DNA methylation signatures reflect aging in patients with nonalcoholic steatohepatitis

Rohit Loomba,1 Yevgeniy Gindin,2 Zhaoshi Jiang,2 Eric Lawitz,3 Stephen Caldwell,4 C. Stephen Djedjos,2 Ren Xu,2 Chuhan Chung,2 Robert P. Myers,2 G. Mani Subramanian,2 Zachary Goodman,3 Michael Charlton,1 Nezam H. Afshari,1 and Anna Mae Diehl1

1University of California, San Diego, La Jolla, California, USA. 2Gilead Sciences Inc., Foster City, California, USA. 3Texas Liver Institute, University of Texas Health Science Center, San Antonio, Texas, USA. 4University of Virginia, Charlottesville, Virginia, USA. 5Inova Fairfax Hospital, Falls Church, Virginia, USA. 6Intermountain Medical Center, Salt Lake City, Utah, USA. 7Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, USA. 8Duke Clinical Research Institute, Durham, North Carolina, USA.

Introduction

Nonalcoholic fatty liver disease (NAFLD) is the most common chronic liver condition globally. Although the majority of patients present with steatosis alone, the presence of inflammation and hepatocellular injury in a subgroup of patients with nonalcoholic steatohepatitis (NASH) signifies a risk of progressive fibrosis. Natural history studies of NAFLD have demonstrated that fibrosis stage, regardless of other histologic features, best predicts liver-related mortality (1–3). The assessment of fibrosis, however, represents only a single snapshot of liver injury, without providing a contextual picture of the biological health of the organ as a whole. Novel surrogate markers that measure the age of the liver may provide additional prognostic information beyond staging of fibrosis alone.

Recently, a molecular prediction model based on epigenetic DNA methylation (DNAm) has been proposed as a measure of human chronological and biological age. We determined peripheral blood DNAm in patients with nonalcoholic steatohepatitis (NASH) and assessed whether accelerated aging occurs in these patients. DNAm signatures were obtained in patients with biopsy-proven NASH and stage 2–3 fibrosis. The DNAm profile from one test and two validation cohorts served as controls. Age acceleration was calculated as the difference between DNAm age and the predicted age based on the linear model derived from controls. Hepatic collagen content was assessed by quantitative morphometry. The Horvath clock accurately predicts the chronological age of the entire cohort. Age acceleration was observed among NASH subjects compared with control data sets and our test controls. Age acceleration in NASH subjects did not differ by fibrosis stage but correlated with hepatic collagen content. A set of 152 differentially methylated CpG islands between NASH subjects and controls identified gene set enrichment for transcription factors and developmental pathways. Patients with NASH exhibit epigenetic age acceleration that correlates with hepatic collagen content.
DNAmethylation (DNAm) signature to gauge chronological age has been validated in multiple cohorts and demonstrates predictive capacity across different tissue sites, including the liver (5, 10, 11).

The intrinsic rate of the epigenetic clock can be altered by diseases and metabolic insults related to progressive liver disease. For instance, HIV and obesity are associated with increased liver fibrosis, and both accelerate the epigenetic clock more than would be expected compared with age-matched control specimens (11–13). Interestingly, obesity altered the epigenetic clock of the liver but not for any other tissue sites (11). Acceleration of this clock has functional implications, as age-accelerated DNAm correlates with poorer performance on physical and mental aptitude tests and higher overall mortality, even after adjusting for known risk factors (14, 15). Thus, the Horvath clock can be assayed from peripheral blood, demonstrates concordance with tissue DNAm, and can be used to calculate accelerated aging to provide prognostic information.

Previous studies of DNAm in NAFLD have examined gene-specific epigenetic modifications that may be involved in disease pathogenesis by targeting genes known to regulate fibrogenesis or metabolic function (16–20). For instance, Hardy et al. examined two specific CpG loci on the PPARγ promoter in circulating DNA and found differences in methylation between NAFLD patients with severe fibrosis and those with mild fibrosis (17). Comprehensive integration of genome-wide methylation and gene expression studies has also been performed on liver tissue and been used to identify differential targets between mild and severe NAFLD (18, 20). Limited data exist, however, on whether patients with NASH exhibit a global methylation profile consistent with accelerated aging compared with healthy controls. Therefore, we designed the current study to address the hypothesis that patients with NASH would exhibit accelerated aging compared with controls using the Horvath clock.

In the current study, we determined that the DNAm signature in NASH patients correlates with the severity of hepatic fibrosis. Using three control cohorts (two test, one validation), NASH patients with documented liver fibrosis demonstrated age acceleration compared with healthy controls. DNAm correlates significantly with hepatic collagen content quantified morphometrically and serum markers of fibrosis. We further define two distinct clusters of NASH patients with functional differences in cellular apoptosis and necrosis. These results identify a peripheral blood DNAm signature that links enhanced age acceleration of NASH patients with tissue injury.

**Results**

The study population included 44 patients with NASH and stage 2–3 fibrosis enrolled in a phase II, randomized, controlled trial of selonsertib (formerly GS-4997; ClinicalTrials.gov, NCT02466516). The median age of the NASH cohort was 56 years; the majority were female (66%) and had diabetes mellitus (70%).

Other demographic and baseline characteristics of the NASH patients and the three control populations are included in Table 1. To validate the DNAm age prediction model, we calculated the biological age of the individuals in our cohort and compared it to their chronological age (Figure 1A). The correlation coefficient between DNAm-based and chronological age is 0.84, with a root-mean-square error of 4.7 years. The slope of this line represents the intrinsic rate of DNAm that occurs with aging — the rate of the epigenetic clock during adulthood (4, 5). We next aimed to determine if the NASH cohort is age accelerated (i.e., vertically shifted) compared with normal controls.

We calculated age acceleration in our cohort by first deriving a linear regression model between DNAm age and chronological age in control samples and then comparing the distributions of residuals of NASH and control samples. The median age acceleration of NASH samples is 2.8 years (interquartile range [IQR]: −0.35–4.68), which is significantly greater than that observed among age- and sex-matched healthy controls ($P = 0.03$; Figure 1B). We then validated these findings using publically available DNAm data sets containing healthy controls: the European Prospective Investigation into Cancer and Nutrition (EPIC; http://epic.iarc.fr/) data set and the Hannum et al. (4) data set. We found that the NASH cohort was age accelerated compared with controls in the EPIC data set (median age acceleration = 5.3 years; $P = 1.33 \times 10^{-5}$; Figure 1C) and in the Hannum et al. data set (median age acceleration = 1.9 years; $P = 0.03$; Figure 1D). Taken together, these results show that biological age, based on DNAm profile, is highly correlated with chronological age. Furthermore, compared with three independent data sets of healthy controls, the NASH patient cohort demonstrates significant biological age acceleration.

Since liver fibrosis is the defining feature that predicts mortality in NASH, we next evaluated whether age acceleration was associated with the severity of liver fibrosis. Compared with control subjects, DNA methylation is significantly associated with hepatic fibrosis. We next evaluated whether this association could be used to predict the severity of hepatic fibrosis. We found that the NASH cohort was age accelerated compared with healthy controls. DNAm correlates significantly with hepatic collagen content quantified morphometrically and serum markers of fibrosis. We further define two distinct clusters of NASH patients with functional differences in cellular apoptosis and necrosis. These results identify a peripheral blood DNAm signature that links enhanced age acceleration of NASH patients with tissue injury.
from NASH patients with stage 2 fibrosis did not demonstrate increased age acceleration ($P = 0.12$), while those with stage 3 fibrosis demonstrated modestly increased age acceleration ($P = 0.04$; Figure 2A). Since the histological staging of liver fibrosis is limited by an ordinal scoring system, we evaluated the associations between age acceleration and two continuous measures of liver fibrosis, namely hepatic collagen content determined morphometrically and the enhanced liver fibrosis (ELF) score, a serum marker of fibrosis. As shown in Figure 2B, a greater proportion of patients with NASH exhibited age acceleration compared with controls in the test cohort (68% [30 of 44] vs. 55% [10 of 18]). In the NASH cohort, age acceleration was significantly associated with hepatic collagen content ($\rho = 0.45; P = 0.004$; Figure 2, C and D). Similarly, serum ELF levels were positively correlated with age acceleration ($\rho = 0.35; P = 0.03$; Supplemental Figure 1; supplemental material available online with this article; https://doi.org/10.1172/jci.insight.96685DS1). Thus, these results identify an association between accelerated aging in NASH patients and liver fibrosis.

To determine whether NASH patients can be stratified based on a methylation signature, we performed genome-wide methylation analysis to examine differences in the status of CpG islands in controls versus NASH patients. After adjusting for multiple testing, this analysis identified 152 differentially methylated CpG islands, 120 of which exhibited greater methylation in NASH patients (Figure 3A). An enrichment analysis of the genes associated with the differentially methylated CpG islands revealed significant enrichment of developmental signaling pathways, including Wnt/\(\beta\)-catenin signaling and thyroid hormone/retinoid X receptor activation (Figure 3B). Hierarchical clustering of NASH and control samples based on the 152-CpG island signature identified two distinct clusters within the NASH cohort (Figure 3A). Since methylation of developmental pathways suggests functional differences in cell survival, we evaluated differences in cellular apoptosis/necrosis based on serum levels of cytokeratin-18 (CK-18) M30 and M65 fractions by ELISA. Levels of CK-18 M30 and M65 at baseline were significantly greater in NASH cluster 2 patients (Figure 3, C and D), indicating that a subset of NASH patients exists with increased cellular apoptosis/necrosis based on their DNAm pattern.

Discussion

The ability to measure a surrogate marker of liver aging from a peripheral blood sample has broad implications for assessing clinically “silent” chronic diseases, such as NASH, and, potentially, their response to interventions. In the current study, we validate the utility of the Horvath clock in measuring age acceleration in a defined cohort of NASH patients with moderate to severe liver fibrosis. Using three different control data sets — two validation and one test cohort — the intrinsic rate of DNAm that occurs with aging was determined. We then ascertained that the DNAm signature of NASH patients was age accelerated approximately 5 years compared with public control data sets and 3 years versus our test control cohort. Further, based on two continuous measures of fibrosis, namely hepatic collagen content assessed morphometrically and the serum ELF test, we identified an association between NASH-related fibrosis and the

### Table 1. Demographic and baseline characteristics of NASH patients and the three control populations

<table>
<thead>
<tr>
<th></th>
<th>NASH ($n = 44$)</th>
<th>Healthy controls ($n = 18$)</th>
<th>Controls (data set 1) ($n = 379$)</th>
<th>Controls (data set 2) ($n = 321$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (y)</td>
<td>56 (48–65)</td>
<td>49 (42–56)</td>
<td>58 (46–71)</td>
<td>55 (45–65)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>European descent, 61%; Hispanic, 30%; Asian, 7%; other, 2%</td>
<td>European descent, 100%</td>
<td>European descent, 48%; Hispanic, 52%; Unavailable</td>
<td></td>
</tr>
<tr>
<td>Male ($n$ [%])</td>
<td>15 (34)</td>
<td>7 (38)</td>
<td>181 (48)</td>
<td>79 (25)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>32 (24–40)</td>
<td>27 (20–34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes ($n$ [%])</td>
<td>31 (70)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liver biochemistry</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALT (U/l)</td>
<td>63 (20–105)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGT (U/l)</td>
<td>62 (14–109)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELF</td>
<td>9.9 (9.0–10.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAS 7–8</td>
<td>9 (20%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3 fibrosis</td>
<td>29 (66%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatic collagen (%)</td>
<td>3.8 (2.7–5.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data presented as median (IQR). ALT, alanine aminotransferase; GGT, \(\gamma\)-glutamyl transpeptidase; ELF, enhanced liver fibrosis; NAS, NAFLD activity score.
degree of age acceleration. In sum, a peripheral DNAm signature in NASH patients identified age acceleration that reflected the extent of liver fibrosis.

The current study adds to the emerging literature examining the role of epigenetics and the NAFLD phenotype. Using liver biopsy tissue from NAFLD patients, Sookoian et al. identified methylation of promoter sites on PPARγ coactivator 1α (PPARGC1A) with markers of insulin resistance such as HOMA-IR and serum insulin (19). Other investigators have demonstrated that specific methylation sites on promoters of genes are linked to more severe liver fibrosis (17, 18). Although these reports establish a potential relationship between fibrosis and specific epigenetic modifications, targeted individual CpG sites may not accurately reflect the complex interaction between causal and compensatory measures in chronic diseases such as NASH. Thus, utilization of an aggregate DNAm profile obtained from peripheral blood that has been evaluated across different tissue types may provide a valuable measure of overall liver health. Here, we provide confirmation that the Horvath epigenetic clock, a measure of DNAm age, correlates with the degree of liver fibrosis in biopsy-proven NASH.

The results presented herein are consistent with studies of other systemic diseases that are characterized by age acceleration. For instance, HIV infection has been found to significantly accelerate age-related methylation by nearly 14 years in some studies (12, 21). These findings may have relevance to the current study, since HIV is a major determinant of liver fibrosis progression in chronic liver diseases, such as hepatitis C virus (HCV) infection. Prospective studies of patients with HCV and HIV-HCV coинфекion confirm more rapid fibrosis

![Figure 1. NASH patients are age accelerated compared with healthy controls. (A) DNAm age as a function of chronological age. (B) Comparison of age acceleration between NASH and test control cohorts. (C) Acceleration of DNAm between NASH subjects and the validation control cohort derived from the EPIC data set (http://epic.iarc.fr/). (D) Comparison of age acceleration between NASH and control cohorts derived from the Hannum et al. data set (4). P values were computed using a 2-tailed Mann-Whitney U test.](https://insight.jci.org)
progression in coinfected patients using paired liver biopsies (22, 23). Other conditions, such as obesity, also contribute to age acceleration but to a lesser extent (11). Therefore, DNAm age acceleration appears disease specific and modified by environmental factors. Given this complexity, a global signature of the health of the individual reflected in DNAm may serve as an indicator of the chronic injury found in liver diseases, such as NAFLD. Whether a DNAm signature captures those NASH patients at high risk of developing future complications, such as hepatocellular carcinoma, remains unknown and warrants investigation.

The findings in the current work are in contrast to an earlier study that found no association of DNAm with the NAFLD activity score or stage of liver fibrosis in patients with NASH (11). Importantly, that study assessed liver fibrosis based on conventional histological staging only, using the ordinal METAVIR classification. Similarly, we also found no difference in age acceleration between patients with stage 2 and 3 fibrosis according to the NASH Clinical Research Network (CRN) classification. On the contrary, by evaluating two continuous measures of fibrosis (hepatic collagen content by morphometry and the serum ELF test), which have a greater dynamic range than traditional histological staging, we found that patients with higher age acceleration have increased hepatic fibrosis. These findings illustrate the utility of continuous liver fibrosis scales that cannot be ascertained with traditional pathological staging. Whether DNAm is associated with other noninvasive measures of liver fibrosis is unknown and warrants investigation.

Figure 2. Age acceleration is associated with liver fibrosis and collagen content. (A) Age acceleration in control and NASH data sets according to liver fibrosis score (F2 or F3). (B) The magnitude and direction of age acceleration by disease status show enrichment of NASH subjects in those with positive age acceleration. Each bar represents a sample in the cohort; color corresponds to disease status. (C) Relationship between percentage collagen at baseline (y axis) and age acceleration (x axis). Each dot represents a sample in the NASH cohort. (D) The magnitude and direction of age acceleration by disease status. Each bar represents a sample in the cohort; color corresponds to disease status. P values were computed using a 2-tailed Mann-Whitney U test.
Figure 3. NASH patients have a distinct methylation profile. (A) Heatmap of methylation values at CpG islands (rows) by samples (columns). Two distinct NASH sample clusters are annotated. (B) Pathway enrichment analysis of differentially methylated CpG islands between NASH and control cohorts. (C and D) Relationship between M30 and M65 ELISA values and NASH sample clusters 1 and 2, respectively. *P* values were computed using a 2-tailed Mann-Whitney *U* test.
Limitations of our study are its relatively small sample size, its inclusion of predominately subjects of European descent, and the absence of specimens representing the full spectrum of liver fibrosis, early (F1) fibrosis and cirrhosis (F4). This study, however, included patients who were part of a clinical trial in which protocol-obtained biopsies were read by a central pathologist (ZG) and morphometric quantification of collagen standardized. Recent work indicates that the intrinsic pace of DNAm does vary by ethnicity (24). Thus, our study represents an important initial study in a population of mostly European descent with NASH, which identifies age acceleration and its correlation with fibrosis content. Future studies that examine the full spectrum of the NASH phenotype in different racial groups may identify specific DNAm aging patterns based on ethnicity.

In summary, a DNAm signature obtained from the peripheral blood of NASH patients identifies epigenetic age acceleration that correlates with increased liver fibrosis. These findings have potential prognostic implications and may serve as a surrogate marker of therapeutic efficacy in NASH.

Methods

Patients and control data sets. The study population was derived from a phase II study of selonsertib in patients with biopsy-proven NASH and stage 2 or 3 fibrosis according to the NASH CRN classification (ClinicalTrials.gov, NCT02466516). Liver biopsies were obtained at baseline. Age- and sex-matched healthy control peripheral blood mononuclear cells (PBMCs) were acquired from Conversant Biologics.

The data sets for validating the control population were obtained from previously published work (4, 25). The data set by Hannum et al. (4) was downloaded from Gene Expression Omnibus (GSE40279) as β values using the GEOquery R package (26). The EPIC data set was downloaded from ArrayExpress (E-GEOD-51032) as IDAT files and processed in the same manner as data generated for this study (25). We restricted the subjects’ ages in the Hannum and EPIC data sets to between 40 and 69 years to reflect the age range of the NASH subjects in our cohort. For the EPIC data set, we excluded all subjects that have been annotated with a disease category.

Sample collection and methylation analysis. DNA extracted from PBMCs was assayed for cytosine methylation using the Infinium Methylation Assay protocol for the Infinium Human Methylation450K BeadChip (Illumina). In brief, DNA was treated with sodium bisulfite to convert unmethylated cytosines to uracil, leaving methylated cytosines unchanged. The treated DNA sample was then denatured, neutralized, and isothermally amplified. The amplified DNA was fragmented, precipitated with isopropanol, and resuspended prior to hybridization onto BeadChips. The converted and nonconverted amplified DNAs were hybridized to their corresponding probes, and excess DNA was washed away. Hybridized DNAs then underwent single-base extension and staining for labeling, followed by scanning on an Illumina iScan instrument for detection. Scanned images were analyzed using Illumina Genome Studio software.

Data processing. Data were processed using the minfi R Bioconductor package version 1.2 (27) and watermelon package version 1.18 (28).

Epigenetic age and determination of differential DNAm. Epigenetic age, based on the Horvath model, was calculated from the methylation β values using the minfi’s age function. Age acceleration was calculated as the difference between an individual’s DNAm age and one calculated from a linear model derived using the control samples (5). Differential DNAm analysis was restricted to genomic locations that do not overlap known SNPs and those that lie within CpG islands. Methylation β values were first quantile normalized and then converted to M values. CpG island methylation values were calculated using an average value of individual probes that span it. Tests for differential methylation were carried out using the limma R package version 3.3 (29). An FDR-adjusted P value of 0.05 was used to identify differentially methylated CpG islands. Associated between genes and CpG islands was done by genomic distance.

Ingenuity pathway analysis. Genomic DNAm and the relationship between different patterns of gene methylation were analyzed using Ingenuity Pathway Analysis software (Ingenuity Systems). Networks of methylated genes were algorithmically generated based on their connectivity and assigned a score. Scores were used to rank signaling and developmental pathway networks according to their methylation content.

Hepatic collagen determination. Sections of each liver biopsy stained with Picrosirius red were used for hepatic collagen content determination as previously described (30). A digital image of the entire stained section was acquired with an Aperio Scanscope CS (Aperio Technologies Inc.). The pixel count was obtained according to the manufacturer’s instructions to quantify the fraction of positively stained pixels. Results were expressed as a fraction of the entire biopsy specimen that was positive for the red staining. The central pathologist (ZG) reviewed biopsy specimens in a blinded fashion.
Serum markers of fibrosis. The ELF panel was determined as previously described (31). Approximately 20 ml blood was obtained under fasting conditions. Procollagen III N-terminal peptide (PIIINP), hyaluronic acid (HA), and tissue inhibitor of metalloproteinase-1 (TIMP-1) were measured by an automated clinical immunochemistry analyzer that performs magnetic separation enzyme immunoassay tests (ADVIA Centaur, Siemens Healthcare Diagnostics). The ELF score was calculated using the algorithm (ELF = 2.278 + 0.851 ln(HA) + 0.751 ln(PIIINP) + 0.394 ln(TIMP-1)).

CK-18 M30 and M65 ELISA. CK-18 M30 Apoptosense and M65 ELISA kits were obtained from PEVIVA AB and were performed according to the manufacturer’s instructions as previously described (32). Values are expressed in units, where 1 U/l is equivalent to 1.24 pM of a synthetic 21–amino acid fragment of CK18 containing both M30 (e.g., M30) and M65 (e.g., M5) epitopes. Patient plasma samples were run in duplicate, and values were plotted against standard curve.

Statistics. A 2-tailed Mann-Whitney test was used to test the statistical significance of age acceleration in two groups. A \( P \) value of less than 0.05 was considered significant.

Study approval. The study was approved by the institutional review boards of the University of California, San Diego, Gilead Sciences Inc., Texas Liver Institute, University of Virginia, Inova Fairfax Hospital, Intermountain Medical Center, Beth Israel Deaconess Medical Center, Harvard Medical School, and Duke Clinical Research Institute. Written informed consent was obtained from each individual prior to his or her participation in the study.

Author contributions
RL, GMS, RPM, and AMD conceived the project and designed experiments. YG, ZJ, CC, and RPM performed analysis. RL, YG, CC, RPM, and GMS wrote the manuscript. RL, YG, ZJ, EL, SC, CSD, RX, CC, RPM, GMS, ZG, MC, NHA, and AMD reviewed the data and the manuscript.

Acknowledgments
This work was supported by Gilead Sciences Inc. We thank Lakshmi Alaparthi and Fanny Monge for assistance with the morphometry analyses.

Address correspondence to: Rohit Loomba, Division of Gastroenterology, 9500 Gilman Drive, University of California, San Diego, La Jolla, California, USA. Phone: 858.534.2624; Email: roloomba@ucsd.edu.

References