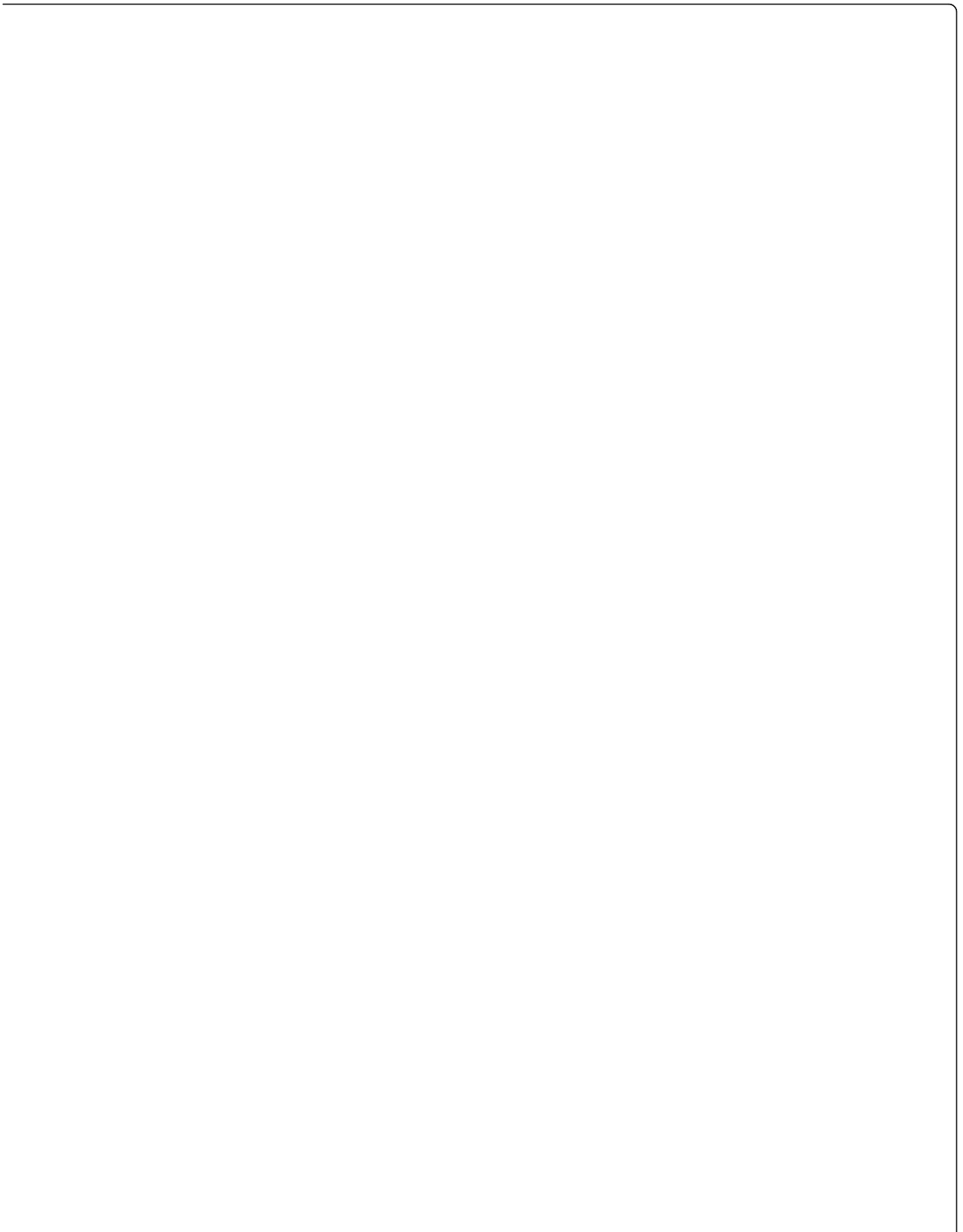


3, 36

ates





\mathbb{A}^1 (resp. \mathbb{A}^n) is a \mathbb{A}^1 -fibration (resp. \mathbb{A}^1 -fibration) over \mathbb{A}^1 (resp. \mathbb{A}^n) with fiber \mathbb{A}^1 (resp. \mathbb{A}^1).

The total space of the \mathbb{A}^1 -fibration $\mathbb{A}^1 \rightarrow \mathbb{A}^1$ is \mathbb{A}^1 . The total space of the \mathbb{A}^1 -fibration $\mathbb{A}^n \rightarrow \mathbb{A}^n$ is \mathbb{A}^n .

The first part of the proof is to show that the set $\{T_1, T_2, \dots, T_n\}$ is linearly independent. Suppose that there are scalars $\alpha_1, \alpha_2, \dots, \alpha_n$ such that $\alpha_1 T_1 + \alpha_2 T_2 + \dots + \alpha_n T_n = 0$. We will show that $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.

Consider the first component of the vector $\alpha_1 T_1 + \alpha_2 T_2 + \dots + \alpha_n T_n$. This component is $\alpha_1 + \alpha_2 + \dots + \alpha_n$. Since this component is zero, we have $\alpha_1 + \alpha_2 + \dots + \alpha_n = 0$.

Now consider the second component of the vector $\alpha_1 T_1 + \alpha_2 T_2 + \dots + \alpha_n T_n$. This component is $\alpha_1 + \alpha_2 + \dots + \alpha_n$. Since this component is zero, we have $\alpha_1 + \alpha_2 + \dots + \alpha_n = 0$.

In general, the i -th component of the vector $\alpha_1 T_1 + \alpha_2 T_2 + \dots + \alpha_n T_n$ is $\alpha_1 + \alpha_2 + \dots + \alpha_n$. Since this component is zero, we have $\alpha_1 + \alpha_2 + \dots + \alpha_n = 0$.

Therefore, we have $\alpha_1 + \alpha_2 + \dots + \alpha_n = 0$ for all $i = 1, 2, \dots, n$. This implies that $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.

Thus, the set $\{T_1, T_2, \dots, T_n\}$ is linearly independent.

Next, we will show that the set $\{T_1, T_2, \dots, T_n\}$ spans the space. Let v be an arbitrary vector in the space. We will show that v can be written as a linear combination of T_1, T_2, \dots, T_n .

Consider the first component of the vector v . This component is v_1 . We can write v_1 as $\alpha_1 + \alpha_2 + \dots + \alpha_n$, where $\alpha_1, \alpha_2, \dots, \alpha_n$ are scalars.

Now consider the second component of the vector v . This component is v_2 . We can write v_2 as $\alpha_1 + \alpha_2 + \dots + \alpha_n$, where $\alpha_1, \alpha_2, \dots, \alpha_n$ are scalars.

In general, the i -th component of the vector v is v_i . We can write v_i as $\alpha_1 + \alpha_2 + \dots + \alpha_n$, where $\alpha_1, \alpha_2, \dots, \alpha_n$ are scalars.

Therefore, we have $v = \alpha_1 T_1 + \alpha_2 T_2 + \dots + \alpha_n T_n$.

Thus, the set $\{T_1, T_2, \dots, T_n\}$ spans the space.

Since the set $\{T_1, T_2, \dots, T_n\}$ is linearly independent and spans the space, it is a basis for the space.

The first part of the proof is to show that the set (\mathbf{p}, \mathbf{r}) is a solution of the system $(\mathbf{p}, \mathbf{r}) \in \mathbf{A}$, where $\mathbf{A} = \{(\mathbf{p}, \mathbf{r}) \in \mathbb{R}^n \times \mathbb{R}^n : \mathbf{p} \geq \mathbf{0}, \mathbf{r} \geq \mathbf{0}, \mathbf{p} + \mathbf{r} = \mathbf{1}\}$.

Let $(\mathbf{p}, \mathbf{r}) \in \mathbf{A}$. Then $\mathbf{p} \geq \mathbf{0}$, $\mathbf{r} \geq \mathbf{0}$, and $\mathbf{p} + \mathbf{r} = \mathbf{1}$. We have $\mathbf{p} \geq \mathbf{0}$ and $\mathbf{r} \geq \mathbf{0}$ by definition. To show $\mathbf{p} + \mathbf{r} = \mathbf{1}$, we note that $\mathbf{p} + \mathbf{r} = \mathbf{1}$ is equivalent to $\mathbf{p} = \mathbf{1} - \mathbf{r}$. Since $\mathbf{r} \geq \mathbf{0}$, we have $\mathbf{p} = \mathbf{1} - \mathbf{r} \leq \mathbf{1}$. Also, since $\mathbf{p} \geq \mathbf{0}$, we have $\mathbf{p} = \mathbf{1} - \mathbf{r} \geq \mathbf{0}$. Therefore, $\mathbf{p} = \mathbf{1} - \mathbf{r}$ and $\mathbf{p} + \mathbf{r} = \mathbf{1}$.

The second part of the proof is to show that the set (\mathbf{p}, \mathbf{r}) is a solution of the system $(\mathbf{p}, \mathbf{r}) \in \mathbf{B}$, where $\mathbf{B} = \{(\mathbf{p}, \mathbf{r}) \in \mathbb{R}^n \times \mathbb{R}^n : \mathbf{p} \geq \mathbf{0}, \mathbf{r} \geq \mathbf{0}, \mathbf{p} + \mathbf{r} = \mathbf{1}, \mathbf{p} \mathbf{A} \mathbf{p} = \mathbf{0}, \mathbf{r} \mathbf{A} \mathbf{r} = \mathbf{0}\}$.

Let $(\mathbf{p}, \mathbf{r}) \in \mathbf{A}$. Then $\mathbf{p} \geq \mathbf{0}$, $\mathbf{r} \geq \mathbf{0}$, and $\mathbf{p} + \mathbf{r} = \mathbf{1}$. We have $\mathbf{p} \mathbf{A} \mathbf{p} = \mathbf{0}$ and $\mathbf{r} \mathbf{A} \mathbf{r} = \mathbf{0}$ by definition. Therefore, $(\mathbf{p}, \mathbf{r}) \in \mathbf{B}$.

Discussion

epigenetic clock theory of aging,

epigenetic clock theory of aging,

..... "....."
.....

Conclusions

Epigenetic maintenance systems play a critical role in the development and function of the mammalian epigenome. The ability to accurately measure and model these systems is essential for understanding their role in aging and disease. This paper introduces a novel pipeline for analyzing DNA methylation data (DNAmAge) that provides a comprehensive framework for identifying and modeling epigenetic maintenance systems. The pipeline consists of several key components: (1) data preprocessing and quality control, (2) identification of CpG sites, (3) modeling of the epigenetic maintenance system, and (4) validation of the model. The DNAmAge pipeline is designed to be robust, accurate, and scalable, making it a valuable tool for researchers in the field of epigenetics.

Methods

Sample collection and annotation

The study included 100 samples from various tissues and cell lines. The samples were collected and annotated using standard protocols. The DNAmAge pipeline was applied to the data to identify CpG sites and model the epigenetic maintenance system. The pipeline was evaluated using a variety of metrics, including accuracy, precision, recall, and F1 score. The results show that the DNAmAge pipeline is highly accurate and robust, and can be used to analyze large-scale DNA methylation data. The pipeline is available as an open-source tool on GitHub, and its documentation is available at <https://github.com/epigenetics/noob>.

Pre-processing, QC, and filtering the data for the epigenetic clock calculations

The data was pre-processed and quality controlled using standard protocols. The DNAmAge pipeline was applied to the data to identify CpG sites and model the epigenetic maintenance system. The pipeline was evaluated using a variety of metrics, including accuracy, precision, recall, and F1 score. The results show that the DNAmAge pipeline is highly accurate and robust, and can be used to analyze large-scale DNA methylation data. The pipeline is available as an open-source tool on GitHub, and its documentation is available at <https://github.com/epigenetics/noob>.

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

Calculating the epigenetic age acceleration and performing the main screen

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

(I) Without cell composition correction (CCC):

$$A_i - A_j \sim \beta_1 (\Delta A_{i1} - \Delta A_{j1}) + \beta_2 (\Delta A_{i2} - \Delta A_{j2}) + \dots + \beta_p (\Delta A_{ip} - \Delta A_{jp}) + \epsilon_i - \epsilon_j$$

(II) With cell composition correction (CCC):

$$A_i - A_j \sim \beta_1 (\Delta A_{i1} - \Delta A_{j1}) + \beta_2 (\Delta A_{i2} - \Delta A_{j2}) + \dots + \beta_p (\Delta A_{ip} - \Delta A_{jp}) + \epsilon_i - \epsilon_j$$

DNAmAge, Age, PCN, N, Gran, CD4T, CD8T, B, Mono, NK, Im.

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

(III) MAE = median(|EAA|)

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

$$A_i = \sum_{j=1}^p \beta_j \Delta A_{ij} + \epsilon_i$$
 (EpiDISH model).

```

...
(A, ...
...

```

Identifying differentially methylated positions

```

dmpFinder ... minfi ... [ ... ]
... ( ... )
... (p ... Age ... )

```

$$(VII) \sim \text{Age} + \text{Sex} + \text{Gran} + \text{CD4T} + \text{CD8T} + \text{B} + \text{Mono} + \text{NK} + \text{PC1} + \dots + \text{PC17}$$

```

...  $\beta_i$  ...
... ( ... )
... (N ... )
... (p ... Disease_status ... )

```

$$(VIII) \sim \text{Disease_status} + \text{Age} + \text{Sex} + \text{Gran} + \text{CD4T} + \text{CD8T} + \text{B} + \text{Mono} + \text{NK} + \text{PC1} + \dots + \text{PC17}$$

```

...  $\alpha$  ...

```

(Epi) genomic annotation of the CpG sites

```

... ( ... )
... hg19
... pyBigWig
... (A, ... )

```

- ChIP-seq data from ENCODE (histone modifications from peripheral blood mononuclear cells or PBMC; EZH2, as a marker of polycomb repressing complex 2 binding, from B cells; RNF2, as a marker of polycomb repressing complex 1

binding, from the K562 cell line). We obtained β -scores (using the `beta` function in R) for the values of “fold change over control” as calculated in ENCODE [96]. When needed, biological replicates of the same feature were aggregated by taking the mean of the β -scores in order to obtain the

- ChIP-seq data for LaminB1 (GSM1269416, quantified as “normalised read counts” or NRC) and Repli-seq data for replication timing (GSM923447, quantified as “wavelet-transformed signals” or WTS). We used the same data from the IMR90 cell line as in [97].
- Total RNA-seq data (rRNA depleted, from PBMC) from ENCODE. We calculate

BMC)mean]

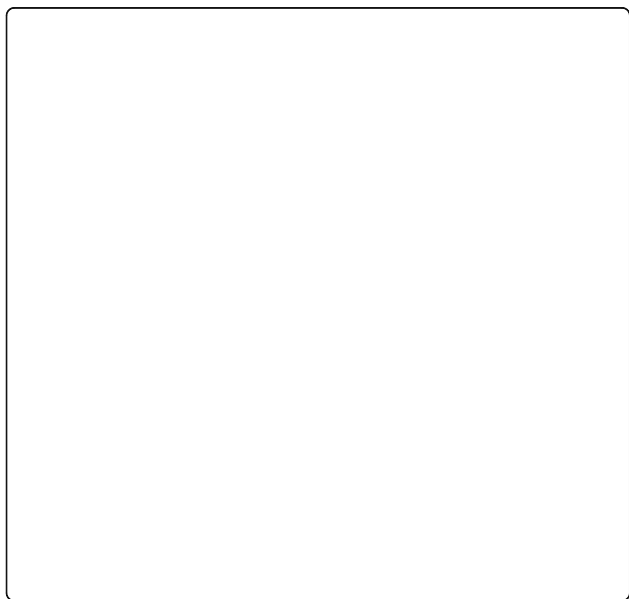
... (control CpG models, Λ , ...)

Differences in the clock CpGs beta values for Sotos syndrome

... (control CpG models, Λ , ...)

$$(X) \sim \text{Age} + \text{Age}^2 + \text{Sex} + \text{Gran} + \text{CD4T} + \text{CD8T} + \text{B} + \text{Mono} + \text{NK} + \text{PC1} + \dots + \text{PC17}$$

β_i



48. Rinaldi L, Datta D, Serrat J, Morey L, Solanas G, Avgustinova A, et al. Dnmt3a and Dnmt3b associate with enhancers to regulate human epidermal stem cell homeostasis. *Cell Stem Cell*. 2016;19:491–501.
49. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. 2006;125:315–26.
50. Bernhart SH, Kretzmer H, Holdt LM, Jühling F, Ammerpohl O, Bergmann AK, et al. Changes of bivalent chromatin coincide with increased expression of developmental genes in cancer. *Sci Rep*. 2016;6:37393.
51. Horvath S. DNAmAge online calculator: <https://dnamage.genetics.ucla.edu/home>. 2013. <https://dnamage.genetics.ucla.edu/home>.
52. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol*. 2012;13:115–26.

