BASIC AND TRANSLATIONAL—LIVER

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,[3] MELANIE E. GARRETT,[4] ALLISON ASHLEY–KOCH,[4] AYAKO SUZUKI,[3] HANS L. TILLMANN,[3] MICHAEL A. HAUSER,[4] and ANNA MAE DIEHL[3]

Departments of [1]Obstetrics and Gynecology; [2]Bioinformatics and Biostatistics; [3]Division of Gastroenterology and Hepatology; and [4]Center 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.[2] 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

Abbreviations used in this paper: AHCY, adenosyl homocysteine; ALDH1L1, aldehyde dehydrogenase 1 family, member L1; BMI, body mass index; CGI, CpG islands; CASP1, caspase 1; DM, differentially methylated; FGFR2, fibroblast growth factor receptor 2; gene expression, (GEx); MAT1A, methionine adenosyl methyltransferase 1A; MTHFD, methylenetetrahydrofolate dehydrogenase 2; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; TSS, transcription start site; UTR, untranslated region.

0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2013.07.047

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,[3,4] 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,[5–8] while replenishing methyl stores avoids all of these outcomes.[9] 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.[10] 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.[11]

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[10] (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)[12] 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.