Published in final edited form as:

Hepatology. 2019 February ; 69(2): 639–652. doi:10.1002/hep.30211.

Integrating the Epigenome to Identify Novel Drivers of Hepatocellular Carcinoma

DR. RA Hlady¹, A Sathyanarayan¹, JJ Thompson¹, D Zhou¹, Q Wu², K Pham², DR. JH Lee³, C Liu², and KD Robertson¹

¹Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN 55905, USA

²Department of Pathology and Laboratory Medicine, New Jersey Medical School, Rutgers, The State University of New Jersey, Newark, NJ 07103, USA

³Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, 55905 Mayo Clinic, Rochester, MN 55905, USA

Abstract

Disruption of epigenetic mechanisms has been intimately linked to the etiology of human cancer. Understanding how these epigenetic mechanisms (including DNA methylation (5mC), hydroxymethylation (5hmC), and histone post-translational modifications) work in concert to drive cancer initiation and progression still remains unknown. Hepatocellular carcinoma (HCC) is increasing in frequency in western countries but lacks efficacious treatments. The epigenome of HCC remains understudied. To better understand the epigenetic underpinnings of HCC, we performed a genome-wide assessment of 5mC, 5hmC, four histone modifications linked to promoter/enhancer function (H3K4me1, H3K27ac, H3K4me3, and H3K27me3), and transcription across normal, cirrhotic, and HCC liver tissue. Implementation of bioinformatic strategies integrated these epigenetic marks with each other and with transcription to provide a comprehensive epigenetic profile of how and when the liver epigenome is perturbed during progression to HCC. Our data demonstrate significant deregulation of epigenetic regulators combined with disruptions in the epigenome hallmarked by profound loss of 5hmC, locus-specific gains in 5mC and 5hmC, and markedly altered histone modification profiles, particularly remodeling of enhancers. Data integration demonstrates that these marks collaborate to influence transcription (*e.g.* hyper-5hmC in HCC-gained active enhancers is linked to elevated expression) of genes regulating HCC proliferation. Two such putative epigenetic driver loci identified through our integrative approach, *COMT* and *FMO3*, increase apoptosis and decrease cell viability in liver-derived cancer cell lines when ectopically re-expressed.

Conclusion: Altogether, integration of multiple epigenetic parameters is a powerful tool for identifying epigenetically regulated drivers of HCC and elucidating how epigenome deregulation contributes to liver disease and HCC.

Keywords

cirrhosis; liver cancer; epigenome; ChIP-seq; RRBS; TET-assisted RRBS; DNA methylation; DNA hydroxymethylation; enhancer; bivalent; RNA-seq

Hepatocellular carcinoma (HCC), the most prevalent primary tumor of the liver, is projected to become the third leading cause of cancer mortality worldwide by 2020 (1). HCC almost always occurs in the background of liver cirrhosis. Cirrhosis and HCC are driven by several etiologic ‘environmental’ exposures including chronic alcohol abuse, hepatitis B and hepatitis C viral infection, and increasingly due to metabolic disorders from high fat western diets (2). Direct acting antivirals are effective in eradicating HCV, although hurdles to their widespread use due to surveillance, cost, and residual disease following treatment, remain. In addition, elimination of HCV infection or alcohol abuse does not completely eliminate risk for HCC in patients that have developed cirrhosis (3). Despite intensive research, the five-year survival rate for HCC has remained unchanged at <12% due to a lack of early detection strategies and effective therapies (4). Thus, it is crucial to gain a more complete understanding of the pathological processes driving cirrhosis and HCC.

While it is likely that etiologic risk exposures drive HCC through both genetic and epigenetic mechanisms, epigenetic modifications appear to be profoundly disrupted preceding detectable malignancy (5, 6). Epigenetic marks on the DNA, including 5-methylcytosine (5mC, mediated by the DNA methyltransferases DNMT1, DNMT3A, and DNMT3B) and 5-hydroxymethylcytosine (5hmC, mediated by the ten-eleven translocation family TET1, TET2, and TET3), are critical regulators of development that frequently become deregulated in and directly contribute to tumorigenesis (7). 5mC and 5hmC influence transcription through their differential enrichment at regulatory elements like promoters and enhancers (8, 9). These patterns are markedly deregulated during transformation as manifested by global losses in 5mC accompanied by region-specific gains (7). 5mC and 5hmC interface with epigenetic post-translational modifications of the core histones, including methylation and acetylation, to establish proper growth control. Little is known about how deregulated histone modifications influence liver disease. Recent studies show that elevated expression of histone modifiers like EZH2 (writer for the H3K27me3 repressive mark), correlate with HCC stage and survival (10). Thus, alterations in histone marks may be equally widespread and contribute to HCC, necessitating that an integrative approach be taken to elucidate the role of epigenetic deregulation in driving HCC.

In the current study, we took such an integrative approach by deeply profiling a panel of DNA and histone marks that are key regulators of transcription, in a clinically well-annotated panel of human normal, cirrhotic, and HCC liver tissues, to determine how the epigenome evolves during disease progression and identify driver events linked to epigenetic deregulation. Our results reveal distinctive alterations in all investigated epigenetic marks in cirrhotic and HCC liver tissue, with many such changes linked to transcriptome signatures specific to disease stage. Analysis of histone modifications reveals stage-specific gains in enhancers that correlate with elevated expression of disease-relevant pathways. A marked genome-wide loss of 5hmC occurs during cirrhosis, a phenomenon that is largely maintained

in HCC indicating that 5hmC deregulation precedes transformation. Interestingly, a modest number of 5hmC gain events are observed and are linked to gained enhancers and elevated gene expression. In contrast to 5hmC, *de novo* 5mC events are extensive and frequently correlate with gene repression. Finally, integration of epigenetic and transcriptomic data yields novel loci (including *COMT* and *FMO3*) that directly impact HCC cell line proliferation, thus revealing potential new drivers of HCC that are controlled through epigenetic mechanisms.

Methods

DNA methylation/hydroxymethylation

Reduced representation bisulfite sequencing (RRBS) was performed as described (11). Libraries were prepared following a standard RRBS protocol from size selected *MspI* digested genomic DNA isolated from liver tissues. Sequencing on an Illumina HiSeq 2500 was performed at the Medical Genome Facility at Mayo Clinic Rochester followed by QC, processing, and alignment through the SAAP-RRBS pipeline (Supplemental Table S1 (12)). Additional information can be found in the Supplemental Materials section. Methylated DNA immunoprecipitation (DIP)-qPCR was performed as described previously (11).

Gene expression

RNA-seq was performed as previously described (11). Gene expression is represented as Fragments per Kilobase of Transcript per Million mapped reads (FPKM). All qRT-PCR analyses were performed in triplicate, as in (13). Primers are listed in Supplemental Table S2 and in (14). Additional detail is provided in the supplemental material.

Histone modifications

ChIP-seq was performed as previously described (11, 14). Briefly, samples were fixed with 1% PFA and digested with micrococcal nuclease to isolate nucleosomes. Sonication was performed using a Diagenode Bioruptor. Chromatin input was measured using a Qubit Fluorimeter (Thermo Fisher). The following antibodies were used: H3K4me1 (ab8896, Abcam), H3K27ac (ab4729, Abcam), H3K4me3 (ab8580, Abcam), and H3K27me3 (#9733, CST). Samples were 51bp paired-end sequenced on the Illumina HiSeq2500. Peaks were called using SICER (15).

Functional assays

Three liver cell lines (Hu1545, Hep3B, and Huh7) were used to investigate the effect of ectopic re-expression of putative epigenetic driver genes on cell growth. Hu1545 is an hTERT-immortalized normal hepatocyte, which serves as a 'normal' reference. Hep3B and Huh7 are HCC-derived. Cell viability is assessed using PrestoBlue Cell Viability Reagent (A13261, Thermo Fisher) according to the manufacturer's specifications, using a BMG FLUOstar Omega plate reader. Colorimetric measurement of caspase 3 (CASP3C, Sigma-Aldrich) is used as a surrogate for activation of the apoptosis pathway using standard protocols. Cell invasion assays (Cytoselect, Cell Biolabs, Inc) were performed according to manufacturer's instructions. All assays are performed in triplicate. Additional information is contained in the Supplemental Materials.

Software packages and statistical analysis

Ingenuity Pathway Analysis (IPA, Qiagen) is used for gene ontology and comparative analyses. Basic summarization, visualization, and statistical analysis are performed in R. Phyloepigenetic trees are constructed using the R package ‘ape’. ChromHMM is used to construct histone modification state models (16). Browser views were obtained from the UCSC genome browser. Additional description is provided in the Supplemental Materials.

Results

Defining the normal liver epigenome

To elucidate the role of the epigenome in driving liver disease and HCC, ten fresh-frozen patient liver samples were macrodissected and deeply profiled for epigenetic marks and expression, including two normal tissues (no cirrhosis or HCC, NL-1/NL-2), and four matched cirrhosis-HCC pairs (Cirr1–4/HCC1–4, Supplemental Tables S1 and S3). A series of genome-wide methods was then performed on each of the samples, including: RRBS for 5mC and a TET-assisted modification of RRBS (termed here TAB-RRBS(17)) for 5hmC, chromatin immunoprecipitation (ChIP-seq) for four histone modifications: H3K4me1 (poised enhancers), H3K27ac (active enhancers when co-localized with H3K4me1, and active promoters), H3K4me3-active promoters, and H3K27me3 (polycomb repressed regions or bivalent loci when co-localized with H3K4me3), and RNA-seq for gene expression. A global summary of these data is shown as a Circos plot (Supplemental Figure 1A (18)) and in numerical form in Table 1.

We first compared the data derived from our normal liver samples against each other and with publicly available data to validate our methods, and to establish a reliable normal liver reference epigenome/transcriptome to permit accurate quantification of disease-related changes. Datasets from normal tissues are obtained from ENCODE (Liver; E066, Brain; E071, Small intestine; E109, and Lung; E096). Based on gene expression, our normal liver samples are highly similar to one another ($R=0.98$), but substantially different from other tissues, as expected (Supplemental Figure 1B). Similarly, more than 80% of enhancer-associated H3K4me1 and H3K27ac peaks from our normal liver samples reciprocally overlap with ChIP-seq derived from ENCODE normal liver (Supplemental Figure 1C), consistent with previous reports demonstrating that enhancers are tissue specific (19). Furthermore, 5mC and 5hmC show a high degree of correlation in normal liver ($R=0.93$ and 0.83 , respectively) between samples in the RRBS/TAB-RRBS data (Supplemental Figures 1D-E), thus establishing a normal liver reference epigenome (Supplemental Figure 1F shows a representative region). We also compared results derived from different genome-wide 5mC/5hmC measurement platforms from our normal liver tissues, including the Illumina Infinium MethylationEPIC (850k) and 450k BeadChip arrays (Supplemental Figures 2A-B), which revealed good concordance. Examination of individual loci with significant overlap between Infinium 850k and RRBS data within four genes (*LCAT*, *C1ORF159*, *PVRL4*, and *TNR*) demonstrates a high degree of conservation between deep sequencing and array-based measurements for both methylation and hydroxymethylation (Supplemental Figures 2C-F).

Gene expression

We next assessed gene expression profiles in our normal, cirrhotic, and HCC liver samples using RNA-seq (Table 1, Supplemental Tables S1 and S3). Independent validation of RNA-seq by qRT-PCR shows good agreement between methods (Supplemental Figure 3A). A substantial proportion of expression changes in cirrhotic liver (n=1,262, Figure 1A, Table 1) are increases relative to normal liver (89%). The number of altered genes further increases in HCC (to 2,080 genes altered, Figures 1B-C), again largely overexpression events (83%) relative to normal liver. Due to the expression differences between disease stages, gene expression alone separated samples into three groups corresponding to their disease stage (Supplemental Figure 3B). Many of the genes upregulated in cirrhosis remain upregulated during HCC, often demonstrating a stepwise increase in expression (Figures 1A-C, Supplemental Figure 3C). Differentially expressed loci involved in pathways related to immune function (e.g. lymphocyte migration) are preferentially enriched in cirrhotic tissue, while HCC-specific gene expression changes are more strongly associated with terms including 'invasion of tumor' based on Ingenuity Pathway Analysis (IPA, Figure 1D). However, several upstream regulators are enriched in both disease stages; many of these are associated with general hepatic function and are intimately linked to liver disease (e.g. WNT1, TGFB1, Supplemental Figure 3D). Interestingly, many inflammation-related genes upregulated in cirrhosis are associated with HCC patient survival differences (Supplemental Table 4).

Given the epigenetic marks we are examining, and the moderate frequency of chromatin regulator mutations in HCC (20), we performed a focused expression analysis on a set of relevant epigenetic writers, readers, and erasers. Many of these genes demonstrate changes in expression, including a stepwise loss of TET2 (Figure 1E, $p=0.07$) representing at least one potential contributor to deregulation of the 5hmC mark in liver disease (as described later). DNMTs are upregulated in both cirrhosis and HCC, consistent with previous reports (21). The H3K27me3 writer EZH2 is stepwise upregulated in cirrhosis and HCC and significantly associated with poor survival (Figure 1E (10)). Frequently mutated regulators of chromatin in HCC ($\geq 6\%$), including *ARID1A*, *ARID2*, *SETD2*, and *RB1*, show minimal changes in expression. Examination of expression in the same panel of epigenetic regulators (Figure 1E) in the TCGA LIHC dataset reveals several striking correlations. For example, EZH2, which is upregulated in our RNA-seq data set in HCC, is also upregulated in 98% of HCCs in the TCGA dataset (367/373 HCCs). Similarly, TET2, which is downregulated in our dataset, is downregulated in 73% of HCCs from TCGA. DNMT1, DNMT3A, and DNMT3B are upregulated in our data set, consistent with upregulation in 88%, 91%, and 91% of TCGA HCC samples, respectively. Of the genes covered in our study and TCGA, 1,469 (91%) of our 1,606 upregulated genes show concordant expression changes in the TCGA data set (Figure 1F). Similarly, 155 (67%) of the 231 downregulated genes are consistent between the two data sets. This analysis further confirms that our results, derived from a smaller sample set, reflect liver disease and that differences we identify are likely broadly applicable to HCC biology.

Histone modifications

To address the gap in knowledge regarding the dynamics of histone modifications in liver disease, we performed ChIP-seq for four histone marks key to the control and definition of promoters and enhancers in a single normal liver tissue and two cirrhosis-HCC pairs previously analyzed by RNA-seq (Figure 2A, Table 1, Supplemental Tables 1 and 3). By assessing combinations of marks, we can define regulatory element classes most affected across disease stages. Bivalent loci (marked by H3K4me3 and H3K27me3, generally lowly expressed but poised for activation or repression (22)) for example, are relatively unique to each disease stage (Figure 2B, left). In contrast, active enhancers (marked by H3K4me1 and H3K27ac) are much more highly conserved across the three stages (Figure 2B, right).

To elucidate how histone modifications evolve during hepatocarcinogenesis, we performed ChromHMM (16) using our four histone marks (Figure 2C). Each combination of histone marks is linked to regulatory elements and/or genome features, derived from a prior study (23). ChromHMM reveals a number of interesting findings; for example, new active enhancers (state 7; presence of H3K4me1 and H3K27ac) are enriched in cirrhotic tissue. To examine the dynamics between two stages of disease, we plotted the ten states in normal liver and determined their corresponding states in cirrhosis and HCC (Figure 2D, Supplemental Figures 4A-C). As each state in normal liver may remain unchanged in disease or transition to a different state, we can understand the dynamics of histone modifications genome-wide by simplifying them into ten possible configurations (*i.e.* states). For example, genes associated with quiescent regions (unmarked, state 1 in Figure 2C) in normal liver generally remain quiescent in tumor (89% of the total regions in this category; Figure 2D; 1st bar, Supplemental Figure 4B; 1st bar). However, 5.6% of genes that were formerly quiescent (Figure 2D; 1st bar) gain enhancer marks in liver disease (state 7; H3K4me1 and H3K27ac, and state 8; H3K4me1 for poised enhancers). Many bivalent loci (state 3, 3rd bar; H3K4me3 and H3K27me3 together) become epigenetically silenced (14.6%) following loss of H3K4me3. State 2 frequently loses repressive H3K27me3 to become unmarked (59%), while the H3K27ac-associated state 6 frequently acquires H3K4me1 to become an active enhancer (41%, state 7). Thus, while much of the epigenome remains unchanged in cirrhosis and HCC, many specific regions display marked epigenetic switching, particularly those linked to enhancers.

Within individual cirrhosis-tumor pairs (e.g. Supplemental Figures 4A-B) enhancers are major targets for epigenome remodeling. For example, regions with H3K4me1 or H3K27ac alone (6th and 8th bars) become active enhancers at high frequency in cirrhosis and HCC. To test whether these state transitions impact transcription, we examined matched RNA-seq data for each sample to determine whether specific state transitions impact expression. Indeed, gain of active enhancer marks in normally quiescent loci is associated with increased transcription (Figure 2E) and is linked to “hematologic system development and function” using IPA. In addition, when repressive H3K27me3 is removed from a bivalent promoter to transition to an active promoter, we observe increases in transcription of loci associated with “generation of tumor” (Figure 2F). Finally, gain of the activating mark combination H4K4me3 and H3K27ac (state 5), is enriched in liver disease and is linked to upregulation of genes associated with these regions (Figures 2C-D, Supplemental Figures 4A-D). Taken

together, this analysis demonstrates that histone modifications are significantly altered in cirrhosis and HCC, with particular enrichment for changes in enhancers leading to widespread gene upregulation.

DNA methylation/hydroxymethylation

We next performed RRBS to assess DNA methylation across the genome. Given that standard RRBS measures the summation of 5mC and 5hmC, and because 5hmC is emerging as a mark particularly sensitive to environmental perturbations (24), we also performed TET-assisted RRBS (TAB-RRBS (17)) to specifically interrogate 5hmC. We observe genome-wide increases in 5mC in both cirrhosis and HCC (Figures 3A-B upper panels). However, 5hmC is markedly depleted in cirrhotic tissues, and remains low in HCC, with only a small fraction of CpG sites gaining hydroxymethylation (~0.5% of CpGs linked to genes, Figures 3A-B lower panels). Comparing the dynamics between cirrhosis and HCC, we observe that 5mC is gained preferentially at regions that already possess moderate to high levels of 5mC in normal liver. However, HCC also displays a clear increase in *de novo* hypermethylation events (Figure 3A, top panels). Gains in 5mC are prominent at transcription start sites (TSS), with a stepwise increase from normal to cirrhosis to HCC (Figure 3B, top panel). 5hmC loss is a genome-wide effect, with little evidence for changes specifically focused on the TSS, as these regions are already depleted for 5hmC in normal liver (Figure 3B, bottom panel). The loss of 5hmC is further supported by mass spectrometry (LC-MS/MS) analysis of genomic DNA (Supplemental Figures 5A-B). Mass spectrometry quantification shows a modest overall reduction in total 5mC levels, while RRBS shows an increase in average methylation. This is likely due to 5mC loss in intergenic and repetitive regions in MS measurements, as has been observed in many cancers (7, 25), while RRBS has an inherent bias towards CG-rich regions including promoters (26).

We recently reported that DNA methylation displays substantial heterogeneity within individual patient liver tissues that correlates with cell proliferation (5, 27). A subset of cirrhotic regions is epigenetically more ‘aggressive’, which correlates highly with an “epigenetic progression” score derived from phyloepigenetic tree reconstruction (5). Applying a similar concept to samples in this study, we observe that 5mC robustly segregates normal, cirrhosis, and HCC ($p=0.05$; cirrhosis vs HCC, Figure 3C; left), while 5hmC shows no such separation, but both disease states segregate from normal liver based on 5hmC ($p=0.89$; cirrhosis vs HCC, Figure 3C; right). To address whether the DNA mark changes we observe represent “switching” between 5mC and 5hmC during disease progression, we plotted changes in each mark against one another in promoter regions (Figure 3D). Since there are relatively few hyperhydroxymethylation events, it is not surprising that relatively few CpGs (~0.3%) gain 5hmC with or without changes in 5mC (within 200 bp upstream of the TSS; TSS200) in cirrhosis and HCC relative to normal liver (Figure 3D, upper right quadrants). However, a significant number of promoter-associated CpGs gain 5mC at the expense of 5hmC (~3,000 sites in cirrhosis and ~3,500 in HCC, Figure 3D, upper left quadrants). We performed IPA on genes that switched from 5hmC to 5mC in the TSS200 region in HCC and identified pathways associated with liver tumorigenesis as among the most enriched (Supplemental Figure 5C). As a specific example, we selected a gene from the 5hmC to 5mC switching TSS200 ontology, *lecithin-*

cholesterol acyltransferase (LCAT), to examine whether switching correlates with expression at loci identified in Figure 3D. As Figure 3E shows, 5mC progressively increases at the *LCAT* promoter during disease progression with a corresponding decrease in expression ($R^2=0.55$, Supplemental Figures 5D-E). Querying the much larger TCGA HCC sample set reveals that lower *LCAT* expression correlates with poorer survival (Figure 3F, left). *LCAT* 5mC data from the TSS200 region is also significantly associated with survival (Figure 3F, right). Further examination of *LCAT* reveals that it is among the top ten most differentially expressed genes based on HCC pathologic stage from TCGA (Kruskal-Wallis $P=1.15E-05$), further solidifying an important role for *LCAT* in the prognosis of HCC, which was revealed through an integrated epigenomic analysis.

Integration of all -omics platforms

Since epigenetic marks typically collaborate to establish functional outcomes, we simplified each of the parameters relative to important regulatory functions through ChromHMM (Figure 4A (19)). The data was parsed into ten states (Figure 2) and additional parameters (e.g. expression, 5mC, 5hmC) were linked to each state (Figure 4A, top four panels). States associated with promoter regions including state 5 and state 10, have the highest level of expression and the lowest level of 5mC (Figure 4A). After merging the ChromHMM model, we examined differences in state calls between normal liver and HCC to identify regions where a particular state is unique to a disease stage. TCGA survival data was then linked to genes within these regions (Figure 4B). We observe that a greater proportion of genes associated with state 7 only in tumor are more likely to be associated with a survival difference than those where state 7 appears in both normal and tumor tissue (Figure 4C, left). Thus, regions that gain the active enhancer mark combination (H3K4me1 and H3K27ac) during hepatocarcinogenesis, are more likely to be linked to patient prognosis, suggesting the epigenetic change has functional consequences for the patient. A similar trend is also observed for active promoter regions (state 10, Figure 4C, right).

Although there are few hyper-hydroxymethylation events observed during cirrhosis and HCC progression, we nonetheless examined whether they co-localized with other marks within the genome using ChromHMM. As shown in Figure 4D, regions marked by H3K4me1 alone (poised enhancer, state 8) or H3K4me1 and H3K27ac (active enhancer, state 7), contain the most hyper-5hmC events, while promoter regions including state 5 (H3K4me3 and H3K27ac) and state 10 (H3K4me3, H3K27ac, and H3K4me1) are depleted for these events, suggesting that enhancer regions are particularly important targets of aberrant epigenetic regulation by 5hmC (Figure 4D, Supplemental Figures 6A-C). Since we observe enrichment of 5hmC gains in state 7, and both states 7 and 10 are enriched in HCC, we further examined these states. Expression was quantified at genes where active enhancers are gained in tumor and a significant association with patient survival was identified. Interestingly, of the 94 genes linked to a survival difference for HCC based on TCGA LIHC data, 84% are more highly expressed in the HCC samples relative to normal liver (Figure 4E). One such gene, *COMT*, displays a clear inverse correlation between 5mC and gene expression level (Figure 4F, Supplemental Figure 6D, right). Low expression of *COMT* is associated with poor HCC-specific prognosis (Supplemental Figure 6D, left). Similar trends are observed for other loci including *FMO3* (Supplemental Figure 6E), while an inverse

trend is observed for genes like *Annexin A2* (*ANXA2*), where high expression and low methylation of *ANXA2* correlate with poor survival in HCC (Supplemental Figure 6F).

TCGA 450k data does not distinguish between 5mC and 5hmC, potentially confounding functional implications of an observed '5mC' change in disease. To address this, we examined genes showing significant increases in promoter 5hmC in our dataset that overlapped with TCGA hypermethylated promoters (Supplemental Figures 7A-B). A significant proportion of genes that sustain hyperhydroxymethylation in our dataset also display elevated DNA modifications in TCGA data. Interestingly, when we examine expression from the same cohort of TCGA samples, these same loci are in general upregulated, consistent with the positive links between 5hmC and transcription (Supplemental Figure 7C). Examination of a specific locus, the *TBX15* promoter, demonstrates that this gene is hyperhydroxymethylated in HCC, with relatively modest changes in 5mC (Supplemental Figures 7D-E, left compared to middle panel). However, TCGA data suggests that the *TBX15* promoter is hypermethylated and upregulated (Supplemental Figures 7C, 7D right panel, 7E right panel), when in fact this more likely results from elevated 5hmC that is concomitantly measured with 5mC using the Infinium 450k array.

We next integrated our epigenetic mark maps with our transcription data and coupled these to publicly available data to identify genes likely to be involved in epigenetically driving hepatocarcinogenesis. As a representative example, *COMT* displays DNA-based epigenetic changes at regions that lose active enhancer histone modifications in HCC relative to normal liver (Figure 5A). Further examination of *COMT*, as well as *FMO3*, another locus showing coordinated changes in multiple epigenetic marks and transcription, using TCGA data demonstrates that both genes are downregulated in HCC and display an inverse relationship between expression and 5mC (Supplemental Figures 6D-E). High levels of *COMT/FMO3* expression correlate with good prognosis in HCC patients, while high levels of methylation in these genes are associated with poor prognosis (Figure 5B and Supplemental Figures 6D-E). Moreover, *COMT* and *FMO3* are significantly downregulated in a panel of 50 non-tumor adjacent/HCC pairs (Figure 5C). Consistent with this, a majority of HCCs demonstrate promoter hypermethylation relative to their non-tumor adjacent counterparts for both *COMT* and *FMO3* (Supplemental Figure 6I). To determine whether these epigenetically driven silencing events (or putative epigenetic drivers of HCC) impact cell growth/survival, we ectopically expressed *COMT* and *FMO3* in a panel of HCC-derived cell lines and measured cell viability (Figure 5D, Supplemental Figure 5H) and caspase-3 activity (as a measure of apoptosis, Figure 5E, Supplemental Figure 5H, left). Ectopic expression of *COMT* or *FMO3* significantly reduces cell viability compared to a GFP control in HCC cell lines (Huh7 and Hep3B, where the endogenous gene is epigenetically hypermethylated (Supplemental Figure 6G) (25)) but has minimal impact on viability of an immortalized hepatocyte line (Hu1545). In keeping with viability, caspase-3 activity is elevated upon re-expression of *COMT/FMO3* relative to the GFP control for cancer cell lines (Figure 5E, Supplemental Figure 5H, right). Finally, re-expression of either *FMO3* or *COMT* in Huh7 cells reduces their ability to invade an ECM coated membrane (Figure 5F). These data therefore demonstrate that coordinated cancer-specific alterations in histone modifications, 5mC, 5hmC, and expression target

genes relevant to HCC development and/or progression, and the integrated approach reveals novel drivers of liver cancer.

Discussion

In this study, we defined the epigenome in normal, cirrhotic, and HCC liver tissue to decipher how epigenetic marks relevant to promoter and enhancer regulation are altered individually and combinatorially during liver disease progression. We observe that histone modifications are particularly volatile, with bivalent regions being relatively unique among disease stages, while active enhancer regions show both significant unique and shared localization among disease states. *De novo* acquisition of active enhancers is frequent in both cirrhosis and HCC, correlates with cancer-related pathways, and is significantly enriched for genes associated with prognosis. This is in contrast to the rest of the profiled genome where 5hmC is markedly depleted during cirrhosis and remains suppressed in HCC. Conversely, hypermethylation is abundant and displays distinct patterns between cirrhosis and HCC, where HCC is marked mainly by *de novo* hypermethylation events. Indeed, many CpGs undergo “switching” from 5hmC to 5mC, leading to gene repression. Integration of transcriptional and epigenetic data unveils putative epigenetically regulated drivers of HCC, including *COMT* and *FMO3*. Both of these genes, when re-expressed in HCC cell lines, significantly decrease cell viability, increase apoptosis, and decrease invasive capacity. Overall, these results demonstrate that integration of epigenetic data is a viable strategy for identifying key genes involved in hepatocarcinogenesis, potentially unveiling new therapeutic targets and/or biomarkers for this notoriously hard-to-treat cancer.

Recent work has begun to implicate histone marks and their associated regulators (such as the H3K27me3 writer *EZH2*) in liver disease progression. High expression of *EZH2* strongly correlates with poor prognosis ($p=0.004$, $n=417$ HCCs from TCGA). In addition, H3K4me2 (a mark linked to transcription factor binding) is depleted in HCC while the H3K4me3 writer *SETD1B* is overexpressed (consistent with our data, Figure 1E). Targeting enhancer function to treat cancer is quickly gaining traction with the BRD4 inhibitor JQ1, which interferes with enhancer activity, demonstrating anti-tumor activity in blood, brain, and lung cancers (28). Interestingly, aberrant enhancer function in *MLL-AF9* leukemia can be targeted therapeutically by three different mechanisms: BRD4 inhibition to broadly inhibit enhancer function (28), disruption of a specific enhancer region by genome editing (29), or repression of enhancer target genes (30), highlighting the importance of enhancers in cancer and their potential as therapeutic targets. We observe that active enhancers are one of the most dramatically remodeled histone modification switching events in both cirrhosis and HCC, suggesting that genes regulated by these enhancers play important roles in the initiation (*de novo* enhancers in cirrhosis) and/or progression (*de novo* enhancers in HCC and conserved enhancer deregulation between cirrhosis and HCC), and represent an efficient way to pinpoint new epigenetic driver events.

We have demonstrated here that defining the aberrant HCC epigenetic landscape using an integrated approach is a powerful tool for identifying new potential driver events regulated by epigenetic mechanisms. Such events could also represent novel therapeutic targets. Both *COMT* and *FMO3* were identified through analysis of histone modifications, 5mC/5hmC,

transcription, and TCGA data, and were functionally assessed. *COMT* methylates catecholamines, leading to their degradation, is highly expressed in normal liver, and regulates estrogen and androgen production and inactivation. *COMT* polymorphisms are associated with higher HCC risk in women (31) as well as for breast (32), and gastric cancers (33). High *COMT* expression is associated with better prognosis in TCGA data and it demonstrates a robust inverse correlation between expression and methylation (Supplemental Figure S6D). Low levels of *COMT* promoter methylation are also associated with favorable prognosis (Figure 5B). Functional data support the links between methylation and expression with survival, as ectopic expression of *COMT* in HCC cell lines leads to reduced viability/invasion and increased apoptosis (Figures 5D-F), suggesting that *COMT* functions as a tumor suppressor in HCC. *FMO3* is the most abundant flavin-containing monooxygenase family member in liver and is responsible for degradation of xenobiotics including chemotherapeutic agents like desatinib, imatinib, and vandetanib (34). Furthermore, *FMO3* levels are elevated in HCV-positive HCC, leading to increased carcinogenic metabolites (35). Our data here support a role for *FMO3* in HCC prognosis (Figure 5B, Supplemental Figure 6E), and as a putative tumor suppressor in HCC (Figures 5D-F, Supplemental Figure 6H-I). *ANXA2* is another gene derived from our dataset as having a potential role in hepatocarcinogenesis (Supplemental Figure 6F). *ANXA2* is upregulated in a variety of cancers including HCC (36). Inhibition of *ANXA2* decreases migration and invasion of HCC cells (37) and is a putative biomarker for HCC (38). Taken together, the integrated approach used here resulted in identification of loci that have been firmly linked to hepatocarcinogenesis (e.g. *ANXA2*) and that represent novel functionally relevant epigenetically regulated genes (*COMT* and *FMO3*).

Hepatocellular carcinoma is a notoriously hard to treat cancer, with few therapies available for late stage disease. Genetic changes contributing to HCC have been thoroughly characterized, and the frequently mutated driver genes are well established (e.g. *TERT*, *CTNNB1*, *TP53*) (20, 39). Most of these highly mutated genes, however, are not targetable, spurring interest in alternative approaches to identify genes/pathways critical for the initiation, progression, and metastatic potential of HCC. Epigenetic alterations remain understudied in liver cancer despite mounting evidence of their importance, including frequent mutation of epigenetic regulators *ARID1A*, *IDH1*, and *BAP1*. Data presented in this study identify global epigenetic defects during liver disease progression while also revealing specific loci that impact HCC biology, through integration of multiple epigenetic marks. These results therefore emphasize the importance of accounting for epigenetic changes in understanding the biology of HCC, demonstrate how epigenetic mechanisms converge to drive hepatocarcinogenesis, and serve as a model for integration of these marks to pinpoint novel epigenetically regulated HCC driver genes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support

This work was supported by the American Association for the Study of Liver Diseases (RAH - Clinical and Translational Fellowship Award in Liver Disease) and the National Institutes of Health (KDR/CL - R01DK110024).

Abbreviations

DNMT	DNA methyltransferase
TET	Ten-eleven translocation
COMT	Catechol-O-methyltransferase
FMO3	Flavin containing monooxygenase 3
EZH2	Enhancer of zeste homolog 2
LCAT	Lecithin-cholesterol acyltransferase
ANXA2	Annexin A2
HCC	Hepatocellular carcinoma
5mC	DNA 5-methylcytosine
5hmC	DNA 5-hydroxymethylcytosine
HCV	Hepatitis C virus
TCGA LIHC	The Cancer Genome Atlas liver hepatocellular carcinoma
RRBS	Reduced representation bisulfite sequencing
TAB-RRBS	TET-assisted reduced representation bisulfite sequencing
ChromHMM	Chromatin Hidden Markov modeling
TNM	Tumor node metastasis
FPKM	Fragments per kilobase of transcript per million mapped reads

References

1. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 2014;74:2913–2921. [PubMed: 24840647]
2. Njei B, Rotman Y, Ditah I, Lim JK. Emerging trends in hepatocellular carcinoma incidence and mortality. *Hepatology* 2015;61:191–199. [PubMed: 25142309]
3. Seitz HK, Stickel F. Molecular mechanisms of alcohol-mediated carcinogenesis. *Nat Rev Cancer* 2007;7:599–612. [PubMed: 17646865]
4. El-Serag HB, Kanwal F. Epidemiology of hepatocellular carcinoma in the United States: where are we? Where do we go? *Hepatology* 2014;60:1767–1775. [PubMed: 24839253]
5. Hlady RA, Zhou D, Puszyk W, Roberts LR, Liu C, Robertson KD. Initiation of aberrant DNA methylation patterns and heterogeneity in precancerous lesions of human hepatocellular cancer. *Epigenetics* 2017;12:215–225. [PubMed: 28059585]
6. Villanueva A, Portela A, Sayols S, Battiston C, Hoshida Y, Mendez-Gonzalez J, Imbeaud S, et al. DNA methylation-based prognosis and epdrivers in hepatocellular carcinoma. *Hepatology* 2015;61:1945–1956. [PubMed: 25645722]

7. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nature reviews. Cancer* 2011;11:726–734. [PubMed: 21941284]
8. Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev* 2014;28:2103–2119. [PubMed: 25223896]
9. Putiri EL, Tiedemann RL, Liu C, Choi JH, Robertson KD. Impact of human MLL/COMPASS and polycomb complexes on the DNA methylome. *Oncotarget* 2014;5:6338–6352. [PubMed: 25071008]
10. Hung SY, Lin HH, Yeh KT, Chang JG. Histone-modifying genes as biomarkers in hepatocellular carcinoma. *Int J Clin Exp Pathol* 2014;7:2496–2507. [PubMed: 24966962]
11. Zhou D AB, Li S, Hlady RA, Thompson JJ, Schroeder MA, Lee J- H, Qiu J, Schwartz PH, Sarkaria JN, Robertson KD. Distinctive epigenomes characterize glioma stem cells and their response to differentiation cues. *Genome Biology* 2018.
12. Sun Z, Baheti S, Middha S, Kanwar R, Zhang Y, Li X, Beutler AS, et al. SAAP-RRBS: streamlined analysis and annotation pipeline for reduced representation bisulfite sequencing. *Bioinformatics* 2012;28:2180–2181. [PubMed: 22689387]
13. Tiedemann RL, Putiri EL, Lee JH, Hlady RA, Kashiwagi K, Ordog T, Zhang Z, et al. Acute depletion redefines the division of labor among DNA methyltransferases in methylating the human genome. *Cell Rep* 2014;9:1554–1566. [PubMed: 25453758]
14. Tiedemann RL, Hlady RA, Hanavan PD, Lake DF, Tibes R, Lee JH, Choi JH, et al. Dynamic reprogramming of DNA methylation in SETD2-deregulated renal cell carcinoma. *Oncotarget* 2016;7:1927–1946. [PubMed: 26646321]
15. Zang C, Schones DE, Zeng C, Cui K, Zhao K, Peng W. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* 2009;25:1952–1958. [PubMed: 19505939]
16. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat Protoc* 2017;12:2478–2492. [PubMed: 29120462]
17. Yu M, Hon GC, Szulwach KE, Song CX, Jin P, Ren B, He C. Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat Protoc* 2012;7:2159–2170. [PubMed: 23196972]
18. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, et al. Circos: an information aesthetic for comparative genomics. *Genome Res* 2009;19:1639–1645. [PubMed: 19541911]
19. Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, Zhang X, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 2011;473:43–49. [PubMed: 21441907]
20. Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Nguyen HH, Aoki M, et al. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat Genet* 2012;44:760–764. [PubMed: 22634756]
21. Oh BK, Kim H, Park HJ, Shim YH, Choi J, Park C, Park YN. DNA methyltransferase expression and DNA methylation in human hepatocellular carcinoma and their clinicopathological correlation. *Int J Mol Med* 2007;20:65–73. [PubMed: 17549390]
22. Voigt P, Tee WW, Reinberg D. A double take on bivalent promoters. *Genes Dev* 2013;27:1318–1338. [PubMed: 23788621]
23. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;518:317–330. [PubMed: 25693563]
24. Tammen SA, Park LK, Dolnikowski GG, Ausman LM, Friso S, Choi SW. Hepatic DNA hydroxymethylation is site-specifically altered by chronic alcohol consumption and aging. *Eur J Nutr* 2017;56:535–544. [PubMed: 26578530]
25. Hlady RA, Tiedemann RL, Puszyk W, Zendejas I, Roberts LR, Choi JH, Liu C, et al. Epigenetic signatures of alcohol abuse and hepatitis infection during human hepatocarcinogenesis. *Oncotarget* 2014;5:9425–9443. [PubMed: 25294808]
26. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res* 2005;33:5868–5877. [PubMed: 16224102]

27. Hlady RA, Robertson KD. Genetic and epigenetic heterogeneity in normal liver homeostasis and its implications for liver disease and hepatocellular cancer. *Semin Liver Dis* 2018;38:41–50. [PubMed: 29471564]
28. Carretta M, Brouwers-Vos AZ, Bosman M, Horton SJ, Martens JHA, Vellenga E, Schuringa JJ. BRD3/4 inhibition and FLT3-ligand deprivation target pathways that are essential for the survival of human MLL-AF9+ leukemic cells. *PLoS One* 2017;12:e0189102. [PubMed: 29240787]
29. Bahr C, von Paleske L, Uslu VV, Remeseiro S, Takayama N, Ng SW, Murison A, et al. A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. *Nature* 2018;553:515–520. [PubMed: 29342133]
30. Jiang X, Huang H, Li Z, Li Y, Wang X, Gurbuxani S, Chen P, et al. Blockade of miR-150 maturation by MLL-fusion/MYC/LIN-28 is required for MLL-associated leukemia. *Cancer Cell* 2012;22:524–535. [PubMed: 23079661]
31. Yin PH, Lee HC, Chau GY, Liu TY, Liu HC, Lui WY, Chi CW. Polymorphisms of estrogen-metabolizing genes and risk of hepatocellular carcinoma in Taiwan females. *Cancer Lett* 2004;212:195–201. [PubMed: 15341023]
32. Long JR, Cai Q, Shu XO, Cai H, Gao YT, Zheng W. Genetic polymorphisms in estrogen-metabolizing genes and breast cancer survival. *Pharmacogenet Genomics* 2007;17:331–338. [PubMed: 17429315]
33. Freedman ND, Ahn J, Hou L, Lissowska J, Zatonski W, Yeager M, Chanock SJ, et al. Polymorphisms in estrogen- and androgen-metabolizing genes and the risk of gastric cancer. *Carcinogenesis* 2009;30:71–77. [PubMed: 19015200]
34. Teo YL, Ho HK, Chan A. Metabolism-related pharmacokinetic drug-drug interactions with tyrosine kinase inhibitors: current understanding, challenges and recommendations. *Br J Clin Pharmacol* 2015;79:241–253. [PubMed: 25125025]
35. Jeong SW, Jang JY, Chung RT. Hepatitis C virus and hepatocarcinogenesis. *Clin Mol Hepatol* 2012;18:347–356. [PubMed: 23323249]
36. Kohli S, Bhardwaj A, Kumari R, Das S. SIRT6 Is a target of regulation by UBE3A that contributes to liver tumorigenesis in an ANXA2-dependent manner. *Cancer Res* 2018;78:645–658. [PubMed: 29217762]
37. Zhang W, Zhao P, Xu XL, Cai L, Song ZS, Cao DY, Tao KS, et al. Annexin A2 promotes the migration and invasion of human hepatocellular carcinoma cells in vitro by regulating the shedding of CD147-harboring microvesicles from tumor cells. *PLoS One* 2013;8:e67268. [PubMed: 23950866]
38. El-Abd N, Fawzy A, Elbaz T, Hamdy S. Evaluation of annexin A2 and as potential biomarkers for hepatocellular carcinoma. *Tumour Biol* 2016;37:211–216. [PubMed: 26189841]
39. Cancer Genome Atlas Research Network. Electronic address: wheeler@bcm.edu. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell* 2017;169:1327–1341 e1323. [PubMed: 28622513]

Figure 1:

Dynamic gene expression across liver disease and cancer. A. Volcano plots comparing gene expression between cirrhotic and normal liver, B. HCC and normal liver, and C. HCC and cirrhosis. More highly expressed genes in each comparison (n=2 normal liver, n=4 cirrhosis, n=4 HCC) are colored in red, while downregulated genes are colored green. A threshold of (1.5-fold change * -log p-value) is used for defining significant changes. D. IPA enrichment for differentially expressed genes in HCC and cirrhosis, relative to normal liver. A color bar is shown for the predicted activation Z-score (purple; down, orange; up). E. Heatmap of epigenetic regulator expression in liver disease. A color bar is shown with lowly expressed genes represented by white and highly expressed genes in blue. The average of two normal, four cirrhosis, and four HCCs is shown. F. Overlap of differentially expressed genes derived from our RNA-seq dataset with those identified in TCGA LIHC tissue samples. Up- and down-regulated genes in our dataset are defined as in part A. A total of 50 non-tumor adjacent and 373 HCCs are used for TCGA data. The overlapping pool of genes used in the analysis is 16,502.

Figure 2:

Genome-wide alterations in histone modifications in liver disease. A. The number of called peaks for each mark or combination of marks (enhancer; H4K4me1 + H3K27ac, bivalent; H3K4me3 + H3K27me3) for normal liver (n=1), cirrhosis (n=2), and HCC (n=2) samples. B. Venn diagrams depicting unique and overlapping bivalent (left panel) and enhancer (right panel) marks in normal (green), cirrhotic (blue), and HCC (red) liver samples. Numbers of peaks are given as $\times 1000$ ('k'). C. ChromHMM modeling of four histone marks with ten states (left), a description of their regulatory function (middle), and the distribution of each state across the genome in normal, cirrhotic, and HCC liver tissues (right). In ChromHMM, darker blue represents higher enrichment in the left panel, and darker blue represents more genes in the right panel. D. Stacked bar chart displaying the changes in ChromHMM states between normal and HCC for the ten called states in part C. Each color corresponds to a specific state in HCC, as shown by the key on the right. For example, regions colored in red are active enhancers in HCC, which are observed in regions that are active_tss, active, active_enhancer, and poised_enhancer in normal liver. E. Bar chart highlighting gene expression differences between normal and cirrhotic liver in the state transition 'quiescent to active enhancer' (left) and the associated IPA for loci enriched in this state change (right). The same samples used for ChromHMM are shown here. F. Bar charts of gene expression for the state transition of 'bivalent to active promoter' across normal (green), cirrhosis

(blue), and HCC (red) liver (left). IPA comparing HCC relative to normal liver for this state transition is shown on the right.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3:

Defining DNA methylation and hydroxymethylation alterations during hepatocarcinogenesis. A. Binned scatterplots of 5mC (top two panels) and 5hmC changes (bottom two panels) in cirrhosis (left) and HCC (right) relative to normal liver. The change in each DNA mark is plotted against the average of the mark for involved samples (e.g. two normal and four cirrhotic livers in top left panel) at a given CpG site. A scale is shown at the right where red shading indicates a higher density of CpGs. B. Tag density plots for 5mC (top) and 5hmC (bottom) levels flanking the transcription start site (TSS). The average for normal liver (NL1-2) is shown in green, cirrhosis (Cirr1-4) in blue, and HCC (HCC1-4) in red. C. Phylogenetic trees based on 5mC (left) and 5hmC (right). Samples are color coded based on disease state as in B. P-values for cirrhosis versus HCC epigenetic progression (EP)-score5 are shown for 5mC and 5hmC. D. Binned scatterplots of 'switching' between 5mC and 5hmC at a given CpG site in regions surrounding the transcription start site (TSS200) in cirrhosis (left) and HCC (right) relative to normal liver. E. 5mC and expression status of LCAT, showing an inverse correlation between these parameters in cirrhosis and HCC. A trend line is shown ($R=0.55$). F. Kaplan-Meier survival curve based on LCAT expression data derived from TCGA LIHC (left panel). Expression data from 329 TCGA samples was used (high $n=165$, low $n=164$). Kaplan-Meier survival curve based on DNA methylation of a CpG within the TSS200 region of LCAT from the same patient dataset (right panel). Infinium 450k data from 408 TCGA samples was used (high $n=204$, low $n=204$). P-values are shown to indicate a significant survival difference between low expression/methylation and high expression/methylation.

Figure 4:

Integration of transcriptome and epigenome data reveals potential novel drivers of HCC. A. Bar charts summarizing gene expression (top), DNA methylation (dark red bars), DNA hydroxymethylation (pink bars), and number of CpGs in each state (black bars), segregated by ChromHMM state (blue, using ten states). B. Relative enrichment of genes associated with survival differences, based on TCGA LIHC data, across the ten ChromHMM states identified normal liver and HCC (blue) or unique to HCC (red). C. Bar chart of significant p-values for survival differences in state 7 (active enhancer; left) and state 10 (active promoter; right) for uniquely identified states in normal liver, tumor, or shared (between normal and HCC). The height of the bar corresponds to the relative number of genes with significant differences in survival from TCGA. D. Bar chart showing the relative enrichment of regions that gain 5hmC in HCC relative to normal liver by ChromHMM state, demonstrating localization of such events to active and poised enhancers (states 7 and 8, red arrows) and depletion of this change in active promoter regions (states 5 & 10, green arrows). E.

Spaghetti plot of expression changes between normal liver and HCC for genes linked to gains in active enhancer status in HCC (state 7), showing that most enhancer gains are associated with increased expression of the nearby gene(s). F. Scatterplot of promoter 5mC and expression levels for the COMT gene, demonstrating an inverse correlation. Normal (n=2, green), cirrhosis (n=4, blue), and HCC liver tissues (n=4, red) are indicated.

Figure 5:

Putative epigenetic drivers COMT and FMO3 modulate growth of HCC cell lines. A. UCSC genome browser views for the COMT gene in normal (left) and HCC liver (right) tissues. Histone modifications are shown in green (H3K4me1), blue (H3K27ac), purple (H3K4me3), and red (H3K27me3) with input (black) as a reference. 5mC is shown in red, 5hmC in pink. ChromHMM states are shown below (colors correspond to states identified in Figure 2C). Active enhancers that are lost in HCC are highlighted in red and blue boxes. B. DNA methylation-derived survival curves for COMT (left) and FMO3 (right) from TCGA 450k data. C. Gene expression change for COMT (blue) and FMO3 (orange) for 50 non-tumor adjacent/HCC pairs derived from TCGA. D. Cell viability assay at 0, 24, and 48 hours after ectopic re-expression of GFP, COMT, or FMO3 in the indicated cell lines. Samples are normalized to empty GFP vector transfected cells for each cell line. E. Line plot of apoptosis based on caspase 3 activity in cell lines ectopically expressing GFP, COMT, or FMO3 over 72 hours. Samples are normalized to an empty GFP vector transfected control for each cell

line. Significance is denoted as * ($p < 0.05$). F. Cell invasion assay for COMT and FMO3 re-expression in normal hepatocyte (Hu1545, blue) or HCC (Huh7, orange) cell lines.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1: Summary of the number of changes in each of the indicated parameters relative to normal liver.

Mark	5mC		5hmC		RNA-seq		H3K4me1		H3K27ac		H3K4me3		H3K27me3	
	gain	loss	gain	loss	up	down	gain	loss	gain	loss	gain	loss	gain	loss
Cirrhosis	63941	5253	7130	135582	1126	136	27425	11849	27006	13653	26515	60566	55413	68679
HCC	90148	26096	4554	169342	1723	357	35492	26393	24107	31561	33303	57898	51497	65264