Adapting Epigenetic Clocks for Cell-Free DNA High-Throughput Sequencing Data

Guangyu Li [1,6], Weilai Huang [2,6], Xuechen Zhao [2,6], Jun Wu [2], Yuanfang Guo [2], Liying Chen [3], Xingqi Cao [3], Zipeng Yang [2], Siqi Jiang [2], Bin Hu [2], Yijun Wang [4], Dinson Tan [2], Vicky Tong [2], Changdong Tang [2], Xiaowen Feng [2], Xiaomin Hu [5], Chuan Ouyang [2,*], Guangyu Zhou [1,2,*]

[1] Liangzhu Laboratory, Zhejiang University School of Medicine, Hangzhou, 310000, China

[2] Regenerative Bio Inc., Hangzhou, Zhejiang, China

[3] Department of General Practice, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

[4] School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, United States

[5] Department of Medical Research Center, Peking Union Medical College Hospital, Chinese Academy of Medical Science & Peking Union Medical College, Beijing, China

[6] These authors contributed equally: Guangyu Li, Weilai Huang, Xuechen Zhao

[*] Corresponding authors

E-mail: gavin.zhou@regen-bio.com; orcharry@gmail.com

Abstract

Cell-free DNA (cfDNA) methylation sequencing holds promise for developing epigenetic aging clocks. However, current clocks—primarily trained on array-based data—do not readily generalize to high-throughput sequencing (HTS) cfDNA profiles. Using datasets with technical replicates encompassing HTS data from both cfDNA and gDNA, alongside gDNA methylation array data, we systematically assessed factors influencing clock accuracy and reproducibility. We identified key strategies to overcome HTS-specific challenges: maintaining ≥10× mean target depth, applying elastic net regression with strong L2 regularization, and imputing unreliable

beta-values. Further, transfer learning more effectively corrected platform biases than DNA type-related biases, and it enhanced performance robustly across multiple independent cohorts in aging and disease-related applications. Our findings demonstrate that array-derived epigenetic clocks can be effectively adapted to cfDNA sequencing data. This work offers critical methodological insights and practical guidelines, advancing the feasibility of minimally invasive aging assessment using cfDNA.

Introduction

Recent studies advocated `biological age` (BA) over `chronological age` (CA, the time elapsed since birth) for a more accurate indicator of an individual's physiological aging state. BA is estimated using `Biomarkers of Aging` (BoAs), which are quantified via mathematical models known as aging clocks [1–3]. Among diverse BoAs ranging from clinical parameters [4,5] and telomere length [6] to various "omic's" data and facial images [7–11], epigenetic clocks—based on the methylome (DNA methylation profiles)—have distinguished themselves by their outstanding predictive power [4,9,12–20], achieving high accuracy even with linear models [4,9,12,19,21,22]. Due to their superior performance, epigenetic clocks are now widely used in health monitoring, age-related disease assessment, and longevity interventions studies [2,17,23–25].

The circulating cell-free DNA (cfDNA or ccfDNA), primarily found in body fluids, enabled minimally invasive testing for early cancer detection and disease monitoring [26–29]. Its utility in epigenetic aging clocks is now being actively explored [30–32]. However, utilizing cfDNA for age prediction requires adapting epigenetic clocks, which are predominantly trained with methylation arrays (e.g. Illumina 450K and EPIC/EPICv2) data. Array platforms exhibit limited sensitivity to low-abundance methylation changes, narrow dynamic range, and inability to capture intra-sample methylation heterogeneity [33]. Furthermore, the low yield of cfDNA (typically 0.1–10 ng/sample, compared to the recommended 200 ng input for arrays [34]) and its small fragment size present significant challenges for array-based platforms [35]. Even though methods such as preamplification can reduce the required input amount [36], the resulting potential distortion of beta-values still hampers array-based application. High-throughput sequencing (HTS), particularly targeted sequencing with hybridization probe capture, overcomes these limitations

[33,37–43]. HTS offers a technically robust alternative solution for cfDNA methylation profiling [44,45], thereby providing a viable path toward cfDNA-based epigenetic aging and disease detection.

However, current epigenetic clocks have not been fully adapted to HTS data, particularly for cfDNA samples. Current clocks are still predominantly trained on array-based genomic DNA (gDNA) methylation data, with only a few exceptions trained on HTS gDNA data (such as TIME-seq) [18,30,33,46,47]. Previous studies illustrated that direct application of array-based gDNA clocks on cfDNA data resulted in impaired performance [13,33,42,46]. Datasets that allow comprehensive analysis of such performance differences are still lacking especially for cfDNA, although similar resource exists for buccal gDNA [42]. This gap is critical because cfDNA exhibits unique fragmentation patterns and multi-tissue origins [30,32,48]. Furthermore, methylation datasets with technical replicates are still scarce. This hinders assessment of clock reproducibility and reliability, particularly as no such evaluation has been conducted for cfDNA-based epigenetic clock applications. Finally, as cfDNA remains a novel biomarker for epigenetic clock applications, to our best knowledge, practical guidelines for experimental design, data processing, model selection, and benchmarking are still lacking.

Adaptation of epigenetic clocks across different data types has been explored under various rationales; however, none have specifically addressed cfDNA applications. Batch correction and harmonization methods can mitigate beta-value discrepancies between methylation arrays [49–51], though often at the cost of introducing false positives or artifacts [52]. Domain adaptation approaches enabled the transfer of array data across tissues and cohorts [53,54], but their reliance on empirical estimation of beta-value distributions is incompatible with the dynamic nature of cfDNA. Some epigenetic clocks have been designed with data adaptability in mind by selecting features robust to tissue-specific variance [55], but they are unlikely to generalize well for pan-tissue age prediction. Deep learning-based clocks [15,56] may achieve similar adaptability when trained on sufficiently comprehensive datasets, but cfDNA methylation data generated for age prediction remain limited. Principal component analysis (PCA) preprocessing has been widely applied and recognized for improving clock reproducibility in both array- and HTS-based datasets [22,42,57], albeit with a trade-off of reduced interpretability [55]. In summary, a

comprehensive and robust adaptation of array-based epigenetic clocks for HTS-based data, particularly for cfDNA samples, is urgently needed.

Motivated by these knowledge gaps, we generated a dataset specifically designed for comparison between cfDNA/gDNA sample types and HTS/methylation array technologies. This dataset enabled systematic benchmark of reproducibility and accuracy of array-based epigenetic clocks on HTS-based cfDNA methylation data. Furthermore, we delved into exploration of experimental and data aspects that can affect epigenetic accuracy and reproducibility. Finally, we developed and evaluated multiple improvement techniques informed by both experimental and computational considerations, for establishing practical experimental guidelines and machine-learning strategies for adapting epigenetic clocks to diverse cfDNA-based aging and disease applications. Overall, this study contributes to the adaptation of epigenetic clocks for cfDNA methylation data, paves the road for future minimally invasive aging evaluation.

Results

Methylation profiling technologies and epigenetic clocks exhibited heterogeneous CpG coverage

To evaluate the adaptability of array-trained epigenetic clocks for HTS-based methylation data, we compared four profiling methods: two array-based methods—the Illumina Human Methylation Screening Array (MSA) and the Illumina Methylation EPIC v2.0 BeadArray (EPICv2)—and two targeted sequencing-based methods—the iGeneTech BisCap CpG Galaxy Panel (Galaxy) and the Twist Human Methylome Panel (Twist). These methods were selected to represent both established and emerging methylation profiling technologies (Methods). For comparative analysis, 24 out of the 165 participants in the Sir Run Run Shaw Hospital (SRRSH) cohort were randomly selected (the SRRSH-24 cohort) and profiled using all four methods on both plasma cfDNA and buffy coat gDNA, yielding six datasets: gDNA_MSA, gDNA_EPICv2, gDNA_Galaxy, gDNA_Twist, cfDNA_Galaxy, and cfDNA_Twist. Two technical replicates were included in each of the six datasets for assessment of reproducibility. The remaining 141 participants (the SRRSH-141 cohort) were reserved for validation. Fig. 1a shows the overall study design and workflow.

The heterogeneous CpG sets of the four methods (MSA, EPICv2, Galaxy and Twist) (0.28 to 4 million CpGs, Fig. S1a) lead to various degrees of feature coverage (Fig. S1b) of the 53 epigenetic clocks (Methods, Table S1). Twist showed the highest overall CpG coverage. Although Galaxy and EPICv2 share overlapping CpG sets by design, the discrepancy between reference and observed coverages was likely contributed from (reduced EPICv2 coverage) and off-target reads (increased Galaxy coverage). MSA exhibited limited coverage (< 70 %) in 25 of 53 clocks, potentially limiting its utility. Regarding feature coverage, two clock categories were identified among first- and second-generation clocks applicable to human adults: 20 Hi-Cov-All clocks with ≥ 80 % coverage across all four methods, and another 12 Hi-Cov-nonMSA clocks with ≥ 80 % coverage except for MSA (Fig. S1b, Table S1). Notably, six clocks (pchannum, pchorvath2013, pcphenoage, pcskinandblood, pcgrimage and zhangblup) were differently labelled as Hi-Repro despite their sufficiency in Hi-Cov-nonMSA criteria due to outstanding age prediction performance as discussed later.

Regarding the information carried by CpG loci, both Method-shared sites (the 136,437 CpGs shared by all four methods) and Clock-shared sites (the 2,727 CpGs shared by Hi-Cov-All clocks) encoded methylation signals to distinguish DNA type and method-induced variations (Fig. 1b-c, Fig. S1c-d). However, Clock-shared sites likely captured DNA-type information more effectively than Method-shared CpGs (Fig. S1e-f), implying that a proper selection of CpGs contributes to clock robustness against DNA-type variances.

Together, these results highlight that CpG coverage is a critical factor for epigenetic clock adaptation across DNA types and methylation profiling methods.

Reduced CpG consistency and ICC posed challenges in HTS-based age prediction consistency and reproducibility

An ideal methylation profiling method required both reproducibility (agreement between technical replicates) and consistency (agreement across other methods). This motivated us to assess the 6 SRRSH-24 datasets to compare reproducibility and consistency across DNA types and methylation profiling methods.

For reproducibility, we calculated per-CpG intraclass correlation coefficients (ICCs) [58] across datasets, which quantify and compare intra- versus inter-group variance. As expected, HTS-

based datasets (Galaxy and Twist) exhibited lower ICCs compared to array-based datasets (MSA and EPICv2) for both Method-shared CpGs (Fig. 1d) and Clock-shared CpGs (Fig. S1g). Consistent with a previous study [22], CpGs with low ICCs were enriched near extreme beta-values (near 0 or 1) regardless of methylation profiling methods (Fig. S2).

For consistency, we calculated pairwise Pearson's R between datasets using the beta-value means of Method-shared CpGs (Fig. 1e). Array-based datasets exhibited notably higher correlation (R = 0.9947), than HTS-based datasets (R = 0.9125 - 0.9781). A similar pattern was also observed in Clock-shared CpGs (Fig. S1h).

These findings suggested that HTS-based datasets in general exhibited lower consistency and reproducibility compared with array-based methods.

CpG reproducibility and consistency influenced clock reproducibility and consistency

With respect to reproducibility, array-based datasets show higher age prediction ICCs (0.469 – 0.999; median 0.950) relative to HTS-based datasets (-0.188 – 0.997; median 0.721) (Fig. 1g), aligning with their higher CpG ICCs as discussed previously. These observations are consistent with prior studies linking ICC to CpG reliability [22,42]. Six first- and second-generation clocks, including five PCA-preprocessed clocks [22] and zhangblup [21] were designated as Hi-Repro clocks for their markedly higher reproducibility (ICC ≥ 0.9) (Fig. 1f, Fig. S3) and lower replicate prediction discrepancy (Fig. S4) than other recognized clocks.

With respect to consistency, the datasets showing higher CpG consistency with each other also exhibited high age prediction ICCs (Fig. 1h). Surprisingly, high age prediction ICC between gDNA datasets was only observed with Hi-Repro clocks while not with other clock categories (Fig. 1h), indicating that Hi-Repro clocks are more consistent and resilient against method-related methylation variations. Notably, 17 out of the 53 clocks (including all Hi-Repro clocks) exhibited Pearson's R ≥ 0.7 (Fig. S5), with pcgrimage achieving the highest.

Collectively, these results confirm that CpG reproducibility and consistency critically impacts clock prediction reproducibility and consistency. Amongst, Hi-Repro clocks are preferable than others for applications with HTS-based datasets for their superior reproducibility and consistency.

HTS-based dataset proposed challenges in age prediction accuracy

The Hi-Repro, Hi-Cov-All, and Hi-Cov-nonMSA clocks were comprehensively evaluated for age prediction accuracy across the six SRRSH-24 datasets. The gDNA_EPICv2 dataset exhibited the lowest mean absolute error (MAE) (Fig. 2a&c, S6a), possibly due to its best match with the original training data of these clocks (gDNA samples measured via methylation arrays). CfDNA samples consistently exhibited higher MAEs and lower correlations with CA than gDNA samples, particularly with Hi-Cov-All clocks (Fig. 2b-c), emphasizing the necessity of clock adaptation between dataset derived from different DNA types or profiling methods. This consisted with a previous study which reported reduced reliability and accuracy when applying array-trained clocks on HTS-based data, although tested clocks were limited [42]. Among all clocks, zhangblup consistently demonstrated the lowest (MAE ≤ 5.2 years) across datasets except for gDNA_MSA (Fig. 2a, S6b), where it suffered from poor (25%) CpG coverage (Fig. S1b).

To test the reverse scenario, i.e. if HTS-trained clocks also suffer reduced accuracy on array-based data, we incorporated two independent whole-blood gDNA datasets: GSE245628 (N = 24, Illumina MethylationEPIC array) and GSE232346 (N = 24, HTS-based) [18]. As shown in Fig. 2d-e, the 4 array-trained clocks (hannum, horvath2013, skinandblood, dnamphenoage) and one HTS-trained clock (timeseq) displayed opposite accuracy patterns. This observation indicates that accuracy loss occurs both when applying array-trained clocks on HTS data, and also when applying HTS-trained clocks on array data [30,33,46].

In summary, these comparisons highlight the importance of matching data type and clock design, particularly regarding DNA source and measurement technology (array- vs. HTS-based). These findings further underscore the need to adapt epigenetic clocks for use across different DNA types and measurement platforms.

Epigenetic clock reproducibility indicates performance

Interestingly, significant correlations were identified between clock prediction ICCs and two other metrics, MAE and Pearson's R (correlation with CA). These results suggested that clock reproducibility serves as an alternative indicator of epigenetic clock performance (Fig. 2f-g).

Age prediction replicate difference (RD) is negatively associated with number of CpGs in linear regression-based epigenetic clocks

Noticing that all clocks in the Hi-Repro category had high numbers of input CpGs (≥ 70,000), we investigated 20 linear-architecture clocks from the Hi-Cov-All and Hi-Repro categories (collectively termed Hi-Cov/Repro-Linear, or HCRL clocks for short) to explore the relationship between CpG count and reproducibility. Negative associations between the number of CpGs and age prediction RD were observed, consistent across all 11 tested datasets, including the six SRRSH-24 datasets and five publicly available ones (GSE83944, N = 48; GSE247193, 22 replicated samples; GSE247195, N = 22; GSE247197, N = 46; GSE55763, 72 replicated samples) (Table 1, Fig. S7).

These results align with earlier findings showing that clocks with more CpGs exhibit greater robustness to replicate variance, particularly PCA-preprocessed clocks [22,42,59]. These findings motivated in-depth investigation of how epigenetic clock model complexity contributes to clock reproducibility, which can inform the development and selection of more reliable epigenetic clocks.

L2 regularization strength contributed to low RD in independent testing

We trained a series of elastic net models with varying CpG counts to further examine the relationship between CpG count and RD. The elastic net framework was selected for its widespread use in epigenetic age modeling. The variation in CpG count was achieved by tuning the two hyperparameters: $\alpha \in [0, +\infty)$ (the overall regularization strength) and $\lambda \in [0,1]$ (the mixing parameter that balances L1 and L2 penalties). All models were trained on GSE55763 (2711 replicated and unreplicated samples) and tested on the six SRRSH-24 datasets.

As shown in Fig. 3a, α exhibited a threshold effect around 0.01, beyond which the MAE increased sharply. We therefore focused on $\alpha \leq 0.01$ range in subsequent analyses to better follow model optimization behavior. The negative correlation between CpG count and RD was consistently reproduced in all test cases (Fig. 3b), confirming that this trend is intrinsic to elastic nets [22] rather than coincidences of the Hi-Repro clocks.

Results revealed that λ had a stronger and more significant association (Pearson's R^2 = 0.67 - 0.78, $p = 1.6\times10^{-47}$ - 7.3×10^{-33}) than α (Pearson's R^2 = 0.05 - 0.21, $p = 9.7\times10^{-9}$ - 8.2×10^{-3}) (Fig.

3c-d). Notably, the positive correlation between λ and RD indicates that skewing towards L2 penalty (relative to L1) reduces RD. This observation partially explains the negative association between CpG count and RD, as L2 penalty (ridge regression) tends to retain a large number of small coefficients, while L1 (LASSO) favors sparse coefficients.

Conclusively, these results suggest that elastic net models emphasizing L2 regularization (approaching the limiting case of ridge regression) enhance epigenetic clock reproducibility against replicate-level variability.

Ensuring epigenetic clock accuracy and reproducibility requires 10× minimum, 20× recommended mean target depth

In HTS-based datasets, the beta-value of each CpG is calculated as the ratio of two integers: methylation counts divided by total counts. The sequencing depth hence introduces an additional source of variance that can affect observed beta-values, and further, clock accuracy and reproducibility. Two potential effects were examined in this study: (1) The stochasticity in observed beta value using a binomial model; and (2) The resolution limit posed by presenting both methylation and total counts as integers. Two sequencing depth thresholds at approximately 20× and 10× were identified respectively, above which MAE and RD stabilized around their baseline value in a given dataset (Fig. S8).

To validate these findings under more practical conditions, we conducted randomized downsampling tests on five high-depth samples from the SRRSH-141 cohort (with ≥ 50 Gb). MAE convergence was achieved at approximately 10-20× coverage (Fig. 4a) across the 20 HCRL clocks as well as 12 others (Fig. S9). Notably, the fraction of CpGs with ≥ 20× coverage increased rapidly under 20 - 30× downsampling depth, but slowed thereafter (Fig. 4b, Fig. S9). Taken together, these results suggest a minimum mean target depth of 10×, optimally ≥ 20×, for HTS-based epigenetic age prediction. All our four SRRSH-24 HTS-based datasets and the SRRSH-141 dataset generated in this study comply with this standard (Fig. S10). In addition, this standard is comparable with a previous recommendation for WGBS, recommending 2 replicates with combined 30× coverage using a different rationale [60].

HTS-based methylation datasets are more prone to unreliable beta-values than methylation array datasets

Previous studies and our earlier findings (Fig. S2) both suggested that CpGs with beta-values near 0 or 1 are unreliable, as reflected by their low ICCs. Unfortunately, HTS-based datasets are particularly susceptible to their occurrence. Across eight tested datasets in Fig. S11, 18.4 - 61.3 % of CpGs (median 32.9 %) exhibited beta-values exactly equal to 0 or 1 in at least one sample. Approximately 1/3 to 3/4 of them were associated with low sequencing depth (0 depth for missing), except for the Buccal dataset due to explicit filtering (≤10× coverage) [42].

Further analyses revealed that these extreme beta-values are likely non-biological artifacts, as they displayed distinct depth distributions compared with CpGs having non-extreme values (Fig. S12-S13). The observation that a significant proportion of these CpGs were dataset-wise missing further supported this hypothesis (Fig. S14). Importantly, these CpGs overlapped with 2.0 - 34.7 % of the features in the 20 HCRL clocks (Fig. 4c, Table S3), although their impact on clock accuracy remains unclear. These observations highlight the widespread presence of artifactual beta-values in HTS-based data and motivated us to explore strategies for evaluating and improving model performance in the light of these beta-values.

Imputation on unreliable beta-values improved model prediction reproducibility

Multiple imputation strategies to correct unreliable beta-values were tested at varying cut-off depths ranging from 0 to 20×, assessed by age prediction MAE and RD changes (ΔMAE and ΔRD, negative for improvement and positive for deterioration). The cut-off depth defines the depth threshold below which beta-values are marked for imputation.

We first tested the K-nearest neighbor (KNN) method, one of the most widely used imputation approaches, revealing possible distinct responses to imputation between cfDNA and gDNA samples. For cfDNA, 1 - 5 years of MAE improvements were observed, whereas no significant changes in gDNA regardless of the capture panel (Twist or Galaxy) (Fig. 4d). For RD, improvements were consistently observed in both datasets across all cut-off depths (Fig. 4e), being more effective in cfDNA (0.7 - 2.1 years) than gDNA (0.7 - 1.1 years). These RD improvements likely stem from correcting CpGs with unreliable beta-values in one replicate but not in the other.

Other common imputation techniques, mean, median, and 0-filling, were also assessed at 20× cut-off (Fig. 4f-g). Mean and median produced comparable outcomes to KNN, resulting in overall significant RD improvements, accompanied by significant MAE deterioration only in cfDNA. In contrast, 0-filling caused no significant change in either metric, coincide with a previous report [42]. This distinction is unsurprising as 0-fill doesn't infer missing values from other observed data. Interestingly, while performance change varied across clocks (Fig. S15), zhangblup was the only clock with consistent improvements across all four datasets.

Collectively, these results indicate that KNN, mean, and median imputations serve as conditional strategies, providing reliable RD improvements particularly for gDNA samples due to unlikeliness of simultaneous MAE deterioration.

Transfer learning significantly reduced age prediction MAE

We tested distillation as a transfer learning (TL) technique for adaptation of epigenetic clocks to HTS-based methylation data. Distillation was chosen for its accessibility, as it can be implemented for nearly any existing or future epigenetic clock without requiring access to the original algorithm, parameters, or training data. In this study, the original model (teacher) outputs were incorporated as an additional feature to train an elastic net (student), which was then used for age prediction and validation. All 20 HCRL clocks were individually tested as teachers. In addition, PCA (fitted independently from GSE55763) was performed before TL to reduce potential overfitting and training due to limited sample size compared to CpG counts.

The 5-fold cross-validation results in Fig. 5a demonstrated substantial MAE improvements in a subset of HCRL clocks covering clocks belong to 1st, 2nd generations, and the previously identified best (zhangblup) as teachers. The effect on RD, however, varied by teacher model (Fig. 5b). Nonetheless, the varying range of mean RD across clocks decreased, suggesting that distillation tends to normalize RD performance, improving clocks with high initial RD (e.g., dnamphenoage) while reducing those with low initial RD (e.g., zhangblup). Overall, distillation showed consistent MAE improvements, promising to be a highly effective adaptation strategy for enhancing accuracy. However, its effects remained heterogeneous per clock, calling for careful consideration based on the specific reproducibility requirements of the application.

We further evaluated student model performance on datasets whose DNA type or methylation profiling method differed from those used during adaptation. As shown in Fig. 5c, prediction accuracies were significantly higher only when the testing dataset shared the same DNA type as the adaptation data regardless of methylation platforms, but not otherwise. This suggests that distillation corrects platform-related variation more effectively than DNA type-specific signals, implying potential applications on HTS-based datasets from different capture panels without further tuning as long as the DNA type consists.

Validation of transfer learning as a promising tool for epigenetic clocks adaptation

We then systematically evaluated all eight possible combinations of three proposed strategies for adapting 20 HCRL clocks to HTS-based data: depth-based filtering (DF), unreliable beta-value imputation (IM), and TL, sequentially applied in DF → IM → TL order. Evaluation on the Buccal cohort (N = 100) showed consistent reductions in MAE and RD across all strategies incorporating TL, indicating that TL is a particularly effective adaptation step (Fig. S16). Notably, TL alone emerged as the optimal overall strategy, yielding MAE reductions of 1.5 - 42.0 years (median 10.3) and increases in CA correlation of 0.02 - 0.78 (median 0.15) in Pearson's R (Fig. 6a). Detailed improvements per clock are shown in Fig. S17.

Because CA is an imperfect proxy for biological aging, we further assessed adaptation performance against PhenoAge [4] in the SRRSH-141 cohort. PhenoAge provides a clinically oriented measure of aging by integrating CA with nine plasma biomarkers. Consistent with the Buccal analysis, reaffirmed the effectiveness of TL in improving concordance between BA and CA (Fig. S18), as well as with PhenoAge (Fig. S19). In this case study, the combined IM+TL strategy performed best against PhenoAge, achieving MAE reductions of 1.0 - 23.4 years (median 6.6) and Pearson's R improvements of 0.03 - 0.77 (median 0.17) (Fig. 6b, Fig. S20). Taken together, these findings highlight TL as a critical component for adapting epigenetic clocks to HTS-based data, enabling improved alignment with both CA-based and clinically inferred aging metrics.

Finally, we evaluated the proposed adaptation strategies in a disease detection setting using a public cfDNA methylation dataset comprising 12 normal controls and 12 amyotrophic lateral sclerosis (ALS) samples (GSE164600). Previous studies have shown that epigenetic delta age

(the deviation between BA and CA) is a potential biomarker for ALS [61,62]. Fig. 6c illustrates delta age changes before and after IM+TL, with the median delta ages of controls prior to adaptation as references. Despite relatively low CpG coverage (46.5 - 73.9%, median 58.8%), increased Jensen-Shannon divergence (JSD) between diseased and normal samples was observed for a subset of clocks, most notably all PCA-preprocessed clocks (Fig. 6c-d). Because higher JSD indicates more distinct and disjoint distributions between diseased and healthy samples, it is unsurprising that clocks with largest ΔJSDs collectively improved disease classification performance (Fig. 6e). Classification was performed using a support vector machine with a linear kernel. The highest area under the curve (AUC) was achieved using delta age from as few as the top four adapted clocks, yielding an AUC improvement of 0.125 relative to the non-adapted inputs (Fig. 6f). Collectively, these results demonstrate that IM+TL adaptation enhances the utility of epigenetic clocks for disease biomarker analysis.

Discussion

In this study, we generated a unique, comprehensive dataset encompassing both array- and HTS-based methylation data. Public datasets including both replicate measurements and cross-platform profiling are currently scarce. While Shokhirev et al. [42] focused on buccal gDNA, our dataset extended assessments regarding different DNA types and methylation profiling technologies. The six SRRSH-24 datasets generated in this study expand the limited resources available for systematic comparisons between array- and HTS-based methylation technologies. Moreover, they contribute valuable replicate data critical for evaluating the reproducibility of methylation profiling methods and epigenetic clocks.

Imperfect reproducibility of epigenetic measurements inherently limits the accuracy of age prediction by epigenetic clocks. Conventional accuracy metrics such as MAE or RMSE aggregate two sources of variation: (1) the intrinsic quantitative relationship between methylation features and biological/chronological age, and (2) the stochastic effects specific to sample measurements, Consequently, these metrics already encode reproducibility to some extent. Although these two components can be mathematically modeled, the precise biological disentanglement and quantification of their respective contributions remain challenging. Coupled with finite sample size, this modeling imperfection imposes an upper bound on the practical reproducibility of a clock, and thereby, an upper bound on its achievable predictive accuracy.

Our results show that HTS-based data pose additional challenges for epigenetic clocks as the suboptimal reliability and reproducibility of the HTS input signal can attenuate both the accuracy and reproducibility of the age predictions. In addition, we found that L2 regularization in elastic net clocks empirically enhances reproducibility. Related approaches, including PCA-preprocessed clocks and zhangblup (see Supplementary Discussion), achieved mean RD as low as 1.6 years across HTS-based datasets (Fig. S3, Fig. S4a). We therefore recommend L2 regularization as a practical and promising strategy for improving reproducibility of linear-architecture clocks.

HTS technologies utilize distinct signal processing principles compared to methylation arrays [37,63,64], which can affect CpG-level reliability and reproducibility. Among tested adaptation strategies, imputing unreliable beta-values consistently improved clock reproducibility, but causing accuracy loss particularly in cfDNA samples. In contrast, gDNA samples were minimally affected by this accuracy loss, likely due to their higher cellular homogeneity, a factor known to enhance imputation accuracy [41]. Although not tested, these findings suggest that performing cell-type deconvolution prior to imputation could mitigate the MAE trade-offs observed in cfDNA samples. Additionally, statistical modeling and downsampling experiments suggest ≥ 10× mean HTS target depth (ideally ≥ 20×) for epigenetic clock applications. These recommendations, comparable to previous WGBS guidelines [60,65], aim to support more robust and standardized implementations of epigenetic clocks.

Finally, our systematic evaluation and validation of multiple adaptation strategies (DF, IM and TL) demonstrated that TL offered the most critical component of epigenetic clock adaptation. This effect may stem from the correction of systematic pattern differences. Particularly, the trained correction model can potentially be applied across capture panels without retraining when the DNA type consists. While previous studies have focused on comparability between methylation arrays [66–68], our results extend these observations to different capture panels and gDNA/cfDNA types. Efficacy of TL-like dataset harmonization was also recently reported although from different rationales than our proposed methods [54,69]. Distillation, the TL method tested in this study, shares characteristics with error correction and model recalibration techniques. A recent study demonstrated that model retraining and recalibration can compensate for the missing probes in array data [49], and our study effectively extends this utility to HTS-

based data. Lastly, we validated the adaptation strategies with clinically derived aging status and disease classification applications. These findings underscore future development and refinement of epigenetic clock adaptation methods for cfDNA applications.

Our study has several limitations. The type and size of our datasets prevented assessment of aging biomarkers beyond methylation and limited exploration of novel epigenetic clocks [14,16,70,71]. Due to the limited sample size and lack of data, no sex- or ethnicity/race-related effects was evaluated. Besides, all datasets were cross-sectional without longitudinal and intervention-oriented analysis, precluding insights into temporal methylation dynamics. Lacking these analyses may hamper the generalization of results found this study, although our dataset sizes are comparable among studies with similar goals. Primarily, we focused on linear relationships between methylation data and clock predictions to balance result interpretability and model complexity. While linear models can achieve high accuracy [9,12,19,21,22], the true biological relationships between methylation and aging may be complex. Dataset size also limited our ability to investigate tissue heterogeneity in cfDNA, which is important for validating our hypotheses about imputation accuracy. For transfer learning, we only tested distillation under the consideration of minimal requirements of prior knowledge about model structure, parameters, and original training data; consequently, alternative methods such as regularization sharing [72], transfer component analysis [73], or deep learning-based approaches [74] remain unexplored. In addition, PCA preprocessing prior to TL was applied to improve training efficiency and reduce overfitting, though it may independently influence TL performance [22,42,75]. The effects are not adjusted by confounders such as lifestyle factors due to the lack of related information, which may limit the explanation of mechanistic insights. Finally, HTS-based cfDNA methylation datasets encompassing both diseased and normal samples and with sufficient sequencing depth are still scarce, which hampered further generalizability analysis of our proposed adaptation methods.

Overall, our study systematically compared the characteristics of HTS-based and array-based methylation data using dedicated datasets. We tested and validated three machine learning techniques to adapt epigenetic clocks for HTS data, thereby demonstrating these clocks' broad applicability in aging and disease classification applications. We also provided crucial technical

recommendations and experimental guidelines, significantly advancing the field for future HTS-based and cfDNA-oriented epigenetic clock applications.

Methods

The SRRSH cohorts and ethics declaration

The ethics committee of the Sir Run Run Shaw Hospital (SRRSH, affiliated with Zhejiang University School of Medicine) approved this study (approval number 2024-0547). Written informed consent was obtained from all participants.

The recruitment and data collection for all volunteers in the cohort were conducted under strict protocols. The following criteria were applied for recruiting eligible volunteers: (1) healthy adults aged 20 - 84 years; (2) Within the past 5 years the subject has completed at least two routine health evaluations that included complete blood count, urinalysis, fasting plasma glucose, lipid profile, liver-function tests, renal-function tests, and at least one set of imaging studies: abdominal ultrasound (liver, gall-bladder, pancreas, spleen, both kidneys), chest radiograph, thyroid ultrasound, prostate ultrasound (men), uterus and adnexa ultrasound (women). All reported results must be within normal limits and free of any diagnosed disease; (3) no smoking history within 5 years before screening and agreement to remain abstinent from tobacco for 1 year after consent; (4) body-mass index 18.5-28.0 kg/m^2; (5) without acute infection or active inflammation within 14 days before screening; (6) without cancer-related signs or symptoms (e.g., hemoptysis, rectal bleeding, altered bowel habit, hematuria, dysphagia, palpable mass) within 30 days before screening; (7) without previous organ transplantation or any non-autologous bone-marrow or stem-cell transplantation; (8) without blood transfusion within 7 days before screening; (9) without use of therapeutic drugs for any illness (nutritional supplements excluded) within 30 days before screening; (10) without any surgical procedure within 5 years before screening; (11) pregnant or breastfeeding women were excluded.

Methylation array and capture panel selection

Four methylation profiling methods were included in this study, including two array-based methods and two HTS-based methods with different hybridization probe capture panels: Illumina

Human Methylation Screening Array (referred to as MSA), Illumina Methylation EPIC v2.0 BeadArray (referred to as EPICv2), targeted sequencing with iGeneTech BisCap CpG Galaxy Panel for capturing (referred to as Galaxy) and targeted sequencing with Twist Human Methylome Panel for capturing (referred to as Twist). EPICv2 was included because it encompasses the largest CpG set among currently available methylation arrays, which serve as the basis for training most existing epigenetic clocks. MSA is a newer methylation array with a smaller CpG set than EPICv2, yet is claimed by the manufacturer to be suitable for epigenetic clock applications. Galaxy was included because it is designed as an alternative to EPICv2, covering over 95% of its CpG sites. Lastly, Twist was included as a mature solution for HTS-based methylation profiling with appropriate CpG coverage in the current marketplace. Overall, these methods were selected to cover mature, widely used, and emerging methods in both methylation array and HTS categories.

Experimental design

In total, 165 participants (100 % East Asian / Pacific Islander; aged 20 - 77 years old; 41.8 % female) were recruited as the SRRSH cohort. Among which, 24 subjects were randomly selected (22 - 77 years old; 54.2 % female) for investigating and evaluating the adaptation of epigenetic clocks for cfDNA-based methylation data, referred to as the SRRSH-24 cohort. Though no sample size calculations were conducted *a priori*, a previous study suggested a similar sample size is sufficient to identify 80% of reliable probes [67]. Their cfDNA was extracted for targeted DNA methylation sequencing, and buffy coat gDNA was extracted for both methylation profiling. This resulted in six datasets: gDNA_MSA, gDNA_EPICv2, gDNA_Galaxy, gDNA_Twist, cfDNA_Galaxy, and cfDNA_Twist. Two technical replicates per sample were generated via independent library preparations from the same DNA aliquot. At last, the SRRSH-141 dataset contained Twist-profiled plasma cfDNA methylation data for the remaining 141 participants. In addition, nine phenoage-related [4] plasma biomarkers were collected from the SRRSH-141 cohort, namely albumin, creatinine, glucose, CRP, lymphocyte percent, mean cell volume, red blood cell distribution width, alkaline phosphatase and white blood cell count.

Epigenetic clock selection

A total of 53 epigenetic clocks were included in this study (Table S1). Although we comprehensively surveyed available epigenetic clocks, some clocks were excluded for their

incompatibility with our data. We used 47 clocks integrated into the `pyaging` package (version 0.1.27) [76] for convenience, which greatly simplified multiple clock predictions due to a uniform interface. Six GP-age clocks (https://github.com/mirivar/GP-age, version 1.0) were also included. In addition, TIME-seq (https://github.com/patricktgriffin/TIME-Seq, Commitment 140622f), was included as a representative of clocks trained by HTS data [18].

Biological sample pretreatment and storage

An 8-10 mL whole-blood sample was collected from each volunteer via intravenous collection, after an overnight fast. One anticoagulant tube (Streck, Cell-Free DNA BCT 10ml) was immediately transported to the laboratory for subsequent treatment. The blood was centrifuged at 4 ˚C (1900× g, 10 min) to obtain plasma and buffy coat. All biological samples were stored at -80 ˚C for long-term preservation.

DNA extraction

Cell-free DNA was extracted from 1.0 - 1.5 mL plasma using the QIAamp MinElute ccfDNA Mini Kit (Qiagen, Cat# 55204) following the manufacturer's instructions. Genomic DNA was isolated from approximately 500 μL of thawed buffy coat using TIANamp Genomic DNA Kit (TIANGEN, Cat# DP304) following the manufacture's protocol. The concentration of purified DNA was quantified using a Thermo Fisher Qubit 4.0 fluorometer (Qubit™ dsDNA HS Assay Kit, Cat# Q32854).

Library preparation and sequencing

For genomic DNA, 0.5% (by mass) of unmethylated lambda DNA and 0.05% (by mass) of CpG-methylated pUC19 DNA were spiked into 400-800 nanograms of buffy coat genomic DNA, followed by dilution to 130 μL of 10 mM Tris 0.1 mM EDTA, pH 8.0, following the manufacturer's recommendations. Internal control DNA underwent the same fragmentation procedure as gDNA samples. The DNA mixture was then transferred to a Covaris microTUBE and sheared to an average size of 100-200 bp (peak = 150bp) using a Covaris™ S220 focused ultrasonicator (Covaris, MA, USA).

For cfDNA, unmethylated lambda DNA and CpG methylated pUC19 DNA were respectively sheared to an average size of 100-200 bp (peak = 150bp) in advance. Next, 0.5% (by mass) of

fragmented unmethylated lambda DNA and 0.05% (by mass) of fragmented CpG methylated pUC19 DNA were added into 1-10 ng cfDNA samples.

Both cfDNA and sonicated gDNA samples, along with spike-in internal control DNA, were both used to construct EM-Seq libraries with the NEBNext Enzymatic Methyl-seq Kit (New England Biolabs, Cat# E7120) following the manufacturer's instructions. Libraries were quantified using a Thermo Fisher Qubit 4.0 fluorometer. Then capture-based human methylome enrichment was performed according to Twist's (South San Francisco, USA) protocols for their Twist Human Methylome Panel (Cat# 105521), or according to iGeneTech's (Beijing, China) protocols for their BisCap CpG Galaxy Panel (Cat# PE4000175). The final libraries were analyzed and quantified using TapeStation (Agilent Technologies). The whole-genome libraries were 150 bp paired-end sequenced using the Illumina NovaSeq X Plus sequencer on a 10B or 25B flow cell (Illumina) with 12.5 % PhiX spiked-in, to an average yield of 36.3Gbp raw base per library (Fig S15).

EPICv2 and MSA

Buffy coat genomic DNA samples were sent to Novogene (Tianjin, China) to perform Illumina Infinium MethylationEPIC v2.0 BeadChip array (EPICv2; 937,691 probes), or sent to WeGene (Shenzhen, China) to perform Infinium Methylation Screening array (MSA; 284,318 probes). For each technical replicate sample, 500 nanograms (for EPICv2) or 200 nanograms (for MSA) of genomic DNA were bisulfite treated using the EZ DNA Methylation Kit (Zymo Research, Cat# D5001) following the manufacturer's recommendations for Infinium assays.

As the EPICv2 and MSA arrays have a capacity of 8 and 48 samples per array, respectively, we included two replicates (2 replicates × 24 gDNA samples = 48 positions).

Methylation array data processing

Raw IDAT files from both Infinium MethylationEPIC v2.0 array and Infinium Methylation Screening Array (MSA) were processed using the `minfi` R package (version 1.54.1) according to established protocols. A rigorous quality control and filtering pipeline was applied: probes located on sex chromosomes were excluded, as well as those with a detection p-value greater than 0.01 regarding as unreliable signals. Furthermore, probes known to harbor single nucleotide polymorphisms (SNPs) at or near the CpG site, or those that did not uniquely map to the

reference genome, were also removed. Following these filtering steps, the raw intensity data underwent intra-array normalization using the `Noob` method. Finally, beta-values were calculated as the primary methylation metric and prepared for downstream quantitative analyses.

High-throughput sequence (HTS) methylation data processing

Prior to downstream processing, raw paired-end FASTQ files for all samples were downsampled to standardize sequencing depth at approximately 30 Gbp (1 Gbp = 10^9 base pairs). This step was performed to mitigate potential biases stemming from varying sequencing depths and to ensure fair comparisons among samples. Subsequently, these trimmed files underwent adapter removal using `fastp` (version 0.23.4) and were aligned to a comprehensive reference genome, comprising GRCh38, pUC19, and phage lambda sequences, using `bwa-meth` (version 0.2.7). Alignment files were then sorted and indexed using `Samtools` (version 1.21), followed by the identification and marking of duplicate reads using `Samblaster` (version 0.1.26). Methylation levels across CpG contexts were determined and extracted using the mbias and extract modules of `MethylDackel` (version 0.6.1).

Genomic coordinate alignment and probe annotation

Manifest files for the MSA, HM450K, EPICv1, and EPICv2 arrays were obtained from the manufacturers' official websites. Since our analysis was based on the hg38 reference genome, genomic coordinates for the HM450K probes were converted from hg19 to hg38 using the UCSC `LiftOver` tool. The four manifest files were then merged to create a comprehensive probe-to-coordinate reference file.

For sequencing data (Twist and Galaxy), the per-sample output files from `MethylDackel` (https://github.com/dpryan79/MethylDackel, version 0.6.1) were processed. Genomic coordinates and beta-values were extracted for each CpG sites and then merged using `mergeContext` function in `MethyDackel`. These files were subsequently merged into a matrix using `bedtools unionbedg` (version 2.31.1). Finally, the `bedtools intersect` function was employed to annotate the sequencing-derived CpG sites with corresponding probe IDs from the reference file, based on overlapping genomic coordinates.

Dataset benchmarking calculations

Principal component analysis (PCA) was performed using `prcomp()` from R version 4.5.1 (save below). Pearson correlation analysis was performed using `cor()`. Intraclass correlation coefficient (ICC) was calculated using `icc()` from the `irrR` package (version 0.84.1) based on a single-rater, absolute agreement, two-way random-effects model. Missing values (`NaN`) were removed before these analyses. Benchmarking results were visualized using R packages `ggplot2` (version 4.0.0), `ComplexHeatmap` (version 2.22.0), and `linkET` (version 0.0.7.4) respectively.

CpG coverage analyses were conducted in two levels: (1) across platform; and (2) clock CpGs covered by each platform. Reference CpG sites for each platform were extracted from manufacturer-provided datasheets. In addition, for MSA and EPICv2 arrays, only CpGs with defined genomic coordinates were retained. Input CpGs for each of the 53 clocks were extracted from clock models respectively. Results were visualized using R functions `UpSet()` (`ComplexHeatmap` version 2.22.0) and `pheatmap()` (`pheatmap` version 1.0.13).

Statistical analysis

The comparisons between means and medians of two groups of data were conducted using permutation test (implemented by custom script) by randomly permuting observations between the two groups and comparing the mean/median after permutation with the observed value. The p-values are calculated as the fraction of permutation rounds with its mean/median after permutation being greater than the observed value. Permutation tests make no assumptions of distributions and can be applied to both small and large number of observations.

To account for multiple comparisons, the p-values were adjusted for false discovery rate (FDR) control using the Benjamini-Hochberg (BH) method (`statsmodels.stats.multitest.multipletests(method="fdr_bh")` in the `statsmodels` package, version 0.14.5).

Public datasets

To ensure reproducibility and external validation, we included 11 publicly available datasets. GSE55763 (N = 2711, Illumina HumanMethylation 450K) [78] was downloaded from the Gene Expression Omnibus (GEO) database for model and PCA training for its large size, and was also

used to assess age prediction RD using the 72 replicated samples. GSE83944 (N = 48, Illumina HumanMethylation 450K) [79], GSE247193 (N = 23 including one subject without replicate, Illumina Infinium EPIC Human Methylation Beadchip) [80], GSE247195 (N = 23 including one subject without replicate, Illumina Infinium EPIC Human Methylation Beadchip) [80] and GSE247197 (N = 46, Illumina HumanMethylation 450K) [80] were downloaded from the GEO database to assess age prediction RD as they all contain technical replicates. HTS-based datasets GSE144691 (N = 29) [81] and GSE86832 (N = 4) [66] were downloaded from the GEO database for assessment of unreliable CpGs for not containing replicates. Two additional datasets from the GEO database were downloaded for assessment of age prediction MAE and RD, namely GSE232346 (N = 24, TimeSeq) [18] and GSE245628 (N = 24, Illumina Infinium EPIC Human Methylation Beadchip) [18]. Finally, we also downloaded the Buccal cohort (Shokhirev & Johnson, 2025, N = 100, Twist and Illumina Infinium EPIC Human Methylation Beadchip v2) [42] and GSE164600 (N = 24, cfDNA WGBS) [82] datasets for various validation steps. The ethnicity and sex information can be found in the source dataset but was not considered in this study.

Correlation between clock number of CpGs and RD

20 HCRL clocks were selected to predict biological age with `pyaging` (version 0.1.27). No additional data preprocessing was performed except for filling missing values (`NaN`) with KNN imputation. RD were calculated as the mean absolute difference between the predicted ages between replicates. The linear regression was performed using `scipy.stats.linregress` in the `scipy` package (version 1.15.3). Result p-values were corrected for multiple tests via the Benjamini-Hochberg procedure using custom scripts.

Elastic net model training for independent verification of the correlation between number of CpGs and and RD

A series of elastic net models were trained on the dataset GSE55763 (N = 2711) with `sklearn.linear_model.ElasticNet` implemented in the `scikit-learn` package (version 1.6.1). The 90,394 CpGs shared between GSE55763 and the 6 datasets (gDNA_EPICv2, gDNA_MSA, gDNA_Galaxy, gDNA_Twist, cfDNA_Galaxy and cfDNA_Twist) were extracted, and missing values (`NaN`) were filled with KNN imputation implemented in `pyaging`. No additional data cleaning was performed. The elastic net models were trained with 30 logarithmically spaced α values ranging from 10^{-5} to 10^{2}, and 15 linearly spaced `l1_ratio` (λ)

values ranging from 0.0-1.0, resulting in a total of 450 trained models. All other parameters were kept default. Linear regression was performed using `scipy.stats.linregress`.

Minimal sequencing investigation based on binomial stochasticity.

Randomized samples were generated via multi-dimensional binomial model, where the observed methylation counts at each CpG follows independent binomial distributions. For each binomial distribution, the probability parameter was set to the observed beta-value in the original dataset, and the count parameter set to the sequencing depth, given by:

$$beta \sim \frac{1}{depth}\, Binomial(depth, beta_{obs})$$

Where:

depth denotes the simulated sequencing depth under investigation.

$beta_{obs}$ is the empirically estimated true beta-value, calculated as the mean in a replicate pair. Value of $beta_{obs}$ differs across subjects and datasets.

To reduce randomness, 1,000 random datasets were generated at each simulated sequencing depth. Finally, clock prediction metrics (MAE and RD) were calculated with respect to each simulated sequencing depth.

Minimal sequencing investigation based on beta resolution.

This investigation evaluates the impact of uncertainty related to the finite sequencing depths, when beta-values cannot be represented at arbitrary precision. For example, a sequencing depth of 4 can only represent beta-values among 0, 0.25, 0.5, 0.75 and 1.0. To simulate this mechanism, the beta-values were rounded under a given sequencing depth using the following formula before age prediction:

$$beta = \frac{\left\lfloor beta_{obs} \times depth + \frac{1}{2} \right\rfloor}{depth}$$

Where:

depth denotes the simulated sequencing depth under investigation.

$beta_{obs}$ is the empirically estimated true beta-value, calculated as the mean in a replicate pair.

$\lfloor \cdot \rfloor$ is the floor function that maps a real number to the greatest integer less than or equal to that number.

Clock prediction metrics (MAE and RD) were then calculated with respect to each simulated sequencing depth. No repeated tests were necessary since this process is deterministic.

Minimal sequencing investigation based on downsampling.

A downsampling analysis was conducted on five samples, each having an initial sequencing result size ≥50Gb, to assess the effect of varying data volumes on biological age estimation. Raw sequencing data were downsampled at proportions ranging from 0.025 to 0.9 (namely 0.025, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, ..., 0.9). At each downsampling point, biological age was estimated using various established models.

Unreliable beta-value imputation at different cut-off depths

For each combination of selected cut-off depth and imputation method, the beta matrix in steps in the following order:

(1) Missing value imputation (first pass): Missing values (`NaN`) were imputed using the chosen method implemented in `pyaging.preprocess`. Briefly, the KNN imputation method replaces missing values with the mean of k nearest neighbors, where distances between samples are computed based on features with non-missing beta-values. The mean and median methods replace missing entries with the mean or median value of the respective feature, calculated from non-missing observations. The 0-filling method assigns 0 to all missing entries. Notably, among these methods, KNN is the only one that considers information across all features rather than only the imputed one.

(2) Unreliable beta-value removal: Selected beta-values were set to `NaN` when both of the following conditions were met: (a) the observation was 0 or 1, and (b) the sequencing depth at the corresponding CpG site was below the defined cut-off.

(3) Missing value imputation (second pass): The selected imputation method was re-applied to impute these newly removed unreliable CpG sites.

This process was repeated for cut-off depths ranging from 1 to 20 across all four HTS-based datasets generated in this study (gDNA_Galaxy, gDNA_Twist, cfDNA_Galaxy, and cfDNA_Twist).

Transfer learning of classic epigenetic models

We used distillation to adopt classic epigenetic models for HTS methylation data. The classic epigenetic models are teacher models. Each HTS dataset were first PCA-transformed using PCA models trained from public dataset GSE55763 (N = 2711) with an output dimensionality of 2500. The transformed dataset was then combined with the output of a teacher model to train and test a student elastic net model using a 5-fold cross-validation (5-fold CV) implemented in the `scikit-learn` package (version 1.6.1). The final evaluation of age prediction MAE and RD were summarized from all testing splits. The hyperparameters (`alpha` and `l1_ratio`) was optimized using an independent 5-fold cross-validation (`sklearn.linear_models.ElasticNetCV(cv=5)`).

PhenoAge calculation.

The PhenoAge of the samples from the SRRSH-141 cohort was calculated from collected clinical biomarkers using the phenoage clock [4] integrated in `pyaging`.

JSD calculation.

Jenson-Shannon Divergence (JSD) were calculated on delta age predicted by each clock individually. The distribution of delta age probability density functions (p.d.f.) were estimated using Gaussian kernel density estimation (KDE) (`scipy.stats.gaussian_kde`) independently for normal and diseased samples. JSD was then calculated between the normal and disease sample p.d.f.'s using `scipy.spatial.distance.jensonshannon`. The bandwidth parameter in KDE is automatically determined using the "scott" method.

ALS classification.

The ALS classification was performed via support vector machine (SVM) classifier with a linear kernel (`sklearn.svm.SVC(kernel="linear")`) using the delta ages calculated before and after model adaptation respectively. The classification predicts if a sample is ALS (1) or normal (0). A stratified 12-fold cross-validation was used to evaluate the classification

performance while reducing potential overfitting. Briefly, 1 normal and 1 ALS sample were randomly chosen (without replacement) as testing samples, and the rest were for training. Classification performance was evaluated after 12 iterations, i.e. all samples have been predicted as testing cases using the area under receiver operating characteristic curve (AUC) metric (`sklearn.metrics.auc_score`).

Acknowledgement

Special thanks to the high-performance computing cluster provided by Liangzhu Laboratory, Zhejiang University for computation support.

Data availability

The de-identified data generated in this study, including processed methylation matrices, sequencing depth matrices and non-sensitive subject metadata are available at zendoro.org under accession 17453844 (DOI: 10.5281/zenodo.17453844). Raw data available at GSA for Human (https://ngdc.cncb.ac.cn/gsa-human/) under accession HRA014554.

Code availability

The code related to this project are publicly accessible in at https://github.com/regen-bio/cfDNA-HTS-Clock-Analysis.git.

Author contributions

C.O. and G.Z. conceived and supervised the study. G.L., C.O., and G.Z. managed the overall project. Cohort management: L.C., X.C., Z.Y., D.T., X.F., X.H., and V.T. organized the cohort, collected human blood samples, and coordinated resources. X.H. and V.T. also provided essential medical and technical support. Data generation: C.O., Y.G., S.J., B.H., and C.T. performed

sample collection, DNA extraction, and library preparation. Bioinformatics analysis: J.W. and W.H. developed the bioinformatics pipeline. J.W. preprocessed HTS and methylation array data and performed downsampling analysis. W.H. conducted reproducibility, consistency, and accuracy analyses. G.L. investigated clock complexity, unreliable beta filtering, transfer learning, and combination strategies. X.Z. contributed to statistical and machine learning analyses, with advisory support from Y.W. Manuscript preparation: G.L., W.H., X.Z., J.W., and C.O. wrote the manuscript.

All authors reviewed and approved the final manuscript.

Competing interests

W.H., X.Z., J.W., Y.G., Z.Y., S.J., B.H., D.T., V.T., C.T., X.F., C.O. and G.Z. are affiliated with Regenerative Bio Inc., a global biotechcompany powered by AI, providing organ-specific aging assessment models and longevity-focused health solutions. Regenerative Bio Inc. was the study sponsor. The other authors declare no competing interests.

Funding

This research received no grants from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] S. Horvath, K. Raj, *Nat. Rev. Genet.* **2018**, *19*, 371.

[2] M. Moqri, C. Herzog, J. R. Poganik, J. Justice, D. W. Belsky, A. Higgins-Chen, A. Moskalev, G. Fuellen, A. A. Cohen, I. Bautmans, M. Widschwendter, J. Ding, A. Fleming, J. Mannick, J.-D. J. Han, A. Zhavoronkov, N. Barzilai, M. Kaeberlein, S. Cummings, B. K. Kennedy, L. Ferrucci, S. Horvath, E. Verdin, A. B. Maier, M. P. Snyder, V. Sebastiano, V. N. Gladyshev, *Cell* **2023**, *186*, 3758.

[3] Biomarkers of Aging Consortium, C. M. S. Herzog, L. J. E. Goeminne, J. R. Poganik, N. Barzilai, D. W. Belsky, J. Betts-LaCroix, B. H. Chen, M. Chen, A. A. Cohen, S. R. Cummings, P. O. Fedichev, L. Ferrucci, A. Fleming, K. Fortney, D. Furman, V.

Gorbunova, A. Higgins-Chen, L. Hood, S. Horvath, J. N. Justice, D. P. Kiel, G. A. Kuchel, J. Lasky-Su, N. K. LeBrasseur, A. B. Maier, B. Schilling, V. Sebastiano, P. E. Slagboom, M. P. Snyder, E. Verdin, M. Widschwendter, A. Zhavoronkov, M. Moqri, V. N. Gladyshev, *Nat. Aging* **2024**, *4*, 1372.

[4] M. E. Levine, A. T. Lu, A. Quach, B. H. Chen, T. L. Assimes, S. Bandinelli, L. Hou, A. A. Baccarelli, J. D. Stewart, Y. Li, E. A. Whitsel, J. G. Wilson, A. P. Reiner, A. Aviv, K. Lohman, Y. Liu, L. Ferrucci, S. Horvath, *Aging* **2018**, *10*, 573.

[5] L. Sagers, L. Melas-Kyriazi, C. J. Patel, A. K. Manrai, *Aging* **2020**, *12*, 7626.

[6] M. A. Blasco, *Nat. Chem. Biol.* **2007**, *3*, 640.

[7] E. Bobrov, A. Georgievskaya, K. Kiselev, A. Sevastopolsky, A. Zhavoronkov, S. Gurov, K. Rudakov, M. Del Pilar Bonilla Tobar, S. Jaspers, S. Clemann, *Aging* **2018**, *10*, 3249.

[8] J. Li, M. Xiong, X.-H. Fu, Y. Fan, C. Dong, X. Sun, F. Zheng, S.-W. Wang, L. Liu, M. Xu, C. Wang, J. Ping, S. Che, Q. Wang, K. Yang, Y. Zuo, X. Lu, Z. Zheng, T. Lan, S. Wang, S. Ma, S. Sun, B. Zhang, C.-S. Chen, K.-Y. Cheng, J. Ye, J. Qu, Y. Xue, Y.-G. Yang, F. Zhang, W. Zhang, G.-H. Liu, *Med* **2023**, *4*, 825.

[9] S. Horvath, *Genome Biol.* **2013**, *14*, 3156.

[10] M. J. Peters, R. Joehanes, L. C. Pilling, C. Schurmann, K. N. Conneely, J. Powell, E. Reinmaa, G. L. Sutphin, A. Zhernakova, K. Schramm, Y. A. Wilson, S. Kobes, T. Tukiainen, NABEC/UKBEC Consortium, M. A. Nalls, D. G. Hernandez, M. R. Cookson, R. J. Gibbs, J. Hardy, A. Ramasamy, A. B. Zonderman, A. Dillman, B. Traynor, C. Smith, D. L. Longo, D. Trabzuni, J. Troncoso, M. Van Der Brug, M. E. Weale, R. O'Brien, R. Johnson, R. Walker, R. H. Zielke, S. Arepalli, M. Ryten, A. B. Singleton, Y. F. Ramos, H. H. H. Göring, M. Fornage, Y. Liu, S. A. Gharib, B. E. Stranger, P. L. De Jager, A. Aviv, D. Levy, J. M. Murabito, P. J. Munson, T. Huan, A. Hofman, A. G. Uitterlinden, F. Rivadeneira, J. Van Rooij, L. Stolk, L. Broer, M. M. P. J. Verbiest, M. Jhamai, P. Arp, A. Metspalu, L. Tserel, L. Milani, N. J. Samani, P. Peterson, S. Kasela, V. Codd, A. Peters, C. K. Ward-Caviness, C. Herder, M. Waldenberger, M. Roden, P. Singmann, S. Zeilinger, T. Illig, G. Homuth, H.-J. Grabe, H. Völzke, L. Steil, T. Kocher, A. Murray, D. Melzer, H. Yaghootkar, S. Bandinelli, E. K. Moses, J. W. Kent, J. E. Curran, M. P. Johnson, S. Williams-Blangero, H.-J. Westra, A. F. McRae, J. A. Smith, S. L. R. Kardia, I. Hovatta, M. Perola, S. Ripatti, V. Salomaa, A. K. Henders, N. G. Martin, A. K. Smith, D. Mehta, E. B. Binder, K. M. Nylocks, E. M. Kennedy, T. Klengel, J. Ding, A. M. Suchy-Dicey, D. A. Enquobahrie, J. Brody, J. I. Rotter, Y.-D. I. Chen, J. Houwing-Duistermaat, M. Kloppenburg, P. E. Slagboom, Q. Helmer, W. Den Hollander, S. Bean, T. Raj, N. Bakhshi, Q. P. Wang, L. J. Oyston, B. M. Psaty, R. P. Tracy, G. W. Montgomery, S. T. Turner, J. Blangero, I. Meulenbelt, K. J. Ressler, J. Yang, L. Franke, J. Kettunen, P. M. Visscher, G. G. Neely, R. Korstanje, R. L. Hanson, H. Prokisch, L. Ferrucci, T. Esko, A. Teumer, J. B. J. Van Meurs, A. D. Johnson, *Nat. Commun.* **2015**, *6*, 8570.

[11] M. B. Álvez, S. Bergström, J. Kenrick, E. Johansson, M. Åberg, M. Akyildiz, O. Altay, H. Sköld, K. Antonopoulos, E. Apostolakis, Y. H. Balcioglu, A. Bergström, G. Bergström, S. Björkander, S. E. Brage, P. Brodin, L. Butler, S. Cajander, H. Danielsson, M. Dayangac, G. Dinler-Doganay, L. Doğanay, G. Enblad, M. Enblad, L. Fagerberg, S. Falck-Jones, A. Färnert, M. Forsberg, L. Gonzalez, A. Gummesson, K. Gunnarsson, I. Gunnarsson, U. Gyllensten, G. Hesselager, A. Hober, M. Höglund, M. Holmqvist, B. Horuluoglu, R. Hultgren, M. J. Iglesias, H. Janols, F. Johansson, A. Johnsson, L. Klareskog, D. Kotol, I. Kull, M. Kvarnström, M. J. Lautenbach, U. Liljedahl, H. Lindman, C. Lindskog, M.

Lipcsey, I. E. Lundberg, A. Mardinoglu, E. Melén, L. Meng, A.-S. Merritt, J. Mulder, M. T.-H. Nguyen, J. Nordlund, A. Norrby-Teglund, A. Notarnicola, P. Nowak, J. Odeberg, P. Oksvold, T. Olsson, L. Padyukov, K. Pauksens, F. Piehl, E. Pin, F. Pontén, N. Rameika, A. Reepalu, J. Roy, J. M. Schwenk, M. Sen, A. Siika, O. E. Simonson, Å. Sivertsson, T. Sjöblom, E. Sjöstedt, L. Skoglund, A. Smed-Sörensen, K. Sondén, A. Sönnerborg, K. Stålberg, K. Strålin, J. Sundén-Cullberg, C. Sundling, T. Sutantiwanichkul, F. C. Svedman, M. Svensson, E. Svenungsson, T. Lakshmikanth, K. H. Tran-Minh, H. Türkez, C. Unge, P. Venge, M. Wahren-Herlenius, J. Woessmann, H. Yang, U. H. Yeşilkaya, M. Yuan, M. Zeybel, C. Zhang, W. Zhong, M. Zwahlen, K. Von Feilitzen, P. Nilsson, F. Edfors, M. Uhlén, *Science* **2025**, eadx2678.

[12] A. E. Teschendorff, S. Horvath, *Nat. Rev. Genet.* **2025**, *26*, 350.

[13] M. Varshavsky, G. Harari, B. Glaser, Y. Dor, R. Shemer, T. Kaplan, *Cell Rep. Methods* **2023**, *3*, 100567.

[14] K. Ying, J. Song, H. Cui, Y. Zhang, S. Li, X. Chen, H. Liu, A. Eames, D. L. McCartney, R. E. Marioni, J. R. Poganik, M. Moqri, B. Wang, V. N. Gladyshev, *MethylGPT: a foundation model for the DNA methylome*, Systems Biology **2024**.

[15] L. P. De Lima Camillo, L. R. Lapierre, R. Singh, *Npj Aging* **2022**, *8*, 4.

[16] F. Galkin, V. Naumov, S. Pushkov, D. Sidorenko, A. Urban, D. Zagirova, K. M. Alawi, A. Aliper, R. Gumerov, A. Kalashnikov, S. Mukba, A. Pogorelskaya, F. Ren, A. Shneyderman, Q. Tang, D. Xiao, A. Tyshkovskiy, K. Ying, V. N. Gladyshev, A. Zhavoronkov, *Precious3GPT: Multimodal Multi-Species Multi-Omics Multi-Tissue Transformer for Aging Research and Drug Discovery*, Bioinformatics **2024**.

[17] R. Sehgal, Y. Markov, C. Qin, M. Meer, C. Hadley, A. H. Shadyab, R. Casanova, J. E. Manson, P. Bhatti, A. Z. Moore, E. M. Crimmins, S. Hagg, T. L. Assimes, E. A. Whitsel, A. T. Higgins-Chen, M. Levine, *Nat. Aging* **2025**, *5*, 1880.

[18] P. T. Griffin, A. E. Kane, A. Trapp, J. Li, M. Arnold, J. R. Poganik, R. J. Conway, M. S. McNamara, M. V. Meer, N. Hoffman, J. A. Amorim, X. Tian, M. R. MacArthur, S. J. Mitchell, A. L. Mueller, C. Carmody, D. L. Vera, C. Kerepesi, K. Ying, N. Noren Hooten, J. R. Mitchell, M. K. Evans, V. N. Gladyshev, D. A. Sinclair, *Nat. Aging* **2024**, *4*, 261.

[19] H. Tong, V. B. Dwaraka, Q. Chen, Q. Luo, J. A. Lasky-Su, R. Smith, A. E. Teschendorff, *Nat. Aging* **2024**, *4*, 886.

[20] A. Haghani, A. T. Lu, Q. Yan, J. C. I. Belmonte, P. Reddy, V. Cheng, X. W. Yang, N. Wang, K. Mozhui, K. Murach, A. Ocampo, R. W. Williams, M. Jucker, C. Bergmann, J. R. Poganik, B. Zhang, V. N. Gladyshev, S. Horvath, *GeroScience* **2025**.

[21] Q. Zhang, C. L. Vallerga, R. M. Walker, T. Lin, A. K. Henders, G. W. Montgomery, J. He, D. Fan, J. Fowdar, M. Kennedy, T. Pitcher, J. Pearson, G. Halliday, J. B. Kwok, I. Hickie, S. Lewis, T. Anderson, P. A. Silburn, G. D. Mellick, S. E. Harris, P. Redmond, A. D. Murray, D. J. Porteous, C. S. Haley, K. L. Evans, A. M. McIntosh, J. Yang, J. Gratten, R. E. Marioni, N. R. Wray, I. J. Deary, A. F. McRae, P. M. Visscher, *Genome Med.* **2019**, *11*, 54.

[22] A. T. Higgins-Chen, K. L. Thrush, Y. Wang, C. J. Minteer, P.-L. Kuo, M. Wang, P. Niimi, G. Sturm, J. Lin, A. Z. Moore, S. Bandinelli, C. H. Vinkers, E. Vermetten, B. P. F. Rutten, E. Geuze, C. Okhuijsen-Pfeifer, M. Z. Van Der Horst, S. Schreiter, S. Gutwinski, J. J. Luykx, M. Picard, L. Ferrucci, E. M. Crimmins, M. P. Boks, S. Hägg, T. T. Hu-Seliger, M. E. Levine, *Nat. Aging* **2022**, *2*, 644.

[23] E. L. S. Conole, J. A. Robertson, H. M. Smith, S. R. Cox, R. E. Marioni, *Nat. Rev. Neurol.* **2025**, *21*, 411.
[24] E. A. Klein, D. Richards, A. Cohn, M. Tummala, R. Lapham, D. Cosgrove, G. Chung, J. Clement, J. Gao, N. Hunkapiller, A. Jamshidi, K. N. Kurtzman, M. V. Seiden, C. Swanton, M. C. Liu, *Ann. Oncol.* **2021**, *32*, 1167.
[25] M. Moqri, C. Herzog, J. R. Poganik, K. Ying, J. N. Justice, D. W. Belsky, A. T. Higgins-Chen, B. H. Chen, A. A. Cohen, G. Fuellen, S. Hägg, R. E. Marioni, M. Widschwendter, K. Fortney, P. O. Fedichev, A. Zhavoronkov, N. Barzilai, J. Lasky-Su, D. P. Kiel, B. K. Kennedy, S. Cummings, P. E. Slagboom, E. Verdin, A. B. Maier, V. Sebastiano, M. P. Snyder, V. N. Gladyshev, S. Horvath, L. Ferrucci, *Nat. Med.* **2024**, *30*, 360.
[26] Q. Gao, Q. Zeng, Z. Wang, C. Li, Y. Xu, P. Cui, X. Zhu, H. Lu, G. Wang, S. Cai, J. Wang, J. Fan, *The Innovation* **2022**, *3*, 100259.
[27] J. E. Medina, N. C. Dracopoli, P. B. Bach, A. Lau, R. B. Scharpf, G. A. Meijer, C. L. Andersen, V. E. Velculescu, *J. Immunother. Cancer* **2023**, *11*, e006013.
[28] K. Zhang, R. Fu, R. Liu, Z. Su, *Trends Cancer* **2024**, *10*, 161.
[29] D. C. Bruhm, N. A. Vulpescu, Z. H. Foda, J. Phallen, R. B. Scharpf, V. E. Velculescu, *Nat. Rev. Cancer* **2025**, *25*, 341.
[30] S.-J. Li, X. Gao, Z.-H. Wang, J. Li, L.-T. Zeng, Y.-M. Dang, Y.-Q. Ma, L.-Q. Zhang, Q.-Y. Wang, Y.-M. Zhang, H.-L. Liu, R.-M. Qi, J.-P. Cai, *Epigenomics* **2024**, *16*, 715.
[31] A. V. Sergeev, O. V. Kisil, A. A. Eremin, A. S. Petrov, M. E. Zvereva, *Biochem. Mosc.* **2025**, *90*, S342.
[32] M. Shtumpf, S. Jeong, M. Bikova, H. Mamayusupova, L. Ruje, V. B. Teif, *Aging Cell* **2024**, *23*, e14100.
[33] C. Hu, Y. Li, L. Li, N. Zhang, X. Zheng, *Bioinformatics* **2024**, *40*, btae656.
[34] Infinium MethylationEPIC v2.0 BeadChip, .
[35] B. Poggiali, M. E. Dupont, M.-L. Kampmann, A. Vidaki, V. Pereira, C. Børsting, J. Tfelt-Hansen, J. D. Andersen, *Biol. Proced. Online* **2025**, *27*, 34.
[36] S. M. Lee, C. E. Loo, R. D. Prasasya, M. S. Bartolomei, R. M. Kohli, W. Zhou, *Nucleic Acids Res.* **2024**, *52*, e38.
[37] C. Shu, X. Zhang, B. E. Aouizerat, K. Xu, *Epigenetics Chromatin* **2020**, *13*, 51.
[38] J. A. Heiss, K. J. Brennan, A. A. Baccarelli, M. M. Téllez-Rojo, G. Estrada-Gutiérrez, R. O. Wright, A. C. Just, *Epigenetics* **2020**, *15*, 174.
[39] L.-Y. Zhao, J. Song, Y. Liu, C.-X. Song, C. Yi, *Protein Cell* **2020**, *11*, 792.
[40] K. De Ridder, H. Che, K. Leroy, B. Thienpont, *Nat. Commun.* **2024**, *15*, 4134.
[41] T. Zhu, J. Liu, S. Beck, S. Pan, D. Capper, M. Lechner, C. Thirlwell, C. E. Breeze, A. E. Teschendorff, *Nat. Methods* **2022**, *19*, 296.
[42] M. N. Shokhirev, A. A. Johnson, *GeroScience* **2025**.
[43] A. R. De Abreu, J. Ibrahim, V. Lemonidis, L. Mateiu, G. Van Camp, K. Op De Beeck, *Epigenetics Chromatin* **2025**, *18*, 57.
[44] L. Liu, J. M. Toung, A. F. Jassowicz, R. Vijayaraghavan, H. Kang, R. Zhang, K. M. Kruglyak, H. J. Huang, T. Hinoue, H. Shen, N. S. Salathia, D. S. Hong, A. Naing, V. Subbiah, S. A. Piha-Paul, M. Bibikova, G. Granger, B. Barnes, R. Shen, K. Gutekunst, S. Fu, A. M. Tsimberidou, C. Lu, C. Eng, S. L. Moulder, E. S. Kopetz, R. N. Amaria, F. Meric-Bernstam, P. W. Laird, J.-B. Fan, F. Janku, *Ann. Oncol.* **2018**, *29*, 1445.

[45] D. N. Buckley, J. P. Lewinger, G. Gooden, M. Spillman, M. Neuman, X. M. Guo, B. Y. Tew, H. Miller, V. U. Khetan, L. P. Shulman, L. Roman, B. Salhia, *Clin. Cancer Res.* **2023**, *29*, 5196.
[46] S. Komaki, M. Nagata, E. Arai, R. Otomo, K. Ono, Y. Abe, H. Ohmomo, S. Umekage, N. O. Shinozaki, T. Hachiya, Y. Sutoh, Y. Otsuka-Yamasaki, Y. Arai, N. Hirose, A. Yoneyama, H. Okano, M. Sasaki, Y. Kanai, A. Shimizu, *Lancet Healthy Longev.* **2023**, *4*, e83.
[47] O. Bl, N. D, C. D, P. A, P. S, G. R. O, H. A, L. N, V. M, R. R, S. I, F. Y, H. H, G. B, D. Y, K. T, S. R, *Cell Rep.* **2025**.
[48] Y. Jin, K. N. Conneely, W. Ma, R. K. Naviaux, T. Siddique, E. G. Allen, S. Guingrich, R. M. Pascuzzi, P. Jin, *Cell Biosci.* **2025**, *15*, 26.
[49] L. D. Garma, M. Quintela-Fandino, *Genome Med.* **2024**, *16*, 116.
[50] J. Wang, *ComBat-met: Adjusting Batch Effects in DNA Methylation Data*, Bioinformatics **2024**.
[51] K. Ying, S. Paulson, A. Eames, A. Tyshkovskiy, S. Li, N. Eynon, M. Jacques, R. Grolaux, K. Seale, E. Jacques, L. J. E. Goeminne, A. Cipriano, M. Perez-Guevara, M. Emamifar, M. Casas Martínez, D. Kwon, A. Kosheleva, M. Snyder, D. Gobel, C. Herzog, D. L. McCartney, R. E. Marioni, J. Lasky-Su, J. R. Poganik, M. Moqri, V. N. Gladyshev, *Nat. Aging* **2025**.
[52] T. Zindler, H. Frieling, A. Neyazi, S. Bleich, E. Friedel, *BMC Bioinformatics* **2020**, *21*, 271.
[53] L. Handl, A. Jalali, M. Scherer, N. Pfeifer, *Partially blind domain adaptation for age prediction from DNA methylation data*, arXiv **2016**.
[54] L. Luo, L. Shang, J. M. Goodrich, K. E. Peterson, P. X.-K. Song, *Bridging the gap: Enhancing the generalizability of epigenetic clocks through transfer learning*, Epidemiology **2025**.
[55] A. Tomusiak, A. Floro, R. Tiwari, R. Riley, H. Matsui, N. Andrews, H. G. Kasler, E. Verdin, *Commun. Biol.* **2024**, *7*, 934.
[56] F. Galkin, P. Mamoshina, K. Kochetov, D. Sidorenko, A. Zhavoronkov, *Aging Dis.* **2021**, *12*, 1252.
[57] S. Fong, K. Pabis, D. Latumalea, N. Dugersuren, M. Unfried, N. Tolwinski, B. Kennedy, J. Gruber, *Nat. Aging* **2024**, *4*, 1137.
[58] T. K. Koo, M. Y. Li, *J. Chiropr. Med.* **2016**, *15*, 155.
[59] R. Sehgal, D. Borrus, J. Gonzalez, Y. Markov, ADNI, A. Higgins-Chen, *Biological versus Technical Reliability of Epigenetic Clocks and Implications for Disease Prognosis and Intervention Response*, Bioinformatics **2025**.
[60] M. J. Ziller, K. D. Hansen, A. Meissner, M. J. Aryee, *Nat. Methods* **2015**, *12*, 230.
[61] Y. Zhao, X. Li, K. Wang, G. Iyer, S. A. Sakowski, L. Zhao, S. Teener, K. M. Bakulski, J. F. Dou, B. J. Traynor, A. Karnovsky, S. A. Batterman, E. L. Feldman, M. A. Sartor, S. A. Goutman, *eBioMedicine* **2024**, *109*, 105383.
[62] P. Merbaum, R. Zwamborn, P. Hop, P. M. A. Sequencing Consortium, W. Van Rheenen, J. Veldink, *Epigenomics* **2025**, *17*, 1153.
[63] T. Wang, W. Guan, J. Lin, N. Boutaoui, G. Canino, J. Luo, J. C. Celedón, W. Chen, *Epigenetics* **2015**, *10*, 662.
[64] I. Raluseviciute, F. Drabløs, M. B. Rye, *Clin. Epigenetics* **2019**, *11*, 193.
[65] L. Zhou, H. K. Ng, D. I. Drautz-Moses, S. C. Schuster, S. Beck, C. Kim, J. C. Chambers, M. Loh, *Sci. Rep.* **2019**, *9*, 10383.

[66] R. Pidsley, E. Zotenko, T. J. Peters, M. G. Lawrence, G. P. Risbridger, P. Molloy, S. Van Djik, B. Muhlhausler, C. Stirzaker, S. J. Clark, *Genome Biol.* **2016**, *17*, 208.
[67] K. Sugden, E. J. Hannon, L. Arseneault, D. W. Belsky, D. L. Corcoran, H. L. Fisher, R. M. Houts, R. Kandaswamy, T. E. Moffitt, R. Poulton, J. A. Prinz, L. J. H. Rasmussen, B. S. Williams, C. C. Y. Wong, J. Mill, A. Caspi, *Patterns* **2020**, *1*, 100014.
[68] J. H. Tay, Y. E. Chew, W. Wang, Z. M. Lim, L. Guan, R. Dorajoo, B. K. Kennedy, R. Brooke, J. Gordevicius, S. Horvath, E. Sandalova, A. B. Maier, *Commun. Biol.* **2025**, *8*, 654.
[69] D. S. Borrus, R. Sehgal, J. F. Armstrong, J. Gonzalez, G. Zou, J. Kasamoto, Y. Markov, J. Lasky-Su, A. Higgins-Chen, *TranslAGE: A Comprehensive Platform for Systematic Validation of Epigenetic Aging Biomarkers*, Genomics **2025**.
[70] L. P. De Lima Camillo, R. Sehgal, J. Armstrong, A. T. Higgins-Chen, S. Horvath, B. Wang, *CpGPT: a Foundation Model for DNA Methylation*, Systems Biology **2024**.
[71] Y. Zhang, Y. Yao, Y. Tang, Y. Cheng, Y. Xu, Y. He, Y. Qi, L. Jin, *Nat. Comput. Sci.* **2026**.
[72] T. Evgeniou, M. Pontil, in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, Seattle WA USA **2004**, pp. 109–117.
[73] S. J. Pan, I. W. Tsang, J. T. Kwok, Q. Yang, *IEEE Trans. Neural Netw.* **2011**, *22*, 199.
[74] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, *J Mach Learn Res* **2016**, *17*, 2096.
[75] J. D. Faul, J. K. Kim, M. E. Levine, B. Thyagarajan, D. R. Weir, E. M. Crimmins, *Proc. Natl. Acad. Sci.* **2023**, *120*, e2215840120.
[76] L. P. De Lima Camillo, *Bioinformatics* **2024**, *40*, btae200.
[77] M. J. Aryee, A. E. Jaffe, H. Corrada-Bravo, C. Ladd-Acosta, A. P. Feinberg, K. D. Hansen, R. A. Irizarry, *Bioinformatics* **2014**, *30*, 1363.
[78] B. Lehne, A. W. Drong, M. Loh, W. Zhang, W. R. Scott, S.-T. Tan, U. Afzal, J. Scott, M.-R. Jarvelin, P. Elliott, M. I. McCarthy, J. S. Kooner, J. C. Chambers, *Genome Biol.* **2015**, *16*, 37.
[79] G. Oh, K. Koncevičius, S. Ebrahimi, M. Carlucci, D. E. Groot, A. Nair, A. Zhang, A. Kriščiūnas, E. S. Oh, V. Labrie, A. H. C. Wong, J. Gordevičius, P. Jia, M. Susic, A. Petronis, *Genome Biol.* **2019**, *20*, 2.
[80] K. Koncevičius, A. Nair, A. Šveikauskaitė, A. Šeštokaitė, A. Kazlauskaitė, A. Dulskas, A. Petronis, *Aging Cell* **2024**, *23*, e14170.
[81] H. A. Abdelsamed, C. C. Zebley, H. Nguyen, R. L. Rutishauser, Y. Fan, H. E. Ghoneim, J. C. Crawford, F. Alfei, S. Alli, S. P. Ribeiro, A. H. Castellaw, M. A. McGargill, H. Jin, S. K. Boi, C. Speake, E. Serti, L. A. Turka, M. E. Busch, M. Stone, S. G. Deeks, R.-P. Sekaly, D. Zehn, E. A. James, G. T. Nepom, B. Youngblood, *Nat. Immunol.* **2020**, *21*, 578.
[82] C. Caggiano, B. Celona, F. Garton, J. Mefford, B. L. Black, R. Henderson, C. Lomen-Hoerth, A. Dahl, N. Zaitlen, *Nat. Commun.* **2021**, *12*, 2717.

Tables

Table 1| Linear regression results between the age prediction replicate difference (RD) and the number of CpGs in 20 HCRL clocks. HCRL denotes 20 linear-architecture clocks classified as Hi-Cov-All or High-Repro. Negative correlations indicate larger number of CpGs is associated with lower age prediction discrepancies between technical replicates. P-values were adjusted for multiple testing using the Benjamini-Hochberg procedure. Full regression outputs for each dataset are provided in Fig. S7.

Dataset	Slope	Intercept	Pearson's R	p
gDNA_MSA	-0.36	2.34	-0.94	**3.65E-10**
gDNA_EPICv2	-0.24	1.30	-0.87	**4.67E-07**
gDNA_Twist	-0.28	2.87	-0.90	**4.91E-08**
gDNA_Galaxy	-0.31	3.01	-0.93	**3.49E-09**
cfDNA_Twist	-0.29	3.04	-0.91	**2.69E-08**
cfDNA_Galaxy	-0.28	3.01	-0.90	**5.86E-08**
GSE83944	-0.18	0.75	-0.83	**7.78E-05**
GSE247193	-0.19	1.60	-0.84	**5.65E-05**
GSE247195	-0.22	1.14	-0.85	**2.99E-05**
GSE247197	-0.14	0.55	-0.67	**4.62E-03**
GSE55763	-0.18	0.93	-0.81	**1.65E-04**

Figure Legend

Fig. 1| Study design and evaluation of methylation data reproducibility and consistency across six SRRSH-24 datasets. a, The overall design of the study. **b-c**, Principal component analysis (PCA) of all samples colored by datasets (**b**) or replicates (**c**). Beta-values of Method-shared CpGs (N = 136,437) were used for PCA calculation. **d**, Intraclass Correlation Coefficients (ICCs) of Method-shared CpGs across six datasets. **e**, Pearson correlations among datasets computed using the mean beta-values of Method-shared CpGs. **f**, ICCs of 53 epigenetic clocks benchmarked in the study. Data are presented as ICC estimates with 95 % confidence intervals and colored by grouped ICC values as follows: Excellent: $0.9 \leq$ ICC, Good: $0.75 \leq$ ICC < 0.9, Moderate: $0.5 \leq$ ICC < 0.75, Poor: ICC < 0.5. Clocks were classified in two different rationales: (1) By generations: First, Chronological age prediction; Second, Phenotypic age or mortality prediction; Third, Pace of aging prediction; Fourth, Causal clocks; Gestational, Gestational age prediction; Other: Telomere length or mitotic age prediction. (2) By CpG coverage across the four profiling methods (MSA, EPICv2, Galaxy, Twist): Hi-Cov-All, first- or second-generation clocks, applicable to adults and have > 80 % CpGs covered by all four methods; Hi-Repro, first- or second-generation clocks with excellent ICC across datasets; Hi-Cov-nonMSA, first- or second-generation clocks, applicable to adults, have > 80 % CpGs covered by three methods except MSA, and not classified as Hi-Repro; Other, all remaining clocks. **g**, Comparisons of per clock ICCs of all tested epigenetic clocks (N = 53). **h**, Pearson's correlations across six datasets calculated using mean of clock predictions for three clock categories: Hi-Repro, Hi-Cov-All, and Hi-Cov-nonMSA. P-values in (**d**) and (**g**) were calculated by two-tailed block permutation test with 10,000 permutations, then adjusted for multiple comparisons using the Benjamini-Hochberg procedure. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; N.S., not significant. The center line of each boxplot marks the median, the box body marks the 1st and 3rd quartile, the caps and/or the span of whisker mark the min and max of RD values respectively.

Fig. 2| Decline of epigenetic clock performance in 6 SRRSH-24 HTS-based datasets. a, Lollipop plot showing Pearson's R and MAE (between predicted age and chronological age) for 6 Hi-Repro clocks. Each dot represents a replicate, and color represents R. **b-c**, Boxplots showing comparisons and statistical significance of R (**b**) and MAE (**c**) for Hi-Cov-All clocks across datasets. The center line of each boxplot marks the median, the box body marks the 1st and 3rd quartile, and the caps mark the min and max of RD values respectively. **d-e**, Performance evaluation (R and MAE) of 4 array-trained clocks (hannum, horvath2013, skinandblood and dnamphenoage) and a HTS data-trained clock (timeseq) cross-applied to array-based dataset (GSE245628, labelled as EPICv1) and HTS-based dataset (GSE232346, labelled as TimeSeq). Only the 24 shared samples in both GSE245628 and GSE232346 were incorporated. **f**, Scatter plot showing the correlation between clock ICC and MAE. **g**, Scatter plot showing the correlation between clock ICC and R. Each dot represents one clock from Hi-Cov-All category, color represents replicate group. Lines represent linear regression fits for each replicate group respectively. The p-values in (**b**) and (**c**) were calculated by two-tailed block permutation test with 10,000 permutations, then adjusted for multiple comparisons using the Benjamini-Hochberg procedure. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; N.S., not significant; The R and p-values in (**f**) and (**g**) were calculated by Pearson correlation at replicate level.

Fig. 3| Investigation of the negative correlation between number of CpGs (log-transformed) and the age prediction replicate difference (RD) (log-transformed) shows that the association is largely driven by the λ hyperparameter (L1-regularisation mixing ratio). a, A series of elastic nets were trained on dataset GSE55763 (N = 2711) and tested with the 6 datasets from this study. The models were trained under different hyperparameter (α and λ) values to obtain models with different number of CpGs. Each point represents a model trained with a distinct α and λ pair. **b**, The negative correlation between number of CpGs and age prediction RD was validated in all testing datasets with high significance ($p \leq 4.7\times10^{-10}$). **c-d**, Compared with

hyperparameter α, λ showed a stronger and more significant correlation with RD (p: 9.7×10^{-9} - 8.2×10^{-3} vs. 1.6×10^{-47} - 7.3×10^{-33}; Pearson's R^2: 0.05 - 0.21 vs. 0.67 - 0.78). R^2 and p-values were calculated by Pearson correlation.

Fig. 4| Investigating the impact of sampling depth, unreliable beta-values on 20 HCRL clocks and imputation-based correction techniques. HCRL denotes 20 linear-architecture clocks classified as Hi-Cov-All or High-Repro. **a-b**, Investigation of the sequencing depth requirements with downsampling simulations with 5 qualified samples from the SRRSH cohort with sequencing result size ≥50Gb. **a**, Age prediction mean absolute error converged to the dataset baseline at 10-20× downsampling depths. **b**, The increasing of the fraction of CpGs with ≥ 20× slowed down above ≥ 20× downsampling depth, although it could continue climbing up to 50 - 80× downsampling depth. **c**, HCRL clocks are prone to unreliable beta-values originating from high-throughput sequencing (HTS)-based methylation profiling technologies, with CpG coverage varying from 2.0% to 34.6% across the clocks examined (Details are shown in Table S3). The center line of each boxplot marks the median, the box body marks the 1st and 3rd quartile, and the caps mark the min and max values respectively. **d-e**, The K-nearest-neighbor (KNN) imputation technique was tested at varying depth cut-offs; only CpG sites below the cut-off were imputed. **f-g**, Similarly, mean and median imputation significantly improved RD but simultaneously increased MAE in cfDNA datasets. The horizontal strokes mark the medians of each respective distribution. No significant increase of MAE was observed in gDNA datasets. 0-filling yielded neither significant improvement nor deterioration in MAE or RD. P-values were obtained by two-sided permutation tests (10,000 permutations, custom script) comparing pre- and post-imputation predictions. Multiple testing results were adjusted using the Benjamini-Hochberg procedure. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; NS, not significant.

Fig. 5| Classic epigenetic clock prediction accuracy can be improved using transfer learning techniques. Distillation, selected for its minimal requirement for original parameters and training data, incorporates the original model (teacher model) output as an additional input feature to train an elastic net student model for prediction and validation. All models were trained and tested with 5-fold cross-validation. Prior to transfer learning, datasets were transformed by PCA (loadings derived from GSE55763; N = 2711) to reduce dimensionality and accelerate training. **a**, Distillation improved age prediction mean absolute error (MAE) with 6 selected epigenetic clocks as teacher models. MAE values for the teacher model and for the student model without transfer learning are shown for reference. MAE was assessed after summarizing all predicted splits after a full pass of 5-cold cross validation. **b**, Distillation reduced age prediction replicate difference (RD) between teacher clocks, however, the direction and magnitude of improvement varied across datasets and teacher model selections. RD values for the teacher model and the student model without transfer learning are shown for reference. RD was assessed after summarizing all predicted splits after a full pass of 5-cold cross validation. The center line of each boxplot marks the median, the box body marks the 1st and 3rd quartile, and the caps mark the min and max of RD values respectively. **c**, Generalizability test of distillation across six SRRSH-24 datasets, by testing the distillation model on a dataset of DNA type or methylation profiling technology different from its training settings. Boxplots through **b-c** are with boxes indicating the 1st -3rd quartile range, caps indicating the minimum and maximum values, and the center strokes for median values.

Fig. 6| Validating epigenetic clock adaptations strategies. a-b, Transfer learning (TL) exhibited as a critical component in adapting array-based epigenetic clocks on HTS-based data, reflected by significant improvements in age prediction mean absolute error (MAE) and correlation. Performance tested on 20 HCRL clocks with both (**a**) chronological age (CA) (on the Buccal cohort, N = 100) and (**b**) clinical biomarker derived PhenoAge (on the SRRSH-141 cohort, N = 141) as age references. HCRL denotes 20 linear-architecture clocks classified as Hi-Cov-All or High-Repro. Method abbreviations: Direct, no improvement

techniques used; DF, depth filtering with a minimal sequencing depth of 20×; IM, imputing unreliable (0 and 1) beta-values; TL, incorporated distillation as transfer learning technique. These techniques were sequentially applied in DF, IM, TL order. **c-f**, Validation of epigenetic clock adaptation in amyotrophic lateral sclerosis (ALS) classification applications using public dataset GSE164600 (N = 24). **c**, The distributional change of delta age predicted by 20 HCRL clocks before and after clock adaptation (IM+TL). The center line in each violin plot marks the median. For each model, the delta age change was plotted in reference to the median delta age of normal samples prior to adaptation. **d**, The Jenson-Shannon Divergence (JSD) calculated to evaluate the distributional distances between ALS and normal sample. The models in (**c**) were sorted in descending order per change in (ΔJSD) with respect to clock adaptation. Each box marks the 1st -3rd quartile range, cap marks minimum and maximum values, and the center line marks the median. **e**, Collectively using the delta age predicted by top-ranked models (by ΔJSD) improved ALS classification performance. Classification was conducted using support machine classifier with a linear kernel and was evaluated using area under the receiver operating characteristic curve (AUC). **f**, The maximum AUC could be achieved using as few as 4 top ΔJSD clocks, leading the non-adapted prediction by 0.125 AUC.

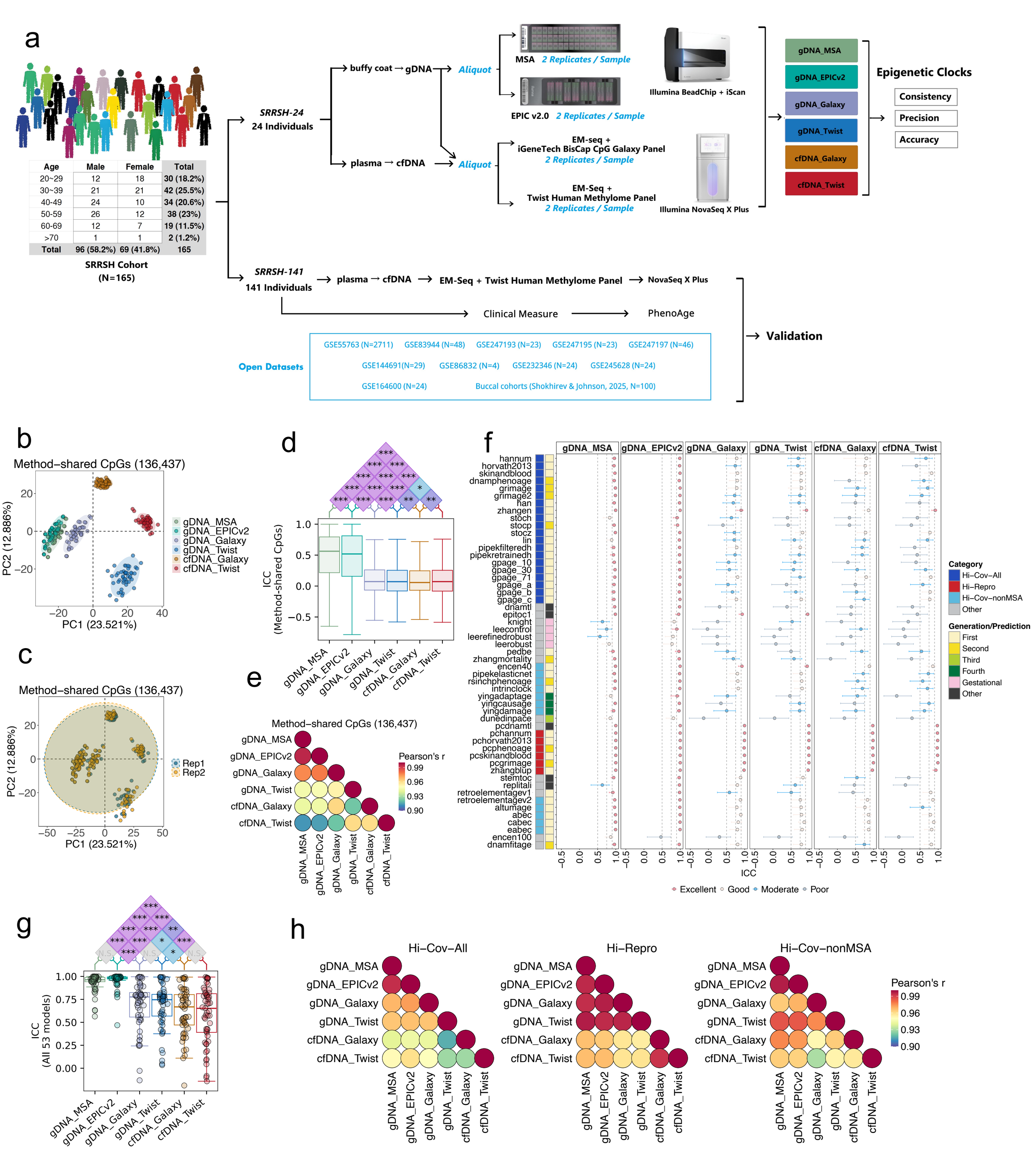
a
Age | Male | Female | Total
20~29 | 12 | 18 | 30 (18.2%)
30~39 | 21 | 21 | 42 (25.5%)
40-49 | 24 | 10 | 34 (20.6%)
50-59 | 26 | 12 | 38 (23%)
60-69 | 12 | 7 | 19 (11.5%)
>70 | 1 | 1 | 2 (1.2%)
Total | 96 (58.2%) | 69 (41.8%) | 165
SRRSH Cohort (N=165)
SRRSH-24
24 Individuals
buffy coat → gDNA
plasma → cfDNA
Aliquot
MSA 2 Replicates / Sample
EPIC v2.0 2 Replicates / Sample
Illumina BeadChip + iScan
EM-seq + iGeneTech BisCap CpG Galaxy Panel
2 Replicates / Sample
EM-Seq + Twist Human Methylome Panel
2 Replicates / Sample
Illumina NovaSeq X Plus
gDNA_MSA
gDNA_EPICv2
gDNA_Galaxy
gDNA_Twist
cfDNA_Galaxy
cfDNA_Twist
Epigenetic Clocks
Consistency
Precision
Accuracy
SRRSH-141
141 Individuals
plasma → cfDNA → EM-Seq + Twist Human Methylome Panel → NovaSeq X Plus
Clinical Measure → PhenoAge
Open Datasets
GSE55763 (N=2711) GSE83944 (N=48) GSE247193 (N=23) GSE247195 (N=23) GSE247197 (N=46)
GSE144691(N=29) GSE86832 (N=4) GSE232346 (N=24) GSE245628 (N=24)
GSE164600 (N=24) Buccal cohorts (Shokhirev & Johnson, 2025, N=100)
Validation
b
Method−shared CpGs (136,437)
PC2 (12.886%)
PC1 (23.521%)
gDNA_MSA
gDNA_EPICv2
gDNA_Galaxy
gDNA_Twist
cfDNA_Galaxy
cfDNA_Twist
c
Method−shared CpGs (136,437)
PC2 (12.886%)
PC1 (23.521%)
Rep1
Rep2
d
ICC (Method-shared CpGs)
e
Method−shared CpGs (136,437)
Pearson's r
0.99
0.96
0.93
0.90
f
ICC
Excellent
Good
Moderate
Poor
Category
Hi−Cov−All
Hi−Repro
Hi−Cov−nonMSA
Other
Generation/Prediction
First
Second
Third
Fourth
Gestational
Other
g
ICC (All 53 models)
h
Hi−Cov−All
Hi−Repro
Hi−Cov−nonMSA
Pearson's r

Fig. 1

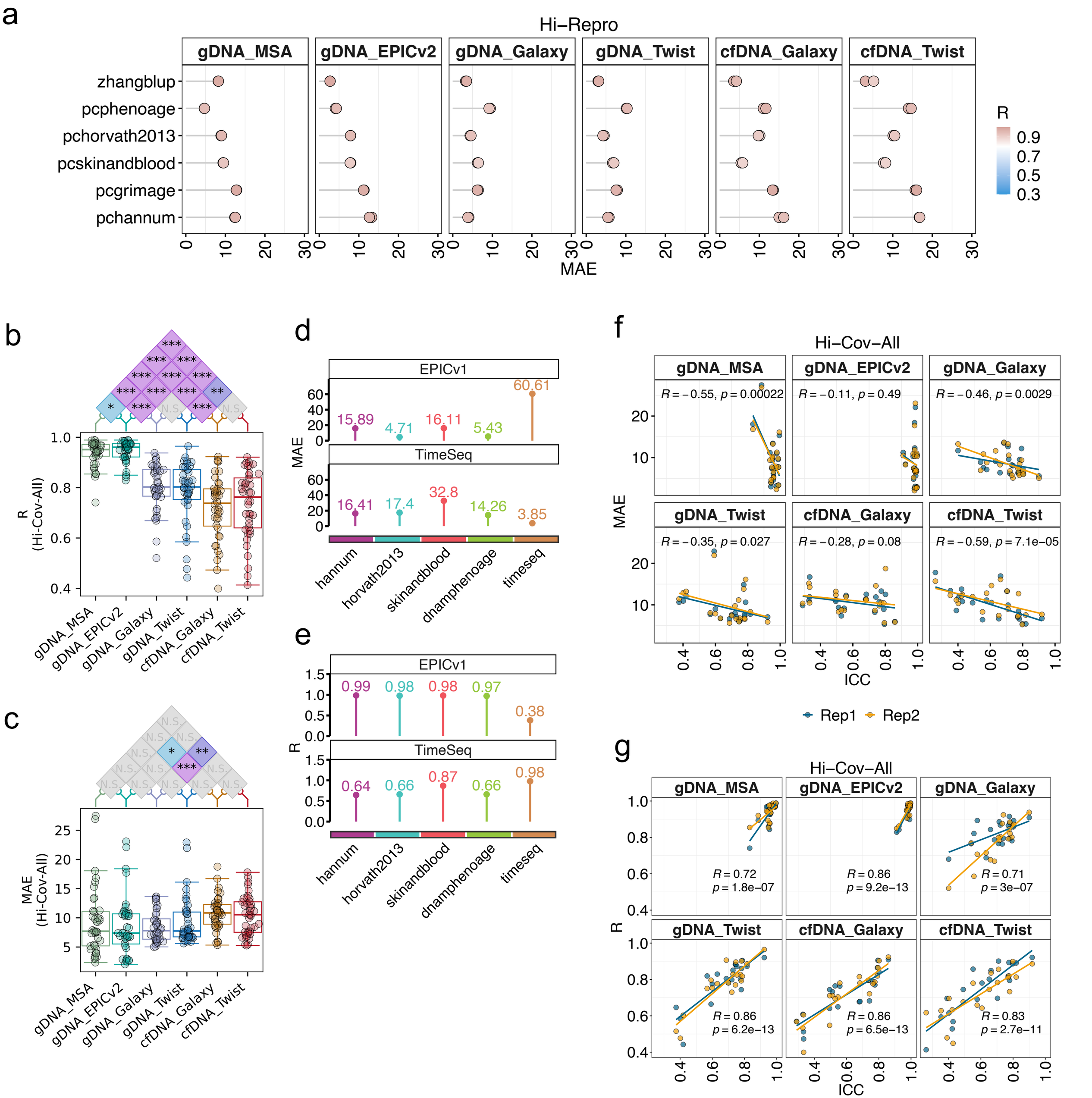
a
Hi−Repro
gDNA_MSA
gDNA_EPICv2
gDNA_Galaxy
gDNA_Twist
cfDNA_Galaxy
cfDNA_Twist
zhangblup
pcphenoage
pchorvath2013
pcskinandblood
pcgrimage
pchannum
MAE
R
0.9
0.7
0.5
0.3
b
R (Hi-Cov-All)
c
MAE (Hi-Cov-All)
d
EPICv1
TimeSeq
MAE
15.89
4.71
16.11
5.43
60.61
16.41
17.4
32.8
14.26
3.85
hannum
horvath2013
skinandblood
dnamphenoage
timeseq
e
R
0.99
0.98
0.98
0.97
0.38
0.64
0.66
0.87
0.66
0.98
f
Hi−Cov−All
R = −0.55, p = 0.00022
R = −0.11, p = 0.49
R = −0.46, p = 0.0029
R = −0.35, p = 0.027
R = −0.28, p = 0.08
R = −0.59, p = 7.1e−05
ICC
Rep1
Rep2
g
R = 0.72
p = 1.8e−07
R = 0.86
p = 9.2e−13
R = 0.71
p = 3e−07
R = 0.86
p = 6.2e−13
R = 0.86
p = 6.5e−13
R = 0.83
p = 2.7e−11

Fig. 2

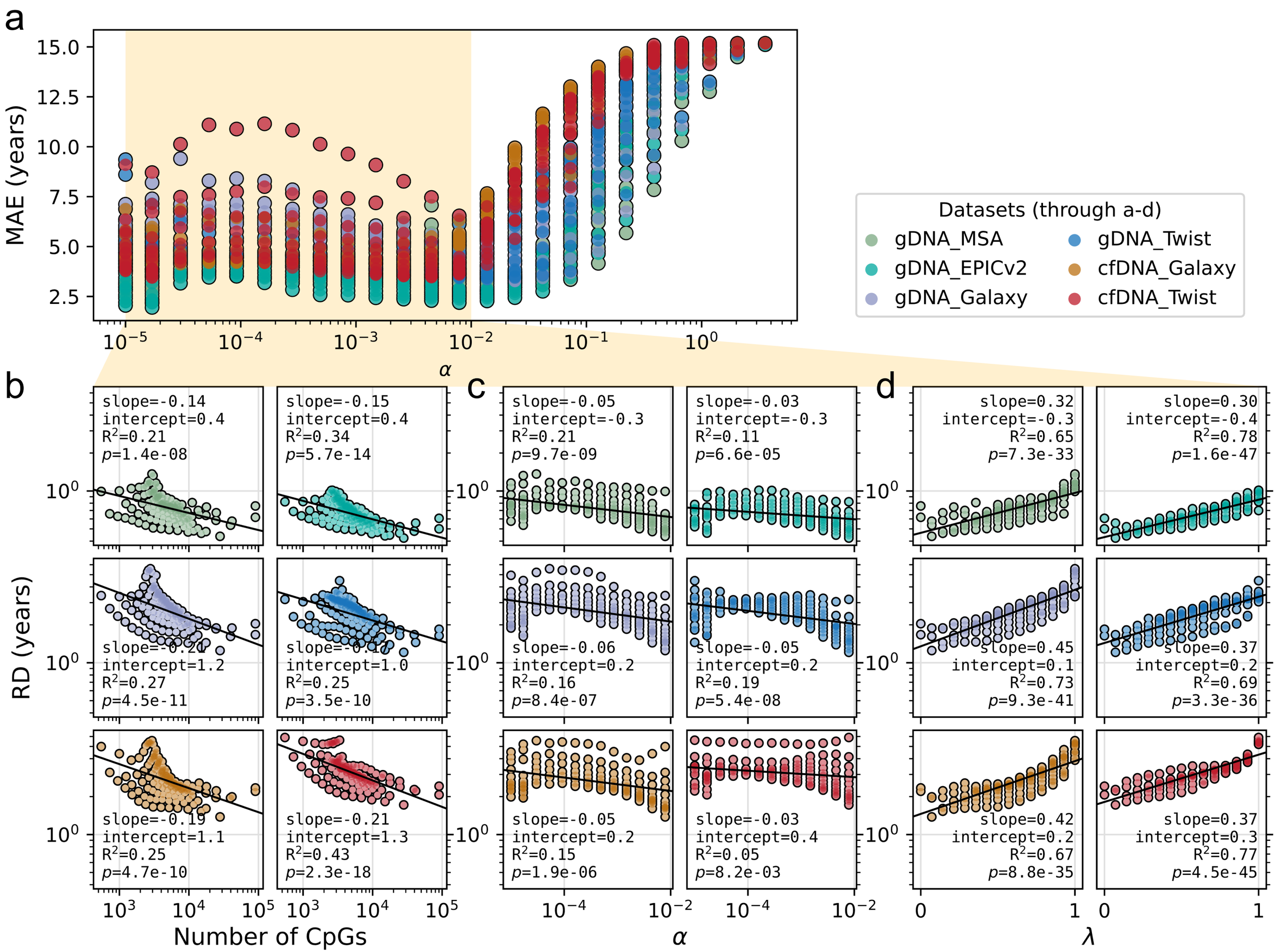
a
MAE (years)
15.0
12.5
10.0
7.5
5.0
2.5
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
α
Datasets (through a-d)
gDNA_MSA
gDNA_EPICv2
gDNA_Galaxy
gDNA_Twist
cfDNA_Galaxy
cfDNA_Twist
b
c
d
RD (years)
slope=-0.14
intercept=0.4
R²=0.21
p=1.4e-08
slope=-0.15
intercept=0.4
R²=0.34
p=5.7e-14
slope=-0.05
intercept=-0.3
R²=0.21
p=9.7e-09
slope=-0.03
intercept=-0.3
R²=0.11
p=6.6e-05
slope=0.32
intercept=-0.3
R²=0.65
p=7.3e-33
slope=0.30
intercept=-0.4
R²=0.78
p=1.6e-47
slope=-0.20
intercept=1.2
R²=0.27
p=4.5e-11
slope=-0.17
intercept=1.0
R²=0.25
p=3.5e-10
slope=-0.06
intercept=0.2
R²=0.16
p=8.4e-07
slope=-0.05
intercept=0.2
R²=0.19
p=5.4e-08
slope=0.45
intercept=0.1
R²=0.73
p=9.3e-41
slope=0.37
intercept=0.2
R²=0.69
p=3.3e-36
slope=-0.19
intercept=1.1
R²=0.25
p=4.7e-10
slope=-0.21
intercept=1.3
R²=0.43
p=2.3e-18
slope=-0.05
intercept=0.2
R²=0.15
p=1.9e-06
slope=-0.03
intercept=0.4
R²=0.05
p=8.2e-03
slope=0.42
intercept=0.2
R²=0.67
p=8.8e-35
slope=0.37
intercept=0.3
R²=0.77
p=4.5e-45
10^{0}
10^{3}
10^{4}
10^{5}
Number of CpGs
10^{-4}
10^{-2}
α
0
1
λ

Fig. 3

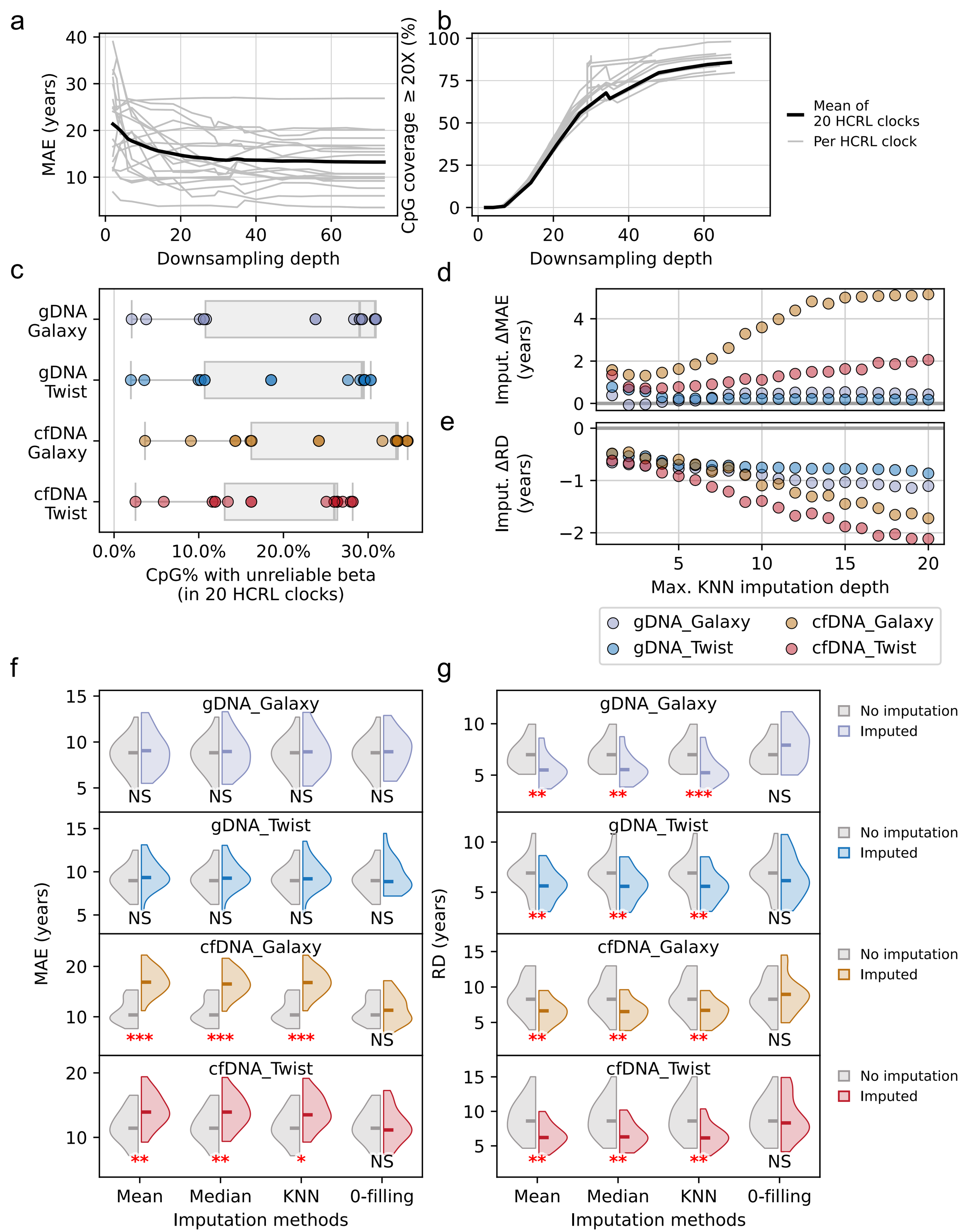
a
MAE (years)
Downsampling depth
b
CpG coverage ≥ 20X (%)
Downsampling depth
Mean of 20 HCRL clocks
Per HCRL clock
c
gDNA Galaxy
gDNA Twist
cfDNA Galaxy
cfDNA Twist
CpG% with unreliable beta (in 20 HCRL clocks)
d
Imput. ΔMAE (years)
Imput. ΔRD (years)
Max. KNN imputation depth
gDNA_Galaxy
gDNA_Twist
cfDNA_Galaxy
cfDNA_Twist
e
f
MAE (years)
gDNA_Galaxy
gDNA_Twist
cfDNA_Galaxy
cfDNA_Twist
NS
Mean
Median
KNN
0-filling
Imputation methods
g
RD (years)
No imputation
Imputed
Imputation methods

Fig. 4

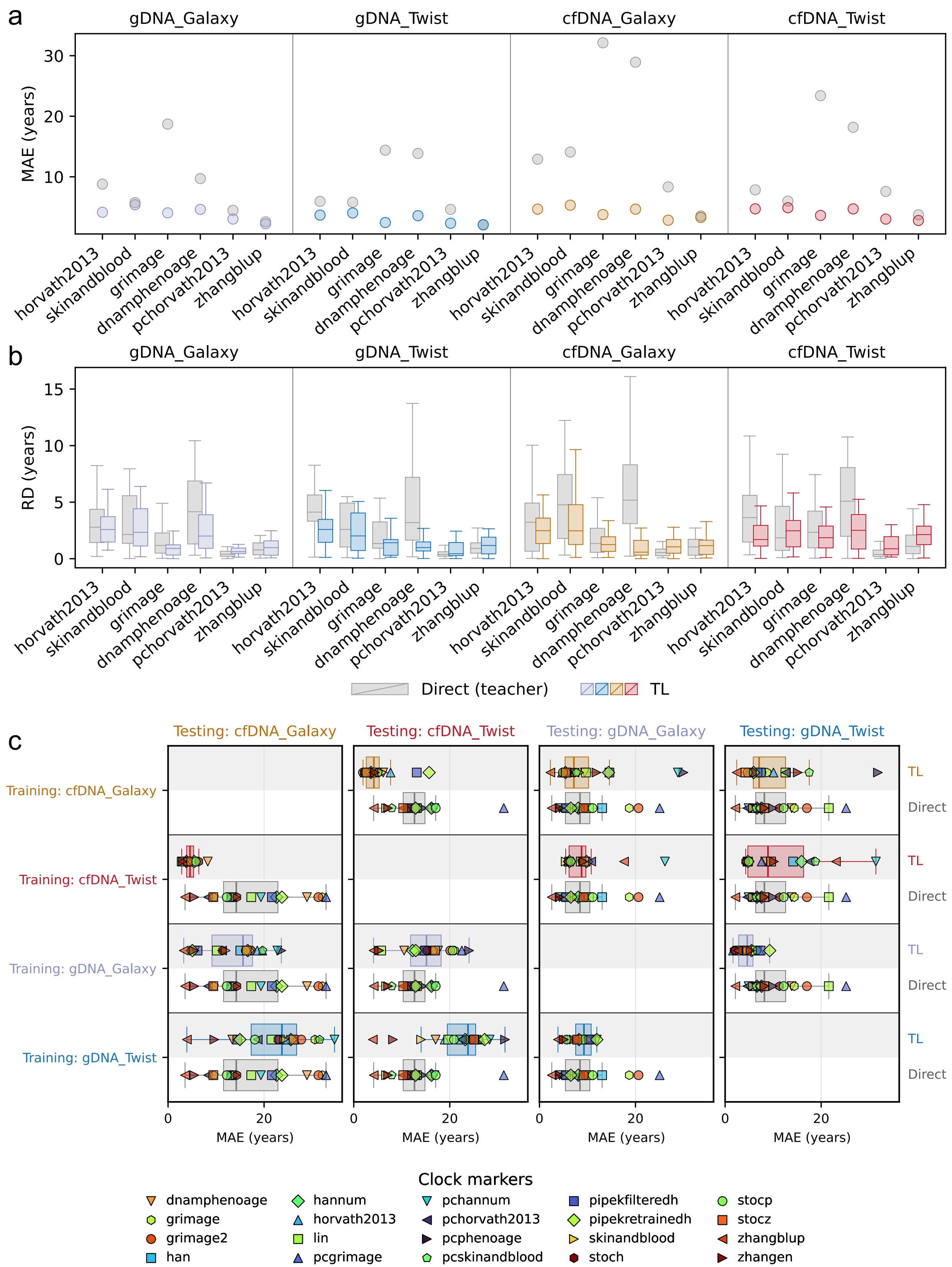
a
gDNA_Galaxy
gDNA_Twist
cfDNA_Galaxy
cfDNA_Twist
MAE (years)
horvath2013
skinandblood
grimage
dnamphenoage
pchorvath2013
zhangblup
b
RD (years)
Direct (teacher)
TL
c
Testing: cfDNA_Galaxy
Testing: cfDNA_Twist
Testing: gDNA_Galaxy
Testing: gDNA_Twist
Training: cfDNA_Galaxy
Training: cfDNA_Twist
Training: gDNA_Galaxy
Training: gDNA_Twist
TL
Direct
MAE (years)
Clock markers
dnamphenoage
grimage
grimage2
han
hannum
horvath2013
lin
pcgrimage
pchannum
pchorvath2013
pcphenoage
pcskinandblood
pipekfilteredh
pipekretrainedh
skinandblood
stoch
stocp
stocz
zhangblup
zhangen

Fig. 5

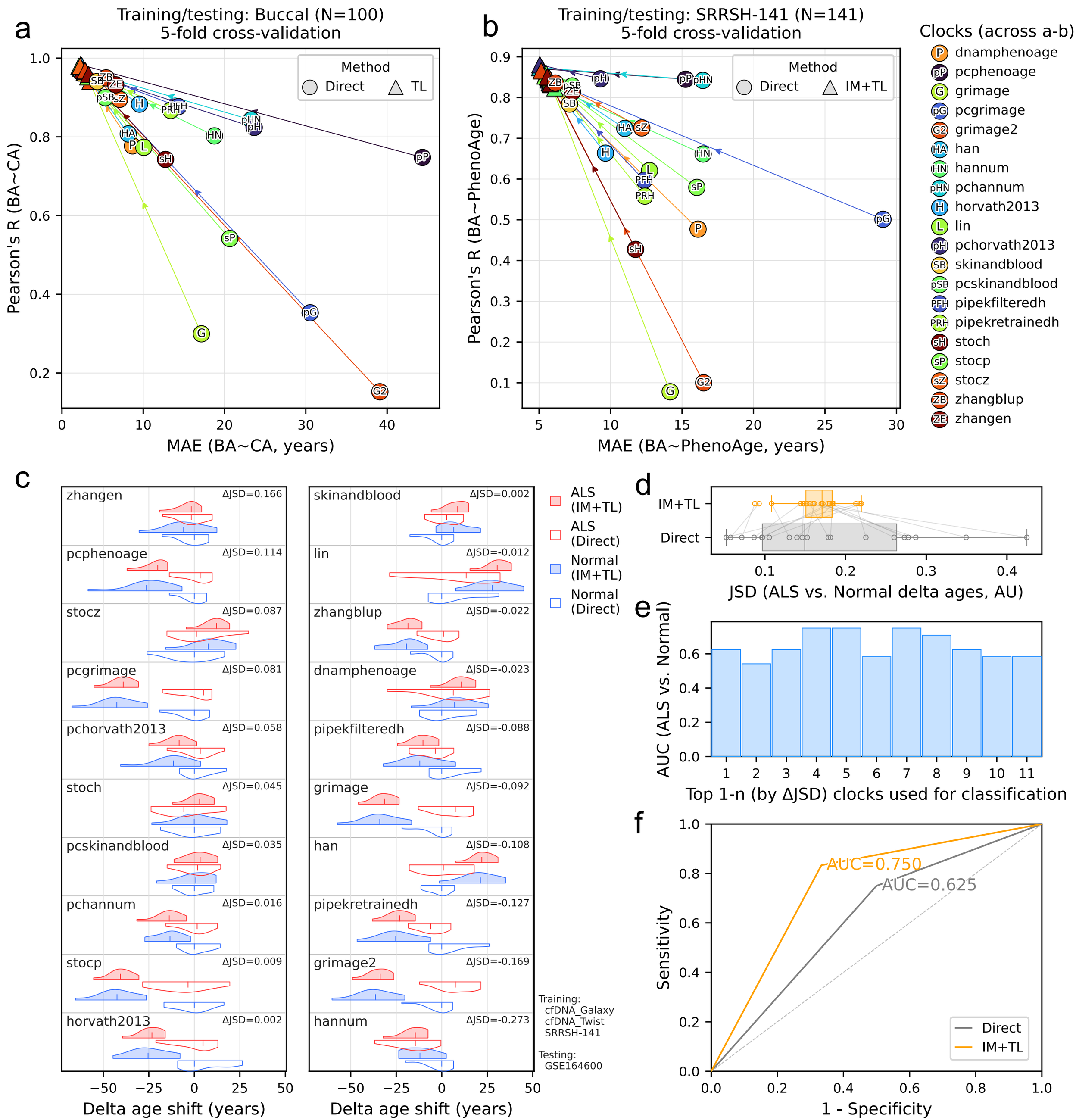
a
Training/testing: Buccal (N=100)
5-fold cross-validation
Method
Direct
TL
Pearson's R (BA~CA)
MAE (BA~CA, years)
b
Training/testing: SRRSH-141 (N=141)
5-fold cross-validation
Method
Direct
IM+TL
Pearson's R (BA~PhenoAge)
MAE (BA~PhenoAge, years)
Clocks (across a-b)
dnamphenoage
pcphenoage
grimage
pcgrimage
grimage2
han
hannum
pchannum
horvath2013
lin
pchorvath2013
skinandblood
pcskinandblood
pipekfilteredh
pipekretrainedh
stoch
stocp
stocz
zhangblup
zhangen
c
ALS (IM+TL)
ALS (Direct)
Normal (IM+TL)
Normal (Direct)
zhangen ΔJSD=0.166
pcphenoage ΔJSD=0.114
stocz ΔJSD=0.087
pcgrimage ΔJSD=0.081
pchorvath2013 ΔJSD=0.058
stoch ΔJSD=0.045
pcskinandblood ΔJSD=0.035
pchannum ΔJSD=0.016
stocp ΔJSD=0.009
horvath2013 ΔJSD=0.002
skinandblood ΔJSD=0.002
lin ΔJSD=-0.012
zhangblup ΔJSD=-0.022
dnamphenoage ΔJSD=-0.023
pipekfilteredh ΔJSD=-0.088
grimage ΔJSD=-0.092
han ΔJSD=-0.108
pipekretrainedh ΔJSD=-0.127
grimage2 ΔJSD=-0.169
hannum ΔJSD=-0.273
Training:
cfDNA_Galaxy
cfDNA_Twist
SRRSH-141
Testing:
GSE164600
Delta age shift (years)
d
IM+TL
Direct
JSD (ALS vs. Normal delta ages, AU)
e
AUC (ALS vs. Normal)
Top 1-n (by ΔJSD) clocks used for classification
f
AUC=0.750
AUC=0.625
Sensitivity
1 - Specificity
Direct
IM+TL

Fig. 6