmnet (Friedman et al., 2009), while the full 450k model was performed using sci-kit learn's ElasticNetCV (Fabian Pedregosa et al., 2011). In both cases, 10-fold cross-validation modeling was used with the L1 vs L2 selection parameter at 0.5 (elastic net), and the selectivity parameter adjusted based on cross-validation.

Age-related locus analysis was performed using OLS regression in the model methylation ~ age + tissue + age:tissue. After FDR multiple testing correction, 676 loci were left for downstream analyses.

PhenoAge predictions were determined as described in Levine, et al. (Levine et al., 2018).

## 4.5 | Genomic Feature and Motif Enrichment

Identifying gene, gene-related, and CpG island-related enrichments were performed by using bedtools to annotate sets of loci with features from UCSC. Feature enrichments were computed based on hypergeometric tests comparing the various groups with a 10% FDR correction for multiple testing.

Motif enrichments were performed using both HOMER (Heinz et al., 2010) and MEME (Machanick & Bailey, 2011). Fasta files were constructed with the 50 bp sequence of each probe on the Illumina 450k Array for all loci, and age-related loci from mutual information, elastic net, and OLS regression models. These loci were then analyzed for known and *de novo* motifs, with enrichments calculated against the background of all sites on the array.

Identifying enrichments in genomic features were based on data from UCSC genome browser. CpG islands and genic regions were defined using UCSC's annotated island bodies and genes, with promoters being 2kb upstream of the TSS and shore/shelves being defined as 2kb blocks up and downstream of the island body. Simultaneous comparison of all regulatory elements was performed using ORegAnno regulatory elements (Griffith et al., 2007). Vista enhancers were used to look specifically at known human enhancers (Visel et al., 2006). eQTLs were taken from GTEx (Lonsdale et al., 2013). Chromatin density was inferred using DNase hypersensitivity data from ENCODE DNase-seq (Consortium & E.P., 2004).

### ACKNOWLEDGEMENTS

The authors would like to thank the four anonymous peer-reviewers for their helpful comments and feedback.

### CONFLICT OF INTEREST

The authors declare no conflicts of interest.

### AUTHOR CONTRIBUTIONS

H.L.P., W.M.F., and J.D.W. conceived and designed the study. W.M.F., C.G., and J.D.W. supervised the study. H.L.P., C.A.B., X.R., and C.B.G. produced code for sample processing and analysis. H.L.P. and C.G. performed statistical analyses. H.L.P., W.M.F., and J.D.W. wrote the manuscript. All authors discussed the results and commented on the manuscript.

### DATA AVAILABILITY STATEMENT

The datasets analyzed in the current study are available in NCBI GEO. These datasets are all from GPL13534 and can be obtained from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL13534>.

### ORCID

Hunter L. Porter  <https://orcid.org/0000-0002-2495-2423>

Jonathan D. Wren  <https://orcid.org/0000-0003-2776-3545>



## REFERENCES

- Bell, C. G., Lowe, R., Adams, P. D., Baccarelli, A. A., Beck, S., Bell, J. T., Christensen, B. C., Gladyshev, V. N., Heijmans, B. T., Horvath, S., Ideker, T., Issa, J. P. J., Kelsey, K. T., Marioni, R. E., Reik, W., Relton, C. L., Schalkwyk, L. C., Teschendorff, A. E., Wagner, W., ... Rakyanc, V. K. (2019). DNA methylation aging clocks: challenges and recommendations. *Genome Biology*, 20(1), 249. <https://doi.org/10.1186/s13059-019-1824-y>
- Calude, C. S. (2017). Longo, G, *The Deluge of Spurious Correlations in Big Data. Foundations of Science*, 22, 595–612. <https://doi.org/10.1007/s10699-016-9489-4>
- Chen, B. H., Marioni, R. E., Colicino, E., Peters, M. J., Ward-Caviness, C. K., Tsai, P.-C., Roetker, N. S., Just, A. C., Demerath, E. W., Guan, W., Bressler, J., Fornage, M., Studenski, S., Vandiver, A. R., Moore, A. Z., Tanaka, T., Kiel, D. P., Liang, L., Vokonas, P., ... Horvath, S. (2016). DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY)*, 8(9), 1844–1865. <https://doi.org/10.18632/aging.101020>
- Consortium, E.P. (2004). The ENCODE (Encyclopedia of DNA elements) project. *Science*, 306(5696), 636–640.
- Davis, S., & Meltzer, P. S. (2007). GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*, 23(14), 1846–1847. <https://doi.org/10.1093/bioinformatics/btm254>
- Esteban-Cantos, A., Rodríguez-Centeno, J., Barruz, P., Alejos, B., Saiz-Medrano, G., Nevado, J., Martín, A., Gayá, F., De Miguel, R., Bernardino, J. I., Montejano, R., Mena-Garay, B., Cadiñanos, J., Florence, E., Mulcahy, F., Banhegyi, D., Antinori, A., Pozniak, A., Wallet, C., ... Arribas, J. R. (2021). Epigenetic age acceleration changes 2 years after antiretroviral therapy initiation in adults with HIV: a substudy of the NEAT001/ANRS143 randomised trial. *Lancet HIV*, 8(4), e197–e205. [https://doi.org/10.1016/S2352-3018\(21\)00006-0](https://doi.org/10.1016/S2352-3018(21)00006-0)
- Fabian Pedregosa, G. V., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. Scikit-learn: Machine Learning in Python. *JMLR*, 2011:2825–2830
- Fahy, G. M., Brooke, R. T., Watson, J. P., Good, Z., Vasanawala, S. S., Maecker, H., Leipold, M. D., Lin, D. T. S., Kobar, M. S., & Horvath, S. (2019). Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell*, e13028. <https://doi.org/10.1111/accel.13028>
- Fransquet, P. D., Wrigglesworth, J. O., Woods, R. L., Ernst, M. E., & Ryan, J. (2019). The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. *Clin Epigenetics*, 11(1), 62. <https://doi.org/10.1186/s13148-019-0656-7>
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1–22. <http://dx.doi.org/10.18637/jss.v033.i01>
- Giles, C. B., Brown, C. A., Ripperger, M., Dennis, Z., Roopnarinesingh, X., Porter, H., Perz, A., & Wren, J. D. (2017). ALE: automated label extraction from GEO metadata. *BMC Bioinformatics*, 18(Suppl 14), 509. <https://doi.org/10.1186/s12859-017-1888-1>
- Griffith, O. L., Montgomery, S. B., Bernier, B., Chu, B., Kasaiyan, K., Aerts, S., Mahony, S., Sleumer, M. C., Bilenyk, M., Haeussler, M., Griffith, M., Gallo, S. M., Gardine, B., Hooghe, B., Van Loo, P., Blanco, E., Ticoll, A., Lithwick, S., Portales-Casamar, E., ... Jones, S. J. (2007). ORegAnno: an open-access community-driven resource for regulatory annotation. *Nucleic Acids Research*, 36(suppl\_1), D107–D113. <https://doi.org/10.1093/nar/gkm96>
- Hadad, N., Masser, D. R., Logan, S., Wronowski, B., Mangold, C. A., Clark, N., & Freeman, W. (2016). Absence of genomic hypomethylation or regulation of cytosine-modifying enzymes with aging in male and female mice. *Epigenetics Chromatin*, 9, 30.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., & Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell*, 38(4), 576–589. <https://doi.org/10.1016/j.molcel.2010.05.004>
- Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, 14(10), R115. <https://doi.org/10.1186/gb-2013-14-10-r115>
- Horvath, S., & Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics*, 19(6), 371. <https://doi.org/10.1038/s41576-018-0004-3>
- Koh, K. P., & Rao, A. (2013). DNA methylation and methylcytosine oxidation in cell fate decisions. *Current Opinion in Cell Biology*, 25(2), 152–161. <https://doi.org/10.1016/j.jceb.2013.02.014>
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, 28(6), 882–883. <https://doi.org/10.1093/bioinformatics/bts034>
- Levine, A. J., Quach, A., Moore, D. J., Achim, C. L., Soontornniyomkij, V., Masliah, E., Singer, E. J., Gelman, B., Nemanim, N., & Horvath, S. (2016). Accelerated epigenetic aging in brain is associated with pre-mortem HIV-associated neurocognitive disorders. *J Neurovirol*, 22(3), 366–375. <https://doi.org/10.1007/s13365-015-0406-3>
- Levine, M. E., Lu, A. T., Bennett, D. A., & Horvath, S. (2015). Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. *Aging (Albany NY)*, 7(12), 1198–1211. <https://doi.org/10.18632/aging.100864>
- Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., Hou, L., Baccarelli, A. A., Stewart, J. D., Li, Y., Whitsel, E. A., Wilson, J. G., Reiner, A. P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L., & Horvath, S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*, 10(4), 573–591. <https://doi.org/10.18632/aging.101414>
- Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., & Foster, B. (2013). The genotype-tissue expression (GTEx) project. *Nature Genetics*, 45(6), 580.
- Lowe, D., Horvath, S., & Raj, K. (2016). Epigenetic clock analyses of cellular senescence and ageing. *Oncotarget*, 7(8), 8524. <https://doi.org/10.18632/oncotarget.7383>
- Macaluso, M., & Giordano, A. (2004). How does DNA methylation mark the fate of cells? *Tumori Journal*, 90(4), 367–372.
- Machanic, P., & Bailey, T. L. (2011). MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics*, 27(12), 1696–1697. <https://doi.org/10.1093/bioinformatics/btr189>
- Marioni, R. E., Suderman, M., Chen, B. H., Horvath, S., Bandinelli, S., Morris, T., Beck, S., Ferrucci, L., Pedersen, N. L., Relton, C. L., Deary, I. J., & Hägg, S. (2019). Tracking the Epigenetic Clock Across the Human Life Course: A Meta-analysis of Longitudinal Cohort Data. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 74(1), 57–61. <https://doi.org/10.1093/gerona/gly060>
- McKinney, B. C., Lin, H., Ding, Y., Lewis, D. A., & Sweet, R. A. (2018). DNA methylation age is not accelerated in brain or blood of subjects with schizophrenia. *Schizophrenia Research*, 196, 39–44. <https://doi.org/10.1016/j.schres.2017.09.025>
- Perna, L., Zhang, Y., Mons, U., Holleczek, B., Saum, K. U., & Brenner, H. (2016). Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin Epigenetics*, 8, 64.
- Sargent, J. (2015). Obesity turns the hepatic epigenetic clock forward. *Nature Reviews Endocrinology*, 11(1), 2–3. <https://doi.org/10.1038/nrendo.2014.198>
- Shireby, G. L., Davies, J. P., Francis, P. T., Burrage, J., Walker, E. M., Neilson, G. W. A., Dahir, A., Thomas, A. J., Love, S., Smith, R. G., Lunnon, K., Kumari, M., Schalkwyk, L. C., Morgan, K., Brookes, K., Hannon, E., & Mill, J. (2020). Recalibrating the epigenetic clock:



- implications for assessing biological age in the human cortex. *Brain*, 143(12), 3763–3775. <https://doi.org/10.1093/brain/awaa334>
- Slieker, R. C., van Iterson, M., Luijk, R., Beekman, M., Zhernakova, D. V., Moed, M. H., Mei, H., van Galen, M., Deelen, P., Bonder, M. J., Zhernakova, A., Uitterlinden, A. G., Tigchelaar, E. F., Stehouwer, C. D. A., Schalkwijk, C. G., van der Kallen, C. J. H., Hofman, A., van Heemst, D., de Geus, E. J., ... Heijmans, B. T. (2016). Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms. *Genome Biology*, 17(1), 191. <https://doi.org/10.1186/s13059-016-1053-6>
- Starnawska, A., Tan, Q., Lenart, A., McGue, M., Mors, O., Børjglum, A. D., Christensen, K., Nyegaard, M., & Christiansen, L. (2017). Blood DNA methylation age is not associated with cognitive functioning in middle-aged monozygotic twins. *Neurobiology of Aging*, 50, 60–63. <https://doi.org/10.1016/j.neurobiolaging.2016.10.025>
- Thompson, M. J., Chwiałkowska, K., Rubbi, L., Lusi, A. J., Davis, R. C., Srivastava, A., Korstanje, R., Churchill, G. A., Horvath, S., & Pellegrini, M. (2018). A multi-tissue full lifespan epigenetic clock for mice. *Aging (Albany NY)*, 10(10), 2832–2854. <https://doi.org/10.18632/aging.101590>
- Tsika, R. W., Ma, L., Kehat, I., Schramm, C., Simmer, G., Morgan, B., Fine, D. M., Hanft, L. M., McDonald, K. S., Molkentin, J. D., Krenz, M., Yang, S., & Ji, J. (2010). TEAD-1 overexpression in the mouse heart promotes an age-dependent heart dysfunction. *Journal of Biological Chemistry*, 285(18), 13721–13735. <https://doi.org/10.1074/jbc.M109.063057>
- Tsygankov, D., Liu, Y., Sanoff, H. K., Sharpless, N. E., & Elston, T. C. (2009). A quantitative model for age-dependent expression of the p16INK4a tumor suppressor. *Proc Natl Acad Sci U S A*, 106(39), 16562–16567. <https://doi.org/10.1073/pnas.0904405106>
- Tuttle, C. S. L., Waaijer, M. E. C., Slee-Valentijn, M. S., Stijnen, T., Westendorp, R., & Maier, A. B. (2020). Cellular senescence and chronological age in various human tissues: A systematic review and meta-analysis. *Aging Cell*, 19(2), e13083. <https://doi.org/10.1111/accel.1308>
- Visel, A., Minovitsky, S., Dubchak, I., & Pennacchio, L. A. (2006). VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Research*, 35(suppl\_1), D88–D92. <https://doi.org/10.1093/nar/gkl822>
- Wang, T., Tsui, B., Kreisberg, J. F., Robertson, N. A., Gross, A. M., Yu, M. K., Carter, H., Brown-Borg, H. M., Adams, P. D., & Ideker, T. (2017). Epigenetic aging signatures in mice livers are slowed by dwarfism, calorie restriction and rapamycin treatment. *Genome Biology*, 18(1), 57. <https://doi.org/10.1186/s13059-017-1186-2>
- Ward-Caviness, C. K., Nwanaji-Enwerem, J. C., Wolf, K., Wahl, S., Colicino, E., Trevisi, L., Kloog, I., Just, A. C., Vokonas, P., Cyrus, J., Gieger, C., Schwartz, J., Baccarelli, A. A., Schneider, A., & Peters, A. (2016). Long-term exposure to air pollution is associated with biological aging. *Oncotarget*, 7(46), 74510–74525. <https://doi.org/10.18632/oncotarget.12903>
- Weidner, C., Lin, Q., Koch, C., Eisele, L., Beier, F., Ziegler, P., Bauerschlag, D., Jöckel, K.-H., Erbel, R., Mühleisen, T., Zenke, M., Brümmendorf, T., & Wagner, W. (2014). Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biology*, 15(2), R24. <https://doi.org/10.1186/gb-2014-15-2-r24>
- Xie, Q. I., Chen, J., Feng, H., Peng, S., Adams, U., Bai, Y., Huang, L. I., Li, J. I., Huang, J., Meng, S., & Yuan, Z. (2013). YAP/TEAD-mediated transcription controls cellular senescence. *Cancer Research*, 73(12), 3615–3624. <https://doi.org/10.1158/0008-5472.CAN-12-3793>
- Yang, Z., Wong, A., Kuh, D., Paul, D. S., Rakyen, V. K., Leslie, R. D., Zheng, S. C., Widschwendter, M., Beck, S., & Teschendorff, A. E. (2016). Correlation of an epigenetic mitotic clock with cancer risk. *Genome Biology*, 17(1), 205. <https://doi.org/10.1186/s13059-016-1064-3>
- Zheng, S. C., Widschwendter, M., & Teschendorff, A. E. (2016). Epigenetic drift, epigenetic clocks and cancer risk. *Epigenomics*, 8(5), 705–719. <https://doi.org/10.2217/epi-2015-0017>
- Zhu, Y., Davis, S., Stephens, R., Meltzer, P. S., & Chen, Y. (2008). GEOmetadb: powerful alternative search engine for the Gene Expression Omnibus. *Bioinformatics*, 24(23), 2798–2800. <https://doi.org/10.1093/bioinformatics/btn520>

#### SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Porter, H. L., Brown, C. A., Roopnarinesingh, X., Giles, C. B., Georgescu, C., Freeman, W. M., & Wren, J. D. (2021). Many chronological aging clocks can be found throughout the epigenome: Implications for quantifying biological aging. *Aging Cell*, 20, e13492. <https://doi.org/10.1111/accel.13492>