Profiling epigenetic age in single cells

Alex Trapp, Csaba Kerepesi, Vadim N. Gladyshev
Division of Genetics, Department of Medicine
Brigham and Women's Hospital, Harvard Medical School

Abstract

DNA methylation dynamics emerged as a promising biomarker of the mammalian aging process, with multivariate machine learning models ('epigenetic clocks') enabling measurement of biological age in bulk tissue samples. However, intrinsically sparse and binarized methylation profiles in individual cells have so far precluded the assessment of aging in single-cell data. Here, we introduce *scAge*, a novel statistical framework for epigenetic age profiling at single-cell resolution. Our approach recapitulates chronological age in mouse tissues, while simultaneously uncovering heterogeneity among cells. We show accurate tracking of aging in single hepatocytes and uncover attenuated epigenetic aging in muscle stem cells. We also use *scAge* to reveal, at the single-cell level, a natural rejuvenation event occurring during early embryogenesis. We report our framework as a resource to dissect heterogeneous epigenetic aging trajectories at single-cell resolution.

Introduction & Methods

Current single-cell sequencing workflows produce sparse, binary methylation profiles with limited CpG overlap across different cells **(Figure 1)**, precluding the use or application of conventional elastic-net regression methods[1]. To overcome these limitations, we designed a novel framework that leverages the relationship between methylation and age at certain CpG sites in bulk data, enabling accurate epigenetic age predictions in individual cells **(Figure 2)**. First, we intersect single-cell methylomes with a reference set of pre-computed linear models, selecting only highly age-associated CpG sites common between the reference and any given single cell **(Figure 3)**. Once a filtered methylation profile is obtained, we employ a statistical algorithm that computes the probability of observing a binary methylation state at one CpG for any given age across a wide range. Then, we iterate our algorithm across many age-associated CpGs, producing an age-likelihood profile for every cell **(Figure 4)**. Finally, we pick the age of maximum likelihood as the predicted epigenetic age of the cell.

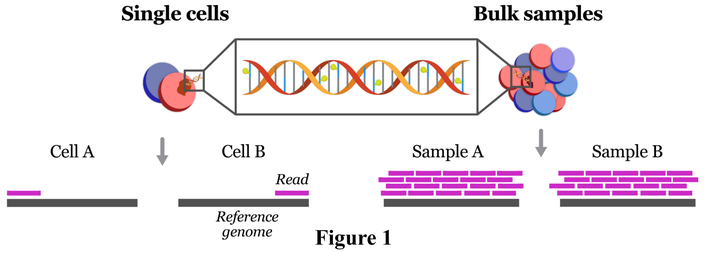

Figure 1
Schematic comparison of single-cell and bulk methylation sequencing outputs

Framework

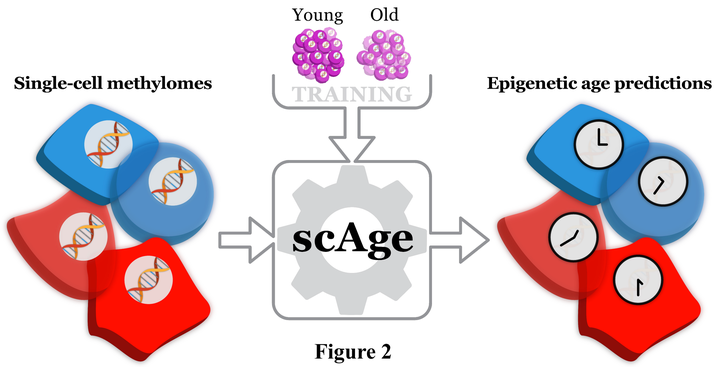

Figure 2
Schematic workflow of *scAge*

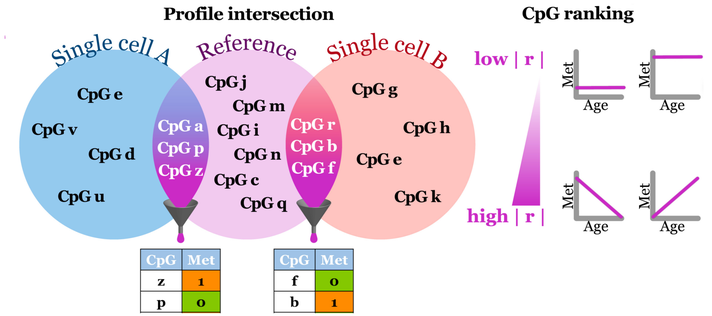

Figure 3
Intersection (left) and age-associated ranking (right) steps of the *scAge* framework

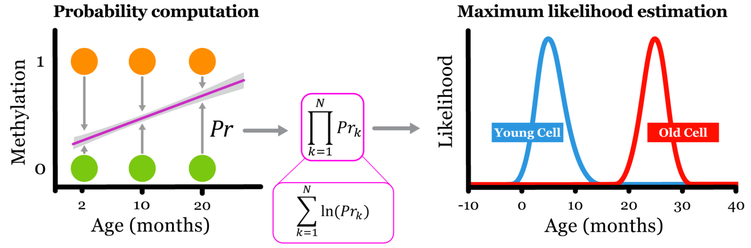

Figure 4
Schematic of the linear model-based probability computation algorithm (left), which is leveraged across many age-associated CpGs (middle) to produce distinct age-likelihood profiles (right) for every cell

Results & Discussion

We first validated our *scAge* approach on hepatocytes and mouse embryonic fibroblasts (MEFs)[2] **(Figure 5)**. Remarkably, our framework trained on bulk liver samples was able to discern between hepatocytes from young (4-month-old) and aged (26-month-old) livers with very high accuracy ($r = 0.95$). Additionally, epigenetic age predictions for MEFs were close to 0, corroborating their embryonic origin.

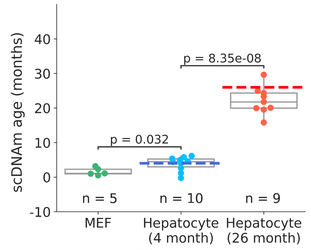

Figure 5
scAge predictions in hepatocytes and MEFs

Next, we applied *scAge* trained on bulk muscle samples to individual muscle stem cells from young (1.5-month-old) and aged (26-month-old) mice[3] **(Figure 6)**. Interestingly, we observed attenuated epigenetic aging in these cells, corroborating earlier pseudo-bulk predictions[3]. This suggests deep heterogeneity in the individual epigenetic aging trajectories of distinct cell types in muscle tissue.

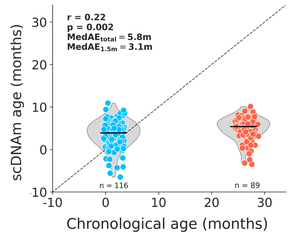

Figure 6
scAge predictions in muscle stem cells

Lastly, we applied our approach to single mouse embryonic cells around the time of gastrulation[4] **(Figure 7)**. This revealed a natural rejuvenation event during embryogenesis, where biological age is reset from parent to offspring. These findings are in line with recent application of bulk epigenetic clocks to mouse embryonic data[5].

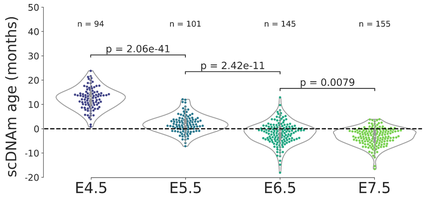

Figure 7
scAge predictions in embryonic cells

Overall, we report *scAge*, a novel framework for epigenetic age profiling at single-cell resolution. We validate our method across multiple mouse tissues of different ages, revealing heterogeneity in the aging processes of individual cells. Moreover, we observe promising performance of our statistical approach in predicting biological age, with diverse applications at the intersection of aging research, rapidly evolving single-cell technologies, and emerging rejuvenation therapies.

References:
1. Bell, C. G. *et al.* DNA methylation aging clocks: challenges and recommendations. *Genome Biol.* **20**, 249 (2019).
2. Gravina, S., Dong, X., Yu, B. & Vijg, J. Single-cell genome-wide bisulfite sequencing uncovers extensive heterogeneity in the mouse liver methylome. *Genome Biol.* **17**, 150 (2016).
3. Hernando-Herraez, I. *et al.* Ageing affects DNA methylation drift and transcriptional cell-to-cell variability in mouse muscle stem cells. *Nat. Commun.* **10**, 4361 (2019).
4. Argelaguet, R. *et al.* Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* **576**, 487–491 (2019).
5. Kerepesi, C., Zhang, B., Lee, S.-G., Trapp, A. & Gladyshev, V. N. Epigenetic clocks reveal a rejuvenation event during embryogenesis followed by aging. *Sci. Adv.* **7**, eabg6082 (2021).

Acknowledgements & Competing Interests:
We are grateful to Tiamat Fox and Adit Ganguly for help with schematic figures. We also thank the entirety of the Gladyshev Lab for helpful discussion, and NIA for funding to Vadim N. Gladyshev. Brigham and Women's Hospital is the sole owner of a provisional patent application regarding this work, which lists Alexandre Trapp, Csaba Kerepesi and Vadim N. Gladyshev as inventors.