Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Mitchell Lee Date

Associations of DNA methylation with biomarkers of liver function among male veterans living with HIV

By

Mitchell Lee
Master of Science in Public Health

Epidemiology

Yan Sun, PhD
Committee Chair

Associations of DNA methylation with biomarkers of liver function among male veterans living with HIV

By

Mitchell Lee

B.S.
Duke University
2017

Thesis Committee Chair: Yan Sun, PhD

An abstract of
a thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Epidemiology
2021

Abstract

Associations of DNA methylation with biomarkers of liver function
among male veterans living with HIV

By Mitchell Lee

Background: Chronic liver disease is a prominent cause of morbidity and mortality for people with HIV. The etiology of liver disease and function in people with HIV, however, is not fully understood. This thesis presents epigenome-wide association studies (EWAS) to identify associations between DNA methylation and seven biomarkers of liver function—serum levels of aspartate transaminase, alanine transaminase, and albumin, and total bilirubin, platelet count, FIB-4 score, and APRI score—among a cohort of male US veterans with HIV.

Methods: Blood samples and clinical data were obtained for 960 HIV+ male veterans from the Veterans Aging Cohort Study (VACS). Blood DNA methylation was assessed using the HumanMethylation450 (450K) array or the HumanMethylationEPIC (EPIC) array from Illumina, which cover over 450,000 and 870,000 DNA methylation (DNAm) sites, respectively. Associations between DNAm age acceleration (AA) and selected liver biomarkers were assessed by regressing biomarker values on IEAA, EEAA, PhenoAA, and GrimAA in linear models controlling for covariates. Associations between individual DNAm sites and selected liver biomarkers were assessed by separate EWAS of the EPIC and 450K sub-cohorts using mixed effect models controlling for covariates and batch effects. For DNAm sites measured by both platforms, meta-analysis of the separate sub-cohort EWAS results for each liver biomarker was performed.

Results: Significant association was observed between PhenoAA and serum albumin ($\beta = -0.007$, *P*-value = 8.6×10^{-4}) among all AA measurements and liver biomarkers. Nine DNAm sites annotated to the *TMEM49*, *SOCS3*, *FKBP5*, *ZEB2*, and *SAMD14* genes were significantly associated with serum albumin in the meta-analysis of the EPIC and 450K EWAS results. Beta coefficients from the separate EPIC and 450K EWAS results for those DNAm sites were positively correlated, indicating consistency between the EPIC and 450K cohorts. No significant associations were detected for the six other biomarkers after meta-analysis.

Conclusion: The EWAS results suggest that the *TMEM49*, *SOCS3*, *FKBP5*, *ZEB2*, and *SAMD14* genes might be linked to liver function through serum albumin. Since this is the first EWAS of liver function among people with HIV, further replication analyses in independent cohorts are warranted to confirm the epigenetic mechanisms underlying liver function.

Associations of DNA methylation with biomarkers of liver function among male veterans living with HIV

By

Mitchell Lee

B.S.
Duke University
2017

Thesis Committee Chair: Yan Sun, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Epidemiology
2021

Acknowledgements

Completion of this thesis despite the many challenges imposed by the COVID pandemic would not have been possible without the support of many people who provided essential methodological guidance and constant reassurance throughout trying circumstances.

I am indebted to my thesis advisor, Dr. Yan Sun, for this thoughtful assistance with identifying and framing the aims of this thesis, and for his deep methodological expertise that enabled me to conduct rigorous work.

I am also grateful to Qin Hui and Zeyuan Wang, who provided much-appreciated direct assistance with data management and quality control tasks. Without their help, this thesis would have been built on a column of sand.

I am also sincerely grateful to the other members of the Sun lab for sharing in the learning process. They helped me with my mistakes and openly shared your own so that we could all improve together.

I also owe sincere thanks to Keven Haynes, who worked tirelessly to grace me with access to a powerful new HPC cluster at Emory and who assisted time and time again with my issues as I learned to deploy statistical programs in parallel on that cluster. Without his help, my analyses would still be running.

Finally, I thank my family for offering unshakeable support and encouragement during this process—for hugging me when I was down, kicking me into gear when I needed it, and supplying so many late-night cups of tea.

Cheers!

Table of Contents

INTRODUCTION

Chronic liver disease (CLD)—defined as long-lasting diminished capacity of the liver to perform its many essential metabolic functions (1, 2)—is a prominent cause of morbidity, mortality, and healthcare spending for people living with HIV (3-7) (see literature review below for more details). Although CLD can often be prevented or intervened through behavioral changes and alleviation of treatable underlying causes (8-10), liver transplantation remains the only treatment option for patients suffering liver failure as a result of CLD (11-13). As such, and because the number of people living with HIV and their life expectancies are projected to increase in coming decades (14-16), there is an urgent need to further improve and develop methods to prevent, predict, diagnose, and treat CLD. Along with more epidemiologic studies to identify modifiable risk factors, further expansion of biomedical knowledge of the physiology and etiology of CLD will benefit that effort.

Progress toward a complete understanding of the biological mechanisms underlying CLD is encumbered by the complexity of the liver's metabolic roles (17-21), which make disentangling the underlying mechanisms difficult, and the subtly of the effects that many genes and regulatory mechanisms have on controlling its activity (22, 23), which can make detection of those effects difficult. As a discipline that applies statistical analysis of large datasets to enable detection of subtle associations at the population level between biological phenomena and the human genome and its regulation (24-26), epigenetic epidemiology has great potential to help overcome those barriers, potentially leading to the generation of new hypotheses that yield actionable insights into prevention and treatment of CLD.

Epigenetic epidemiology can be especially useful in elucidating biological roles and effects of DNA methylation, wherein DNA becomes chemically modified with methyl groups that can repress gene expression by inhibiting transcription machinery and stimulating repressive

condensation of DNA structure (27, 28). Epigenome-wide association studies (EWAS) can identify subtle effects of epigenetic regulation in a multifaceted condition like CLD, but require large sample sizes and high-throughput computational methods to detect effects from vast arrays of methylation events that happen throughout the human genome (29-31). Epigenomic epidemiologic studies have, therefore, endeavored to identify DNA methylation events that associate with CLD (see literature review below). HIV populations are, however, underrepresented in those studies. Moreover, those studies have focused on CLD as an overall clinical diagnosis and largely ignored looking for potential associations between DNA methylation and specific metabolic functions of the liver as represented by biomarkers that are regulated by those functions. Because such biomarkers sometimes directly indicate capacity to perform specific independent aspects of liver function, and because their modulation also sometimes constitutes the negative downstream consequences of liver disease (32, 33), such gaps in the current body of research might impede specific and detailed understanding of the mechanisms underlying CLD and its deleterious downstream consequences.

In recognition of the need to fill the gap, this thesis presents efforts to identify associations between DNA methylation and seven markers of liver function—aspartate transaminase, alanine transaminase, serum albumin, total bilirubin, platelet count, FIB-4 score, and APRI score—that are commonly used in clinical settings to assess specific facets of liver function, as reviewed below (34-37). Specifically, this thesis presents results from two primary avenues of investigation: (1) linear association studies of whether acceleration of DNA methylation age relative to chronological age associates with variation in levels of the selected liver biomarkers; and (2) a series of epigenome-wide association studies (EWAS) to identify specific CpG sites across the autosomal portion of the human genome where DNA methylation associates with differences in

levels of the selected liver biomarkers. We anticipate identifying genes involved in the physiology of liver function that could be included as candidates in further, more focused studies of the genetic, molecular, and cellular bases of liver dysfunction, specifically in the context of people living with HIV, who suffer higher burden to CLD.

LITERATURE REVIEW

Burden of Chronic Liver Disease Among People with HIV

Epidemiologic studies have demonstrated for decades that liver disease is a common comorbidity of HIV infection that accounts for a substantial proportion of mortality among people living with HIV (38). Evidence of this association first emerged as early as 1997, when a study of non-AIDS mortality after seroconversion among European drug injectors living with HIV found that 12.3% of deaths resulted from liver disease (39). A more recent study of mortality among a large cohort of people living with HIV in Europe, Australia, and the US during the early 2000s shows that the problem continues to grow, with 14.5% of deaths resulting from liver disease (40). The mortality among people living with HIV caused by liver disease is likely even higher today across the globe, where HIV and liver disease are both currently estimated to affect tens of millions of people with no anticipation for shrinkage of their prevalence (41-46).

The physical and psychological suffering that liver disease imposes on people living with HIV is compounded by the significant financial costs that caring for CLD incurs for affected individuals, healthcare institutions, and insurance companies. A review of the economic burden imposed by cirrhosis estimates that more than 7 billion dollars were spent on care for chronic liver disease in the US alone in 2014 (47). According to that same review, that overall spending can translate to tens of thousands of dollars in cost per year for individuals (47), a burden that can be increased by the need for liver transplantation, which can cost hundreds of thousands of additional dollars (48). For people living with HIV, who already often struggle to carry the economic burden imposed by paying for anti-retroviral therapy and clinical care (49), the added costs imposed by CLD can make pursuing treatment unfeasible, impeding progress toward reducing the prevalence of CLD and preventing subsequent mortality.

Physiology and Etiology of Chronic Liver Disease

The physiology of CLD has been reviewed extensively (50-52). Rare genetic disorders can lead to CLD by disrupting metabolic pathways that perform specific metabolic activities, thereby impairing specific functions of an otherwise health liver (53). In Wilson's disease, for example, mutation of the *ATP7B* gene that encodes a copper transport protein leads to a toxic accumulation of copper in the body that damages liver cells (54). The majority of clinically relevant CLD, however, results from either progressive scarring due to chronic injury of liver cells or to the development of liver cancer (50-52). In early stages, scarring of the liver, known as fibrosis, often causes no or minor symptoms and is usually reversible if the cause of the chronic damage is removed (55-58). Sustained scarring, however, can progress to a generally irreversible stage known as cirrhosis, where scar tissue accumulates and displaces healthy tissue until too few hepatic cells remain to sustain liver function (59-61). The most common form of liver cancer is hepatocellular carcinoma, which also inhibits liver function by displacing healthy cells (62-64). With both cirrhosis and hepatocellular carcinomas, the resulting disruption of liver function can intensify to the point of total liver failure, which is followed quickly by death (65, 66) and thus accounts for much mortality among people with CLD.

Many causes for chronic liver damage that instigate progression through hepatic fibrosis to cirrhosis have been identified (67). One that has been long-recognized is excessive, long-term consumption of alcohol, which harms the liver when the hepatic cells that metabolize ethanol are damaged by its dehydrating effects and the toxic byproducts of its metabolism (68-70). Another well-established cause is infection with hepatitis viruses, particularly the hepatitis C virus (71-74). A final major source that has become increasingly prevalent in recent decades is non-alcoholic fatty liver disease (NAFLD), wherein adipose tissue accumulates in and on the liver, leading to

replacement of healthy functional liver cells and damaging inflammation of those that remain (75, 76). Why NAFLD occurs is not clear and continues to attract intense research.

People living with HIV face additional drivers of CLD that are specific to their infection with HIV (77). Infection with HIV associates with increased risk of viral hepatitis, likely because the hepatitis viruses share a common infection route with HIV and cannot be cleared as effectively by the weakened immune systems of people living with HIV (78-80). Thus, people living with HIV show increased prevalence of CLD attributable to viral hepatitis (81). In fact, a study found that 76% of the deaths from liver disease among people living with HIV occurred in individuals who were coinfected with HBV and/or HCV, indicating that the hepatitis viruses contribute substantially to the burden of liver disease among people living with HIV (40). Indeed, another study of mortality among people living with HIV in France in 2000 that found that 97% of those who died of liver disease were infected with HBV and/or HCV (82). Moreover, the HIV virus itself can directly harm liver cells by infecting certain hepatic mesenchymal cells and by triggering harmful inflammatory immune responses within the liver that bring collateral damage (77). Finally, many anti-retroviral medications prescribed to treat HIV can have hepato-toxic side effects that can contribute to the progression of fibrosis (83, 84). Contextualizing the findings presented in this thesis within what is already known epidemiologically and biomedically as described above will be critical.

DNA Methylation as a Mechanism of Gene Regulation

DNA methylation involves the chemical bonding of a methyl group to a nucleotide in a strand of DNA (27, 28). In mammals, DNA methylation occurs predominantly at cytosine nucleotides followed by guanosine nucleotides, which together form a locus known as a CpG site (27, 28).

When CpG sites are located near or within the promoter or coding sequences of a gene, methylation at that CpG site can suppress expression of that gene by sterically inhibiting binding of transcriptional enzymes to the DNA (85) and by stimulating recruitment of molecular machinery that reorganizes DNA into compact chromatin that cannot be accessed by molecular transcription machinery (86). As such, DNA methylation can contribute to disease by instigating deleterious repression of expression of genes that play vital roles in maintaining healthy biological processes. Indeed, repression of gene expression by DNA methylation has been shown to contribute to range of disease, especially congenital defects arising during fetal development (87-90). As such, and because DNA methylation is estimated to occur at nearly 80% of CpG sites (27, 28), investigating how DNA methylation might be involved in liver dysfunction and disease among people living with HIV might contribute substantially to understanding the biological mechanisms underlying liver disease and dysfunction.

DNA Methylation Age Acceleration in People with HIV

Several studies have demonstrated that people infected with HIV exhibit accelerated DNA methylation characterized by global levels of DNA methylation that are elevated above the average level exhibited by non-infected people of equivalent chronological age. Rickabaugh *et al.* (2015), for example, demonstrated that HIV infection associates significantly with increased methylation of a select set of CpG sites in blood cells sampled from older and younger men who have sex with men, even when controlling for age (91). Horvath *et al.* (2015) similarly demonstrated that HIV infection associates significantly with increased DNA methylation in samples of brain tissue, whole blood, and leukocytes taken from several different populations of men and women (92). Horvath *et al.* (2018) also demonstrated that perinatally acquired HIV associates with increased

DNA methylation among adolescents from South Africa (93). The repeated observation of this association between HIV infection and accelerated DNA methylation has led to the hypothesis that accelerated DNA methylation contributes to the etiology by which HIV associates with the relatively early onset of many conditions among people living with HIV.

Relative acceleration of DNA methylation is often measured by four metrics: intrinsic epigenetic age acceleration (IEAA), extrinsic epigenetic age acceleration (EEAA), phenotypic age acceleration (PhenoAA), and Grim age acceleration (GrimAA). IEAA is calculated as the residual of a person's DNA methylation level in blood cells across a set of CpG sites proposed by Horvath (92) after regressing DNA methylation level on chronological age while controlling for proportions of different leukocytes within the sample (94). EEAA is calculated similarly, except that a different set of CpG sites proposed by Hanum (95) are used while controlling for proportions of a different set of cell types (96). PhenoAA, in contrast, is calculated by first estimating phenotypic age (in years) with a linear regression model that uses clinical variables as inputs, then estimating DNA methylation age (in years) based on methylation levels at a set of CpG sites proposed by Levine *et al.* (2018), and finally calculating the residual that results from regressing calculated DNA methylation age on calculated phenotypic age (97). GrimAA is calculated as with PhenoAA, except that levels of plasma biomarkers indicative of physiological stress are used to estimate phenotypic age and a different set of CpG sites are used to assess DNA methylation age (98). Each of these four measures of relative acceleration of DNA methylation age performs better than the others for predicting different outcomes. As such, testing all four metrics has become standard practice for DNA methylation studies.

Associations of DNA Methylation with Liver Function

Research on the role of DNA methylation in chronic liver disease has focused predominantly on NAFLD, as clear etiologies for NAFLD cannot be as easily identified as with CLD resulting from alcoholism or viral hepatitis. There is therefore a need to conduct exploratory epidemiological studies to identify associations with DNA methylation that might hint at genes involved in underlying biological mechanisms for CLD. Zhang *et al.* (2020) recently conducted an exhaustive literature review and summarized results from eleven studies that investigated associations between DNA methylation and NAFLD (99). Two of those studies investigated the association between NAFLD and global DNA methylation using methylation-specific PCR or bisulphite pyrosequencing and samples of liver tissue obtained from South American and European participants (100, 101). Highlights from the results include that hypermethylation of the mitochondrial NAHD dehydrogenase gene and of the *LINE-1* transposable element is associated with NAFLD. Four of the studies investigated the association between NAFLD and methylation in or near specific candidate genes using bisulphite pyrosequencing or methylation-specific PCR and samples of liver tissue or peripheral blood (102-105). Highlights from the studies that used peripheral blood samples include the observation that hypomethylation of mitochondrial DNA relative to genomic DNA associates with NAFLD, while methylation of *GSTT1* and *GSTP1* did not associate with NAFLD. Highlights from the studies that used samples of liver tissue include detection of an association of NAFLD with hypermethylation at the *PNPLA3* gene, and with hypomethylation at the *PARVB*, *MATIA*, *CAPS1*, and *FGFR2* genes. Four of the remaining five studies conducted EWAS for NAFLD using the Infinium HumanMethylation450 array (Illumina) and samples of liver tissue or whole blood (106-109). Collectively, the studies that used liver tissue identified associations between methylation at the *PAPLN*, *LBH*, *DPYSL3*, *JAG1*, *NPC1L1*,

STARD, and *GRHL* genes, while the studies that used whole blood samples identified significant associations between NAFLD and *SLC7A11* and *LINC00649*. The final remaining study found that NAFLD associates with DNA methylation at the fibrosis-associated *PPARα,* TGFβ1, *Collagen 1A1*, and *PDGF*α genes using bisulphite pyrosequencing of liver biopsy samples (110).

Studies of associations between alcoholic liver disease and DNA methylation have focused on searching for genes where DNA methylation associates with incidence and progression of hepatocellular carcinomas. Chronic alcohol consumption has been shown to associate with global hypomethylation (111) and with methylation of at least ninety-four oncogenes for which aberrant expression has been shown to associate with hepatocellular carcinoma or other liver cancers (112-114). Notable inclusions among those oncogenes are *RASSF1, APC*, and *CDKN2A*, which encode tumor suppressor proteins that are expressed at high levels in the liver, and the *GSTP1, MGMT*, and *CHRNA3* genes, for which hypermethylation is associated with alcohol-attributed hepatocellular carcinoma (112-114). Additionally, in a study to investigate whether methylation of CpG islands can be used to predict hepatocellular carcinomas, Wen *et al.* (2015) identified 41 CpG islands where hypermethylation strongly predicts hepatocellular carcinomas (ROC AUC > 0.9 for tumors > 3 cm, and < 0.8 for tumors ≤ 3 cm) (115). Importantly, however, whether methylation within or near those CpG islands, or near the genes identified in the other studies, plays a causal role in hepatocellular carcinoma remains unclear.

Some epigenetic studies have also been conducted to identify DNA methylation events that associate with CLD in patients for whom the primary driver of disease is chronic viral hepatitis (116). Su *et al.* (2007) led the way by finding that hypermethylation of the promoter that regulates E-cadherin expression associates with hepatocellular carcinomas in patients with chronic hepatitis B infections (117). In a study of patients from the Renji and Shanghai Chest hospitals in Shanghai,

Zhao *et al.* (2014) expanded that knowledge by determining that changes in DNA methylation at the global scale and near transcription-start sites associate with progression of CLD in patients with chronic hepatitis B infections (118). They also identified 10 CpG sites near three genes—*SHISA7*, *ZNF300*, and *SLC22A20*—where DNA methylation associates with progression of CLD due to chronic hepatitis B infection (118). Finally, in another study of 4,841 Taiwanese men with current or previous hepatitis B infections, Kao *et al.* (2017) identified and validated 22 CpG sites located in genes enriched in immune regulation, development, and proteasome degradation where DNA methylation associates with hepatocellular carcinoma (119). No other studies have been conducted that investigated associations between DNA methylation and CLD specifically in the context of chronic hepatitis infections.

In contrast to the many studies that have been conducted to identify associations between DNA methylation and liver fibrosis and CLD as clinical diagnoses, only one study has been reported that sought to identify associations between DNA methylation and specific markers of liver function. In that study of 731 participants in the Rotterdam Study and another 719 non-overlapping individuals, Nano *et al.* (2017) identified no CpG sites that are significantly associated with ALT after Bonferroni correction, and only one CpG site annotated to the *SLC7A1* gene that was significantly associated with AST and a decreased risk of steatohepatitis (120). Thus, essentially no associations between DNA methylation and the specific markers of liver function reviewed above have been identified for any population, let alone people living with HIV. The absence of data on the relationship between DNA methylation and specific markers of liver function constitutes a stark gap in the literature that potentially results in the missing of insights that might benefit efforts to improve prevention, diagnosis, and treatment of CLD.

Clinical Biomarkers of Liver Function and Disease

The liver performs a large and complex array of metabolic functions that can be perturbed independently and in concert by a variety of issues (17-21). Assessing liver function in a clinical or research setting therefore requires incorporating information from many biomarkers that generally reflect overall liver health and/or directly indicate its ability to perform specific metabolic functions (34-37). Several reviews have been written about which biomarkers are routinely used by clinicians to assess liver health and how abnormalities in those biomarkers should be interpreted and synthesized to accurately and completely understand liver condition (34-37). While associations between every one of those clinically relevant markers is worthwhile given that each might uniquely capture a distinct component of liver function, the analyses presented in this thesis are limited to markers for which data from the Veterans Aging Cohort Study were already available: aspartate transaminase (AST), alanine transaminase (ALT), serum albumin, total bilirubin, platelet count, fibrosis-4 (FIB-4) score, and the ast-to-platelet-ratio index (APRI) score. As a result, many other markers that are used in clinical settings to assess liver health and function, and that directly reflect harmful downstream consequences of liver disease and dysfunction, are excluded from this study. Still, the new evidence presented for the biomarkers that are assessed for this thesis might direct researchers toward new hypotheses that expand current understanding of how liver disease progresses and contributes to mortality, both for people living with HIV and broader populations. A summary of the clinical and physiological significances of these markers regarding liver health and function are presented below.

AST & ALT

Aspartate transaminase (AST) and alanine transaminase (ALT) are enzymes expressed primarily in liver cells that play important roles in protein synthesis and gluconeogenesis by catalyzing the interconversion of L-aspartate and L-glutamate, and of L-alanine and L-glutamate, respectively (121). Normally, AST and ALT are not secreted by a healthy liver and only appear at low levels in blood serum (0-35 units/L and 0-45 units/L, respectively) as a result of their release by decomposing liver cells that die as a normal part of cell turnover that is essential to maintaining healthy tissues in the body (122). From a clinical perspective, elevated levels of AST and ALT therefore indicate physical damage to the liver (122), and persistent elevation of AST and ALT is considered a universal warning signal for liver disease of every etiology (122). That universality also presents a limitation, however, as AST and ALT levels cannot be used to distinguish between potential causes of liver damage (e.g., alcohol abuse vs. viral hepatitis). Clinicians therefore always evaluate AST and ALT along with other biomarkers to diagnose specific causes for observed liver disease (34-37).

Serum Albumin

Serum albumin (also called total albumin) is a protein that is synthesized by the liver and released into the blood, where it serves to maintain proper osmotic pressure and to carry important hydrophobic macromolecules that would not otherwise travel well through a hydrophilic environment (123). Because osmotic pressure must be delicately balanced, the liver normally carefully controls the levels of serum albumin in the blood by tightly regulating its synthesis (123). Elevations or reductions in serum albumin are therefore interpreted to represent dysfunction in the liver's biosynthetic capabilities (34-37). Such dysfunction can result from flaws specifically in the

metabolic pathway that generates serum albumin, and thus may not represent a more general disease of the liver. Concurrent elevation of AST and ALT, however, is interpreted to indicate that the synthetic dysfunction is a result of liver damage (34-37).

Total Bilirubin

Bilirubin is a normal byproduct of the normal breakdown of dead red blood cells that is metabolically prepared by the liver for excretion from the body (124). Excessive accumulation of bilirubin in blood serum is therefore interpreted in clinical settings to represent a depression of the liver's metabolic capacity (34-37). Such suppression might result from issues specific to the metabolic pathways responsible for the breakdown of red blood cells. When observed in combination with abnormal elevation of direct indicators of liver damage like AST and ALT, however, elevated total bilirubin is interpreted to result from liver damage, especially if the elevation is prolonged (34-37). Serum bilirubin is therefore included among the tests that are routinely used to assess liver health and function and serves to provide information about the liver's metabolic activity that complements the information on the physical integrity of the liver and its biogenic capacity offered by AST/ALT and serum albumin, respectively.

Platelet Count

Thrombocytopenia, defined as a low platelet count, is strongly associated with liver disease (125,126). While the etiology and physiology of thrombocytopenia is complex, with many potential causes and contributing factors that are unrelated to the liver (126), liver dysfunction can contribute by decreasing production of the hematopoietic growth factor thrombopoietin, which drives platelet cell production in marrow, and by inhibiting other metabolic hepatic processes that

support platelet synthesis (126). Platelet count is therefore routinely included in the panel of clinical tests that are used to assess liver health, and thrombocytopenia concurrent with abnormalities in other liver markers is interpreted to indicate liver disease resulting from liver damage (34-37).

FIB-4 Score and APR Index

The Fibrosis-4 (FIB-4) scoring index and aminotransferase-to-platelet ratio index (APRI) are metrics for assessing the liver health that incorporate information on a patient's age, AST value, ALT value, and platelet count using similar equations (127, 128). Originally developed to detect assess the degree of liver damage incurred by patients with HCV infection, both metrics have proven effective at identifying liver disease of other etiologies as well, including alcohol abuse and non-alcoholic fatty liver disease (34-37). These metrics have therefore become standard orders when liver disease is suspected and provide a general overview of liver health that is supplemented by the specific information offered by other biomarkers to pace disease progression and identify specific causes on a case-by-case basis (34-37).

METHODS

Blood and Phenotype Sampling

Phenotypic and epigenetic data were obtained from the Veterans Aging Cohort Study (VACS), a prospective observational cohort study of health outcomes among veterans living with HIV who receive healthcare from the Veterans Health Administration throughout the United States (129). The VACS was approved by Yale's human research protection program and by the institutional review board of the West Haven campus of the Connecticut veteran healthcare system. All participants provided written consent for the use of their data.

The sample consisted of 960 male veterans living with HIV who were cancer-free at the time of blood collection and for whom phenotypic and epigenetic data are available through VACS. Information on age, race, smoking status, BMI, diabetes status, alcohol use, ever infection with HBV or HCV, and anti-retroviral (ART) use at time of blood draw were obtained by review of medical records or by survey when relevant records were not available. AST, ALT, serum albumin, total bilirubin, platelet count, CD4 count, and viral load were measured for each patient using standard clinical tests. Participants were categorized as virally suppressed if their viral load was <200 copies/mL or as unsuppressed otherwise. A Fibrosis-4 (FIB-4) score and an AST-to-platelet ratio index (APRI) score were calculated for each participant following the standard equations (127, 128).

Associations Among Selected Liver Biomarkers

To identify correlations among the biomarkers of liver function selected for this study, a Pearson correlation coefficient and corresponding *P*-value were calculated for every possible combination of two selected biomarkers using data from all 960 study participants as input for the *rcorr()*

function from the *Hmisc* package in R. To normalize the distributions of strongly right-skewed biomarkers for this and all subsequent analyses, a natural log transformation was performed on values for ALT, AST, FIB-4 score, and APRI score. Distributions of the biomarkers before and after transformation, when needed, are presented in Supplementary Figure 1 in the log2 scale for easier interpretation. Patterns in those correlation coefficients were then visualized via a correlation matrix generated using the *corrplot()* function from the *corrplot* package in R. Statistical significance of obtained correlation coefficients was determined against a Bonferroni-corrected α of 0.05.

DNAm Data Generation, Processing, and Quality Control

Blood samples for clinical tests and methylation analyses were collected simultaneously. Genomic DNA for epigenetic analysis was extracted using FlexiGene DNA extraction kits from (130) and PAXGene collection tubes (131) from QIAGEN. Genome-wide DNA methylation levels for 473 of the participants were assessed using the Infinium HumanMethylation450 (450K) array platform (Illumina), while those for the remaining 487 participants were assessed using the Infinium HumanMethylationEPIC (EPIC) array platform (Illumina) at the Yale Center for Genomic Analysis.

To remove samples and CpG sites with lower data quality, the following quality control steps were performed on the data from both array platforms: 1) probes within 10 nucleotides of a known SNP, or that map to multiple genomic locations, were removed; 2) intensity values with a detection *P*-value of ≥0.001 were set to missing, and any CpG missing >5% of intensity scores across samples were subsequently removed; 3) any sample missing an intensity score for more than 5% of the CpGs covered by the platform on which it was analyzed was removed. As a result

of those steps, 412,583 CpGs remained for analysis in the dataset acquired with the 450K array while 846,604 CpGs remained for analysis in the dataset acquired with the EPIC array. Within both datasets, raw methylation intensity values were then quantile normalized using the *limma* package in R in keeping with the control probe scaling procedure from Illumina. The resulting normalized intensity values were finally used to calculate a DNA methylation score ranging in value from 0 (completely unmethylated) to 1 (completely methylated) for each CpG site following this equation: β = methylated intensity value / (unmethylated intensity value + methylated intensity value + 100).

Because differences in the proportions of the six main leukocyte cell types present in whole blood (CD4+ T, CD8+ T, monocytes, B cells, granulocytes, and natural killer cells) across samples are well-known confounders of associations between DNA methylation in the blood and many phenotypes, the proportions of those six cell types for each participant were determined based on the top 100 cell-type-specific DNA methylation sites in a reference panel of known proportions following the standard algorithm through the *minfi* package in R (132). The estimated cell type proportions for each participant were then controlled for in all EWAS analyses.

Associations Between EWAS Covariates and Selected Liver Biomarkers

The association between each covariate included in the EWAS model—race, smoking status, BMI, diabetes status, hazardous alcohol use, ever HCV infection, ever HBV infection, ART use, CD4 count, viral suppression, and leukocyte cell-type proportions—and each liver marker selected for this study was assessed using a linear model controlling for chronological age. Associations with chronological age were also assessed using a model that did not control for chronological age. Regardless of *P*-value, all covariates listed above were included in the final EWAS model to

account for the reasonable possibility that they might confound the relationship between DNA methylation and the liver markers.

Associations of DNAm Age Acceleration with Selected Liver Biomarkers

DNA methylation age acceleration for each participant was measured using the IEAA, EAA, PhenoAA, and GrimAA metrics as specified by original reporting articles (94, 96-98). The association between DNA methylation age acceleration, as measured by each of those four metrics, and each of the selected biomarkers of liver function was then assessed by linear regression while controlling for all covariates included in the final EWAS model.

Principal Component Analysis

To adjust for population structures that might confound observed associations between DNA methylation events and the selected biomarkers of liver function among participants profiled with the 450K array or those profiled with the EPIC array, separate principal component analyses were performed for both groups following the analytical approach developed by Barfield *et al.* (2014) (133). To begin, a pruned version of the dataset from each group was created that contained only CpG sites within 50 base pairs of known SNPs to approximate the genome-wide genetic variation traditionally used in principal component analyses. Then, principal components were calculated for both groups based on methylation beta values of the pruned set of CpG sites using the *prcomp* function from the *Factoextra* package in R. The top ten resulting principal components of the 450K group were then included in all subsequent epigenetic association models applied to data from the 450K group, while the top ten principal components of the EPIC group were included in all epigenetic association models applied to data from the EPIC group.

Statistical Methods for DNAm EWAS for Selected Liver Biomarkers

To investigate the association between each selected biomarker of liver function and methylation at individual CpG sites across the autosomal chromosomes, data from the 450K group was analyzed in parallel with data from the EPIC group using the same model, and a meta-analysis was conducted for CpG sites covered by both platforms as described later. To control for systematic variation in methylation intensity values that might have occurred across groups of co-processed samples because of batch effects caused by deviations in sample processing and array performance, a linear mixed effect model controlling for random batch effects across arrays was employed. That final model was adjusted for age, race, current smoking, BMI, diabetes, hazardous alcohol use, ever infection with HCV, ever infection with HBV, ART use, CD4 count, viral load suppression, leukocyte cell-type proportions, and the top ten principal components for the group being analyzed. When needed for a given combination of CpG site and biomarker, participants missing values needed by the model were excluded from the analysis for that combination. CpG sites for which the applied model did not converge were also excluded. Furthermore, because the number of covariates in the model often led to destabilizing over-sparsity when the sample size was <100 people, CpG sites for which the sample size used by the model was <100 people due to missing values were excluded from the results and further analysis.

A variance-weighted meta-analysis was performed for 385,062 CpG sites that are covered by both the 450K and EPIC arrays. For CpG sites not covered by both platforms, results were obtained only from the cohort profiled by the covering platform and no meta-analysis was conducted. Separate false-discovery rate adjustments ($Q < 0.05$) were conducted for each liver biomarker, where one adjustment was conducted on *P-values* from the meta-analysis, and the second was conducted on *P*-values that could not be included in the meta-analysis.

RESULTS

Participant Characteristics

After data processing and quality control, the analysis dataset for those profiled using the 450K platform included observations from 473 individuals and the dataset for those profiled using the EPCI platform included observations from 487 individuals. Relevant characteristics of the two groups are summarized in Table 1. All participants were male veterans living with HIV who were never diagnosed with cancer and had an average age of 51.2 ± 7.5 years. The average difference in chronological age between the 450K and EPIC cohorts was significant but mall (Δ = 1.1 years, *P*-value = 0.031). The two cohorts did not differ significantly (*P*-value > 0.05) in race (Black vs. non-Black) or in prevalence of current smoking, diabetes, hazardous alcohol use, ever infection with HBV, ever infection with HCV, or viral load suppression. They also did not differ significantly in their average BMI or CD4 count. Average biomarker values from the 450K cohort differed significantly from those from the EPIC cohort for ALT (Δ = 4.7 units/L, *P*-value = 1.93×10^{-3}), AST (Δ = 6.4 units/L, *P*-value = 7.34×10^{-5}), serum albumin (Δ = -0.11 mg/dL, *P*-value = 2.89×10^{-4}), FIB-4 score (Δ = 0.24, *P*-value = 8.49×10^{-4}), and APRI score (Δ = 0.112, *P*-value = 8.35×10^{-4}). Total bilirubin and platelet count did not differ significantly between the two sub-cohorts.

Correlations Among Selected Liver Biomarkers

Observed values for ALT, AST, FIB-4 score, and APRI score all correlated positively and strongly with each other ($\rho \geq 0.6$, FDR $Q < 0.05$), except for ALT and FIB-4 score, which also correlated positively but less strongly (ρ = 0.36, FDR $Q < 0.05$) (Figure 1). Platelet count correlates negatively with AST, ALT, FIB-4 score, and APRI score to varying degrees ($Q < 0.05$) and weakly

with albumin ($\rho = 0.13$, FDR $Q < 0.05$), but not with total bilirubin ($\rho = -0.06$, $Q > 0.05$). Serum albumin correlated negatively with AST, APRI score, and FIB-4 score ($-0.25 \leq \rho \leq -0.18$, $Q < 0.05$) but not with ALT ($\rho = -0.01$, $Q > 0.05$). Total bilirubin did not correlate significantly with serum albumin, platelet count, FIB-4 score, or APRI score ($Q > 0.05$) and correlated only weakly with AST and ALT ($\rho < 0.14$, FDR $Q < 0.05$).

Covariate Associations with Selected Liver Biomarkers

Associations between covariates included in the final EWAS model and the selected liver biomarkers are presented in Table 2. Only ever infection with HCV was associated with all seven selected liver health markers when controlling for age, while diabetes status was not associated with any of the selected markers. All liver markers associated with at least four of the covariates included in the model excluding cell type proportions. Regardless of the *P*-value for the observed association between each covariate and each liver biomarker, all covariates were included in the final epigenetic analysis model to account for the reasonable possibility that each might confound associations between DNA methylation events and liver phenotypes.

Associations of DNAm Age Acceleration with Selected Liver Biomarkers

DNA methylation age acceleration as measured by IEAA, EEAA, PhenoAA, and GrimAA was not significantly associated with any of the selected biomarkers of liver function, except in the case of serum albumin, which associated significantly with PhenoAA ($\beta = -0.014$, *P*-value = 1.6×10^{-13}) in an unadjusted model (Figure 2). That association remained significant after FDR adjustment to $Q < 0.05$ in a model that adjusted for all covariates included in the final EWAS model ($\beta = -0.007$, *P*-value = $8.6\text{x}10^{-4}$) (Table 3), while no other metric of DNA methylation age

acceleration was associated with any other selected biomarker of liver function in the adjusted model (Table 3).

DNAm EWAS for Selected Liver Biomarkers

Meta-analysis of the separate EWAS results obtained from the EPIC and 450K cohorts for each selected liver biomarker and each CpG site included in both platforms reveals that increased DNA methylation at nine CpG sites associates significantly with increased serum albumin among male veterans living with HIV after adjusting for all covariates in the final EWAS model and an FDR correction against $Q < 0.05$ (Table 4, Figure 3). A regional plot for the section of chromosome 17 that contains five of those CpG sites is presented in Supplementary Figure 2. Quantile-quantile analysis of the expected and observed *P*-values from the meta-analysis for serum albumin reveals no global inflation of the unadjusted *P*-values (Figure 4), so no further corrections were applied. Notably, among the one hundred CpG sites with the lowest unadjusted *P*-values for association between DNA methylation and serum albumin after meta-analysis, the beta coefficients obtained from the separate EWAS analyses of the EPIC and 450K cohorts for those CpG sites show a strong, positive correlation ($\rho = 0.89$, *P*-value < 0.001) (Figure 5).

Four of the nine CpG sites where methylation associates positively with serum albumin—cg16936953, cg18942579, cg01409343, and cg12054453—are annotated to be located within a codon of the *TMEM49* gene (Illumina). 10% increases in methylation of those sites correspond to average increases in serum albumin of 0.12 g/dL (β 95% CI: 0.08, 0.16), 0.14 g/dL (β 95% CI: 0.09, 0.19), 0.17 g/dL (β 95% CI: 0.11, 0.23), and 0.10 g/dL (β 95% CI: 0.06, 0.14), respectively. Four of the remaining five CpG sites—cg18181703, cg03546163, cg20995564, and cg23966214—are annotated to correspond to the *SOCS3*, *FKBP5*, *ZEB2*, and *SAMD14* genes,

respectively (Illumina). 10% increases in methylation at those CpG sites correspond to average increases in serum albumin of 0.19 g/dL (β 95% CI: 0.12, 0.26), 0.11 g/dL (β 95% CI: 0.07, 0.15), 0.12 g/dL (β 95% CI: 0.07, 0.17), and 0.35 g/dL (β 95% CI: 0.21, 0.49). The final remaining CpG site, cg12992827, is not annotated to correspond to any known coding sequence. A 10% increase in methylation at that site corresponds to an average increase in serum albumin of 0.19 g/dL (β 95% CI: 0.11, 0.26). No other meta-analyzed CpG sites associate significantly with a selected liver biomarker after FDR correction to $Q < 0.05$. A quantile-quantile plot and Manhattan plot of the FDR-unadjusted *P*-values obtained from meta-analysis of the separate EWAS results from the EPIC and 450K cohorts are presented in Supplementary Figures 3-8 for each selected liver biomarker except serum albumin. Additionally, statistics and annotations for the ten CpG sites with the smallest FDR-unadjusted *P*-values after meta-analysis are presented in Supplementary Table 1 for each selected liver marker.

Among the CpG sites that are included in only one of the two methylation assay platforms used for this study, and that therefore cannot be included in the meta-analysis, DNA methylation was not significantly associated with any of the selected liver biomarkers. Statistics and annotations for the ten CpG sites with the lowest unadjusted *P*-values among those that could not be meta-analyzed are presented for each liver biomarker in Supplementary Table 2.

DISCUSSION

To expand on the contributions of studies that have identified associations between DNA methylation and clinical liver disease as part of a larger effort to better understand the etiology of liver function and disease among people living with HIV, the EWAS presented in this thesis provide the first data to evaluate whether specific DNA methylation events associate with seven commonly-used biomarkers of liver function—aspartate transaminase, alanine transaminase, serum albumin, total bilirubin, platelet count, FIB-4 score, and APRI score—among male US veterans living with HIV based on data from VACS.

The liver biomarkers selected for this study show intercorrelations that justify performing a separate EWAS for each biomarker. The positive correlations among AST, ALT, FIB-4 score, and APRI score are expected because AST and ALT are both released into the blood upon physical damage to liver cells and thus should increase together, and because AST is included in the numerator of the equations used to calculate FIB-4 and APRI scores. The negative correlations between platelet count and FIB-4 and APRI scores are also expected because platelet count is included in the denominator of both equations. While strong intercorrelations might suggest that conducting a separate EWAS for each marker is unnecessary, the correlations observed are imperfect enough to indicate that separate EWAS are warranted. Following that same logic, the absence of even moderate correlations with serum albumin and total serum bilirubin suggests that those markers might have mechanisms of regulation that are independent of those for the other selected liver biomarkers, thus warranting the inclusion of separate EWAS to probe for genes involved in those separate mechanisms that EWAS for the other uncorrelated liver markers might not detect.

Despite the slight association observed between PhenoAA and serum albumin, the absence of any other significant associations for the four tested metrics of accelerated DNA methylation aging—PhenoAA, GrimAA, IEAA, and EEAA—suggests that differential biological aging does not affect the selected biomarkers of liver function. While these findings largely do not support that global changes in DNA methylation are linked to changes in liver function, they leave open the possibility that specific sites of DNA methylation might regulate the expression of genes in a manner that influences behavior of the selected biomarkers of liver health, either by influencing the progression of conditions that generally damage the liver and inhibit function—like fibrosis or hepatocellular carcinoma—or by affecting pathways that are specifically involved in regulation of a particular liver biomarkers.

Nine CpG sites were identified where DNA methylation is associated with changes in serum albumin. Four of those sites are annotated to be located near the *TMEM49* gene (also known as *VMP1*), which encodes a transmembrane protein that drives cell autophagy (134) and is overexpressed in pancreatitis-affected acinar cells (135). Cell autophagy, therefore, might be linked to regulation of serum albumin. Indeed, studies have demonstrated that serum albumin suppresses cell autophagy via mTOR activation and that depriving cultured cells of serum albumin induces autophagy that is thought to protect against the accumulation of harmful reactive oxygen species (ROS) that are released by damaged mitochondria in response to low serum albumin (136, 137). Given those findings, the positive association observed between serum albumin and DNA methylation near *TMEM49* might reflect that DNA methylation is involved in a mechanism by which serum albumin suppresses autophagy through repression of *TMEM49* expression. Further studies are needed to investigate that possible relationship.

One site where DNA methylation associates positively with serum albumin levels is located near the *SOCS3* gene, which encodes a protein that inhibits cytokine signaling along the STAT pathway in response to elevated cytokine levels (138, 139). Studies have established that repression of *SOCS3* enables liver fibrosis by allowing for increased fibrogenic signaling via STAT-3-mediated upregulation of TGF-β (140-142). The positive association between increased methylation near *SOCS3* and increased serum albumin observed here might therefore reflect that DNA methylation is linked to that relationship. Additionally, obesity has been shown to associate with downregulation of *SOCS3*, which draws attention to the possibility that an association between methylation near *SOCS3* and liver dysfunction might result from the well-observed association between obesity and liver dysfunction. Indeed, cg18181703 has been shown to be associated with obesity in other EWAS (143). As such, further exploration of how *SOCS3* might contribute to the etiology of CLD, and of the role that DNA methylation might play, is warranted.

Another site where DNA methylation associates positively with serum albumin levels in the EWAS presented here is located near the *FKBP5* gene, which encodes an immunoregulatory protein that contributes to basic protein folding and trafficking (144). *FKBP5* has been shown to positively regulate stress response and to drive acquisition of metabolic disorders including obesity, insulin resistance, and diabetes (145). Moreover, *FKBP5* is thought to contribute to liver dysfunction (146). Indeed, deletion of *FKBP5* protected knock-out mice from fatty liver disease despite high-fat diets (147). Thus, the positive association of DNA methylation near *FKBP5* with serum albumin levels might reflect that *FKBP5* expression should be investigated as the underlying mechanism of the association.

Yet another site where DNA methylation associates positively with serum albumin levels in the EWAS presented here is located near the *ZEB2* gene, which encodes a zinc-finger DNA-

binding protein known to be expressed in hepatocytes (148). Studies have demonstrated that repression of *ZEB2* expression by microRNAs induces apoptosis in hepatocytes (149), suggesting that *ZEB2* might influence preservation of liver integrity in the face of hepatic cell damage. Further investigation of the observed association between DNA methylation near *ZEB2* and serum albumin might therefore focus on the potential effects that DNA methylation might have on *ZEB2* expression and the downstream consequences of those effects for hepatic damage.

The remaining CpG site where DNA methylation associates positively with serum albumin levels is annotated to be located near genes encoding the *SMAD14* gene. The function of the *SMAD14* gene is largely unknown, but it has been shown to associate with gastric cancer and has therefore been hypothesized to act as a tumor suppressor (150). No studies have been conducted that directly link *SMAD14* or its tumor suppressor properties to either general liver health and function or specifically to serum albumin regulation. As such, further studies are needed.

Strengths and Weaknesses

A strength of this study is that it combined data from two different microarray platforms for assessing genome-wide DNA methylation levels that only partially overlap in their coverage of CpG sites, thereby enabling analysis of DNA methylation at a wider range of CpG sites than either microarray platform alone would have allowed. A corresponding weakness, however, is that meta-analysis could only be performed for the subset of CpG sites that are included in both platforms. Still, results from the two cohorts at CpG sites that could be meta-analyzed were highly consistent among top hits for serum albumin, suggesting that combining data from the two platforms successfully augments power to detect significant associations.

Another strength is that this study simultaneously examined multiple biomarkers that, combined, generally reflect liver function but also individually directly indicate the liver's capacity to perform specific metabolic functions that are not perfectly correlated and thus might be regulated by independent mechanisms or affected by liver disease via distinct pathways. As a result, this study enables probing for associations with DNA methylation across a wider range of genes that might only influence specific aspects of liver function and therefore might be missed if a single microarray platform were used or liver disease as a general diagnosis were assessed.

Another strength of this study is that, unlike with any EWAS study of liver function or disease conducted thus far, the cohort used consists exclusively of people living with HIV. As a result, this study can still contribute to the identification of genes that are involved in liver health and function for all people, but also contribute to the identification of genes that affect liver function specifically in the context of HIV infection—something that previous studies that did not exclusively include people with HIV may not have been powered to accomplish. Restriction of the study cohort to only veterans living with HIV who were recruited to the VACS study also constitutes a potential weakness, however, since veterans tend to face certain stressors and medical conditions more frequently than the general population. Moreover, the veterans recruited for this study represent an especially unhealthy section of veterans living with HIV. These factors jeopardize the degree to which the cohorts included in this study represent broader populations of people living with HIV in the US, especially younger members of those populations who never served in the military. As a result, the associations identified in this study might not generalize well to those external populations, and associations might exist in those external populations that were not detected in this cohort.

Another weakness is that this study used DNA methylation levels measured in blood cells to assess associations between DNA methylation and biomarkers of liver function, which is potentially problematic because DNA methylation patterns in blood cells may not represent DNA methylation patterns in the liver cells that are responsible for regulating the selected biomarkers. Notably, however, blood DNA methylation levels have been shown to associate with many health outcomes that extend beyond the blood, including obesity, diabetes, and certain cancers (151-154). As such, using blood DNA methylations in the studies presented here is a reasonable first step toward understanding the relationship between DNA methylation and liver function, especially given that EWAS studies are meant to generate hypotheses and identify genes that can be included in candidate gene studies to confirm their observed association with liver function and further elucidate the role of DNA methylation in modulating that relationship.

Finally, this study failed to identify significant associations between DNA methylation events at specific CpG sites and six of the seven outcomes assessed. The failure to detect significant associations between DNA methylation events and aspartate transaminase levels, alanine transaminase levels, total serum bilirubin, platelet count, FIB-4 score, and APRI score might result from an absence of true associations between DNA methylation events and those outcomes. The failure might also result, however, from insufficient power to detect true associations with the data collected. Increasing the sample size and selecting from a generally healthier population of HIV+ veterans—who are more likely to have less variation in their liver health and, thus, their markers of liver health and function—would improve the power of this study. Nevertheless, this study generates several hypotheses regarding the roles of several genes in regulating either general liver function or specific metabolic functions of the liver that have not yet been directly investigated.

Future Work

This thesis demonstrates that EWAS can reveal DNA methylation events that associate with biomarkers of liver function, suggesting that EWAS might be employed as a tool to identify candidate genes for more rigorous causal studies to uncover biological mechanisms that affect liver function by either regulating overall liver health or influencing pathways only involved in a specific liver function reflected by a biomarker of interest. Such application of EWAS could be beneficial if it broadens and deepens understanding the complex activity of the liver. Critically, future work seeking to identify genes that influence liver function in people living with HIV should include participants who do not have HIV to allow for evaluation of interaction between HIV and DNA methylation at identified CpG sites. Evaluation of such interaction would be useful because it would contribute to discrimination between CpG sites where DNA methylation influences liver function for all people, and CpG sites where methylation only influences liver function among people living with HIV. Genes that only influence liver function among those living with HIV would be especially interesting targets for those deploying limited resources to discover drug targets to improve the health of people living with HIV.

This study offers little to substantiate the possibility that methylation at CpG sites could be measured to predict risk of CLD, either among people living with HIV or other populations, especially because none of the four metrics of accelerated DNA methylation were significantly associated with changes in the seven tested biomarkers of liver function, except with PhenoAA and serum albumin. Future work on this topic should remain responsive to identification of patterns in DNA methylation that might serve as early warning indicators of CLD, especially as more DNA methylation clocks are validated and proposed. Indeed, work could be done to specifically generate a set of CpG sites that can be used to predict or estimate liver function. Moreover, future

researchers should look out for patterns of DNA methylation that can be used to distinguish between subtypes of CLD that are defined by which specific metabolic functions are impeded. Defining such collections of CpG sites might enable clinicians to prioritize treatments that seek to correct specific dysfunctions in addition to correcting general health, thus giving the patient optimal outcome chances in accordance with the vision of precision medicine that is so urgently sought.

REFERENCES

1. Goodman ZD. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. Journal of hepatology. 2007 Oct 1;47(4):598-607.

2. Clark JM, Diehl AM. Defining nonalcoholic fatty liver disease: implications for epidemiologic studies. Gastroenterology. 2003 Jan 1;124(1):248-50.

3. Propst A, Propst T, Zangerl G, Öfner D, Judmaier G, Vogel W. Prognosis and life expectancy in chronic liver disease. Digestive diseases and sciences. 1995 Aug;40(8):1805-15.

4. Vong S, Bell BP. Chronic liver disease mortality in the United States, 1990–1998. Hepatology. 2004 Feb;39(2):476-83.

5. Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi Z, Sebastiani G, Ekstedt M, Hagstrom H, Nasr P, Stal P. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. Hepatology. 2017 May;65(5):1557-65.

6. Paik JM, Henry L, De Avila L, Younossi E, Racila A, Younossi ZM. Mortality related to nonalcoholic fatty liver disease is increasing in the United States. Hepatology communications. 2019 Nov;3(11):1459-71.

7. Mantovani A, Scorletti E, Mosca A, Alisi A, Byrne CD, Targher G. Complications, morbidity and mortality of nonalcoholic fatty liver disease. Metabolism. 2020 Oct 1;111:154170.

8. Riley T, Bhatti AM. Preventive strategies in chronic liver disease: part I. Alcohol, vaccines, toxic medications and supplements, diet and exercise. American family physician. 2001 Nov 1;64(9):1555.

9. Riley T, Bhatti AM. Preventive strategies in chronic liver disease: part II. Cirrhosis. American family physician. 2001 Nov 15;64(10):1735.

10. Wiegand J, Berg T. The etiology, diagnosis and prevention of liver cirrhosis: part 1 of a series on liver cirrhosis. Deutsches Ärzteblatt International. 2013 Feb;110(6):85.

11. Terpstra OT, Schalm SW, Weimar W, Willemse PJ, Baumgartner D, Groenland TH, ten Kate FW, Porte RJ, Rave SD, Reuvers CB, Stibbe J. Auxiliary partial liver transplantation for end-stage chronic liver disease. New England Journal of Medicine. 1988 Dec 8;319(23):1507-11.

12. Roth K, Lynn J, Zhong Z, Borum M, Dawson NV. Dying with end stage liver disease with cirrhosis: insights from SUPPORT. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatment. Journal of the American Geriatrics Society. 2000 May 1;48(S1):S122-30.

13. Silva AS, Santos LL, Passos AD, Sankarankutty AK, Martinelli AD, Castro e Silva OD. Chronic liver disease prevention strategies and liver transplantation. Acta Cirúrgica Brasileira. 2006;21:79-84.

14. Julien J, Ayer T, Bethea ED, Tapper EB, Chhatwal J. Projected prevalence and mortality associated with alcohol-related liver disease in the USA, 2019–40: a modelling study. The Lancet Public Health. 2020 Jun 1;5(6):e316-23.

15. Adams LA, Roberts SK, Strasser SI, Mahady SE, Powell E, Estes C, Razavi H, George J. Nonalcoholic fatty liver disease burden: Australia, 2019–2030. Journal of gastroenterology and hepatology. 2020 Sep;35(9):1628-35.

16. Estes C, Chan HL, Chien RN, Chuang WL, Fung J, Goh GB, Hu TH, Huang JF, Jang BK, Jun DW, Kao JH. Modelling NAFLD disease burden in four Asian regions—2019-2030. Alimentary pharmacology & therapeutics. 2020 Apr;51(8):801-11.

17. Gebhardt R. Metabolic zonation of the liver: regulation and implications for liver function. Pharmacology & therapeutics. 1992 Jan 1;53(3):275-354.

18. Campbell I. Liver: metabolic functions. Anaesthesia & Intensive Care Medicine. 2006 Feb 1;7(2):51-4.

19. Anty R, Lemoine M. Liver fibrogenesis and metabolic factors. Clinics and research in hepatology and gastroenterology. 2011 Jun 1;35:S10-20.

20. Mitra V, Metcalf J. Metabolic functions of the liver. Anaesthesia & Intensive Care Medicine. 2012 Feb 1;13(2):54-5.

21. Eriksen PL, Sørensen M, Grønbæk H, Hamilton-Dutoit S, Vilstrup H, Thomsen KL. Non-alcoholic fatty liver disease causes dissociated changes in metabolic liver functions. Clinics and research in hepatology and gastroenterology. 2019 Oct 1;43(5):551-60.

22. Bathum L, Petersen HC, Rosholm JU, Hyltoft Petersen P, Vaupel J, Christensen K. Evidence for a substantial genetic influence on biochemical liver function tests: results from a population-based Danish twin study. Clinical chemistry. 2001 Jan 1;47(1):81-7.

23. Rahmioglu N, Andrew T, Cherkas L, Surdulescu G, Swaminathan R, Spector T, Ahmadi KR. Epidemiology and genetic epidemiology of the liver function test proteins. PloS one. 2009 Feb 11;4(2):e4435.

24. Foley DL, Craig JM, Morley R, Olsson CJ, Dwyer T, Smith K, Saffery R. Prospects for epigenetic epidemiology. American journal of epidemiology. 2009 Feb 15;169(4):389-400.

25. Relton CL, Smith GD. Epigenetic epidemiology of common complex disease: prospects for prediction, prevention, and treatment. PLoS Med. 2010 Oct 26;7(10):e1000356.

26. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. Nature Reviews Genetics. 2013 Aug;14(8):585-94.

27. Singal R, Ginder GD. DNA methylation. Blood, The Journal of the American Society of Hematology. 1999 Jun 15;93(12):4059-70.

28. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013 Jan;38(1):23-38.

29. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nature Reviews Genetics. 2011 Aug;12(8):529-41.

30. Michels KB, Binder AM, Dedeurwaerder S, Epstein CB, Greally JM, Gut I, Houseman EA, Izzi B, Kelsey KT, Meissner A, Milosavljevic A. Recommendations for the design and analysis of epigenome-wide association studies. Nature methods. 2013 Oct;10(10):949.

31. Birney E, Smith GD, Greally JM. Epigenome-wide association studies and the interpretation of disease-omics. PLoS genetics. 2016 Jun 23;12(6):e1006

32. Doweiko JP, Nompleggi DJ. The role of albumin in human physiology and pathophysiology, Part III: Albumin and disease states. JPEN. Journal of parenteral and enteral nutrition. 1991 Jul 1;15(4):476-83.

33. Lin JP, Vitek L, Schwertner HA. Serum bilirubin and genes controlling bilirubin concentrations as biomarkers for cardiovascular disease. Clinical chemistry. 2010 Oct 1;56(10):1535-43.

34. Woreta TA, Alqahtani SA. Evaluation of abnormal liver tests. Medical Clinics. 2014 Jan 1;98(1):1-6.

35. Agrawal S, Dhiman RK, Limdi JK. Evaluation of abnormal liver function tests. Postgraduate medical journal. 2016 Apr 1;92(1086):223-34.

36. Kwo PY, Cohen SM, Lim JK. ACG clinical guideline: evaluation of abnormal liver chemistries. Official journal of the American College of Gastroenterology| ACG. 2017 Jan 1;112(1):18-35.

37. Schreiner AD, Rockey DC. Evaluation of abnormal liver tests in the adult asymptomatic patient. Current opinion in gastroenterology. 2018 Jul 1;34(4):272-9.

38. Joshi D, O'Grady J, Dieterich D, Gazzard B, Agarwal K. Increasing burden of liver disease in patients with HIV infection. The Lancet. 2011 Apr 2;377(9772):1198-209.

39. Prins M, Aguado IH, Brettle RP, Robertson JR, Broers B, Carré N, Goldberg DJ, Zangerle R, Coutinho RA, van den Hoek A. Pre-AIDS mortality from natural causes associated with HIV disease progression: evidence from the European Seroconverter Study among injecting drug users. Aids. 1997 Nov 15;11(14):1747-56.

40. Collins S, Mertenskoetter T, Loeliger E, Tressler R, Weller I, Friis-Møller N, Worm SW, Sabin CA, Sjøl A, Lundgren JD, Sawitz A. Liver-related deaths in persons infected with the human immunodeficiency virus: the D: A: D study. Archives of internal medicine. 2006 Aug 28;166(15):1632-41.

41. Frank TD, Carter A, Jahagirdar D, Biehl MH, Douwes-Schultz D, Larson SL, Arora M, Dwyer-Lindgren L, Steuben KM, Abbastabar H, Abu-Raddad LJ. Global, regional, and national incidence, prevalence, and mortality of HIV, 1980–2017, and forecasts to 2030, for 195 countries and territories: a systematic analysis for the Global Burden of Diseases, Injuries, and Risk Factors Study 2017. The lancet HIV. 2019 Dec 1;6(12):e831-59.

42. Sayyah M, Rahim F, Kayedani GA, Shirbandi K, Saki-Malehi A. Global View of HIV Prevalence in Prisons: A Systematic Review and Meta-Analysis. Iranian journal of public health. 2019 Feb;48(2):217.

43. Asrani SK, Devarbhavi H, Eaton J, Kamath PS. Burden of liver diseases in the world. Journal of hepatology. 2019 Jan 1;70(1):151-71.

44. Xiao J, Wang F, Wong NK, He J, Zhang R, Sun R, Xu Y, Liu Y, Li W, Koike K, He W. Global liver disease burdens and research trends: analysis from a Chinese perspective. Journal of hepatology. 2019 Jul 1;71(1):212-21.

45. Ye Q, Zou B, Yeo YH, Li J, Huang DQ, Wu Y, Yang H, Liu C, Kam LY, Tan XX, Chien N. Global prevalence, incidence, and outcomes of non-obese or lean non-alcoholic fatty liver disease: a systematic review and meta-analysis. The Lancet Gastroenterology & Hepatology. 2020 Aug 1;5(8):739-52.

46. Ye Q, Zou B, Yeo YH, Li J, Huang DQ, Wu Y, Yang H, Liu C, Kam LY, Tan XX, Chien N. Global prevalence, incidence, and outcomes of non-obese or lean non-alcoholic fatty liver disease: a systematic review and meta-analysis. The Lancet Gastroenterology & Hepatology. 2020 Aug 1;5(8):739-52.

47. Desai AP, Mohan P, Nokes B, Sheth D, Knapp S, Boustani M, Chalasani N, Fallon MB, Calhoun EA. Increasing economic burden in hospitalized patients with cirrhosis: analysis of a national database. Clinical and translational gastroenterology. 2019 Jul;10(7).

48. Cook M, Zavala E. The finances of a liver transplant program. Current opinion in organ transplantation. 2019 Apr 1;24(2):156-60.

49. Ong KJ, van Hoek AJ, Harris RJ, Figueroa J, Waters L, Chau C, Croxford S, Kirwan P, Brown A, Postma MJ, Gill ON. HIV care cost in England: a cross-sectional analysis of antiretroviral treatment and the impact of generic introduction. HIV medicine. 2019 Jul;20(6):377-91.

50. Brunt EM. Pathology of fatty liver disease. Modern Pathology. 2007 Feb;20(1):S40-8.

51. Brunt EM. Pathology of nonalcoholic fatty liver disease. Nature reviews Gastroenterology & hepatology. 2010 Apr;7(4):195-203.

52. Yeh MM, Brunt EM. Pathological features of fatty liver disease. Gastroenterology. 2014 Oct 1;147(4):754-64.

53. Morrison ED, Kowdley KV. Genetic liver disease in adults: early recognition of the three most common causes. Postgraduate medicine. 2000 Jan 1;107(2):147-59.

54. Ala A, Walker AP, Ashkan K, Dooley JS, Schilsky ML. Wilson's disease. The Lancet. 2007 Feb 3;369(9559):397-408.

55. Afdhal NH, Nunes D. Evaluation of liver fibrosis: a concise review. American Journal of Gastroenterology. 2004 Jun 1;99(6):1160-74.

56. Bataller R, Brenner DA. Liver fibrosis. The Journal of clinical investigation. 2005 Feb 1;115(2):209-18.

57. Hernandez-Gea V, Friedman SL. Pathogenesis of liver fibrosis. Annual review of pathology: mechanisms of disease. 2011 Feb 28;6:425-56.

58. Ebrahimi H, Naderian M, Sohrabpour AA. New concepts on pathogenesis and diagnosis of liver fibrosis; a review article. Middle East journal of digestive diseases. 2016 Jul;8(3):166.

59. Schuppan D, Afdhal NH. Liver cirrhosis. The Lancet. 2008 Mar 8;371(9615):838-51.

60. Pinzani M, Rosselli M, Zuckermann M. Liver cirrhosis. Best practice & research Clinical gastroenterology. 2011 Apr 1;25(2):281-90.

61. Tsochatzis EA, Bosch J, Burroughs AK. Liver cirrhosis. The Lancet. 2014 May 17;383(9930):1749-61.

62. Yang JD, Roberts LR. Hepatocellular carcinoma: a global view. Nature reviews Gastroenterology & hepatology. 2010 Aug;7(8):448.

63. Bruix J, Gores GJ, Mazzaferro V. Hepatocellular carcinoma: clinical frontiers and perspectives. Gut. 2014 May 1;63(5):844-55.

64. Balogh J, Victor III D, Asham EH, Burroughs SG, Boktour M, Saharia A, Li X, Ghobrial RM, Monsour Jr HP. Hepatocellular carcinoma: a review. Journal of hepatocellular carcinoma. 2016;3:41.

65. Bernal W, Auzinger G, Dhawan A, Wendon J. Acute liver failure. The Lancet. 2010 Jul 17;376(9736):190-201.

66. Bernal W, Wendon J. Acute liver failure. New England Journal of Medicine. 2013 Dec 26;369(26):2525-34.

67. Losser MR, Payen D. Mechanisms of liver damage. InSeminars in liver disease 1996 (Vol. 16, No. 04, pp. 357-367). 1996 by Thieme Medical Publishers, Inc.

68. Dey A, Cederbaum AI. Alcohol and oxidative liver injury. Hepatology. 2006 Feb;43(S1):S63-74.

69. McKillop IH, Schrum LW. Alcohol and liver cancer. Alcohol. 2005 Apr 1;35(3):195-203.

70. Grewal P, Viswanathen VA. Liver cancer and alcohol. Clinics in liver disease. 2012 Nov 1;16(4):839-50.

71. Scheuer PJ, Ashrafzadeh P, Sherlock S, Brown D, Dusheiko GM. The pathology of hepatitis C. Hepatology. 1992 Apr;15(4):567-71.

72. Dhillon AP, Dusheiko GM. Pathology of hepatitis C virus infection. Histopathology. 1995 Apr;26(4):297-309.

73. Geller SA. Hepatitis B and hepatitis C. Clinics in liver disease. 2002 May 1;6(2):317-34.

74. Fiel MI. Pathology of chronic hepatitis B and chronic hepatitis C. Clinics in liver disease. 2010 Nov 1;14(4):555-75.

75. Michelotti GA, Machado MV, Diehl AM. NAFLD, NASH and liver cancer. Nature reviews Gastroenterology & hepatology. 2013 Nov;10(11):656-65.

76. Byrne CD, Targher G. NAFLD: a multisystem disease. Journal of hepatology. 2015 Apr 1;62(1):S47-64.

77. Kaspar MB, Sterling RK. Mechanisms of liver disease in patients infected with HIV. BMJ open gastroenterology. 2017 Oct 1;4(1):e000166.

78. Koziel MJ, Peters MG. Viral hepatitis in HIV infection. New England Journal of Medicine. 2007 Apr 5;356(14):1445-54.

79. Soriano V, Vispo E, Labarga P, Medrano J, Barreiro P. Viral hepatitis and HIV co-infection. Antiviral research. 2010 Jan 1;85(1):303-15.

80. Sulkowski MS, Mehta SH, Torbenson MS, Higgins Y, Brinkley SC, De Oca RM, Moore RD, Afdhal NH, Thomas DL. Rapid fibrosis progression among HIV/hepatitis C virus-co-infected adults. Aids. 2007 Oct 1;21(16):2209-16.

81. Soriano V, Barreiro P, Sherman KE. The changing epidemiology of liver disease in HIV patients. AIDS reviews. 2013 Jan 1;15(1):25-31.

82. Salmon-Ceron D, Lewden C, Morlat P, Bévilacqua S, Jougla E, Bonnet F, Héripret L, Costagliola D, May T, Chêne G. Liver disease as a major cause of death among HIV infected patients: role of hepatitis C and B viruses and alcohol. Journal of hepatology. 2005 Jun 1;42(6):799-805.

83. Soriano V, Puoti M, Garcia-Gasco P, Rockstroh JK, Benhamou Y, Barreiro P, McGovern B. Antiretroviral drugs and liver injury. Aids. 2008 Jan 2;22(1):1-3.

84. Kontorinis N, Dieterich D. Hepatotoxicity of antiretroviral therapy. AIDS reviews. 2003 Jan 1;5(1):36-43.

85. Newell-Price J, Clark AJ, King P. DNA methylation and silencing of gene expression. Trends in Endocrinology & Metabolism. 2000 May 1;11(4):142-8.

86. Mierzejewska K, Bochtler M, Czapinska H. On the role of steric clashes in methylation control of restriction endonuclease activity. Nucleic acids research. 2016 Jan 8;44(1):485-95.

87. Robertson KD, Wolffe AP. DNA methylation in health and disease. Nature Reviews Genetics. 2000 Oct;1(1):11-9.

88. Richardson B. DNA methylation and autoimmune disease. Clinical Immunology. 2003 Oct 1;109(1):72-9.

89. Robertson KD. DNA methylation and human disease. Nature Reviews Genetics. 2005 Aug;6(8):597-610.

90. Bergman Y, Cedar H. DNA methylation dynamics in health and disease. Nature structural & molecular biology. 2013 Mar;20(3):274-81.

91. Rickabaugh TM, Baxter RM, Sehl M, Sinsheimer JS, Hultin PM, Hultin LE, Quach A, Martínez-Maza O, Horvath S, Vilain E, Jamieson BD. Acceleration of age-associated methylation patterns in HIV-1-infected adults. PloS one. 2015 Mar 25;10(3):e0119201

92. Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. The Journal of infectious diseases. 2015 Nov 15;212(10):1563-73.

93. Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. The Journal of infectious diseases. 2015 Nov 15;212(10):1563-73.

94. Levine ME, Hosgood HD, Chen B, Absher D, Assimes T, Horvath S. DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. Aging (Albany NY). 2015 Sep;7(9):690.

95. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R. Genome-wide methylation profiles reveal quantitative views of human aging rates. Molecular cell. 2013 Jan 24;49(2):359-67.

96. Levine ME, Lu AT, Bennett DA, Horvath S. Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. Aging (Albany NY). 2015 Dec;7(12):1198.

97. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018 Apr;10(4):573.

98. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, Whitsel EA. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY). 2019 Jan 31;11(2):303.

99. Zhang X, Asllanaj E, Amiri M, Portilla-Fernandez E, Bramer WM, Nano J, Voortman T, Pan Q, Ghanbari M. Deciphering the role of epigenetic modifications in fatty liver disease: A systematic review. European Journal of Clinical Investigation. 2020 Dec 22:e13479.

100. Pirola CJ, Gianotti TF, Burgueño AL, Rey-Funes M, Loidl CF, Mallardi P, San Martino J, Castaño GO, Sookoian S. Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. Gut. 2013 Sep 1;62(9):1356-63.

101. de Mello VD, Matte A, Perfilyev A, Männistö V, Rönn T, Nilsson E, Käkelä P, Ling C, Pihlajamäki J. Human liver epigenetic alterations in non-alcoholic steatohepatitis are related to insulin action. Epigenetics. 2017 Apr 3;12(4):287-95.

102. Sookoian S, Rosselli MS, Gemma C, Burgueño AL, Fernández Gianotti T, Castaño GO, Pirola CJ. Epigenetic regulation of insulin resistance in nonalcoholic fatty liver disease: Impact of liver methylation of the peroxisome proliferator–activated receptor γ coactivator 1α promoter. Hepatology. 2010 Dec;52(6):1992-2000.

103. Kordi-Tamandani DM, Hashemi M, Birjandian E, Bahari A, Valizadeh J, Torkamanzehi A. Lack of association of GSTT1 and GSTP1 genes methylation and their expression profiles with risk of NAFLD in a sample of Iranian patients. Clinics and research in hepatology and gastroenterology. 2011 May 1;35(5):387-92.

104. Murphy SK, Yang H, Moylan CA, Pang H, Dellinger A, Abdelmalek MF, Garrett ME, Ashley–Koch A, Suzuki A, Tillmann HL, Hauser MA. Relationship between methylome and transcriptome in patients with nonalcoholic fatty liver disease. Gastroenterology. 2013 Nov 1;145(5):1076-87.

105. Kitamoto T, Kitamoto A, Ogawa Y, Honda Y, Imajo K, Saito S, Yoneda M, Nakamura T, Nakajima A, Hotta K. Targeted-bisulfite sequence analysis of the methylation of CpG islands in genes encoding PNPLA3, SAMM50, and PARVB of patients with non-alcoholic fatty liver disease. Journal of hepatology. 2015 Aug 1;63(2):494-502.

106. Hotta K, Kikuchi M, Kitamoto T, Kitamoto A, Ogawa Y, Honda Y, Kessoku T, Kobayashi K, Yoneda M, Imajo K, Tomeno W. Identification of core gene networks and hub genes associated with progression of non-alcoholic fatty liver disease by RNA sequencing. Hepatology Research. 2017 Dec;47(13):1445-58.

107. Mwinyi J, Boström AE, Pisanu C, Murphy SK, Erhart W, Schafmayer C, Hampe J, Moylan C, Schiöth HB. NAFLD is associated with methylation shifts with relevance for the expression of genes involved in lipoprotein particle composition. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids. 2017 Mar 1;1862(3):314-23.

108. Nano J, Ghanbari M, Wang W, de Vries PS, Dhana K, Muka T, Uitterlinden AG, van Meurs JB, Hofman A, Franco OH, Pan Q. Epigenome-wide association study identifies methylation sites associated with liver enzymes and hepatic steatosis. Gastroenterology. 2017 Oct 1;153(4):1096-106.

109. Ma J, Nano J, Ding J, Zheng Y, Hennein R, Liu C, Speliotes EK, Huan T, Song C, Mendelson MM, Joehanes R. A peripheral blood DNA methylation signature of hepatic fat reveals a potential causal pathway for nonalcoholic fatty liver disease. Diabetes. 2019 May 1;68(5):1073-83.

110. Zeybel M, Hardy T, Robinson SM, Fox C, Anstee QM, Ness T, Masson S, Mathers JC, French J, White S, Mann J. Differential DNA methylation of genes involved in fibrosis progression in non-alcoholic fatty liver disease and alcoholic liver disease. Clinical epigenetics. 2015 Dec;7(1):1-1.

111. Zakhari S. Alcohol metabolism and epigenetics changes. Alcohol research: current reviews. 2013;35(1):6.

112. Varela-Rey M, Woodhoo A, Martinez-Chantar ML, Mato JM, Lu SC. Alcohol, DNA methylation, and cancer. Alcohol research: current reviews. 2013;35(1):25.

113. Rosen AD, Robertson KD, Hlady RA, Muench C, Lee J, Philibert R, Horvath S, Kaminsky ZA, Lohoff FW. DNA methylation age is accelerated in alcohol dependence. Translational psychiatry. 2018 Sep 5;8(1):1-8.

114. French SW. Epigenetic events in liver cancer resulting from alcoholic liver disease. Alcohol research: current reviews. 2013;35(1):57.

115. Wen L, Li J, Guo H, Liu X, Zheng S, Zhang D, Zhu W, Qu J, Guo L, Du D, Jin X. Genome-scale detection of hypermethylated CpG islands in circulating cell-free DNA of hepatocellular carcinoma patients. Cell research. 2015 Nov;25(11):1250-64.

116. Nishida N, Nagasaka T, Nishimura T, Ikai I, Boland CR, Goel A. Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. Hepatology. 2008 Mar;47(3):908-18.

117. Su PF, Lee TC, Lin PJ, Lee PH, Jeng YM, Chen CH, Liang JD, Chiou LL, Huang GT, Lee HS. Differential DNA methylation associated with hepatitis B virus infection in hepatocellular carcinoma. International journal of cancer. 2007 Sep 15;121(6):1257-64.

118. Zhao Y, Xue F, Sun J, Guo S, Zhang H, Qiu B, Geng J, Gu J, Zhou X, Wang W, Zhang Z. Genome-wide methylation profiling of the different stages of hepatitis B virus-related hepatocellular carcinoma development in plasma cell-free DNA reveals potential biomarkers for early detection and high-risk monitoring of hepatocellular carcinoma. Clinical epigenetics. 2014 Dec;6(1):1-8.

119. Kao WY, Yang SH, Liu WJ, Yeh MY, Lin CL, Liu CJ, Huang CJ, Lin SM, Lee SD, Chen PJ, Yu MW. Genome-wide identification of blood DNA methylation patterns associated with early-onset hepatocellular carcinoma development in hepatitis B carriers. Molecular carcinogenesis. 2017 Feb;56(2):425-35.

120. Nano J, Ghanbari M, Wang W, de Vries PS, Dhana K, Muka T, Uitterlinden AG, van Meurs JB, Hofman A, Franco OH, Pan Q. Epigenome-wide association study identifies methylation sites associated with liver enzymes and hepatic steatosis. Gastroenterology. 2017 Oct 1;153(4):1096-106.

121. Karmen A, Wróblewski F, LaDue JS. Transaminase activity in human blood. The Journal of clinical investigation. 1955 Jan 1;34(1):126-33.

122. McGill MR. The past and present of serum aminotransferases and the future of liver injury biomarkers. EXCLI journal. 2016;15:817.

123. Rothschild MA, Oratz M, Schreiber SS. Serum albumin. Hepatology. 1988 Mar;8(2):385-401.

124. Fevery J. Bilirubin in clinical practice: a review. Liver International. 2008 May;28(5):592-605.

125. Afdhal N, McHutchison J, Brown R, Jacobson I, Manns M, Poordad F, Weksler B, Esteban R. Thrombocytopenia associated with chronic liver disease. Journal of hepatology. 2008 Jun 1;48(6):1000-7.

126. Mitchell O, Feldman DM, Diakow M, Sigal SH. The pathophysiology of thrombocytopenia in chronic liver disease. Hepatic medicine: evidence and research. 2016;8:39.

127. Vallet-Pichard A, Mallet V, Nalpas B, Verkarre V, Nalpas A, Dhalluin-Venier V, Fontaine H, Pol S. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. Hepatology. 2007 Jul;46(1):32-6.

128. Loaeza-del-Castillo A, Paz-Pineda F, Oviedo-Cárdenas E, Sánchez-Avila F, Vargas-Voráckova F. AST to platelet ratio index (APRI) for the noninvasive evaluation of liver fibrosis. Annals of hepatology. 2008;7(4):350-7.

129. Justice AC, Dombrowski E, Conigliaro J, Fultz SL, Gibson D, Madenwald T, Goulet J, Simberkoff M, Butt AA, Rimland D, Rodriguez-Barradas MC. Veterans aging cohort study (VACS): overview and description. Medical care. 2006 Aug;44(8 Suppl 2):S13.

130. Obersteller S, Neubauer H, Hagen RM, Frickmann H. Comparison of five commercial nucleic acid extraction kits for the PCR-based detection of Burkholderia Pseudomallei DNA in formalin-fixed, paraffin-embedded tissues. European Journal of Microbiology and Immunology. 2016 Sep;6(3):244-52.

131. Chai V, Vassilakos A, Lee Y, Wright JA, Young AH. Optimization of the PAXgene™ blood RNA extraction system for gene expression analysis of clinical samples. Journal of clinical laboratory analysis. 2005;19(5):182-8.

132. Houseman EA, Molitor J, Marsit CJ. Reference-free cell mixture adjustments in analysis of DNA methylation data. Bioinformatics. 2014 May 15;30(10):1431-9.

133. Barfield RT, Almli LM, Kilaru V, Smith AK, Mercer KB, Duncan R, Klengel T, Mehta D, Binder EB, Epstein MP, Ressler KJ. Accounting for population stratification in DNA methylation studies. Genetic epidemiology. 2014 Apr;38(3):231-41.

134. Molejon MI, Ropolo A, Re AL, Boggio V, Vaccaro MI. The VMP1-Beclin 1 interaction regulates autophagy induction. Scientific reports. 2013 Jan 11;3(1):1-1.

135. Vaccaro MI, Grasso D, Ropolo A, Iovanna JL, Cerquetti MC. VMP1 expression correlates with acinar cell cytoplasmic vacuolization in arginine-induced acute pancreatitis. Pancreatology. 2003 Jan 1;3(1):69-74.

136. Liu SY, Chen CL, Yang TT, Huang WC, Hsieh CY, Shen WJ, Tsai TT, Shieh CC, Lin CF. Albumin prevents reactive oxygen species-induced mitochondrial damage, autophagy, and apoptosis during serum starvation. Apoptosis. 2012 Nov;17(11):1156-69.

137. Shi Y, Su C, Cui W, Li H, Liu L, Feng B, Liu M, Su R, Zhao L. Gefitinib loaded folate decorated bovine serum albumin conjugated carboxymethyl-beta-cyclodextrin nanoparticles enhance drug delivery and attenuate autophagy in folate receptor-positive cancer cells. Journal of nanobiotechnology. 2014 Oct;12(1):1-1.

138. Rottenberg ME, Carow B. SOCS3, a major regulator of infection and inflammation. Frontiers in immunology. 2014 Feb 19;5:58.

139. Jo D, Liu D, Yao S, Collins RD, Hawiger J. Intracellular protein therapy with SOCS3 inhibits inflammation and apoptosis. Nature medicine. 2005 Aug;11(8):892-8.

140. Ogata H, Chinen T, Yoshida T, Kinjyo I, Takaesu G, Shiraishi H, Iida M, Kobayashi T, Yoshimura A. Loss of SOCS3 in the liver promotes fibrosis by enhancing STAT3-mediated TGF-β 1 production. Oncogene. 2006 Apr;25(17):2520-30.

141. Dees C, Pötter S, Zhang Y, Bergmann C, Zhou X, Luber M, Wohlfahrt T, Karouzakis E, Ramming A, Gelse K, Yoshimura A. TGF-β–induced epigenetic deregulation of SOCS3 facilitates STAT3 signaling to promote fibrosis. The Journal of clinical investigation. 2020 Apr 6;130(5).

142. Jadid FZ, Chihab H, Alj HS, Elfihry R, Zaidane I, Tazi S, Badre W, Marchio A, El Filali KM, Tahiri M, Saile R. Control of progression towards liver fibrosis and hepatocellular carcinoma by SOCS3 polymorphisms in chronic HCV-infected patients. Infection, Genetics and Evolution. 2018 Dec 1;66:1-8.

143. Mori H, Hanada R, Hanada T, Aki D, Mashima R, Nishinakamura H, Torisu T, Chien KR, Yasukawa H, Yoshimura A. Socs3 deficiency in the brain elevates leptin sensitivity and confers resistance to diet-induced obesity. Nature medicine. 2004 Jul;10(7):739-43.

144. Zannas AS, Wiechmann T, Gassen NC, Binder EB. Gene–stress–epigenetic regulation of FKBP5: clinical and translational implications. Neuropsychopharmacology. 2016 Jan;41(1):261-74.

145. Sidibeh CO, Pereira MJ, Abalo XM, Boersma GJ, Skrtic S, Lundkvist P, Katsogiannos P, Hausch F, Castillejo-López C, Eriksson JW. FKBP5 expression in human adipose tissue: potential role in glucose and lipid metabolism, adipogenesis and type 2 diabetes. Endocrine. 2018 Oct;62(1):116-28.

146. Kusumanchi P, Liang T, Zhang T, Ross RA, Han S, Chandler K, Oshodi A, Jiang Y, Dent AL, Skill NJ, Huda N. Stress-responsive gene FKBP5 mediates alcohol-induced liver injury through the hippo pathway and CXCL1 signaling. Hepatology. 2020 Nov 1.

147. Stechschulte LA, Qiu B, Warrier M, Hinds Jr TD, Zhang M, Gu H, Xu Y, Khuder SS, Russo L, Najjar SM, Lecka-Czernik B. FKBP51 null mice are resistant to diet-induced obesity and the PPARγ agonist rosiglitazone. Endocrinology. 2016 Oct 1;157(10):3888-3900.

148. Cai MY, Luo RZ, Chen JW, Pei XQ, Lu JB, Hou JH, Yun JP. Overexpression of ZEB2 in peritumoral liver tissue correlates with favorable survival after curative resection of hepatocellular carcinoma. PLoS One. 2012 Feb 29;7(2):e32838.

149. Zhao YX, Sun YY, Huang AL, Li XF, Huang C, Ma TT, Li J. MicroRNA-200a induces apoptosis by targeting ZEB2 in alcoholic liver disease. Cell Cycle. 2018 Jan 17;17(2):250-62.

150. Xu X, Chang X, Xu Y, Deng P, Wang J, Zhang C, Zhu X, Chen S, Dai D. SAMD14 promoter methylation is strongly associated with gene expression and poor prognosis in gastric cancer. International journal of clinical oncology. 2020 Mar 21:1-0.

151. Willmer T, Johnson R, Louw J, Pheiffer C. Blood-based DNA methylation biomarkers for type 2 diabetes: potential for clinical applications. Frontiers in endocrinology. 2018 Dec 4;9:744.

152. Samblas M, Milagro FI, Martínez A. DNA methylation markers in obesity, metabolic syndrome, and weight loss. Epigenetics. 2019 May 4;14(5):421-44.

153. Das PM, Singal R. DNA methylation and cancer. Journal of clinical oncology. 2004 Nov 15;22(22):4632-42.

154. Li L, Choi JY, Lee KM, Sung H, Park SK, Oze I, Pan KF, You WC, Chen YX, Fang JY, Matsuo K. DNA methylation in peripheral blood: a potential biomarker for cancer molecular epidemiology. Journal of epidemiology. 2012 Sep 5;22(5):384-94.

TABLES

Table 1. Characteristics of participants grouped by platform used for genome-wide DNA methylation profiling.

Characteristic	450K (N = 473)	EPIC (N = 488)	Overall (N = 961)	*P*-value
ALT (units/L)	38.8 (24.9)	34.1 (21.2)	36.4 (23.2)	1.9×10^{-3}
AST (units/L)	43.4 (28.1)	37.0 (20.1)	40.2 (24.6)	7.3×10^{-5}
Serum Albumin (mg/dL)	3.84 (0.489)	3.95 (0.457)	3.90 (0.476)	2.9×10^{-4}
Total Serum Bilirubin (mg/dL)	0.78 (0.485)	0.76 (0.513)	0.77 (0.500)	0.61
Platelet Count ($x10^9$/L)	224 (71.3)	232 (69.4)	228 (70.4)	8.0×10^{-2}
FIB-4 Score	1.81 (1.22)	1.57 (0.943)	1.69 (1.09)	8.5×10^{-4}
APRI Score	0.58 (0.602)	0.47 (0.393)	0.52 (0.509)	8.4×10^{-4}
Age (years)	51.8 (7.53)	50.7 (7.48)	51.2 (7.52)	3.1×10^{-2}
Race				1.00
Black	402 (85.0%)	391 (80.1%)	793 (82.5%)	
Non-Black	71 (15.0%)	97 (19.9%)	168 (17.5%)	
Smoking Status				1.00
Not Current	203 (42.9%)	210 (43.0%)	413 (43.0%)	
Current	270 (57.1%)	278 (57.0%)	548 (57.0%)	
Diabetes Status				1.00
No	390 (82.5%)	406 (83.2%)	796 (82.8%)	
Yes	83 (17.5%)	82 (16.8%)	165 (17.2%)	
BMI	25.4 (4.44)	25.9 (4.43)	25.6 (4.44)	0.14
Missing	8 (1.7%)	9 (1.8%)	17 (1.8%)	
Alcohol Use				1.00
Non-Hazardous	195 (41.2%)	191 (39.1%)	386 (40.2%)	
Hazardous	276 (58.4%)	297 (60.9%)	573 (59.6%)	
Ever Hep C Infection				1.00
No	241 (51.0%)	312 (63.9%)	553 (57.5%)	
Yes	232 (49.0%)	176 (36.1%)	408 (42.5%)	
Ever Hep B Infection				1.00
No	399 (84.4%)	430 (88.1%)	829 (86.3%)	
Yes	52 (11.0%)	40 (8.2%)	92 (9.6%)	
CD4 Count	411 (252)	442 (263)	427 (258)	6.4×10^{-2}
Viral Load				1.00
Suppressed	196 (41.4%)	204 (41.8%)	400 (41.6%)	
Unsuppressed	274 (57.9%)	284 (58.2%)	558 (58.1%)	

Statistics for numeric variables are presented as mean (sd), while those for categorical variables are presented as count (%). Counts across levels of some categorical variables may not sum to the corresponding total due to missing values.

Table 2. Linear association of covariates with selected liver markers

	Liver Biomarker											
	ln(AST)			ln(ALT)			ln(FIB-4)			ln(APRI)		
Covariate	**β**	**SE**	***P***	**β**	**SE**	***P***	**β**	**SE**	***P***	**β**	**SE**	***P***
Age	$3.0x10^{-3}$	$2.2x10^{-3}$	0.17	$6.3x10^{-4}$	$2.5x10^{-3}$	0.80	$2.4x10^{-2}$	$2.2x10^{-3}$	$4.1x10^{-25}$	$3.6x10^{-3}$	$3.0x10^{-3}$	0.23
Race	$-6.3x10^{-2}$	$4.3x10^{-2}$	0.14	$5.1x10^{-2}$	$4.9x10^{-2}$	0.30	$-7.4x10^{-2}$	$4.4x10^{-2}$	$9.0x10^{-2}$	$-4.0x10^{-2}$	$5.9x10^{-2}$	0.50
Current Smoking	$9.1x10^{-3}$	$3.3x10^{-2}$	0.78	$-3.1x10^{-2}$	$3.7x10^{-2}$	0.41	$5.0x10^{-2}$	$3.3x10^{-2}$	0.14	$2.6x10^{-2}$	$4.5x10^{-2}$	0.57
Diabetes	$-8.1x10^{-3}$	$4.4x10^{-2}$	0.85	$2.9x10^{-2}$	$5.0x10^{-2}$	0.57	$-3.6x10^{-2}$	$4.4x10^{-2}$	0.41	$-2.1x10^{-2}$	$6.0x10^{-2}$	0.72
BMI	$-8.1x10^{-3}$	$3.7x10^{-3}$	0.03	$9.2x10^{-3}$	$4.2x10^{-3}$	$2.8x10^{-2}$	$-1.3x10^{-2}$	$3.8x10^{-3}$	$8.3x10^{-4}$	$8.1x10^{-3}$	$5.1x10^{-2}$	0.11
Hazardous Alcohol Use	0.13	$3.3x10^{-2}$	$7.5 x10^{-5}$	0.11	$3.8x10^{-2}$	$2.9x10^{-2}$	$6.8x10^{-2}$	$3.4x10^{-2}$	$4.3x10^{-2}$	0.12	$4.5x10^{-2}$	$1.1x10^{-2}$
Ever HCV	0.48	$3.0x10^{-2}$	$3.7x10^{-51}$	0.48	$3.5x10^{-2}$	$2.2 x10^{-39}$	0.37	$3.2x10^{-2}$	$4.2x10^{-29}$	0.61	$4.1x10^{-2}$	$8.7x10^{-45}$
Ever HBV	$9.4x10^{-2}$	$5.6x10^{-2}$	$9.5x10^{-2}$	0.10	$6.3x10^{-2}$	0.11	0.18	$5.6x10^{-2}$	$1.5x10^{-3}$	0.23	$7.6x10^{-2}$	$2.5x10^{-3}$
ART Use	$-7.0 x10^{-3}$	$4.3x10^{-2}$	0.87	$9.7x10^{-2}$	$4.9x10^{-2}$	$4.8x10^{-2}$	$-7.1x10^{-2}$	$4.4x10^{-2}$	0.11	$-2.1x10^{-2}$	$5.9x10^{-2}$	0.72
CD4 Count	$-1.7x10^{-4}$	$6.3x10^{-5}$	$8.0x10^{-3}$	$2.7x10^{-5}$	$7.2x10^{-5}$	0.71	$-4.7x10^{-4}$	$6.2x10^{-5}$	$6.7x10^{-14}$	$-4.1x10^{-4}$	$8.6x10^{-5}$	$2.5x10^{-6}$
Viral Load	$-6.7x10^{-2}$	$3.3x10^{-2}$	$4.7x10^{-2}$	$7.2x10^{-2}$	$3.8x10^{-2}$	$5.7x10^{-2}$	-0.17	$3.3x10^{-2}$	$7.0x10^{-7}$	-0.13	$4.6x10^{-2}$	$4.6x10^{-3}$
CD8 T Prop	-0.16	0.16	0.31	-0.16	0.18	0.38	$7.5x10^{-2}$	0.16	0.65	$6.2x10^{-3}$	0.22	0.98
CD4 T Prop	-0.23	0.29	0.42	0.44	0.33	0.18	-0.79	0.30	$8.1x10^{-3}$	-0.45	0.40	0.26
NK Prop	0.79	0.28	$4.8x10^{-3}$	0.63	0.31	$4.6x10^{-2}$	0.95	0.28	$8.4x10^{-4}$	1.2	0.38	$1.9x10^{-2}$
B Cell Prop	1.5	0.37	$4.9x10^{-5}$	1.9	0.41	$2.4x10^{-6}$	0.92	0.38	$1.5x10^{-2}$	1.8	0.51	$4.9x10^{-4}$
Mono Prop	1.6	0.41	$1.1x10^{-4}$	0.71	0.47	0.13	2.3	0.42	$2.3x10^{-8}$	2.6	0.57	$4.1x10^{-6}$
Gran Prop	-0.55	0.14	$1.2x10^{-4}$	-0.63	0.16	$8.4x10^{-5}$	-0.57	0.14	$8.7x10^{-5}$	-0.86	0.19	$9.3x10^{-6}$

Table 2 (Continued)

	Liver Biomarker								
	ALB			TBILI			Platelet Count		
Covariate	**β**	**SE**	***P***	**β**	**SE**	***P***	**β**	**SE**	***P***
Age	-1.4×10^{-3}	2.1×10^{-3}	0.49	3.4×10^{-3}	2.2×10^{-3}	0.12	-0.14	0.31	0.64
Race	0.24	4.0×10^{-2}	3.8×10^{-9}	0.12	4.3×10^{-2}	5.4×10^{-3}	-10.6	6.0	7.7×10^{-2}
Current Smoking	-0.13	3.1×10^{-2}	4.9×10^{-5}	-1.5×10^{-2}	3.3×10^{-2}	0.65	0.50	4.6	0.91
Diabetes	-4.2×10^{-2}	4.2×10^{-2}	0.31	2.5×10^{-2}	4.4×10^{-2}	0.56	3.4	6.2	0.58
BMI	1.3×10^{-2}	3.5×10^{-3}	3.3×10^{-4}	2.2×10^{-3}	3.7×10^{-3}	0.56	0.16	0.52	0.77
Hazardous Alcohol Use	-0.13	3.1×10^{-2}	2.0×10^{-5}	-3.0×10^{-2}	3.3×10^{-2}	0.37	3.1	4.7	0.51
Ever HCV	-0.19	3.1×10^{-2}	2.1×10^{-9}	8.1×10^{-2}	3.4×10^{-2}	1.7×10^{-2}	-23.7	4.7	4.2×10^{-7}
Ever HBV	-0.14	5.3×10^{-2}	7.4×10^{-3}	7.9×10^{-2}	5.6×10^{-2}	0.16	-18.3	7.8	2.0×10^{-2}
ART Use	0.18	4.1×10^{-2}	1.5×10^{-5}	0.13	4.3×10^{-2}	2.6×10^{-3}	5.9	6.1	0.33
CD4 Count	4.0×10^{-4}	5.9×10^{-5}	1.2×10^{-11}	2.6×10^{-5}	6.4×10^{-5}	0.69	5.7×10^{-2}	8.8×10^{-3}	1.4×10^{-10}
Viral Load	0.25	3.1×10^{-2}	3.4×10^{-15}	0.17	3.3×10^{-2}	6.3×10^{-7}	14.2	4.7	2.4×10^{-3}
CD8 T Prop	-8.9×10^{-2}	0.15	0.56	2.4×10^{-2}	0.16	0.88	-8.8	22.6	0.70
CD4 T Prop	1.7	0.27	6.4×10^{-10}	0.12	0.30	0.69	78.5	41.2	5.7×10^{-2}
NK Prop	-0.81	0.26	2.3×10^{-3}	-0.20	0.28	0.47	-117.7	39.1	2.7×10^{-3}
B Cell Prop	0.41	0.35	0.25	0.44	0.37	0.23	-51.4	51.6	0.32
Mono Prop	-0.55	0.39	0.16	0.46	0.41	0.27	-183.8	57.8	1.5×10^{-3}
Gran Prop	-3.0×10^{-2}	0.14	0.83	-0.22	0.14	0.13	49.2	20.1	1.4×10^{-2}

Linear models are adjusted for age, except when age is the outcome. Results that remained significant after FDR correction to $Q < 0.05$ are highlighted in yellow. Abbreviations: AST, aspartate transaminase (units/L); ALT, alanine transaminase (units/L); ALB, serum albumin (mg/dL); TBILI, total bilirubin (mg/dL); PLT, platelet count (cells/μL).

Table 3. Associations between DNA methylation age acceleration and selected biomarkers of liver function

Liver Marker		Model			
		Liver Marker ~ PhenoAA + Covariates	Liver Marker ~ GrimAA + Covariates	Liver Marker ~ IEAA + Covariates	Liver Marker ~ EEAA + Covariates
ln(AST)	β	2.5×10^{-3}	7.9×10^{-3}	2.6×10^{-3}	1.9×10^{-3}
	SE	2.1×10^{-3}	4.0×10^{-3}	2.8×10^{-3}	2.8×10^{-3}
	P	0.23	0.05	0.36	0.51
ln(ALT)	β	1.9×10^{-4}	3.4×10^{-3}	1.3×10^{-3}	5.5×10^{-4}
	SE	2.5×10^{-3}	4.7×10^{-3}	3.3×10^{-3}	3.3×10^{-3}
	P	0.94	0.47	0.69	0.87
ALB	β	-7.0×10^{-3}	-7.2×10^{-3}	-4.1×10^{-3}	-4.3×10^{-3}
	SE	2.1×10^{-3}	4.0×10^{-3}	2.8×10^{-3}	2.8×10^{-3}
	P	8.6×10^{-4}	0.07	0.15	0.12
TBILI	β	-4.9×10^{-3}	-1.8×10^{-3}	2.1×10^{-3}	2.0×10^{-3}
	SE	2.3×10^{-3}	4.4×10^{-3}	3.2×10^{-3}	3.1×10^{-3}
	P	0.04	0.69	0.50	0.52
PLT	β	-0.48	-0.12	-0.21	-0.72
	SE	0.32	0.61	0.44	0.43
	P	0.14	0.84	0.64	0.09
ln(FIB-4)	β	1.4×10^{-3}	-5.5×10^{-4}	-1.4×10^{-3}	1.8×10^{-3}
	SE	2.2×10^{-3}	4.1×10^{-3}	2.9×10^{-3}	2.9×10^{-3}
	P	0.53	0.89	0.63	0.54
ln(APRI)	β	1.3×10^{-3}	3.2×10^{-3}	2.6×10^{-4}	3.7×10^{-3}
	SE	2.9×10^{-3}	5.5×10^{-3}	3.9×10^{-3}	3.9×10^{-3}
	P	0.64	0.56	0.95	0.34

Biomarker values were ln-transformed when needed to achieve a more normal distribution as indicated on the left and then regressed on each DNA methylation age acceleration metric in linear models that control for the covariates included in the EWAS model. Associations that remained significant after FDR adjustment at $Q < 0.05$ are highlighted in yellow. Abbreviations: AST, aspartate aminotransferase (units/L); ALT, alanine aminotransferase (units/L); ALB, serum albumin (mg/dL); TBILI, total bilirubin level (mg/dL); PLT, platelet count (cells/μL); FIB-4, FIB-4 score; APRI, APRI score; PhenoAA, phenotype age acceleration; GrimaAA, Grim age acceleration; IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration.

Table 4. CpG sites significantly associated with selected liver markers after meta-analysis and FDR correction

				Serum Albumin								
				Meta-analysis			EPIC			450K		
CpG	Chr.	Pos.	Gene	β	SE	*P*	β	SE	*P*	β	SE	*P*
cg16936953	17	57915665	*TMEM49*	1.2	0.21	$4.2x10^{-9}$	1.0	0.29	$6.6x10^{-4}$	1.4	0.29	$1.7x10^{-6}$
cg18181703	17	76354621	*SOCS3*	1.9	0.35	$9.5x10^{-8}$	2.3	0.45	$9.9x10^{-7}$	1.3	0.55	$2.1x10^{-2}$
cg03546163	6	35654363	*FKBP5*	1.1	0.20	$1.1x10^{-7}$	1.2	0.26	$9.2x10^{-6}$	0.94	0.33	$4.5x10^{-3}$
cg18942579	17	57915773	*TMEM49*	1.4	0.26	$1.1x10^{-7}$	1.2	0.34	$7.3x10^{-4}$	1.7	0.41	$3.2x10^{-5}$
cg01409343	17	57915740	*TMEM49*	1.7	0.32	$1.2x10^{-7}$	1.7	0.43	$1.4x10^{-4}$	1.8	0.48	$3.1x10^{-4}$
cg20995564	2	145172035	*ZEB2*	1.2	0.23	$1.9x10^{-7}$	1.5	0.29	$3.0x10^{-7}$	0.66	0.39	$9.1x10^{-2}$
cg23966214	17	48203188	*SAMD14*	3.5	0.70	$4.7x10^{-7}$	3.2	0.81	$8.7x10^{-5}$	4.4	1.4	$1.6x10^{-3}$
cg12054453	17	57915717	*TMEM49*	1.0	0.20	$5.0x10^{-7}$	0.7	0.29	$1.1x10^{-2}$	1.3	0.28	$8.4x10^{-6}$
cg12992827	3	101901234	-	1.9	0.38	$8.0x10^{-7}$	2.2	0.47	$4.0x10^{-6}$	1.3	0.66	$5.5x10^{-2}$

Results for CpG sites where DNA methylation remained significantly associated with serum albumin after meta-analysis and FDR correction to $Q < 0.05$ are presented, including statistics from the meta-analysis and from the separate EWAS of the two cohorts. No significant associations were identified after meta-analysis and FDR correction for AST, ALT, total serum albumin, platelet count, FIB-4 score, or APRI score, so those markers are excluded from the table. β coefficients represent the average change in serum albumin (mg/dL) expected for a increase in DNA methylation from 0% to 100%.

FIGURES

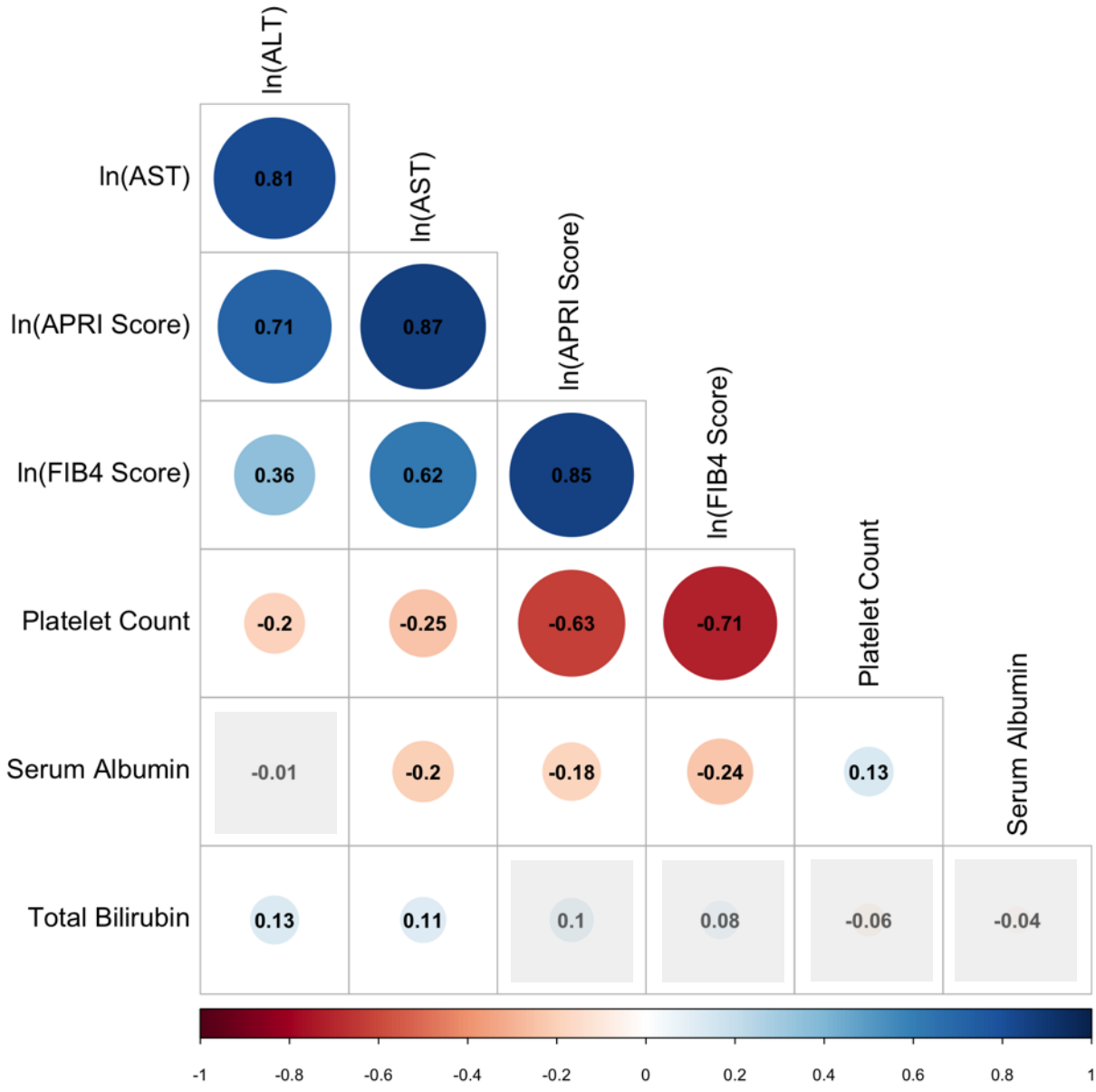


Figure 1: Correlation matrix based on Pearson coefficients for selected liver markers among all study participants. The strength and direction of each correlation is indicated by the size and color of its representative circle in the matrix. A legend for the colors is included beneath the matrix. Values of the Pearson correlation coefficients are also presented within the circles. Grayed-out squares indicate that the observed correlation coefficient is not statistically significant after Bonferroni correction to an overall $\alpha = 0.05$.

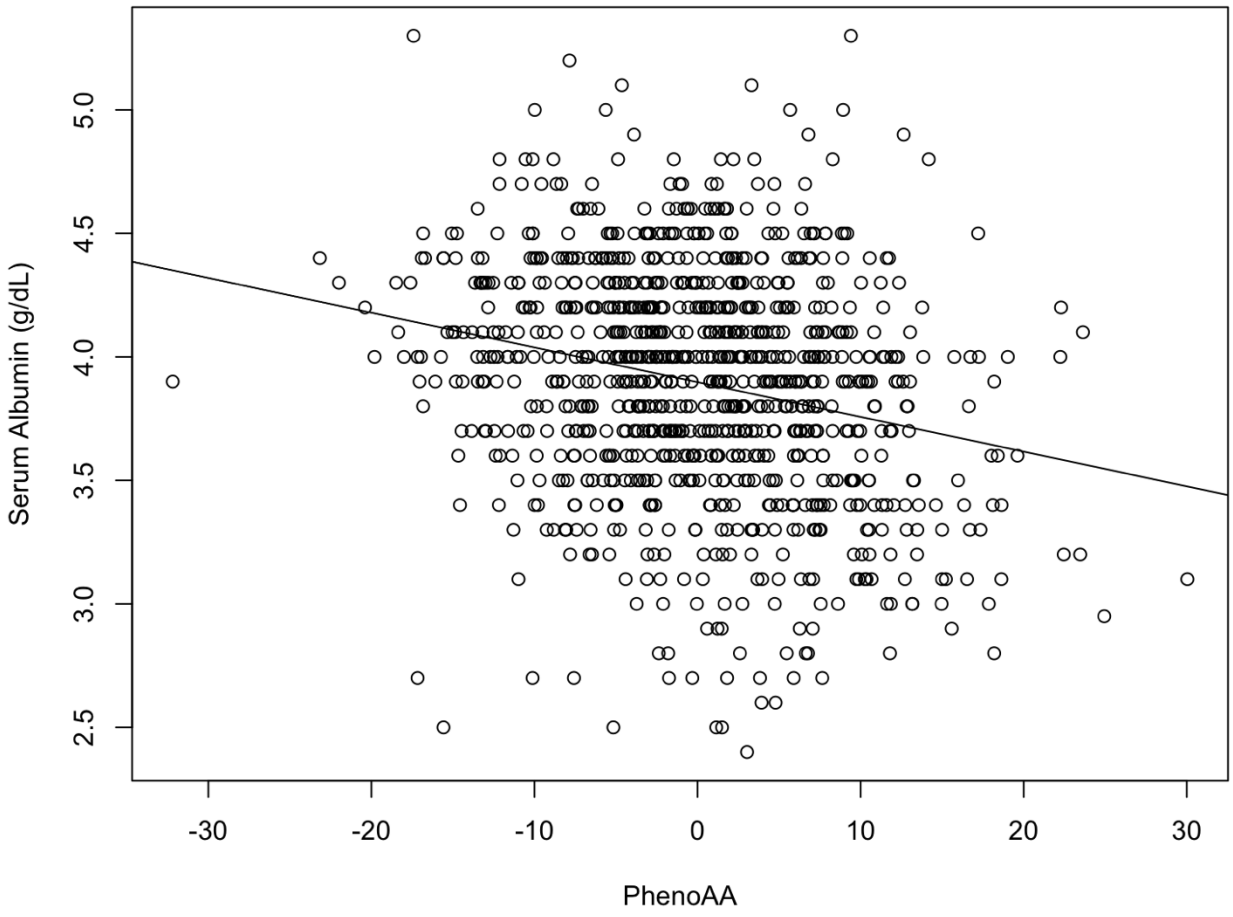


Figure 2: **Scatter plot of serum albumin values across values for phenotypic age acceleration (PhenoAA) among the entire cohort.** An unadjusted best-fit line is included (β = -0.014, *P*-value = 1.6×10^{-13}).

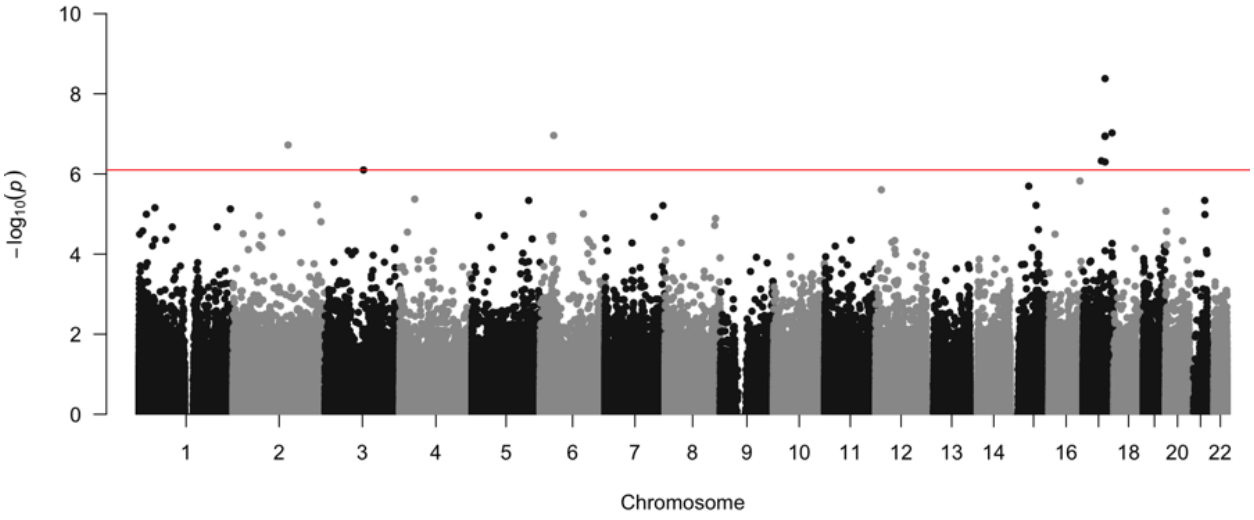


Figure 3. Manhattan plot of unadjusted *P*-values from meta-analysis of EWAS results for serum albumin. The red line indicates the unadjusted *P*-value that corresponds to a threshold for FDR significance at Q < 0.05. DNA methylation was not significantly associated with serum albumin at any CpG site after a more restrictive Bonferroni correction to an overall $\alpha = 0.05$.

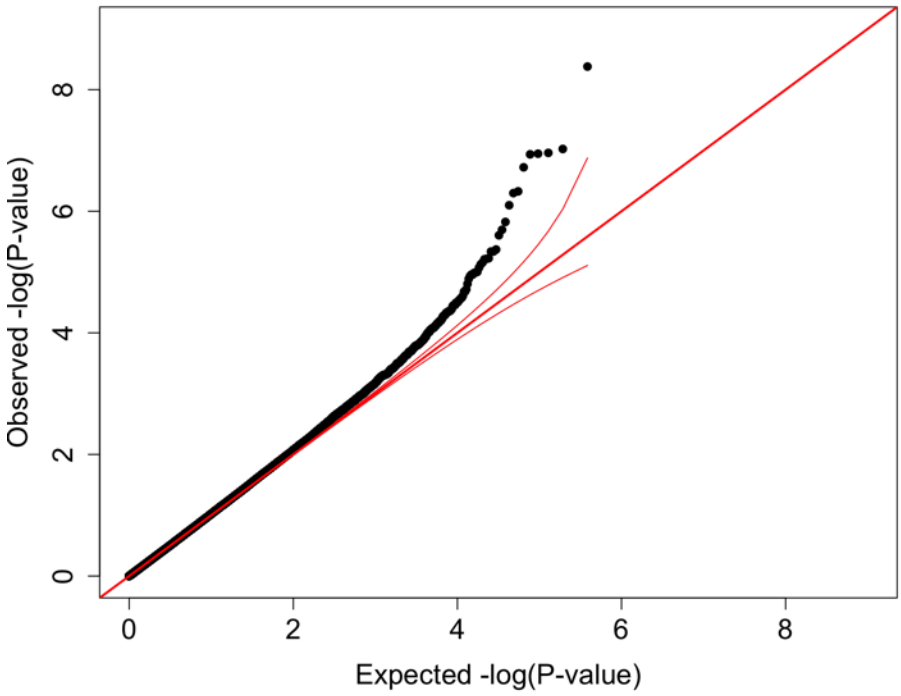


Figure 4: **Quantile-quantile plot of unadjusted *P*-values from meta-analysis of EWAS results for serum albumin.** The global inflation factor was 1.02. Red lines represent a perfect 1:1 association with a 95% confidence interval.

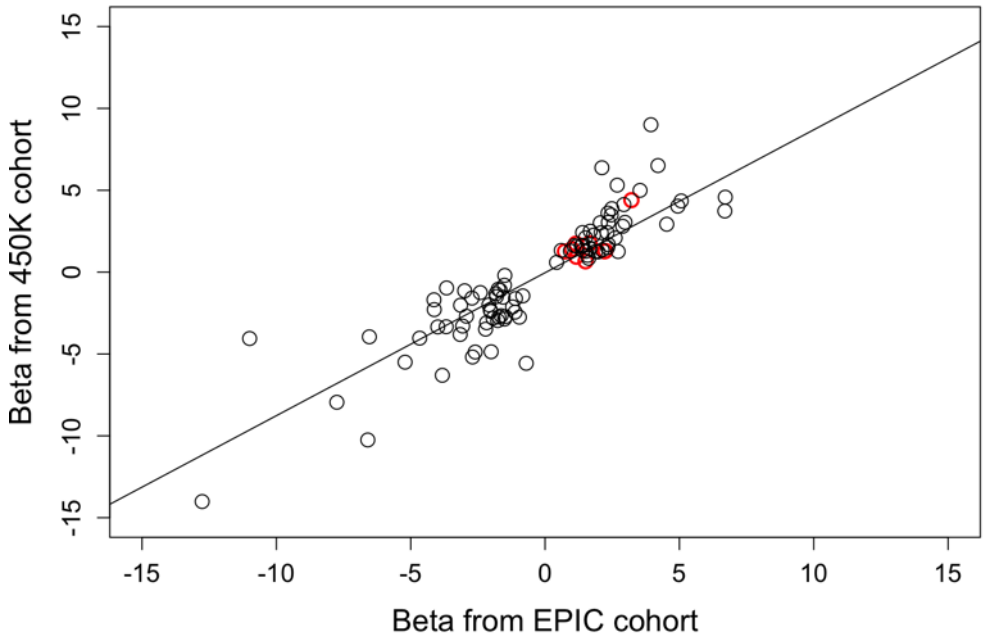


Figure 5: **Correlation of β coefficients obtained from separate EWAS of serum albumin among the EPIC and 450K cohorts for the 100 CpG sites with the smallest unadjusted *P*-values after meta-analysis.** Beta coefficients from the separate analyses are positively correlated ($\rho = 0.89$, *P*-value < 0.001), indicating that the signs of observed beta coefficient are consistent between the two datasets. Circles representing *P*-value pairs that were significant after meta-analysis and FDR correction to $Q < 0.05$ are highlighted in red.

Supplementary Tables & Figures

Supplementary Table 1: Top meta-analyzed CpG sites with smallest unadjusted *P-values* for association between methylation and each selected biomarker.

ln(AST)

				Meta			EPIC			450K		
CpG	Chr	Pos	Gene	β	SE	*P*	β	SE	*P*	β	SE	*P*
cg14218844	18	56709536	*LOC390858*	3.4	0.66	$2.7x10^{-7}$	4.1	0.76	$1.9x10^{-7}$	1.4	1.3	0.27
cg26312951	21	42797847	*MX1*	-0.87	0.18	$1.1x10^{-6}$	-0.86	0.28	$2.3x10^{-3}$	-0.89	0.23	$1.9x10^{-4}$
cg19939130	1	158978468	*IFI16*	-2.9	0.59	$1.2x10^{-6}$	-2.3	0.70	$8.6x10^{-4}$	-4.3	1.1	$1.7x10^{-4}$
cg18759102	1	3100343	*PRDM16*	-4.2	0.87	$1.9x10^{-6}$	-4.4	1.0	$3.0x10^{-5}$	-3.6	1.6	$2.6x10^{-2}$
cg06188083	10	91093005	*IFIT3*	-0.76	0.16	$3.2x10^{-6}$	-0.60	0.22	$7.9\ x10^{-3}$	-0.95	0.24	$9.2x10^{-5}$
cg11829870	22	50988451	*KLHDC7B*	-1.7	0.38	$5.5x10^{-6}$	-1.1	0.48	$2.8x10^{-2}$	-2.8	0.61	$7.5x10^{-6}$
cg02247863	22	50983415	-	-1.9	0.43	$7.1x10^{-6}$	-1.4	0.56	$1.6x10^{-2}$	-2.8	0.68	$4.7x10^{-5}$
cg15796173	22	26934515	*TPST2*	5.4	1.2	$9.9x10^{-6}$	4.3	1.5	$3.3x10^{-3}$	7.9	2.2	$4.3x10^{-4}$
cg14793819	3	196669628	*LOC152217 NCBP2*	2.9	0.66	$1.0x10^{-5}$	3.5	0.79	$1.0x10^{-5}$	1.4	1.2	0.24
cg21366673	6	30459512	*HLA-E*	-1.5	0.35	$1.1x10^{-5}$	-1.2	0.40	$9.4x10^{-3}$	-3.0	0.70	$2.1x10^{-5}$

ln(ALT)

				Meta			EPIC			450K		
CpG	Chr	Pos	Gene	β	SE	*P*	β	SE	*P*	β	SE	*P*
cg11367159	12	110044531	-	-1.6	0.31	$5.5x10^{-7}$	-2.0	0.42	$3.1x10^{-6}$	-1.0	0.47	$2.7x10^{-2}$
cg03140421	16	67200181	*HSF4*	-2.7	0.56	$2.3x10^{-6}$	-2.0	0.78	$1.2x10^{-2}$	-3.4	0.82	$3.3x10^{-5}$
cg20204986	11	32448067	*WT1*	-2.1	0.45	$3.2x10^{-6}$	-2.0	0.56	$3.4x10^{-4}$	-2.2	0.76	$3.4x10^{-3}$
cg06780216	16	2374397	*ABCA3*	8.0	1.7	$3.8x10^{-6}$	8.9	2.8	$2.0x10^{-3}$	7.5	2.2	$6.7x10^{-4}$
cg04363281	2	121224648	*LOC84931*	-2.4	0.52	$5.6x10^{-6}$	-2.2	0.68	$1.2x10^{-3}$	-2.6	0.82	$1.6x10^{-3}$
cg19867481	5	1282984	*TERT*	4.0.	0.88	$6.2x10^{-6}$	3.8	1.2	$1.8x10^{-3}$	4.2	1.3	$1.3x10^{-3}$
cg16904639	1	224033060	*TP53BP2*	6.0	1.3	$6.3x10^{-6}$	2.9	1.8	0.11	9.3	1.9	$1.9x10^{-6}$
cg27127351	8	670293	*ERICH1*	16	3.6	$7.1x10^{-6}$	16.9	3.9	$1.6x10^{-5}$	11	9.9	0.26
cg26312951	21	42797847	*MX1*	-0.94	0.21	$8.5x10^{-6}$	-0.96	0.35	$6.1x10^{-3}$	-0.92	0.26	$5.4x10^{-4}$
cg13946956	9	899635	*DMRT1*	-1.5	0.34	$8.6x10^{-6}$	-2.4	0.44	$2.4x10^{-7}$	-0.35	0.52	0.51

Supplementary Table 1 (Continued)

ALB

				Meta			EPIC			450K		
CpG	Chr	Pos	Gene	β	SE	*P*	β	SE	*P*	β	SE	*P*
cg16936953	17	57915665	*TMEM49*	1.2	0.21	$4.2x10^{-9}$	1.0	0.29	$6.6x10^{-4}$	1.4	0.29	$1.7x10^{-6}$
cg18181703	17	76354621	*SOCS3*	1.9	0.35	$9.5x10^{-8}$	2.3	0.45	$9.9x10^{-7}$	1.3	0.55	$2.1x10^{-2}$
cg03546163	6	35654363	*FKBP5*	1.1	0.20	$1.1x10^{-7}$	1.2	0.26	$9.2x10^{-6}$	0.94	0.33	$4.5x10^{-3}$
cg18942579	17	57915773	*TMEM49*	1.4	0.26	$1.1x10^{-7}$	1.2	0.34	$7.3x10^{-4}$	1.7	0.41	$3.2x10^{-5}$
cg01409343	17	57915740	*TMEM49*	1.7	0.32	$1.2x10^{-7}$	1.7	0.43	$1.4x10^{-4}$	1.8	0.48	$3.1x10^{-4}$
cg20995564	2	145172035	*ZEB2*	1.2	0.23	$1.9x10^{-7}$	1.5	0.29	$3.0x10^{-7}$	0.66	0.39	$9.1x10^{-2}$
cg23966214	17	48203188	*SAMD14*	3.5	0.70	$4.7x10^{-7}$	3.2	0.81	$8.7x10^{-5}$	4.4	1.4	$1.6x10^{-3}$
cg12054453	17	57915717	*TMEM49*	1.0	0.20	$5.0x10^{-7}$	0.74	0.29	$1.1x10^{-2}$	1.3	0.28	$8.4x10^{-6}$
cg12992827	3	101901234	-	1.9	0.38	$8.0x10^{-7}$	2.2	0.47	$4.0x10^{-6}$	1.3	0.66	$5.5x10^{-2}$
cg16583186	16	81526361	*CMIP*	1.4	0.29	$1.5x10^{-6}$	1.2	0.44	$7.5x10^{-3}$	1.6	0.39	$6.4x10^{-5}$

TBILI

				Meta			EPIC			450K		
CpG	Chr	Pos	Gene	β	SE	*P*	β	SE	*P*	β	SE	*P*
cg21497060	2	152590396	*NEB*	-3.6	0.71	$3.0x10^{-7}$	-4.0	0.79	$1.1x10^{-6}$	-2.4	1.6	0.13
cg23091255	6	64345855	-	3.3	0.68	$9.5x10^{-7}$	5.2	0.85	$2.3x10^{-9}$	-0.08	1.4	0.94
cg23747525	20	37064089	*LOC388796 SNORA71D*	6.9	1.4	$1.1x10^{-6}$	6.9	1.6	$1.6x10^{-5}$	6.8	3.2	$3.5x10^{-2}$
cg06882901	10	25013879	*ARHGAP21*	-11	2.5	$2.7x10^{-6}$	-12	3.4	$3.7x10^{-4}$	-11	3.6	$2.6x10^{-3}$
cg00617975	22	46644463	*C22orf40*	3.5	0.75	$3.5x10^{-6}$	3.4	0.80	$4.0x10^{-5}$	4.2	2.0	$3.9x10^{-2}$
cg12723809	4	87813066	*C4orf36*	3.7	0.80	$3.8x10^{-6}$	4.6	1.0	$8.5x10^{-6}$	2.2	1.3	$9.1x10^{-2}$
cg07134589	8	110551993	*EBAG9*	5.3	1.2	$6.0x10^{-6}$	5.7	1.2	$3.9x10^{-6}$	0.15	4.3	0.97
cg06783121	5	176314628	*HK3*	-4.1	0.91	$7.4x10^{-6}$	-4.9	1.3	$1.3x10^{-4}$	-3.2	1.3	$1.6x10^{-2}$
cg03293770	11	61197075	*SDHAF2 CPSF7*	3.9	0.87	$7.6x10^{-6}$	3.9	1.6	$1.4x10^{-2}$	3.9	1.0	$2.2x10^{-4}$
cg20754634	17	7218651	*GPS2*	6.7	1.5	$8.9x10^{-6}$	5.1	2.0	$1.3x10^{-2}$	8.8	2.3	$1.3x10^{-4}$

Supplementary Table 1 (Continued)

PLT

				Meta			EPIC			450K		
CpG	Chr	Pos	Gene	β	SE	*P*	β	SE	*P*	β	SE	*P*
acg26563141	2	88124876	*RGPD2*	-254	49	$2.0x10^{-7}$	-234	67	$5.2x10^{-4}$	-277	72	$1.3x10^{-4}$
cg15971496	17	695156	*RNMTL1*	-407	85	$1.7 x10^{-6}$	-563	118	$2.7x10^{-6}$	-237	123	$5.6x10^{-2}$
cg20752553	3	43527232	*ANO10*	1006	216	$3.4 x10^{-6}$	1213	263	$5.6x10^{-6}$	569	382	0.14
cg20819397	9	123964132	*RAB14*	270	60	$6.6x10^{-6}$	242	86	$5.2x10^{-3}$	296	83	$4.4x10^{-4}$
cg26346197	5	53813541	*SNX18*	-1798	408	$1.1x10^{-5}$	-2288	654	$5.3 x10^{-4}$	-1485	523	$4.8 x10^{-3}$
cg00092309	4	185934797	-	177	41	$1.3x10^{-5}$	223	54	$4.8x10^{-5}$	118	61	$5.4x10^{-2}$
cg26110653	13	38924412	*UFM1*	454	104	$1.4x10^{-5}$	344	125	$6.4x10^{-3}$	705	189	$2.3x10^{-4}$
cg11296553	5	612297	*CEP72*	1134	265	$1.8x10^{-5}$	1207	354	$7.4x10^{-4}$	1042	398	$9.2x10^{-3}$
cg06812840	17	7492524	*SOX15*	-341	80	$2.2x10^{-5}$	-422	99	$2.9x10^{-5}$	-191	136	0.16
cg17941109	19	17407198	*ABHD8*	-513	122	$2.5x10^{-5}$	-658	157	$3.4x10^{-5}$	-292	193	0.13

ln(FIB-4)

				Meta			EPIC			450K		
CpG	Chr	Pos	Gene	β	SE	*P*	β	SE	*P*	β	SE	*P*
cg22955628	19	45752130	*NA*	-19	3.8	$9.9x10^{-7}$	-23.55	5.21	$8.8x10^{-6}$	-13	5.7	$2.1x10^{-2}$
cg05283597	3	52812410	*ITIH1*	2.3	0.49	$2.1x10^{-6}$	1.45	0.63	$2.3x10^{-2}$	3.6	0.76	$4.4x10^{-6}$
cg19393677	15	67841407	*MAP2K5*	2.9	0.61	$2.5x10^{-6}$	2.84	0.73	$1.1x10^{-4}$	2.9	1.1	$9.1x10^{-3}$
cg21946195	2	86012225	*ATOH8*	2.2	0.47	$4.4x10^{-6}$	2.22	0.57	$1.2x10^{-4}$	2.0	0.83	$1.6x10^{-2}$
cg06204735	14	74892986	*TMEM90A*	-2.4	0.52	$6.4x10^{-6}$	-2.41	0.58	$3.7x10^{-5}$	-2.1	1.2	$8.9x10^{-2}$
cg20249469	12	11801908	*ETV6*	-19	4.2	$6.8x10^{-6}$	-19.43	48	$5.8x10^{-5}$	-17	8.8	$5.6x10^{-2}$
cg10493186	1	3134756	*PRDM16*	1.5	0.33	$7.1x10^{-6}$	1.24	0.47	$8.6x10^{-3}$	1.7	0.47	$2.5x10^{-4}$
cg04603976	19	4052706	*ZBTB7A*	23	0.51	$8.2x10^{-6}$	1.92	0.68	$5.0x10^{-3}$	2.7	0.76	$4.7x10^{-4}$
cg26596719	7	142986429	*CASP2*	1.3	0.30	$9.5x10^{-6}$	1.73	0.40	$2.6x10^{-5}$	0.85	0.45	$6.3x10^{-2}$
cg22822630	13	27692257	*USP12*	6.2	1.4	$1.0x10^{-5}$	6.37	1.66	$1.5x10^{-4}$	5.8	2.6	$3.0x10^{-2}$

Supplementary Table 1 (Continued)

				ln(APRI)								
				Meta			EPIC			450K		
CpG	Chr	Pos	Gene	β	SE	*P*	β	SE	*P*	β	SE	*P*
cg22955628	19	45752130	-	-24	5.1	$2.3x10^{-6}$	-27	6.9	$9.5x10^{-5}$	-20	7.6	$7.9x10^{-3}$
cg26312951	21	42797847	*MX1*	-1.1	0.24	$2.9x10^{-6}$	-1.6	0.38	$2.7x10^{-5}$	-0.81	0.32	$1.2x10^{-2}$
cg06204735	14	74892986	*TMEM90A*	-3.2	0.69	$4.8x10^{-6}$	-3.3	0.76	$2.0x10^{-5}$	-2.6	1.7	0.12
cg21366673	6	30459512	*HLA-E*	-2.2	0.48	$5.4x10^{-6}$	-1.5	0.55	$7.2x10^{-3}$	-4.3	0.96	$1.4x10^{-5}$
cg12674760	5	178577498	*ADAMTS2*	-5.3	1.2	$6.9x10^{-6}$	-5.3	1.6	$1.3x10^{-3}$	-5.3	1.7	$2.0x10^{-3}$
cg01740002	2	45475777	-	5.1	1.1	$7.0x10^{-6}$	6.8	1.5	$1.1x10^{-5}$	3.0	1.7	$8.2x10^{-2}$
cg20298895	18	42260130	*SETBP1*	12	2.6	$8.7x10^{-6}$	12	3.1	$9.3x10^{-5}$	9.6	4.6	$3.7x10^{-2}$
cg20249469	12	11801908	*ETV6*	-25	5.6	$8.9x10^{-6}$	-26	6.3	$3.7x10^{-5}$	-19	12	0.11
cg07455789	2	232419317	-	3.8	0.87	$1.0x10^{-5}$	3.6	1.2	$3.6x10^{-3}$	4.0	1.2	$1.1x10^{-3}$
cg07011778	8	144105570	-	-4.6	1.1	$1.7x10^{-5}$	-4.3	1.3	$5.7x10^{-4}$	-5.1	2.0	$1.1x10^{-2}$

For each biomarker of liver function, the ten CpG sites where DNA methylation associates with the biomarker with the smallest *P*-values after meta-analysis are presented along with genomic annotations and EWAS statistics. CpG sites that were not included in both platforms and thus could not be meta-analyzed are excluded. Abbreviations: AST, aspartate aminotransferase (units/L); ALT, alanine aminotransferase (units/L); ALB, serum albumin (mg/dL); TBILI, total bilirubin level (mg/dL); PLT, platelet count (cells/μL).

Supplementary Table 2: Top non-meta-analyzed CpG sites with smallest unadjusted *P-values* for association between methylation and each selected biomarker.

ln(AST)

CpG	Chr	Pos	Gene	Platform	β	SE	*P*
cg09836914	6	33449182	-	EPIC	-1.9	0.39	$8.9x10^{-7}$
cg04948230	19	39402846	*CCER2*	EPIC	-1.8	0.37	$2.8x10^{-6}$
cg04543048	15	61055941	*RORA*	EPIC	3.0	0.64	$6.7x10^{-6}$
cg00115090	1	113242166	*MOV10*	EPIC	-5.2	1.2	$1.1x10^{-5}$
cg15015917	19	39402823	*CCER2*	EPIC	-1.6	0.35	$1.1x10^{-5}$
cg18428234	12	52652742	-	450K	2.8	0.62	$1.2x10^{-5}$
cg15870857	14	104742172	-	EPIC	1.8	0.41	$1.4x10^{-5}$
cg14429846	10	75991175	*ADK*	EPIC	-4.3	0.99	$1.8x10^{-5}$
cg14029100	14	75518411	*MLH3*	EPIC	-2.1	0.48	$1.9x10^{-5}$
cg10874977	4	4696317	*STX18-AS1*	EPIC	3.0	0.69	$2.0x10^{-5}$

ln(ALT)

CpG	Chr	Pos	Gene	Platform	β	SE	*P*
cg06347851	5	124347050	-	EPIC	-4.0	0.75	$2.2x10^{-7}$
cg09836914	6	33449182	-	EPIC	-2.5	0.48	$4.4x10^{-7}$
cg01754642	3	39150244	*GORASP1 TTC21AT*	EPIC	-4.2	0.83	$7.1x10^{-7}$
cg08756327	4	46895394	*COX7B2*	EPIC	3.8	0.81	$3.1x10^{-6}$
cg20535253	4	122875768	-	EPIC	6.1	1.3	$6.8x10^{-6}$
cg03196177	10	93862017	*CPEB3*	EPIC	6.5	1.4	$7.2x10^{-6}$
cg01087242	20	33030037	*ITCH*	EPIC	6.9	1.5	$8.2x10^{-6}$
cg23299862	1	53326448	*ZYG11A*	EPIC	-4.0	0.87	$8.8x10^{-6}$
cg10239476	1	159568244	-	EPIC	6.2	1.4	$9.1x10^{-6}$
cg20829688	13	42621527	*DGKH*	EPIC	-2.1	0.46	$1.1x10^{-5}$

Supplementary Table 2 (Continued)

ALB

CpG	Chr	Pos	Gene	Platform	β	SE	*P*
cg15020309	20	23848170	-	EPIC	2.1	0.45	$3.5x10^{-6}$
cg19700645	15	79705656	-	EPIC	-3.3	0.70	$3.6x10^{-6}$
cg25121721	4	160218463	*RAPGEF2*	EPIC	1.6	0.33	$4.4x10^{-6}$
cg02491794	2	47204734	*TTC7A*	EPIC	4.0	0.87	$4.9x10^{-6}$
cg22898516	12	3011514	*TULP3*	EPIC	3.0	0.65	$6.5x10^{-6}$
cg04615481	8	141074581	*TRAPPC9*	EPIC	4.6	1.0	$6.7x10^{-6}$
cg01532746	20	5510187	-	EPIC	1.1	0.23	$8.8x10^{-6}$
cg17144326	3	193963847	*LOC100505920*	EPIC	3.3	0.73	$9.0x10^{-6}$
cg11211173	1	160102000	*ATP1A2*	EPIC	3.4	0.75	$9.4x10^{-6}$
cg21682939	4	34260762	-	EPIC	1.6	0.36	$1.1x10^{-5}$

TBILI

CpG	Chr	Pos	Gene	Platform	β	SE	*P*
cg20717792	17	17111358	-	EPIC	-4.0	0.72	$5.6x10^{-8}$
cg26070749	8	129182593	-	EPIC	-5.6	1.2	$2.3x10^{-6}$
cg09704072	3	31872626	*OSBPL10*	EPIC	-3.4	0.72	$3.4x10^{-6}$
cg16082566	1	45355704	*EIF2B3*	EPIC	-2.6	0.56	$6.7x10^{-6}$
cg07091392	4	170533728	*NEK1*	EPIC	3.2	0.71	$7.0x10^{-6}$
cg11062835	2	55390630	-	EPIC	-4.6	1.0	$7.9x10^{-6}$
cg24762192	1	154292657	*AQP10*	EPIC	-4.9	1.1	$9.4x10^{-6}$
cg15586258	8	129806984	-	EPIC	9.5	2.1	$1.2x10^{-5}$
cg14087804	5	67101830	-	EPIC	-3.9	0.88	$1.5x10^{-5}$
cg13671412	19	5335254	*PTPRS*	EPIC	-4.0	0.72	$5.6x10^{-5}$

Supplementary Table 2 (Continued)

PLT

CpG	Chr	Pos	Gene	Platform	β	SE	*P*
cg02314501	3	14639595	-	EPIC	380	80	$3.1x10^{-6}$
cg03026689	6	74363884	*SLC17A5*	EPIC	2697	582	$5.3x10^{-6}$
cg12960616	12	95638334	*VEZT*	EPIC	-483	105	$6.3x10^{-6}$
cg06696623	17	20956123	-	EPIC	-816	178	$6.8x10^{-6}$
cg21627016	5	73704114	*LINC01331*	EPIC	-291	64	$8.2x10^{-6}$
cg15632659	6	42931860	*PEX6*	EPIC	-502	111	$8.7x10^{-6}$
cg17427349	15	38988004	*C15orf53*	EPIC	-514	114	$9.1x10^{-6}$
cg14535019	4	129433853	-	EPIC	711	159	$1.0x10^{-5}$
cg05284887	1	147233306	*GJA5*	EPIC	312	70	$1.2x10^{-5}$
cg10998631	10	134331816	-	EPIC	281	63	$1.2x10^{-5}$

ln(FIB-4)

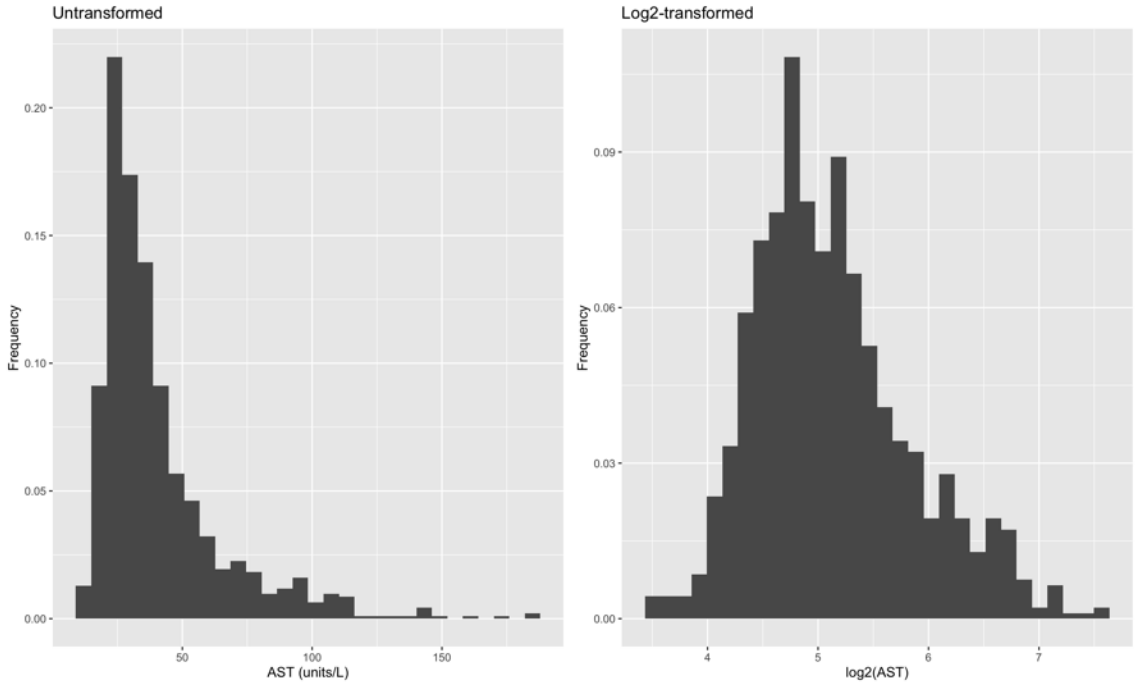
CpG	Chr	Pos	Gene	Platform	β	SE	*P*
cg12356793	17	13299004	-	EPIC	-2.5	0.55	$1.3x10^{-5}$
cg10913741	10	121244710	-	EPIC	-3.8	0.85	$1.4x10^{-5}$
cg11009066	13	25592074	-	EPIC	-2.6	0.61	$2.1x10^{-5}$
cg08034051	17	48949161	-	EPIC	1.9	0.44	$2.5x10^{-5}$
cg16130802	11	118773983	*BCL9L*	450K	12	2.6	$2.6x10^{-5}$
cg14397434	1	237288473	*RYR2*	EPIC	-2.5	0.57	$2.6x10^{-5}$
cg00096756	1	14678278	-	EPIC	-2.0	0.46	$2.8x10^{-5}$
cg23481873	20	464190	*CSNK2A1*	EPIC	-2.4	0.56	$3.3x10^{-5}$
cg24001468	2	62901049	*EHBP1*	450K	-4.6	1.1	$4.2x10^{-5}$
cg26102728	7	28218933	*JAZF1*	450K	-3.7	0.90	$4.4x10^{-5}$

Supplementary Table 2 (Continued)

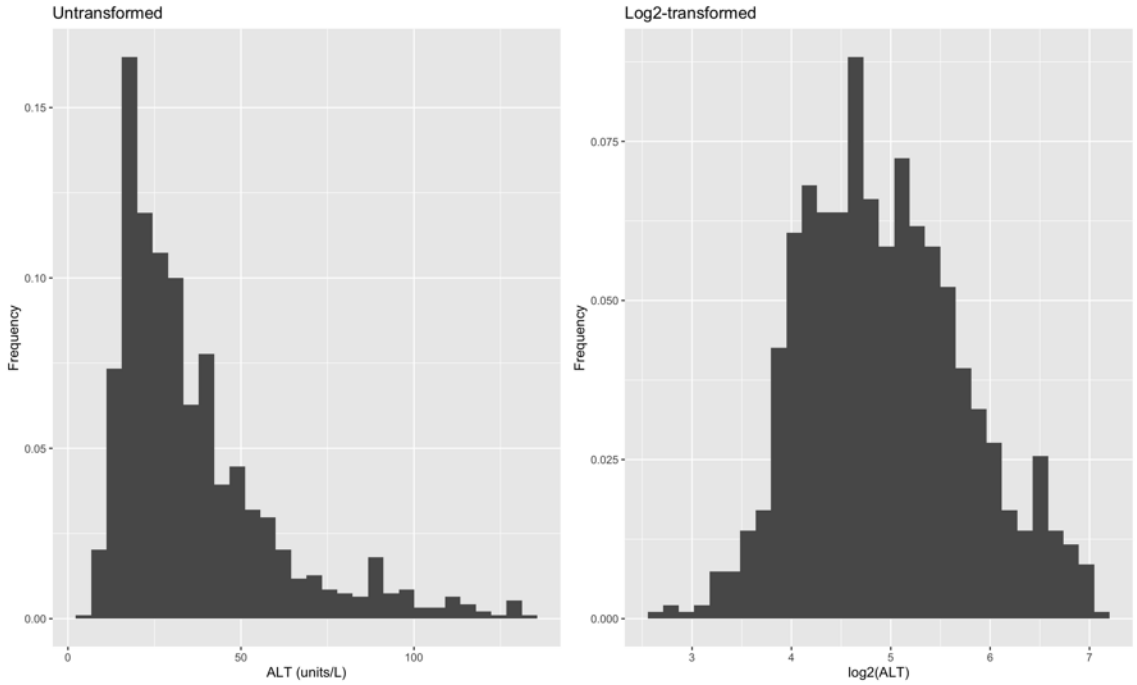
					ln(APRI)		
CpG	Chr	Pos	Gene	Platform	β	SE	*P*
cg08034051	17	48949161	-	EPIC	2.9	0.57	7.8×10^{-7}
cg04057469	3	16472844	*RFTN1*	EPIC	-6.7	1.4	2.8×10^{-6}
cg03196177	10	93862017	*CPEB3*	EPIC	7.3	1.6	4.9×10^{-6}
cg16706690	21	15135580	-	EPIC	-2.6	0.56	5.3×10^{-6}
cg22547004	20	2822133	*PCED1A VPS16*	EPIC	5.9	1.3	8.8×10^{-6}
cg16002826	3	123011707	*ADCY5*	EPIC	-2.3	0.54	1.9×10^{-5}
cg19306317	3	156654260	*LEKR1*	EPIC	1.5	0.34	2.6×10^{-5}
cg06427816	17	1090326	*ABR*	EPIC	3.2	0.76	2.8×10^{-5}
cg05198507	5	138735248	*SPATA24*	EPIC	-3.9	0.92	2.9×10^{-5}
cg26044128	1	33842402	*PHC2*	EPIC	-3.8	0.91	3.0×10^{-5}

For each biomarker of liver function, the ten CpG sites not included in both platforms (that therefore could not be meta-analyzed) where DNA methylation associates with the biomarker with the smallest *P*-values are presented along with genomic annotations and EWAS statistics. CpG sites that could be meta-analyzed are excluded. Abbreviations: AST, aspartate aminotransferase (units/L); ALT, alanine aminotransferase (units/L); ALB, serum albumin (mg/dL); TBILI, total bilirubin level (mg/dL); PLT, platelet count (cells/μL).

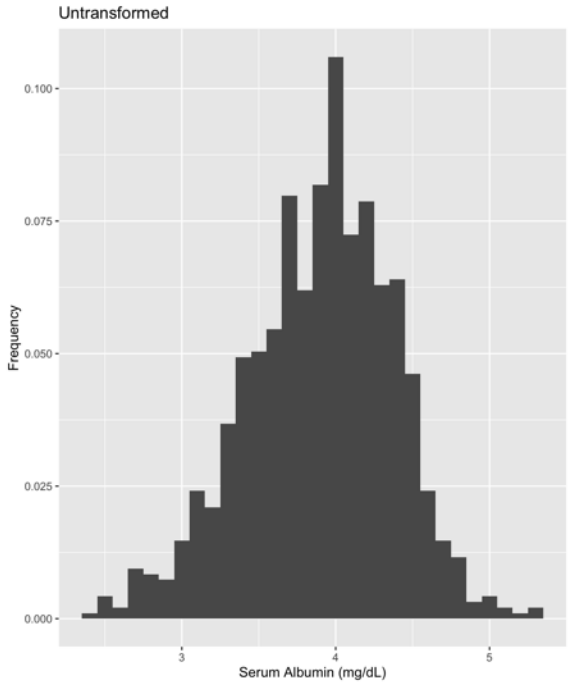
A


Untransformed
Frequency
0.20
0.15
0.10
0.05
0.00
50
100
150
AST (units/L)
Log2-transformed
Frequency
0.09
0.06
0.03
0.00
4
5
6
7
log2(AST)

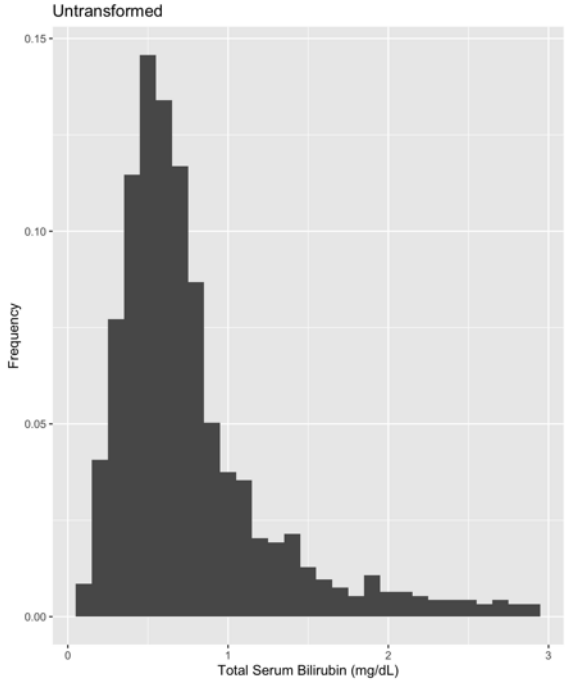
B


Untransformed
Frequency
0.15
0.10
0.05
0.00
0
50
100
ALT (units/L)
Log2-transformed
Frequency
0.075
0.050
0.025
0.000
3
4
5
6
7
log2(ALT)

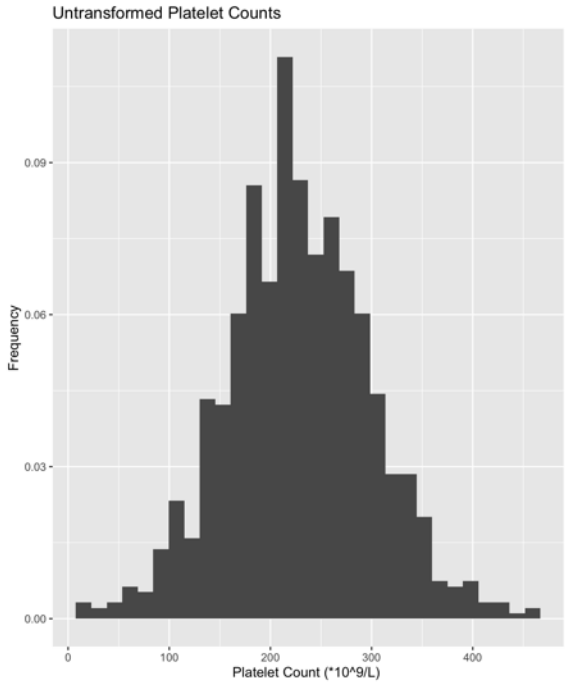
C


Untransformed
0.100
0.075
0.050
0.025
0.000
Frequency
3
4
5
Serum Albumin (mg/dL)

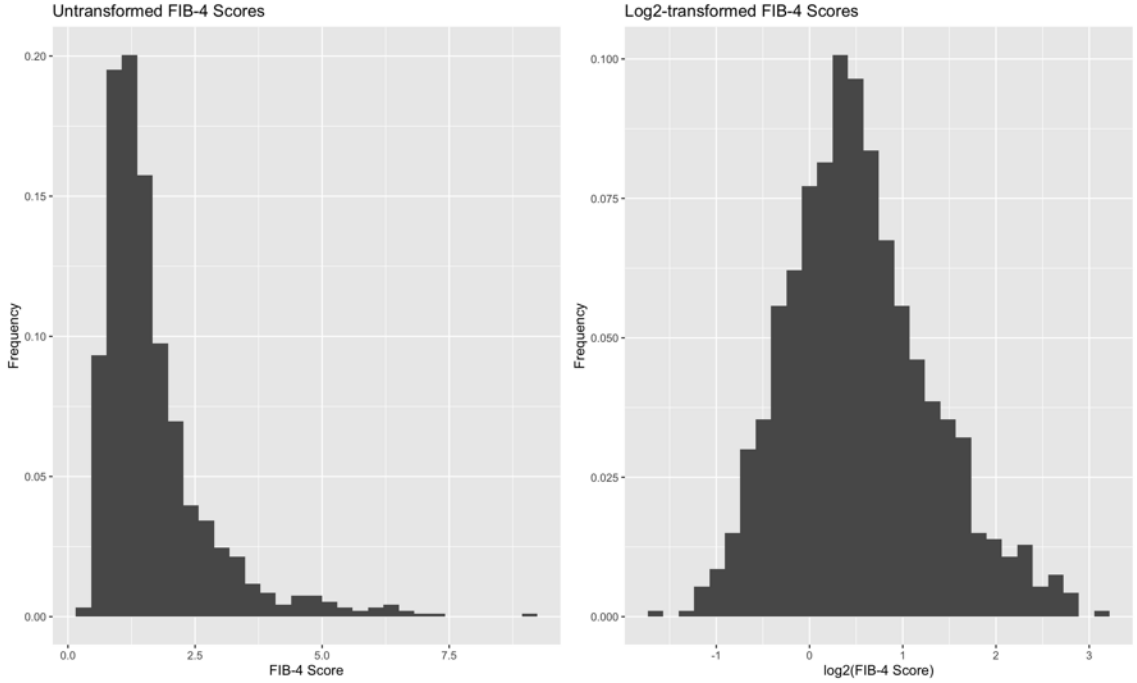
D


Untransformed
0.15
0.10
0.05
0.00
Frequency
0
1
2
3
Total Serum Bilirubin (mg/dL)

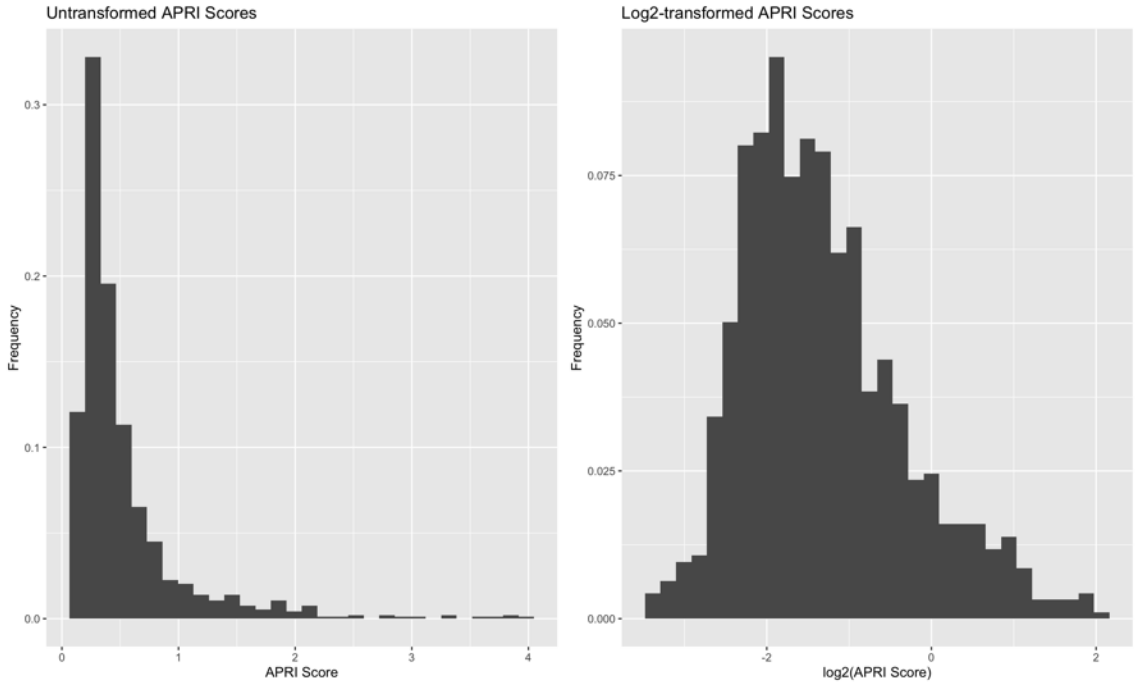
E


Untransformed Platelet Counts
Frequency
0.09
0.06
0.03
0.00
0
100
200
300
400
Platelet Count (*10^9/L)

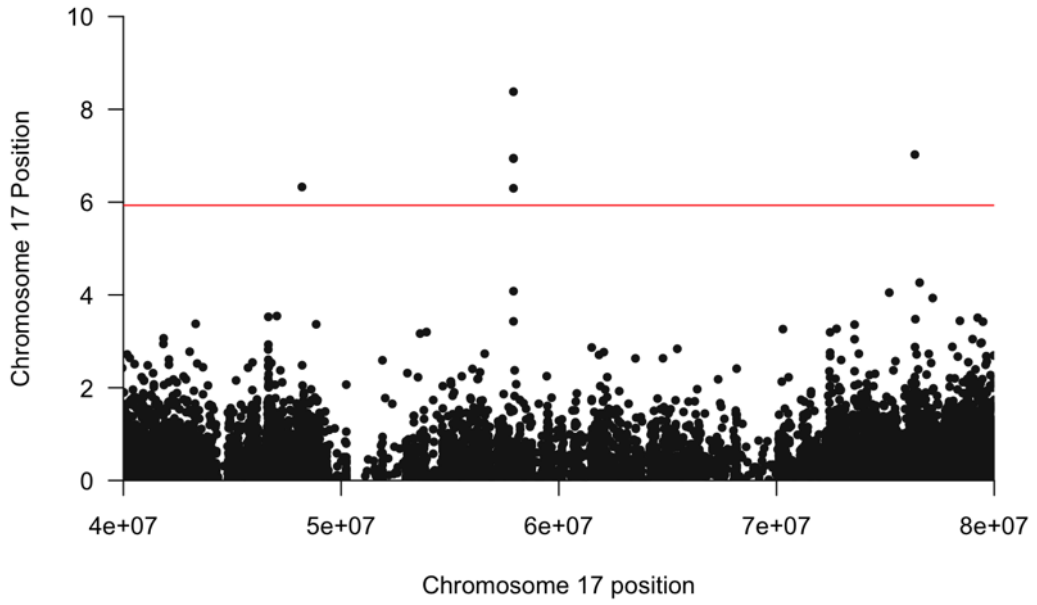
F


Untransformed FIB-4 Scores
Frequency
0.20
0.15
0.10
0.05
0.00
0.0
2.5
5.0
7.5
FIB-4 Score
Log2-transformed FIB-4 Scores
Frequency
0.100
0.075
0.050
0.025
0.000
-1
0
1
2
3
log2(FIB-4 Score)

G

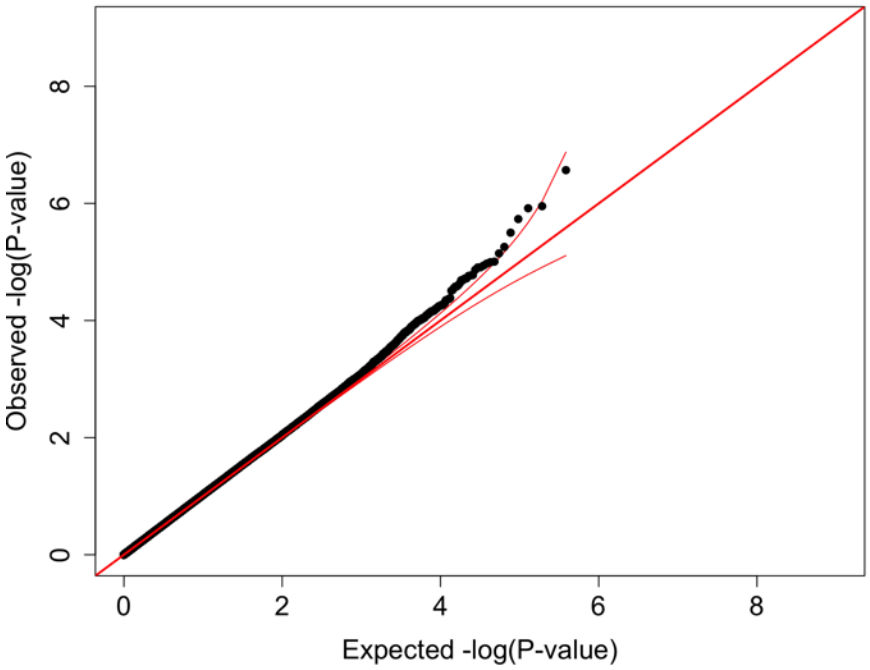


Supplementary Figure 1: Distributions of values for selected biomarkers of liver function among the entire cohort. Only biomarkers with highly right-skewed distributions were log-transformed to produce more normal distributions that allow for inference from linear modelling. For this figure, log2 transformation was used for easy interpretation. For EWAS, a natural log transformation was applied when needed. **A)** Distribution of AST before and after log transformation. **B)** Distribution of ALT before and after log transformation. **C)** Distributions of ALB without log transformation. **D)** Distributions of ALB without log transformation. **E)** Distributions of ALB without log transformation. **F)** Distribution of FIB-4 scores before and after log transformation. **G)** Distribution of APRI scores before and after log transformation. Abbreviations: AST, aspartate aminotransferase level (units/L); ALT, alanine aminotransferase level (units/L); ALB, serum albumin level (mg/dL); TBILI, total bilirubin level (mg/dL); PLT, platelet count (cells/μL); FIB-4, FIB-4 score; APRI, APRI score.

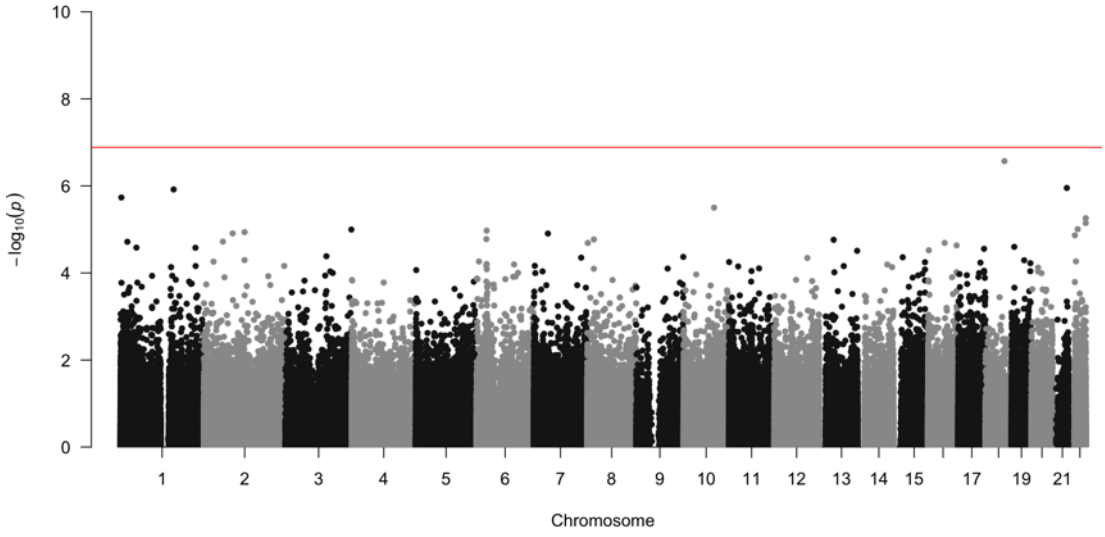


Supplementary Figure 2: **Regional manhattan plot of unadjusted *P*-values from meta-analysis of EWAS results for CpG sites in chromosome 17 and serum albumin**. The red line indicates the unadjusted *P*-value that corresponds to a threshold for FDR significance (among all CpG sites from across entire genome) at Q < 0.05. DNA methylation was not significantly associated with serum for any CpG site after a more restrictive Bonferroni correction to an overall $\alpha = 0.05$.

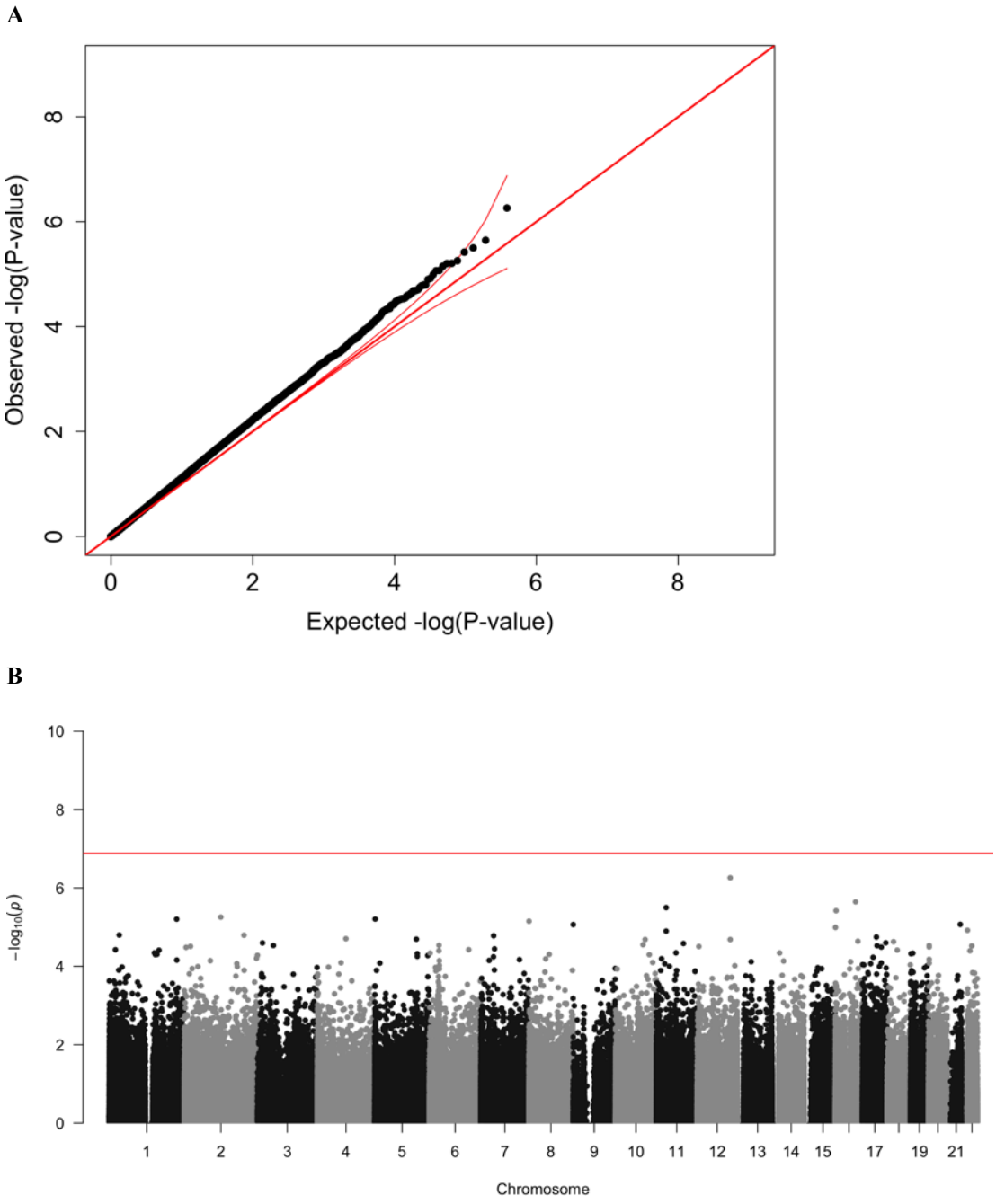
A



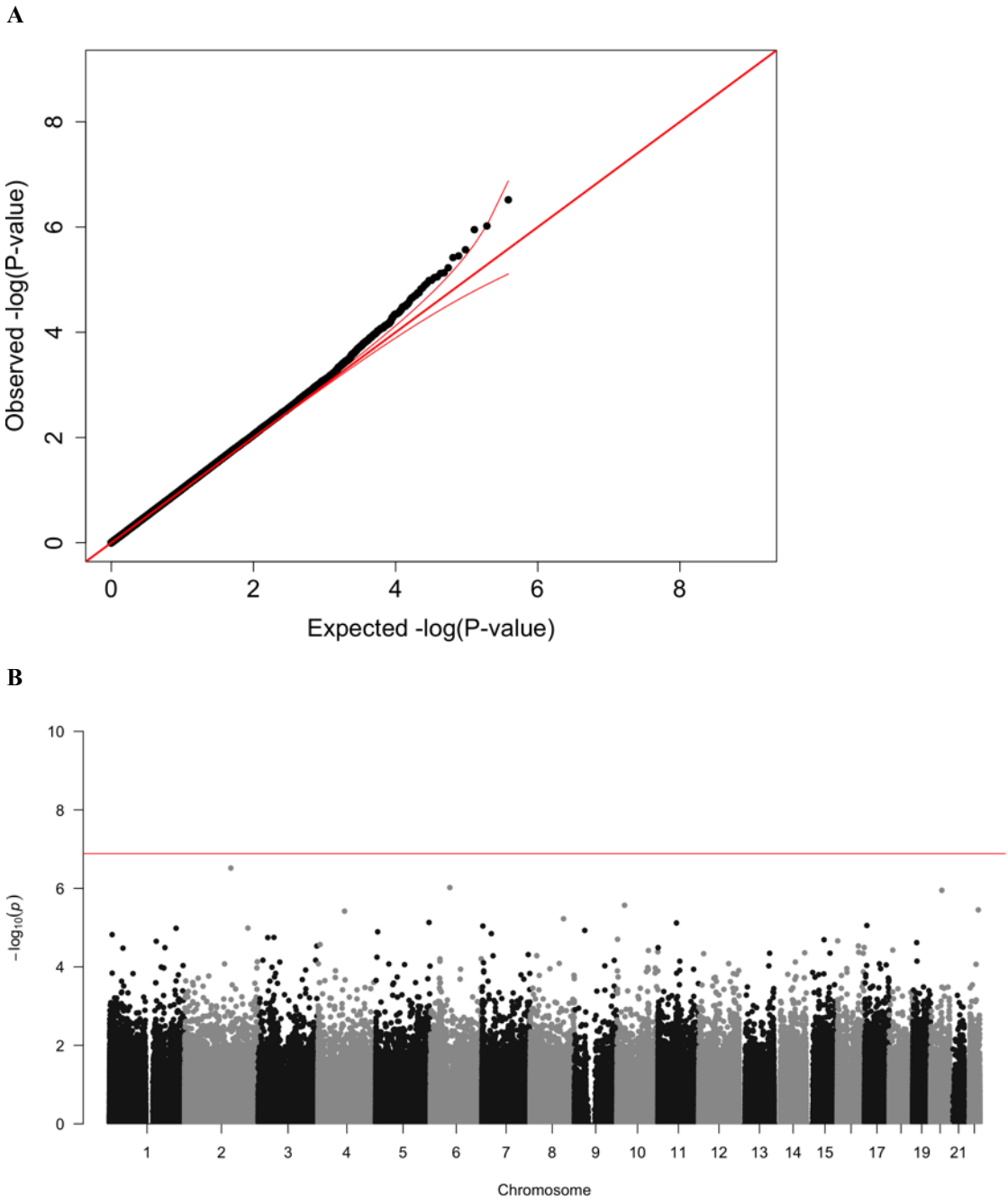
B



Supplementary Figure 3: **Distribution of unadjusted *P*-values from meta-analysis of EWAS results from the EPIC and 450K cohorts for AST.** (**A**) Quantile-quantile plot of unadjusted *P*-values from meta-analysis of EWAS results for AST. The global inflation factor was 1.04, so no further adjustment for global inflation was performed. Red lines represent a perfect 1:1 association with a 95% confidence interval. **(B)** Manhattan plot of unadjusted *P*-values from meta-analysis of EWAS results for AST. The red line indicates the unadjusted *P*-value that corresponds to a threshold for FDR significance at $Q < 0.05$. DNA methylation was not significantly associated with AST for any CpG site after a more restrictive Bonferroni correction to an overall $\alpha = 0.05$.

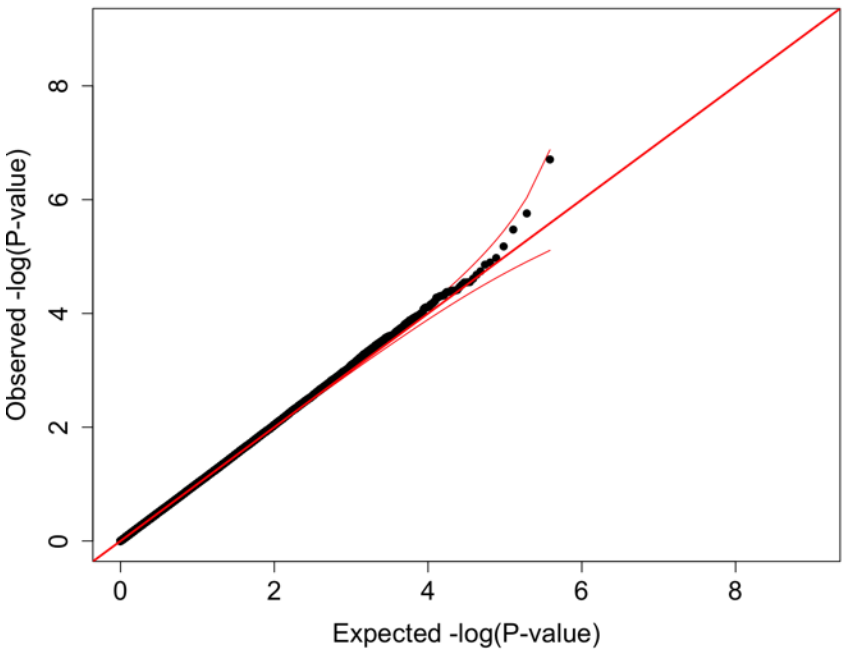


Supplementary Figure 4: **Distribution of unadjusted *P*-values from meta-analysis of EWAS results from the EPIC and 450K cohorts for ALT.** (**A**) Quantile-quantile plot of unadjusted *P*-values from meta-analysis of EWAS results for ALT. The global inflation factor was 1.14, so no further adjustment for global inflation was performed. Red lines represent a perfect 1:1 association with a 95% confidence interval. **(B)** Manhattan plot of unadjusted *P*-values from meta-analysis of EWAS results for ALT. The red line indicates the unadjusted *P*-value that corresponds to a threshold for FDR significance at $Q < 0.05$. DNA methylation was not significantly associated with ALT for any CpG site after a more restrictive Bonferroni correction to an overall $\alpha = 0.05$.

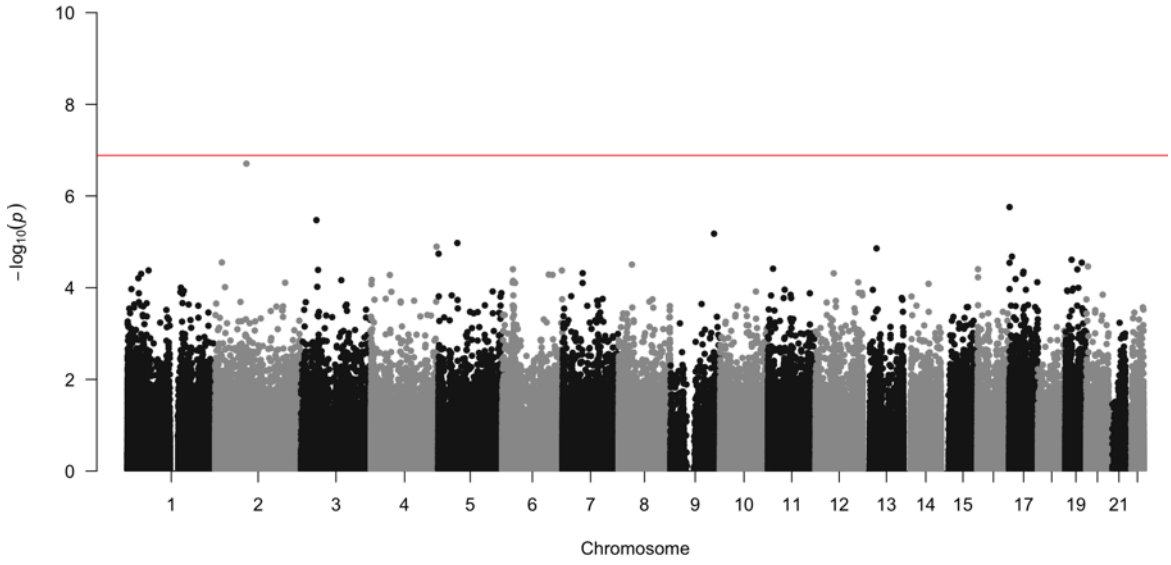


Supplementary Figure 5: **Distribution of unadjusted *P*-values from meta-analysis of EWAS results from the EPIC and 450K cohorts for total bilirubin.** (**A**) Quantile-quantile plot of unadjusted *P*-values from meta-analysis of EWAS results for total bilirubin. The global inflation factor was 1.04, so no further adjustment for global inflation was performed. Red lines represent a perfect 1:1 association with a 95% confidence interval. **(B)** Manhattan plot of unadjusted *P*-values from meta-analysis of EWAS results for total bilirubin. The red line indicates the unadjusted *P*-value that corresponds to a threshold for FDR significance at Q < 0.05. DNA methylation was not significantly associated with total bilirubin for any CpG site after a more restrictive Bonferroni correction to an overall $\alpha = 0.05$.

A

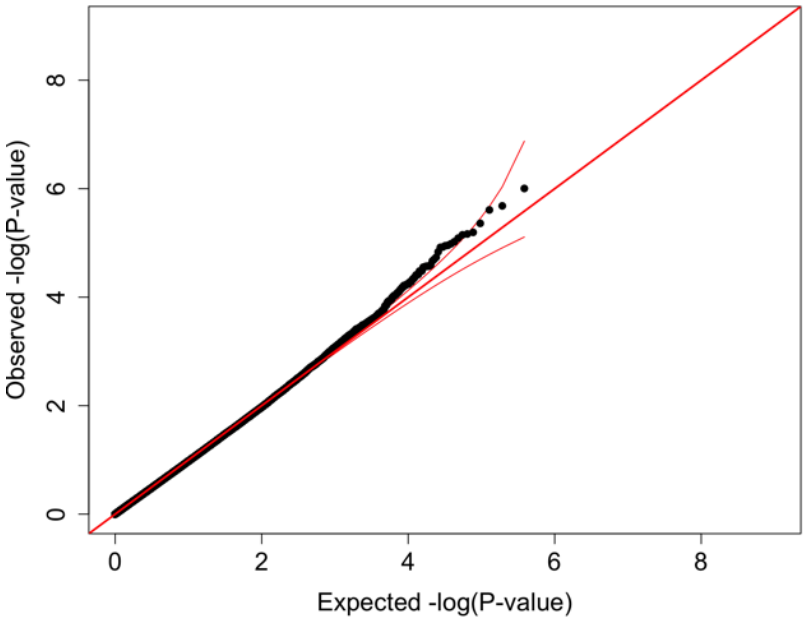


B

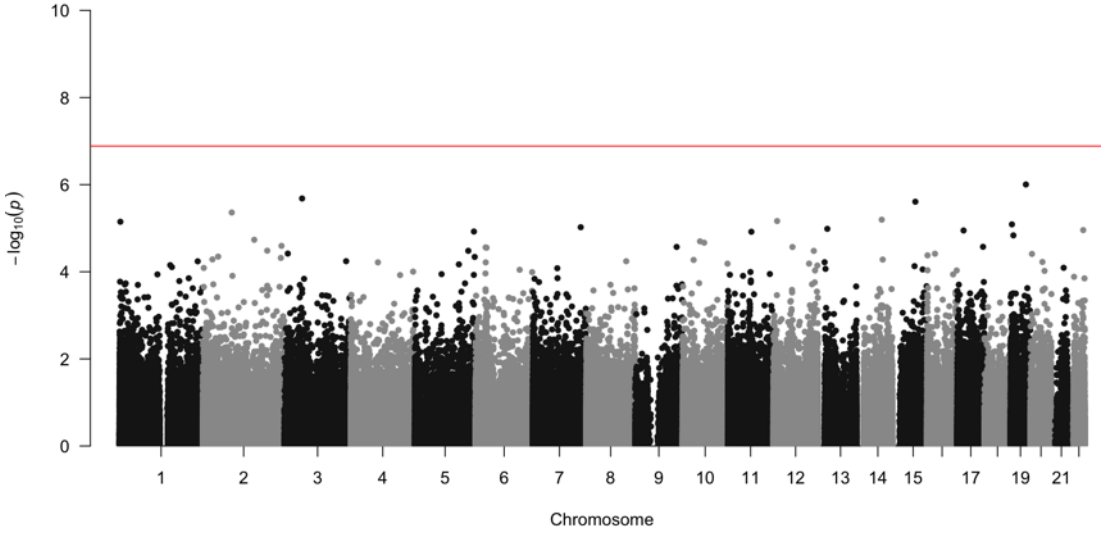


Supplementary Figure 6: **Distribution of unadjusted *P*-values from meta-analysis of EWAS results from the EPIC and 450K cohorts for platelet count.** (**A**) Quantile-quantile plot of unadjusted *P*-values from meta-analysis of EWAS results for platelet count. The global inflation factor was 1.02, so no further adjustment for global inflation was performed. Red lines represent a perfect 1:1 association with a 95% confidence interval. **(B)** Manhattan plot of unadjusted *P*-values from meta-analysis of EWAS results for platelet count. The red line indicates the unadjusted *P*-value that corresponds to a threshold for FDR significance at Q < 0.05. DNA methylation was not significantly associated with platelet count for any CpG site after a more restrictive Bonferroni correction to an overall $\alpha = 0.05$.

A

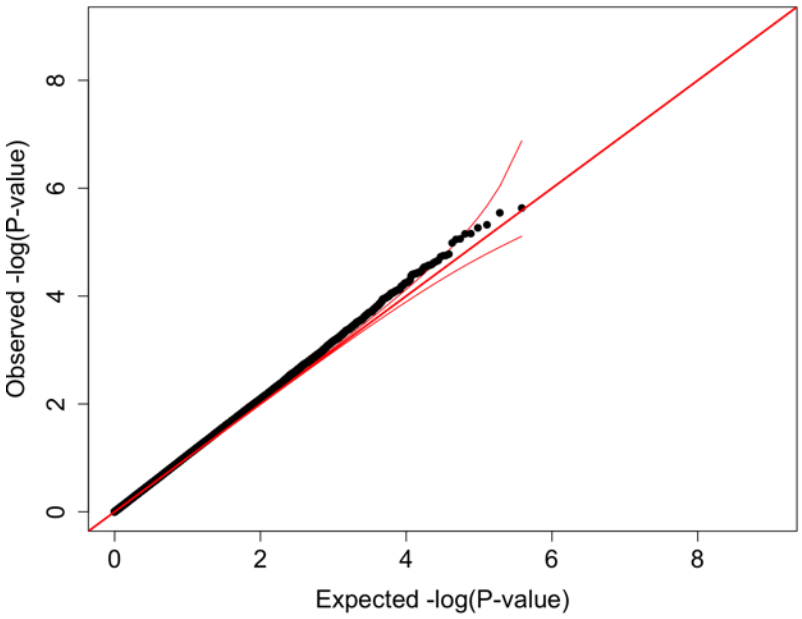


B

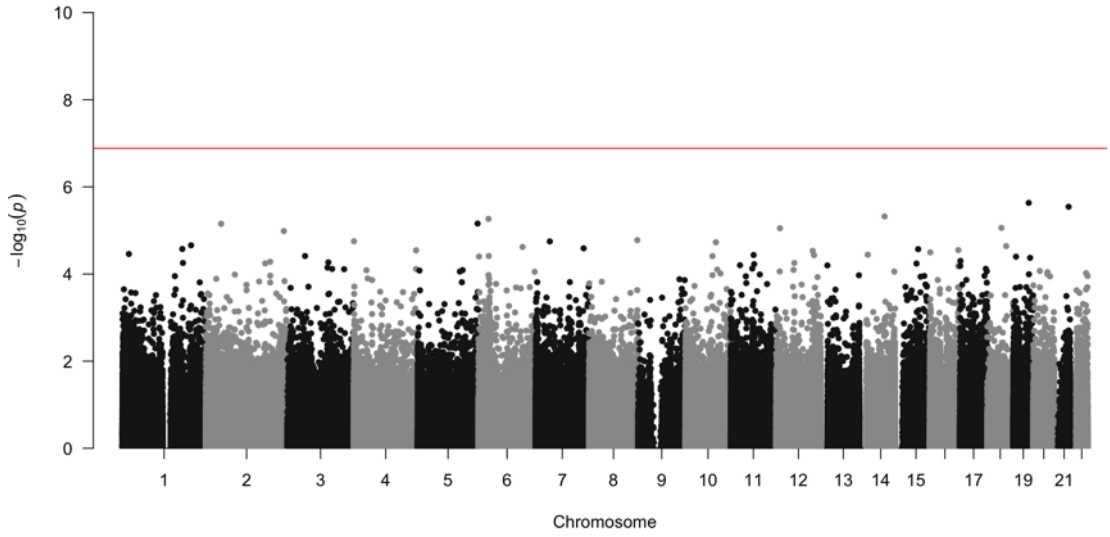


Supplementary Figure 7: **Distribution of unadjusted *P*-values from meta-analysis of EWAS results from the EPIC and 450K cohorts for FIB-4 score.** (**A**) Quantile-quantile plot of unadjusted *P*-values from meta-analysis of EWAS results for FIB-4 score. The global inflation factor was 0.96, so no further adjustment for global inflation was performed. Red lines represent a perfect 1:1 association with a 95% confidence interval. **(B)** Manhattan plot of unadjusted *P*-values from meta-analysis of EWAS results for FIB-4 score. The red line indicates the unadjusted *P*-value that corresponds to a threshold for FDR significance at Q < 0.05. DNA methylation was not significantly associated with FIB-4 score for any CpG site after a more restrictive Bonferroni correction to an overall $\alpha = 0.05$.

A



B



Supplementary Figure 8: **Distribution of unadjusted *P*-values from meta-analysis of EWAS results from the EPIC and 450K cohorts for APRI score.** (**A**) Quantile-quantile plot of unadjusted *P*-values from meta-analysis of EWAS results for APRI score. The global inflation factor was 1.08, so no further adjustment for global inflation was performed. Red lines represent a perfect 1:1 association with a 95% confidence interval. **(B)** Manhattan plot of unadjusted *P*-values from meta-analysis of EWAS results for APRI score. The red line indicates the unadjusted *P*-value that corresponds to a threshold for FDR significance at Q < 0.05. DNA methylation was not significantly associated with APRI score for any CpG site after a more restrictive Bonferroni correction to an overall $\alpha = 0.05$.