

## Development of a novel epigenetic clock resistant to changes in immune cell composition

Alan Tomusiak<sup>1,2</sup>, Ariel Floro<sup>1,2</sup>, Ritesh Tiwari<sup>1</sup>, Rebeccah Riley<sup>1</sup>, Hiroyuki Matsui<sup>1</sup>, Nicolas Andrews<sup>1</sup>, Herbert G. Kasler<sup>1</sup> and Eric Verdin<sup>1\*</sup>

<sup>1</sup>\*Buck Institute for Research on Aging, 8001 Redwood Blvd, Novato, 94945, California, USA.

<sup>2</sup>Department of Gerontology, University of Southern California, 3715 McClintock Ave, Los Angeles, 90089, California, USA.

\*Corresponding author: [everdin@buckinstitute.org](mailto:everdin@buckinstitute.org)

Contributing authors: [atomusiak6@gmail.com](mailto:atomusiak6@gmail.com); [afloro@buckinstitute.org](mailto:afloro@buckinstitute.org);  
[rtiwari@buckinstitute.org](mailto:rtiwari@buckinstitute.org); [rriley@buckinstitute.org](mailto:rriley@buckinstitute.org); [hmatsui@buckinstitute.org](mailto:hmatsui@buckinstitute.org);  
[nicolasandrews1@gmail.com](mailto:nicolasandrews1@gmail.com); [hkasler@buckinstitute.org](mailto:hkasler@buckinstitute.org)

## Abstract

Epigenetic clocks are age predictors that use machine-learning models trained on DNA CpG methylation values to predict chronological or biological age. Increases in predicted epigenetic age relative to chronological age (epigenetic age acceleration) are connected to aging-associated pathologies, and changes in epigenetic age are linked to canonical aging hallmarks. However, epigenetic clocks rely on training data from bulk tissues whose cellular composition changes with age. We found that human naive CD8<sup>+</sup> T cells, which decrease during aging, exhibit an epigenetic age 15–20 years younger than effector memory CD8<sup>+</sup> T cells from the same individual. Importantly, homogenous naive T cells isolated from individuals of different ages show a progressive increase in epigenetic age, indicating that current epigenetic clocks measure two independent variables, aging and immune cell composition. To isolate the age-associated cell intrinsic changes, we created a new clock, the IntrinsicClock, that did not change among 10 immune cell types tested. IntrinsicClock showed a robust predicted epigenetic age increase in a model of replicative senescence *in vitro* and age reversal during OSKM-mediated reprogramming.

**Keywords:** aging, epigenetic clock, immune, inflammaging, differentiation, DNA methylation

## Introduction

Epigenetic clocks, age predictors based on DNA methylation levels at selected CpG loci, have grown in popularity as a tool to study aging and predict health outcomes in humans. The first epigenetic clocks developed by Hannum et al.<sup>1</sup> and Horvath<sup>2</sup> showed remarkably high accuracy ( $R > .90$ ) in predictions of chronological age. These “first-generation” epigenetic clocks provide unique biological insights into the aging process. For example, some but not all forms of senescence accelerate epigenetic clock age predictions<sup>3</sup>. Using a later clock trained on chronological age, Kabacik et al.<sup>4</sup> identified nutrient sensing, mitochondrial activity and stem cell composition as being associated with epigenetic aging but not telomere attrition or genomic instability. A recent report demonstrated the development of an epigenetic clock effective at predicting age across a variety of species, providing evidence for a shared mammalian aging program<sup>5</sup>.

More recently, “second-generation” clocks designed to predict phenotypic aging measures have been developed. These clocks, including PhenoAge<sup>6</sup> and GrimAge<sup>7</sup>, show strong associations with diseases, such as depression<sup>8</sup> and mortality<sup>9</sup>. DunedinPACE is a similar marker of phenotypic aging that captures the pace of aging rather than the accumulation of aging<sup>10</sup>. These clocks show promise as markers of physiological aging, but their two-step construction methodology (training a DNA methylation predictor on measures of phenotypic rather than chronological age) adds a secondary layer of complexity to interpretation.

Given the ability of epigenetic clocks to detect aging phenotypes across species and levels of organization that include cells, tissues, and organs, there is significant interest in understanding the underlying mechanism(s) enabling their function. Recent preprints have been released on this topic, notably including one by Levine et al.<sup>11</sup> that suggests epigenetic clocks are composites of different modules characterized by their changes during aging and reprogramming. Novel epigenetic clocks have been developed that seek to capture the aging phenomenon in more defined ways, including by identifying CpG sites predicted to be causal by Mendelian randomization<sup>12</sup> or those capturing purely stochastic variation<sup>13</sup>. These clocks are informative about aspects of the aging process and have the potential to be particularly well-suited for certain use cases.

One major challenge in understanding the mechanism(s) underlying epigenetic clocks is the

confounding effect of age-related changes in cell-type composition of many tissues. While changes in cell-type composition are an important part of aging, they can make interpreting epigenetic clocks more difficult as the relevant CpG sites may be cell-type-specific markers rather than those affecting cell-intrinsic aging. Most epigenetic clocks are trained largely on blood, which sees a drop in naïve CD8<sup>+</sup> T cells with age and a corresponding increase in more terminally differentiated memory T-cell types<sup>14</sup>. Some clocks may be more impacted by changes in cell-type composition than others, depending on how they were constructed<sup>15</sup>. T-cell and NK (natural killer) cell activation have been implicated as major drivers in epigenetic clock progression<sup>16</sup>.

Other approaches have been explored to create epigenetic age predictions that are less sensitive to changes in cell type composition. Most notably, residuals from regression models that include epigenetic age and proportions of several blood cell types have been used to generate an “intrinsic epigenetic age acceleration” measure<sup>17</sup>. While the resulting measure is cell-type independent, it becomes challenging to biologically interpret as the underlying signal is derived from a mixture of CpG sites that can be either cell type-independent or cell type-dependent. Other modern approaches include the development of single-cell epigenetic clocks<sup>18,19</sup>, though the underlying technology will require further maturing before it can match the sensitivity and accuracy of bulk measurement-based clocks.

In this work, we report our analysis of the differences in epigenetic age predictions derived from four epigenetic clocks (Hannum<sup>1</sup>, Horvath<sup>2</sup>, Horvath Skin and Blood<sup>20</sup>, and PhenoAge<sup>6</sup>) for cytotoxic CD8<sup>+</sup> T cells at different stages of differentiation. We found that human naïve CD8<sup>+</sup> T cells, which decrease in humans during aging, exhibit an epigenetic age 15–20 years younger than effector memory CD8<sup>+</sup> T cells isolated from the same individual. Interestingly, naïve T cells isolated from individuals of different ages still show a progressive increase in epigenetic age. Based on these observations, which indicate, as predicted, that current epigenetic clocks measure two independent variables, aging and immune cell composition, we created a new clock, the IntrinsicClock, that does not change among 10 immune cell types tested. Remarkably, this clock shows an increase in a model of replicative senescence *in vitro* and shows decreased aging during OSKM reprogramming. Lastly, we investigate the IntrinsicClock’s applicability for use in studying and detecting the effects of cell-intrinsic perturbations on aging.

## Results

**Existing epigenetic clock age predictions depend on CD8<sup>+</sup> T-cell differentiation state.** In humans, CD8<sup>+</sup> T cells decrease in frequency, with a particularly pronounced loss of naïve T cells during aging<sup>21</sup>. We used a negative bead-based selection method to isolate total T cells from seven donors of varying ages, all of whom were positive for cytomegalovirus (CMV<sup>+</sup>). We then used FACS to isolate CD8<sup>+</sup> naïve (CD8<sup>+</sup> CD28<sup>+</sup> CD45RO<sup>-</sup>), CD8<sup>+</sup> central memory (CD8<sup>+</sup> CD28<sup>+</sup> CD45RO<sup>+</sup>), CD8<sup>+</sup> effector memory (CD8<sup>+</sup> CD28<sup>-</sup> CD45RO<sup>+</sup>), and CD8<sup>+</sup> terminal effector memory RA<sup>+</sup> (CD8<sup>+</sup> CD28<sup>-</sup> CD45RO<sup>-</sup>) cells (Figure 1A). After DNA isolation and profiling using the Illumina Infinium MethylationEPIC<sup>TM</sup> platform, we noted a distinct clustering of CD8<sup>+</sup> naïve cells away from CD8<sup>+</sup> central memory (CM), effector memory (EM), and terminal effector memory RA<sup>+</sup> cells (TEMRA) (Figure 1B) in UMAP analysis. Horvath clock epigenetic ages were measured in each of the CD8 T-cell subsets and found to correlate with age across every subset. However, strikingly, naïve T cells consistently showed a significantly younger epigenetic age than other CD8<sup>+</sup> subsets (Figure 1C). This result suggests that epigenetic clock measurements are affected by CD8<sup>+</sup> T-cell differentiation. Equally interestingly, naïve CD8<sup>+</sup> T cells from individuals of different chronological age showed an increase in epigenetic age that was parallel to chronological age but consistently lower than the chronological age (Figure 1C). The

same observation was made for CMs, EMs and TEMRAs except that these cells' epigenetic age appeared closer to the chronological age of the donors.

Next, using differential methylation analysis on methylation M-values, we identified 22,963 CpGs that changed with age and 370,383 CpGs that changed between naive CD8<sup>+</sup> T cells and CD8<sup>+</sup> CM, CD8<sup>+</sup> EM, or CD8<sup>+</sup> TEMRA cells. Of the 22,963 aging-related CpGs, 9,992 were also affected by differentiation (Figure 1D). To understand how this could affect epigenetic clock predictions, we investigated the proportion of CpG sites used for epigenetic age prediction in the Hannum, Horvath, Horvath Skin and Blood, and PhenoAge clocks that we identified were affected by CD8<sup>+</sup> T-cell differentiation. In all four clocks, more than a third of the predictive sites were changed with differentiation (Figure 1E), and all four had a difference in age acceleration for CD8<sup>+</sup> T-cell subsets. In all clocks, CD8<sup>+</sup> TEMRA and CD8<sup>+</sup> EM cells were predicted to be older than CD8<sup>+</sup> CM cells, which were predicted to be older than CD8<sup>+</sup> naive cells (Figures 1F-1I). The differences in epigenetic ages among the CD8<sup>+</sup> T-cell subsets varied among clocks. For example, PhenoAge predicts CD8<sup>+</sup> naive cells to be over 60 years younger than the donor chronological age, but the difference was much smaller for both Horvath clocks with an epigenetic age prediction of only approximately 12 years lower than chronological age (Figures 1F - 1I).

**Development of a novel epigenetic clock (IntrinClock) resistant to changes in immune cell composition.** Given the overlap of DNA methylation signatures of cellular aging and CD8<sup>+</sup> differentiation, we sought to create a new epigenetic clock that is unaffected by changes in immune cell composition. We began by generating a database of 14,601 DNA methylation samples from 71 different datasets<sup>1,22-90</sup>, generated on either the Illumina Infinium™ HumanMethylation450 (450K) or the Illumina Infinium™ MethylationEPIC (EPIC) array, all sourced from the Gene Expression Omnibus (GEO) database or the Genotype-Tissue Expression project (GTEx) (Supplementary Table 1). The number of samples per dataset ranged from six to 1,218, with a mean number of samples per dataset of 213 (Figure 4A). The distribution of sexes was approximately equal (Figure S4B). Samples were derived from a variety of tissues with the majority from blood (Figure S4C), and the DNA methylation assay platform was split roughly evenly between the 450K and the EPIC array. (Figure S4D).

Once the database of samples was assembled, we performed a series of filtering and quality control steps. We filtered out all samples that were missing more than 10% of CpG sites measured by the 450K array, those that were derived from cancerous tissue, and those that were derived from germline tissues. We then removed outliers, defining outliers as those with principal components more than two interquartile ranges away from the mean (Figure 2B). After performing a random 75-25 training/test split, 9104 samples were used to train the model and 2994 were used to validate it.

Given the unique methylation pattern (Figure 1B) and quiescent biology<sup>91</sup> of naive CD8<sup>+</sup> T cells, we aimed to use them as a basis on which to eliminate CpGs linked to CD8<sup>+</sup> T-cell differentiation and performed additional filtering steps. When constructing our database of DNA methylation data, we initially collected all CpG sites measured by the 450K array for all samples. To increase reliability, we first filtered out CpG sites that were present in fewer than 90 percent of samples. To ensure forward compatibility, we also included only CpG sites that were present on the Illumina Infinium™ MethylationEPICv2.0 array. Next, we opted to remove any CpG sites that were correlated with a sample being a naive CD8<sup>+</sup> sample ( $R > .3$ ) within our CD8<sup>+</sup> subset data (i.e., CpG sites whose methylation patterns were distinct in CD8<sup>+</sup> naive cells as compared to CD8<sup>+</sup> CM/EM/TEMRA cells). We also opted to include only those CpG sites correlated with age ( $R > .3$ ) (Figure 2C), to decrease the search space for the elastic net algorithm to identify age-predictive sites. Interestingly, we observed a negative correlation-of-correlations between the age correlation and naive CD8<sup>+</sup> correlation of CpG sites ( $R = -$

.45) (Figure 2C), indicating that CpG sites that are hypermethylated with age tend to be hypomethylated in naive CD8<sup>+</sup> cells, and vice-versa. We utilized the elastic net algorithm on the remaining 55,896 CpGs to generate a new epigenetic clock based on 410 CpG sites. To increase accuracy and reduce the number of necessary prediction sites, we used a novel approach whereby we employed the elastic net algorithm a second time on the training data filtered only on the 410 CpG sites used for the clock. This reduced the number of predictive CpG sites in the final model (*IntrinClock*) to 381, and reduced error by approximately 3 months (Figure S5).

**IntrinClock is accurate across tissues, and its age predictions are not affected by adaptive immune cell compositional changes.** Next, we tested the *IntrinClock* on a variety of tissues in the test set and observed high overall prediction accuracy ( $R \sim .972$ , mean absolute error (MAE)  $\sim 3.83$ ) (Figure 2D). Age prediction errors on blood and saliva were particularly low (MAE  $\sim 3.25$ , MAE  $\sim 3.21$ , respectively) (Figure 2E, 2G). Tissues with less immune infiltration also had high epigenetic age correlations with chronological age ( $R \sim .944$  for brain,  $R \sim .841$  for skin). We were interested in discovering whether the *IntrinClock* would predict chronological age in semen samples, as previous epigenetic clocks have shown significant age deceleration in sperm<sup>2</sup>. We found that epigenetic age predictions of semen had only a weak correlation with chronological age ( $R \sim .32$ ), and the predicted age of sperm samples, using a previously generated dataset<sup>92</sup>, appears to consistently be  $\sim 12$  (Figure 2I).

Importantly and as expected, *IntrinClock* applied to our generated CD8<sup>+</sup> DNA methylation data showed no epigenetic age prediction differences among CD8<sup>+</sup> T-cell subsets (Figure 3A). As these samples were included in the training set for clock construction, we validated our approach on two external datasets<sup>93,94</sup> with CD8<sup>+</sup> naive and CD8<sup>+</sup> EM DNA methylation data and found no differences in epigenetic age (paired t-test p-value  $> .05$ ) (Figure 3B). We also tested whether our clock could find a shift in epigenetic age between CD4<sup>+</sup> naive and CD4<sup>+</sup> CM cells, as the proportion of CD4<sup>+</sup> naive cells also decreases with age<sup>95</sup>. Using two external data sets<sup>96,97</sup>, we discovered no evidence for a shift in epigenetic age between CD4<sup>+</sup> naive and CM cells (Figure 3C) (paired t-test p-value  $> .05$ ), despite our filtering strategy being based only on CD8<sup>+</sup> cells.

We also tested whether the *IntrinClock* would be similarly unperturbed in other immune cell types, particularly naive and memory B cells, which change in frequency with age<sup>98</sup>. We sorted CD8<sup>+</sup> naive (CD8<sup>+</sup>CD28<sup>+</sup>CD45RO<sup>-</sup>), CD8<sup>+</sup> CM (CD8<sup>+</sup>CD28<sup>+</sup>CD45RO<sup>+</sup>), CD8<sup>+</sup> combined EM/TEMRA (CD8<sup>+</sup>CD28<sup>+</sup>), CD4<sup>+</sup> naive (CD4<sup>+</sup>CD28<sup>+</sup>CD45RO<sup>-</sup>), CD4<sup>+</sup> CM (CD4<sup>+</sup>CD28<sup>+</sup>CD45RO<sup>+</sup>), B-cell naive (CD3<sup>-</sup>CD19<sup>+</sup>CD27<sup>+</sup>IgD<sup>+</sup>), class-switched B cells (CD3<sup>-</sup>CD19<sup>+</sup>CD27<sup>+</sup>IgD<sup>-</sup>), CD16<sup>+</sup>CD56<sub>dim</sub> NK cells (CD3<sup>-</sup>CD19<sup>-</sup>CD56<sub>dim</sub>CD16<sup>+</sup>), classical monocytes (CD3<sup>-</sup>CD19<sup>-</sup>HLADR<sup>+</sup>CD14<sup>+</sup>CD16<sub>dim</sub>), and whole-peripheral blood mononuclear cell (PBMC) samples from nine donors aged 30–68 and collected DNA for methylation analysis. To increase cell recovery, we performed two sequential rounds of positive selection for CD8<sup>+</sup> and then CD4<sup>+</sup> cells using magnetic enrichment kits prior to flow sorting, similar to a published strategy<sup>33</sup>. Concurrently, we analyzed the PBMC samples using high-parameter spectral flow cytometry to empirically determine whether changes in immune cell composition of the PBMC samples would impact predicted epigenetic age of the whole PBMC fraction.

As predicted, we found no evidence for an association between cell subset and epigenetic age prediction (Figure 3D) or between cell subset and epigenetic age acceleration (ANOVA p-value  $> .05$ ) (Figure S6A). This remained consistent whether epigenetic age acceleration was defined as the difference between predicted age and chronological age or as the residual after regressing predicted epigenetic age on chronological age. In contrast, cell subset and epigenetic age acceleration were significantly correlated, according to the Hannum (Figure S6B), Horvath (Figure S6C), Horvath Skin and Blood (Figure S6D), and PhenoAge (Figure S6E) clocks. To further investigate how resistant

IntrinClock is to the change in immune cell composition, we analyzed the correlation between the PBMC epigenetic age and percentage of several PBMC subsets. As expected, we identified no significant relationship between the PBMC epigenetic age acceleration and percentage of CD8<sup>+</sup> EM cells (Figure 3E), CD4<sup>+</sup> CM cells (Figure 3F), class-switched B cells (Figure 3G), CD16<sup>+</sup> CD56<sub>dim</sub> NK cells (Figure 3H), or classical monocytes (Figure 3I), relative to their parent populations (Pearson's correlation p-value > .05). Combined with our observations of the IntrinClock's high accuracy across many tissues, these observations indicate that shifts in immune cell composition do not impact IntrinClock age predictions.

**IntrinClock is highly enriched for CpG sites upstream of transcription start sites, and its sites are enriched for motifs whose TFs are implicated in cancer.** One central challenge in understanding epigenetic clocks comes from a lack of knowledge regarding to what extent epigenetic clocks are tracking a cell-autonomous or, conversely, a cell-ensemble phenomenon<sup>99</sup>. Our data provide evidence that current epigenetic clocks represent a composite of at least two variables, change in DNA methylation associated with aging in a cell intrinsic manner (IntrinClock), and a change in cell composition associated with aging. Due to the IntrinClock's resistance to changes in immune cell composition, the CpG sites that constitute the clock may have more readily interpretable cell-autonomous biology as they are less likely to track markers of changing immune cell composition. This prediction could be particularly helpful in the context of identifying a functional or causal relationship between epigenetic clock sites and aging. We found that the sites in the IntrinClock that are hypermethylated with age are enriched within the region 200–1500 bp upstream of gene transcription start sites, and correspondingly strongly depleted in sites distant from genes (25% vs. 15%) (Figure 4A). In sites that are hypomethylated with age, there was a significant enrichment within the first exon of genes (8% vs. 5%) (Figure 4B). DNA methylation changes within 1500 bp of the transcription start site are most closely linked to alterations in gene expression<sup>100</sup>. Similarly, IntrinClock CpGs are enriched for being located near CpG islands (45% vs. 31%) and are depleted from open sea regions (20% vs. 36%) (Figure 4C).

Transcription factor activity and DNA methylation are biologically connected both directly, as in the case of E2F family transcription factors requiring methylated DNA to bind<sup>101</sup>, and indirectly, as in the case of passive methylation from lack of TF binding<sup>102,103</sup>. We investigated regions within 40 bp of IntrinClock CpG sites and used HOMER<sup>104</sup> to identify enriched motifs associated with transcription factor-binding sites (Figure 4D). Motifs associated with TFAP2C, ZNF341, ZFP57, RUNX1, E2F3, HOXA1, SP4, MYB, GRHL2, MGA, IRF3, and INSM1 binding were significantly enriched, compared to a 40-bp background of basepairs surrounding CpG sites that are assayed by both Illumina Infinium HumanMethylation450K and MethylationEPIC chips. Aberrant activity of each corresponding transcription factor has been associated with cancer development or worsened prognosis<sup>79,105–115</sup>. Some of these, such as E2F3<sup>116</sup> and IRF3<sup>117</sup>, have been associated with aging-related diseases, whereas a connection for others has yet to be discovered.

We were interested in exploring general patterns of shifts in IntrinClock CpGs with age. To avoid uneven distribution of tissue samples across age groups, we focused our analysis on blood samples. Given that a linear regression model was used to build the IntrinClock, we were not surprised that the two most prevalent patterns were a linear decrease and increase, respectively, of DNA methylation with age (Figure S7). However, we also found several CpGs (Clusters 4, 5, and 6) where the CpGs reverse their age-related direction of DNA methylation around the age of 21-30. This indicates that, for a subset of CpGs in the IntrinClock, there is a distinction between aging prior to and post sexual maturity. Interestingly, these CpGs were 2.3-fold (34% vs. 14.9%) enriched for being located 200-1500bp

upstream of a TSS, and 2-fold (19.4% vs. 10%) enriched for being located on a genomic south shore region (Figure S8), which are stronger enrichments than identified for IntrinsicClock sites generally (Figures 4A – 4C).

**IntrinsicClock epigenetic age is accelerated in models of intrinsic hallmarks of aging and in HIV<sup>+</sup> individuals.** HIV was one of the earliest conditions to be associated with acceleration of epigenetic age<sup>74</sup>. HIV infection is associated with a plethora of clinical manifestations and morbidities consistent with accelerated aging. However, HIV also causes major changes in immune cell composition<sup>118</sup>, which could skew previous versions of epigenetic clocks. As a result, it is unclear whether early results showcasing epigenetic age acceleration during HIV infection are due to changes in blood cell composition or an accelerated intrinsic rate of aging. Using the IntrinsicClock on previously generated data from HIV<sup>+</sup> individuals and controls, we identified an HIV-associated increase in epigenetic age of two years, supporting the model that HIV leads to accelerated aging independently of shifts in immune cell composition (Figure 5A). We also sought to investigate whether the IntrinsicClock would be accelerated by other acute immune-related diseases. Using a dataset primarily generated in 2020, we found that the IntrinsicClock age prediction was not affected by COVID-19 (Figure 5B), contrary to findings in other epigenetic clocks where COVID-19 infection was associated with an increase in epigenetic age<sup>119</sup>. As the data analyzed in this study were generated early in the COVID-19 pandemic, most individuals would have been acutely, rather than chronically, ill with COVID-19. It remains to be seen whether the IntrinsicClock will predict a higher epigenetic age in those who are infected with COVID-19 for a prolonged period (i.e., long COVID).

One application of epigenetic clocks is in tracking the effect of rejuvenating or aging interventions on cells. As the IntrinsicClock was developed on sites that are not shifting due to immune cell compositional changes, we reasoned it may be more sensitive to such interventions. Consistent with this idea, we used an external dataset<sup>120</sup> to find that the IntrinsicClock is sensitive to Yamanaka factor-mediated reprogramming in fibroblasts. The study authors sorted cells positive for TRA-1-60<sup>+</sup>, a marker for de-differentiation, at six time points after initiation of reprogramming. We investigated IntrinsicClock epigenetic age predictions at each time point and found that, from an initial mean predicted epigenetic age of 31, the age prediction decreased to 20 after 11 days of OSKM-mediated reprogramming. A mean age of 0 was reached after 20 total days of reprogramming (Figure 5C). Conversely, using publicly available data using an in vitro fibroblast model of replicative cellular senescence<sup>121</sup>, we found that the IntrinsicClock was progressively accelerated with cell divisions as cells become progressively more senescent. IntrinsicClock increased from a baseline predicted age of 10 to 15 after 14 population doublings, and then further increased to 20 after another 14 population doublings (Figure 5D). This effect was comparable to that seen using the PhenoAge clock, and stronger relative to the Hannum, Horvath, and Horvath Skin & Blood clocks (Figure S9).

## Discussion

Epigenetic clocks hold great promise for the study of longevity due to their high correlation with age and (particularly for second-generation clocks) association with aging-related disease state. As diagnostic tools, they have the potential to serve as important predictive biomarkers for assessing biological age, determining risk for age-associated diseases, and assessing the efficacy of interventions that target the aging process<sup>122–126</sup>. Recent technical advances, such as the development of principal component clocks<sup>127</sup> and novel techniques for cost reduction<sup>128</sup>, promise to increase reliability and usability further. However, their current status as a composite of multiple aging signals makes them

difficult to interpret and to link to specific biological processes. As an example, a recent study in patients post-COVID 19 infection demonstrated a significant PhenoAge epigenetic age acceleration in individuals over the age of 50, but an epigenetic age reversal for those under the age of 50<sup>119</sup>. Further, the manner in which clocks track healthspan is not fully overlapping, as clocks can be independently predictive of mortality even when analyzed jointly<sup>129</sup>. This challenge in interpretation is equally important for cellular models of the hallmarks of aging. In models of senescence or reprogramming, the sensitivity or even direction of the perturbation on predicted epigenetic age can dramatically differ, depending on the epigenetic clock used. For example, in this study, we identified the Hannum clock as predicting an age *reversal* in a fibroblast model of cellular replicative senescence (Figure S9).

The immune system changes dramatically with aging, and its decline can exacerbate or lead to many aging-related pathologies<sup>130</sup>. Clocks built solely on inflammatory markers can be used to predict age and risk of multimorbidity<sup>131</sup>. However, the presence of CpG sites that track primarily with immune cell markers makes epigenetic clocks applied to cell-intrinsic effects (e.g., cellular reprogramming in fibroblast cell culture) difficult to understand. Such sites can introduce background noise to the resulting measurement.

Here, using sorted CD8<sup>+</sup> T-cell subsets, we observed that naive T cells consistently showed a younger epigenetic age than other CD8<sup>+</sup> subsets (Figure 1C), ranging from a 10-year average age underprediction in some clocks to as high as a 60-year underprediction in others. This suggests that epigenetic clock measurements are significantly affected by CD8<sup>+</sup> T-cell differentiation. These observations reinforce the finding that current epigenetic clocks represent the integration of at least two variables: cell intrinsic aging and changes in immune composition during aging.

To isolate these variables, we developed a novel epigenetic clock that is based on CpG sites that do not change with CD8<sup>+</sup> T-cell differentiation (IntrinClock). We further observed that this clock predicts the same age in each individual across a wide variety of immune cell types. Interestingly, a filtering step based on naive CD8<sup>+</sup> T cells can generate a clock that is not affected by differentiation in cells from different lineages, such as CD4<sup>+</sup> cells or even B cells. This indicates part of a unique “CD8<sup>+</sup> naive” signal may, in fact, be a conserved quiescence program shared by a variety of immune cells. This observation is supported by our finding that methylation patterns associated with naive CD8<sup>+</sup> T cells have a negative correlation with those changing with aging (Figure 2C). A connection between quiescence and aging is found in a wide variety of cell types, including neural stem cells<sup>132</sup>.

The IntrinClock’s higher proportion of sites near transcription start sites and CpG islands and its expected relationship with reprogramming and senescence suggest that it is tracking an intrinsic cellular aging program. Enrichment of IntrinClock CpG sites within motifs bound by transcription factors linked to cancer progression is consistent with a recent review investigating the connection between epigenetic clocks, global hypomethylation, cancer, and aging<sup>133</sup>. It will be important in the future to test whether acceleration of the IntrinClock is linked to particular disease states. This application could be a novel tool used to distinguish age-related diseases caused by aberrant cell-to-cell interactions from those caused by intrinsic cellular dysfunction.

The approach described here reduces the potential of cellular composition changes to be a confounder, particularly in blood or saliva samples, and will likely increase our understanding of biological aging and age-associated diseases. The IntrinClock holds the promise of being more sensitive to cell-intrinsic rejuvenation approaches, as its constituent CpG sites are not affected by immune cell composition. It may also be more closely linked to CpG sites with a functional or even causal relationship with the aging process. Overall, IntrinClock represents a new instrument to add to the aging biomarker toolkit, with a potential wide variety of applications and uses.

## Methods

**Ethics approval.** NIH provided approval for use of phs000424/GRU (GTEx) age data via the dbGaP database approval system. Ethics approval was not required for other datasets generated.

**Immune cell isolation, sorting, and DNA extraction.** PBMCs were extracted from leukopheresis chambers from CMV<sup>+</sup> donors. Blood was first diluted 1:1 with PBS with 2% FBS. Diluted blood was slowly layered on top of 12 mL of Ficoll in a 50-mL Falcon conical tube. The tube was then centrifuged for 30 minutes at 2000 rpm at 21°C without applying a break. The layer containing white blood cells was removed, diluted with FBS-supplemented PBS, and centrifuged for 3 minutes at 2500 rpm. The cell pellet was re-suspended in 15 mL of ACK lysis buffer and incubated for 3 minutes. The cells were topped up with PBS with 2% FBS, centrifuged, and resuspended.

For the initial CD8<sup>+</sup> epigenetic clock characterization experiment, an EasySep<sup>TM</sup> Human T Cell Isolation kit was used to extract T cells from the PBMC fraction. T cells were then washed, stained with 1:500 LIVE/DEAD<sup>TM</sup> Fixable Near-IR Dead Cell staining kit, washed, stained with an antibody cocktail (Supplementary Table 2), and washed again. FACS was performed on a BD FACS Aria<sup>TM</sup> II instrument. DNA was isolated using a Zymo Quick-DNA<sup>TM</sup> Microprep Plus kit.

For the second comprehensive immune cell-sorting experiment, 2 million PBMCs were frozen immediately after extraction. The remaining cells were then positively selected for a CD4 fraction using the EasySep<sup>TM</sup> Human CD4 Positive Selection Kit II. The CD4 cells were stained with 1:500 LIVE/DEAD<sup>TM</sup> Fixable Near-IR Dead Cell staining kit, washed, and stained with CD4/CD8 antibody cocktail (Supplementary Table 2), and the remaining cells were positively selected for a CD8 fraction using the EasySep<sup>TM</sup> Human CD8 Positive Selection Kit II. Both CD8<sup>+</sup> cells and remaining PBMCs were washed, stained with 1:500 LIVE/DEAD<sup>TM</sup> Fixable Near-IR Dead Cell staining kit and washed again. CD8<sup>+</sup> cells were stained with a CD4/CD8 antibody cocktail (Supplementary Table 2), and the remaining PBMCs were stained with a B Cell/NK Cell/Monocyte antibody cocktail (Supplementary Table 3), after blocking with human IgG. All three fractions were then subjected to FACS analysis using a BD FACS Aria<sup>TM</sup> II instrument. DNA was isolated using a Zymo Quick-DNA/RNA<sup>TM</sup> Microprep Plus kit.

For both experiments, DNA was quantified using Qubit<sup>TM</sup> HS dsDNA quantification reagents. Bisulfite conversion and DNA methylation assessment were performed by Diagenode. For all experiments involving FACS, post-sort validations were performed to verify cell sort purity by analyzing sorted populations via flow cytometry. The Clock Foundation assisted with facilitating DNA methylation assessment and data transfer for the initial CD8<sup>+</sup> experiment.

**High-dimensional flow cytometry.** PBMCs were transferred to a 96-well V-bottom plate. Cells were re-suspended in a 1:500 dilution of LIVE/DEAD<sup>TM</sup> Fixable Blue Dead Cell Stain kit in cold PBS and incubated for 30 minutes in the dark. Cells were then washed and blocked with human IgG for 30 minutes. They were then washed twice and stained with a PBMC phenotyping antibody cocktail (Supplementary Table 4). Cell phenotyping was performing on a Cytek Aurora<sup>TM</sup> instrument and analyzed using FlowJo<sup>TM</sup>.

**DNA methylation analysis and pre-processing.** .idat files were converted into beta values by using the *minfi* R package<sup>134</sup>, with a functional normalization pre-processing step<sup>135</sup>. For differential methylation analyses, beta values were converted to M-values through the formula  $M = \log_2(B / (1-B))$ . The R package *umap* was used for UMAP dimensionality reduction<sup>136</sup>.

**Dataset collection and pre-processing.** All datasets used to build the novel epigenetic clock were either generated in this study or downloaded from GEO. Exact ages were obtained for GTEx data through dbGaP<sup>137</sup>, as exact chronological ages of tissues were required. For constructing the clock, the assembled database of DNA methylation data was first culled of any samples that had more than 10% of CpGs missing and of any CpGs that had more than 10% samples missing. All samples derived from cancer tissues were removed. To ensure forward compatibility, we filtered out CpGs that were not on the Infinium MethylationEPIC v2.0 array. Based on our CD8<sup>+</sup> DNA methylation data, we tested the correlation of each CpG methylation with age and with naive CD8<sup>+</sup> T cells. To assess whether CpGs were correlated with naive CD8<sup>+</sup> cells, we binarized each naive sample as “1” and each non-naive (CM, EM, TEMRA) sample as “0” and then used the R *cor* function to compute a Spearman’s correlation between methylation and naive T-cell state. All CpGs with an absolute value correlation of .3 or greater with naive T-cell state were removed, and all CpGs with an absolute value correlation of .3 or less with age were removed.

Once CpGs and samples were filtered, the samples were split 75% for the training set and 25% for the test set. Imputation of missing was performed separately between training sets and test sets, and separately between different tissues within training sets and test sets (imputation performed using the *impute* R package<sup>138</sup>). Outliers were detected and removed using the *outlyx* function in the R *watermelon* package<sup>139</sup>. Untransformed beta values were used for model creation and age prediction. Prior to training the model, ages were transformed using Horvath’s formula used in his original epigenetic clock<sup>2</sup>. An elastic net model using *glmnet*<sup>140</sup> was used to develop the IntrinClock, with alpha value set at .5. Once the first model was generated, the training data were a subset of only those CpGs with non-zero coefficients, which were used for training the final model.

**Statistical methods.** For comparisons between two measurements from one individual, as in Figures 1B and 1C, paired t-tests were used for assessment of significant changes. For multiple comparisons between a group and a background reference, as in Figures 4A, 4B, and 4C, one-sample proportional tests using the *prop.test* function from the R *stats* package were utilized with Bonferroni multiple-comparisons correction. For samples of multiple measurements, repeated measures ANOVA implemented via the *statix* packages<sup>141</sup> was used to test significance. Most graphs and figures were created with aid of the *ggplot* R package<sup>142</sup>.

**Motif enrichment and pattern analyses.** For motif enrichment analysis, the HOMER software tool was utilized<sup>104</sup>. To define sequences of interest, we investigated 40-bp windows surrounding the 381 CpG sites that compose the IntrinClock. As a background, we investigated 40-bp windows around CpG sites in our dataset immediately before removal of CpG sites associated with naive CD8<sup>+</sup> cells and those not associated with aging. For investigating patterns of IntrinClock CpG shifts with age, beta values from blood samples were converted to M values, after which the *degPatterns* function from the *DEGreport* R package<sup>143</sup> was utilized. Patterns with fewer than 10 CpG sites were discarded from analysis. Ages were binned into groups of 10 (0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81<sup>+</sup>). Each bin was confirmed to have at least 100 samples.

**Epigenetic age acceleration analysis.** To compute DNA methylation age for each epigenetic clock, the R *methylclock* package was utilized<sup>144</sup>. For experiments containing a limited number of donors or cell types, epigenetic age acceleration was defined as the difference between epigenetic age prediction and chronological age. For larger studies, epigenetic age acceleration was defined as the residual after regressing predicted epigenetic age on chronological age.

## **Data availability**

DNA methylation profiles generated in this study will be submitted to the GEO public repository prior to publication. Code used to generate the results in this study will be made public on Github prior to publication.

## **Acknowledgments**

This research was primarily funded by the Michael Antonov Foundation and Buck institutional support. Cloud computing support was provided by the Amazon Web Services Activate program, facilitated by the On Deck Longevity Biotech fellowship. Furthermore, we extend our warm appreciation to Dr. Steve Horvath and the Clock Foundation for assisting us with sample handling and data interpretation during the early phases of this work.

## **Author contributions**

N.A., H.M., and H.K. assisted in flow panel design and optimization. A.F. planned and conducted some of the experiments. R.R. and E.V. assisted in manuscript drafting. All authors contributed to manuscript editing and revising. A.T. designed, performed, and analyzed the experiments. A.T. drafted and wrote the manuscript.

## **Competing interests**

A.T. and E.V. are listed co-inventors on pending patents relating to work disclosed in this manuscript.

## References

1. Hannum, G. *et al.* Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Mol. Cell* **49**, 359–367 (2013).
2. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
3. Perna, L. *et al.* Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin. Epigenetics* **8**, 64 (2016).
4. Kabacik, S. *et al.* The relationship between epigenetic age and the hallmarks of aging in human cells. *Nat. Aging* **2**, 484–493 (2022).
5. Consortium, M. M. *et al.* Universal DNA methylation age across mammalian tissues. 2021.01.18.426733 Preprint at <https://doi.org/10.1101/2021.01.18.426733> (2021).
6. Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
7. Lu, A. T. *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303–327 (2019).
8. Protsenko, E. *et al.* “GrimAge,” an epigenetic predictor of mortality, is accelerated in major depressive disorder. *Transl. Psychiatry* **11**, 1–9 (2021).
9. McCrory, C. *et al.* GrimAge Outperforms Other Epigenetic Clocks in the Prediction of Age-Related Clinical Phenotypes and All-Cause Mortality. *J. Gerontol. A. Biol. Sci. Med. Sci.* **76**, 741–749 (2021).
10. Belsky, D. W. *et al.* DunedinPACE, a DNA methylation biomarker of the pace of aging. *eLife* **11**, e73420 (2022).
11. Levine, M. E., Higgins-Chen, A., Thrush, K., Minteer, C. & Niimi, P. Clock Work: Deconstructing the Epigenetic Clock Signals in Aging, Disease, and Reprogramming. 2022.02.13.480245 Preprint at <https://doi.org/10.1101/2022.02.13.480245> (2022).
12. Ying, K. *et al.* Causal Epigenetic Age Uncouples Damage and Adaptation. 2022.10.07.511382 Preprint at <https://doi.org/10.1101/2022.10.07.511382> (2022).
13. Schumacher, B. & Meyer, D. Accurate aging clocks based on accumulating stochastic variation. Preprint at <https://doi.org/10.21203/rs.3.rs-2351315/v1> (2023).
14. Goronzy, J. J., Fang, F., Cavanagh, M. M., Qi, Q. & Weyand, C. M. Naïve T cell maintenance and function in human aging. *J. Immunol. Baltim. Md 1950* **194**, 4073–4080 (2015).
15. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat. Rev. Genet.* **19**, 371–384 (2018).
16. Jonkman, T. H. *et al.* Functional genomics analysis identifies T and NK cell activation as a driver of epigenetic clock progression. *Genome Biol.* **23**, 24 (2022).
17. Chen, B. H. *et al.* DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging* **8**, 1844–1859 (2016).
18. Bonder, M. J. *et al.* Single cell DNA methylation ageing in mouse blood. 2023.01.30.526343 Preprint at <https://doi.org/10.1101/2023.01.30.526343> (2023).
19. Trapp, A., Kerepesi, C. & Gladyshev, V. N. Profiling epigenetic age in single cells. *Nat. Aging* **1**, 1189–1201 (2021).
20. Horvath, S. *et al.* Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. *Aging* **10**, 1758–1775 (2018).
21. Lazuardi, L. *et al.* Age-related loss of naïve T cells and dysregulation of T-cell/B-cell interactions in human lymph nodes. *Immunology* **114**, 37–43 (2005).
22. Policicchio, S. *et al.* Genome-wide DNA methylation meta-analysis in the brains of suicide

completers. *Transl. Psychiatry* **10**, 1–13 (2020).

- 23. Pihlstrøm, L. *et al.* Epigenome-wide association study of human frontal cortex identifies differential methylation in Lewy body pathology. *Nat. Commun.* **13**, 4932 (2022).
- 24. Thompson, E. E. *et al.* Cytokine-induced molecular responses in airway smooth muscle cells inform genome-wide association studies of asthma. *Genome Med.* **12**, 64 (2020).
- 25. Tsaprouni, L. G. *et al.* Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics* **9**, 1382–1396 (2014).
- 26. Oelsner, K. T., Guo, Y., To, S. B.-C., Non, A. L. & Barkin, S. L. Maternal BMI as a predictor of methylation of obesity-related genes in saliva samples from preschool-age Hispanic children at-risk for obesity. *BMC Genomics* **18**, 57 (2017).
- 27. Viana, J. *et al.* Schizophrenia-associated methylomic variation: molecular signatures of disease and polygenic risk burden across multiple brain regions. *Hum. Mol. Genet.* **26**, 210–225 (2017).
- 28. Voisin, S. *et al.* An epigenetic clock for human skeletal muscle. *J. Cachexia Sarcopenia Muscle* **11**, 887–898 (2020).
- 29. Wockner, L. F. *et al.* Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. *Transl. Psychiatry* **4**, e339–e339 (2014).
- 30. Xu, H. *et al.* Sex-biased methylome and transcriptome in human prefrontal cortex. *Hum. Mol. Genet.* **23**, 1260–1270 (2014).
- 31. Xu, Z., Sandler, D. P. & Taylor, J. A. Blood DNA Methylation and Breast Cancer: A Prospective Case-Cohort Analysis in the Sister Study. *JNCI J. Natl. Cancer Inst.* **112**, 87–94 (2020).
- 32. Zannas, A. S. *et al.* Epigenetic upregulation of FKBP5 by aging and stress contributes to NF-κB–driven inflammation and cardiovascular risk. *Proc. Natl. Acad. Sci.* **116**, 11370–11379 (2019).
- 33. Roy, R. *et al.* DNA methylation signatures reveal that distinct combinations of transcription factors specify human immune cell epigenetic identity. *Immunity* **54**, 2465–2480.e5 (2021).
- 34. Somineni, H. K. *et al.* Blood-Derived DNA Methylation Signatures of Crohn’s Disease and Severity of Intestinal Inflammation. *Gastroenterology* **156**, 2254–2265.e3 (2019).
- 35. Takeuchi, C. *et al.* Autoimmune gastritis induces aberrant DNA methylation reflecting its carcinogenic potential. *J. Gastroenterol.* **57**, 144–155 (2022).
- 36. Ringh, M. V. *et al.* Methylome and transcriptome signature of bronchoalveolar cells from multiple sclerosis patients in relation to smoking. *Mult. Scler. J.* **27**, 1014–1026 (2021).
- 37. Roos, L. *et al.* Higher Nevus Count Exhibits a Distinct DNA Methylation Signature in Healthy Human Skin: Implications for Melanoma. *J. Invest. Dermatol.* **137**, 910–920 (2017).
- 38. Magnaye, K. M. *et al.* DNA methylation signatures in airway cells from adult children of asthmatic mothers reflect subtypes of severe asthma. *Proc. Natl. Acad. Sci.* **119**, e2116467119 (2022).
- 39. Martino, D. *et al.* Epigenetic dysregulation of naive CD4+ T-cell activation genes in childhood food allergy. *Nat. Commun.* **9**, 3308 (2018).
- 40. Muse, M. E. *et al.* Genome-Scale DNA Methylation Analysis Identifies Repeat Element Alterations that Modulate the Genomic Stability of Melanocytic Nevi. *J. Invest. Dermatol.* **142**, 1893–1902.e7 (2022).
- 41. Nicodemus-Johnson, J. *et al.* DNA methylation in lung cells is associated with asthma endotypes and genetic risk. *JCI Insight* **1**, (2016).
- 42. Nonino, C., NY, Noronha, CF, Nicoletti, & MA, Pinhel. Trait related and differential DNA Methylation in obese and normal weight Brazilian women. (2021).
- 43. Vyas, C. M. *et al.* Pilot study of genome-wide differences in DNA methylation among older adults with normal cognition and mild cognitive impairment, with and without neuropsychiatric symptoms. *Alzheimers Dement.* **17**, e055497 (2021).

44. Oliva, M. *et al.* DNA methylation QTL mapping across diverse human tissues provides molecular links between genetic variation and complex traits. *Nat. Genet.* **55**, 112–122 (2023).
45. Pai, S. *et al.* Differential methylation of enhancer at IGF2 is associated with abnormal dopamine synthesis in major psychosis. *Nat. Commun.* **10**, 2046 (2019).
46. Renauer, P. A., Coit, P. & Sawalha, A. H. The DNA methylation signature of human TCR $\alpha\beta$ +CD4–CD8– double negative T cells reveals CG demethylation and a unique epigenetic architecture permissive to a broad stimulatory immune response. *Clin. Immunol.* **156**, 19–27 (2015).
47. Ringh, M. V. *et al.* Tobacco smoking induces changes in true DNA methylation, hydroxymethylation and gene expression in bronchoalveolar lavage cells. *eBioMedicine* **46**, 290–304 (2019).
48. Ishak, M. *et al.* Genome-Wide Open Chromatin Methylome Profiles in Colorectal Cancer. *Biomolecules* **10**, 719 (2020).
49. Islam, S. A. *et al.* Integration of DNA methylation patterns and genetic variation in human pediatric tissues help inform EWAS design and interpretation. *Epigenetics Chromatin* **12**, 1 (2019).
50. Jiang, W. *et al.* Genome-Wide Identification of a Methylation Gene Panel as a Prognostic Biomarker in Nasopharyngeal Carcinoma. *Mol. Cancer Ther.* **14**, 2864–2873 (2015).
51. Johnson, K. C., Houseman, E. A., King, J. E. & Christensen, B. C. Normal breast tissue DNA methylation differences at regulatory elements are associated with the cancer risk factor age. *Breast Cancer Res.* **19**, 81 (2017).
52. Kananen, L. *et al.* Aging-associated DNA methylation changes in middle-aged individuals: the Young Finns study. *BMC Genomics* **17**, 103 (2016).
53. Kandaswamy, R. *et al.* DNA methylation signatures of adolescent victimization: analysis of a longitudinal monozygotic twin sample. *Epigenetics* **16**, 1169–1186 (2021).
54. Kasuga, Y. *et al.* DNA methylation analysis of cord blood samples in neonates born to gestational diabetes mothers diagnosed before 24 gestational weeks. *BMJ Open Diabetes Res. Care* **10**, e002539 (2022).
55. Kho, M. *et al.* Epigenetic loci for blood pressure are associated with hypertensive target organ damage in older African Americans from the genetic epidemiology network of Arteriopathy (GENOA) study. *BMC Med. Genomics* **13**, 131 (2020).
56. Konigsberg, I. R. *et al.* Host methylation predicts SARS-CoV-2 infection and clinical outcome. *Commun. Med.* **1**, 1–10 (2021).
57. Langevin, S. M. *et al.* CpG island methylation profile in non-invasive oral rinse samples is predictive of oral and pharyngeal carcinoma. *Clin. Epigenetics* **7**, 125 (2015).
58. Lewis, S. K. *et al.* DNA Methylation Analysis Validates Organoids as a Viable Model for Studying Human Intestinal Aging. *Cell. Mol. Gastroenterol. Hepatol.* **9**, 527–541 (2020).
59. Kular, L. *et al.* DNA methylation as a mediator of HLA-DRB1\*15:01 and a protective variant in multiple sclerosis. *Nat. Commun.* **9**, 2397 (2018).
60. Lunnon, K. *et al.* Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat. Neurosci.* **17**, 1164–1170 (2014).
61. Gasparoni, G. *et al.* DNA methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex. *Epigenetics Chromatin* **11**, 41 (2018).
62. Gopalan, S. *et al.* Trends in DNA Methylation with Age Replicate Across Diverse Human Populations. *Genetics* **206**, 1659–1674 (2017).
63. Witte, L. D. de *et al.* Contribution of Age, Brain Region, Mood Disorder Pathology, and Interindividual Factors on the Methylome of Human Microglia. *Biol. Psychiatry* **91**, 572–581 (2022).

64. Hannon, E. *et al.* DNA methylation meta-analysis reveals cellular alterations in psychosis and markers of treatment-resistant schizophrenia. *eLife* **10**, e58430 (2021).
65. Hearn, N. L., Chiu, C. L. & Lind, J. M. Comparison of DNA methylation profiles from saliva in Coeliac disease and non-coeliac disease individuals. *BMC Med. Genomics* **13**, 16 (2020).
66. Hong, S. R. *et al.* DNA methylation-based age prediction from saliva: High age predictability by combination of 7 CpG markers. *Forensic Sci. Int. Genet.* **29**, 118–125 (2017).
67. Horvath, S. *et al.* Obesity accelerates epigenetic aging of human liver. *Proc. Natl. Acad. Sci.* **111**, 15538–15543 (2014).
68. Horvath, S. *et al.* An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol.* **17**, 171 (2016).
69. Xiao, C., Yi, S. & Huang, D. Genome-wide identification of age-related CpG sites for age estimation from blood DNA of Han Chinese individuals. *Electrophoresis* **42**, 1488–1496 (2021).
70. McEwen, L. M. *et al.* Systematic evaluation of DNA methylation age estimation with common preprocessing methods and the Infinium MethylationEPIC BeadChip array. *Clin. Epigenetics* **10**, 123 (2018).
71. Huynh, J. L. *et al.* Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. *Nat. Neurosci.* **17**, 121–130 (2014).
72. Clement, J. *et al.* Umbilical cord plasma concentrate has beneficial effects on DNA methylation GrimAge and human clinical biomarkers. *Aging Cell* **21**, e13696 (2022).
73. Garcia-Prieto, C. A. *et al.* Epigenetic Profiling and Response to CD19 Chimeric Antigen Receptor T-Cell Therapy in B-Cell Malignancies. *JNCI J. Natl. Cancer Inst.* **114**, 436–445 (2022).
74. Horvath, S. & Levine, A. J. HIV-1 Infection Accelerates Age According to the Epigenetic Clock. *J. Infect. Dis.* **212**, 1563–1573 (2015).
75. Fries, G. R. *et al.* Accelerated hippocampal biological aging in bipolar disorder. *Bipolar Disord.* **22**, 498–507 (2020).
76. Estupiñán-Moreno, E. *et al.* Methylome and transcriptome profiling of giant cell arteritis monocytes reveals novel pathways involved in disease pathogenesis and molecular response to glucocorticoids. *Ann. Rheum. Dis.* **81**, 1290–1300 (2022).
77. Dávalos, V. *et al.* Epigenetic profiling linked to multisystem inflammatory syndrome in children (MIS-C): A multicenter, retrospective study. *eClinicalMedicine* **50**, (2022).
78. Quintivano, J., Aryee, M. J. & Kaminsky, Z. A. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics* **8**, 290–302 (2013).
79. Liu, Y. *et al.* Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat. Biotechnol.* **31**, 142–147 (2013).
80. Martino, D. *et al.* Longitudinal, genome-scale analysis of DNA methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance. *Genome Biol.* **14**, R42 (2013).
81. Arpón, A. *et al.* Epigenome-wide association study in peripheral white blood cells involving insulin resistance. *Sci. Rep.* **9**, 2445 (2019).
82. Bacalini, M. G. *et al.* Identification of a DNA methylation signature in blood cells from persons with Down Syndrome. *Aging* **7**, 82–96 (2014).
83. Bartlett, T. E. *et al.* Antiprogestins reduce epigenetic field cancerization in breast tissue of young healthy women. *Genome Med.* **14**, 64 (2022).
84. Bauer, M. A. *et al.* Genome-Wide DNA Methylation Signatures Predict the Early Asymptomatic Doxorubicin-Induced Cardiotoxicity in Breast Cancer. *Cancers* **13**, 6291 (2021).

85. Brennan, K. *et al.* NSD1 mutations deregulate transcription and DNA methylation of bivalent developmental genes in Sotos syndrome. *Hum. Mol. Genet.* **31**, 2164–2184 (2022).
86. Cerapio, J. P. *et al.* Global DNA hypermethylation pattern and unique gene expression signature in liver cancer from patients with Indigenous American ancestry. *Oncotarget* **12**, 475–492 (2021).
87. Cullell, N. *et al.* DNA Methylation and Ischemic Stroke Risk: An Epigenome-Wide Association Study. *Thromb. Haemost.* 1767–1778 (2022) doi:10.1055/s-0042-1749328.
88. Li, M. *et al.* Genomic methylation variations predict the susceptibility of six chemotherapy related adverse effects and cancer development for Chinese colorectal cancer patients. *Toxicol. Appl. Pharmacol.* **427**, 115657 (2021).
89. Cobben, J. M. *et al.* DNA methylation abundantly associates with fetal alcohol spectrum disorder and its subphenotypes. *Epigenomics* **11**, 767–785 (2019).
90. Charlton, J. *et al.* Methylome analysis identifies a Wilms tumor epigenetic biomarker detectable in blood. *Genome Biol.* **15**, 434 (2014).
91. Bennett, T. J., Udupa, V. A. V. & Turner, S. J. Running to Stand Still: Naive CD8+ T Cells Actively Maintain a Program of Quiescence. *Int. J. Mol. Sci.* **21**, 9773 (2020).
92. Jenkins, T. *et al.* The impact of zinc and folic acid supplementation on sperm DNA methylation: results from the folic acid and zinc supplementation randomized clinical trial (FAZST). *Fertil. Steril.* **117**, 75–85 (2022).
93. Schlums, H. *et al.* Cytomegalovirus Infection Drives Adaptive Epigenetic Diversification of NK Cells with Altered Signaling and Effector Function. *Immunity* **42**, 443–456 (2015).
94. Rodriguez, R. M. *et al.* Epigenetic Networks Regulate the Transcriptional Program in Memory and Terminally Differentiated CD8+ T Cells. *J. Immunol.* **198**, 937–949 (2017).
95. Li, M. *et al.* Age related human T cell subset evolution and senescence. *Immun. Ageing* **16**, 24 (2019).
96. Pitaksalee, R. *et al.* Differential CpG DNA methylation in peripheral naïve CD4+ T-cells in early rheumatoid arthritis patients. *Clin. Epigenetics* **12**, 54 (2020).
97. Garaud, S. *et al.* FOXP1 is a regulator of quiescence in healthy human CD4+ T cells and is constitutively repressed in T cells from patients with lymphoproliferative disorders. *Eur. J. Immunol.* **47**, 168–179 (2017).
98. Chong, Y. *et al.* CD27+ (memory) B cell decrease and apoptosis-resistant CD27– (naive) B cell increase in aged humans: implications for age-related peripheral B cell developmental disturbances. *Int. Immunol.* **17**, 383–390 (2005).
99. Bell, C. G. *et al.* DNA methylation aging clocks: challenges and recommendations. *Genome Biol.* **20**, 249 (2019).
100. Schlosberg, C. E., VanderKraats, N. D. & Edwards, J. R. Modeling complex patterns of differential DNA methylation that associate with gene expression changes. *Nucleic Acids Res.* **45**, 5100–5111 (2017).
101. Campanero, M. R., Armstrong, M. I. & Flemington, E. K. CpG methylation as a mechanism for the regulation of E2F activity. *Proc. Natl. Acad. Sci.* **97**, 6481–6486 (2000).
102. Medvedeva, Y. A. *et al.* Effects of cytosine methylation on transcription factor binding sites. *BMC Genomics* **15**, 119 (2014).
103. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75–82 (2012).
104. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).
105. Wang, X. *et al.* TFAP2C promotes stemness and chemotherapeutic resistance in colorectal cancer

via inactivating hippo signaling pathway. *J. Exp. Clin. Cancer Res. CR* **37**, 27 (2018).

106.Yan, D. *et al.* Developing ZNF Gene Signatures Predicting Radiosensitivity of Patients with Breast Cancer. *J. Oncol.* **2021**, e9255494 (2021).

107.Chen, L. *et al.* ZFP57 suppress proliferation of breast cancer cells through down-regulation of MEST-mediated Wnt/β-catenin signalling pathway. *Cell Death Dis.* **10**, 1–15 (2019).

108.Tuo, Z. *et al.* RUNX1 is a promising prognostic biomarker and related to immune infiltrates of cancer-associated fibroblasts in human cancers. *BMC Cancer* **22**, 523 (2022).

109.Feng, Z. *et al.* E2F3 promotes cancer growth and is overexpressed through copy number variation in human melanoma. *Oncotargets Ther.* **11**, 5303–5313 (2018).

110.Hedrick, E., Cheng, Y., Jin, U.-H., Kim, K. & Safe, S. Specificity protein (Sp) transcription factors Sp1, Sp3 and Sp4 are non-oncogene addiction genes in cancer cells. *Oncotarget* **7**, 22245–22256 (2016).

111.Cicirò, Y. & Sala, A. MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. *Oncogenesis* **10**, 1–15 (2021).

112.Xiang, X. *et al.* Grhl2 Determines the Epithelial Phenotype of Breast Cancers and Promotes Tumor Progression. *PLOS ONE* **7**, e50781 (2012).

113.Mathsyaraja, H. *et al.* Loss of MGA repression mediated by an atypical polycomb complex promotes tumor progression and invasiveness. *eLife* **10**, e64212 (2021).

114.Tian, M. *et al.* IRF3 prevents colorectal tumorigenesis via inhibiting the nuclear translocation of β-catenin. *Nat. Commun.* **11**, 5762 (2020).

115.Rocha, R. & Henrique, R. Insulinoma-Associated Protein 1 (INSM1): Diagnostic, Prognostic, and Therapeutic Use in Small Cell Lung Cancer. *J. Mol. Pathol.* **3**, 140–167 (2022).

116.Ki, S. *et al.* Global Transcriptional Profiling Reveals Distinct Functions of Thymic Stromal Subsets and Age-Related Changes during Thymic Involution. *Cell Rep.* **9**, 402–415 (2014).

117.Zhang, X. *et al.* Interferon Regulatory Factor 3 Deficiency Induces Age-Related Alterations of the Retina in Young and Old Mice. *Front. Cell. Neurosci.* **13**, 272 (2019).

118.Douek, D. C. *et al.* Changes in thymic function with age and during the treatment of HIV infection. *Nature* **396**, 690–695 (1998).

119. Cao, X. *et al.* Accelerated biological aging in COVID-19 patients. *Nat. Commun.* **13**, 2135 (2022).

120.Ohnuki, M. *et al.* Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 12426–12431 (2014).

121.Xie, W. *et al.* DNA Methylation Patterns Separate Senescence from Transformation Potential and Indicate Cancer Risk. *Cancer Cell* **33**, 309–321.e5 (2018).

122.Oblak, L., van der Zaag, J., Higgins-Chen, A. T., Levine, M. E. & Boks, M. P. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing Res. Rev.* **69**, 101348 (2021).

123.Duan, R., Fu, Q., Sun, Y. & Li, Q. Epigenetic clock: A promising biomarker and practical tool in aging. *Ageing Res. Rev.* **81**, 101743 (2022).

124.Fransquet, P. D., Wrigglesworth, J., Woods, R. L., Ernst, M. E. & Ryan, J. The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. *Clin. Epigenetics* **11**, 62 (2019).

125.Simpson, D. J. & Chandra, T. Epigenetic age prediction. *Aging Cell* **20**, e13452 (2021).

126.Noroozi, R. *et al.* DNA methylation-based age clocks: From age prediction to age reversion. *Ageing Res. Rev.* **68**, 101314 (2021).

127. Higgins-Chen, A. *et al.* A Computational Solution to Bolster Epigenetic Clock Reliability for Clinical Trials and Longitudinal Tracking. *Innov. Aging* **5**, 5 (2021).

128. Ultra-cheap and scalable epigenetic age predictions with TIME-Seq | bioRxiv.  
<https://www.biorxiv.org/content/10.1101/2021.10.25.465725v1.full>.

129. Li, X. *et al.* Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. *eLife* **9**, e51507 (2020).

130. Weyand, C. M. & Goronzy, J. J. Aging of the Immune System. Mechanisms and Therapeutic Targets. *Ann. Am. Thorac. Soc.* **13**, S422–S428 (2016).

131. Sayed, N. *et al.* An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging. *Nat. Aging* **1**, 598–615 (2021).

132. Audesse, A. J. & Webb, A. E. Mechanisms of enhanced quiescence in neural stem cell aging. *Mech. Ageing Dev.* **191**, 111323 (2020).

133. Johnstone, S. E., Gladyshev, V. N., Aryee, M. J. & Bernstein, B. E. Epigenetic clocks, aging, and cancer. *Science* **378**, 1276–1277 (2022).

134. Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).

135. Fortin, J.-P. *et al.* Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.* **15**, 503 (2014).

136. Konopka, T. umap: Uniform Manifold Approximation and Projection. (2023).

137. THE GTEx CONSORTIUM. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).

138. Hastie, T., Tibshirani, R., Narasimhan, B. & Chu, G. impute: Imputation for microarray data. (2023) doi:10.18129/B9.bioc.impute.

139. Schalkwyk, L. C. *et al.* waterRmelon: Illumina 450 and EPIC methylation array normalization and metrics. (2023) doi:10.18129/B9.bioc.wateRmelon.

140. Friedman, J. *et al.* glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models. (2022).

141. Kassambara, A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. (2023).

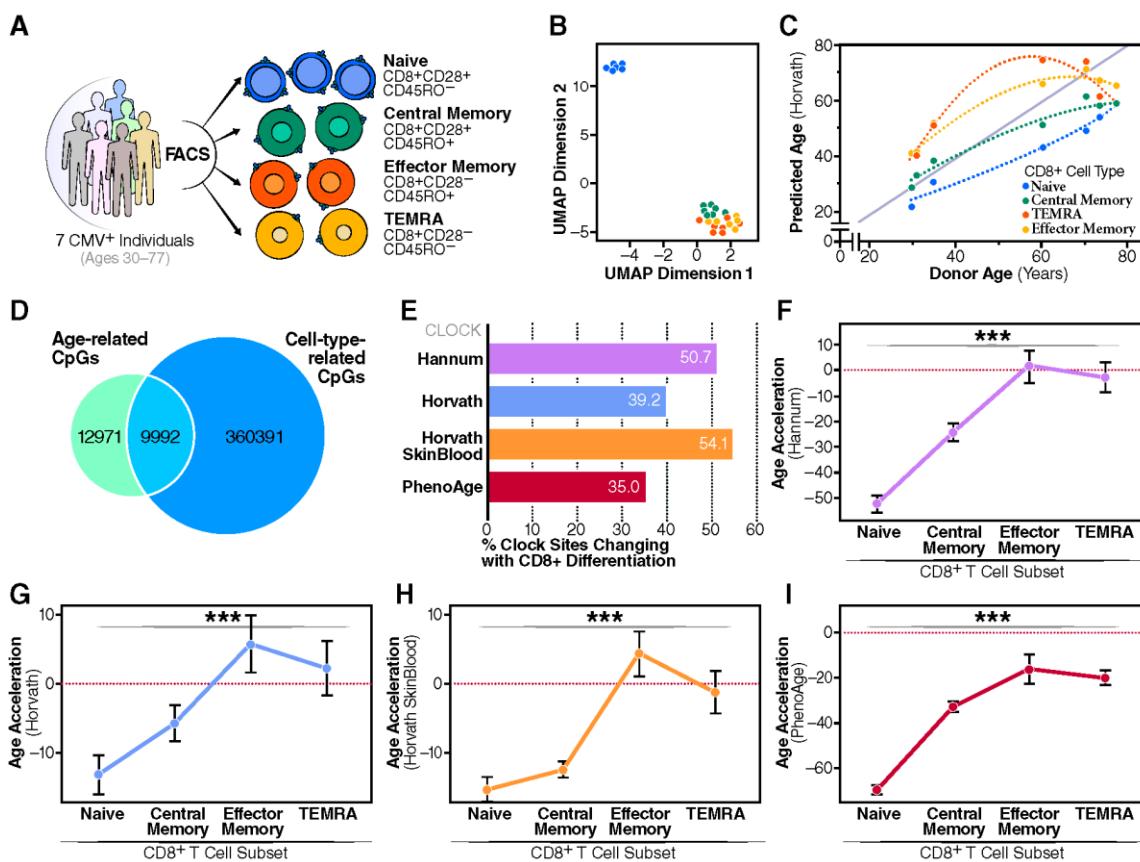
142. Wickham, H. *et al.* ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. (2022).

143. Pantano, L. *et al.* DEGreport: Report of DEG analysis. (2023) doi:10.18129/B9.bioc.DEGreport.

144. Pelegri-Siso, D. & Gonzalez, J. R. Methylclock - DNA methylation-based clocks. (2023) doi:10.18129/B9.bioc.methylclock.

## Figures

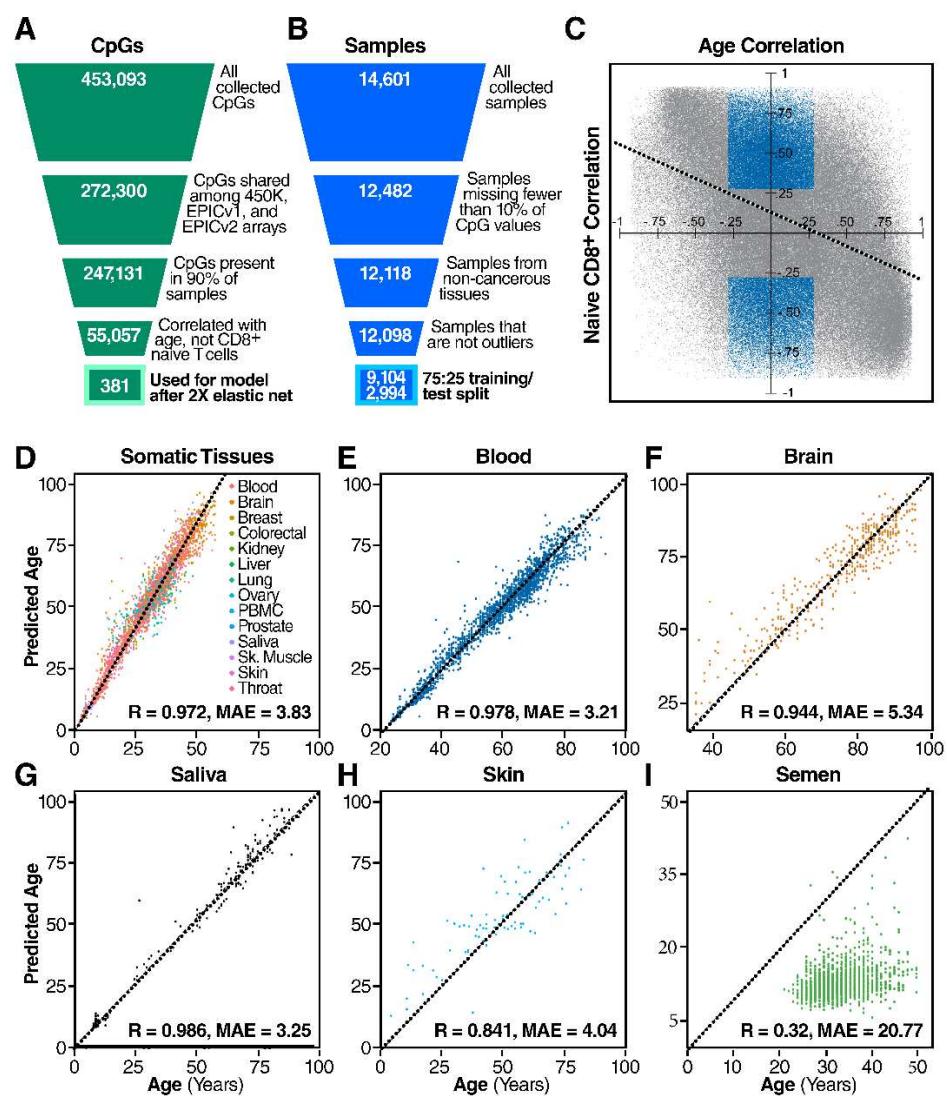
Figure 1



Verdin 23.03A

**Figure 1. CpG site changes during T-cell differentiation.** **a**, Experimental design for determining impact of CD8<sup>+</sup> differentiation on epigenetic clock age prediction. **b**, UMAP dimensionality reduction of CD8<sup>+</sup> DNA methylation profiles. **c**, Differences between predicted epigenetic age as a function of donor age and CD8<sup>+</sup> T-cell subset. **d**, Comparison of shared CpG site changes between age in CD8<sup>+</sup> T cells and CD8<sup>+</sup> cell subset. **e**, Percent of sites in four epigenetic clocks that are altered by CD8<sup>+</sup> T-cell differentiation. **f-i**, Comparison of the (f) Hannum, (g) Horvath, (h) Horvath skin and blood, and (i) PhenoAge epigenetic age acceleration predictions for four CD8<sup>+</sup> T-cell subsets. \*\*\* ANOVA p-value less than .001.

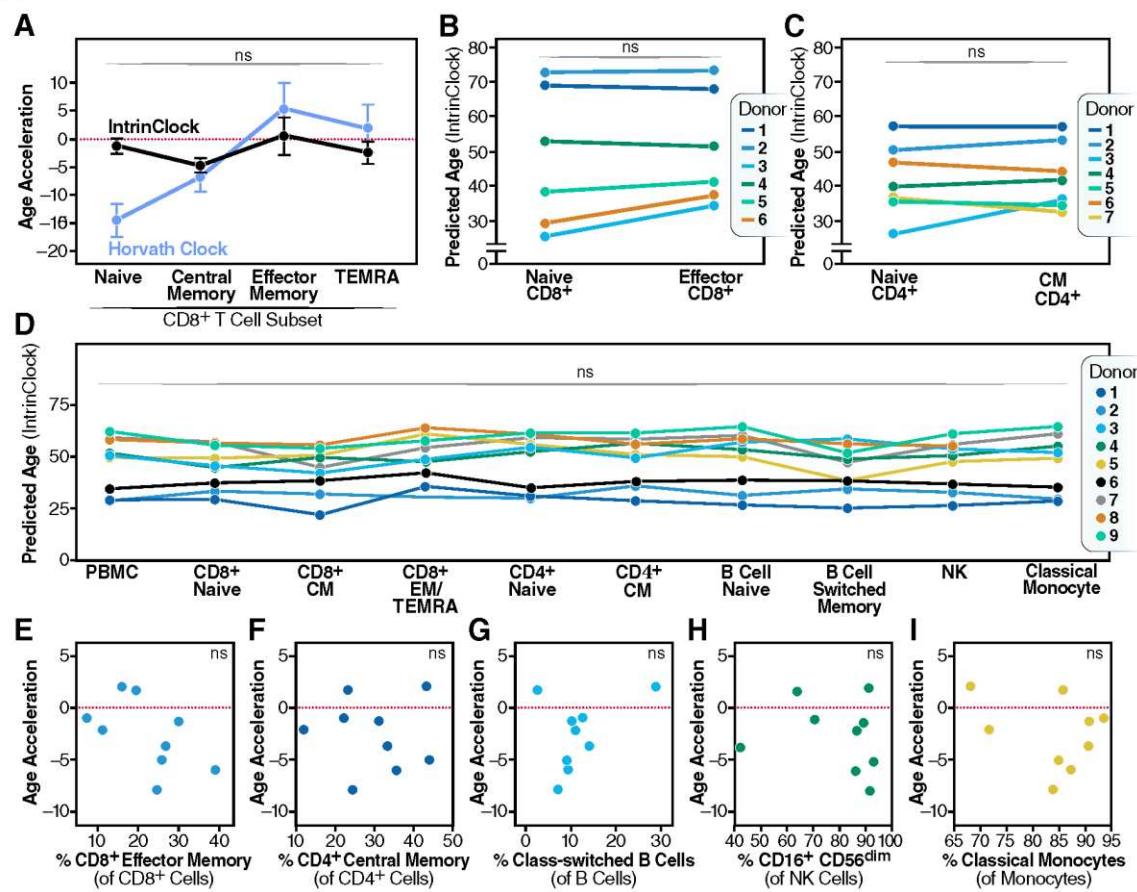
Figure 2



Verdin 23.03B

**Figure 2. Preparation of tissues.** **a**, Filtering strategy for CpG sites. **b**, Filtering strategy for samples. **c**, Visualization of the filtering process for differentiation-independent age-related CpGs. Blue CpGs (those correlated with age but not with being a naïve cell) were included in the feature set, whereas gray CpGs were not. Green dashed line indicates linear least-squared regression line of relationship between CpG age correlation and CpG CD8<sup>+</sup> naïve cell correlation. **d**, Correlation between age and IntrinClock predicted age in a variety of tissues from the test set. **e-h**, Individual correlation plots for specific tissues in the test set. **i**, Epigenetic age vs. chronological age correlation plot for semen samples.

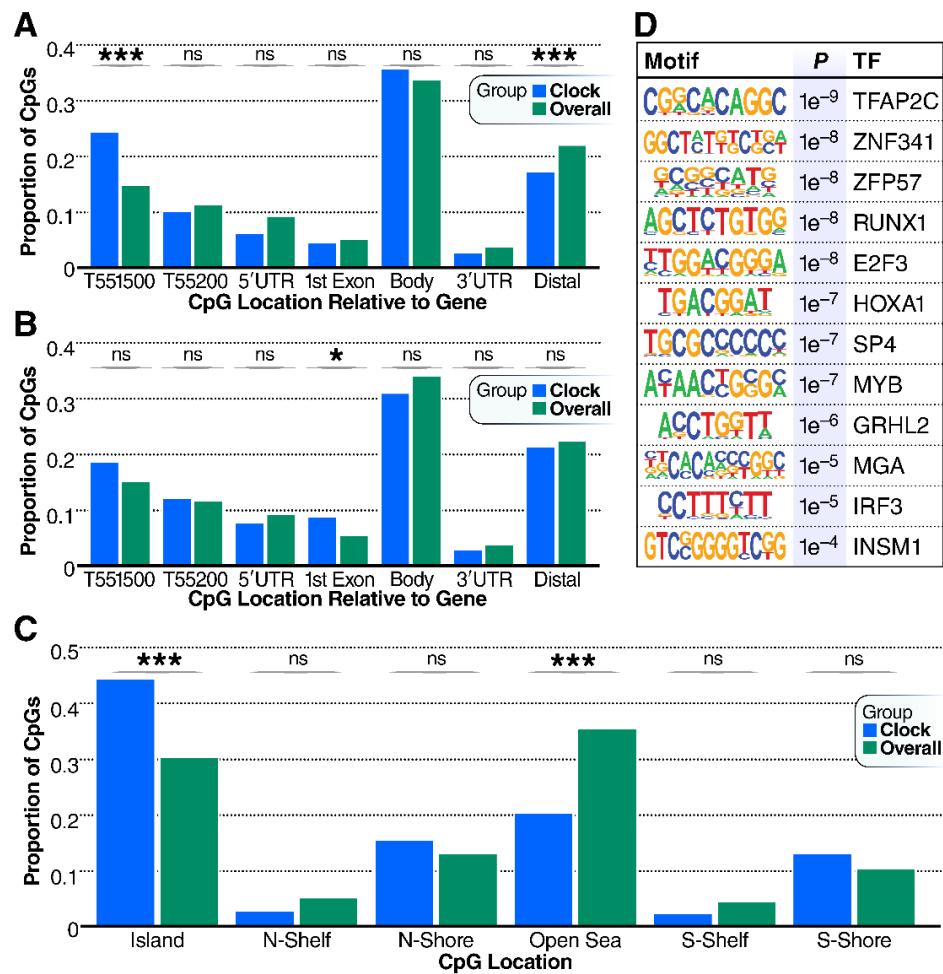
Figure 3



Verdin 23.03C

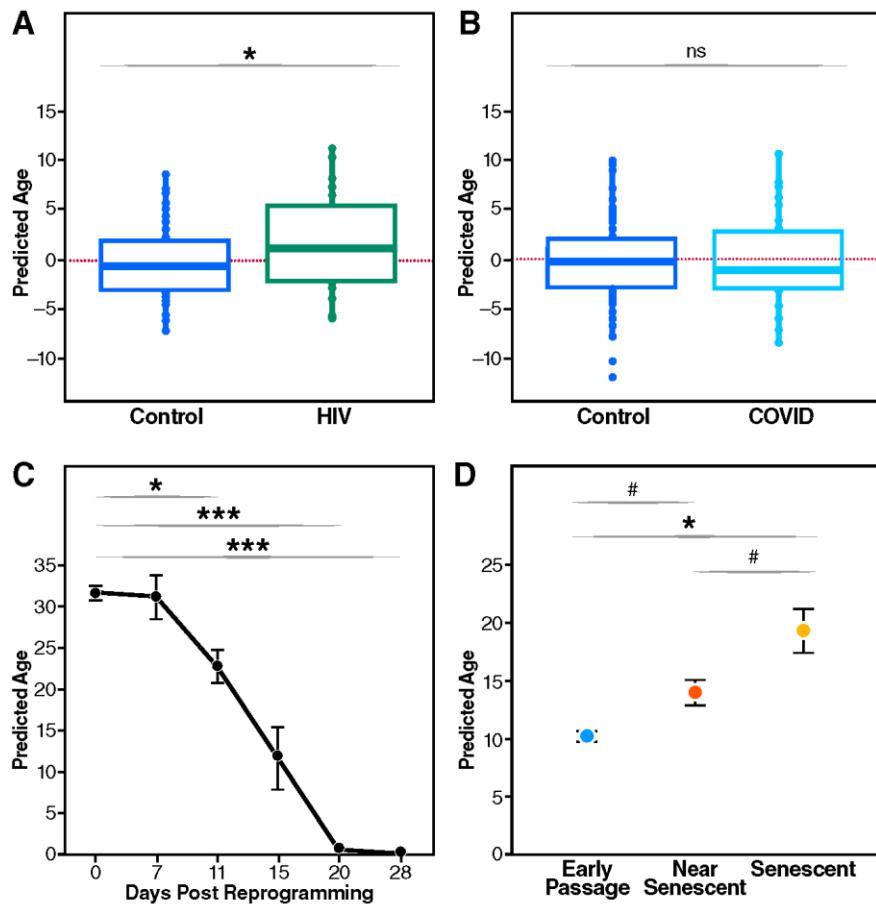
**Figure 3. Epigenetic age accelerations by different clocks. a**, Differences in epigenetic age accelerations in different CD8<sup>+</sup> subsets generated in this study. Horvath clock predictions overlaid in light gray. **b**, Epigenetic ages of CD8<sup>+</sup> naive cells and effector memory cells, based on data from GSE66564 and GSE83156. **c**, Epigenetic ages of CD4<sup>+</sup> naive cells and central memory cells, based on data from GSE121192 and GSE71825. **d**, Epigenetic ages of PBMCs, CD8<sup>+</sup> naive, CD8<sup>+</sup> central memory, CD8<sup>+</sup> combined effector and TEMRA, CD4<sup>+</sup> naive, CD4<sup>+</sup> central memory, B-cell naive, B-cell switched memory, CD16<sup>+</sup>CD56<sup>dim</sup> NK, and classical monocyte cells. **e-i**, Association of percentage of e, effector memory CD8<sup>+</sup> cells, f, central memory CD4<sup>+</sup> cells, g, class-switched B cells, h, CD16<sup>+</sup> CD56<sup>dim</sup> NK cells, and i, classical monocytes with epigenetic age acceleration.

**Figure 4**



**Figure 4. Distributions of CpG positions.** **a**, Distributions of CpG positions relative to genes in IntrinClock sites that are hyper-methylated with age relative to background. **b**, Distributions of CpG positions relative to genes in IntrinClock sites that are hypo-methylated with age relative to background. **c**, Genomic distribution of IntrinClock CpG positions. **d**, HOMER analysis of the top 12 motifs enriched within 19bp on either side (5' or 3') of IntrinClock sites (40 bp total). \*\*\* one-sample proportion t-test p-value < .001; \* < .05

**Figure 5**



Verdin 23.03E

**Figure 5. IntrinsicClock in HIV-infected individuals.** **a**, Increase of IntrinsicClock epigenetic age in HIV<sup>+</sup> individuals, DNA methylation data from GSE67751. **b**, No increase of IntrinsicClock epigenetic age due to COVID, DNA methylation data from GSE167202. **c**, Epigenetic reprogramming affects fibroblast predicted IntrinsicClock age. DNA methylation data from GSE54848. **d**, Induced replicative senescence in fibroblasts leads to an increase in IntrinsicClock predicted age. DNA methylation data from GSE91069. T-test p-values # < .10; \* < .05; \*\*\* < .001.