A New Monocyte Epigenetic Clock Reveals Effects of Alcohol Consumption on Epigenetic Aging in Three Independent Cohorts

Xiaoyu Liang[1,2,9], Rajita Sinha[1,7,8], Amy C. Justice [2,3], Mardge H. Cohen[4], Bradley E. Aouizerat[5,6], and Ke Xu [1,2*]

[1]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA;

[2]VA Connecticut Healthcare System, West Haven, CT, USA;

[3]Yale University School of Medicine, New Haven Veterans Affairs Connecticut Healthcare System, New Haven, CT, USA;

[4]Department of Medicine, Stroger Hospital of Cook County, Chicago, IL, USA;

[5]Bluestone Center for Clinical Research, College of Dentistry, New York University, New York, NY, USA;

[6]Department of Oral and Maxillofacial Surgery, College of Dentistry, New York University, New York, NY, USA;

[7]Child Study Center, Yale School of Medicine, New Haven, CT, USA;

[8]Department of Neuroscience, Yale School of Medicine, New Haven, CT, USA

[9]Department of Preventive Medicine, University of Tennessee Health Science Center, Memphis, TN, USA

*All corresponding to

Ke Xu, MD, PhD, Associate Professor of Psychiatry

Department of Psychiatry, Yale School of Medicine

VA Connecticut Healthcare System

950 Campbell Avenue

West Haven, CT 06516

Telephone: 203-932-5711

Email: ke.xu@yale.edu

Abstract

Background: Excessive alcohol consumption increases the risk of aging-related comorbidities and mortality. Assessing the impact of alcohol consumption on biological age is important for clinical decision-making and prevention. Evidence shows that alcohol alters monocyte function, and age is associated with DNA methylome and transcriptomic changes among monocytes. However, no monocyte-based epigenetic clock is currently available. In this study, we developed a new monocyte-based DNA methylation clock (MonoDNAmAge) by using elastic net regularization. The MonoDNAmAge was validated by benchmarking using epigenetic age acceleration (EAA) in HIV infection. Using MonoDNAmAge clock as well as four established clocks (i.e., HorvathDNAmAge, HannumDNAmAge, PhenoDNAmAge, GrimDNAmAge), we then evaluated the effect of alcohol consumption on biological aging in three independent cohorts (N=2,242).

Results: MonoDNAmAge, comprised of 186 CpG sites, was highly correlated with chronological age ($r_{training}$=0.96, p<2.20E-16; $r_{testing}$=0.86, p=1.55E-141). The MonoDNAmAge clock predicted an approximately 10-year age acceleration from HIV infection in two cohorts. Quadratic regression analysis showed a nonlinear relationship between MonoDNAmAge and alcohol consumption in the Yale Stress Center Community Study (YSCCS, p_{model}=4.55E-08, p_{x^2}=7.80E-08) and in the Veteran Aging Cohort Study (VACS, p_{model}=1.85E-02, p_{x^2}=3.46E-02). MonoDNAmAge and light alcohol consumption showed a negative linear relationship in the Women's Interagency HIV Study (WIHS, β=-2.63, p_x=2.82E-06). Heavy consumption increased EAA$_{MonoDNAmAge}$ up to 1.60 years in the VACS while light consumption decreased

EAA$_{MonoDNAmAge}$ to 2.66 years in the WIHS. These results were corroborated by the four established epigenetic clocks.

Conclusions: We observed a nonlinear effect of alcohol consumption on epigenetic age that is estimated by a novel monocyte-based "clock" in three distinct cohorts, highlighting the complex effects of alcohol consumption on biological age.

Keywords: Monocyte-based epigenetic clock, alcohol consumption, epigenetic age acceleration.

Background

Alcohol consumption has significant adverse effects on health and contributes to increased morbidity and mortality [1]. Biological aging has been proposed as an indicator of alcohol's adverse effect on health and frailty [2, 3]. Recently developed epigenetic "clocks" employ cellular DNA methylation (DNAm) as a measure of the aging process. DNAm-based epigenetic clocks were shown to be more sensitive and precise measures of cellular age than other genomic measures (e.g., transcriptome, telomere length) [4]. However, the study of the impact of alcohol consumption on epigenetic age is still in its infancy.

To date, more than a dozen DNAm-based epigenetic clocks have been reported including four well-established DNAm-based age estimators. The Horvath clock (HorvathDNAmAge) is based on 353 CpG sites capture multi-tissue biological age estimation [5]. The Hannum clock (HannumDNAmAge) is derived from 71 CpGs in leukocytes [6]. The Levine clock (PhenoDNAmAge) employs 513 CpGs predicting life span [7]. Lu's GrimAge clock (GrimDNAmAge) is a linear combination of chronological age, sex, and 1,030 CpG sites modeled as surrogate biomarkers for seven plasma proteins and smoking pack-years, predicting age at death [8]. These clocks have been applied as biomarkers of the aging process to understand how numerous medical or psychiatric conditions impact biological age. For example, HIV infection has shown a 5 to 10 year accelerated biological age [9, 10].

These established "clocks" have been applied to examine alcohol use on biological age. Epigenetic age acceleration (EAA) in heavy alcohol use and in children with fetal alcohol spectrum disorder have been reported [11, 12]. People with alcohol use disorder

5

show a trend in age acceleration in liver tissue [13] and have a 2.22-year acceleration in blood relative to healthy individuals [14]. On the other hand, light to moderate alcohol use appears to no change or even slow extrinsic epigenetic age acceleration [15, 16]. These observations suggest that the impact of alcohol consumption on the biological aging process may differ by the quantity of use and by tissue types.

One limitation of previous studies using available epigenetic clocks is the lack of tissue and cell-type specificity to estimate the biological age of alcohol use, thus they may offer limited insight into the mechanisms of tissue-specific cellular aging. Alcohol consumption changes immunity and inflammatory functions that may result from functional alteration of immune cells [17]. For example, moderate or heavy alcohol use changes monocyte function [18-20]. Interestingly, Szabo et al [19] reported that alcohol consumption had a biphasic effect on interferon-inducibility and could affect monocyte-derived inflammatory cytokine production, which changes the course of the aging process. In addition, monocytes show aging-related gene dysfunction in metabolism, immune function, and inflammation [21, 22], as well as DNAm and transcriptomic alterations [23, 24]. Therefore, it is reasonable to posit that a set of CpG sites selected from monocytes can serve as an indicator of the impact of alcohol consumption on biological age and provide insight into its underlying mechanisms.

In this study, we aimed to characterize the effect of a range of alcohol consumption on monocyte epigenetic age. We developed a novel "clock" derived from the human monocyte DNA methylome (termed "MonoDNAmAge") by using Elastic Net Regularization (ENR). We evaluated the performance of the MonoDNAmAge in estimating HIV-associated age acceleration as a benchmark. Finally, we assessed the

6

impact of alcohol consumption on MonoDNAmAge, EAA, and apparent methylation age rate (AMAR) in three distinct cohorts: Yale Stress Center Community Cohort (YSCCS) [25]; Veteran Aging Study Cohort (VACS) [26]; and Women's Interagency HIV Study (WIHS) [27, 28]. Demographic and phenotypic information for each cohort is presented in **Table 1 and Supplementary Information**. Four well-established clocks (i.e., HorvathDNAmAge, HannumDNAmAge, PhenoDNAmAge, GrimDNAmAge) were also used to assess the impact of alcohol consumption on measures of biological age (**Fig. 1**).

Results

MonoDNAmAge clock and biological interpretation

MonoDNAmAge was derived from CD14+ monocyte methylome in Multi-Ethnic Study of Atherosclerosis (MESA) (N=1,202) (GSE56046) [24]. We selected a set of 186 age-associated CpGs using ENR. The correlation between DNAm age and chronological age was 0.96 (p<2.20E-16) in the training set and was 0.86 (p=1.55E-141) in the testing set (**Fig. 2a**). These 186 CpG sites mapped to 135 genes (**Supplementary Information: Table S1**), including well-established genes associated with age (e.g. *KLF14*), transcription factors (e.g. *RUNX3*), and inflammatory function (e.g. *IL17RC*).

Interestingly, the 186 CpG sites were enriched for 25 complex traits with p<5.00E-03 in the Epigenome-Wide Association Study (EWAS) Atlas database [29] (**Fig. 2b**). The top significant traits included aging (p=2.57E-159), smoking (p=2.39E-38), breast cancer

7

(p=2.71E-09), and alcohol consumption (p=8.95E-05). Three established clocks (HorvathDNAmAge, HannumDNAmAge, and PhenoDNAmAge) were enriched for multiple traits including age and smoking, but except HorvathDNAmAge (p=2.45E-03), none of the other two clocks was enriched for alcohol consumption (**Supplementary Information: Fig. S1**). One CpG site, *SCGN* cg06493394 was shared across the four clocks (**Fig. 2c**). Identifiers of CpG sites for GrimDNAmAge were not publicly available for comparisons. The 135 genes harboring the 186 CpG sites were enriched on biological pathways relevant to aging (e.g. regulation of multicellular organismal process) by performing Database for Annotation, Visualization and Integrated Discovery (DAVID) pathway enrichment analysis (http://david.niaid.nih.gov) (**Supplementary Information: Fig. S2**) [30].

MonoDNAmAge was significantly correlated with chronological age in all three cohorts (YSCCS: r=0.90, p=3.12E-181; VACS: r=0.54, p=1.75E-96; WIHS: r=0.66, p=1.50E-60) (**Fig. 2d**). We also estimated the correlations of four established clocks with chronological age in these three cohorts. As expected, all clocks showed significant correlations with chronological age in each cohort (YSCCS: p=8.42E-215~4.69E-168; VACS: p=6.82E-205~3.94E-135; WIHS: p=2.25E-175~1.75E-91) (**Supplementary Information: Fig. S3**).

Benchmarking MonoDNAmAge against HIV infection shows epigenetic age acceleration

EAA analysis showed an average age acceleration of 10.14-years in the VACS (p$_{VACS}$=1.17E-24) and 12.17-years (p$_{WIHS}$=2.07E-28) in the WIHS in HIV+ participants

compared to HIV-negative (HIV-) participants (**Fig. 3a, Supplementary Information: Table S2**). To avoid a bias due to different sample sizes of HIV+ and HIV- participants, we randomly selected HIV+ participants to match the number of HIV- participants and compared the EAA value between the groups in both cohorts in **Fig. 3b**. As shown in **Fig. 3b**, the majority of HIV+ participants displayed positive EAA values while the majority of HIV- participants displayed negative EAA values. The methylation age rate (i.e., AMAR) of monocyte biological age was significantly greater in HIV-positive (HIV+) than HIV- participants in both cohorts (MD_{VACS}=0.21, p_{VACS}=1.00E-27; MD_{WIHS}=0.34, p_{WIHS}=6.51E-37) (**Fig. 3c, Supplementary Information: Table S2**).

Additionally, we tested age acceleration in HIV infection using four established epigenetic clocks in VACS and WIHS. The EAAs estimated by all four clocks were significantly different between HIV+ and HIV- participants in both VACS (p=3.26E-11~1.07E-02) and WIHS (p=4.62E-14~1.34E-04). For example, the EAA differences between HIV+ and HIV- were as great as 2.53 in VACS and 3.60 in WIHS measured by HorvathDNAmAge (**Supplementary Information: Fig. S4 and Table S2**). AMAR was also greater in HIV+ than HIV- participants except for HannumDNAmAge in VACS and GrimDNAmAge in VACS and WIHS (**Supplementary Information: Fig. S5 and Table S2**).

Nonlinear effects of alcohol consumption on DNA methylation age

ANOVA model comparison showed that the quadratic regression model fits the data better than the linear regression model for all five clocks in the YSCCS (p=6.27E-09~2.41E-06) and VACS (p=8.32E-04~3.91E-02) but not in the WIHS cohort (p=2.47E-

01~4.38E-01) (**Supplementary Information: Table S3**). Then, the quadratic regression was applied to YSCCS (**Fig. 4a**) and VACS (**Fig. 4b**). In the YSCCS cohort, we observed a nonlinear relationship between MonoDNAmAge and Alcohol Use Diagnosis Identification Test (AUDIT) score (p_{model} =4.55E-08, p_{x^2}=7.80E-08) (**Fig. 4a**). The other four DNAm clocks also showed nonlinear associations between DNAm age and AUDIT score (HorvathDNAmAge: p_{model}=2.82E-07, p_{x^2}=2.37E-07; HannumDNAmAge: p_{model}=5.54E-06, p_{x^2}=2.41E-06; PhenoDNAmAge: p_{model}=4.08E-10, p_{x^2}=6.27E-09; GrimDNAmAge: p_{model}=1.80E-07, p_{x^2}=6.12E-08).

To further investigate the relationship between MonoDNAmAge and different levels of alcohol consumption, we performed two linear regression analyses at the inflection point of the nonlinear distribution at the AUDIT score of 13 in YSCCS, separately (**Supplementary Information: Fig. S6**). We found that MonoDNAmAge was negatively associated with consumption at AUDIT<13 (p=5.36E-06) and positively associated with heavy drinking of AUDIT>13 (p=3.70E-03). Every unit change in the AUDIT score<13 was associated with a 1.20-year decrease in MonoDNAmAge in the non-heavy alcohol drinking (non-HAD) group and every unit change in AUDIT score>13 was associated with a 0.36-year increase in MonoDNAmAge in the heavy alcohol drinking (HAD) group.

In the VACS cohort, quadratic regression also showed a significant association of MonoDNAmAge on alcohol consumption measured by the natural logarithm of phosphatidylethanol (ln(PEth)) (p_{model}=1.85E-02, p_{x^2}=3.46E-02) (**Fig. 4b**). The results from the three established clocks provided estimates consistent to MonoDNAmAge: HannumDNAmAge (p_{model}=3.74E-03, p_{x^2}=8.32E-04), PhenoDNAmAge (p_{model}=4.97E-

04, p_{x^2}=8.60E-04), GrimDNAmAge (p_{model}=3.40E-02, p_{x^2}=2.21E-02). Here, the inflection point for the curve of PEth levels was close to the PEth cutoff for the definition of HAD. We found that the slope below the PEth inflection point (non-HAD: PEth<20, that is ln(PEth)<2.996) showed no significant (p=6.79E-01) association with MonoDNAmAge while the slope above the PEth inflection point (HAD: PEth≥20, that is ln(PEth)≥2.996), showed a significant positive association with MonoDNAmAge (p=4.39E-02) (**Supplementary Information: Fig. S6**). Every unit change of ln(PEth) above the inflection point was associated with a 1.31-year increase in DNAm age. However, we found that no significant association of DNAm age on AUDIT-Consumption (AUDIT-C, first 3 items of AUDIT) score for all five clocks (p_{model}=1.79E-01~9.01E-01) (**Supplementary Information: Fig. S7**), suggesting that PEth is a more accurate measure of alcohol consumption on DNAm age than self-reported AUDIT-C.

In the WIHS, MonoDNAmAge was linearly associated with light alcohol consumption measured by the number of drinks per week (NDRNKWK) (β=-2.63, p_{model}=2.82E-06, p_x=2.82E-06) (**Fig. 4c**). The negative linear relationship between DNAm age and light alcohol consumption was also observed with the four established clocks (HorvathDNAmAge: β=-1.41, p_{model}=1.33E-05, p_x=1.33E-05; HannumDNAmAge: β=-1.38, p_{model}=4.99E-05, p_x=4.99E-05; PhenoDNAmAge: β=-1.70, p_{model}=7.74E-05, p_x=7.74E-05; GrimDNAmAge: β=-1.03, p_{model}=1.23E-03, p_x=1.23E-03).

Alcohol consumption alters epigenetic age acceleration

In the YSCCS cohort, alcohol consumption was not associated with EAA estimated using MonoDNAmAge or the four established clocks (**Fig. 4d and Table 2**). However,

11

AMAR from HorvathDNAmAge, HannumDNAmAge, and GrimDNAmAge showed greater DNAm age than chronological age and were associated with alcohol consumption (β=0.003~0.005, p=3.31E-05~7.29E-03).

In the VACS cohort, EAA analysis showed that alcohol consumption significantly accelerated MonoDNAmAge by 1.60 years (β=0.477, p=3.99E-03) (**Fig. 4e and Table 2**). EAA estimated by HannumDNAmAge and PhenoDNAmAge also showed significant acceleration (β=0.237, p=2.0E-02 and β=0.375, p=3.76E-03, respectively). AMAR estimated by MonoDNAmAge was associated with alcohol consumption (β=0.012, p=3.55E-04). Similarly, PhenoDNAmAge showed a similar association (β=0.011, p=9.16E-05). HannumDNAmAge and GrimDNAmAge were also significantly positively correlated with alcohol consumption.

In the WIHS cohort, MonoDNAmAge EAA was correlated with alcohol consumption in an inverse relationship (β=-1.627, p=1.56E-04, MD$_{EAA}$=-2.656) (**Fig. 4f and Table 2**). All five clocks showed decelerations of DNAm age with an average MD$_{EAA}$ of -1.24 years. AMAR for MonoDNAmAge was negatively correlated with alcohol consumption (MonoDNAmAge: β=-0.046, p=2.54E-05). Similarly, each AMAR for HorvathDNAmAge, HannumDNAmAge, and PhenoDNAmAge, was significantly negatively correlated with alcohol consumption.

DNAm-based age estimated by the MonoDNAmAge clock and four other clocks were significantly correlated with each of the six cell types in all three cohorts (r=0.54~0.93, p=8.42E-215~1.50E-60) (**Fig. 5**). Correlations were calculated for each of the five clocks with six estimated cell types in the HAD and moderate alcohol drinking (MAD)

groups separately in YSCCS and VACS, and light alcohol drinking (LAD) and non-alcohol drinking (NAD) in WIHS. The correlation patterns were similar between the two alcohol use groups across the three cohorts, suggesting the effect of cell-type composition on the performance of the five biological age estimators does not differ significantly between HAD and MAD, or between LAD and NAD groups.

Discussion

Our results demonstrate that a set of DNAm CpG sites in monocytes are predictive of biological age, enabling the detection of the impact of alcohol consumption on DNAm age validated in three distinct cohorts. MonoDNAmAge clock, benchmarked against HIV infection, showed an approximate 10-year acceleration in HIV-positive participants. More importantly, this novel MonoDNAmAge clock detected a nonlinear relationship between DNAm-based biological age and alcohol consumption in both a healthy community cohort (YSCCS) and a clinic-based cohort (VACS) that included heavy alcohol drinkers (~33%). The impact of alcohol consumption on MonoDNAmAge was corroborated using four established epigenetic clocks. Thus, employing a comprehensive approach (i.e., one novel and four established epigenetic clocks, three independent cohorts, different but commonly employed measures of alcohol exposure), we identified for the first time that alcohol consumption appears to have a complex, nonlinear relationship with DNAm-based estimates of biological age.

Different from previous studies, here we examined the effect of alcohol consumption on epigenetic age by applying a novel cell-type-specific epigenetic clock related to the

13

biological mechanism of alcohol consumption. The monocyte methylome plays an important role in epigenetic aging. Recently, differentially methylated regions in CD14 monocytes between young (24-30 years) and older (57-70 years) individuals have been reported [31]. Those age-associated CpG sites or DNAm regions have been linked to transcriptomic changes with aging [31]. Therefore, a CD14 monocyte epigenetic clock may provide a more accurate measure of the biological age of genes that compose the MonoDNAmAge clock that have functional implications for the aging process in monocyte-related phenotypes such as HIV infection and alcohol consumption. Indeed, the 186 MonoDNAmAge CpG sites are enriched in alcohol consumption and other related traits (e.g. smoking, cancer). Interestingly, although MonoDNAmAge and the established epigenetic clocks are composed of different sets of CpGs with the exception of one site (*SCGN* cg06493394)*,* these clocks showed reasonably consistent but not identical performance in estimating alcohol's effect on epigenetic age. We speculate that each epigenetic clock is estimating different but related facets of biological aging, inferred in part by the overlapping traits that show enrichment for the CpG sites that compose each DNAm-based epigenetic clock.

The nonlinear association between alcohol consumption and biological age is interesting. The beneficial and harmful effects of alcohol consumption have been well documented in the literature [1, 32-35]. While any amount of alcohol is harmful to health [1], some studies showed that alcohol consumption had a U-shaped relationship to biomarkers including HDL, LDL, VLDL, and inflammation [36, 37], underscoring the bidirectional effects on cardiovascular and metabolic disease risks. In a line of this evidence, we observed a nonlinear association between alcohol consumption and

biological age in two cohorts, YSCCS and VACS, in which the numbers of heavy and non-heavy alcohol consumers were relatively balanced. In the WIHS cohort, the predominance of none-to-light drinkers may explain why the negative association with biological aging was similarly observed; insufficient observations were available to test the relationship between heavy drinkers and biological aging. The results of EAA showed heavy consumption measured by PEth accelerated epigenetic age of 1.60 years, consistent with a previous report of a 2-year acceleration in individuals with Alcohol Use Disorder (AUD) [14]. On the other hand, EAA was inversely associated with light to moderate alcohol consumption, which is consistent with a previous report that self-reported moderate alcohol use (i.e., number of drinks per week) showed a negative correlation with biological age estimated using PhenoDNAmAge [7].

A possible explanation for the observed nonlinear relationship between alcohol consumption and biological aging is that individuals with light to moderate alcohol use are more likely to follow a healthier lifestyle. For example, in the YSCCS and VACS, individuals not reporting HAD showed lower smoking rates and cannabis use, than those with HAD, suggesting that a healthy lifestyle may slow or potentially even reverse the epigenetic aging process due to alcohol consumption. Another possibility is due to underreported alcohol consumption among participants. For example, in the VACS cohort, 17.02% of participants reported AUDIT-C as 0 but PEth level among those participants are greater than 8ng/mL (**Supplementary Information: Fig. S8**), which indicates active alcohol use [38]. Inaccurately self-reported alcohol consumption may result in biased findings towards the slow acceleration of light-to-moderate drinkers on biological age. These observations need further investigations with accurately assessed

15

phenotype or in longitudinal cohorts. Whereas the overall harm of alcohol drinking consumption is outweighed the benefit, causality has remained elusive [32].

We acknowledge some limitations of this study. A recent study suggests that the accuracy of DNAm-based epigenetic clock estimation is affected by sample size [39]. Validation of the MonoDNAmAge clock in a large independent sample is warranted. DNAm in the three cohorts was measured in DNA from whole blood. We expect that DNAm derived solely from monocytes may be a more accurate predictor of the impact of alcohol consumption. We are unable to examine DNAm age in past alcohol users, for different patterns of alcohol consumption, or for comorbidity with other substance or drug use. Furthermore, while differences in the effects of alcohol use between women and men are well documented, the study lacked sufficient power to evaluate for sex differences in alcohol exposure on biological aging.

Conclusion

We found that alcohol use impacts epigenetic aging in a nonlinear fashion with heavy consumption accelerating while non-heavy use decreasing the acceleration of biological age. The use of cell-type-specific epigenetic clocks that are known to be directly impacted by alcohol exposure may provide more precise information than provided by more holistic epigenetic clocks that estimate more global biological aging effects. Our study expands previous knowledge and provides new insights on the impact of a spectrum of alcohol use on epigenetic aging.

Methods

Study cohorts and phenotype assessments

Multi-Ethnic Study of Atherosclerosis (MESA) (N=1,202) [24]**.** DNAm data of the MESA Epigenomics and Transcriptomics Study (GSE56046) was used to construct the MonoDNAmAge clock. The DNA samples were from CD14+ monocyte samples, collected from 1,202 individuals with ages ranging from 44 to 83 years.

The following three cohorts were used for assessing alcohol's effects on DNAm aging (**Table 1,** see details of study cohorts and phenotype assessment in **Supplementary Information**).

Yale Stress Center Cohort Study (YSCCS) (N=502) [25]. The cohort served as a community-based sample to examine the impact of alcohol consumption on DNAm age among healthy participants. The 10-item AUDIT scale was used to assess alcohol use. HAD was defined as AUDIT≥8 for men and AUDIT≥7 for women (N=148) [40]. The average AUDIT score was 5.69 among all participants in the cohort.

Veterans Aging Cohort Study (VACS) (N=1,259) [26]. The cohort served as a clinic-based sample to benchmark HIV infection on MonoDNAmAge and to examine the effects of alcohol consumption on DNAm age. The participants included both HIV+ (N=1,151) and HIV- (N=104) individuals. A majority of HIV+ participants were on antiretroviral therapy and were virally suppressed (63.66%). Alcohol consumption was assessed by using PEth, a biomarker for alcohol use [22] that is positively correlated with AUDIT scores [41, 42]. HAD was defined as PEth≥20 (N=299) according to a

17

previous study [43]. The average PEth was 41.7 ng/ml. Self-reported AUDIT-C was also collected for each participant.

***Women's Interagency HIV Study (WIHS)* (N=481)** [27, 28]. The cohort served as a clinic- and community-based sample to examine HIV infection and alcohol consumption on DNAm age. This study included HIV-positive (N=272, 90.44% virally suppressed) and HIV-negative (N=209) participants. The cohort predominantly reported light alcohol use with an average NDRNKWK of 0.7. The LAD was defined as 0<NDRNKWK≤7 (N=196), and the NAD was defined as NDRNKWK=0 (N=255) [44].

DNA methylation and data quality control (QC)

Epigenome-wide CpG methylation was profiled by using either the Illumina HumanMethylation450 BeadChip (HM450K) (San Diego, CA, USA) in MESA (monocyte), YSCCS (blood), and VACS (blood) (57.2% of the sample) or Illumina HumanMethylation EPIC BeadChip (EPIC) (San Diego, CA, USA) in 42.8% of the VACS samples and the WIHS (blood) samples. All samples in the three study cohorts (YSCCS, VACS, and WIHS) were processed at the Yale Center for Genomic Analysis [45]. We applied the method described by Houseman et al. [46, 47] to estimate proportions of CD4+ T cell, CD8+ T cell, NK T cell, B cell, monocyte, and granulocyte DNAm in each cohort. The QC for the MESA cohort was previously reported [24]. For the YSCCS, VACS, and WIHS cohorts, we applied the same QC process as reported in our previous studies [45, 48, 49].

MonoDNAmAge clock development

As depicted in **Fig. 1**, we applied ENR to build a MonoDNAmAge clock. All samples in MESA were divided into two sets: a training (N=721) and a testing (N=481) set.

Briefly, the top 1,000 significant CpG features were preselected from monocyte methylome based on the EWAS on age in the MESA. In the training set, the ENR combined with random sampling was performed to select CpGs and establish the predictive model. The CpGs were filtered and sorted according to their coefficients in ENR. In the testing set, we added CpGs sequentially and calculated the corresponding correlations between the predicted MonoDNAmAge and chronological age. We determined the linear combination of the CpG set with the correlation at the inflection point of the performance curve in the testing set as MonoDNAmAge (**Fig. 2a**). The performance of MonoDNAmAge was evaluated in the three validation data sets (i.e., YSCCS, VACS, and WIHS). Analyses were performed using R software. ENR was performed using the function "cv.glmnet" in the "glmnet" package. The details of feature selection are described below.

Preselection of CpGs. Because a large number of CpGs may introduce noise, DNA methylation of CpGs under the epigenome-wide significance threshold may collectively account for phenotype variation and may improve prediction of a phenotype, we preselected the top 1000 age-associated CpGs from the EWAS on age in the MESA cohort.

Importance ranking of CpGs in the training set. The preselected CpGs were used to establish the predictive model in the training set of MESA. We randomly selected 80% of training subjects without replacement 200 times and constructed a model for each

replication. We only included the CpGs that were present in more than 95% of all replications in the final model and ranked the CpGs based on the coefficients of all replications. For the ENR method in each replication, the 10-fold cross-validation procedure was included in the feature selection to obtain the estimations of ENR tuning parameters. We extracted the coefficients for the model with the lambda value corresponding to the minimum mean cross-validated error. The CpG importance ranking was based on the summation of the absolute value of the coefficients of all replications.

Epigenetic clock construction using ENR in the testing set. The CpG features were selected based on Pearson's correlation coefficient between predicted DNAm age and chronological age in the testing set of MESA. Based on the CpG importance ranking, CpGs were added one at a time, and the correlation coefficients between predicted DNAm age and chronological age were calculated. We selected the CpG set with the correlation at the inflection point of the performance curve.

Evaluation of epigenetic clock performance in the validation dataset. The performance of the CpG features selected from MESA was evaluated in three independent validation cohorts (YSCCS, VACS, and WIHS) using the four measures, EAA, MD_{EAA}, AMAR, and MD_{AMAR} (see Statistics section).

Statistics

Epigenetic clocks and assessments

Pearson correlation coefficients (r) between the DNAm age and chronological age were estimated in each cohort. EAA was defined as the residuals of regressing DNAm age on

20

chronological age [5, 50]. We calculated the mean difference of EAA between heavy users and non-heavy users, between HIV+ and HIV-, and denoted it as MD_{EAA}. AMAR was defined as the ratio of DNAm age to chronological age [6], AMAR>1 represents DNAm age acceleration, and AMAR<1 represents DNAm age deceleration. We also calculated the mean difference of AMAR between heavy and non-heavy users, between HIV+ and HIV-, and denoted as MD_{AMAR}.

Benchmarking the MonoDNAmAge clock using HIV infection

We evaluated whether the MonoDNAmAge was able to predict changes in DNAm age for HIV infection in two cohorts (VACS and WIHS) using linear regression, EAA or AMAR as the dependent and HIV infection as the independent variable. We adjusted potential confounding factors including self-reported ethnicity, sex, tobacco use, body mass index (BMI), and alcohol-related phenotype in the model.

Association between alcohol consumption and DNAm-based biological age

Instead of just assuming a linear relationship between alcohol consumption and DNAm age, we first compared the linear model (reduced model) with the quadratic model (full model) for modeling the relationship between DNAm age and alcohol consumption (AUDIT in YSCCS; ln(PEth) in VACS; and NDRNKWK in WIHS) using ANOVA. The null hypothesis is that the reduced model is as good as the full model. We applied quadratic model when the quadratic model was significantly better than the linear model, otherwise, we applied the linear model.

In the quadratic regression model, DNAmAge=$\beta_2 x^2 + \beta_1 x + \beta_0$, where $\beta_2 \neq 0$ and x represents alcohol consumption. The t-test was used to conduct hypothesis tests on the

regression coefficients obtained in both quadratic regression and simple linear regression. That is, for the quadratic regression model, we tested $H_0: \beta_j = 0$ vs. $\beta_j \neq 0$ using t-test $T_{quadratic} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \sim t(n-3)$, where n is the sample size. We also tested the regression model H_0: the regression model was not significant vs. H_a: the regression model was significant by performing F test, $F_{model} = \frac{MSR}{MSE} = \frac{SSR/2}{SSE/(n-3)} \sim F(2, n-3)$ based on ANOVA test, where MSR was the regression mean square and MSE was the error mean square, $SSR = \sum_{i=1}^{n} n\left(\widehat{DNAmAge}_i - \overline{DNAmAge}\right)^2$ was the corrected sum of squares for regression model and $SSE = \sum_{i=1}^{n} n\left(DNAmAge_i - \widehat{DNAmAge}_i\right)^2$ was the sum of squares for error.

We also examined EAA and AMAR between different groups of alcohol consumption in each cohort, adjusted potential confounding factors including self-reported ethnicity, sex, tobacco use, and BMI.

Correlation between DNAm age and cell types in different alcohol consumption groups

To address potential cell type confounding effects, Pearson correlation coefficients (r) between DNAm-based estimates of biological age and six cell type proportions were estimated in each cohort.

22

Acknowledgments

Conflict of Interest

The authors declare no conflict of interest.

Funding

23

Availability of Data and Materials

Demographic variables, clinical variables and methylation status for the VACS samples are submitted to the GEO dataset (GSE117861) and are available to the public. All codes for analysis are also available upon a request to the corresponding author.

Authors' Contributions

XL was responsible for bioinformatics data processing and statistical analysis. ACJ provided DNA samples and clinical data and contributed to the interpretation of findings and manuscript preparation. BA contributed to the design of the study. RS and BA provided DNA methylation and phenotype data, contributed to the interpretation of findings, and manuscript preparation. MC contributed to the manuscript preparation. KX was responsible for the study design, study protocol, sample preparation, data analysis, interpretation of findings, and manuscript preparation. All authors read and approved the final manuscript.

Table 1. Demographic and clinical characteristics for YSCCS, VACS, and WIHS

Variable	YSCCS			VACS			WIHS		
	HAD	MAD		HAD	MAD		LAD	NAD	
	Men: AUDIT≥8 Women: AUDIT≥7 (N=148)	Men: AUDIT<8 Women: AUDIT<7 (N=354)	p	PEth≥20 (N=299)	PEth<20 (N=738)	p	0<NDRNKWK≤7 (N=196)	NDRNKWK=0 (N=255)	p
age	26.8 ± 7.13	29.8 ± 9.28	1.35E-04	50.7 ± 7.23	51.1 ± 8.14	5.03E-01	40.3 ± 9.13	43.7 ± 9.46	1.13E-04
sex (%male)	64.9%	35.3%	2.18E-09	100%	100%/	NA	0%	0%	NA
race (AA)	12.2%	22.4%	1.29E-02	86.6%	80.2%	1.90E-02	49.0%	49.8%	9.69E-01
smoker	39.9%	12.7%	1.77E-11	68.1%	54.3%	6.27E-05	70.9%	67.3%	8.97E-01
HIV-positive	0%	0%	NA	93.7%	91.3%	2.63E-01	50.5%	63.9%	1.54E-04
$\log_{10}VL$	0	0	NA	2.82 ± 1.23	2.66 ± 1.24	8.44E-02	2.16 ± 0.69	2.05 ± 0.48	1.34E-01
cannabis use	88.4%	56.9%	2.40E-11	84.5%	73.4%	2.21E-04	50.8%	29.1%	3.83E-06
cocaine use	NA	NA	NA	75.3%	66.0%	4.34E-03	NA	NA	NA
stimulant use	NA	NA	NA	36.9%	39.9%	4.08E-01	NA	NA	NA
opiate use	NA	NA	NA	42.3%	44.2%	6.27E-01	NA	NA	NA

YSCCS: Yale Stress Center Cohort Study; VACS: Veterans Aging Cohort Study; WIHS: Women's Interagency HIV Study.
HAD: Heavy Alcohol Drinking; MAD: Moderate Alcohol Drinking; LAD: Light Alcohol Drinking; NAD: Non-Alcohol Drinking.
AA: African American, AUDIT: Alcohol Use Disorders Identification Test, AUDIT-C: Alcohol Use Disorders Identification Test-Consumption, VL: viral load.
Welch's two-sample t-test for testing two groups mean; chi-square test for testing two groups percentage.
Significant phenotypes are in bold.
The cannabis, cocaine, stimulant, and opiate use in VACS were defined as case (>0) and control (=0).

Table 2. Association between EAA/AMAR and alcohol consumption

Measure	Cohort	Method	beta	95% CI	p	MD
EAA	YSCCS	MonoDNAmAge	0.040	(-0.031, 0.112)	2.68E-01	0.129
		HorvathDNAmAge	0.044	(-0.008, 0.096)	1.00E-01	0.415
		HannumDNAmAge	0.028	(-0.026, 0.083)	3.11E-01	0.786
		PhenoDNAmAge	-0.031	(-0.111, 0.049)	4.42E-01	-0.856
		GrimDNAmAge	0.033	(-0.013, 0.079)	1.57E-01	0.767
	VACS	MonoDNAmAge	0.477	(0.153, 0.802)	**3.99E-03**	1.603
		HorvathDNAmAge	0.056	(-0.141, 0.253)	5.77E-01	0.135
		HannumDNAmAge	0.237	(0.038, 0.436)	**2.00E-02**	0.711
		PhenoDNAmAge	0.375	(0.122, 0.628)	**3.76E-03**	1.632
		GrimDNAmAge	0.138	(-0.079, 0.356)	2.13E-01	1.520
	WIHS	MonoDNAmAge	-1.627	(-2.463, -0.791)	**1.56E-04**	-2.656
		HorvathDNAmAge	-0.704	(-1.054, -0.353)	**9.73E-05**	-0.806
		HannumDNAmAge	-0.676	(-1.115, -0.236)	**2.75E-03**	-1.084
		PhenoDNAmAge	-0.843	(-1.374, -0.312)	**1.98E-03**	-0.914
		GrimDNAmAge	-0.406	(-0.658, -0.153)	**1.74E-03**	-0.746
AMAR	YSCCS	MonoDNAmAge	-0.003	(-0.0064, 0.0012)	1.84E-01	-0.056
		HorvathDNAmAge	0.003	(0.0007, 0.0044)	**7.29E-03**	0.026
		HannumDNAmAge	0.004	(0.0021, 0.0065)	**1.33E-04**	0.051
		PhenoDNAmAge	-0.003	(-0.0058, 0.0001)	6.32E-02	-0.052
		GrimDNAmAge	0.005	(0.0025, 0.0070)	**3.31E-05**	0.057
	VACS	MonoDNAmAge	0.012	(0.0054, 0.0185)	**3.55E-04**	0.034
		HorvathDNAmAge	0.002	(-0.0026, 0.0063)	4.19E-01	0.003
		HannumDNAmAge	0.004	(0, 0.0088)	**4.85E-02**	0.012
		PhenoDNAmAge	0.011	(0.0054, 0.0161)	**9.16E-05**	0.031
		GrimDNAmAge	0.010	(0.0044, 0.0157)	**4.89E-04**	0.030
	WIHS	MonoDNAmAge	-0.046	(-0.0668, -0.0247)	**2.54E-05**	-0.092
		HorvathDNAmAge	-0.015	(-0.0241, -0.0062)	**1.00E-03**	-0.013
		HannumDNAmAge	-0.012	(-0.0234, -0.0004)	**4.35E-02**	-0.009
		PhenoDNAmAge	-0.028	(-0.0426, -0.013)	**2.61E-04**	-0.057
		GrimDNAmAge	-0.002	(-0.0099, 0.0052)	5.43E-01	0.002

EAA: Epigenetic Age Acceleration; AMAR: Apparent Methylation Age Rate.
YSCCS: Yale Stress Center Cohort Study; VACS: Veterans Aging Cohort Study; WIHS: Women's Interagency HIV Study.
95% CI: 95% confidence interval for coefficient beta.
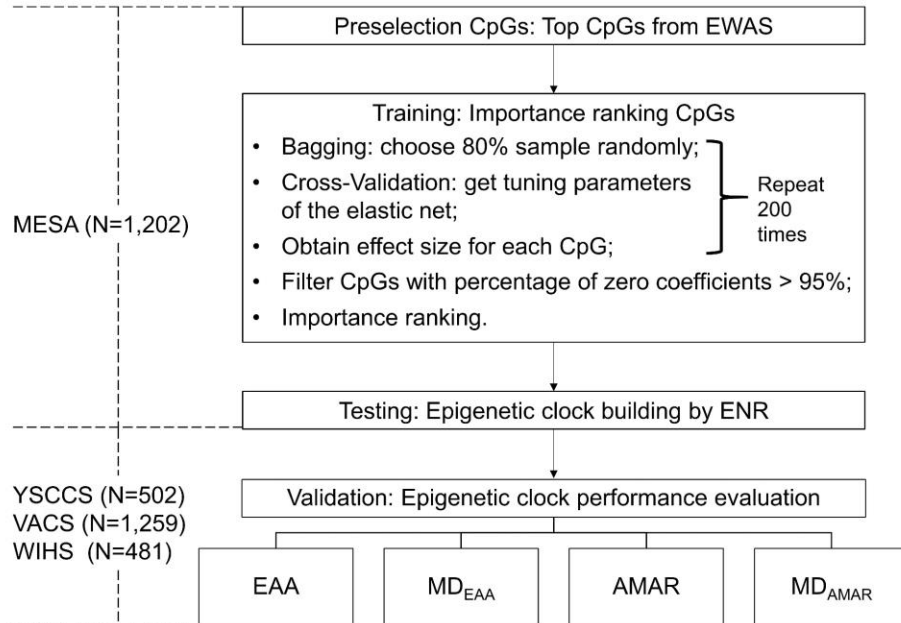The p-values in bold indicate the values less than 5.00E-02.

Fig. 1. The workflow of establishing a monocyte DNA methylation "clock" using a feature selection method. YSCCS: Yale Stress Center Cohort Study; VACS: Veterans Aging Cohort Study; WIHS: Women's Interagency HIV Study; EWAS: Epigenome-Wide Association Studies; ENR: Elastic Net Regularization; EAA: Epigenetic Age Acceleration, the residuals of regressing DNA methylation age on chronological age; MD_{EAA}: The mean difference of EAA between different groups for phenotypes of interest; AMAR: Apparent Methylation Age Rate, the ratio of DNA methylated age to chronological age; MD_{AMAR}: The mean difference of AMAR between different groups for phenotypes of interest.
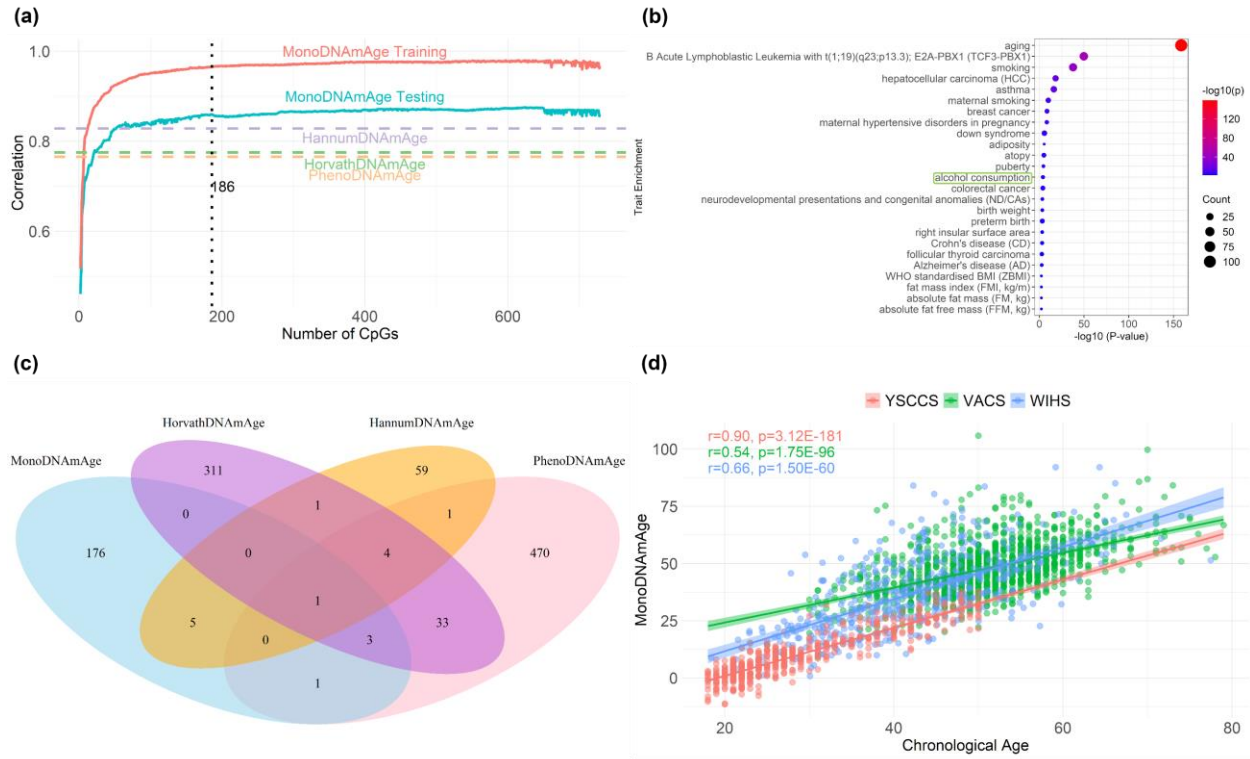
Fig. 2. MonoDNAmAge development and performance. (a) Feature selection using elastic net regularization (ENR) for estimating MonoDNAmAge. The figure shows the Pearson correlation coefficients for the training and the testing data sets from MESA. **(b)** Trait enrichment of the selected 186 CpGs for MonoDNAmAge. **(c)** Overlapping CpG sites among four epigenetic clocks. **(d)** Correlation between MonoDNAmAge and chronological age in three independent cohorts: Yale Stress Center Cohort Study (YSCCS); Veterans Aging Cohort Study (VACS); and Women's Interagency HIV Study (WIHS).
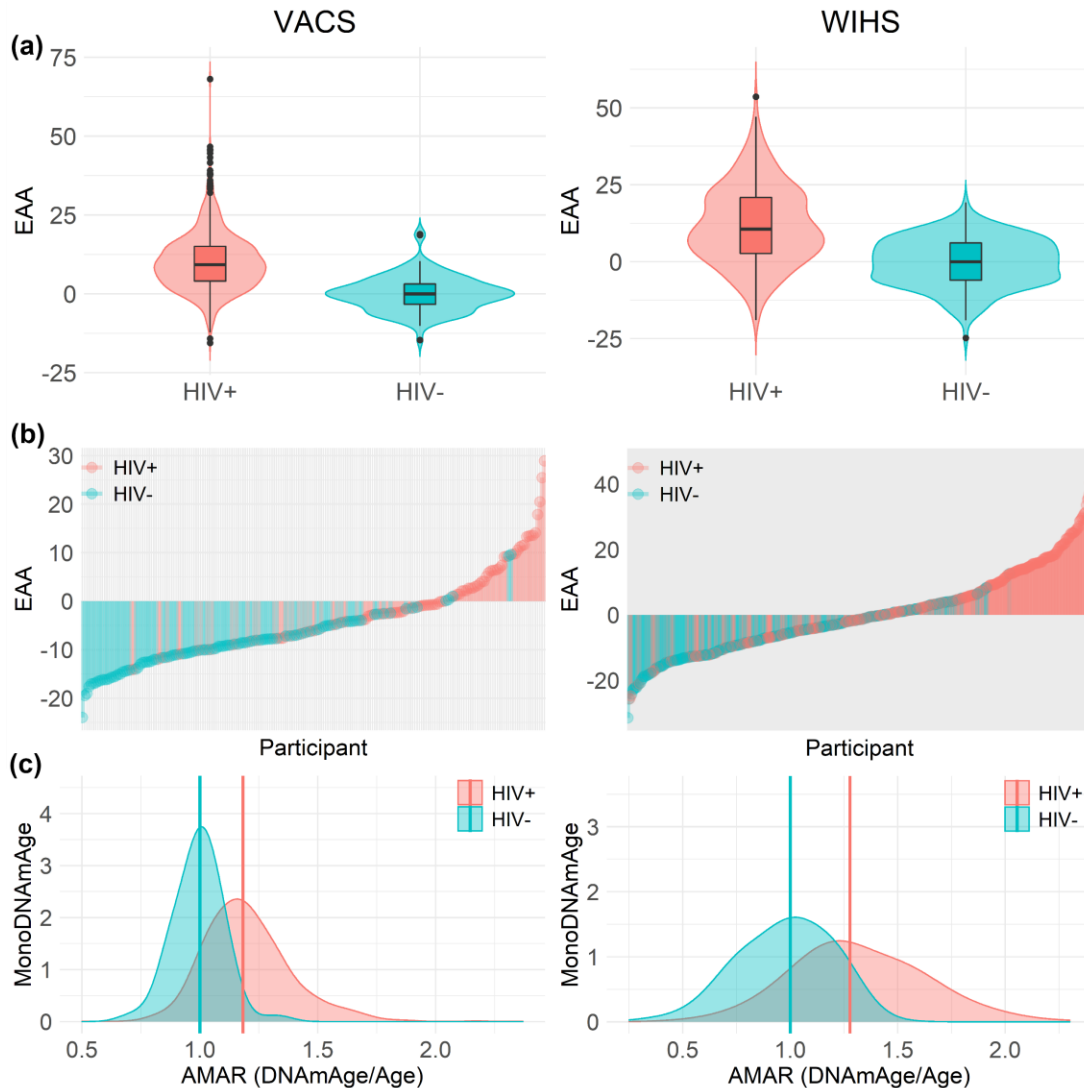
Fig. 3. Comparison of MonoDNAmAge-based estimations of biological aging for HIV-positive and HIV-negative participants. (a) Violin plots showing significant differences in the Epigenetic Age Acceleration (EAA, the residuals of regressing MonoDNAmAge on chronological age) between HIV-positive (HIV+) and HIV-negative (HIV-) groups. **(b)** Lollipop plots for EAA between HIV+ and HIV-. **(c)** Density plots for Apparent Methylation Age Rate (AMAR, the ratio of MonoDNAmAge to chronological age) between HIV+ and HIV-. VACS: Veterans Aging Cohort Study; WIHS: Women's Interagency HIV Study.
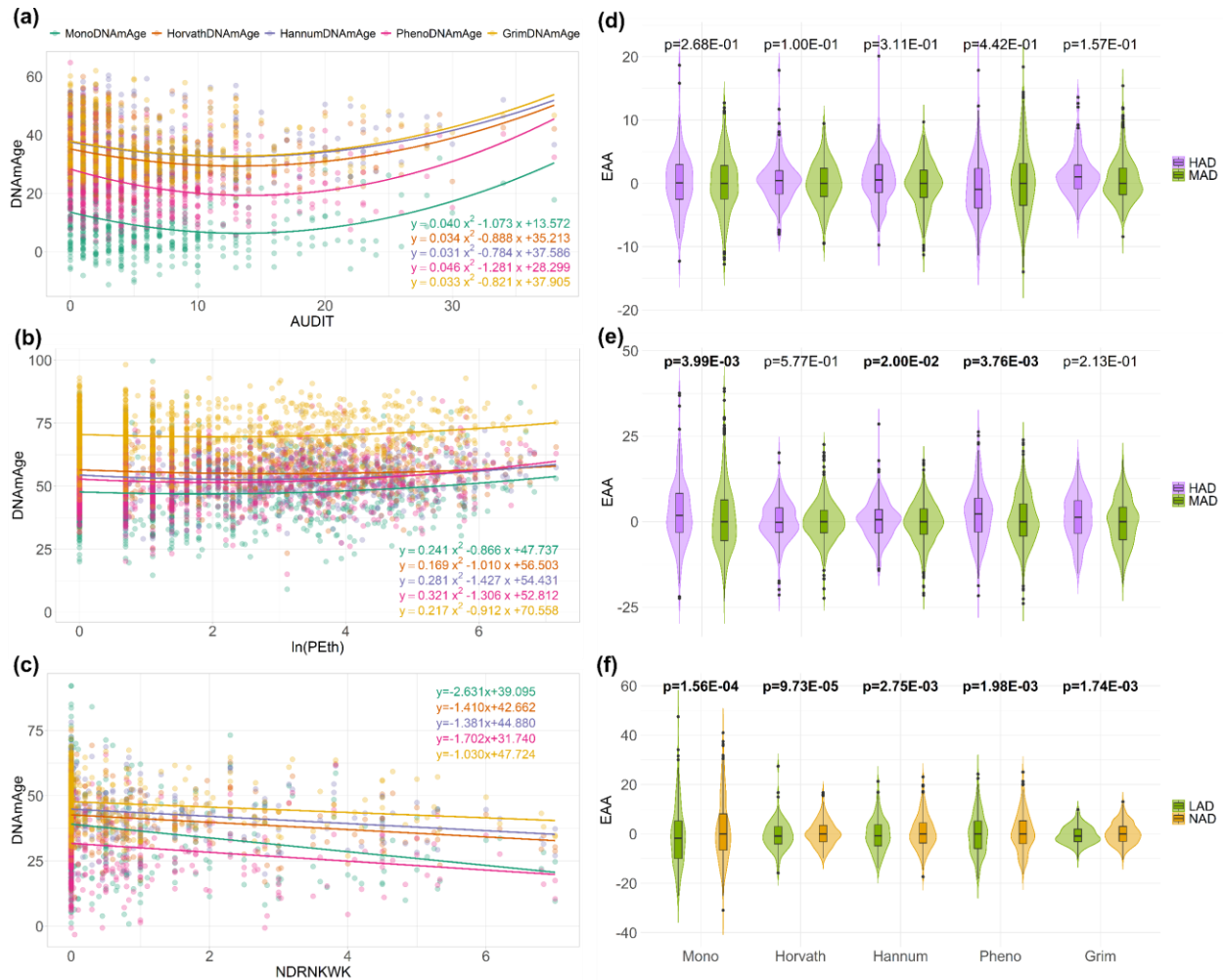
Fig. 4. Relationships between DNA methylation age (DNAmAge) based on different approaches and alcohol consumption. The three rows correspond to Yale Stress Center Cohort Study (YSCCS), Veterans Aging Cohort Study (VACS), and Women's Interagency HIV Study (WIHS). **(a)** The parabola shows the relationship between DNAmAge and the Alcohol Use Disorders Identification Test (AUDIT) score in the YSCCS. **(b)** The parabola shows the relationship between DNAmAge and the natural logarithm of Phosphatidylethanol (ln(PEth)) in the VACS. **(c)** The linear regression line shows the negative relationship between DNAmAge and the number of drinks per week (NDRNKWK) in the WIHS. **(d)-(f)** Violin plots for Epigenetic Age

Acceleration (EAA) (the residuals of regressing DNA methylation age on chronological age) for different groups of alcohol consumption in the three cohorts. **(d)** Comparison of heavy alcohol drinking (HAD) (AUDIT>=7 for men and AUDIT>=8 for women) and moderate alcohol drinking (MAD) (AUDIT<7 for men and AUDIT<8 for women) in the YSCCS. **(e)** Comparison of HAD (PEth>=20) and MAD (PEth<20) in the VACS. **(f)** Comparison of light alcohol drinking (LAD) (0<NDRNKWK<7) and non-alcohol drinking (NAD) (NDRNKWK=0) in the WIHS.
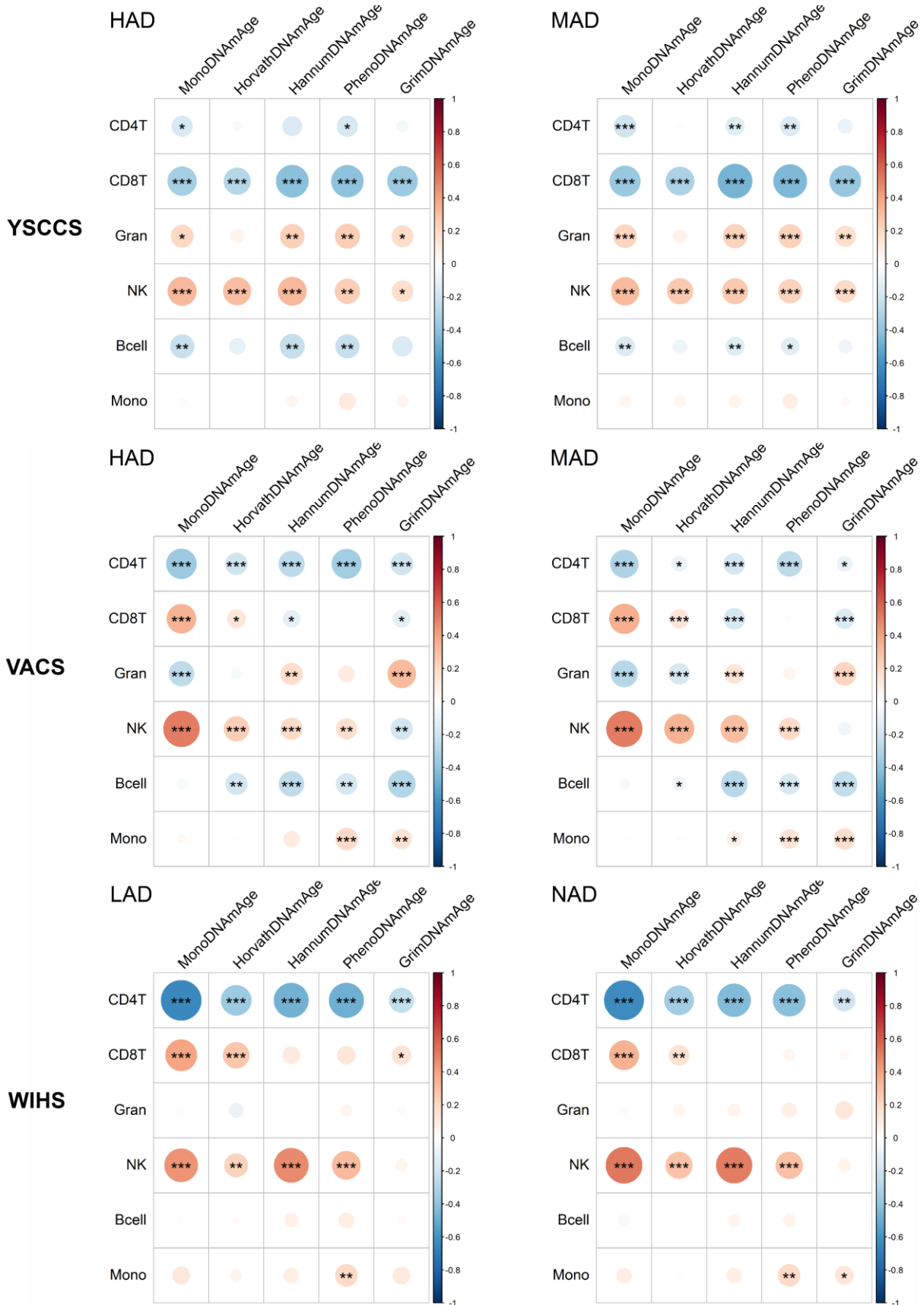
Fig. 5. Correlation between five epigenetic clocks and cell-type proportions.

Comparison of heavy alcohol drinkers (HAD) (AUDIT>=7 for men and AUDIT>=8 for women) and moderate alcohol drinkers (MAD) (AUDIT<7 for men and AUDIT<8 for women) in the Yale Stress Center Cohort Study (YSCCS); HAD (PEth>=20) and MAD (PEth<20) in the Veterans Aging Cohort Study (VACS); light alcohol drinkers (LAD) (0<NDRNKWK<7) and non-alcohol drinkers (NAD) (NDRNKWK=0) in the Women's Interagency HIV Study (WIHS). ***indicates p-value<1.00E-03; ** indicates p-value<1.00E-02; * indicates p-value<5.00E-02.

References

1.   Collaborators GBDA: **Alcohol use and burden for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016.** *Lancet* 2018, **392:**1015-1035.

2.   Strandberg TE, Strandberg AY, Saijonmaa O, Tilvis RS, Pitkala KH, Fyhrquist F: **Association between alcohol consumption in healthy midlife and telomere length in older men. The Helsinki Businessmen Study.** *Eur J Epidemiol* 2012, **27:**815-822.

3.   Shin C, Baik I: **Associations Between Alcohol Consumption and Leukocyte Telomere Length Modified by a Common Polymorphism of ALDH2.** *Alcohol Clin Exp Res* 2016, **40:**765-771.

4.   Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G: **The hallmarks of aging.** *Cell* 2013, **153:**1194-1217.

5.   Horvath S: **DNA methylation age of human tissues and cell types.** *Genome Biol* 2013, **14:**R115.

6.   Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan JB, Gao Y, et al: **Genome-wide methylation profiles reveal quantitative views of human aging rates.** *Mol Cell* 2013, **49:**359-367.

7.   Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, et al: **An epigenetic biomarker of aging for lifespan and healthspan.** *Aging (Albany NY)* 2018, **10:**573-591.

8.   Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, et al: **DNA methylation GrimAge strongly predicts lifespan and healthspan.** *Aging (Albany NY)* 2019, **11:**303-327.

9.   Horvath S, Levine AJ: **HIV-1 Infection Accelerates Age According to the Epigenetic Clock.** *J Infect Dis* 2015, **212:**1563-1573.

10.   Boulias K, Lieberman J, Greer EL: **An Epigenetic Clock Measures Accelerated Aging in Treated HIV Infection.** *Mol Cell* 2016, **62:**153-155.

11. Okazaki S, Otsuka I, Shinko Y, Horai T, Hirata T, Yamaki N, Sora I, Hishimoto A: **Epigenetic Clock Analysis in Children With Fetal Alcohol Spectrum Disorder.** *Alcohol Clin Exp Res* 2021, **45:**329-337.

12. Fiorito G, McCrory C, Robinson O, Carmeli C, Rosales CO, Zhang Y, Colicino E, Dugue PA, Artaud F, McKay GJ, et al: **Socioeconomic position, lifestyle habits and biomarkers of epigenetic aging: a multi-cohort analysis.** *Aging (Albany NY)* 2019, **11:**2045-2070.

13. Rosen AD, Robertson KD, Hlady RA, Muench C, Lee J, Philibert R, Horvath S, Kaminsky ZA, Lohoff FW: **DNA methylation age is accelerated in alcohol dependence.** *Transl Psychiatry* 2018, **8:**182.

14. Luo A, Jung J, Longley M, Rosoff DB, Charlet K, Muench C, Lee J, Hodgkinson CA, Goldman D, Horvath S, et al: **Epigenetic aging is accelerated in alcohol use disorder and regulated by genetic variation in APOL2.** *Neuropsychopharmacology* 2020, **45:**327-336.

15. Robinson O, Chadeau Hyam M, Karaman I, Climaco Pinto R, Ala-Korpela M, Handakas E, Fiorito G, Gao H, Heard A, Jarvelin MR, et al: **Determinants of accelerated metabolomic and epigenetic aging in a UK cohort.** *Aging Cell* 2020, **19:**e13149.

16. Quach A, Levine ME, Tanaka T, Lu AT, Chen BH, Ferrucci L, Ritz B, Bandinelli S, Neuhouser ML, Beasley JM, et al: **Epigenetic clock analysis of diet, exercise, education, and lifestyle factors.** *Aging (Albany NY)* 2017, **9:**419-446.

17. Sureshchandra S, Raus A, Jankeel A, Ligh BJK, Walter NAR, Newman N, Grant KA, Messaoudi I: **Dose-dependent effects of chronic alcohol drinking on peripheral immune responses.** *Sci Rep* 2019, **9:**7847.

18. Badia E, Sacanella E, Fernandez-Sola J, Nicolas JM, Antunez E, Rotilio D, de Gaetano G, Urbano-Marquez A, Estruch R: **Decreased tumor necrosis factor-induced adhesion of human monocytes to endothelial cells after moderate alcohol consumption.** *Am J Clin Nutr* 2004, **80:**225-230.

19. Szabo G: **Monocytes, alcohol use, and altered immunity.** *Alcohol Clin Exp Res* 1998, **22:**216S-219S.

20. Donnadieu-Rigole H, Mura T, Portales P, Duroux-Richard I, Bouthier M, Eliaou JF, Perney P, Apparailly F: **Effects of alcohol withdrawal on monocyte subset defects in chronic alcohol users.** *J Leukoc Biol* 2016, **100:**1191-1199.

21. Saare M, Tserel L, Haljasmagi L, Taalberg E, Peet N, Eimre M, Vetik R, Kingo K, Saks K, Tamm R, et al: **Monocytes present age-related changes in phospholipid concentration and decreased energy metabolism.** *Aging Cell* 2020, **19:**e13127.

22. Viel G, Boscolo-Berto R, Cecchetto G, Fais P, Nalesso A, Ferrara SD: **Phosphatidylethanol in blood as a marker of chronic alcohol use: a systematic review and meta-analysis.** *Int J Mol Sci* 2012, **13:**14788-14812.

23. Metcalf TU, Wilkinson PA, Cameron MJ, Ghneim K, Chiang C, Wertheimer AM, Hiscott JB, Nikolich-Zugich J, Haddad EK: **Human Monocyte Subsets Are Transcriptionally and Functionally Altered in Aging in Response to Pattern Recognition Receptor Agonists.** *J Immunol* 2017, **199:**1405-1417.

24. Reynolds LM, Taylor JR, Ding J, Lohman K, Johnson C, Siscovick D, Burke G, Post W, Shea S, Jacobs DR, Jr., et al: **Age-related variations in the methylome associated with gene expression in human monocytes and T cells.** *Nat Commun* 2014, **5:**5366.

25. Blaine SK, Nautiyal N, Hart R, Guarnaccia JB, Sinha R: **Craving, cortisol and behavioral alcohol motivation responses to stress and alcohol cue contexts and discrete cues in binge and non-binge drinkers.** *Addict Biol* 2019, **24:**1096-1108.

26. Justice AC, Dombrowski E, Conigliaro J, Fultz SL, Gibson D, Madenwald T, Goulet J, Simberkoff M, Butt AA, Rimland D, et al: **Veterans Aging Cohort Study (VACS): Overview and description.** *Med Care* 2006, **44:**S13-24.

27. Adimora AA, Ramirez C, Benning L, Greenblatt RM, Kempf MC, Tien PC, Kassaye SG, Anastos K, Cohen M, Minkoff H, et al: **Cohort Profile: The Women's Interagency HIV Study (WIHS).** *Int J Epidemiol* 2018, **47:**393-394i.

28. Bacon MC, von Wyl V, Alden C, Sharp G, Robison E, Hessol N, Gange S, Barranday Y, Holman S, Weber K, Young MA: **The Women's Interagency HIV**

Study: an observational cohort brings clinical sciences to the bench. *Clin Diagn Lab Immunol* 2005, **12:**1013-1019.

29.   Li M, Zou D, Li Z, Gao R, Sang J, Zhang Y, Li R, Xia L, Zhang T, Niu G, et al: **EWAS Atlas: a curated knowledgebase of epigenome-wide association studies.** *Nucleic Acids Res* 2019, **47:**D983-D988.

30.   Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, Lempicki RA: **DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists.** *Nucleic Acids Res* 2007, **35:**W169-175.

31.   Irina Shchukina1, Juhi Bagaitkar2,13, Oleg Shpynov1,3,13, Ekaterina Loginicheva1, Sofia Porter1,, Denis A. Mogilenko1 EW, Patrick Collins1, German Demidov   1,5,6, Mykyta Artomov   7,8, Konstantin Zaitsev1,12, Sviatoslav Sidorov9, Christina Camell9, Monika Bambouskova1, Laura Arthur1, Amanda Swain   1, Alexandra Panteleeva1, Aleksei Dievskii3, Evgeny Kurbatsky3, Petr Tsurinov1,3, Roman Chernyatchik   1,3, Vishwa Deep Dixit9, Marko Jovanovic4, Sheila A. Stewart   10,, Mark J. Daly7, 11, Sergey Dmitriev3, Eugene M. Oltz1 and Maxim N. Artyomov **Enhanced epigenetic profiling of classical human monocytes reveals a specific signature of healthy aging in the DNA methylome.** *Nature Aging* 2020, **Nov 23**.

32.   Chiva-Blanch G, Badimon L: **Benefits and Risks of Moderate Alcohol Consumption on Cardiovascular Disease: Current Findings and Controversies.** *Nutrients* 2019, **12**.

33.   Bagnardi V, Rota M, Botteri E, Tramacere I, Islami F, Fedirko V, Scotti L, Jenab M, Turati F, Pasquali E, et al: **Light alcohol drinking and cancer: a meta-analysis.** *Ann Oncol* 2013, **24:**301-308.

34.   Costanzo S, Di Castelnuovo A, Donati MB, Iacoviello L, de Gaetano G: **Alcohol consumption and mortality in patients with cardiovascular disease: a meta-analysis.** *J Am Coll Cardiol* 2010, **55:**1339-1347.

35.   Di Castelnuovo A, Costanzo S, Bagnardi V, Donati MB, Iacoviello L, de Gaetano G: **Alcohol dosing and total mortality in men and women: an updated meta-analysis of 34 prospective studies.** *Arch Intern Med* 2006, **166:**2437-2445.

36. Du D, Bruno R, Blizzard L, Venn A, Dwyer T, Smith KJ, Magnussen CG, Gall S: **The metabolomic signatures of alcohol consumption in young adults.** *Eur J Prev Cardiol* 2020, **27:**840-849.

37. Wurtz P, Cook S, Wang Q, Tiainen M, Tynkkynen T, Kangas AJ, Soininen P, Laitinen J, Viikari J, Kahonen M, et al: **Metabolic profiling of alcohol consumption in 9778 young adults.** *Int J Epidemiol* 2016, **45:**1493-1506.

38. Eyawo O, McGinnis KA, Justice AC, Fiellin DA, Hahn JA, Williams EC, Gordon AJ, Marshall BDL, Kraemer KL, Crystal S, et al: **Alcohol and Mortality: Combining Self-Reported (AUDIT-C) and Biomarker Detected (PEth) Alcohol Measures Among HIV Infected and Uninfected.** *J Acquir Immune Defic Syndr* 2018, **77:**135-143.

39. Zhang Q, Vallerga CL, Walker RM, Lin T, Henders AK, Montgomery GW, He J, Fan D, Fowdar J, Kennedy M, et al: **Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing.** *Genome Med* 2019, **11:**54.

40. Lee JH, Kong KA, Lee DH, Choi YH, Jung KY: **Validation and proposal for cut-off values of an abbreviated version of the Alcohol Use Disorder Identification Test using the Korean National Health and Nutrition Examination Survey.** *Clin Exp Emerg Med* 2018, **5:**113-119.

41. Liang X, Justice AC, So-Armah K, Krystal JH, Sinha R, Xu K: **DNA methylation signature on phosphatidylethanol, not on self-reported alcohol consumption, predicts hazardous alcohol consumption in two distinct populations.** *Mol Psychiatry* 2020.

42. Piano MR, Tiwari S, Nevoral L, Phillips SA: **Phosphatidylethanol Levels Are Elevated and Correlate Strongly with AUDIT Scores in Young Adult Binge Drinkers.** *Alcohol Alcohol* 2015, **50:**519-525.

43. Stewart SH, Reuben A, Brzezinski WA, Koch DG, Basile J, Randall PK, Miller PM: **Preliminary evaluation of phosphatidylethanol and alcohol consumption in patients with liver disease and hypertension.** *Alcohol Alcohol* 2009, **44:**464-467.

44. Adams WL, Barry KL, Fleming MF: **Screening for problem drinking in older primary care patients.** *JAMA* 1996, **276:**1964-1967.

45. Zhang X, Hu Y, Justice AC, Li B, Wang Z, Zhao H, Krystal JH, Xu K: **DNA methylation signatures of illicit drug injection and hepatitis C are associated with HIV frailty.** *Nat Commun* 2017, **8:**2243.

46. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT: **DNA methylation arrays as surrogate measures of cell mixture distribution.** *BMC Bioinformatics* 2012, **13:**86.

47. Jaffe AE, Irizarry RA: **Accounting for cellular heterogeneity is critical in epigenome-wide association studies.** *Genome Biol* 2014, **15:**R31.

48. Xu K, Zhang X, Wang Z, Hu Y, Sinha R: **Epigenome-wide association analysis revealed that SOCS3 methylation influences the effect of cumulative stress on obesity.** *Biol Psychol* 2018, **131:**63-71.

49. Zhang X, Justice AC, Hu Y, Wang Z, Zhao H, Wang G, Johnson EO, Emu B, Sutton RE, Krystal JH, Xu K: **Epigenome-wide differential DNA methylation between HIV-infected and uninfected individuals.** *Epigenetics* 2016, **11:**750-760.

50. Horvath S, Raj K: **DNA methylation-based biomarkers and the epigenetic clock theory of ageing.** *Nat Rev Genet* 2018, **19:**371-384.