A decade of epigenetic change in aging twins: genetic and environmental contributions to longitudinal DNA methylation

Chandra A. Reynolds [1]*, Qihua Tan [2], Elizabeth Munoz [1][a], Juulia Jylhävä [3], Jacob Hjelmborg [2], Lene Christiansen [2,4], Sara Hägg [3], and Nancy L. Pedersen [3]

[1] University of California, Riverside, USA

[2] University of Southern Denmark, Odense, Denmark

[3] Karolinska Institutet, Stockholm, Sweden

[4] Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

[a] now at University of Texas at Austin, USA

* Corresponding author: Chandra A. Reynolds, Department of Psychology, University of California Riverside, Riverside, CA, 92833 USA; Chandra.Reynolds@ucr.edu

Summary/Abstract

Background. Epigenetic changes may result from the interplay of environmental exposures and genetic influences and contribute to differences in aging outcomes. However, the etiologies contributing to stability and change in DNA methylation have rarely been examined longitudinally. **Methods**. We considered DNA methylation in whole blood leukocyte DNA across a 10-year span in two samples of same-sex aging twins: (a) Swedish Adoption Twin Study of Aging (SATSA; N = 53 pairs, 53% female; 62.9 and 72.5 years, SD=7.2 years); (b) Longitudinal Study of Aging Danish Twins (LSADT; N = 43 pairs, 72% female, 76.2 and 86.1 years, SD=1.8 years). Joint biometrical analyses were conducted on 359,399 methylation probes in common. Bivariate twin models were fitted, adjusting for age, sex and country. **Results**. Overall, results suggest genetic contributions to DNA methylation across 359,399 sites tended to be small in effect size and lessen across 10 years (broad heritability *M*=23.8% and 18.1%) but contributed to stability across time while person-specific environmental influences explained emergent influences across the decade. Aging-specific sites identified from prior EWAS and methylation age clocks were more heritable than background sites (1.1 to 2.8-fold higher). The 5,172 sites that showed the greatest heritable/familial influences (p<1E-07) were enriched for immune and inflammation pathways and neurotransmitter transporter activity pathways. **Conclusions**. Across time, stability in methylation is primarily due to genetic contributions, while novel experiences and exposures contribute to methylation differences. Elevated genetic contributions at age-related methylation sites suggest that adaptions to aging and senescence may be differentially impacted by genetic background.

Keywords: DNA methylation, heritability, aging, longitudinal

Key messages

- Individual differences in late-life methylation are due partly to genetic influences which contribute to stability in methylation patterns across late life.
- Environmental experiences unique to individuals contribute to new influences on methylation patterns across late life.
- Age-related CpG sites show greater heritable influences on methylation consistent with genetic regulation of biological aging rates.

Introduction

The functional profiles of genes are not static and vary across time, and indeed across the lifespan, in part as a result of different environmental exposures and contexts (1-6). Measurable gene-environment dynamics for behavioral traits are possible due to advances in biotechniques for global epigenetic profiling at, e.g. specific methylation sites in the human genome. Epigenetic changes may be critical to the development of complex diseases, accelerated aging, or steeper declines in cognitive and physical functioning with age (1). Understanding epigenetic changes over time in the elderly may identify pathways of decline or plasticity (e.g., maintenance or even boosts in functioning) during the aging process and help with elucidating the biology of aging and survival.

Epigenetic modifications resulting in altered gene expression may occur due to a number of processes, including direct methylation of DNA (7). DNA methylation results from *nongenetic* (i.e., epigenetic) processes that may arise due to prenatal or early life exposures or at later points in development (8, 9). DNA methylation is characteristically produced by the addition of a methyl group to the DNA molecule cytosine within cytosine-guanine dinucleotides (CpGs), estimated at 28 million sites across the human genome (10). Dense regions of CpGs referred to as 'islands' and represent about 5% of CpGs occurring in the genome (about 20,000 total) and often reside in promotor regions (11); in addition, surrounding 'shores' and 'shelves' to these islands are of interest and may be differentially methylated compared to islands (4). The addition of methylation tags to CpG sites alters gene expression, typically by interfering with or silencing gene transcription although upregulation of gene expression has been documented (12), and

may differentially occur in cells across multiple tissue types including brain, muscle and leukocytes (13). Methylation tags can be removed as a consequence of exposures as well, leading to dynamics in expression across time (9).

Although epigenetic variation is largely attributed to environmental factors, there is evidence for genetic contributions to variation in methylation across the epigenome (5, 6), with an average of 16.5 – 18.0% across sites in the Illumina 450k chip array from whole blood and common environmental influence of 3 to 16.7%, with stronger evidence of common environment in young adulthood (6).

Epigenetic changes may accelerate over time and potentiate the development of health and aging conditions earlier in life. Indeed, methylation is correlated with age (5, 14), is used to define biological clocks that may more closely track biological aging (15), and is associated with mortality (16) and a number of physical and neuropsychiatric health traits (1, 9). Longitudinal studies of twins represent a valuable approach to evaluate genetic and environmental contributions to stability and change in methylation across the methylome (17). Investigations of etiological contributions have relied primarily on cross-sectional data (5, 6) and have addressed age-related differences (5) but not change. We evaluate DNA methylation across a 10-year span in two Scandinavian samples of same-sex aging twins to evaluate the heritable and environmental contributions to stability as well as emergent influences on methylation at individual CpG sites across the methylome. Moreover, we examine whether surrounding 'shores' and 'shelves' may be differentially methylated compared to islands, and, whether epigenome-wide association study (EWAS) sites identified for aging and individual CpG clock sites are differentially heritable.

Methods

Sample. We considered DNA methylation across a 10-year span in 96 pairs of same-sex aging twins (40 monozygotic, MZ pairs; 56 dizygotic, DZ pairs). Across the two samples, the average age at time 1 was 68.89 years (*SD*=8.58) and at time 2 was 78.59 years (*SD*=8.70). Specifically, the Swedish Adoption Twin Study of Aging (SATSA) included 53 pairs (22 MZ, and 31 DZ pairs; 53% female), selected with measurements about 10 years apart (range = 8.00 to 11.82 years) at ages 62.9 and 72.5 years at time 1 and time 2 respectively (SD=7.2). The Longitudinal Study of Aging Danish Twins (LSADT) included 43 pairs (18 MZ, and 25 DZ pairs; 72% female) at ages 76.2 and 86.1 years at time 1 and time 2 (SD=1.8).

Materials. Methylation measurements from the Illumina HumanMethylation450 array (Illumina, San Diego, CA, USA) were preprocessed and normalized with adjustments for cell counts and batch effects. Processing of the SATSA sample probes has been described previously (18, 19) and in brief included: (a) preprocessing with the R package RnBeads (20) where filtering of samples and probes proceeded with a greedy-cut algorithm maximizing false positive rate versus sensitivity at a detection p-value of 0.05; (b) removal of sites that overlap with a known SNP site or reside on sex chromosomes; (c) normalization of data using dasen (21); (d) applying a Sammon mapping method (22) to remove technical variance; (e) adjustment for cell counts (23); (f) correction for batch effects using the ComBat approach in the sva package (24).

Processing of the LSADT data has described previously and in brief included (25): (a) preprocessing with the R-package MethylAid (26) where samples below quality requirements

were excluded and probes with detection p-value>0.01, no signal, or bead count<3 were filtered out; (b) removal of probes with >5% missing values, removal of sites that reside on sex chromosomes or cross-reactive probes; (c) normalization and batch-correction using functional normalization(27) with four principal components.

Although Beta-values are preferred for interpretation of methylation, Beta-value units were translated into M-values via a log2 ratio for improved distributional properties for the analysis of individual differences (28). After performing the preprocessing steps, 390,894 probes remained for SATSA and 452,920 CpG sites remained for LSADT.

Altogether 368,391 sites were in common across the Swedish and Danish samples. After the described QC pre-processing in SATSA, 49 of 53 pairs had methylation data available for both members of each pair at both time points, and 4 pairs of 53 where one cotwin had missing methylation data at time 1 (3 cotwins) or time 2 (1 cotwin). After pre-processing, LSADT sample had methylation data represented for both cotwins among the 43 pairs.

Filtering of sites post- analysis. We conducted additional filtering of probes where model-fitting results evidenced means or variances outside of expected values. Specifically, we filtered based on the typical range of M-values (c.f., 28), with expected mean values falling outside the range -6.25 to 6.25 for time 1 for 1232 sites or time 2 for 632 sites. Likewise, we filtered based on expected standard deviations exceeding 1.5 (28) resulting in 5122 exceeding the threshold for time 1 and 6190 at time 2. The effective reduction is sites was from 368,391 to 359,399 after dropping 8992 unique sites from the analysis set.

Analysis. Bivariate twin models of M-values were fitted, adjusting for centered age (centered at the average age across time = age - 74 years), sex (0=males, 1=females), and

country (0=Sweden, 1=Denmark). Bivariate ACE and ADE Cholesky models evaluated the degree to which additive genetic (A), dominance or non-additive genetic (D), common environmental (C), and unique non-shared environmental influences (E) contributed to variation and covariation in M-values within and across time (see Figure 1).The resolution of the genetic and environmental effects are done by comparing the relative similarity of monozygotic (MZ) twins who share 100% of their genes in common, including all additive effects and dominance deviations, versus dizygotic (DZ) twins who share on average 50% of segregating genes in common leading to expectations of 50% for additive effects and 25% of dominance deviations. Both twin types are presumed to have the same contribution of common environmental effects that contribute to twin similarity. We fitted ADE and ACE models as dominance (D) and common environment (C) could not be simultaneously estimated (see Figure 1).

Fit comparison between the ACE and ADE models was done via Akaike Information Criterion (AIC; (29)). If the fit of the ADE model was as good or better than the ACE model it was retained as 'best' fitting, and otherwise the ACE model was retained as best. We evaluated submodels including AE, CE and E models. Differences in nested model deviance statistics [-2ln(L)] are distributed as chi-square (χ^2) with the difference in the number of parameters between the full and constrained models as the degrees-of-freedom (df). LSADT samples tended to show lower variability in methylation at any given probe compared to SATSA, hence we allowed for scalar differences at each time-point (k_1, k_2) in variances between the two samples (see Figure 1). Thus, the relative contributions of A, C or D, and E were equated across LSADT and SATSA, but the scalar allowed for the variance components to differ by a constant at each assessment.

In comparing relative heritabilities across sites by location, we fitted random effects regression models to age 69 and 79 biometrical estimates using *lme* (version 1.1-21; 30). We allowed for random effects between and within sites, reflecting consistency of effects by CpG sites across age and nonsystematic variation across time.

Enrichment analyses were conducted using Great 3.0.0 (31). Selected sites were mapped to the Human GRCh37 build and default settings were used for association rules (i.e., basal+extension: 5000 bp upstream, 1000 bp downstream, 1000000 bp max extension, curated regulatory domains included). We present results of both binomial and hypergeometric tests where the False Discovery Rate (FDR) achieved $p < .05$ and where fold enrichment (FE) tests exceeded 2.0.

Results

Broad-sense heritable contributions (A+D, N=359,399) were on average small at age 69 years (*M* = 0.238*100 = 23.8%, time 1) and decreased across 10 years (*M* = 0.181*100 = 18.1%, time 2) (see Table 1). Across time, heritabilities showed divergence by location [ADE best (A+D): χ^2 (5) = 943.871, *p* = 8.50E-202; ACE best (A): χ^2 (5) = 562.25, *p* = 2.88E-119], where islands and shelves showed lower heritabilities by about .01 or 1% than opens seas (*p* = 4.73E-02 to 9.15E-96) and shores were the same or higher than opens seas (see Figure 2, Supplemental Table S1). At *p*<1E-07, 5172 or 1.4% of sites showed broad genetic (A, D) or familial effects (A, C) within or across time (*df* = 6), and 36033 or 10.0% of sites met *p*<1E-02.

Among the 368,391 sites, 52% of sites showed the better-fitting model was ADE (N=187,937) while 48% showed ACE as the better-fitting (N=171,462) (see Table 1, Figure 3). In terms of contributions to stability and change in methylation due to genetic or environmental,

across 359,399 sites, 58.5% showed cross-time associations at $p < .05$ (df = 3, where a_{12}=[d_{12} or c_{12}]=e_{12}=0) indicative of stability due either to genetic and/or environmental mechanisms. As shown in Figure 3, the cross-time stability was largely due to genetic effects in both the ADE best and ACE best models which was most often perfect in correlation.

Sites in which E explained all of the variability of M-values (>99%) at both time points included 8,274 total sites (5,525 ADE best, 2,749 ACE best). In all these cases, we observed that either the MZ twin correlations of M-values were negative (< 0), or DZ correlations were sufficiently negative (< -.05), or the difference between MZ and DZ correlations at each time point were sufficiently negative (< -.1).

Aging-related sites. We evaluated the best-fitting ADE and ACE results of two published CpG sets that were identified in EWAS as related to Aging: (I) 1217 sites from Wang et al. (19); (II) 1940 sites from Tan et al. (32). Multilevel regression models compared heritabilities by location from the ADE best or ACE best model, fitted to both age 69 and 79 estimates in set I [ADE best (A+D): χ^2 (5) = 72.86, p = 2.60e-14; ACE best (A): χ^2 (5) = 49.56, p = 1.71e-09], with Islands showing lower heritabilities by .09-.10 or up to a 10% difference than open seas (both $p \leq$ 2.44e-11; see supplemental Table S1). In set II, age 69 and 79 heritability estimates also showed divergence by location [ADE best (A+D): χ^2 (5) 28.04, p = 3.57E-05; ACE best (A): χ^2 (5) = 35.65, p = 1.12E-06], with North and South Shores showing higher heritabilities by about .04 or 4% than open seas (all $p \leq$ 1.71E-03; see supplemental Table S1).

Among the Aging set CpG sites, stronger heritable influences were apparent compared to the total set with up to 1.7- to 2.8-fold higher broad heritabilities in set I compared to the overall respective ADE or ACE best-fit results, and 1.4 to 2.0-fold higher in set II (see

Supplemental Table S2). Higher estimates of common environment (C) were also observed at 1.3 to 1.5-fold higher in the aging sets than the total.

Available CpG sites from three epigenetic-clocks were evaluated in similar fashion: (a) 71 sites Hannum clock (33), (b) 353 sites Horvath clock (34), and (c) 513 sites Levine clock(35). Among the 776 CpG sites from the epigenetic-clocks, where ADE fit best for 490 CpGs and ACE for 286 sites. When multilevel regression models were fitted to age 69 and 79 estimates, Hannum clock sites tended to show stronger genetic and shared environmental contributions from .05 to .07, or 5-7%, than Horvath or Levine clock sites when ACE was the better fitting model ($p \leq 9.32E-03$), but not for ADE ($p = 1.48E-01$) (see supplemental Table S3). The ratio of Intercept variability to total variability (ρ) in heritability estimates was .047 for ADE best models and .010 for ACE best suggesting 4.7% and 1% of the variation in heritability, respectively, was methylation site specific across time though the clear majority of the variation was unique to site and time. Heritability estimates were 1.1- and 1.3-fold greater than the respective total set ADE best and ACE best estimates, respectively (see Supplemental Table S2).

Enrichment analysis. A set of 5172 CpGs achieving $p<1E-07$ when evaluating tests of heritability (AD vs E; N = 2117) or familiality (AC vs E; N=3055) across time were submitted to GREAT 3.0.0 to identify functions of cis-regulatory regions (31). This represents 1.4% of the 359,399 sites tested. Specifically, we report the binomial and hypergeometric tests over genomic regions covered by the 5172 CpGs, reporting those that achieved region-based fold enrichment (FE) > 2 and both binomial and hypergeometric FDR Q-Values < .05 (see Table 2). For full ontology results see Supplemental Tables S4 – S6. The sites that showed the greatest

heritabilities showed enrichment in immune and inflammation pathways and neurotransmitter transporter activity pathways.

Discussion

Overall, results suggest genetic contributions to DNA methylation tended to be small and decrease across a decade, however, genetic influence mainly contributed to stability of methylation. Unique person-specific influences not shared by co-twins were emergent across 10 years suggesting that non-shared environments become more salient to DNA methylation in late life. The sites that showed the greatest heritabilities showed enrichment in immune and inflammation pathways and neurotransmitter transporter activity pathways.

Prior studies report average heritabilities of 16.5 – 18.0% across adulthood (17-79 years) (5, 6), and common environmental influences of 3.0 – 16.7%, with stronger evidence of common environment in young adulthood (6). Our results of weakening heritable influences across age is consistent with the Dutch cross-sectional study reporting average heritabilities of 21% and 18% at ages 25 and 50 assuming an AE model, whereas our estimates of broad heritability under an ADE model are 24% and 18% many decades later at age 69 and 79 years, respectively. Where non-additive genetic effects fit best, the average broad heritability was 24% across age. For sites where including common environment fit best (ACE), lower average heritabilities were observed at 7% and common environment contributed 10% to variation in methylation across age.

CpG sites related to aging show a greater impact of heritable influences consistent with genetic regulation of the rate of biological aging. Sites associated with aging and longevity

generally show higher heritabilities than the total background sites, 1.4 to 2.8-fold higher values. CpGs related to aging, as well as the total set of CpGs, varied in magnitude of heritabilities by location, where 'islands', which often reside in promotor regions (11), typically showed lower heritability than those sites residing in surrounding 'shores' and 'shelves', which have been shown to be differentially methylated compared to islands (4).

Moreover, the set of methylation clocks sites are likewise appreciably heritable, with the Hannum sites 1.1 to 1.9 times more heritable than Horvath and Levine sites. We have recently reported heritability estimates of methylation clock ages of 52% for the Horvath clock and 36% for the Levine clock (36), where, consistent with our current site-specific effects, stability across time was mediated primarily by genetic factors, whereas the person-specific environmental factors contributed to differences across time. The Horvath clock contains 353 CpG sites selected as best predictors of chronological age using multiple tissues (34), the Levine clock 513 sites were selected based on prediction of chronological age and nine biomarkers of phenotypic aging and also trained on multiple tissues (35), whereas the Hannum clock's 71 selected sites best predicted age (adjusted for sex, BMI) based on methylation observed in whole blood. The current findings of moderately higher heritabilities in Hannum versus the other clocks may be in part due to our use of blood tissue.

Enrichment analyses of the 1.4% of sites meeting p<1E-07 suggest immune and inflammation pathways and neurotransmitter transporter activity pathways may feature in sites with strong heritable or familial components. A related study of German and Danish individuals (including an overlapping sample of twins herein) evaluating RNA-sequencing expression patterns and longevity identified expression patterns in biological processes

contributing to immune system and response pathways (37), and observed high heritabilities (30-99%) among 20% of cis-eQTLS. Immunosenescence describes an age-associated decline in elderly individuals' immune functioning, such as mounting less effective responses to vaccines, and lowered resistance to illnesses, with concomitant up-regulation of pro-inflammatory cytokines, among several other cellular and physiological changes in the immune system (38). It has been proposed that heritable factors may be partly associated with differential immune responses (39, 40) and may predict influenza-related susceptibility and mortality (41), for example, and, broadly, successful aging and longevity (39, 40). Hence, differential adaptions to aging processes including immunosenescence reflect gene-environment dynamics with some individuals showing better adaptions than others due to genetic influences.

The current study establishes the extent to which the genetic and environmental influences contribute to site-specific methylation across a 10-year span in a longitudinal sample of Swedish and Danish twins. While stability of methylation was largely due to genetic influences, person-specific environmental influences were emergent across time and explained change. By and large, the dynamics of methylation may be influenced by experiences and exposures, suggesting possible mediation of gene expression; however, the most heritable sites may participate in immune and inflammation pathways and neurotransmitter transporter activity pathways which suggest that adaptions to aging and senescence may be differentially impacted by genetic background.

Funding

SATSA has been supported by the National Institute on Aging (AG04563, AG10175), the MacArthur Foundation Research Network on Successful Aging, the Swedish Council for Working Life and Social Research (FAS) (97:0147:1B, 2009-0795), and the Swedish Research Council (825-2007-7460, 825-2009-6141). DNA extraction supported in part by AG028555, AG17561. Methylation work supported by FORTE 2013-2292, the Swedish Research Council (521-2013-8689, 2015-03255) and a Distinguished Professor Award from the KI to NLP.

LSADT study has been supported by grants from the VELUX FOUNDATION, the U.S. National Institute on Aging (P01-AG08761), the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement no. 259679, and The Danish National Program for Research Infrastructure 2007 [09-063256].

Collaborative work has been supported in part by the National Institute on Aging (AG037985). The manuscript content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Lappe M, Landecker H. How the genome got a life span. New Genet Soc. 2015;34(2):152-76.
2. McClearn GE. Contextual genetics. Trends Genet. 2006;22(6):314-9.
3. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. PLoS Genet. 2012;8(4):e1002629.
4. Jones MJ, Goodman SJ, Kobor MS. DNA methylation and healthy human aging. Aging Cell. 2015;14(6):924-32.
5. van Dongen J, Nivard MG, Willemsen G, Hottenga J-J, Helmer Q, Dolan CV, et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. Nature Communications. 2016;7:11115.
6. Hannon E, Knox O, Sugden K, Burrage J, Wong CCY, Belsky DW, et al. Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. PLoS Genet. 2018;14(8):e1007544.
7. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science. 2001;293(5532):1068-70.
8. Gottesman, II, Hanson DR. Human development: biological and genetic processes. Annu Rev Psychol. 2005;56:263-86.
9. Kanherkar RR, Bhatia-Dey N, Csoka AB. Epigenetics across the human lifespan. Front Cell Dev Biol. 2014;2:49.
10. Lovkvist C, Dodd IB, Sneppen K, Haerter JO. DNA methylation in human epigenomes depends on local topology of CpG sites. Nucleic Acids Res. 2016;44(11):5123-32.
11. Vinson C, Chatterjee R. CG methylation. Epigenomics. 2012;4(6):655-63.
12. Wang C, Chen L, Yang Y, Zhang M, Wong G. Identification of potential blood biomarkers for Parkinson's disease by gene expression and DNA methylation data integration analysis. Clin Epigenetics. 2019;11(1):24.
13. Fernandez AF, Assenov Y, Martin-Subero JI, Balint B, Siebert R, Taniguchi H, et al. A DNA methylation fingerprint of 1628 human samples. Genome Res. 2012;22(2):407-19.
14. Ciccarone F, Tagliatesta S, Caiafa P, Zampieri M. DNA methylation dynamics in aging: how far are we from understanding the mechanisms? Mechanisms of Ageing and Development. 2018;174:3-17.
15. Field AE, Robertson NA, Wang T, Havas A, Ideker T, Adams PD. DNA Methylation Clocks in Aging: Categories, Causes, and Consequences. Mol Cell. 2018;71(6):882-95.
16. Zhang Y, Wilson R, Heiss J, Breitling LP, Saum K-U, Schöttker B, et al. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. Nature Communications. 2017;8:14617.
17. Tan Q, Christiansen L, von Bornemann Hjelmborg J, Christensen K. Twin methodology in epigenetic studies. J Exp Biol. 2015;218(Pt 1):134-9.

18. Jylhävä J, Hjelmborg J, Soerensen M, Munoz E, Tan Q, Kuja-Halkola R, et al. Longitudinal changes in the genetic and environmental influences on the epigenetic clocks across old age: Evidence from two twin cohorts. EBioMedicine. 2019;40:710-6.
19. Wang Y, Karlsson R, Lampa E, Zhang Q, Hedman AK, Almgren M, et al. Epigenetic influences on aging: a longitudinal genome-wide methylation study in old Swedish twins. Epigenetics. 2018;13(9):975-87.
20. Assenov Y, Muller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. Nat Methods. 2014;11(11):1138-40.
21. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14:293.
22. Sammon JW. A Nonlinear Mapping for Data Structure Analysis. IEEE Trans Comput. 1969;18(5):401-9.
23. Jones MJ, Islam SA, Edgar RD, Kobor MS. Adjusting for Cell Type Composition in DNA Methylation Data Using a Regression-Based Approach. In: Haggarty P, Harrison K, editors. Population Epigenetics: Methods and Protocols. New York, NY: Springer New York; 2017. p. 99-106.
24. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882-3.
25. Svane AM, Soerensen M, Lund J, Tan Q, Jylhava J, Wang Y, et al. DNA Methylation and All-Cause Mortality in Middle-Aged and Elderly Danish Twins. Genes (Basel). 2018;9(2).
26. van Iterson M, Tobi EW, Slieker RC, den Hollander W, Luijk R, Slagboom PE, et al. MethylAid: visual and interactive quality control of large Illumina 450k datasets. Bioinformatics. 2014;30(23):3435-7.
27. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol. 2014;15(12):503.
28. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11:587.
29. Akaike H. A new look at the statistical model identification. Selected Papers of Hirotugu Akaike: Springer; 1974. p. 215-22.
30. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software; Vol 1, Issue 1 (2015). 2015.
31. McLean CY, Bristor D Fau - Hiller M, Hiller M Fau - Clarke SL, Clarke Sl Fau - Schaar BT, Schaar Bt Fau - Lowe CB, Lowe Cb Fau - Wenger AM, et al. GREAT improves functional interpretation of cis-regulatory regions. (1546-1696 (Electronic)).
32. Tan Q, Heijmans BT, Hjelmborg JV, Soerensen M, Christensen K, Christiansen L. Epigenetic drift in the aging genome: a ten-year follow-up in an elderly twin cohort. Int J Epidemiol. 2016;45(4):1146-58.
33. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Molecular cell. 2013;49(2):359-67.

34. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
35. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10(4):573-91.
36. Jylhava J, Hjelmborg J, Soerensen M, Munoz E, Tan Q, Kuja-Halkola R, et al. Longitudinal changes in the genetic and environmental influences on the epigenetic clocks across old age: Evidence from two twin cohorts. EBioMedicine. 2019;40:710-6.
37. Häsler R, Venkatesh G, Tan Q, Flachsbart F, Sinha A, Rosenstiel P, et al. Genetic interplay between human longevity and metabolic pathways - a large-scale eQTL study. Aging cell. 2017;16(4):716-25.
38. Accardi G, Caruso C. Immune-inflammatory responses in the elderly: an update. Immun Ageing. 2018;15:11-.
39. Derhovanessian E, Maier AB, Beck R, Jahn G, Hahnel K, Slagboom PE, et al. Hallmark features of immunosenescence are absent in familial longevity. J Immunol. 2010;185(8):4618-24.
40. Ostan R, Bucci L, Capri M, Salvioli S, Scurti M, Pini E, et al. Immunosenescence and immunogenetics of human longevity. Neuroimmunomodulation. 2008;15(4-6):224-40.
41. Poland GA, Ovsyannikova IG, Kennedy RB, Lambert ND, Kirkland JL. A systems biology approach to the effect of aging, immunosenescence and vaccine response. Curr Opin Immunol. 2014;29:62-8.

Table 1. Variance components at time 1 (69 years) and time 2 (79 years).

		A1		D1/C1		E1		A2		D2/C2		E2	
Model	**N sites**	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
ADE	359,399	0.111	0.142	0.127	0.160	0.762	0.176	0.092	0.126	0.089	0.131	0.819	0.159
ADE best	187,937	0.058	0.106	0.218	0.168	0.725	0.182	0.048	0.092	0.152	0.149	0.800	0.169
ACE	359,399	0.151	0.167	0.057	0.083	0.792	0.164	0.109	0.143	0.054	0.080	0.837	0.148
ACE best	171,462	0.077	0.122	0.106	0.093	0.817	0.150	0.056	0.103	0.099	0.092	0.846	0.137

Note. A = Additive genetic, D = Nonadditive genetic (Dominance), C = Common environment, E = Nonshared environment.

Table 2. GREAT 3.0.0 annotations using binomial and hypergeometric tests over genomic regions covered by the 5172 CpGs showing significant heritability/familiality p<1E-07

	Bionomial						Hypergeometric					
Gene Ontology (GO)	Rank	Raw P-Value	FDR Q-Value	Fold Enrichment	Observed Region Hits	Region Set Coverage	Rank	FDR Q-Value	Fold Enrichment	Observed Gene Hits	Total Genes	Gene Set Coverage
GO Molecular Function												
MHC class II receptor activity	1	6.70E-29	2.47E-25	19.1923	31	0.0060	3	5.66E-03	3.4204	11	12	0.0023
neurotransmitter transporter activity	32	8.84E-08	1.02E-05	3.0793	31	0.0060	8	2.35E-02	2.4876	16	24	0.0033
neurotransmitter: sodium symporter activity	54	1.15E-06	7.83E-05	3.0421	26	0.0050	6	1.66E-02	2.7494	14	19	0.0029
GO Biological Process												
interferon-gamma-mediated signaling pathway	13	5.47E-12	4.39E-09	2.6727	65	0.0126	53	1.11E-03	2.0186	33	61	0.0068
establishment of protein localization to plasma membrane	39	7.59E-09	2.03E-06	2.5615	49	0.0095	184	3.25E-02	1.9091	22	43	0.0046
cellular response to interferon-gamma	46	2.89E-08	6.56E-06	2.0169	75	0.0145	55	1.17E-03	1.8657	41	82	0.0085
osteoblast development	101	1.01E-06	1.05E-04	2.3813	40	0.0077	132	1.42E-02	2.5530	13	19	0.0027
Golgi to plasma membrane transport	219	3.71E-05	1.77E-03	2.3871	28	0.0054	138	1.57E-02	2.2962	16	26	0.0033
lens fiber cell development	220	3.78E-05	1.79E-03	2.5893	24	0.0046	108	6.94E-03	3.1094	10	12	0.0021
GO Cellular Component												
MHC class II protein complex	1	1.82E-36	2.31E-33	16.9778	42	0.0081	1	8.92E-07	3.5349	18	19	0.0037
MHC protein complex	2	2.40E-32	1.52E-29	10.9766	47	0.0091	2	2.97E-06	3.0403	22	27	0.0046
integral to lumenal side of endoplasmic reticulum membrane	4	5.51E-25	1.74E-22	9.1918	40	0.0077	8	8.05E-03	2.4876	16	24	0.0033
smooth endoplasmic reticulum	123	1.91E-03	1.97E-02	2.1079	20	0.0039	7	5.01E-03	2.8534	13	17	0.0027

Note. FDR=False Discovery Rate.

Figure 1. Bivariate Cholesky model. *Note*. ACE and ADE models were separately fitted to M-values at two waves 10 years apart.

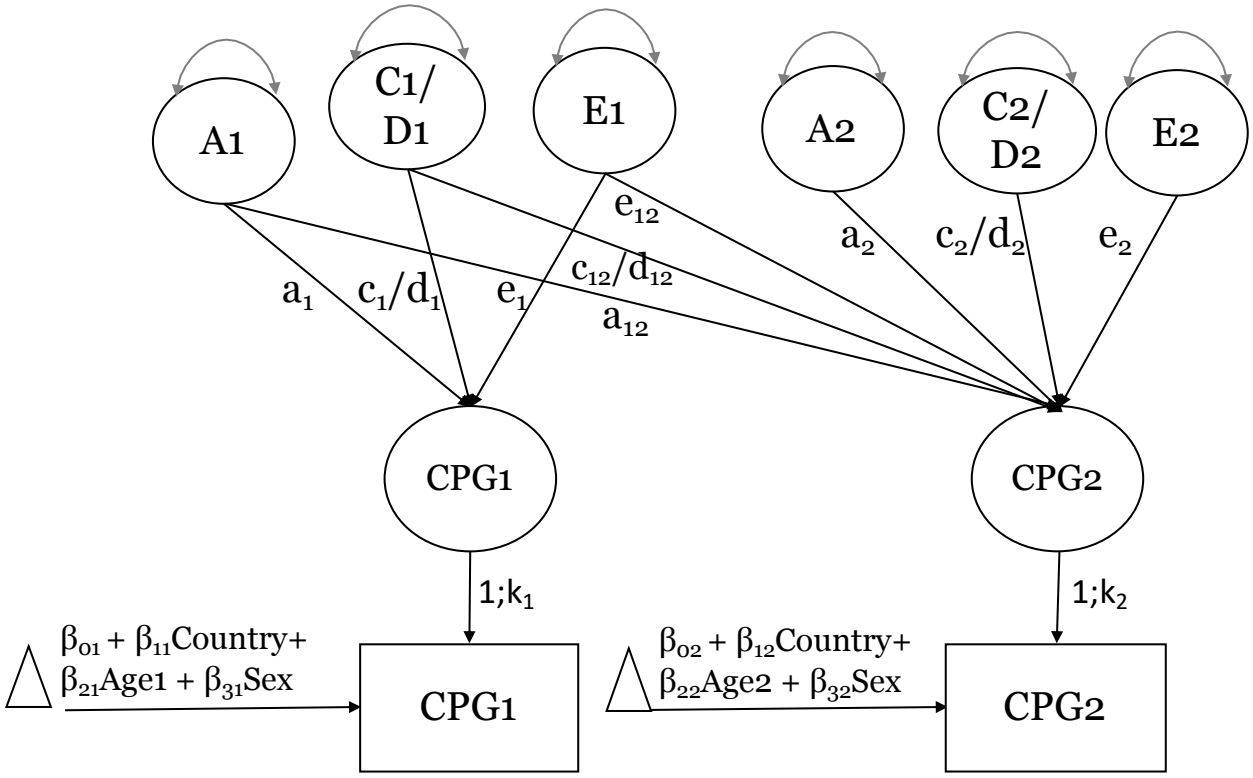

Figure 2. Broad-sense heritability by Location (ADE results, 359,399 CpGs) across 10 years

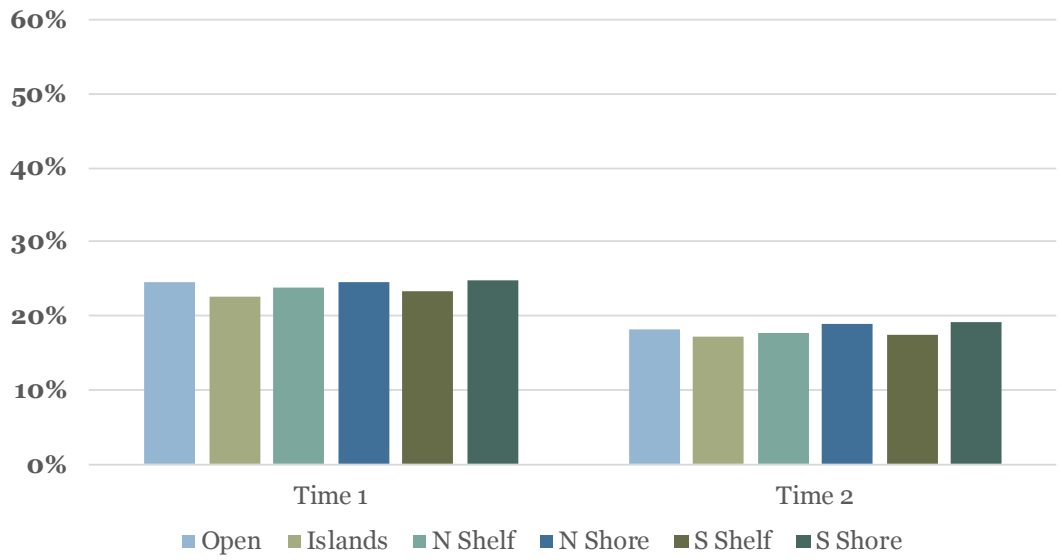

Figure 3. Best-fitting models: ADE (52%) or ACE (48%)

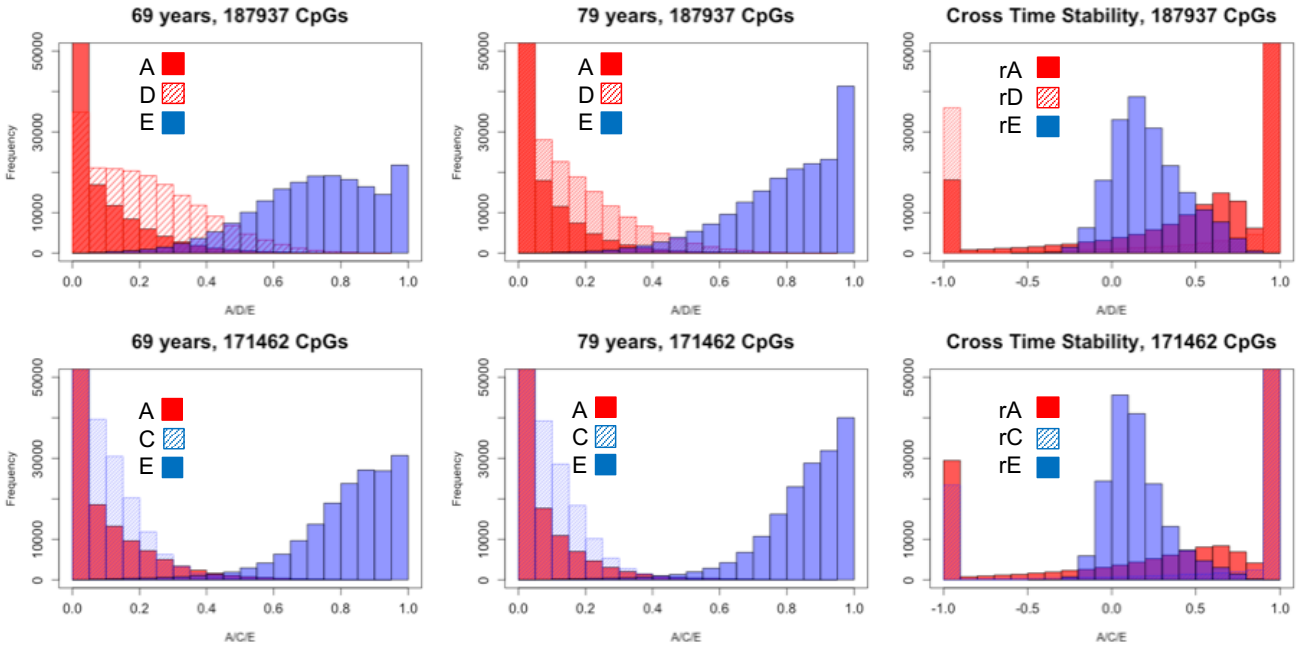

Figure 4. Aging-related CpG Sets: Broad heritability by CpG Location