Epigenetic signature of human immune aging in the GESTALT study

Roshni Roy[1], Pei-Lun Kuo[2], Julián Candia[2], Dimitra Sarantopoulou[1], Ceereena Ubaida-Mohien[2], Dena Hernandez[3], Mary Kaileh[1], Sampath Arepalli[3], Amit Singh[1], Arsun Bektas[2], Jaekwan Kim[1], Ann Zenobia Moore[2], Toshiko Tanaka[2], Julia McKelvey[4], Linda Zukley[4], Cuong Nguyen[5], Tonya Wallace[5], Christopher Dunn[5], William Wood[6], Yulan Piao[6], Christopher Coletta[6], Supriyo De[6], Jyoti Misra Sen[7], Nan-ping Weng[1], Ranjan Sen[1], Luigi Ferrucci[2]*

[1] Laboratory of Molecular Biology and Immunology
[2] Translational Gerontology Branch
[3] Laboratory of Neurogenetics
[4] Clinical Research Core
[5] Flow Cytometry Unit
[6] Laboratory of Genetics & Genomics
[7] Laboratory of Clinical Investigation
National Institute on Aging, Baltimore, MD

* - Corresponding and lead author (correspondence- ferruccilu@grc.nia.nih.gov)

ABSTRACT

Age-associated DNA methylation in blood cells convey information on health status. However, the mechanisms that drive these changes in circulating cells and their relationships to gene regulation are unknown. We identified age-associated DNA methylation sites in six purified blood borne immune cell types (naïve B, naïve $CD4^+$ and $CD8^+$ T cells, granulocytes, monocytes and NK cells) collected from healthy individuals interspersed over a wide age range. Of the thousands of age-associated sites, only 350 sites were differentially methylated in the same direction in all cell types and validated in an independent longitudinal cohort. Genes close to age-associated hypomethylated sites were enriched for collagen biosynthesis and complement cascade pathways, while genes close to hypermethylated sites mapped to neuronal pathways. In-silico analyses showed that in most cell types, the age-associated hypo- and hypermethylated sites were enriched for ARNT (HIF1β) and REST transcription factor motifs respectively, which are both master regulators of hypoxia response. To conclude, despite spatial heterogeneity, there is a commonality in the putative regulatory role with respect to transcription factor motifs and histone modifications at and around these sites. These features suggest that DNA methylation changes in healthy aging may be adaptive responses to fluctuations of oxygen availability.

Keywords- DNA methylation, immune cells, aging, ARNT, REST, hypoxia

INTRODUCTION

Human aging is associated with site-specific changes of DNA methylation. Summary measures of DNA methylation called “epigenetic clocks” are extensively used in aging research to estimate biological age (1, 2, 3). Epigenetic clocks closely approximate chronological age and beyond age, predict adverse health conditions, including frailty (4), Alzheimer’s disease (5) and mortality (6, 7).

Research suggest that changes in DNA methylation with aging are regulated by specific mechanisms rather than by a stochastic drift (8). For example, a loss-of-function mutation in the H3K36 histone methyltransferase has been associated with epigenetic aging in mice (9). In humans, polymorphisms in the telomerase gene (TERT) (10) and age-dependent gain of methylation in the Polycomb repressive complex 2 have been related to accelerated aging(11). However, so far, no sound hypothesis exists that explains the association of DNA methylation with aging and pathology.

A main obstacle in understanding mechanisms driving age-associated changes of DNA methylation is that most human studies were performed in mixed blood cell types. The few studies that investigated select immune circulating cells failed to propose a unifying biological hypothesis explaining predictable changes of DNA methylation with aging (12, 13, 14, 15, 16, 17, 18).

We analyzed age-associated methylation in 6 purified blood-borne cell types sorted from peripheral mononuclear cells (PBMCs) from 55 donors of ages ranging from 22 to 83 years. To minimize the confounding of age-associated pre-clinical and clinical diseases, participants were ascertained to be healthy by trained health professionals

according to strict clinical criteria. We looked for CpGs differentially methylated with aging in the same direction in multiple cell types. Next, in each cell type, we conducted enrichment analyses of genes close to age-associated CpGs. Finally, we looked for chromatin accessibility markers and transcription factor binding sites close to the same age-associated CpGs. Our findings suggest that changes in methylation with aging are related to fluctuation of energetic metabolism during the life course.

RESULTS

Age-associated methylation in individual cell types

A principal component analysis (PCA) was performed on normalized DNA methylation data for all cell types from all the 55 donors (Figure 1A and Supplementary File 1). The PCA showed that clustering by cell types was stronger than by age (Figure 1- figure supplement 1A). The genes associated with the top 500 probes corresponding to PC1, PC2 and PC3 were enriched pathways linked to innate and adaptive lineage development (Supplementary File 2).

Age-associated CpGs were identified through sex-adjusted beta regression models (FDR corrected p-value <0.05). The number of hypo- or hypermethylated sites varied considerably between cell types (Figure 1B) with highest numbers in $CD4^+$ T cells (Figure 1- figure supplement 1B and Supplementary Files 3-4). Using a different approach of comparing between young ($\leq$35 years, 25th percentile) and old ($\geq$70 years, 75th percentile) individuals, we observed >90% overlap with beta regression-derived hypomethylated sites and 70-95% overlap with hypermethylated sites in all cell types

except $CD8^+$ T cells (9-14% overlap) (Figure 1- figure supplement 1C). Having fewer old donors with $CD8^+$ T cells may have contributed to differences (Supplementary File 1).

Like other studies, we found that a significant proportion of age-hypomethylated CpGs were in the intergenic and open-sea (>4kb from CpG island) regions while age-hypermethylated CpGs were in promoters and CpG islands (Chi sq test $p<0.001$) (Figure 1- figure supplement 1D and 1E). Additionally, age-associated differentially methylated sites in PBMC poorly recapitulate age-dependent changes that take place in specific primary immune cells (Figure 1- figure supplement 1E and 1F). These findings point to a wide heterogeneity of age-differential CpG methylation across immune blood cells and suggest that studies in PBMC poorly represents the changes that take place in specific cell types with aging.

Shared age associated methylation across cell types

Only 181 age-associated hypomethylated sites and 169 hypermethylated sites were shared between all 6 cell types. These numbers increased to 776 (age-hypomethylated) and 404 (age-hypermethylated) sites in 5 or more cell types (Figures 1C-D). Thus, most age-related methylation changes are cell-specific. Of note, only 10 of the sites overlap with the 359 CpGs in Horvath's pan-tissue epigenetic clock (19). Several reasons can be attributed to this poor overlap including i) use of methylation array with about 21369 CpGs for development of the clock in contrast to the analyses in this study based on ~850,000 CpG sites ii) use of data from peripheral or whole blood for development these clocks in contrast to data from flow-sorted circulating immune cells in this study. While the number of shared age-hypo or hypermethylated CpGs across cells was relatively small, it was significantly much higher than that expected based on chance

alone, suggesting that common underlying epigenetic mechanisms exist across the considered cell types (Figures 1C & D). For example, CpG sites adjacent to *RCAN1* (calcineurin 1) and *KLF14* (Krüppel-Like Factor 14) show similar age-associated patterns in all cell types (Figures 1E and F).

Next, we wanted to investigate whether the top age-associated genes are the ones which are shared across cell types. For this we arranged the age-associated probes with decreasing order of adjusted p-value and looked at the annotated genes to identify the top 15 genes in each cell type (Figures 2A and B and Figure 2- figure supplements 2A-E and Supplementary File 5). THSD4 and CCDC102B were the most significant age-associated hypomethylated genes shared by 5 or more cell types, while ELOVL2, KLF14, LHFP14 and GPR158 were among the most significant age-hypermethylated genes in 5 or more cell types. This count increased to 5 and 13 genes respectively when the list was expanded to 50 top genes (Supplementary File 5). It is noteworthy that only 13-15% of these "top" age-associated probes overlapped with the list of age-associated probes shared across cell types (181 hypomethylated and 169 hypermethylated probes). These findings suggest that most CpGs with age-associated methylation consistent across cell types undergo moderate (although significant) methylation changes with aging.

Longitudinal validation of age-associated CpG sites

We hypothesized that the age-associated CpGs identified across the six immune cells in this cross-sectional study would also show longitudinal changes of the size and direction predicted. We used DNA methylation data (Illumina 450K microarray on DNA from buffy coats) assessed at baseline and 9- and 13-year follow-up in 699 participants

of the InCHIANTI study (20). Of the 181 hypo-methylated and 169 hypermethylated CpGs with age in all cell types in GESTALT, 72 and 135, respectively, were represented in the 450K microarray. The beta-coefficients for age of the 207 CpG probes (72+135) estimated from the GESTALT study and their corresponding values estimated longitudinally from the InCHIANTI study were highly and significantly correlated (hypomethylated with age CpGs: r=0.49, p=1.2e-09 and hypermethylated with age CpGs: r=0.5, p=6.9e-06 for average beta coefficients across 6 cell types, Figures 2C-D and Figure 2- figure supplement 2). Thus, CpGs identified as differentially methylated with aging across cell types in GESTALT also change longitudinally with aging.

Age-associated probes with opposite trends in different immune cells

Several CpGs showed significant but opposing age-trends in different cell types, especially in B, $CD4^+$ T cells and monocytes (Figure 2- figure supplement 1F and 1G). For example, cg27123256 in the gene body of BCL11B was age-hypomethylated in non-T cells and significantly age-hypermethylated in naïve $CD4^+$ T cells (Figure 2E). Our observations implicate BCL11B in aging-related changes in naïve $CD4^+$ T cell function, distinct from its proposed role in effector cells (14, 21, 22). Conversely, cg03530364 in the body of FAM19A1 gene was hypermethylated in non-T cells but age-hypomethylated in $CD4^+$ T cells (Figure 2F). Of note, none of these CpGs were differentially age-methylated in PBMC. Thus, opposite age-methylation trends in specific cell types may cancel each other and obscure their relevance for aging when mixed cell type samples are assessed.

Pathway analysis of age-associated genes

Gene set enrichment analyses were performed on genes associated with at least one CpG significantly age-hypo or hypermethylated in 5 or more cell types. We identified 30 pathways (q-value<0.05) (Figure 3 and Supplementary File 6). Probes commonly age-hypomethylated in 5 or more cell types (n=776) pointed to genes enriched in collagen biosynthesis, complement cascade and GTPase pathways (left-most column in bottom panel of Figure 3) that highlighted inflammatory and metabolic pathway in aging. Genes associated with shared age-hypermethylated probes (n=404) were enriched for neural pathways previously implicated to brain aging along with G-Protein Coupled Receptors pathways (23, 24) (left-most column in top panel of Figure 3). A recent study by Karagiannis et.al also identified neuronal genes in their PBMC aging data emphasizing a possible interlink between immune-aging and neuronal pathways (25). Other key pathways are highlighted, with associated genes displayed in boxes on the right-hand side.

Functional annotation of age-associated probes

To further interrogate the relationships between DNA methylation and other epigenetic states, we mapped the methylation age-associated sites to cell-specific chromHMM-derived chromatin profiles (26). As controls, we annotated all sites in the EPIC array to the 18-state chromHMM model of respective primary cell type. Granulocytes were excluded from this analysis because reference data were not available.

Age-associated hypomethylated CpGs were significantly enriched for weak/active enhancers (yellow bar, Figure 4A) whereas, confirming previous reports, age-hypermethylated CpGs, were enriched in bivalent/polycomb regions compared to

control set (brown and dark grey bars respectively in Figure 4A). Results for cell type-specific analyses are shown in Figure 4B.

We further mapped the profile of four epigenetic markers from the ENCODE project in and around (± 3kb) age-associated methylation sites. For B and $CD4^+$ T cells, we observed a V-shaped peak-valley-peak pattern of DNase hypersensitivity at sites of age-associated hypomethylation, which is characteristic of promoter sites (Figure 4C) (27). Both age-associated hypo- and hypermethylated sites showed evident H3K4me1 peaks, a marker commonly associated with active and primed enhancers (Figure 4C) (28). No specific trend was observed for H3K4me3 and H3K27ac (Figure 4- figure supplement 1). These patterns were highly consistent across cell types (Figure 4- figure supplement 1) and strongly suggest a functional connection between methylation and chromatin status. However, as the DHS and histone data in the ENCODE database was only available for either one of 2 donors (a 21-year-old male and 37-year-old female), we could not verify whether the patterns observed are stable with change in age.

Pattern of transcription factor binding motifs around age-associated CpGs.

Specific transcription factors (TF) binding may induce loss of DNA methylation or bind DNA that is methylated (29, 30). Through our *de-novo* HOMER analysis, we looked for transcription factor binding motifs in a 200bp window around the age-associated methylated sites in each cell type. We observed that the binding motif for aryl hydrocarbon receptor nuclear translocator (ARNT, also named HIF1β) was associated with age-hypomethylated CpGs across most cell types (Figure 5A). The only exception was naïve $CD8^+$ T cells where the top enriched motif was B-cell lymphoma gene 6 (BCL6). BCL6 code for a zinc finger transcription factor that plays a critical role in the

generation of memory and effector cells in acute infection (31). Another motif associated with age-hypomethylated CpGs across most cell types was chromatin architectural protein CTCF and its closely related gene BORIS. Methylation changes at CTCF sites have been reported to reflect large scale genome reorganization in immune cells in older individuals (32, 33).

Repressor Element 1-Silencing Transcription Factor (REST) was the TF motifs most frequently associated with age-hypermethylated CpGs in 5 of 6 cell types (Figure 5B). Age-hypermethylated sites in PBMCs have been previously shown to be enriched for *REST*, which is known to repress stress response genes and is lost in cognitive impairment and Alzheimer's disease pathology (34, 35). The top enriched TF motif associated with age-hypermethylated sites in monocytes was Arid5A ($p<10^{-27}$) that binds to selective inflammation-related genes, such as IL6 and STAT3 and stabilize their expression (36, 37). We further repeated the analysis with a smaller 50bp window size for TF motif search. Motifs for ARNT, CTCF and REST remained the top hits in most cell types (Supplementary File 7). However, BCL6 and ARID5A were no longer the top motifs in the search indicating that motifs for these TFs appear to be farther from the age-associated CpG sites.

The recurring enrichment of *ARNT* and *REST* with age-associated CpGs observed across multiple cell types, despite relatively few shared genomic region locations, suggests a common mechanism of gene regulation. We found that only 17 and 44 age-associated hypo- and hypermethylated probes, respectively, shared ARNT or REST motifs across all cells (Figure 5- figure supplements 1A-B), suggesting these overlaps are not random and have a specific function (Figure 5- figure supplements 1A-B).

232 Remarkably, ARNT was significantly overexpressed in older age in three of the six cell
233 types and REST showed a significant decrease of expression with age in most cell
234 types (Supplementary Files 8 and 9). These findings suggest that age-associated
235 changes in expression levels of REST and ARNT can affect the epigenetic status of
236 their target genes.

237 Age-related differential methylation and oxygen sensing.

238 ARNT, REST and BCL6, three transcription factors most associated with differentially
239 methylated regions, are implicated in hypoxia response (Figure 5C). ARNT is the beta
240 subunit of Hypoxia Factor 1 (HIF-1), which is stabilized during hypoxia and shuttled to
241 the nucleus where it binds to DNA hypoxia response elements (HRE) and triggers a
242 complex response that include upregulation of angiogenesis and erythropoiesis and
243 reprogramming of energetic metabolism from oxidative phosphorylation to anaerobic
244 glycolysis (38). Hypoxia also upregulates the transcription of REST which is the master
245 regulator of the transcriptional repression arm of the response to hypoxia. Released
246 REST is shuttled to the nucleus where it binds to DNA and regulates approximately
247 20% of the hypoxia-repressed genes, including genes involved in proliferation,
248 translation, and cell cycle progression. We identified 35 genes that were
249 hypomethylated with aging and had close by an ARNT motif in all six cell types
250 (Supplementary File 10). Ten of these genes (right side of Figure 5C, genes under
251 orange headings) have been linked to hypoxia response (39, 40, 41, 42, 43, 44, 45, 46,
252 47, 48). Similarly, we found 26 genes with probes hypermethylated with age and with
253 REST motif in the vicinity in all six cell types (Supplementary File 10). Four of these
254 (right side of Figure 5C, genes under green heading) are known to be downregulated in

hypoxia (49, 50, 51, 52). These results strongly suggest a link between age-associated DNA methylation and oxygen sensing through putative regulation by transcription factors like *ARNT* and *REST* in the various immune cells.

Association with inflammatory cytokines

Low-grade inflammation has been reported to be part of healthy aging. In order to investigate whether age-related pro-inflammatory state may explain the age-related changes in methylation observed in this study, we analyzed the SomaScan protein data of seven pro-inflammatory cytokines (IL6, IL1RN, IL1A, IL1B, TNF, TNFRSF1A, TNFRSF1B) for the same cohort of donors from Tanaka et.al (53). For each cell type, all CpGs reported as significantly hypo- or hyper-methylated with age (in beta-regression analyses adjusted for sex) were reanalyzed by incorporating data on seven pro-inflammatory cytokines (see Materials and Methods for details). Briefly, by comparing a model with a cytokine as explanatory variable (CpG ~ age + sex + cytokine) with another model without it (CpG ~ age + sex), we explored the robustness of age as an explanatory variable of methylation change, as well as possible mediating effects arising from pro-inflammatory cytokines. Detailed results are provided in (https://osf.io/rxw6h/?view_only=82748aad8a6949baae0437b6d9f47681) and summary statistics for each cell type, hypo/hyper-methylation association and pro-inflammatory cytokine are provided in Supplementary File 11. By comparing the results from the two above mentioned regression models, we observed in case of hypermethylated sites in CD4 cells, the number of CpGs dropped by 10% on adding TNFRSF1A to the model, a cytokine that appears significantly associated with 7507 of those CpGs (Column I). In addition, TNFRSF1B appears significantly associated with 9,058 of the age-

hypermethylated CpG sites in CD4 cells. For other cell types like B naïve and monocytes, TNF-alpha was associated with 65-124 age-associated CpG sites respectively. Fewer associations are observed for the remaining analytes. These results suggest a possible link between TNF-alpha signaling pathway, aging and DNA methylation change in circulating immune cells.

DISCUSSION

Novel and important conclusions arise from our observations. First, only few CpG sites are hypomethylated and hypermethylated with aging across all circulating cells while majority of the significant age-associated methylation changes are cell-selective. Indeed, several CpGs show differential age-methylation in opposite directions in different cell types and are unchanged in PBMC, suggesting that they may be missed when studying mixed cell samples. Noteworthy, age-related methylation differences in this cross-sectional study were strongly and significantly correlated with longitudinal age-associated methylation changes in an independent population.

Second, age-associated hypomethylated sites were significantly enriched for active enhancers whereas age-hypermethylated sites were enriched for bivalent/polycomb regions, confirming previous findings in whole blood (34). Age-differential methylation coincided with specific chromatin status and histone markers patterns, suggesting that their position in proximity of promoter and active enhancer regions is connected with chromatic accessibility and potentially modulation of gene expression. Since the ENCODE data was only from 2 donors, it will be worthwhile to see how the histone or chromatin accessibility patterns change with age at and around these age-associated CpG sites.

Third, distinct TF binding motifs co-localize with CpGs differentially methylated with aging despite wide variation in the distribution of such sites across cell types, suggesting a specific regulatory function. Noteworthy, the top age-associated TF identified, *ARNT* and *REST* act in coordination in hypoxia response (54). *BCL6*, another top TF binding motif associated with age-differentially methylated CpG has also been shown to protects cardiomyocyte from damage during hypoxia (55). These findings support the hypothesis that systematic methylation changes with aging may be induced by fluctuations in oxygen availability and energy metabolism. Interestingly, the mRNA encoding *ARNT* significantly increases with age in all cell types except monocytes, while mRNA coding for *REST* declines with aging in 4 cell types and shows no significant change in naïve CD8+ T cells and NK cells. mRNAs coding for *CTCF* showed strong age-association across numerous cell types (Supplementary File 8). The hypothesis that oxygen sensing regulates directly or indirectly DNA methylations is consistent with studies showing that in replicating fibroblasts, biological age estimated by DNA methylation slows down under hypoxia compared to normoxia (56). Further, many genes close by to “shared” age-differentially methylated CpG identified in our analyses play important roles in hypoxia response (Figure 5C).

The specific mechanisms connecting age-related changes in DNA methylation in genes which also contain binding motifs the master hypoxia-response mediators remain unknown. Shahrzad al. reported an inverse correlation between the severity of hypoxia and the degree of DNA methylation (57). There is evidence that hypoxia-induced hypermethylation may be due to reduced TETs activity (58). Our findings add to this literature by suggesting that a direct interaction between hypoxia-related transcription

factors and DNA methylation at specific DNA sites occur with aging, perhaps as an adaptive response triggered by fluctuations in oxygen levels that occur in many age-related conditions. This hypothesis is consistent with oxygen availability been the most important environmental factor that requires physiological adaptation during pregnancy and development and extends this concept in a life course perspective.

A limitation of this study is that we have focused on circulating cells and, therefore, our findings may not apply to age-methylation in other tissues. In addition, our findings were not replicated in an independent cross-sectional study population. Despite these limitations, this study has unique features: a cohort of exceptionally healthy donors and percent methylation was assessed in specific cell types obtained by cytapheresis and sorted by using state-of-the art methods.

CONCLUSION

Age-associated DNA methylation profiles of the six purified primary immune cell populations in the blood show more cell-specificity than sharedness. However, we observe common regulatory features with respect to transcription factor binding motifs and histone modifications. Based on the consistent association of these methylated sites with ARNT and REST, which are master hypoxia regulators, we hypothesize that oxygen sensing and hypoxia drive mechanisms for changes in methylation. This hypothesis should be further explored in animal models with manipulation of oxygen levels and serial measures of DNA methylation in circulating immune cells.

MATERIALS AND METHODS

Ethics statement

GESTALT study was approved by the institutional review board of the National Institutes of Health. Informed consent as well as the consent to publish the data collected was obtained from every participant in the study. Since the study of gene expression and epigenetic regulation are essential aims of GESTALT, all participants were required to consent to DNA/RNA testing and storage at all visits in order to participate in the study. The GESTALT IRB approval number is 15-AG-0063.

Cohort details

Buffy coat, peripheral blood cells (PBMC) and granulocytes were collected from Genetic and Epigenetic Signatures of Translational Aging Laboratory Testing study (GESTALT) study participants (N=55; 34 men and 21 women; age 22-83 years) who were free of diseases (except controlled hypertension or history of cancer silent for > 10 years), not on medications (except one antihypertensive drug), had no physical or cognitive impairments, non-smokers, weighed > 110 lbs, had BMI < 30 kg/m^2 (59) (60). GESTALT was approved by the institutional review board of the National Institutes of Health and participants explicitly consented to participate.

Isolation of PBMC and immune cell populations

PBMCs were isolated from cytapheresis packs by density gradient centrifugation using Ficoll-Paque Plus. Total B, $CD4^+$ and $CD8^+$ T cells were enriched by negative selection using EasySep Negative Human kits specific for each cell type; monocytes were negatively enriched using “EasySep Human Monocyte Enrichment Kit w/o CD16 depletion”. Natural killer cells were negatively enriched by depleting PBMCs with antibodies against CD3, CD4, CD14, CD19 and Glycophorin-A in HBSS buffer.

Enriched cell populations were FACS sorted by flow cytometry as per Human Immunophenotyping Consortium (HIPC) phenotyping panels (61). Gating strategies and post-sort purity were analyzed by FlowJo software (LLC, Ashland, OR)(59). Granulocytes were positively selected from whole blood using EasySep™ Human Whole Blood CD66b Positive Selection Kit. Purified cells and PBMC were washed with PBS, snap frozen and stored at -80° C. All sorted cells were >95% pure by flow cytometry (59).

Assessment of DNA methylation

DNA was isolated from 1–2 million cells using DNAQuik DNA Extraction protocol and the Qiagen DNeasy Kit. 300 ng of DNA was treated with sodium bisulfite using Zymo EZ-96 DNA Methylation Kit. The methylation of ~850,000 CpG sites was determined using Illumina Human MethylationEPIC BeadChip, and data preanalyzed by GenomeStudio 2011.1.

Data processing and functional annotation of CpG sites

Analyses was performed by the R minfi package (62, 63). Probes with low detection p-values (cutoff 0.01) were filtered out(64). Data was normalized using noob and BMIQ(65), batch corrected by ComBat function (sva package), and β values were used for differential methylation analyses. Following the MethylationEPIC probe annotation (IlluminaHumanMethylationEPICanno-.ilm10b2.hg19) to the UCSC RefSeq genes (hg19), we grouped the locations into 3 categories - 1) promoter group- TSS1500 (from 201-1500 bp upstream of TSS), TSS 200 (≤200 bp upstream of TSS), 5'UTR, first exon; 2) genebody- exons (all exons except exon1), exon intron boundary, intron and 3'UTR

and 3) intergenic probes. The first gene in the annotation package was considered. Probes were divided into 3 groups-within CpG islands (CGI), within CpG shore (0-2kb from CGI), CpG shelf (2-4kb from CGI) and open sea (>4kb from CGI).

Definition of age-associated probes

Age and sex adjusted CpG-specific beta regressions were performed on normalized β values using the R *betareg* function. P-values were adjusted for multiple testing (Benjamini-Hochberg (BH) adjusted $p< 0.05$). Probes with FDR $p<0.05$ for age and FDR $p>0.05$ for sex were considered age-differentially methylated CpGs. Beta-regression estimate value was used to group the age-associated probes as hypo- ($Estimate_{age} < 0$) or hypermethylated ($Estimate_{age} >0$). The overlap of probes across multiple combinations of the six cell types was assessed using R package SuperExactTest (v.1.1.0) (66).

Gene Set Enrichment Analysis (GSEA)

Based on the EPICarray annotation, genes were classified as differentially hypo- or hyper-methylated with age. Genes with both age hypo- and hyper-methylated CpGs were removed from the analysis. Enrichment analysis was performed by the tmodHGtest method in the tmod v.0.46.2 R package, comparing a foreground list of genes found in ≥5 cell types against reference gene set collections “Hallmarks” and “Canonical Pathways” (which includes Reactome, KEGG, WikiPathways, PID, and Biocarta gene sets) from the Molecular Signature Database MSigDB (v.7.4) (67).

For the gene enrichment analysis of the principal components, the top 500 CpG probes corresponding to the positive and negative directions along PC1, PC2 and PC3 were

extracted and annotated to nearest gene as per manufacturer's annotation file. The ambiguous genes with probes associated with both positive and negative PC directions were removed from the analysis. The remaining genes were run through the abovementioned enrichment analysis pipeline. A filter based on q-value<0.05 was imposed to find the most significant pathways.

Visualization of histone peaks and DHS peaks

Primary cell DHS and chromatin ChIP-Seq bigwig files were downloaded from ENCODE (https://www.encodeproject.org/) (59). DeepTools was used to visualize DHS and histone peaks in +3kb region surrounding age-associated shared and non-shared methylated sites. For plotting purposes, the order of methylated probes was determined based on descending score of DHS peaks and followed for all histone marks (H3K4me1, H3K4me3 and H3K27ac).

Annotation of age-associated methylated probes using chromHMM

The 18-state chromHMM models (based on 6 chromatin marks H3K4me3, H3K4me1, H3K36me3, H3K27me3, H3K9me3 and H3K27ac) for various immune cells (E032- primary B cell, E038- primary naïve $CD4^+$ T cells, E047- primary naïve $CD8^+$ T cells, E029- monocyte, E046- NK cell) were downloaded from Roadmap epigenomics project (https://egg2.wustl.edu/roadmap/web_portal/chr_state_ learning.html). Bedops tool was used to map the age-associated methylated sites to the respective chromHMM profiles. All Infinium MethylationEPIC array probes were also partitioned using each of the immune cell chromHMM profiles as controls.

Prediction of de-novo transcription factor binding motifs by HOMER

All the age-associated methylation sites were considered for HOMER analysis. A region of ±200bp around each age-associated methylated site was provided as input for analysis in HOMER using de novo setting (68). As a background we used the default background list that HOMER creates by matching the GC% in the input list. The output from the stringent de-novo analysis was considered for downstream data interpretation.

InCHIANTI longitudinal study cohort

InCHIANTI (Invecchiare in Chianti) is a population-based cohort of individuals ≥ 20 years old from the Chianti region of Tuscany, Italy (PMID: 11129752). The Italian National Institute of Research and Care on Aging Institutional Review Board approved the study protocol and all participants explicitly consented to participate. DNA methylation from 699 participants (1841 observations) was used for the analysis. CpG methylation of 485,577 CpGs was determined by the Illumina Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, CA) and data processed by the R package "sesame". Mean rates of change were estimated from 2-3 longitudinal timepoints.

RNA-Seq sample extraction, processing and data analysis

Total RNA was extracted from 2 x10^6 cells, depleted from ribosomal RNA and 50ng was used for cDNA synthesis and library preparation. Libraries were sequenced for 138 cycles on Illumina HiSeq 2500. After adapter removal and end trimming of raw FASTQ files, transcript abundances were quantified with reference to hg19 transcriptome using kallisto 0.44 (with options --single -l 250 -s 25). Transcripts were aggregated to genes with tximport and filtered out if less than 10TPM were detected in more than 33% of the

456 samples. Linear regression models (~ phase + age*sex) were used on TPM normalized
457 expression values to study expression changes of selected transcription factors with
458 age. Only the regression coefficient and p value for the 3 transcription factor genes-
459 *ARNT*, *CTCF* and *REST* were used in this study.

460 Inflammatory cytokine analysis

461 Published SomaScan protein data from the same cohort of donors used in the present
462 study was extracted to look for age-associated changes in seven cytokines (IL6, IL1RN,
463 IL1A, IL1B, TNF, TNFRSF1A, TNFRSF1B) (53). Briefly, plasma proteomics was
464 measured using the 1.3k SomaScan assay (SomaLogic, Boulder, CO) followed by
465 standard quality control and normalization procedures as described in previous
466 publications (69, 70). Normalized data for seven cytokines were extracted (detailed
467 annotation provided in Table S10). To complement the age-association analysis of
468 CpGs adjusted by sex (CpG ~ age + sex), we performed additional beta regression
469 analyses separately including each target pro-inflammatory cytokine as an explanatory
470 variable in the form: CpG ~ age + sex + cytokine. Details about the cytokines are
471 provided in https://osf.io/rxw6h/?view_only=82748aad8a6949baae0437b6d9f47681.

472 Data sharing statement

473 Microarray data are available at GEO under accession number GSE184269.

474 Code sharing

475 Code and data processing scripts are available from the corresponding author upon
476 request.

FIGURE LEGENDS

Figure 1: Study design and identification of age-associated methylation probes A) Study design. B) Age-associated CpG methylation (FDR p<0.05) in 6 cell types. C-D) SuperExactTest circular plots to show the number of age-associated hypo- and hyper-methylated probes shared among different combinations of cell types (indicated by green boxes), respectively. The outermost bars show the number of probes shared among each cell type combination (regardless of other cell types). For example, probes hypomethylated with age in B + CD4 + CD8 + gran + mono (n=222) includes probes also hypomethylated in NK cells (n=181) and probes not hypomethylated with age in NK cells (n=41). Based on the exact probability distributions of multi-set intersections, all the overlaps shown are highly statistically significant ($p<10^{-100}$). E) Graphical representation of age-associated hypomethylation in promoter region of RCAN1 in all 6 cell types. F) Graphical representation of age-associated hypermethylation in promoter region of KLF14. The methylation status in PBMC and buffy coat are also shown. Missing methylation data is represented in white.

Figure 2: Characteristics of age-associated probes. A-B) Manhattan plot of age-associated hypo- and hypermethylated CpG sites in B cells respectively. Most significant genic probes (-log padj >10) are labelled. C) Correlation between beta-regression coefficients of age-differentially methylated CPGs in GESTALT and longitudinal InCHIANTI study. X-axis- InCHIANTI, Y-axis- B cell (Figure 2C) and $CD4^+$ T cell coefficients (Figure 2D). Blue dots - age-hypomethylated CpGs, yellow triangles- age-hypermethylated CpGs. E and F) Scatter plot of age-associated CpGs showing

opposite trends in different immune cells. E) cg27123256 (in BCL11B promoter) is hypomethylated with older age in B, monocytes and NK while is hypermethylated with older age in $CD4^+$ T cells. F) cg03530364 (in FAM19A1 promoter) is hypermethylated with older age in B, granulocytes, monocytes and NK cells while it is hypomethylated with older age in $CD4^+$ T cells.

Figure 3: Pathway analysis of methylated probes. Enrichment analysis of genes annotated to age-associated hypo- and hyper-methylated CpGs in ≥5 cell types (left-most column) and in individual cell types. Red/green shades indicate enrichment scores in hyper- (red) and hypo- (green) methylated genes. Yellow indicates ambiguous pathways associated with both hypo- and hyper-methylated genes in individual cell types. Not significant pathways are shown in grey. Full results in Supplementary File 5.

Figure 4: Functional annotation of age-associated probes along with their grouping based on sharedness. A) ChromHMM annotation of age-associated CpGs. B) Proportion of CpGs mapping to weak/active enhancers (left, orange box), bivalent enhancers/TSS (inset, brown box) and polycomb repressor regions (right, grey box) in age-associated hypo- (blue line), hypermethylated (red line) CpGs as compared to all MethylationEPIC CpGs (grey line). C) DeepTools plots showing the distribution of accessible chromatin (DNase hypersensitive sites) and H3K4me1 histone mark in and around ± 3kb region of age-differentially methylated CpGs. The age-associated sites were divided into shared (blue) (common between 5 or more immune cells) and selective sites (green). The top row shows the pattern for age-associated hypomethylated CpGs while the bottom row is for the age-associated hypermethylated CpGs in B and $CD4^+$ T cells.

Figure 5: Association of transcription factor binding motifs with age-differentially methylated CpGs. A) Top 5 TF motifs at and around (± 200bp) of CpG sites that are hypomethylated with age. All the age-hypomethylated sites were considered for the analysis in each cell type. Recurring motifs like ARNT and CTCF/BORIS are highlighted. B) Top 5 TF motifs at and around (± 200bp) CpG sites that are hypermethylated with age. All the age-hypermethylated sites were considered for the analysis in each cell type. Recurring motifs like REST and Sp100 are highlighted. C) Hypoxia-centric model of age-associated sites with ARNT and REST motifs. CpG sites hypomethylated with aging across 6 different cell types are significantly more likely to host binding motifs for ARNT, the core hub for the hypoxia response. On the contrary, CpG sites hypermethylated with aging are significantly more likely to host binding motifs for REST, a hypoxia response transcriptional repressor. On the right are selected age-associated genes that carry the motifs for ARNT or REST transcription factors.

LIST OF TABLES

Supplementary File 1: Demographic and flow cytometry marker details of the cohort. Details of the age and sex distribution of the healthy donors from the GESTALT study for each of the primary immune cell type population are described. The flow cytometry markers for cell selection are also mentioned.

Supplementary File 2: Pathway enrichment analysis of genes annotated to top 500 probes corresponding to PC1, PC2 and PC3 components of Principal Component Analysis (PCA). The top 500 CpG sites corresponding to PC1, PC2 and PC3 components were annotated to genes followed by gene enrichment analysis. Age-associated genes in each pathway are in column M.

Supplementary File 3: Distribution of slope for probes significantly changing with age in the immune cells. The age-associated probes were identified from beta regression (FDR p<0.05).

Supplementary File 4: List of age-associated probes each of the six primary immune cells. Beta-regression coefficient, FDR p value and genomic annotation of the age-associated probes were identified from beta regression (FDR p<0.05).

Supplementary File 5: List of top age-associated genes in the six immune cell types. The list of top 15 and top 50 age-associated hypo- and hypermethylated genes derived from the most significant age-associated probes in each cell type.

Supplementary File 6: Detailed output of Gene Set Enrichment Analysis. Gene Set Enrichment Analysis was performed on genes based on annotation of age-associated hypo- and hypermethylation probes commonly changing in 5 or more cell types.

Supplementary File 7: Top 5 transcription factor (TF) motifs within ± 50bp of age-associated methylated sites. HOMER de-novo analysis was performed to identify the top 5 TF motifs within 50bp of age-associated hypo- and hypermethylated sites in each of the six cell types.

Supplementary File 8: Average read depths and kallisto TPM normalized read counts of ARNT, CTCF and REST for all the donors. RNASeq data was used to look into the gene expression change of three selected transcription factors (ARNT, CTCF and REST) with age. These transcription factor motifs are most commonly associated with the age-related methylated sites in all immune cells.The mapping rates along with

the Kallisto TPM normalized values for the 3 TFs for each cell type in each of the donors has been shown.

Supplementary File 9: Age-associated differences of transcripts for ARNT, REST and CTCF. FDR p-values derived from the linear regression of expression levels of the 3 TFs with age in each of the 6 cell types.

Supplementary File 10: List of genes with age-associated methylated CpG sites showing ARNT or REST motif within 1kb. The age-associated probes with ARNT or REST motifs within 1kb region were annotated to genes and summarized into a table. For each gene, number of age-associated CpG sites with ARNT/REST motif and number of cell types in which this occurrence has been observed have been mentioned.

Supplementary File 11: Output of beta regression analysis with age and sex and the seven analytes. Summary of two beta regression models have been tabularized. Column C shows the number of age-associated probes from the original model CpG~age+sex with FDR cut-off of adjusted p_{age} <0.05. Columns D-J shows the number of age-associated probes from the model CpG~age+sex+analyte with FDR cut-off of adjusted p_{age} <0.05 and adjusted $p_{analyte}$ <0.05. Finally, columns K-Q represent the number of age-associated probes from the model CpG~age+sex+analyte with FDR cut-off of adjusted $p_{analyte}$ <0.05.

SUPPLEMENTARY FIGURE LEGENDS

Figure 1- figure supplement 1: Characteristics of entire dataset and age-associated methylation data in 6 primary immune cells. A) PCA plot of normalized methylation data of six immune cells in the 55 healthy donors. The cell types are

indicated in different colors, while the three broad age groups (20-40, 40-60 and 60-90 years) are indicated in different shapes (PC1- principal component 1, PC2- principal component 2). B) Distribution of beta-regression coefficient of the age-associated hypo- and hypermethylated probes in all six immune cell types estimated this cross-sectional study. Distribution of coefficient values categorized into groups is shown in Supplementary File 2. C). Age-associated probes from independent sample t-test analysis of young (<=35years, 25th percentile) vs old (>=70years, 75th percentile) donors for each cell type. The pie charts on top show the extent of overlap with results from the beta-regression analysis. D) Distribution of age-associated probes from beta regression into groups based on distance from CpG islands (CGI) (Island- within CGI, shore- within 2kb of CGI, shelf- 2-4kb of CGI, open sea- >4kb from CGI). E) Distribution of age-associated hypo- and hypermethylated probes with respect to location (promoter-1500TSS to 1st exon, genebody-within exons, introns and 3'UTR). F and G) Overlap of age-associated hypo- and hypermethylated probes in the 6 immune cell types with those identified in PBMCs. The first bar indicates the number of age-associated probes identified in PBMC. The following bars show the counts in the other immune cells, the lighter portion of the bars show the number of probes that are shared with PBMCs, and the darker portion indicates non-PBMC cell-specific probes. The numbers above the bars show the log of odd's ratio and its 95% confidence interval to represent the significance of the overlap between age-associated probes in each cell type with PBMC. The p value from odd's ratio analysis is represented as ** as they were all <0.01.

Figure 2- figure supplement 1: Most significant age-associated CpGs in non-B immune cells along with CpGs showing opposite age-associated trends. A) Manhattan plot of age-associated CpGs in $CD4^+$ T cells. The X-axis shows the distribution of significant CpGs (FDR $p<0.05$), and the Y-axis shows the associated negative log of the adjusted p value from beta regression. Positive axis comprises of probes hypermethylated with age while the negative axis shows age-associated hypomethylated probes. The top hits in each group with the most significant p-values are labelled where the orange dot present CpG probes in the gene promoter. B) Manhattan plot of age-associated probes in $CD8^+$ T cells. C) Manhattan plot of age-associated probes in granulocytes. D) Manhattan plot of age-associated probes in NK cells. E) Manhattan plot of age-associated probes in monocytes. F) Count of probes hypomethylated with age in cell type of interest but showing hypermethylation in one or more other cell types. For example, 315 probes are age-hypomethylated in B cells but are significantly hypermethylated with age in one or more other immune cell types. Maximum number of such probes are observed in $CD4^+$ T cells followed by B cells and monocytes. G) Count of probes hypermethylated with age in cell type of interest but showing hypomethylation in one or more other cell types. For example, 282 probes are hypermethylated in B cells but are significantly hypomethylated with age in one or more other immune cell types. Maximum number of such probes are observed in $CD4^+$ T cells followed by B cells and monocytes.

Figure 2- figure supplement 2: Comparison of individual immune cells with InCHIANTI longitudinal study. A-B) Correlation between beta-regression coefficients of age-associated methylation probes in 5 or more cell types in study and beta-

regression coefficients estimated from longitudinal data in the InCHIANTI study. On the X-axis is the data from InCHIANTI longitudinal study cohort while on the Y-axis is cell-specific coefficient values for $CD8^+$ T cells (top left), granulocytes (top right), monocytes (bottom left) and NK cells (bottom right). pink dots are the coefficients of the age-hypomethylated probes (Supplementary Figure 3A) while the blue dots are for age-hypermethylated probes (Supplementary Figure 3B).

Figure 4- figure supplement 1: Functional annotation of age-associated probes with respect to DHS and 3 other histone marks from ENCODE. Rows1-2-The functional annotation of age-associated probes in CD8, monocytes and NK cells with respect to DNase hypersensitivity sites and H3K4me1 peaks. Rows3-6- The functional annotation of age-associated probes in B, CD4, CD8, monocytes and NK cells with respect to H3K4me3 and H3K27ac peaks. Age-associated hypo- or hypermethylated probes were grouped into shared (blue) and selective (green) based on whether they were common across 5 or more of the 6 cell types or not. A region of 3kb of either side of the CpG probes of interest was examined. The DHS and histone bigwig files for the 5 primary cells were obtained from ENCODE. Granulocyte data was not available and hence could not be examined.

Figure 5- figure supplement 1: Count of age-associated hypo- or hypermethylated probes with ARNT or REST motifs within 1kb respectively. ARNT and REST were the top TF motifs associated respectively with all hypo- and hypermethylated age-associated probes in most cell types. To verify whether the same set of probes were present in each cell type with the ARNT(A) or REST(B) motifs, HOMER was used to find the probes that have an ARNT or REST motif within $\pm$ 500bp. The SuperExact test

based circular plots shows the overlap between the different cell types (relevant cell types are indicated in green boxes in the inner circles for each combination). The numbers above the outermost bars indicate the count of probes that have ARNT or REST motifs across various combinations of cells while the color of the outermost bars in the plot indicate the log transformed p values obtained from hypergeometric test to check whether the overlap is significant or not.

Acknowledgements

This work was supported entirely by the Intramural Research Program of the National Institute on Aging. We are grateful to the GESTALT participants and the GESTALT Study Team at Harbor Hospital and NIA.

Conflict of interest disclosure

The authors have no conflict of interest to disclose.

REFERENCES

1. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet. 2018;19(6):371-84.
2. 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. Mol Cell. 2013;49(2):359-67.
3. Bocklandt S, Lin W, Sehl ME, Sanchez FJ, Sinsheimer JS, Horvath S, et al. Epigenetic predictor of age. PLoS One. 2011;6(6):e14821.
4. Gale CR, Marioni RE, Harris SE, Starr JM, Deary IJ. DNA methylation and the epigenetic clock in relation to physical frailty in older people: the Lothian Birth Cohort 1936. Clin Epigenetics. 2018;10(1):101.
5. McCartney DL, Stevenson AJ, Walker RM, Gibson J, Morris SW, Campbell A, et al. Investigating the relationship between DNA methylation age acceleration and risk factors for Alzheimer's disease. Alzheimers Dement (Amst). 2018;10:429-37.

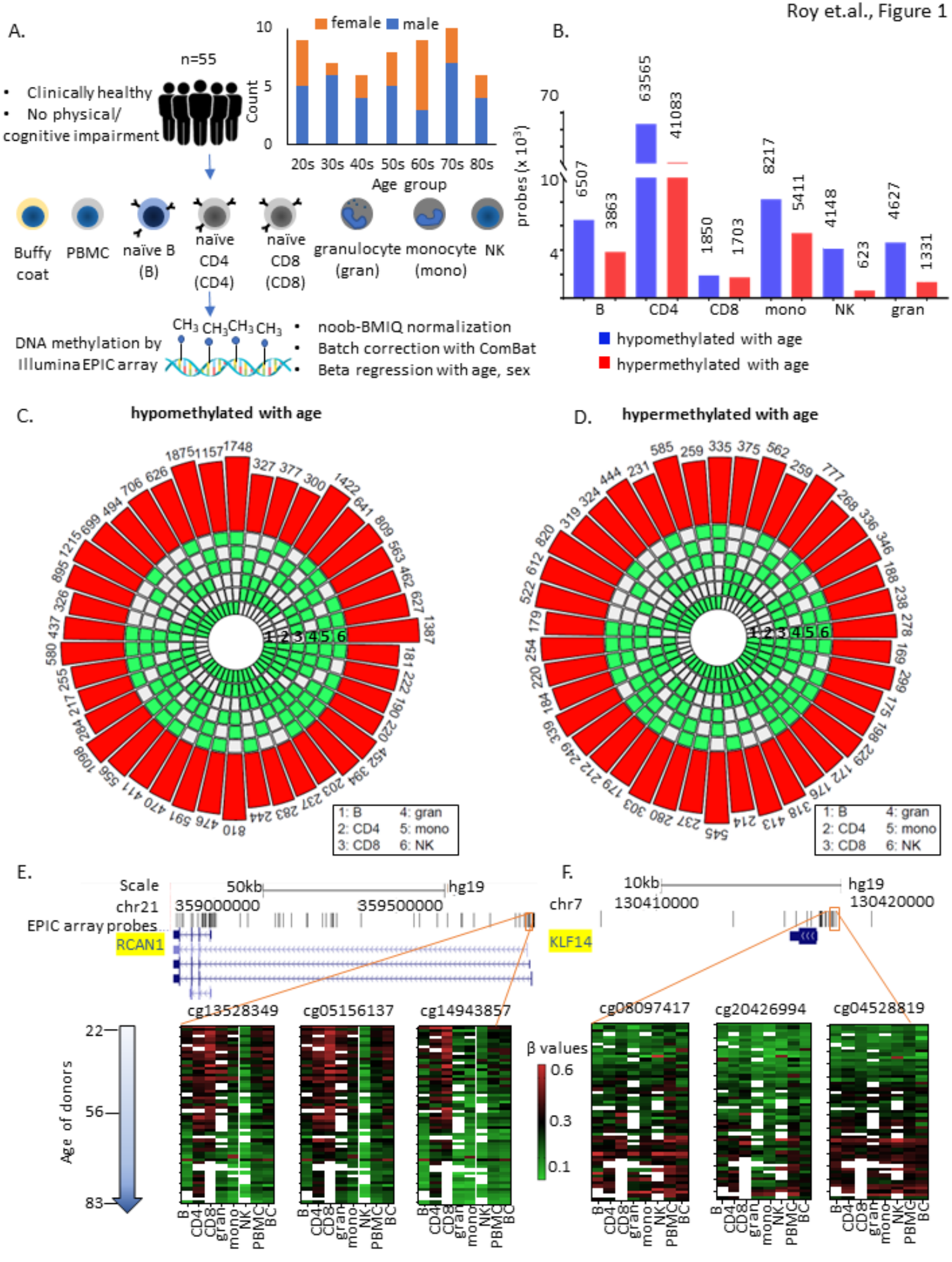
6. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol. 2015;16:25.
7. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging (Albany NY). 2016;8(9):1844-65.
8. Teschendorff AE, West J, Beck S. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? Hum Mol Genet. 2013;22(R1):R7-R15.
9. Martin-Herranz DE, Aref-Eshghi E, Bonder MJ, Stubbs TM, Choufani S, Weksberg R, et al. Screening for genes that accelerate the epigenetic aging clock in humans reveals a role for the H3K36 methyltransferase NSD1. Genome Biol. 2019;20(1):146.
10. Lu AT, Xue L, Salfati EL, Chen BH, Ferrucci L, Levy D, et al. GWAS of epigenetic aging rates in blood reveals a critical role for TERT. Nat Commun. 2018;9(1):387.
11. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res. 2010;20(4):440-6.
12. Dozmorov MG, Coit P, Maksimowicz-McKinnon K, Sawalha AH. Age-associated DNA methylation changes in naive CD4(+) T cells suggest an evolving autoimmune epigenotype in aging T cells. Epigenomics. 2017;9(4):429-45.
13. Reynolds LM, Taylor JR, Ding J, Lohman K, Johnson C, Siscovick D, et al. Age-related variations in the methylome associated with gene expression in human monocytes and T cells. Nat Commun. 2014;5:5366.
14. Tserel L, Kolde R, Limbach M, Tretyakov K, Kasela S, Kisand K, et al. Age-related profiling of DNA methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. Sci Rep. 2015;5:13107.
15. 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.
16. Kananen L, Marttila S, Nevalainen T, Jylhava J, Mononen N, Kahonen M, et al. Aging-associated DNA methylation changes in middle-aged individuals: the Young Finns study. BMC Genomics. 2016;17:103.
17. Acevedo N, Reinius LE, Vitezic M, Fortino V, Soderhall C, Honkanen H, et al. Age-associated DNA methylation changes in immune genes, histone modifiers and chromatin remodeling factors within 5 years after birth in human blood leukocytes. Clin Epigenetics. 2015;7:34.
18. Marttila S, Kananen L, Hayrynen S, Jylhava J, Nevalainen T, Hervonen A, et al. Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. BMC Genomics. 2015;16:179.
19. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
20. Ferrucci L, Bandinelli S, Benvenuti E, Di Iorio A, Macchi C, Harris TB, et al. Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. J Am Geriatr Soc. 2000;48(12):1618-25.
21. Gray SM, Kaech SM, Staron MM. The interface between transcriptional and epigenetic control of effector and memory CD8(+) T-cell differentiation. Immunol Rev. 2014;261(1):157-68.
22. Yui MA, Rothenberg EV. Developmental gene networks: a triathlon on the course to T cell identity. Nat Rev Immunol. 2014;14(8):529-45.
23. de Oliveira PG, Ramos MLS, Amaro AJ, Dias RA, Vieira SI. Gi/o-Protein Coupled Receptors in the Aging Brain. Front Aging Neurosci. 2019;11:89.

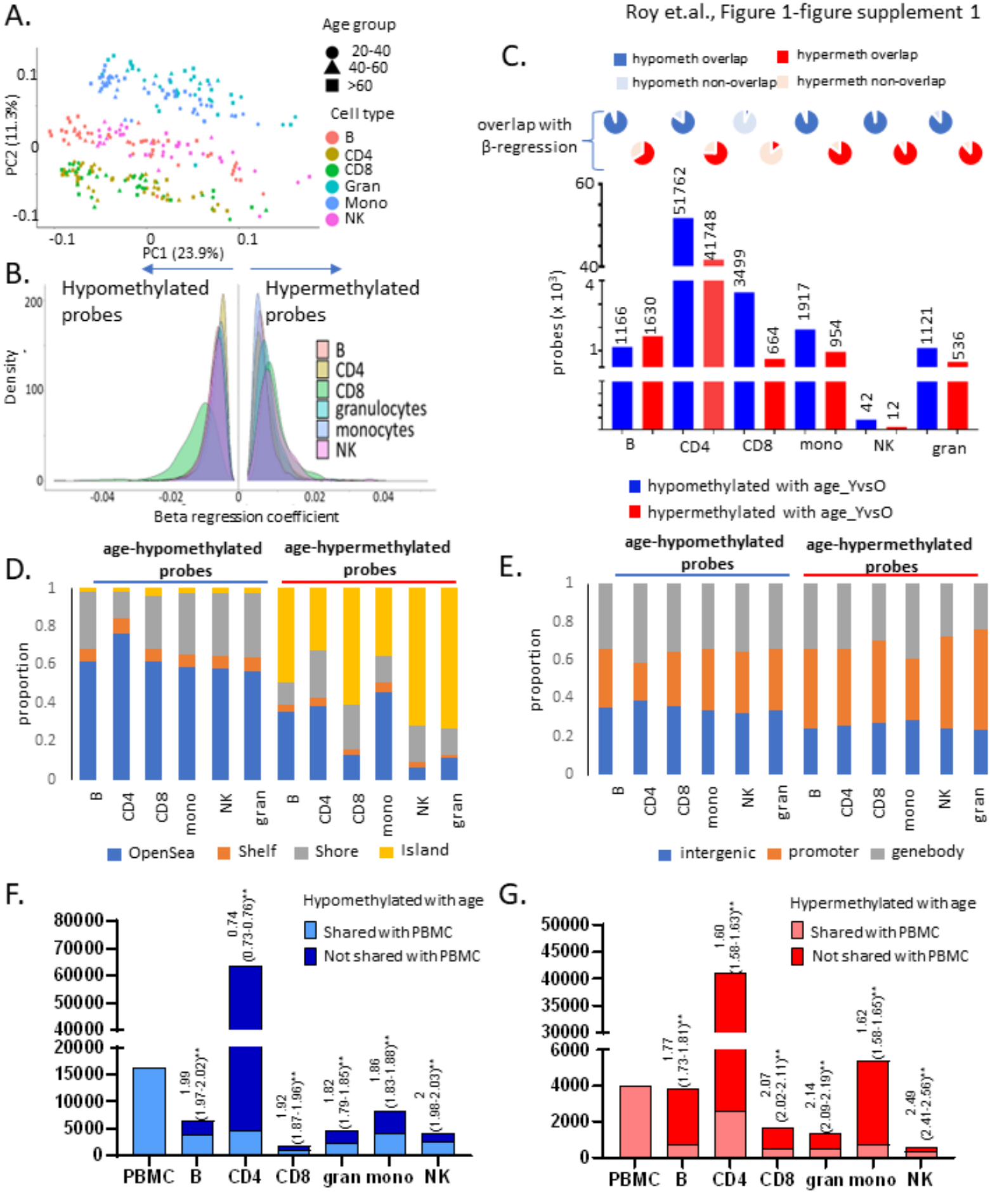
24. Ewing E, Kular L, Fernandes SJ, Karathanasis N, Lagani V, Ruhrmann S, et al. Combining evidence from four immune cell types identifies DNA methylation patterns that implicate functionally distinct pathways during Multiple Sclerosis progression. EBioMedicine. 2019;43:411-23.
25. Karagiannis TT, Dowrey TW, Villacorta-Martin C, Montano M, Reed E, Belkina AC, et al. Multi-modal profiling of peripheral blood cells across the human lifespan reveals distinct immune cell signatures of aging and longevity. EBioMedicine. 2023;90:104514.
26. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 2012;9(3):215-6.
27. Pundhir S, Bagger FO, Lauridsen FB, Rapin N, Porse BT. Peak-valley-peak pattern of histone modifications delineates active regulatory elements and their directionality. Nucleic Acids Res. 2016;44(9):4037-51.
28. Bae S, Lesch BJ. H3K4me1 Distribution Predicts Transcription State and Poising at Promoters. Front Cell Dev Biol. 2020;8:289.
29. Medvedeva YA, Khamis AM, Kulakovskiy IV, Ba-Alawi W, Bhuyan MS, Kawaji H, et al. Effects of cytosine methylation on transcription factor binding sites. BMC Genomics. 2014;15:119.
30. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38(1):23-38.
31. Kim C, Jin J, Weyand CM, Goronzy JJ. The Transcription Factor TCF1 in T Cell Differentiation and Aging. Int J Mol Sci. 2020;21(18).
32. van Ruiten MS, Rowland BD. On the choreography of genome folding: A grand pas de deux of cohesin and CTCF. Curr Opin Cell Biol. 2021;70:84-90.
33. Bhat P, Honson D, Guttman M. Nuclear compartmentalization as a mechanism of quantitative control of gene expression. Nat Rev Mol Cell Biol. 2021;22(10):653-70.
34. Yuan T, Jiao Y, de Jong S, Ophoff RA, Beck S, Teschendorff AE. An integrative multi-scale analysis of the dynamic DNA methylation landscape in aging. PLoS Genet. 2015;11(2):e1004996.
35. Lu T, Aron L, Zullo J, Pan Y, Kim H, Chen Y, et al. REST and stress resistance in ageing and Alzheimer's disease. Nature. 2014;507(7493):448-54.
36. Nyati KK, Zaman MM, Sharma P, Kishimoto T. Arid5a, an RNA-Binding Protein in Immune Regulation: RNA Stability, Inflammation, and Autoimmunity. Trends Immunol. 2020;41(3):255-68.
37. Wilsker D, Patsialou A, Dallas PB, Moran E. ARID proteins: a diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development. Cell Growth Differ. 2002;13(3):95-106.
38. Semenza GL. HIF-1: mediator of physiological and pathophysiological responses to hypoxia. J Appl Physiol (1985). 2000;88(4):1474-80.
39. Craps S, Van Wauwe J, De Moudt S, De Munck D, Leloup AJA, Boeckx B, et al. Prdm16 Supports Arterial Flow Recovery by Maintaining Endothelial Function. Circ Res. 2021;129(1):63-77.
40. Stegmann TJ. FGF-1: a human growth factor in the induction of neoangiogenesis. Expert Opin Investig Drugs. 1998;7(12):2011-5.
41. Chakraborty S, Ain R. Nitric-oxide synthase trafficking inducer is a pleiotropic regulator of endothelial cell function and signaling. J Biol Chem. 2017;292(16):6600-20.
42. Gusdon AM, Zhu J, Van Houten B, Chu CT. ATP13A2 regulates mitochondrial bioenergetics through macroautophagy. Neurobiol Dis. 2012;45(3):962-72.
43. Wang J, Taba Y, Pang J, Yin G, Yan C, Berk BC. GIT1 mediates VEGF-induced podosome formation in endothelial cells: critical role for PLCgamma. Arterioscler Thromb Vasc Biol. 2009;29(2):202-8.
44. Hsu C, Morohashi Y, Yoshimura S, Manrique-Hoyos N, Jung S, Lauterbach MA, et al. Regulation of exosome secretion by Rab35 and its GTPase-activating proteins TBC1D10A-C. J Cell Biol. 2010;189(2):223-32.

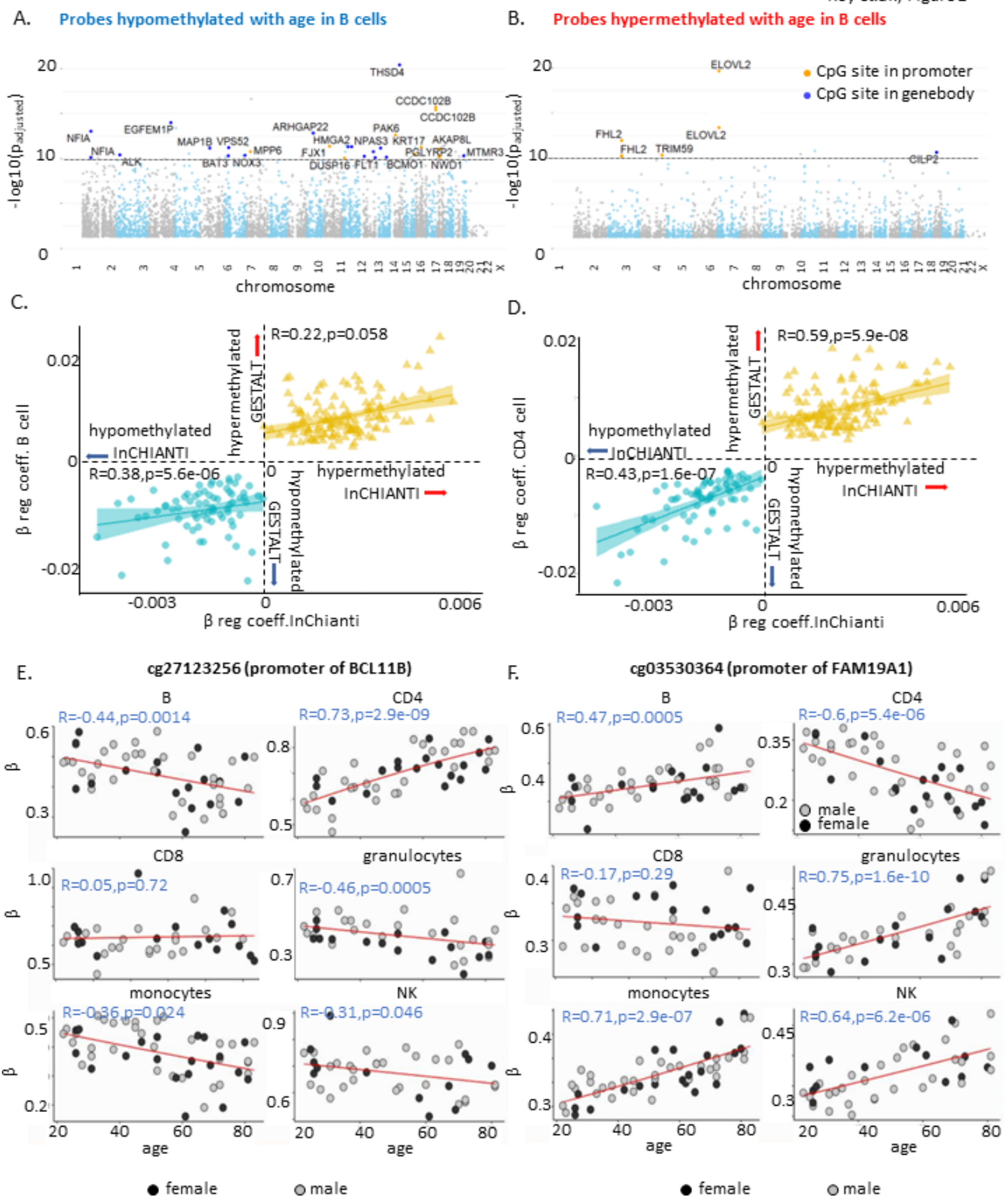
45. Pamenter ME, Hall JE, Tanabe Y, Simonson TS. Cross-Species Insights Into Genomic Adaptations to Hypoxia. Front Genet. 2020;11:743.
46. Pangou E, Befani C, Mylonis I, Samiotaki M, Panayotou G, Simos G, et al. HIF-2alpha phosphorylation by CK1delta promotes erythropoietin secretion in liver cancer cells under hypoxia. J Cell Sci. 2016;129(22):4213-26.
47. Lazarou M, Narendra DP, Jin SM, Tekle E, Banerjee S, Youle RJ. PINK1 drives Parkin self-association and HECT-like E3 activity upstream of mitochondrial binding. J Cell Biol. 2013;200(2):163-72.
48. Cai ZL, Liu C, Yao Q, Xie QW, Hu TT, Wu QQ, et al. The pro-migration and anti-apoptosis effects of HMGA2 in HUVECs stimulated by hypoxia. Cell Cycle. 2020;19(24):3534-45.
49. Tan WS, Lee JJ, Satish RL, Ang ET. Detectability of secretagogin in human erythrocytes. Neurosci Lett. 2012;526(1):59-62.
50. Liu TJ, Yeh YC, Lee WL, Wang LC, Lee HW, Shiu MT, et al. Insulin ameliorates hypoxia-induced autophagy, endoplasmic reticular stress and apoptosis of myocardial cells: In vitro and ex vivo models. Eur J Pharmacol. 2020;880:173125.
51. Dasgupta P. Somatostatin analogues: multiple roles in cellular proliferation, neoplasia, and angiogenesis. Pharmacol Ther. 2004;102(1):61-85.
52. Carmeliet P, Jain RK. Molecular mechanisms and clinical applications of angiogenesis. Nature. 2011;473(7347):298-307.
53. Tanaka T, Biancotto A, Moaddel R, Moore AZ, Gonzalez-Freire M, Aon MA, et al. Plasma proteomic signature of age in healthy humans. Aging Cell. 2018;17(5):e12799.
54. Cavadas MA, Mesnieres M, Crifo B, Manresa MC, Selfridge AC, Keogh CE, et al. REST is a hypoxia-responsive transcriptional repressor. Sci Rep. 2016;6:31355.
55. Gu Y, Luo M, Li Y, Su Z, Wang Y, Chen X, et al. Bcl6 knockdown aggravates hypoxia injury in cardiomyocytes via the P38 pathway. Cell Biol Int. 2019;43(2):108-16.
56. Matsuyama M, WuWong DJ, Horvath S, Matsuyama S. Epigenetic clock analysis of human fibroblasts in vitro: effects of hypoxia, donor age, and expression of hTERT and SV40 largeT. Aging (Albany NY). 2019;11(10):3012-22.
57. Shahrzad S, Bertrand K, Minhas K, Coomber BL. Induction of DNA hypomethylation by tumor hypoxia. Epigenetics. 2007;2(2):119-25.
58. Thienpont B, Steinbacher J, Zhao H, D'Anna F, Kuchnio A, Ploumakis A, et al. Tumour hypoxia causes DNA hypermethylation by reducing TET activity. Nature. 2016;537(7618):63-8.
59. Roy R, Ramamoorthy S, Shapiro BD, Kaileh M, Hernandez D, Sarantopoulou D, et al. DNA methylation signatures reveal that distinct combinations of transcription factors specify human immune cell epigenetic identity. Immunity. 2021.
60. Ubaida-Mohien C, Gonzalez-Freire M, Lyashkov A, Moaddel R, Chia CW, Simonsick EM, et al. Physical Activity Associated Proteomics of Skeletal Muscle: Being Physically Active in Daily Life May Protect Skeletal Muscle From Aging. Front Physiol. 2019;10:312.
61. Maecker HT, McCoy JP, Nussenblatt R. Standardizing immunophenotyping for the Human Immunology Project. Nat Rev Immunol. 2012;12(3):191-200.
62. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-9.
63. Fortin JP, Triche TJ, Jr., Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2017;33(4):558-60.
64. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics. 2016;8(3):389-99.
65. Liu J, Siegmund KD. An evaluation of processing methods for HumanMethylation450 BeadChip data. BMC Genomics. 2016;17:469.

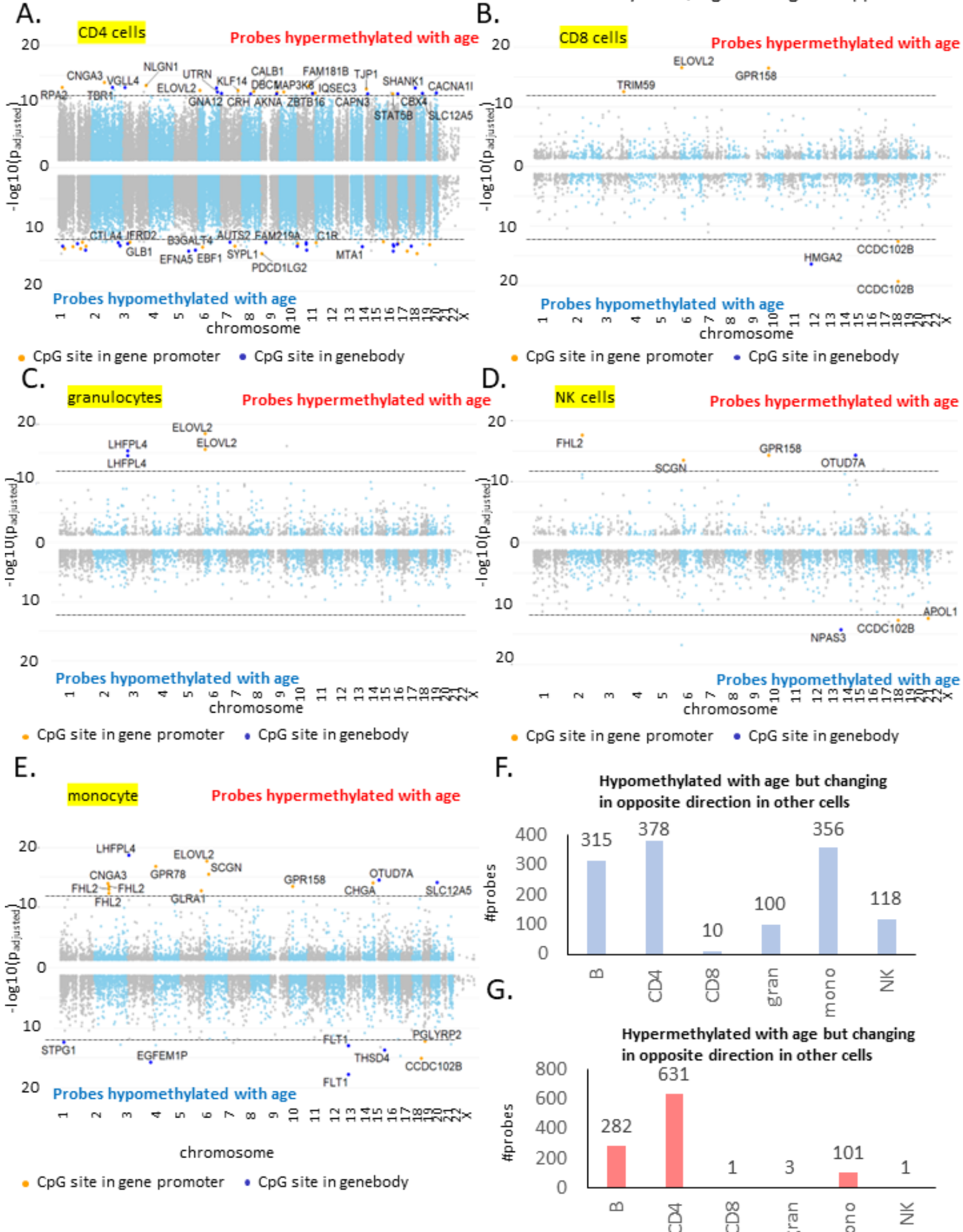
826 66. Wang M, Zhao Y, Zhang B. Efficient Test and Visualization of Multi-Set Intersections. Sci Rep.
827 2015;5:16923.
828 67. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set
829 enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.
830 Proc Natl Acad Sci U S A. 2005;102(43):15545-50.
831 68. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-
832 determining transcription factors prime cis-regulatory elements required for macrophage and B cell
833 identities. Mol Cell. 2010;38(4):576-89.
834 69. Candia J, Cheung F, Kotliarov Y, Fantoni G, Sellers B, Griesman T, et al. Assessment of Variability
835 in the SOMAscan Assay. Sci Rep. 2017;7(1):14248.
836 70. Candia J, Daya GN, Tanaka T, Ferrucci L, Walker KA. Assessment of variability in the plasma 7k
837 SomaScan proteomics assay. Sci Rep. 2022;12(1):17147.

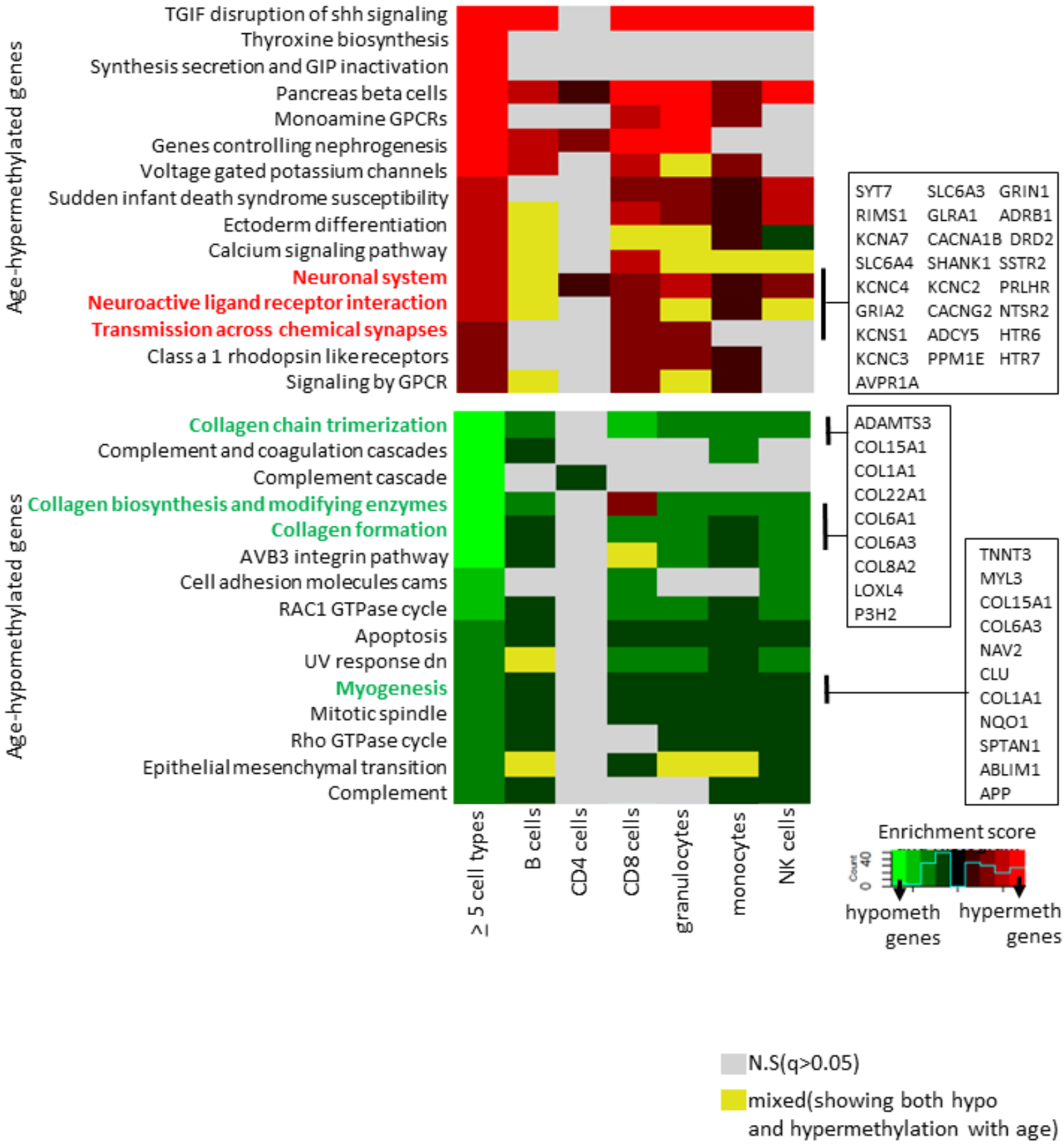
838

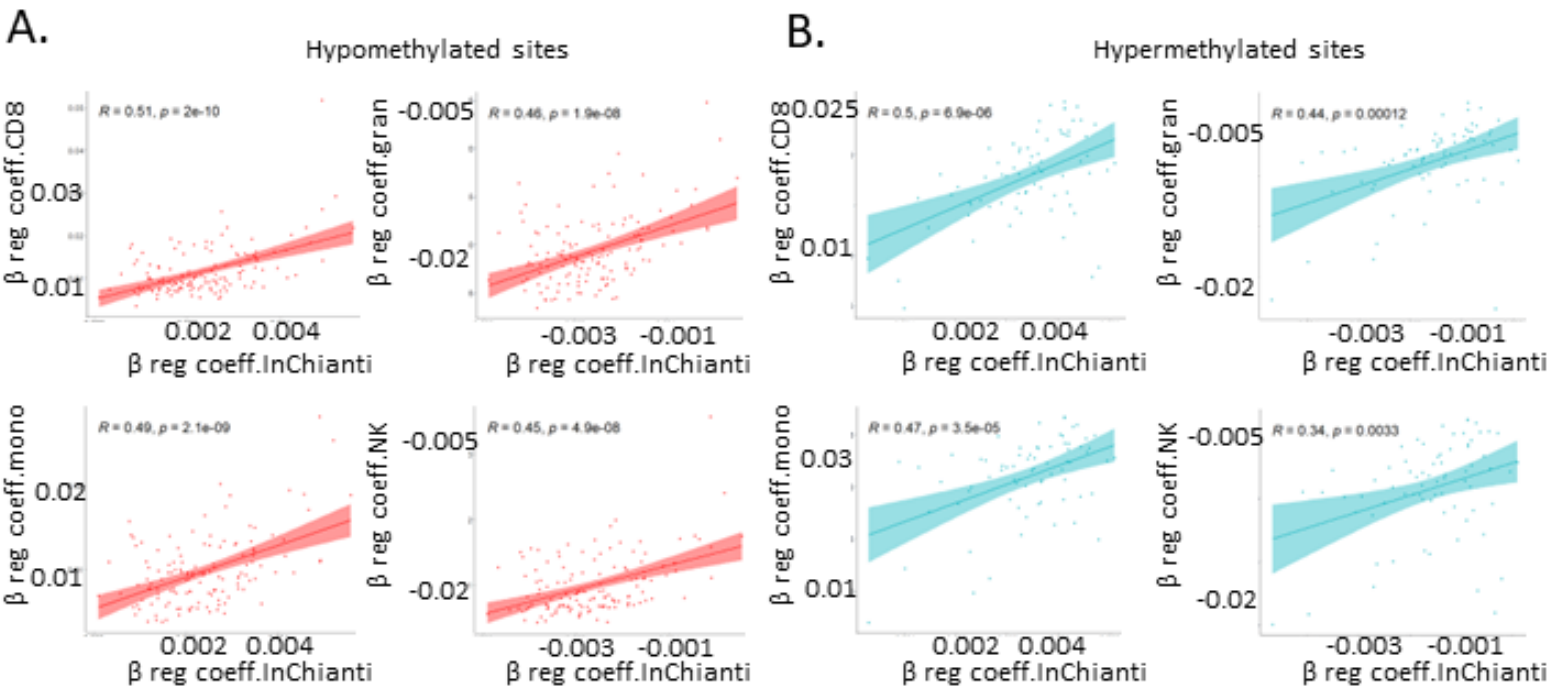

A.
n=55
Clinically healthy
No physical/ cognitive impairment
female
male
Count
Age group
20s 30s 40s 50s 60s 70s 80s
Buffy coat
PBMC
naïve B (B)
naïve CD4 (CD4)
naïve CD8 (CD8)
granulocyte (gran)
monocyte (mono)
NK
DNA methylation by Illumina EPIC array
noob-BMIQ normalization
Batch correction with ComBat
Beta regression with age, sex
B.
probes (x 10^3)
B CD4 CD8 mono NK gran
hypomethylated with age
hypermethylated with age
C.
hypomethylated with age
D.
hypermethylated with age
1: B 4: gran
2: CD4 5: mono
3: CD8 6: NK
E.
Scale
50kb
hg19
chr21
359000000
359500000
EPIC array probes
RCAN1
cg13528349
cg05156137
cg14943857
Age of donors
22
56
83
β values
0.6
0.3
0.1
F.
10kb
hg19
chr7
130410000
130420000
KLF14
cg08097417
cg20426994
cg04528819
B CD4 CD8 gran mono NK PBMC BC

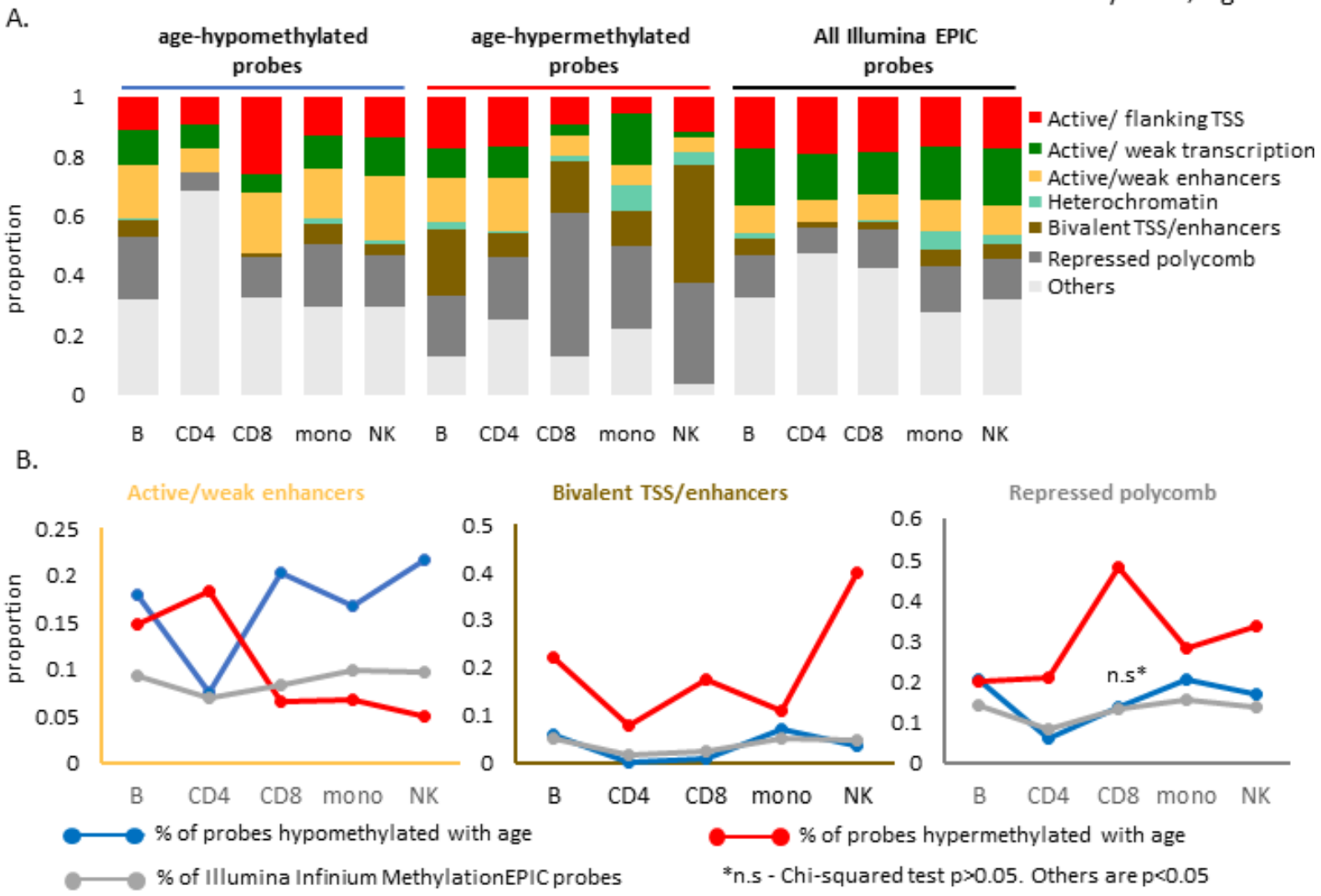

A.
Age group
20-40
40-60
>60
Cell type
B
CD4
CD8
Gran
Mono
NK
PC2 (11.3%)
PC1 (23.9%)
B.
Hypomethylated probes
Hypermethylated probes
Density
Beta regression coefficient
granulocytes
monocytes
C.
hypometh overlap
hypermeth overlap
hypometh non-overlap
hypermeth non-overlap
overlap with β-regression
probes (x 10³)
1166
1630
51762
41748
3499
664
1917
954
42
12
1121
536
B
CD4
CD8
mono
NK
gran
hypomethylated with age_YvsO
hypermethylated with age_YvsO
D.
age-hypomethylated probes
age-hypermethylated probes
proportion
OpenSea
Shelf
Shore
Island
E.
intergenic
promoter
genebody
F.
Hypomethylated with age
Shared with PBMC
Not shared with PBMC
0.74 (0.73-0.76)**
1.99 (1.97-2.02)**
1.92 (1.87-1.96)**
1.82 (1.79-1.85)**
1.86 (1.83-1.88)**
2 (1.98-2.03)**
PBMC
G.
Hypermethylated with age
1.60 (1.58-1.63)**
1.77 (1.73-1.81)**
2.07 (2.02-2.11)**
2.14 (2.09-2.19)**
1.62 (1.58-1.65)**
2.49 (2.41-2.56)**

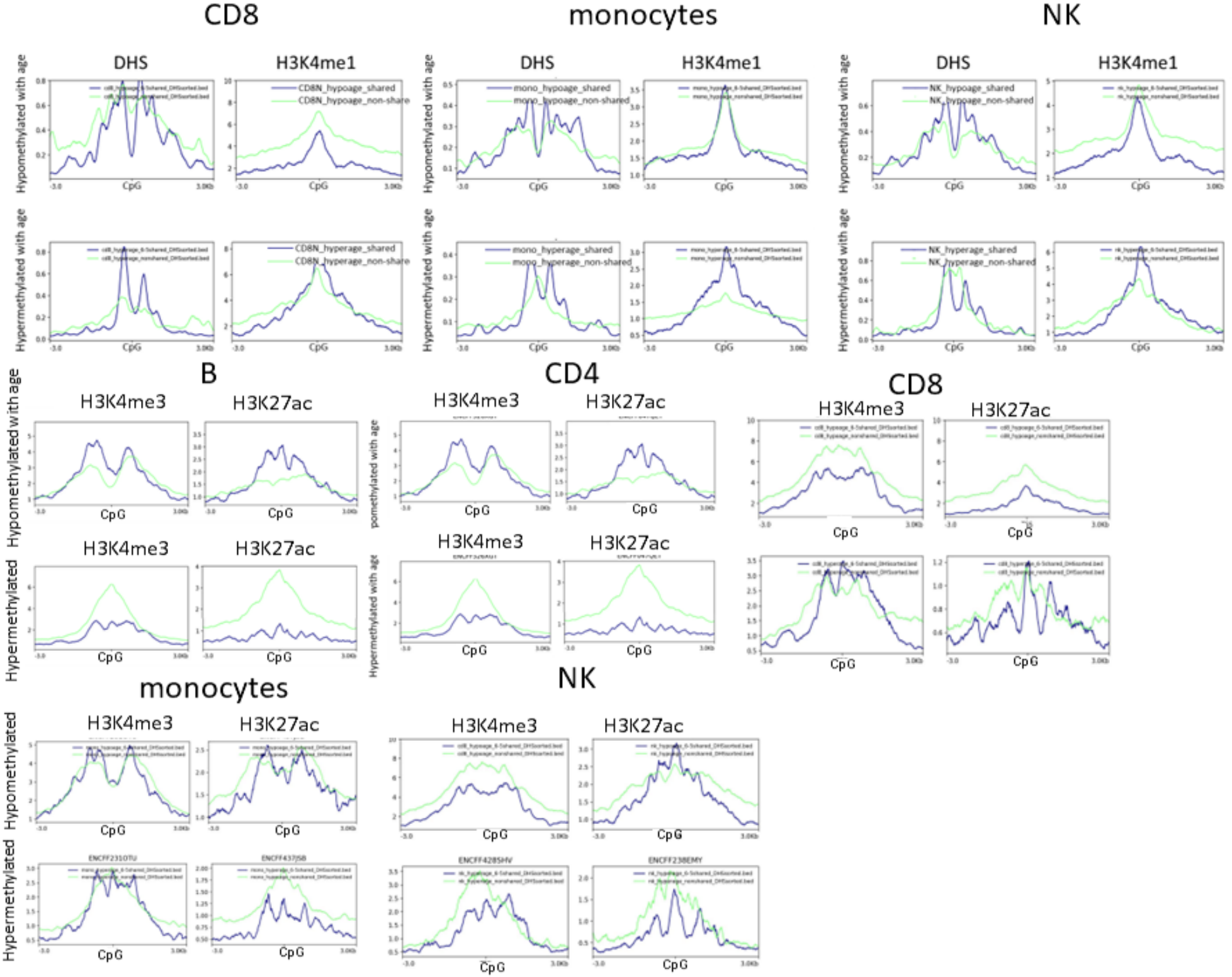

A.
Probes hypomethylated with age in B cells
B.
Probes hypermethylated with age in B cells
CpG site in promoter
CpG site in genebody
-log10(padjusted)
chromosome
THSD4
CCDC102B
EGFEM1P
NFIA
MAP1B
VPS52
ARHGAP22
PAK6
HMGA2
NPAS3
KRT17
AKAP8L
MPP6
FJX1
PGLYRP2
MTMR3
ALK
BAT3
NOX3
DUSP16
FLT1
BCMO1
NWD1
ELOVL2
FHL2
TRIM59
CILP2
C.
D.
β reg coeff. B cell
β reg coeff. CD4 cell
β reg coeff.InChianti
hypomethylated
hypermethylated
InCHIANTI
GESTALT
R=0.22,p=0.058
R=0.38,p=5.6e-06
R=0.59,p=5.9e-08
R=0.43,p=1.6e-07
E.
cg27123256 (promoter of BCL11B)
F.
cg03530364 (promoter of FAM19A1)
B
CD4
CD8
granulocytes
monocytes
NK
R=-0.44,p=0.0014
R=0.73,p=2.9e-09
R=0.05,p=0.72
R=-0.46,p=0.0005
R=-0.36,p=0.024
R=-0.31,p=0.046
R=0.47,p=0.0005
R=-0.6,p=5.4e-06
R=-0.17,p=0.29
R=0.75,p=1.6e-10
R=0.71,p=2.9e-07
R=0.64,p=6.2e-06
β
age
male
female


A.
CD4 cells
Probes hypermethylated with age
Probes hypomethylated with age
-log10(padjusted)
chromosome
CpG site in gene promoter
CpG site in genebody
B.
CD8 cells
Probes hypermethylated with age
Probes hypomethylated with age
-log10(padjusted)
chromosome
CpG site in gene promoter
CpG site in genebody
C.
granulocytes
Probes hypermethylated with age
Probes hypomethylated with age
-log10(padjusted)
chromosome
CpG site in gene promoter
CpG site in genebody
D.
NK cells
Probes hypermethylated with age
Probes hypomethylated with age
-log10(padjusted)
chromosome
CpG site in gene promoter
CpG site in genebody
E.
monocyte
Probes hypermethylated with age
Probes hypomethylated with age
-log10(padjusted)
chromosome
CpG site in gene promoter
CpG site in genebody
F.
Hypomethylated with age but changing in opposite direction in other cells
#probes
315
378
10
100
356
118
B
CD4
CD8
gran
mono
NK
G.
Hypermethylated with age but changing in opposite direction in other cells
#probes
282
631
1
3
101
1
B
CD4
CD8
gran
mono
NK


Age-hypermethylated genes
TGIF disruption of shh signaling
Thyroxine biosynthesis
Synthesis secretion and GIP inactivation
Pancreas beta cells
Monoamine GPCRs
Genes controlling nephrogenesis
Voltage gated potassium channels
Sudden infant death syndrome susceptibility
Ectoderm differentiation
Calcium signaling pathway
Neuronal system
Neuroactive ligand receptor interaction
Transmission across chemical synapses
Class a 1 rhodopsin like receptors
Signaling by GPCR
SYT7 SLC6A3 GRIN1
RIMS1 GLRA1 ADRB1
KCNA7 CACNA1B DRD2
SLC6A4 SHANK1 SSTR2
KCNC4 KCNC2 PRLHR
GRIA2 CACNG2 NTSR2
KCNS1 ADCY5 HTR6
KCNC3 PPM1E HTR7
AVPR1A
Age-hypomethylated genes
Collagen chain trimerization
Complement and coagulation cascades
Complement cascade
Collagen biosynthesis and modifying enzymes
Collagen formation
AVB3 integrin pathway
Cell adhesion molecules cams
RAC1 GTPase cycle
Apoptosis
UV response dn
Myogenesis
Mitotic spindle
Rho GTPase cycle
Epithelial mesenchymal transition
Complement
ADAMTS3
COL15A1
COL1A1
COL22A1
COL6A1
COL6A3
COL8A2
LOXL4
P3H2
TNNT3
MYL3
COL15A1
COL6A3
NAV2
CLU
COL1A1
NQO1
SPTAN1
ABLIM1
APP
≥ 5 cell types
B cells
CD4 cells
CD8 cells
granulocytes
monocytes
NK cells
Enrichment score
Count
0 40
hypometh genes
hypermeth genes
N.S(q>0.05)
mixed(showing both hypo and hypermethylation with age)


A.
Hypomethylated sites
B.
Hypermethylated sites
β reg coeff.CD8
β reg coeff.gran
β reg coeff.mono
β reg coeff.NK
β reg coeff.InChianti
R = 0.51, p = 2e-10
R = 0.46, p = 1.9e-08
R = 0.49, p = 2.1e-09
R = 0.45, p = 4.9e-08
R = 0.5, p = 6.9e-06
R = 0.44, p = 0.00012
R = 0.47, p = 3.5e-05
R = 0.34, p = 0.0033


A.
age-hypomethylated probes
age-hypermethylated probes
All Illumina EPIC probes
proportion
Active/ flanking TSS
Active/ weak transcription
Active/weak enhancers
Heterochromatin
Bivalent TSS/enhancers
Repressed polycomb
Others
B CD4 CD8 mono NK
B.
Active/weak enhancers
Bivalent TSS/enhancers
Repressed polycomb
n.s*
% of probes hypomethylated with age
% of probes hypermethylated with age
% of Illumina Infinium MethylationEPIC probes
*n.s - Chi-squared test p>0.05. Others are p<0.05
C.
hypomethylated
hypermethylated
DHS
H3K4me1
selective
shared
-3kb
CpG
3kb
B cell
CD4 cell



Roy et.al., Figure 4-figure supplement 1

A.

Hypomethylated with age

BN_6507 probes

Motif	de novo	p
	Arnt	1e-97
	CTCF	1e-63
	Stat4	1e-40
	Fra1 (bZIP)	1e-28
	NFIA	1e-24

CD4N_63565 probes

Motif	de novo	p
	Arnt	1e-1084
	Fra1 (bZIP)	1e-424
	Klf7	1e-92
	IRF1 (IRF)	1e-65
	CTCFL	1e-62

CD8N_1850 probes

Motif	de novo	p
	Bcl6	1e-21
	NFIL3	1e-18
	PRDM1	1e-17
	CTCF	1e-17
	NFIA	1e-16

granulocytes_4627 probes

Motif	de novo	p
	Arnt	1e-86
	BORIS (Zf)	1e-65
	ZNF354C	1e-25
	Bhlhb2	1e-23
	RUNX1 (Runt)	1e-21

monocytes_8217 probes

Motif	de novo	p
	Arnt	1e-126
	CTCF (Zf)	1e-113
	Stat3 (Stat)	1e-26
	ETS	1e-25
	ZNF416	1e-20

NK_4148 probes

Motif	de novo	p
	Arnt	1e-56
	CTCF (Zf)	1e-50
	Tcfap2c	1e-18
	NF1	1e-18
	Smad2 (MAD)	1e-18

B.

Hypermethylated with age

BN_3863 probes

Motif	de novo	p
	JUND	1e-28
	Myb	1e-26
	REST	1e-26
	ZBED1	1e-25
	RUNX2	1e-24

CD4N_41083 probes

Motif	de novo	p
	Sp100	1e-359
	ETS	1e-319
	RUNX1	1e-146
	FOXL1	1e-119
	E2F2	1e-114

CD8N_1703 probes

Motif	de novo	p
	REST	1e-27
	Zfp116	1e-24
	Sp100	1e-21
	GRHL1	1e-20
	Lhx3	1e-20

granulocytes_174 probes

Motif	de novo	p
	REST	1e-38
	Sp100	1e-27
	Hoxc9	1e-26
	Sox8	1e-19
	ZEB1 (Zf)	1e-18

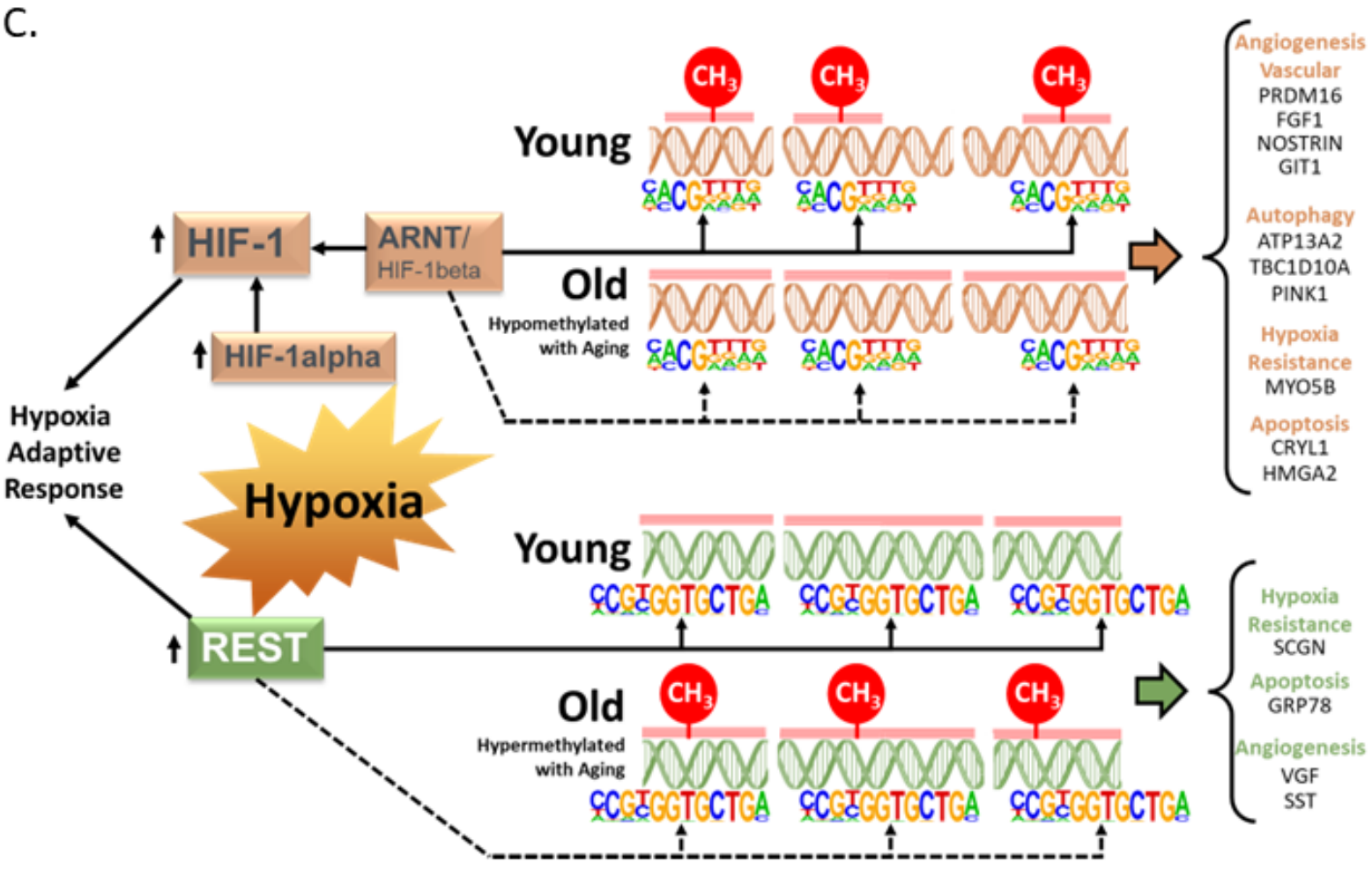
monocytes_5411 probes

Motif	de novo	p
	AridSa	1e-53
	REST	1e-35
	Sp100	1e-28
	HIC1 (Zf)	1e-23
	MYBL2	1e-23

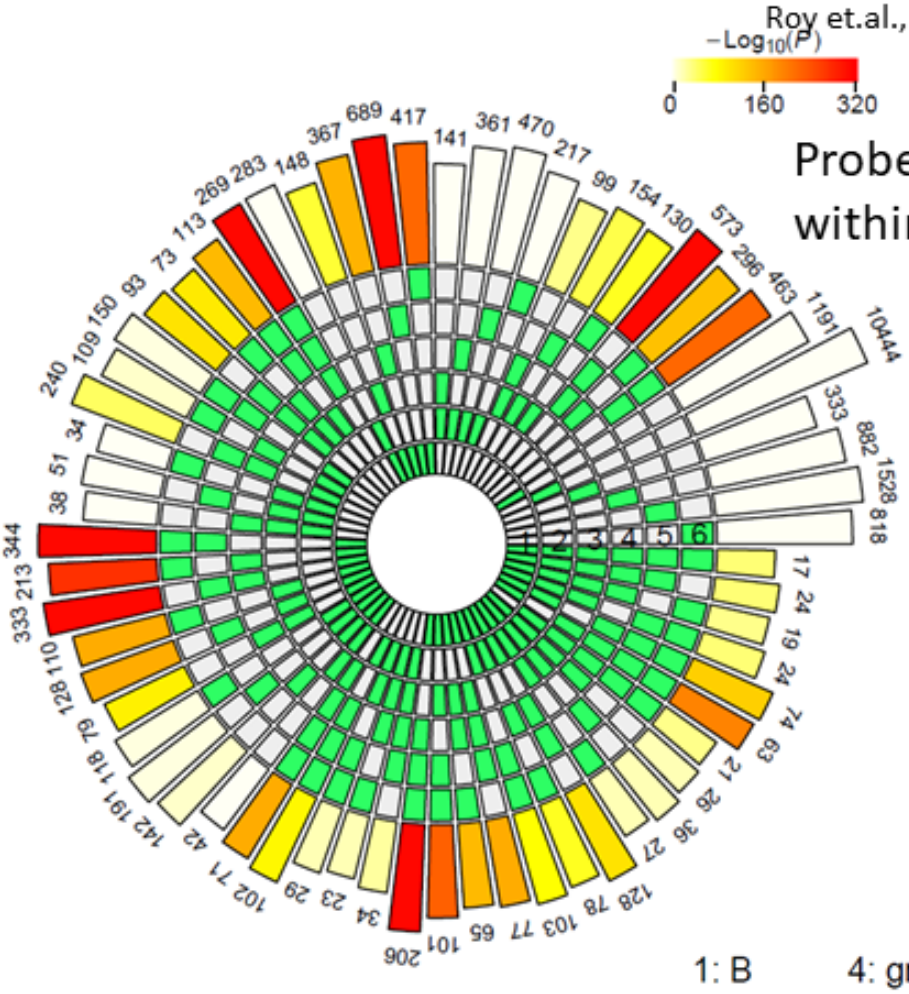
NK_623 probes

Motif	de novo	p
	REST	1e-35
	NRL	1e-20
	Osr2	1e-19
	HIC2	1e-18
	Egr1 (Zf)	1e-17

C.



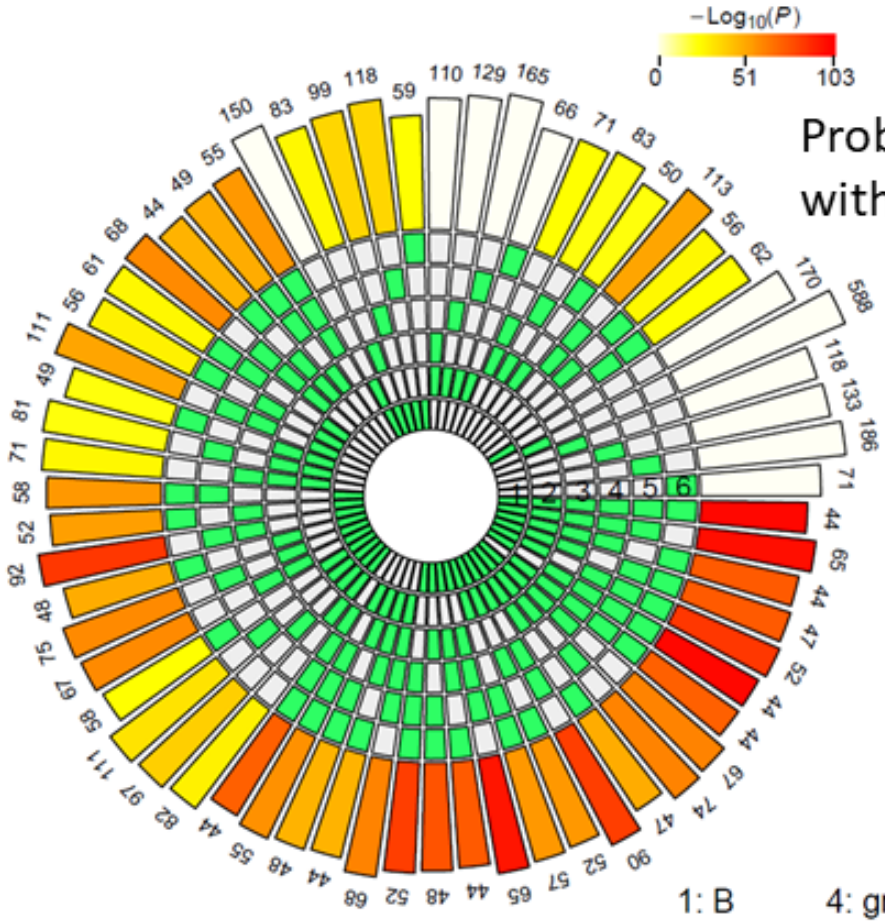
A.



Probes with ARNT motif within 1kb

1: B
2: CD4
3: CD8
4: gran
5: mono
6: NK

B.



Probes with REST motif within 1kb

1: B
2: CD4
3: CD8
4: gran
5: mono
6: NK