Relationship Between Methylome and Transcriptome in Patients With Nonalcoholic Fatty Liver Disease

SUSAN K. MURPHY,1 HYUNA YANG,2 CYNTHIA A. MOYLAN,3 HERBERT PANG,2 ANDREW DELLINGER,2 MANAL F. ABDELMALEK,1 MELANIE E. GARRETT,2 ALLISON ASHLEY-KOCH,4 AYAKO SUZUKI,3 HANS L. TILLMANN,3 MICHAEL A. HAUSER,4 and ANNA MAE DIEHL3

Departments of 1Obstetrics and Gynecology; 2Bioinformatics and Biostatistics; 3Division of Gastroenterology and Hepatology; and 4Center for Human Genetics, Department of Medicine, Duke University, Durham, North Carolina

BACKGROUND & AIMS: Cirrhosis and liver cancer are potential outcomes of advanced nonalcoholic fatty liver disease (NAFLD). It is not clear what factors determine whether patients will develop advanced or mild NAFLD, limiting noninvasive diagnosis and treatment before clinical sequelae emerge. We investigated whether DNA methylation profiles can distinguish patients with mild disease from those with advanced NAFLD, and how these patterns are functionally related to hepatic gene expression. METHODS: We collected frozen liver biopsies and clinical data from patients with biopsy-proven NAFLD (56 in the discovery cohort and 34 in the replication cohort). Samples were divided into groups based on histologic severity of fibrosis: F0–1 (mild) and F3–4 (advanced). DNA methylation profiles were determined and coupled with gene expression data from the same biopsies; differential methylation was validated in subsets of the discovery and replication cohorts. We then analyzed interactions between the methylome and transcriptome. RESULTS: Clinical features did not differ between patients known to have mild or advanced fibrosis based on biopsy analysis. There were 69,247 differentially methylated CpG sites (76% hypomethylated, 24% hypermethylated) in patients with advanced vs mild NAFLD (P < .05). Methylation at fibroblast growth factor receptor 2, methionine adenosyl methyltransferase 1A, and caspase 1 was validated by bisulfite pyrosequencing and the findings were reproduced in the replication cohort. Methylation correlated with gene transcript levels for 7% of differentially methylated CpG sites, indicating that differential methylation contributes to differences in expression. In samples with advanced NAFLD, many tissue repair genes were hypomethylated and overexpressed, and genes in certain metabolic pathways, including 1-carbon metabolism, were hypermethylated and underexpressed. CONCLUSIONS: Functionally relevant differences in methylation can distinguish patients with advanced vs mild NAFLD. Altered methylation of genes that regulate processes such as steatohepatitis, fibrosis, and carcinogenesis indicate the role of DNA methylation in progression of NAFLD.

Keywords: NAFLD; DNA Methylation; Gene Expression; Microarrays.

Nonalcoholic fatty liver disease (NAFLD) is strongly associated with obesity, type 2 diabetes, and the metabolic syndrome. Like these other conditions, NAFLD is increasing in incidence and prevalence. It is now the most common cause of chronic liver disease in the United States and Western Europe.1 NAFLD encompasses a spectrum of liver pathology that is generally characterized by excessive accumulation of fat in hepatocytes (ie, steatosis). Some individuals with fatty hepatocytes also have co-incident hepatic inflammation and increased liver cell death (ie, nonalcoholic steatohepatitis [NASH]). The outcomes of steatosis and steatohepatitis are very different. Individuals with steatosis rarely develop liver fibrosis; they seldom progress to clinically significant liver disease and are considered to have mild NAFLD. In contrast, progressive liver fibrosis occurs in some individuals with NASH, significantly increasing their risk for liver cirrhosis, primary liver cancer, and resultant liver-related morbidity and mortality.1 Because NASH can trigger a fibrogenic repair process that eventually culminates in cirrhosis and/or cancer, it is considered to be part of the spectrum of advanced NAFLD.

Currently, liver biopsy is the only way to reliably stage the severity of liver fibrosis and thereby distinguish individuals with mild NAFLD from those with advanced NAFLD before overt clinical sequelae of liver damage emerge.1 This limits population-based screening, delaying diagnosis of individuals with NAFLD who are at high risk for eventual liver-related morbidity and mortality. Noninvasive biomarkers are needed. Once NASH has been diagnosed, interventions to prevent progression are
necessary, but must have a low risk to benefit ratio, given the high prevalence of NASH and the variable (but generally slow) rates of progression. Success in this endeavor mandates identification of risk factors for advanced NAFLD that are easily (and safely) modified.

DNA methylation is an epigenetic form of gene regulation that is generally associated with transcriptional repression when present at the promoter regions of genes. It works in concert with histone modifications to regulate the activity of genes, and these regulatory mechanisms help guide levels of gene transcription in all tissues. DNA methylation changes are known to modulate susceptibility to obesity, a major risk factor for NAFLD, as well as the outcomes of diet-induced liver injury in rodents. Methyl-depleted diets promote steatohepatitis, cirrhosis, and liver cancer in rats and mice, while replenishing methyl stores avoids all of these outcomes. Changes in the “methylome” are plausible in humans with NAFLD, might differentiate those with mild NAFLD from those with more advanced NAFLD, and might provide novel therapeutic targets. Currently, however, almost nothing is known about the methylome in human NAFLD. To rectify this gap in knowledge, we generated comprehensive liver DNA methylation profiles in a large, well-characterized group of patients with either mild or advanced NAFLD, and correlated differences in liver DNA methylation with differences in liver gene expression (GEx). We discovered that there are numerous, functionally relevant methylation differences that distinguish mild from advanced NAFLD. These results are consistent with a prominent role for methylation in NAFLD progression in humans.

**Methods**

Detailed methods are available in the Supplementary Material.

**Study Samples**

The Duke University Health System NAFLD Clinical Database and Biorepository contains frozen liver biopsies and clinical data from patients with biopsy-proven NAFLD. The biorepository is approved by the Duke University Institutional Review Board. Demographic data and laboratory studies were obtained within 6 months of liver biopsy.

**Generation of Genomic Data**

Generation of GEx data for these samples has been described. The Illumina HumanMethylation450 beadchip platform was used for 33 mild (fibrosis stage, F0–F1) and 23 advanced (fibrosis stage, F3–F4) NAFLD specimens at Expression Analysis (Research Triangle Park, NC).

**DNA Methylation Data Preprocessing**

Data were processed in 1 batch (8 arrays with 12 samples per array). Samples were randomly distributed across the arrays. Principle component analysis was used to identify potential sample artifacts.

**DNA Methylation Data Analysis**

Modified t test statistics identified differentially methylated (DM) CpG sites between advanced and mild NAFLD. Analysis of variance model fitting was used to identify DM CpG sites between advanced and mild NAFLD after controlling for age and sex, with significance defined as q-value < .05. Data analyses were performed using R/Bioconductor statistical packages.

**Integration of DNA Methylation and Expression Data**

Generation of Affymetrix Human Genome U133 Plus 2.0 GeneChip platform (Affymetrix, Santa Clara, CA) expression data has been described (NCBI Accession GSE31803). GEx and methylation data were available for 45 of the 56 patients (27 mild and 18 advanced NAFLD). CpG island (CGI) and expression probe sets were coupled based on information in the University of California Santa Cruz genome browser (GRCh37/hg19) and correlations measured via the Spearman rank correlation coefficient.

**Bioinformatics Analysis**

The Ingenuity Pathways Analysis Tool (Ingenuity Systems, Inc., Redwood City, CA; www.ingenuity.com) was used to determine potential functional significance of methylation-expression relationships.

**Data Access**

Human Methylation450 beadchip data is available from the Gene Expression Omnibus web site (NCBI Accession GSE31803).

**Results**

**Patient and Sample Characteristics**

Genome-scale DNA methylation profiles were obtained for liver tissue of 56 patients, including 33 with mild and 23 with advanced NAFLD. Clinical and laboratory characteristics of patients with advanced NAFLD were not significantly different from patients with mild NAFLD in both the discovery and replication cohorts (Table 1). Histologic characteristics reflecting NAFLD severity differed among advanced and mild NAFLD patients, however, with advanced NAFLD patients having significantly more portal inflammation and ballooning, and therefore significantly higher NAFLD activity scores than those with mild NAFLD (Table 1). These findings confirmed that fibrosis stage provided a reliable surrogate for NAFLD disease severity.

All samples were processed in one batch, but we detected an apparent within-batch positional effect. Signal intensities differed for 3 of the arrays (24 samples) based on principal components analysis (Supplementary Figure 1). Samples were intentionally randomized across positions within the array, such that no one group was solely influenced by this difference. We used a 2-step quantile normalization approach to reduce the artificial noise from this positional effect.

**DNA Methylation Varies Across the Genome in NAFLD**

To determine if NAFLD status influences the distribution of DNA methylation, we first examined methylation profiles according to gene structure.
Table 1. Characteristics of the NAFLD Methylation Discovery and Replication Cohorts

<table>
<thead>
<tr>
<th>Patient demographics and clinical characteristics</th>
<th>Discovery cohort</th>
<th>Replication cohort</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mild NAFLD F0-F1</td>
<td>Advanced NAFLD F3-4</td>
<td>Mild NAFLD F0-F1</td>
</tr>
<tr>
<td>Male sex, n (%)</td>
<td>12 (36.4)</td>
<td>5 (21.7)</td>
<td>10 (52.6)</td>
</tr>
<tr>
<td>Age in years at biopsy, mean (SD)</td>
<td>51.5 (10.3)</td>
<td>51.7 (10.3)</td>
<td>50.4 (3.2)</td>
</tr>
<tr>
<td>Ethnicity, n (%)</td>
<td></td>
<td></td>
<td>17 (85.9)</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>33 (100)</td>
<td>21 (91.3)</td>
<td>17 (85.9)</td>
</tr>
<tr>
<td>Unknown</td>
<td>0 (0)</td>
<td>2 (8.7)</td>
<td>2 (10.5)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>32 (97)</td>
<td>21 (91.3)</td>
<td>15 (78.9)</td>
</tr>
<tr>
<td>Black</td>
<td>1 (3)</td>
<td>2 (8.7)</td>
<td>4 (21.1)</td>
</tr>
<tr>
<td>Body mass index, median (IQR)</td>
<td>32.8 (28.4–40.4)</td>
<td>33.8 (31.3–41.9)</td>
<td>32 (30.5–33.6)</td>
</tr>
<tr>
<td>Hypertension, n (%)</td>
<td>14 (42.4)</td>
<td>15 (65.2)</td>
<td>13 (68.4)</td>
</tr>
<tr>
<td>Diabetes mellitus, n (%)</td>
<td>6 (18.2)</td>
<td>9 (39.1)</td>
<td>6 (31.6)</td>
</tr>
<tr>
<td>Hemoglobin A1c, %, median (IQR)</td>
<td>5.8 (5.5–6.4)</td>
<td>6 (5.8–6.9)</td>
<td>5.8 (5.6–6.2)</td>
</tr>
<tr>
<td>Hyperlipidemia, n (%)</td>
<td>7 (21.2)</td>
<td>4 (17.4)</td>
<td>10 (52.6)</td>
</tr>
<tr>
<td>Current smoking, n (%)</td>
<td>6 (18.2)</td>
<td>5 (22.7)</td>
<td>1 (2.0)</td>
</tr>
<tr>
<td>Laboratory measures, median (IQR)</td>
<td>42 (24–57)</td>
<td>62 (46–85)</td>
<td>38 (30–58)</td>
</tr>
<tr>
<td>Serum ALT, U/Li</td>
<td>47.5 (30–75)</td>
<td>74.5 (46–101)</td>
<td>56 (38–99)</td>
</tr>
<tr>
<td>AST/ALT, U/Li</td>
<td>.8 (.64–1.03)</td>
<td>.92 (.66–1.24)</td>
<td>.72 (.59–.79)</td>
</tr>
<tr>
<td>Histologic characteristics, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steatosis (% &gt;34%)</td>
<td>14 (45.2)</td>
<td>14 (60.9)</td>
<td>10 (52.6)</td>
</tr>
<tr>
<td>Lobular inflammation (% &gt;grade 2)</td>
<td>6 (19.4)</td>
<td>9 (42.9)</td>
<td>1 (5.3)</td>
</tr>
<tr>
<td>Portal inflammation (% &gt;mild)</td>
<td>3 (10.7)</td>
<td>12 (54.6)</td>
<td>6 (31.6)</td>
</tr>
<tr>
<td>Ballooning (% any)</td>
<td>17 (56.7)</td>
<td>22 (95.7)</td>
<td>16 (84.2)</td>
</tr>
<tr>
<td>NAFLD activity score (% ≥5)</td>
<td>7 (23.3)</td>
<td>12 (57.1)</td>
<td>5 (26.3)</td>
</tr>
</tbody>
</table>

ALT, alanine aminotransferase; AST, aspartate aminotransferase; IQR, interquartile range.

aMissing data for 2 patients in discovery cohort.
bMissing data for 12 patients in discovery cohort and 2 in replication cohort.
cMissing data for 9 patients.
dMissing data for 20 patients.
eMissing data for 4 patients in discovery cohort.
fMissing data for 6 patients in discovery cohort.
gMissing data for 3 patients in discovery cohort.
hMissing data for 5 patients in discovery cohort.
iNAFLD Activity Score (range 0–8) is a sum of scores for steatosis, lobular inflammation, and ballooning.
Consistent with the Encyclopedia of DNA Elements data showing that the vast majority of CGI and transcription start sites (TSS) are normally unmethylated, we observed overall hypomethylation at TSS and most CGI in NAFLD tissue (Figure 1A). Autosomal and sex chromosome methylation also differed, with relative hypermethylation of CGIs on the X chromosomes, likely due to X-inactivation in females (64% and 78% of the mild and advanced NAFLD patients, respectively). This pattern was similar for both mild and advanced NAFLD.

Differential methylation in cancer is reported as most evident at CGI “shores” defined as regions located up to 2 kb upstream or downstream from a CGI. When we compared methylation profiles in mild to advanced NAFLD, we also found differential methylation was more pronounced at regions flanking TSSs or CGIs (Figure 1B). Because many CGIs are located at the promoter regions of genes, very close to the TSS, it was not clear whether these findings were due to proximity to the CGI or the TSS. We restricted the analysis to CGI and TSS positioned at least

**Figure 1.** Methylation patterns in NAFLD relative to TSS and CGI. (A) Median methylation for autosomes (solid lines) and the X chromosome in liver tissues from female NAFLD patients as a function of distance from TSSs (left) or from CGIs (right). (B) Probes showing significant differential methylation between mild and advanced NAFLD as a function of distance from TSSs and CGIs, including non-promoter and promoter CGI. Blue lines, $P < .05$; red line, $q < .05$. (C) Median methylation for non-CGI probes with significant differential methylation between mild and advanced NAFLD relative to the distance from TSSs and promoter CGIs. Blue lines, $P < .05$; red line, $q < .05$. 

---

**Figure 1A:**
- **A**
  - Median methylation for autosomes (solid lines) and the X chromosome in liver tissues from female NAFLD patients as a function of distance from TSSs (left) or from CGIs (right).
  - **Auto-mild**
  - **Auto-advanced**
  - **ChrX-mild**
  - **ChrX-advanced**

- **B**
  - Percent of methylated beads
  - Blue lines, $P < .05$; red line, $q < .05$

- **C**
  - Median methylation for non-CGI probes with significant differential methylation between mild and advanced NAFLD relative to the distance from TSSs and promoter CGIs.
  - Blue lines, $P < .05$; red line, $q < .05$. 

---

**Figure 1B:**
- **B**
  - Probes showing significant differential methylation between mild and advanced NAFLD as a function of distance from TSSs and CGIs, including non-promoter and promoter CGI.
  - Blue lines, $P < .05$; red line, $q < .05$.

**Figure 1C:**
- **C**
  - Median methylation for non-CGI probes with significant differential methylation between mild and advanced NAFLD relative to the distance from TSSs and promoter CGIs.
  - Blue lines, $P < .05$; red line, $q < .05$. 

1000 bp apart. Using this reduced dataset, we observed increased differential methylation between mild and advanced NAFLD at CGI shores rather than TSSs (Figure 1C).

The Infinium HumanMethylation450 BeadChip data is obtained using 2 distinct chemical assays (Infinium I and Infinium II probes) with divergent intensity distributions16 (Supplementary Figure 2). We measured the proportion of differential methylation using the type I or type II probes independently, and reached the same conclusion regardless of the assay type (Supplementary Figure 3).

CpG Methylation Differs Between Mild and Advanced NAFLD

A regression model was used to test for differential methylation as the outcome due to NAFLD status after controlling for sex and age. The model identified 69,247 (14.3% of the total CpG sites) DM CpG sites between advanced and mild NAFLD using a q-value threshold of .05. Among the DM CpG sites, advanced NAFLD showed relative hypomethylation at 52,830 CpG sites (11% of total), and hypermethylation at 16,417 sites (3% of total). These results suggest that advanced NAFLD has a greater loss of gene regulatory capacity compared with mild NAFLD.

Functional Methylation-Expression Relationships Identified in NAFLD

There were distinct methylation profiles in NAFLD with regard to genomic position annotations, including intergenic, promoter, 5’ and 3’ untranslated region (UTR), and coding region CpG sites (Figure 2). For example, promoters showed increased hypomethylated sites and 3’UTRs showed increased hypermethylated sites. DNA methylation at gene promoter regions is often associated with transcriptional repression due to interactions between DNA methylation, methyl DNA binding proteins, and histone deacetyltransferases, which contribute to the
formation of heterochromatin. In contrast, promoter hypomethylation is often associated with a euchromatic state and transcriptional permissiveness. We used our pre-existing GEx data for a subset of the NAFLD liver tissues to investigate how DNA methylation might affect GEx in this disease. For each transcript with a defined promoter, we took the median methylation intensities of all the CpG sites located within the region 500 bp upstream to 25 bp downstream of the annotated TSS, and plotted this against the mean transcription levels of the corresponding mRNA sequences. The relationship did not quite reach significance. The methylation levels on the 2 platforms for CASP1 differed substantially, with pyrosequencing showing <6% and the Infinium platform showing >20% methylation for the same tissues (Figure 3F).

We validated methylation findings in an independent group of liver specimens classified as having mild (n = 19) or advanced (n = 15) NAFLD. Normal liver tissues (n = 25) were included for comparison. We again found significant differences between mild and advanced NAFLD for FGFR2 (P = .002) and CASP1 (P = .024), and MAT1A approached significance (P = .088) (Figure 4A–C). Significant differences in the same direction for CASP1 supports the validity of the original findings. Methylation levels in normal livers compared with mild NAFLD showed no differences, and they differed markedly from advanced NAFLD (FGFR2: P < .0001; MAT1A: P = .008; CASP1: P = .001).

**Livers With Advanced NAFLD Show Increased Expression of Hypomethylated Genes**

Next we coupled GEx with methylation status of CpG sites, and identified genes that were both differentially methylated and expressed, and determined if the methylation and expression profiles for these genes went in the expected direction. To do this, we matched each CpG site to its corresponding GEx probe set and identified 76,833 unique CpG–gene expression pairs. We calculated the Spearman rank correlation between methylation of each CpG and the corresponding gene's transcription level, and found 4849 CpG sites with r < −.3 (corresponding to a P value of .02), signifying a significant inverse relationship between the GEx and CpG site methylation. These 4849 sites include hypermethylated down-regulated genes as well as hypomethylated up-regulated genes.

Associations between expression and methylation for which only 1 or 2 CpG sites exhibit a correlative relationship might be due to chance. We therefore restricted focus to genes showing a significant inverse expression–methylation relationship with a unidirectional difference in methylation for ≥3 CpG sites. This conservative approach also protects against the possibility that simple nucleotide polymorphism within, or near to, the methylation probes might solely influence findings. One hundred and forty-three genes were hypermethylated and 365 genes were hypomethylated at ≥3 CpG sites with methylation levels showing a significant inverse relationship with expression in the NAFLD cases (total of 2637 probes) (Supplementary Table 1).

**Many Tissue Repair Genes Are Hypomethylated and Overexpressed in Advanced NAFLD**

Often a group of genes rather than a single gene regulates a phenotype, and network-based approaches have been used to identify sets of functionally related...
We used the Ingenuity Pathways Analysis tool to explore the functions and pathways enriched within the groups of hypomethylated and hypermethylated genes in advanced vs mild NAFLD. Genes with numerous CpG sites showing differential methylation between mild and advanced NAFLD included transmembrane protein 204 (24 CpG sites hypomethylated and up-regulated in advanced NAFLD). This gene encodes a hypoxia-related adherens junction protein involved in cell adhesion and angiogenesis and is aberrantly methylated in pancreatic cancer.

FGFR2 is also hypomethylated (23 CpG sites) with increased expression in advanced NAFLD. FGFR2 functions as a receptor for keratinocyte growth factor, a protein made in chronic liver disease by stellate cells. Other genes encoding matrix molecules and matrix remodeling factors were also hypomethylated with increased expression in advanced vs mild NAFLD, including those associated with fibrosis and cirrhosis (COL1A1, COL1A2, COL4A1, COL4A2, LAMA4, LAMB1, CTGF, and PDGFA), several chemokines (CCR7 and CCL5), and factors related to the inflammasome and pro-inflammatory immune response (STAT1, TNFAIP8, and CASP1).

"Cancer" was the most significant biological function category in pathway analysis with an over-representation of hypomethylated, up-regulated genes in advanced vs mild NAFLD. Top functional annotations included neoplasia, tumorigenesis, and gastrointestinal tract cancer. Other relevant categories were "cellular movement" and "cell death." Significant canonical pathways included "role of tissue factor in cancer" and "hepatic fibrosis/hepatic stellate cell activation."
included “cellular movement, cancer and cell morphology” (score = 44), "cardiovascular system development and function, organ ischemial development and cellular movement” (score = 37), and “lipid metabolism, molecular transport, small molecule biochemistry” (score = 36). Many of these genes, functions and pathways bear relevance to the biology underlying liver fibrosis, providing confidence that this approach has revealed logical methylation-expression relationships in this disease, while also supporting the relevance for other findings that might not be as evident based on what is known about liver fibrosis. The complete lists of functions, canonical pathways, and networks along with the corresponding hypomethylated and transcriptionally activated genes are provided in Supplementary Tables 2–4.

### Certain Metabolic Genes Are Hypermethylated and Down-Regulated in Advanced NAFLD

Biological function categories enriched for the group of hypermethylated, transcriptionally repressed genes included “small molecule biochemistry” (uptake of thyroid hormone, 1.75e-05; oxidation of lipids; 2.85e-05). Canonical pathways enriched for hypermethylated genes were “LPS/IL-1 mediated inhibition of RXR function” (1.32e-06) and “fatty acid metabolism” (2.16e-05). Retinoid X receptors are proteins important in the transmission of signals that regulate cholesterol metabolism, and a high-cholesterol, high-fat diet recapitulates many features of nonalcoholic steatohepatitis in rodent models. Specific hypermethylated and repressed genes within this pathway included APOC4 (lipid metabolism), CYP2C19 (cytochrome P450 family; hepatocellular carcinoma), SLCO1B3 (membrane transporter and role in apoptosis; liver cancer, known target of methylation mediated silencing, including HepG2 cells), SLC10A1 (membrane transport, role in liver cancer; demonstrated to be unmethylated in normal mouse liver tissues), ABCC2 (multidrug resistance, liver neoplasia), ALDH (cell migration, cancer stem cells self-renewal; targeted by methylation in rat hepatoma cells), MGMT (apoptosis, DNA damage response, known target of hypermethylation in liver cancer), ACOX (steatohepatitis, microvesicular hepatic steatosis), and ACSL5 (fatty acid anabolic pathways, colon cancer).

Genes involved in histidine, selenoamino acid, tryptophan, tyrosine, and B-alanine metabolism were also repressed by hypermethylation (all P < 0.004) as were genes involved in the “urea cycle and metabolism of amino groups” (P = 0.001). Networks that contain significant numbers of hypermethylated and repressed genes included “lipid metabolism, small molecule biochemistry, cellular compromise” (score = 44) and “lipid metabolism, small molecule biochemistry, molecular transport” (score = 42). The complete lists of functions, canonical pathways, and networks along with the corresponding hypermethylated and repressed genes are provided in Supplementary Tables 5–7.

### Differential Methylation and Expression of Genes That Regulate Methylation Distinguishes Advanced From Mild NAFLD

Of particular interest is the differential methylation and expression of genes involved in one-carbon metabolism; the pathway that generates methyl groups for methylation reactions. The genes include methylenetetrahydrofolate dehydrogenase 2 (MTHFD2; hypomethylated in advanced NAFLD), adenosyl homocysteine (AHCY; hypermethylated), MAT1A (hypermethylated), and aldehyde dehydrogenase 1 family, member L1 (ALDH1L1; hypermethylated) (Figure 5). MTHFD2 functions in the mitochondria to generate formate, which enters the cytoplasm and is incorporated into tetrahydrofolates for one carbon metabolism. Although MTHFD2 has not been implicated in fibrosis, its protein product promotes cell proliferation. Expression of MTHFD2 has previously only been detected in transformed cells or during embryonic development. A role for AHcy in NAFLD has also not been reported, but Ahcy down-regulation with increased methylation was documented in mice fed a methyl-deficient diet, conditions that lead to NASH.

The MAT1A protein generates S-adenosylmethionine, the universal methyl group donor. Mat1a deletion causes NAFLD in mice and MAT1A is a target of promoter...
Two of the three MAT1A CpG sites identified here are located upstream of the TSS, and the third is approximately 200 bp downstream from the TSS. This is consistent with the report that coding region methylation (+88 from the TSS) is strongly associated with transcriptional regulation of MAT1A in human liver cells.36

Aldehyde dehydrogenase 1 family, member L1 encodes 10-formyltetrahydrofolate dehydrogenase, an enzyme involved in folate metabolism and the one carbon pathway for generation of methyl group donors.37 This gene is also a known target of methylation in liver cells, being unmethylated in normal hepatocytes but highly methylated in hepatocellular carcinoma cells38 with decreased expression an independent predictor of overall survival.39

**Discussion**

We report the first integrated analysis of genome-wide methylation and GEx data of diseased but noncancerous human livers. Our results were generated from percutaneous liver biopsy samples of well-characterized patients with NAFLD, one of the most common types of chronic liver disease in adults.1 Liver histology, the current gold standard diagnostic test,1 was used to segregate these subjects into 2 groups that are predicted to have extremely different liver outcomes at 10 years, based on differences in the severity of liver fibrosis at the time of tissue acquisition.2 As expected, the subgroup with little-to-no fibrosis had mild liver damage (mild NAFLD) and the other subgroup with more severe fibrosis had worse liver damage (advanced NAFLD), although the 2 histologically defined subgroups were clinically indistinguishable. We identified a large number of DM CpG sites between advanced and mild NAFLD, and independently confirmed methylation differences for FGFR2, MAT1A, and CASP1 with replication in an independent NAFLD cohort. FGFR2 was hypomethylated in advanced NAFLD and is involved in epithelial to mesenchymal transition.40,41

The study also revealed a large number of genes whose transcriptional status is significantly correlated with levels of DNA methylation, and showed that this relationship differs in mild vs advanced NAFLD. In addition, many of the differentially affected genes are inter-related not only by their roles in specific pathways and networks, but are highly relevant to the ultimate outcomes of liver injury, based on current understanding of cirrhosis and liver cancer pathogenesis.

Epigenetic processes are now widely recognized as playing prominent roles in many common diseases. Although histone modifications and microRNAs are critically important to regulating gene expression, DNA methylation is a more tractable modification for analysis because of its stability and availability of technologies that allow for relatively rapid assessment of large fractions of the methylome. Combining methylation data with GEx microarray data provides an extremely powerful tool to identify genes with expression levels that correlate with changes in DNA methylation and might be directly influenced by factors that skew normal epigenetic regulation. We used this approach to determine if epigenome-transcriptome interactions differ in individuals with mild NAFLD vs those with advanced NAFLD.
We found that livers with advanced NAFLD are generally hypomethylated relative to those with mild NAFLD. For example, of the genes showing significant inverse relationships between methylation and expression and with at least 3 DM CpG sites, there were more than twice as many genes that were hypomethylated with increased expression in advanced NAFLD relative to mild NAFLD. Hypomethylation is also a characteristic of alcohol-related fatty liver injury.42 Consistent with these results, an earlier study of mice fed a methyl-deficient diet to induce steatohepatitis also demonstrated relative hepatic hypomethylation. That study also noted significantly increased methylation and decreased expression of \textit{Ahcy} in mice with steatohepatitis,33 as was observed in the advanced human NAFLD cases evaluated here. In addition, we found that \textit{MAT1A}, a key gene responsible for generation of methyl groups used in DNA methylation reactions, was relatively hypermethylated and underexpressed in patients with advanced NAFLD. Knockdown of \textit{Mat1a} in mice enhances vulnerability to diet-induced fatty liver, resulting in steatohepatitis, altered lipid and carbohydrate metabolism, and increased incidence of hepatocellular carcinoma.7 Like \textit{MAT1A}, other genes that regulate 1-carbon metabolism were DM and expressed in human NAFLD, with the resultant methylome-transcriptome profiles predicting reduced net efficiency of methylation in advanced NAFLD.

Consistent with that concept, genes in human livers with advanced NAFLD were generally hypomethylated relative to those in livers with mild NAFLD.

Notably, many genes that modulate diverse wound healing responses were hypomethylated and up-regulated in advanced NAFLD, and pathway analysis confirmed that processes involved in carcinogenesis and fibrogenesis were induced. The cumulative data, therefore, suggest that methylome-transcriptome interactions that occur in advanced NAFLD can modulate NAFLD outcomes, including development of cancer and/or cirrhosis (Figure 6). Indeed, changes in DNA methylation at specific genes are well documented in cancer, and have a profound influence on cancer initiation and progression.8 Changes in methylation were recently reported to accompany steatohepatitis and precede emergence of hepatocellular carcinoma after infection with hepatitis C virus.43 Altered DNA methylation has also been reported in liver cirrhosis, primarily in animal models,9 where it is associated with activation of stellate cells and development of fibrosis.3

By identifying plausible methylation-sensitive mechanisms that influence NAFLD progression, our work might lead to development of noninvasive biomarkers that will help identify NAFLD patients at high risk for bad liver outcomes. For example, hypermethioninemia is characterized by abnormally elevated methionine levels in blood and...
has been reported in patients with alcohol-induced cirrhosis.\textsuperscript{44} Known genetic causes include deficiencies of MATIA, glycine N-methyltransferase (GNMT), and AHCY.\textsuperscript{4} It is of interest here that transcription of all three of these hypermethioninemia-causative genes was repressed in the advanced NAFLD patients and this coincided with increased methylation of their respective promoters. Epigenetic repression of transcription of these key genes in advanced NAFLD can mimic the genetic defects known to cause hypermethioninemia, providing a novel noninvasive biomarker for advanced NAFLD. Importantly, evidence for methylome-transcriptome interactions in human NAFLD has immediate therapeutic implications because a methyl donor-rich diet was able to reverse progression of fatty liver damage in rats fed a high-fat sucrose diet,\textsuperscript{45} and others have shown that methyl group donor-rich diets directly affect DNA methylation changes in animal models of obesity.\textsuperscript{46} Epigenetic modifications could provide malleable targets for future interventions that aim to prevent/reverse advanced NAFLD. We acknowledge that our approach cannot determine causality and that the results are largely descriptive, indicating only a general pattern given that neither the expression nor methylation platform used is an exhaustive representation. Also, future research is necessary to confirm that observed DM-related changes in gene expression translate into comparable differences in the respective liver proteins. The provocative data generated by this first-in-human genome-wide integrated methylome-transcriptome analysis of damaged livers provide compelling justification for further research that might alleviate an emerging burden to public health worldwide.

**Supplementary Material**

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://dx.doi.org/10.1053/j.gastro.2013.07.047.

**References**


Author names in bold designate shared co-first authorship.

Received January 22, 2013. Accepted July 26, 2013.

Reprint requests
Address requests for reprints to: Anna Mae Diehl, MD, Department of Gastroenterology, Duke University, Synderman Building (GSRB-1), 595 LaSalle Street, Suite 1073, Durham, North Carolina 27710. e-mail: annamae.diehl@duke.edu; fax: 919-684-4183.

Acknowledgments
The authors thank Daniel Hampton, Zhiqing Huang, and Carole Grenier for technical assistance.

Conflicts of interest
The authors disclose no conflicts.

Funding
Supported by the National Institutes of Health 5RC2-AA019399.