Published in final edited form as:

Hepatology. 2021 July ; 74(1): 183–199. doi:10.1002/hep.31659.

DNA methylation profiling of human hepatocarcinogenesis

Gabriela Hernandez-Meza[1,*], **Johann von Felden**[1,2,*], **Edgar E. Gonzalez-Kozlova**[3], **Teresa Garcia-Lezana**[1], **Judit Peix**[4], **Anna Portela**[5], **Amanda J. Craig**[1], **Sergi Sayols**[5,6], **Myron Schwartz**[7], **Bojan Losic**[3,8,9], **Vincenzo Mazzaferro**[10], **Manel Esteller**[11,14,15,16], **Josep M. Llovet**[1,4,12], **Augusto Villanueva**[1,13]

[1.]Division of Liver Diseases, Liver Cancer Program, Tisch Cancer Institute, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[2.]I. Department of Internal Medicine, University Medical Center Hamburg Eppendorf, Hamburg, Germany.

[3.]Department of Genetics and Genomic Sciences, Cancer Immunology Program, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[4.]Translational Research in Hepatic Oncology, Liver Unit, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS)-Hospital Clínic, Universitat De Barcelona, Catalonia, Spain

[5.]Cancer Epigenetics and Biology Program, Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain.

[6.]Institute of Molecular Biology, Mainz, Germany.

[7.]Department of Surgery, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[8.]Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[9.]Diabetes, Obesity and Metabolism Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[10.]Gastrointestinal Surgery and Liver Transplantation Unit, National Cancer Institute, Milan, Italy.

[11.]Josep Carreras Leukemia Research Institute (IJC), Badalona, Barcelona, Catalonia, Spain

[12.]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia, Spain.

[13.]Division of Hematology and Medical Oncology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

[14.]Centro de Investigacion Biomedica en Red Cancer (CIBERONC), Madrid, Spain.

[15.]Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain.

[16.]Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona (UB), Barcelona, Catalonia, Spain.

Corresponding author: Augusto Villanueva M.D., Ph.D.; Division of Liver Diseases; 1425 Madison Ave, RM 11-70E, Box 1123; 10029 New York, NY (USA), augusto.villanueva@mssm.edu.

*Shared first authorship

Abstract

Background: Mutations in *TERT* promoter are established gatekeepers in early hepatocarcinogenesis, but little is known about other molecular alterations driving this process. Epigenetic deregulation is a critical event in early malignancies. Thus, we aimed: 1) to analyze DNA methylation changes during the transition from preneoplastic lesions to early HCC (eHCC) and to identify candidate epigenetic gatekeepers, and 2) to assess the prognostic potential of methylation changes in cirrhotic tissue.

Methods: Methylome profiling was performed using Illumina HumanMethylation450 (485,000 CpG, 96% of known CpG islands), with data available for a total of 390 samples: 16 normal liver, 139 cirrhotic tissue, 8 dysplastic nodules and 227 HCC samples, including 40 eHCC below 2cm.

Results: A phylo-epigenetic tree derived from the euclidean distances between differentially DNA methylated sites (n=421,997) revealed a gradient of methylation changes spanning normal liver, cirrhotic tissue, dysplastic nodules and HCC with closest proximity of dysplasia to HCC. Focusing on promoter regions, we identified epigenetic gatekeeper candidates with an increasing proportion of hypermethylated samples (beta value >0.5) from cirrhotic tissue (<1%), to dysplastic nodules (>=25%), to eHCC (>50%), and confirmed inverse correlation between DNA methylation and gene expression for *TSPYL5*, *KCNA3*, *LDBH* and *SPINT2* (all p<0.001). Unsupervised clustering of genome-wide methylation profiles of cirrhotic tissue identified two clusters, M1 and M2 with 42% and 58% of patients, respectively, which correlate with survival (p<0.05), independent of etiology.

Conclusions: Genome-wide DNA methylation profiles accurately discriminate the different histological stages of human hepatocarcinogenesis. We report novel epigenetic gatekeepers in the transition between dysplastic nodules and eHCC. DNA methylation changes in cirrhotic tissue correlate with clinical outcomes.

INTRODUCTION

Hepatocellular carcinoma (HCC), the most common form of primary liver cancer, typically occurs in patients with chronic liver disease, mainly cirrhosis. Hepatitis B (HBV) and C (HCV) virus infection, alcohol use disorder, and non-alcoholic fatty liver disease are the major etiologies(1). In the United States (US), the liver cancer death rate increased 43% between 2000 and 2016(2). There is a limited understanding of the molecular events leading to malignant transformation in the context of cirrhosis(3). Specifically, little is known about the molecular mechanisms that govern the transition from cirrhosis to dysplasia and early HCC. *TERT* promoter mutations are the only *bone fide* early molecular events in hepatocarcinogenesis(3–6). They occur in 6–17% of dysplastic nodules, with higher prevalence in high-grade dysplasia compared to low-grade dysplasia, increasing to 35% in early HCC below 2 cm in size and up to 60% in larger HCC(4–6).

Aberrant DNA methylation has been reported as an early event in cancer development in many tumor types(7,8). For example, in colon cancer, DNA methylation has been associated with microsatellite instability caused by promoter epigenetic silencing(8). Expression of microRNAs, which control protein expression pathways post-transcriptionally, can also be modified epigenetically(8). Global DNA hypomethylation in colon cancer can cause

chromosomal instability and facilitate malignant transformation(8). In HCC, in addition to molecular signals from the tumor, we have identified transcriptomic signals from the adjacent non-tumoral cirrhotic tissue that correlated with outcomes in HCC patients treated with resection(9). We have also identified a DNA methylation signature from HCC tumor tissue correlated with survival in patients with early tumors treated with surgical resection with prognostic capacities in terms of survival (10). Previous studies have analyzed DNA methylation changes in HCC(11,12), but little is known about genome-wide DNA methylation changes in cirrhosis or preneoplastic lesions in HCC patients. Considering the contribution of DNA methylation to gene expression and chromosomal stability, we hypothesized that DNA methylation changes contribute to malignant transformation in the liver.

Thus, this study aimed at mapping DNA methylation changes across the different histological stages of human hepatocarcinogenesis. To do so, we analyzed genome-wide DNA methylation profiles of 390 tissue specimens including HCC, cirrhosis, dysplasia and normal liver from 248 patients and integrated gene expression data to identify novel epigenetic HCC gatekeepers. We also described how DNA methylation alterations in the adjacent non-tumoral tissue are able to discriminate HCC patients based on their outcomes.

PATIENTS AND METHODS

Human samples

Human samples were obtained from HCC resection specimens from three institutions of the HCC Genomic Consortium as part of the Heptromic initiative(10). This included the IRCCS Istituto Nazionale Tumori in Milan (Italy), the Hospital Clinic in Barcelona (Spain), and Mount Sinai Hospital in New York (USA). The study was approved by the IRB of each institution and patients provided written consent for tissue analyses. DNA extraction was performed as previously reported(10). Genome-wide methylome profiling was conducted on 144 samples using the Illumina Infinium HumanMethylation450 BeadChip array (Illumina, Inc., San Diego, CA) that interrogates more than 485,000 cytosine-phosphateguanine (CpG) sites covering 96% of known CpG islands. New samples analyzed included 6 normal liver tissue (i.e., adjacent histological normal liver from hepatic resections due to a non-tumoral disease such as hemangioma), 130 cirrhosis, and 8 dysplastic nodules (2 high grade and 6 low grade). The 130 cirrhotic tissue samples were adjacent to tumor tissue previously analyzed with this technology (n= 227 HCC samples)(10). In total, this study analyzed DNA methylation profiles of 390 tissue samples (246 previously reported(10) and 144 newly analyzed for this study) from a total of 248 patients (Fig.1). Sample distribution was: 16 without liver disease (i.e., normal liver), 8 dysplastic nodules, 139 cirrhosis and 227 HCC. We also integrated corresponding RNA expression data from 139 cirrhosis, and 205 HCC tissue samples (total of 344 samples) matching our DNA methylation set. Processing of transcriptome data (i.e., normalization, background correction, and filtering) was conducted as previously reported (13).

Data analysis

DNA methylation and gene expression—In order to study differential DNA methylation between normal liver, dysplastic nodules, cirrhotic tissue and HCC tissue, F-scores directly proportional to the variability of beta values between groups were calculated. We then calculated euclidean distances to derive phylo-epigenetic trees(14) and performed differential gene expression analysis and pathway enrichment analysis. Detailed analysis is available in the supplementary methods.

Identification of candidate epigenetic HCC gatekeepers—To reduce the dimensionality of the dataset we restricted this analysis to probes located in the promoter regions (i.e. TSS200, TSS1500, 5'UTR (untranslated region) and 1st Exon regions) and excluded those in sex chromosomes, leaving n=159,892 probes for further analysis. We defined hypermethylation as beta values greater than 0.5 and hypomethylation as beta values less than 0.3 as per our previous publication(10). Probes hypermethylated in 0% of normal livers, less than 1% of cirrhotic tissue, greater than 25% of dysplastic nodules (always including hypermethylation of the high grade nodules) and greater than 50% of early HCC (size below 2 cm) were further included in the analysis. The same thresholds were used to identify hypomethylated probes. Gene expression of those probes was then correlated to methylation in cirrhosis, early HCC and progressed HCC (size ≥2 cm). Gene expression of dysplastic nodules was not available for analysis. Genes were defined as potential gatekeepers if gene expression and DNA methylation were significantly and negatively correlated (Pearson) at a level equal to or less than −0.3. Given our access to genome-wide DNA methylome data, methylation profiles of previously reported aberrantly methylated or mutated genes in HCC and other cancers (10–12,15–18) were also evaluated in our dataset to test for their alteration in our samples (Suppl. Table S1).

The following datasets were used for external validation: First, an external dataset including six liver cirrhosis, eleven dysplastic nodules, nine early HCC, and six progressed HCC was accessed (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE99036). Secondly, the TCGA hepatocellular carcinoma dataset including 50 normal, and 256 HCC samples with available DNA methylation and gene expression data (50 normal, 212 HCC) was accessed via TCGA Wanderer(19).

DNA methylation-based molecular classes in cirrhosis and outcome analysis —To discover DNA-methylation based molecular classes in cirrhosis, we performed unsupervised clustering of all probes of the cirrhotic samples (130 samples, n=421,997 sites, sex chromosomes excluded) via Kmeans algorithm. Only patients with available DNA methylation and gene expression data were included. Kmeans is a method of cluster analysis that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean while minimizing the variance within clusters(20). The optimal number of clusters was calculated by the gap statistic(21) and average silhouette method(22), which measure the quality of clusters based on the cohesion within one cluster when compared to others and both yield 2 clusters as optimal for our dataset. The same 2 clusters were also identified and visualized via uniform manifold approximation (UMAP) (23). Once the patients were partitioned into two groups (M1 and M2) by unsupervised

clustering, we used log-rank test and Kaplan-Meier curves to assess the relationship between the clusters and overall survival (n=125). Overall survival was defined as the time between sampling (i.e. resection) and death of any cause or censored when alive at last follow-up. Cox's regression was used to model the contribution of clinical variables and methylation clusters to overall survival.

In order to identify differences in DNA methylation between the M1 and M2 clusters, we conducted differentially methylated regions analysis using limma and minfi R packages for pairwise comparisons(24). A threshold of at least a log fold change of 0.1 in the beta values was used to include a probe as significant with adjusted p-values < 0.05, which yielded 4,100 CpG sites. We performed an enrichment analysis for gene ontology terms and KEGG pathways using as input the median methylation status per gene. A heatmap showing the differentially methylated beta values was drawn over the cirrhotic samples, highlighting the differences between the two clusters with the heatmap2 package. The groupings based on gene expression of the cirrhotic tissue by Hoshida et al. are also presented in the same figure, alongside with the etiology of liver disease(9). Differential gene expression analysis between M1 and M2 clusters and gene enrichment was conducted using default negative binomial modeling and moderate t-tests using limma and edgeR packages to perform a pairwise comparison between the clusters previously selected with UMAP and Kmeans.

RESULTS

Aberrant DNA methylation delineates the transition between normal liver to early HCC

We first sought to evaluate differences in genome-wide DNA methylation profiles across the histological spectrum of human hepatocarcinogenesis. In detail, genome-wide DNA methylation data was available for 390 samples, including 16 samples of normal liver tissue, 139 cirrhosis, 8 dysplastic nodules (including 6 low-grade and 2 high-grade dysplastic nodules), and 227 HCC tissue (43 with small tumors defined as less than 2 cm, i.e., early HCC, and 184 with tumors larger than 2 cm, i.e., progressed HCC) (Fig. 1). The clinical and demographic characteristics have been previously reported (10), and included mostly male patients (78%) with a median age of 66 years, chronic hepatitis C as the main etiology (47%), and well preserved liver function (Child-Pugh A in 96%). HCC patients had mostly early stage tumors (87% BCLC 0/A), and 35% had microvascular invasion.

Using euclidean distance between the 100 top differentially methylated CpG sites across the different histological stages, we found that normal tissue is homogeneously linearly clustered (Fig. 2a). Closest to normal, cirrhotic tissue (mean euclidean distance to normal 37.54, standard deviation [SD] 5.74) begins to exhibit some heterogeneity. HCC, which is most different from normal (mean euclidean distance to normal 45.53, SD 4.04), has the highest heterogeneity when compared to the other histological lesions. Dysplastic nodules were interspersed within cirrhotic and HCC tissue (mean euclidean distance to normal 40.03, SD 4.92). Replicating the phyloepigenetic tree by resampling and downsizing the cirrhotic and HCC groups to compensate for the disproportionate sample sizes maintained these changes. Specifically, the stepwise increase in euclidean distance for cirrhotic, dysplastic, and HCC tissue from the normal was maintained and the euclidean distances between normal and cirrhotic tissue and normal and HCC tissue were similar after reducing

sample size (Suppl. Fig. S1). Fig. 2b shows the clustering by tissue type for the first two principal components. After observing the significant stepwise methylation changes that differentiate normal liver from diseased tissue (Fig. 2a, Suppl. Fig S1a), we sought to explore specific differences between histological tissue types. The differential methylation analysis between each condition (normal, cirrhotic, dysplastic, and HCC) revealed the highest number of differentially methylated regions (DMRs) between cirrhotic and HCC with ~75% of probes significantly altered (~45% hypermethylated and ~30% hypomethylated), followed by normal versus HCC samples with ~50% of probes altered (Fig. 2c). Comparisons across cirrhotic, dysplastic and normal tissue samples showed the lowest number of probes significantly altered, under 15% of probes. This is similar to gene expression data, where the highest number of altered genes was found when comparing HCC to preneoplastic stages rather than in the comparison within preneoplastic stages or HCC stages (early vs advanced), respectively(25).

We next performed gene set enrichment analysis to identify altered pathways involved in these transitions (Suppl. Fig. S2). Between cirrhosis and dysplastic nodules we found alterations in several metabolic pathways involved in insulin secretion and glycerolipids, and in transcriptional misregulation in cancer. We also found enrichment in pathways involved in cell cycle, cellular senescence, autophagy, MAPK, Wnt and mTOR signaling. Between cirrhosis and HCC, we found upregulation of ferroptosis and increased carbohydrate absorption, as well as glycosaminoglycan biosynthesis and arginine and proline metabolism. These pathways have been associated with cellular proliferation and wound healing through fibroblast and hepatocyte growth factor signaling(26) and initiation of metastasis in breast cancer(27), respectively. An increase in the methylation of genes associated with cancer and cell cycle was observed in dysplastic nodules compared to normal liver tissue. Differences between HCC and dysplastic nodules were dominated by hypermethylation of genes encoding chemokines and Jak-STAT signaling (STAT1–6, IL4–10). This is in line with previous reports showing downregulation of genes related to the immune response, cytokine–cytokine receptor interactions, and the Jak-STAT pathway when comparing preneoplastic lesions and early HCC (25).

Identification of novel candidate epigenetic gatekeeper in hepatocarcinogenesis

Next we sought to identify potential novel epigenetic gatekeepers during the early transformation of HCC. We looked for genes showing a stepwise increase in hypermethylation from cirrhotic tissue (<1% of samples hypermethylated) to dysplastic nodules (≥25% of samples hypermethylated) and small HCC (≥50% of samples hypermethylated) (Fig. 3a). This approach resembles the stepwise increase in mutation frequency of *TERT* promoter, a *bone fide* gatekeeper for HCC development(4). We identified 30 potential epigenetic gatekeepers (29 CpG sites in a total of 25 genes with stepwise increase in hypermethylation, with 4 genes having two CpG sites each, and 5 CpG sites with stepwise hypomethylation) (Suppl. Fig. S3). We further integrated our 30 epigenetic gatekeeper candidates with corresponding RNA expression data (available for 344 patients, representing 89% of the cohort). When we correlated DNA promoter methylation and gene expression we found a significant (all $p<0.001$) inverse correlation for *TSPYL5* ($r=-0.31$), *KCNA3* ($r=-0.33$), *LDHB* ($r=-0.46$) and *SPINT2* ($r=-0.43$) (Fig. 3b,

relationship to clinicopathological variables on Suppl. Table S2; Suppl. Table S3 for remaining non-significant candidates). These genes are involved in TP53, cAMP, serine protease and NADH regulation.

For validation purposes, we utilized two external datasets. In the first dataset, which included six liver cirrhosis, eleven dysplastic nodules, nine early HCC, and six progressed HCC, we were able to confirm the stepwise increase in hypermethylation for the majority of sites of our gatekeepers throughout hepatocarcinogenesis from liver cirrhosis to dysplasia to HCC (Suppl Fig. S4). Secondly, we used the TCGA hepatocellular carcinoma dataset including 50 normal, and 256 HCC samples, and found that all of our 4 gatekeepers were significantly hypermethylated in HCC compared to normal controls. Importantly, we were also able to confirm downregulation of gene expression of our 4 gatekeepers in HCC (n=212) compared to controls (see new Suppl Fig. S5).

To confirm the internal validity of our dataset, we analyzed the DNA methylation status of reported epidrivers in HCC(10–12,18). We found that *APC, CDKL2, DRD4, EFNB2, FAM196A, FOXE3, NEFH, NKX6–2, TBX15*, and *ZIC1* were significantly hypermethylated in early HCC samples compared to cirrhotic controls (Suppl. Table S4). These genes are related to transcription factors, cell cycle and Wnt/beta-catenin pathway, all frequently deregulated in fully established HCC. However, only *NKX6–2* met our strict criteria to define candidate epigenetic HCC gatekeepers with a stepwise increase of hypermethylation across tissue types (Suppl. Fig. S3). We also explored the DNA methylation status of driver genes commonly mutated, amplified, or deleted in HCC(15) and other solid cancers(16,17) (Suppl. Table S4). Among 159 mutated driver genes previously reported in HCC(15), three presented a stepwise increase in hypermethylation from normal and cirrhotic tissue to dysplastic nodules and small HCC (<2cm) (i.e. *COL16A1, COL7A1,* and *FAT4*). Regarding genes commonly mutated, amplified or deleted among other solid cancers(16,17), two of the 252 tested genes showed this pattern (i.e., *HIST1H3B,* and *IKZF1*).

DNA methylation identifies clinically-relevant subclasses of cirrhotic tumor-adjacent tissue

Given the role of gene expression from adjacent tumor tissue to predict risk of HCC and survival after resection (9,28), we speculated about the role of DNA methylation changes in this context. Specifically, we evaluated if there were different molecular subclasses of cirrhotic adjacent non-tumor tissue based on genome-wide DNA methylation alterations. We included 130 patients with DNA methylation from adjacent non-tumoral tissue (94% of patients with cirrhotic tissue samples). The clinical characteristics of this subset are displayed in Table 1. We identified two robust clusters using Kmeans, which we termed M1 and M2. M1 and M2 contained 55 (42%) and 75 (58%) patients, respectively. This partition was deemed optimal by the average silhouette and gap statistic methods. 31,198 CpG sites were found to be hypomethylated in M1 compared to M2 while 107,187 CpG sites were found to be hypermethylated. 293,953 CpG sites did not significantly differ between the groups (Fig. 4a). These two clusters were also obtained when using UMAP (Uniform Manifold Approximation and Projection) (Fig. 4c). The prognostic capability of DNA methylome from tumor-adjacent cirrhotic tissue was evidenced by the Kaplan-Meier plot

including 125 patients with available outcome data (Fig. 4b), which shows that M1 and M2 are predictive of survival in HCC with shorter overall survival in patients classified as M2 compared to M1 class (mean survival 51.9 versus 60.1 months, p<0.001). Clinical characteristics of the M1 and M2 class are displayed in Table 1. Of note, HCV was more common among the M2 class compared to the M1 class (61% versus 22%, p<0.001) and patients in the M2 class had higher levels of serum bilirubin (1.2 versus 0.8 mg/dl, p<0.001), lower platelet counts (122 versus 178 /mm^3, p<0.001), and smaller tumor nodules 3.5 versus 5 cm, p=0.024) (Table 1). In contrast, age, gender, tumor stage according to BCLC classification, histopathological tumor grading, presence of vascular invasion or satellite nodules, serum AFP levels, and liver function according to Child-Pugh classification were not different between the groups. In univariate Cox regression modeling, our DNA methylation class M2 (HR 1.92, 95% CI 1.16–3.16, p=0.01), number of tumor nodules (multiple versus single, HR 2.24, 95% CI 1.39–3.59, p=0.0008), presence of vascular invasion (HR 1.88, 95% CI 1.19–2.98, p=0.006), and levels of serum albumin (HR 0.52, 95% CI 0.34–0.79, p=0.002) were significant predictors for overall survival, while age, gender, tumor size, presence of satellite nodules, AFP, etiology, and independent markers for advanced liver disease (i.e., platelet count and bilirubin) were not (Table 2). Previous analysis by Hoshida et al. revealed that gene expression of adjacent non-tumoral tissue can be predictive of survival in HCC(9). In our study, we found 65% overlap between our two clusters and the Hoshida classes(9), with M2 containing 87% of the Hoshida poor class. Correlation with the previously published Hoshida Score was only moderate (r^2 = 0.4, p<0.001). As expected, Hoshida's 186-gene signature was predictive of survival in our univariate Cox regression model (HR 1.65, 95% CI 1.03–2.66, p=0.03). However, the Hoshida classification lost significance in multivariate Cox regression modeling including clinical predictors (tumor nodules, vascular invasion, albumin levels) and our methylation classes, while the latter revealed a HR of 1.64 to predict survival (95% CI 0.95–2.84, p=0.07) (Table 2). In fact, in multivariate Cox regression modeling directly comparing the 186-gene expression signature and our methylation classes, the latter revealed a significant HR of 1.8 (95% CI 1.02–3.3, p=0.04), while the Hoshida classification was not significant (p=0.35). We determined the spearman correlation between the Beta value of each promoter associated site and RNA expression. We filtered this list to select for sites with significant correlation after adjusting p-values (67 genes). We then calculated the median correlation per gene and found 46 genes in the Hoshida signature with significant negative correlation between DNA methylation and gene expression, suggesting they are regulated epigenetically via promoter region methylation (Suppl. Fig. S6 and Suppl. Table S5). These findings indicate that DNA methylome changes in non-tumoral tissue might be more sensitive to identify HCC patients with poor prognosis compared to gene expression analysis, potentially by capturing indicators of early tumorigenesis as well as liver deterioration.

Integrated differential DNA methylation and gene expression analysis

We found 31,957 CpG sites (FDR <0.05) differentially methylated between clusters M1 and M2 of cirrhotic tissue. Top differentially methylated genes are displayed in Suppl. Table S6. Fig. 5a depicts the landscape of differential DNA methylation between the groups M1 and M2. These results indicate an increase in autophagy and immune activation of pathways for CD4+ T-cell response and differentiation in cluster M2, especially because this cluster

showed a lower survival probability compared to M1. Additionally, cluster M1 showed a decreased methylation of cholesterol metabolism and neuroactive ligand-receptor interactions. This reflects our poor understanding of the upstream of expression signaling since there is no reported overlap in literature between these signatures and liver cancer.

A differential gene expression analysis for clusters M1 and M2 showed 1,332 genes significantly deregulated with false discovery rate below 0.05 (Fig. 5b, Suppl. Table S7). The logFC and FDR values for M2 were higher compared to M1 indicating that the expression of these genes is increased for M2 (Fig. 5b, right panel). This correlates with the increased methylation in M1, demonstrating that the M2 group has a higher expression and a lower methylation profile compared to M1. It is critical to point out that a gene set analysis for the gene expression results associate regulation of IL10 (an inhibitory interleukin) with higher expression in M1, potentially reducing inflammation and cascading into less cell infiltration and differentiation (Fig. 5b, lower panel). Although our objective was to explore the DNA methylation profiles of cirrhotic samples, we correlated the gene expression profile of clusters M1 and M2 to the Hoshida class. The outcome of this comparison draws a clear picture that patients in group M2 were also in the poor outcome group for Hoshida score (Fig. 5b, left panel). Single sample gene set enrichment analysis revealed that cluster M2 was associated with enrichment of immune, inflammatory, and cell signaling and proliferation signatures, including several cancer hallmark signatures, as well as signatures for poor outcome in liver cancer compared to M1 (significant gene sets are displayed Fig. 6, Suppl. Table S8). Meanwhile, cluster M1 was associated with enrichment of metabolic signatures compared to cluster M2 (Fig. 6). These results are concordant with the differential methylation results which indicated decreased methylation in the cholesterol pathways and increased methylation of immune and inflammatory pathways for M1 compared to M2. These results were also compared to other covariates such as the presence of HBV, HCV and alcohol consumption without significant differences between M1 and M2.

DISCUSSION

Our study provides a comprehensive genome-wide DNA methylation analysis across 390 tissue samples from 248 patients covering the whole spectrum of hepatocarcinogenesis from normal liver tissue and cirrhosis, through dysplastic nodules, to small and progressed HCC. We reveal that DNA methylation accurately discriminates different histological stages during human hepatocarcinogenesis and report on four genes as candidate novel epigenetic gatekeepers in the transition from preneoplastic lesions to early HCC. In addition, we reveal that the DNA methylome of non-tumoral adjacent cirrhotic tissue captures prognostic features in patients with HCC.

When considering DNA methylation data, we detected the greatest differences between normal liver tissue and HCC. HCC samples accumulated the highest variability in terms of DNA methylation changes compared to non-malignant tissue. The gradient of DNA methylation changes observed between normal, cirrhosis, dysplasia and HCC prompted us to investigate potential new epigenetic gatekeepers in hepatocarcinogenesis, particularly in the transition between dysplasia and incipient HCC. Gene mutations do not seem to be key drivers of early transformation except for *TERT* promoter mutations, which are the only

bone fide gatekeeper known in HCC(3–6). We identified four genes (*TSPYL5, KCNA3, LDHB,* and *SPINT2*), involved in TP53, cAMP, serine protease and NADH regulation, with a stepwise increase in promoter hypermethylation and a corresponding decrease in gene expression when comparing cirrhotic non-tumoral tissue with dysplastic nodules and early HCC. *TSPYL5*, *LDHB*, and *SPINT2* have been previously reported as hypermethylated in different cancers, including HCC(12,29–32), while *KCNA3* has only been reported as hypermethylated in colorectal cancer(33). In colorectal cancer, lncRNA KCNA3 was found to inhibit tumor growth through down-regulation of YAP1 which activates genes associated with cell proliferation and suppression of apoptosis(33). Low expression of KCNA3 was associated with increased TNM stage, distant metastasis and shorter overall survival(33). In HCC, *KCNA3* expression has recently been associated with overall survival (34). *LDHB* silencing has been reported as an early transformation event in prostate and pancreatic cancers(31). Silencing of LDHB via DNA methylation has been previously reported in HCC(12). In pancreatic cancer, *LDHB* downregulation results in a glycolytic phenotype, which promotes invasion and migration under hypoxic conditions(35). Our results are aligned with a previous study that failed to find significant differences between *LDHB* expression and gender, age, tumor size, tumor number, AFP status, HBV infection, Child-Pugh score, and BCLC stage in HCC patients(36). The mechanism of action and impact of *LDHB* suppression in HCC has not been fully elucidated, thus the impact of this early metabolic adaptation warrants further study. *SPINT2* is involved in the inhibition of several serine proteases and the formation of active hepatocyte growth factor (HGF) by inhibiting it's activator (37). HGF/SF-MET pathway is used by multiple cancer types to sustain invasive growth, protect against apoptosis and maintenance of cancer stem cell-like phenotype(37). It has been reported as hypermethylated and downregulated in glioblastoma, HCC, renal cell carcinoma and melanoma(37). Increased expression of this inhibitor suppressed invasive growth, suggesting that SPINT2 has tumor suppressor activity. *TSPYL5* is a frequently hypermethylated and silenced gene in different cancers(38,39). Experiments in glioma and gastric cancer cell lines have shown that overexpression of *TSPYL5* suppresses growth, suggesting a role as a tumor suppressor gene(39). Similar to our results, *TSPYL5* methylation has been correlated with serum AFP and tumor stage in HCC patients, but not with gender, cirrhosis and tumor size(29). Our data further strengthens the notion that it might play a role in field cancerization and tumorigenesis. Altogether, low expression of these genes has been associated with increased tumor grade, increased metastatic potential and shorter overall survival in HCC (33,36,37,40). We envision two clinical implications of our findings. First, knowledge on key epigenetic alterations in hepatocarcinogenesis would help to better stratify dysplastic nodules and preneoplastic changes in terms of risk assessment for malignant transformation. Interestingly, a recent study has reported an algorithm including sex, AFP, and 3 methylated DNA markers from cfDNA (including *TSPYL5*) to discriminate early stage HCC from controls at risk(41). Secondly, better understanding of the affected pathways might help to develop targeted approaches for primary prevention of HCC in patients with underlying cirrhosis.

Other studies have previously investigated epigenetic changes in premalignant HCC lesions, but unlike our study, most analyzed a small number of candidate genes. A report covering promoter regions of eight genes found a general increase of DNA methylation from cirrhotic

tissue to dysplastic nodules and early HCC, particularly a gradual increase in *SOCS1(18)*. However, most of these genes were already hypermethylated in cirrhosis (29–56%), which questions their potential function as epigenetic gatekeepers during incipient stages of human hepatocarcinogenesis. Similarly, a small study conducted on eight HBV-HCC patients confirmed a progressive increase in hypermethylation through the transition from cirrhotic tissue to dysplastic nodules and HCC with a peak of hypermethylation in early HCC samples(42). In contrast to these studies and others(10,18,42), our study incorporated the largest dataset of genome-wide DNA methylation and gene expression across different tissue types in the transition towards early HCC, thus allowing a more comprehensive approach in the identification of novel gatekeepers. However, one of the limitations of our study is the relatively small sample size of normal tissue and dysplastic nodules. Patients with normal liver tissue and dysplastic nodules typically do not undergo liver resection and specimens are thus difficult to obtain.

The concept of field effect or cancerization field dates back to 1953, when it was used to explain the preconditioning of benign tissue to the development of multiple primary tumors, local recurrence and multifocal areas of precancerous change due to long exposure to an insult(43). In HCC specifically, the prognostic impact of gene expression in the surrounding non-tumoral tissue of HCC patients has been reported before(9). Broadly, cirrhotic tissue can be classified into two classes with good and poor prognosis based on gene expression(9). In contrast, little is known about the impact of epigenetic changes of non-tumoral adjacent tissue on the prognosis of patients with HCC. In this study, we have conducted genome-wide methylome profiling and identified two methylation clusters within the cirrhotic tissue of 139 patients, M1 and M2, by unsupervised clustering. Importantly, these clusters correlated with overall survival. When comparing these classes in terms of clinical parameters, we did not find differences related to tumor characteristics (e.g., size, stage, vascular invasion) and most were related to liver function. Similar to previous studies(44), we confirm progressive alterations in DNA methylation correlated with the degree of hepatic damage. Further, we evidenced increased enrichment of immune, inflammatory, signaling and proliferation in M2 compared to M1. This pattern of dysregulated immune activity in the group with worse survival may be associated with the effects of the immune-mediated cancer field that have been previously described(45). However, we did not find a significant association of our methylation classes with late recurrence of HCC >2 years after resection (i.e. de novo HCC). Given that late recurrences are rare events, the lack of significance might be due to the limited sample size of our cohort.

Aberrant DNA methylation is frequently described as a key alteration involved in cancer development and progression (46,47). Global hypomethylation has been linked to genome instability and loss of imprinting leading to increased propensity to cancer (48). DNA hypermethylation of promoter regions of tumor suppressors drive the transformation from normal tissue to cancer (49). Further, aberrant DNA methylation in normal tissue exposed to environmental insults (e.g., UV light, smoking) has been suggested as an indicator of propensity to many cancer types by creating an epigenetic cancerization field (50,51). However, the evidence connecting aberrant DNA methylation of the adjacent non-tumoral tissue and clinical outcomes is scant. In prostate cancer, the DNA methylome of the tumour microenvironment was found to correlate with the presence and severity of cancer (52). In

colorectal cancer, aberrant DNA methylation of normal mucosa correlates with cancer development(51) and hypermethylation of specific genes in the normal colonic mucosa is predictive of survival (53). Similarly, our results suggest the DNA methylation profile of the cirrhotic tumor-adjacent tissue is a predictor of survival in HCC. Unfortunately, we could not identify adequate datasets with genome-wide DNA methylation data from cirrhotic tissue and survival data to validate our findings to validate our findings in an external dataset. However, our analysis was unsupervised, thus the risk of overfitting is lower compared to other studies that trained their data on a specific outcome. To understand the biological significance of the DNA methylation-based classes of cirrhotic tissue will require functional studies.

In summary, we describe a stepwise range of DNA methylation changes across all histological stages of human hepatocarcinogenesis spanning normal tissue, cirrhosis, dysplasia and HCC. We report on 4 novel gatekeepers in HCC, which show aberrant methylation and gene expression in dysplasia and early HCC when compared to cirrhotic tissue. Our study also provides evidence of the highly heterogeneous epigenetic landscape of molecular alterations present in non-tumoral adjacent cirrhotic tissue.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support:

JvF is supported by the German Research Foundation (FE1746/1-1) and the Clinician Scientist Program at University Medical Center Hamburg. TGL is supported by the Grant for Studies Broadening from the Spanish Association for the Study of the Liver (Asociación Española para el Estudio del Hígado, AEEH).

AJC is supported by the National Cancer Institute Ruth L. Kirschstein NRSA Institutional Research Training Grant (CA078207). BL is supported by the Icahn Institute of Genomics and Multiscale Biology. JML is supported by grants from the European Commission (EC) Horizon 2020 Program (HEPCAR, proposal number 667273-2), the US Department of Defense (CA150272P3), the National Cancer Institute (P30 CA196521), the Samuel Waxman Cancer Research Foundation, the Spanish National Health Institute (MICINN, SAF-2016-76390 and PID2019-105378RB-I00), through a partnership between Cancer Research UK, Fondazione AIRC and Fundación Científica de la Asociacion Española Contra el Cáncer (HUNTER, Ref. C9380/A26813), and by the Generalitat de Catalunya (AGAUR, SGR-1358). AV is supported by the U.S. Department of Defense (CA150272P3).

Competing interest statement:

JML is receiving research support from Bayer HealthCare Pharmaceuticals, Eisai Inc., Bristol Myers Squibb, Boehringer-Ingelheim and Ipsen, and consulting fees from Eli Lilly, Bayer HealthCare Pharmaceuticals, Bristol-Myers Squibb, Eisai Inc, Celsion Corporation, Exelixis, Merck, Ipsen, Genentech, Roche, Glycotest, Leerink Swann LLC, Fortress Biotech, Nucleix, Can-Fite Biopharma, Sirtex, Mina Alpha Ltd and AstraZeneca. AV has received consulting fees from Guidepoint and Fujifilm; consulting fees from Exact Sciences, Gilead, Nucleix and NGM Pharmaceuticals; and research support from Eisai Pharmaceuticals.

Data availability statement:

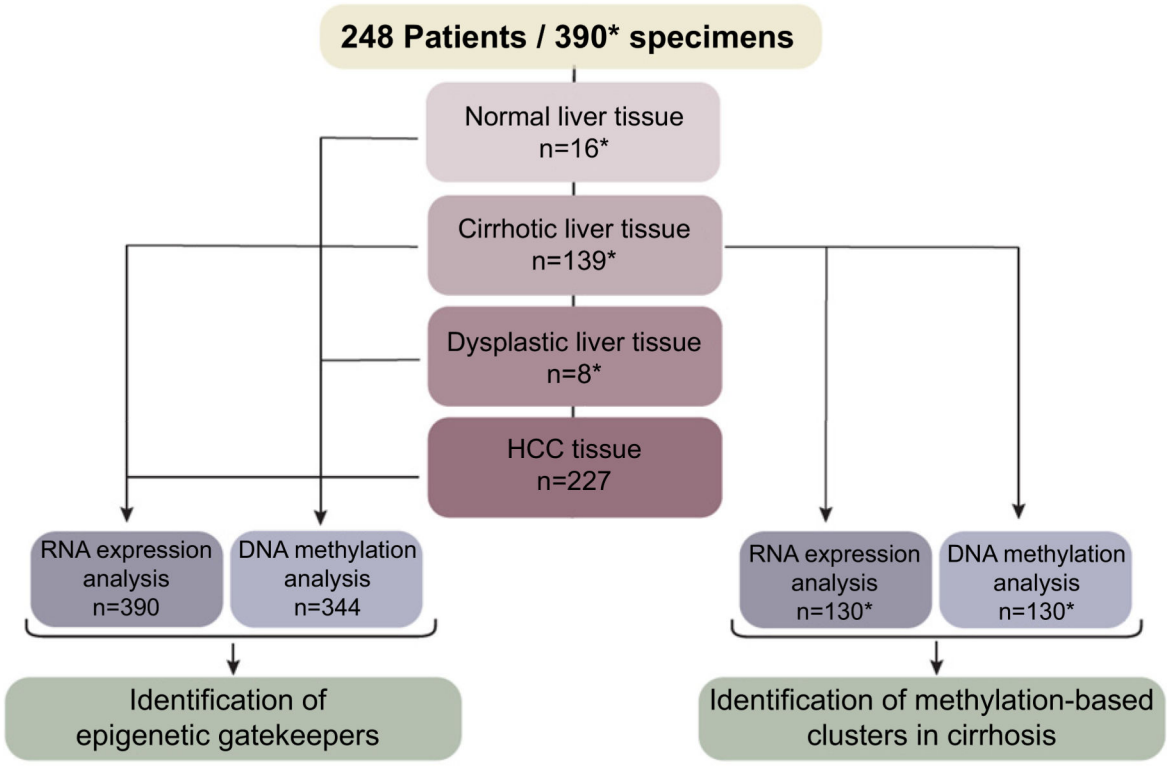
Raw sequencing data will be made publicly available upon acceptance of the article under GEO accession number GSE157973.

REFERENCES

1. Villanueva A Hepatocellular Carcinoma. N. Engl. J. Med 2019;380:1450–1462. [PubMed: 30970190]
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin 2018;68:394–424. [PubMed: 30207593]
3. Craig AJ, von Felden J, Garcia-Lezana T, Sarcognato S, Villanueva A. Tumour evolution in hepatocellular carcinoma. Nat. Rev. Gastroenterol. Hepatol 2020;17:139–152. [PubMed: 31792430]
4. Nault J-C, Ningarhari M, Rebouissou S, Zucman-Rossi J. The role of telomeres and telomerase in cirrhosis and liver cancer. Nat. Rev. Gastroenterol. Hepatol. [Internet]. 2019;Available from: 10.1038/s41575-019-0165-3
5. Torrecilla S, Sia D, Harrington AN, Zhang Z, Cabellos L, Cornella H, et al. Trunk mutational events present minimal intra- and inter-tumoral heterogeneity in hepatocellular carcinoma. J. Hepatol 2017;67:1222–1231. [PubMed: 28843658]
6. Nault JC, Calderaro J, Di Tommaso L, Balabaud C, Zafrani ES, Bioulac-Sage P, et al. Telomerase reverse transcriptase promoter mutation is an early somatic genetic alteration in the transformation of premalignant nodules in hepatocellular carcinoma on cirrhosis. Hepatology. 2014;60:1983–1992. [PubMed: 25123086]
7. Herman JG, Civin CI, Issa JP, Collector MI, Sharkis SJ, Baylin SB. Distinct patterns of inactivation of p15INK4B and p16INK4A characterize the major types of hematological malignancies. Cancer Res 1997;57:837–841. [PubMed: 9041182]
8. Jung G, Hernández-Illán E, Moreira L, Balaguer F, Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. Nat. Rev. Gastroenterol. Hepatol 2020;17:111–130. [PubMed: 31900466]
9. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N. Engl. J. Med 2008;359:1995–2004. [PubMed: 18923165]
10. Villanueva A, Portela A, Sayols S, Battiston C, Hoshida Y, Méndez-González J, et al. DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma. Hepatology. 2015;61:1945–1956. [PubMed: 25645722]
11. Song M-A, Tiirikainen M, Kwee S, Okimoto G, Yu H, Wong LL. Elucidating the landscape of aberrant DNA methylation in hepatocellular carcinoma. PLoS One. 2013;8:e55761. [PubMed: 23437062]
12. Revill K, Wang T, Lachenmayer A, Kojima K, Harrington A, Li J, et al. Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma. Gastroenterology. 2013;145:1424–35.e1–25. [PubMed: 24012984]
13. Villanueva A, Hoshida Y, Battiston C, Tovar V, Sia D, Alsinet C, et al. Combining clinical, pathology, and gene expression data to predict recurrence of hepatocellular carcinoma. Gastroenterology. 2011;140:1501–12.e2. [PubMed: 21320499]
14. Mazor T, Pankov A, Johnson BE, Hong C, Hamilton EG, Bell RJA, et al. DNA Methylation and Somatic Mutations Converge on the Cell Cycle and Define Similar Evolutionary Histories in Brain Tumors. Cancer Cell. 2015;28:307–317. [PubMed: 26373278]
15. Schulze K, Imbeaud S, Letouzé E, Alexandrov LB, Calderaro J, Rebouissou S, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat. Genet 2015;47:505–511. [PubMed: 25822088]
16. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013;339:1546–1558. [PubMed: 23539594]
17. Garraway LA, Lander ES. Lessons from the cancer genome. Cell. 2013;153:17–37. [PubMed: 23540688]
18. Um T-H, Kim H, Oh B-K, Kim MS, Kim KS, Jung G, et al. Aberrant CpG island hypermethylation in dysplastic nodules and early HCC of hepatitis B virus-related human multistep hepatocarcinogenesis. J. Hepatol 2011;54:939–947. [PubMed: 21145824]

19. Díez-Villanueva A, Mallona I, Peinado MA. Wanderer, an interactive viewer to explore DNA methylation and gene expression data in human cancer. Epigenetics Chromatin. 2015;8:22. [PubMed: 26113876]
20. Lloyd S Least squares quantization in PCM [Internet]. IEEE Transactions on Information Theory. 1982;28:129–137. Available from: 10.1109/tit.1982.1056489
21. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic [Internet]. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2001;63:411–423. Available from: 10.1111/1467-9868.00293
22. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis [Internet]. Journal of Computational and Applied Mathematics. 1987;20:53–65. Available from: 10.1016/0377-0427(87)90125-7
23. McInnes L, Healy J, Saul N, Großberger L. UMAP: Uniform Manifold Approximation and Projection [Internet]. Journal of Open Source Software. 2018;3:861. Available from: 10.21105/joss.00861
24. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11:587. [PubMed: 21118553]
25. Wurmbach E, Chen Y-B, Khitrov G, Zhang W, Roayaie S, Schwartz M, et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology. 2007;45:938–947. [PubMed: 17393520]
26. Trowbridge JM, Gallo RL. Dermatan sulfate: new functions from an old glycosaminoglycan. Glycobiology. 2002;12:117R–25R.
27. Elia I, Broekaert D, Christen S, Boon R, Radaelli E, Orth MF, et al. Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. Nat. Commun 2017;8:15267. [PubMed: 28492237]
28. Hoshida Y, Villanueva A, Sangiovanni A, Sole M, Hur C, Andersson KL, et al. Prognostic Gene Expression Signature for Patients With Hepatitis C–Related Early-Stage Cirrhosis. Gastroenterology. 2013;144:1024–1030. [PubMed: 23333348]
29. Shen J, LeFave C, Sirosh I, Siegel AB, Tycko B, Santella RM. Integrative epigenomic and genomic filtering for methylation markers in hepatocellular carcinomas. BMC Med. Genomics. 2015;8:28.
30. Qiu X, Hu B, Huang Y, Deng Y, Wang X, Zheng F. Hypermethylation of ACP1, BMP4, and TSPYL5 in Hepatocellular Carcinoma and Their Potential Clinical Significance. Dig. Dis. Sci 2016;61:149–157. [PubMed: 26386860]
31. Mishra D, Banerjee D. Lactate Dehydrogenases as Metabolic Links between Tumor and Stroma in the Tumor Microenvironment. Cancers [Internet]. 2019;11. Available from: 10.3390/cancers11060750
32. Nishida N, Nagasaka T, Nishimura T, Ikai I, Boland CR, Goel A. Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. Hepatology. 2008;47:908–918. [PubMed: 18161048]
33. Zhong X, Lü M, Wan J, Zhou T, Qin B. Long noncoding RNA kcna3 inhibits the progression of colorectal carcinoma through down-regulating YAP1 expression. Biomed. Pharmacother 2018;107:382–389. [PubMed: 30099342]
34. Tian Z, Wang Z, Chen Y, Qu S, Liu C, Chen F, et al. Bioinformatics Analysis of Prognostic Tumor Microenvironment-Related Genes in the Tumor Microenvironment of Hepatocellular Carcinoma. Med. Sci. Monit 2020;26:e922159. [PubMed: 32231177]
35. Cui J, Quan M, Jiang W, Hu H, Jiao F, Li N, et al. Suppressed expression of LDHB promotes pancreatic cancer progression via inducing glycolytic phenotype. Med. Oncol 2015;32:143. [PubMed: 25807933]
36. Chen R, Zhou X, Yu Z, Liu J, Huang G. Low Expression of LDHB Correlates With Unfavorable Survival in Hepatocellular Carcinoma: Strobe-Compliant Article. Medicine. 2015;94:e1583. [PubMed: 26426634]
37. Kataoka H, Kawaguchi M, Fukushima T, Shimomura T. Hepatocyte growth factor activator inhibitors (HAI-1 and HAI-2): Emerging key players in epithelial integrity and cancer. Pathol. Int 2018;68:145–158. [PubMed: 29431273]

38. Kim T-Y, Zhong S, Fields CR, Kim JH, Robertson KD. Epigenomic profiling reveals novel and frequent targets of aberrant DNA methylation-mediated silencing in malignant glioma. Cancer Res 2006;66:7490–7501. [PubMed: 16885346]
39. Jung Y, Park J, Bang Y-J, Kim T-Y. Gene silencing of TSPYL5 mediated by aberrant promoter methylation in gastric cancers. Lab. Invest 2008;88:153–160. [PubMed: 18059362]
40. Lou C, Du Z, Yang B, Gao Y, Wang Y, Fang S. Aberrant DNA methylation profile of hepatocellular carcinoma and surgically resected margin. Cancer Sci 2009;100:996–1004. [PubMed: 19385975]
41. Chalasani NP, Ramasubramanian TS, Bhattacharya A, Olson MC, Edwards DK V, Roberts LR, et al. A Novel Blood-based Panel of Methylated DNA and Protein Markers for Detection of Early-Stage Hepatocellular Carcinoma. Clin. Gastroenterol. Hepatol. [Internet]. 2020;Available from: 10.1016/j.cgh.2020.08.065
42. Czauderna C, Poplawski A, Castven D, Heilmann-Heimbach S, Odenthal M, Amer W, et al. Landscape of (epi-) genetic alterations during malignant transformation in liver cancer. Zeitschrift für Gastroenterologie. 2019;57:P4–6.
43. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer. 1953;6:963–968. [PubMed: 13094644]
44. Massey V, Cabezas J, Bataller R. Epigenetics in Liver Fibrosis. Semin. Liver Dis 2017;37:219–230. [PubMed: 28847033]
45. Moeini A, Torrecilla S, Tovar V, Montironi C, Andreu-Oller C, Peix J, et al. An Immune Gene Expression Signature Associated With Development of Human Hepatocellular Carcinoma Identifies Mice That Respond to Chemopreventive Agents. Gastroenterology [Internet]. 2019;Available from: 10.1053/j.gastro.2019.07.028
46. Suzuki K, Suzuki I, Leodolter A, Alonso S, Horiuchi S, Yamashita K, et al. Global DNA demethylation in gastrointestinal cancer is age dependent and precedes genomic damage. Cancer Cell. 2006;9:199–207. [PubMed: 16530704]
47. Feinberg AP. The epigenetics of cancer etiology [Internet]. Seminars in Cancer Biology. 2004;14:427–432. Available from: 10.1016/j.semcancer.2004.06.005 [PubMed: 15489135]
48. Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh C-L, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. Cancer Res 2002;62:6442–6446. [PubMed: 12438232]
49. Tsai H-C, Baylin SB. Cancer epigenetics: linking basic biology to clinical medicine. Cell Res 2011;21:502–517. [PubMed: 21321605]
50. Ramírez N, Bandrés E, Navarro A, Pons A, Jansa S, Moreno I, et al. Epigenetic events in normal colonic mucosa surrounding colorectal cancer lesions [Internet]. European Journal of Cancer. 2008;44:2689–2695. Available from: 10.1016/j.ejca.2008.09.004 [PubMed: 18938072]
51. Worthley DL, Whitehall VLJ, Buttenshaw RL, Irahara N, Greco SA, Ramsnes I, et al. DNA methylation within the normal colorectal mucosa is associated with pathway-specific predisposition to cancer. Oncogene 2010;29:1653–1662. [PubMed: 19966864]
52. Lawrence MG, Pidsley R, Niranjan B, Papargiris M, Pereira BA, Richards M, et al. Alterations in the methylome of the stromal tumour microenvironment signal the presence and severity of prostate cancer. Clin. Epigenetics. 2020;12:48. [PubMed: 32188493]
53. Hsu C-H, Hsiao C-W, Sun C-A, Wu W-C, Yang T, Hu J-M, et al. Novel methylation gene panel in adjacent normal tissues predicts poor prognosis of colorectal cancer in Taiwan. World J. Gastroenterol 2020;26:154–167. [PubMed: 31988582]



*Includes a total of 144 new samples. Per category: 6 normal liver, 130 cirrhotic, 8 dysplastic.

Figure 1.
Flow chart of samples and analysis.

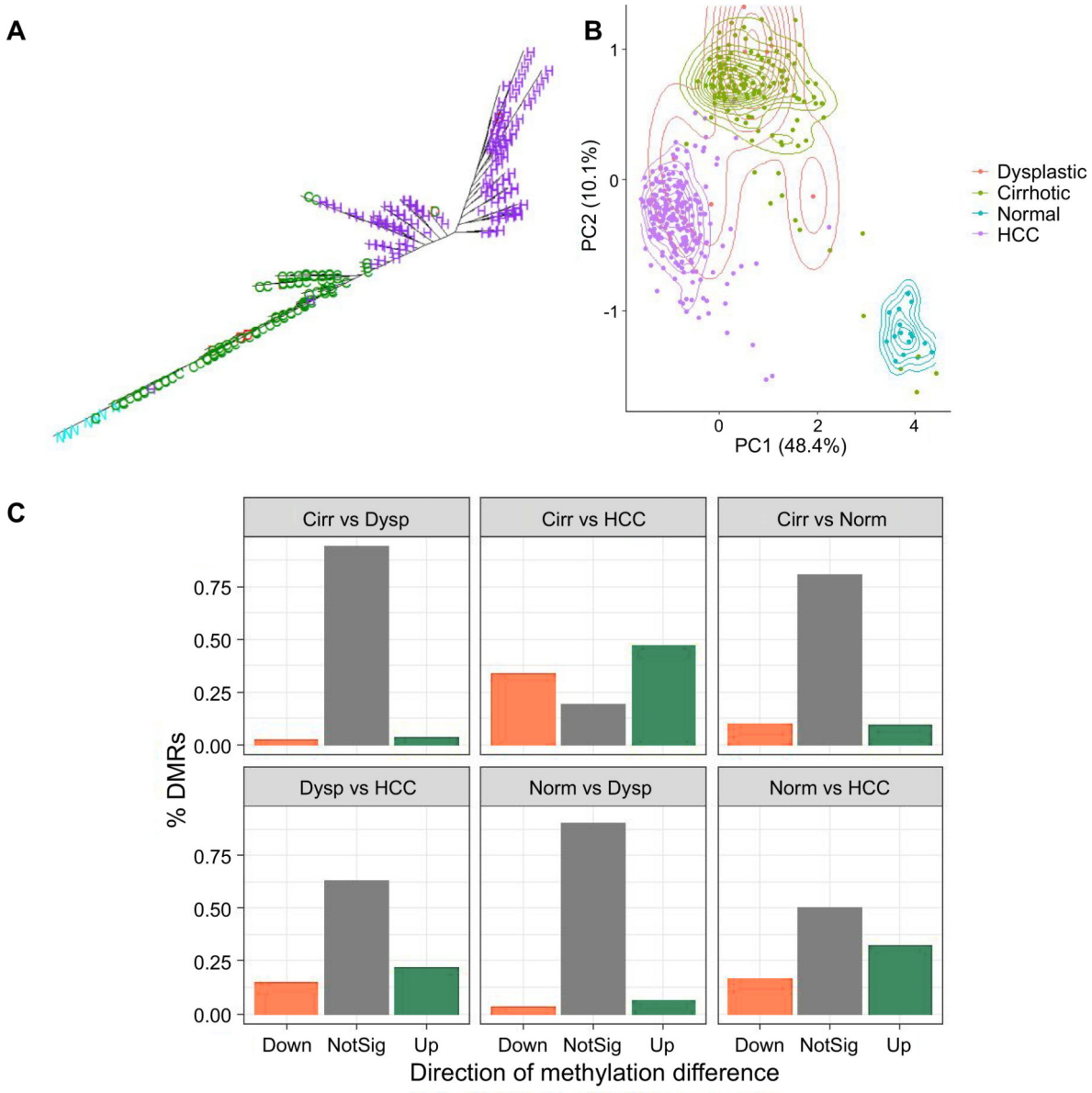


Figure 2. Differential methylation across tissue types.

(**A**) Phyloepigenetic tree showing gradients of methylation changes that span normal tissue to hepatocellular carcinoma (HCC). (**B**) Principal component plot of normal, cirrhosis, dysplasia and HCC methylation. (**C**) Differentially methylated regions (DMRs) for the comparisons between each tissue type.

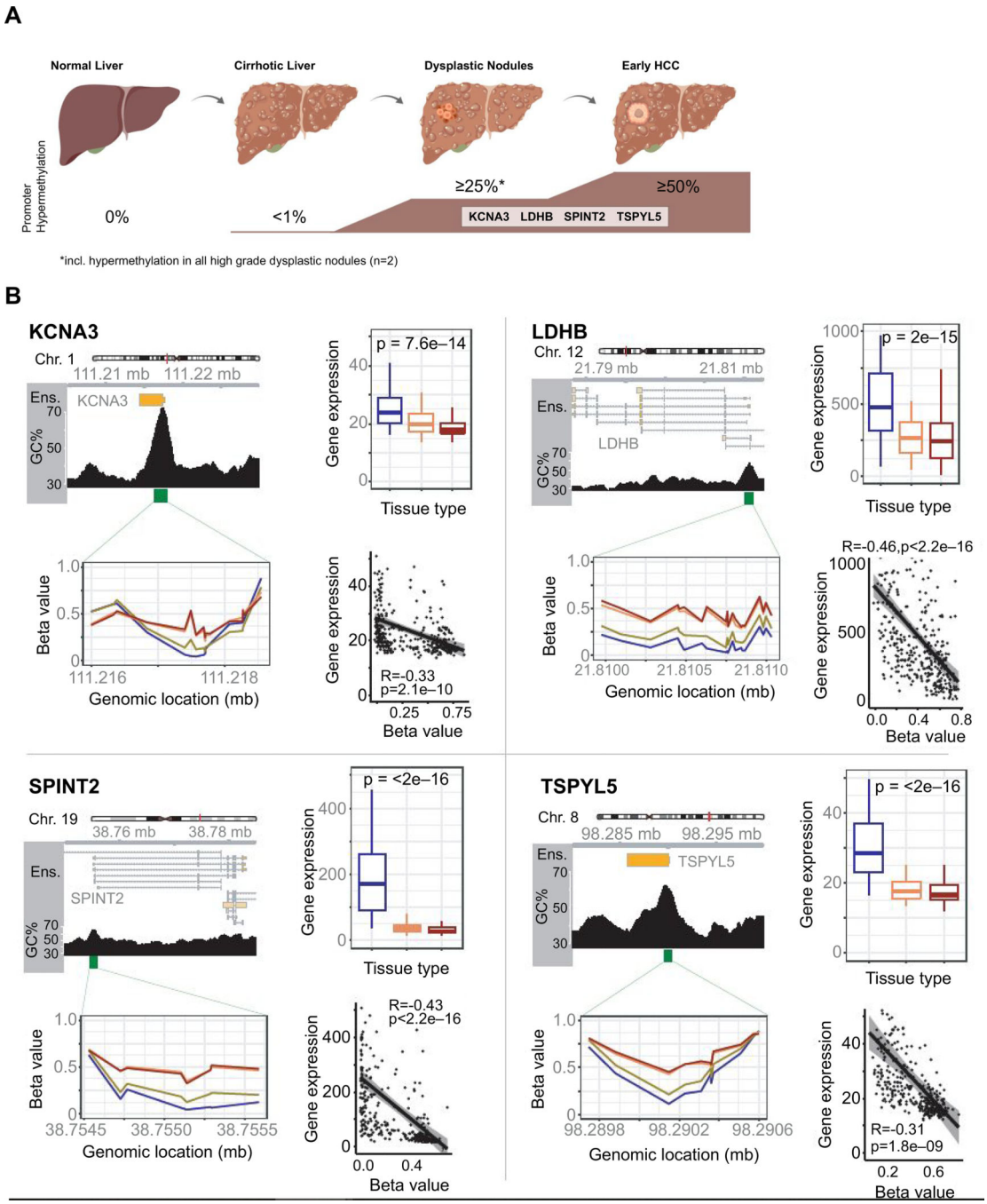


Figure 3. Epigenetic gatekeepers.

(**A**) Selection criteria for gatekeeper discovery. (**B**) Displayed are details for all four gatekeepers of hepatocellular carcinoma as indicated: KCNA3, LDHB, SPINT2, and TSPYL5. Top left panel: Genomic location. Bottom left panel: Beta values for the CpG promoter regions in cirrhotic, dysplastic, early HCC (eHCC) and progressed HCC (pHCC). Top right panel: RNA expression in cirrhotic, eHCC and pHCC tissue. Bottom right panel: Scatter plot for correlation of RNA expression and methylation (beta value).

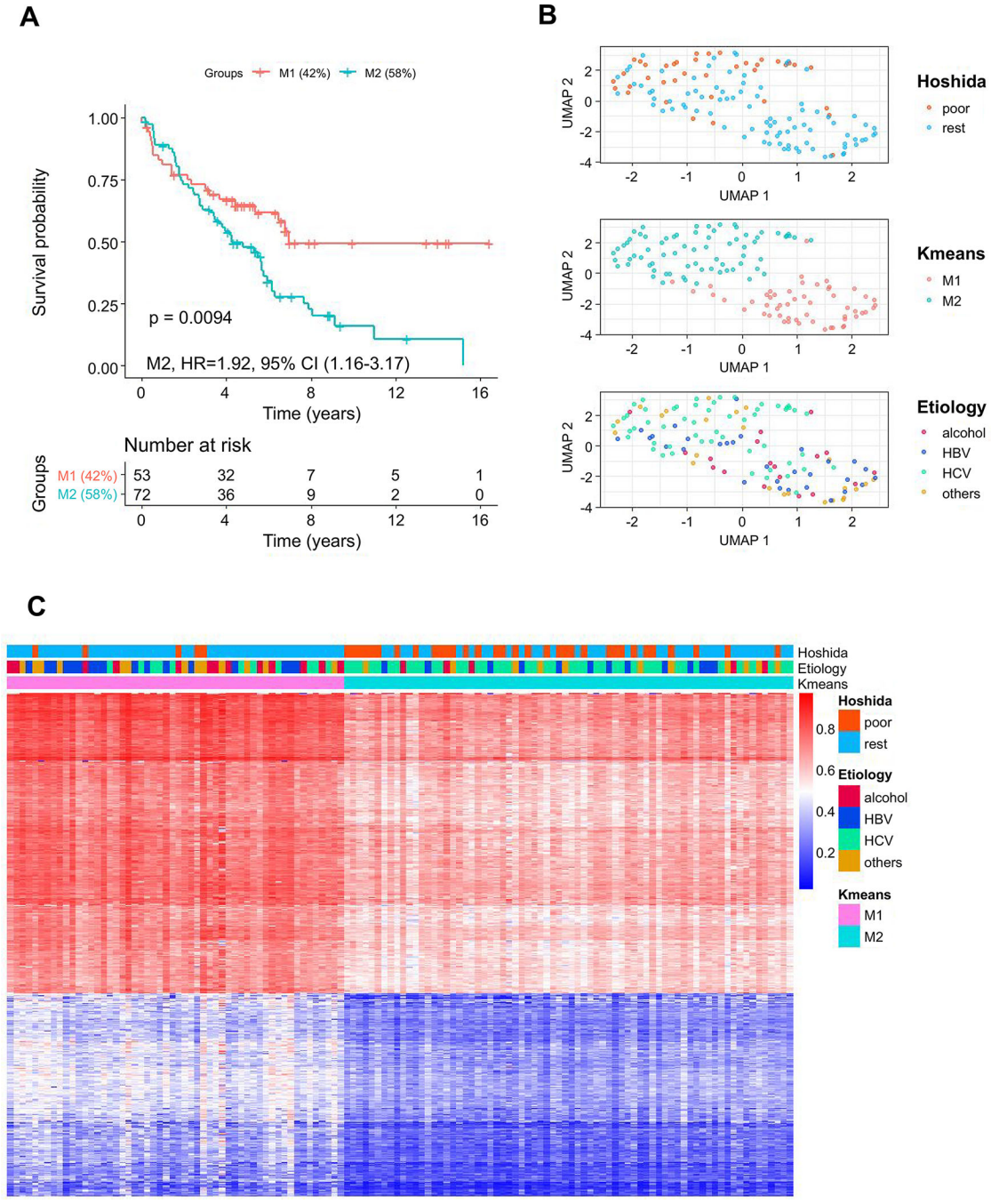


Figure 4. Clustering of cirrhotic tissue based on methylome.

(**A**) Heatmap of cirrhotic tissue showing Kmeans clustering, Hoshida groups, and etiology. (**B**) Kaplan-Meier plot showing significantly different overall survival for Kmeans M1 vs M2. (**C**) UMAP plots with superimposed Hoshida score, Kmeans clusters and etiology.

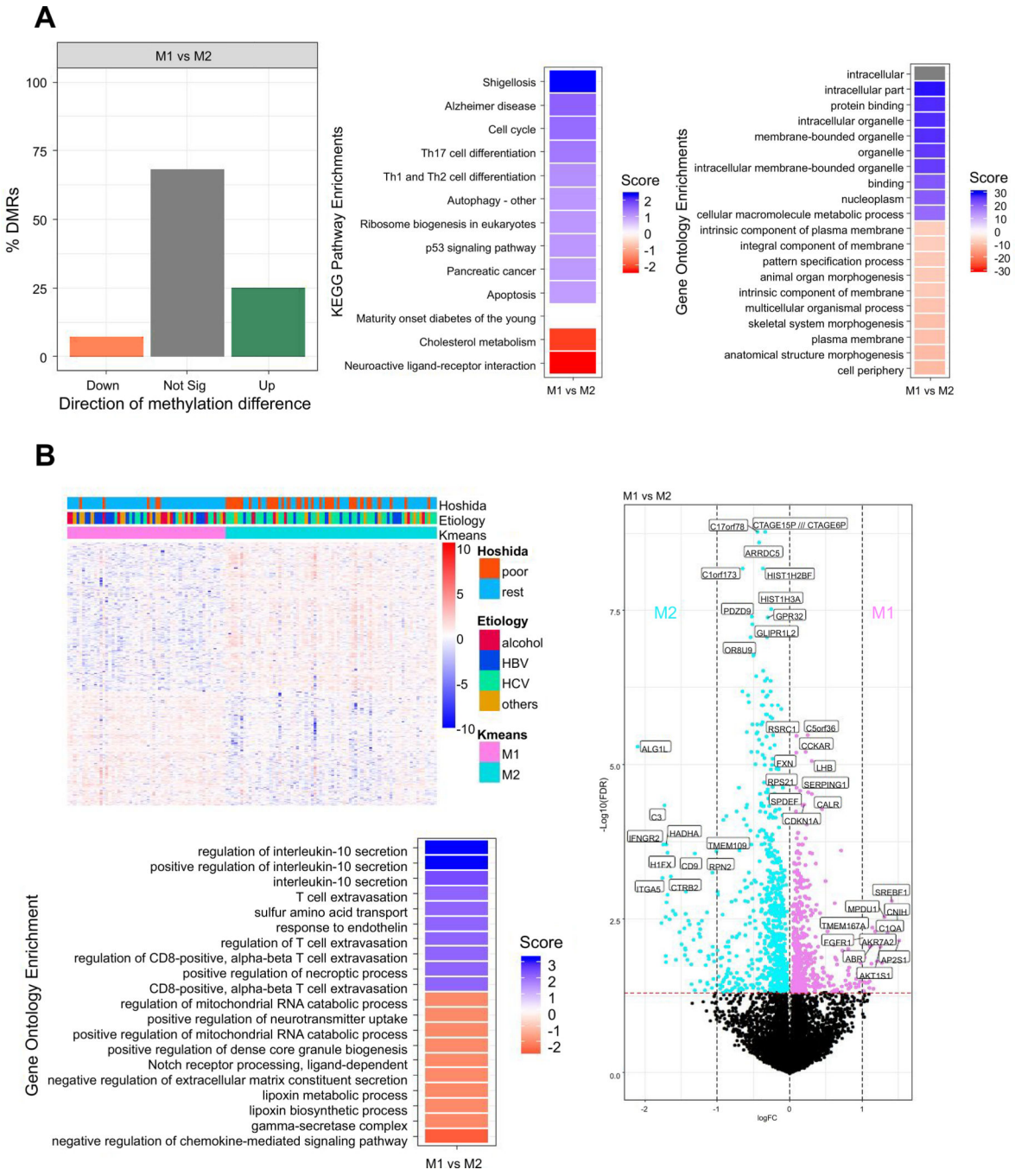


Figure 5. Differential DNA methylation and gene expression analysis by clusters M1 and M2.

(**A**) Left: Differentially methylated regions for M1 vsM2. Middle: Top 10 KEGG pathways for DMRs. Right: Top 10 Gene ontology enrichment for DMRs (Upregulated pathways in M1 compared to M2 are displayed in blue, downregulated pathways in red). (**B**) Top left: Heatmap of 1,300 differentially expressed genes in M1 vs M2, showing grouping by Kmeans clusters, Hoshida score and etiology. Bottom left: Gene ontology enrichment for M1 vs M2 differential gene expression. Right: Volcano plot showing the significantly altered differentially expressed genes.

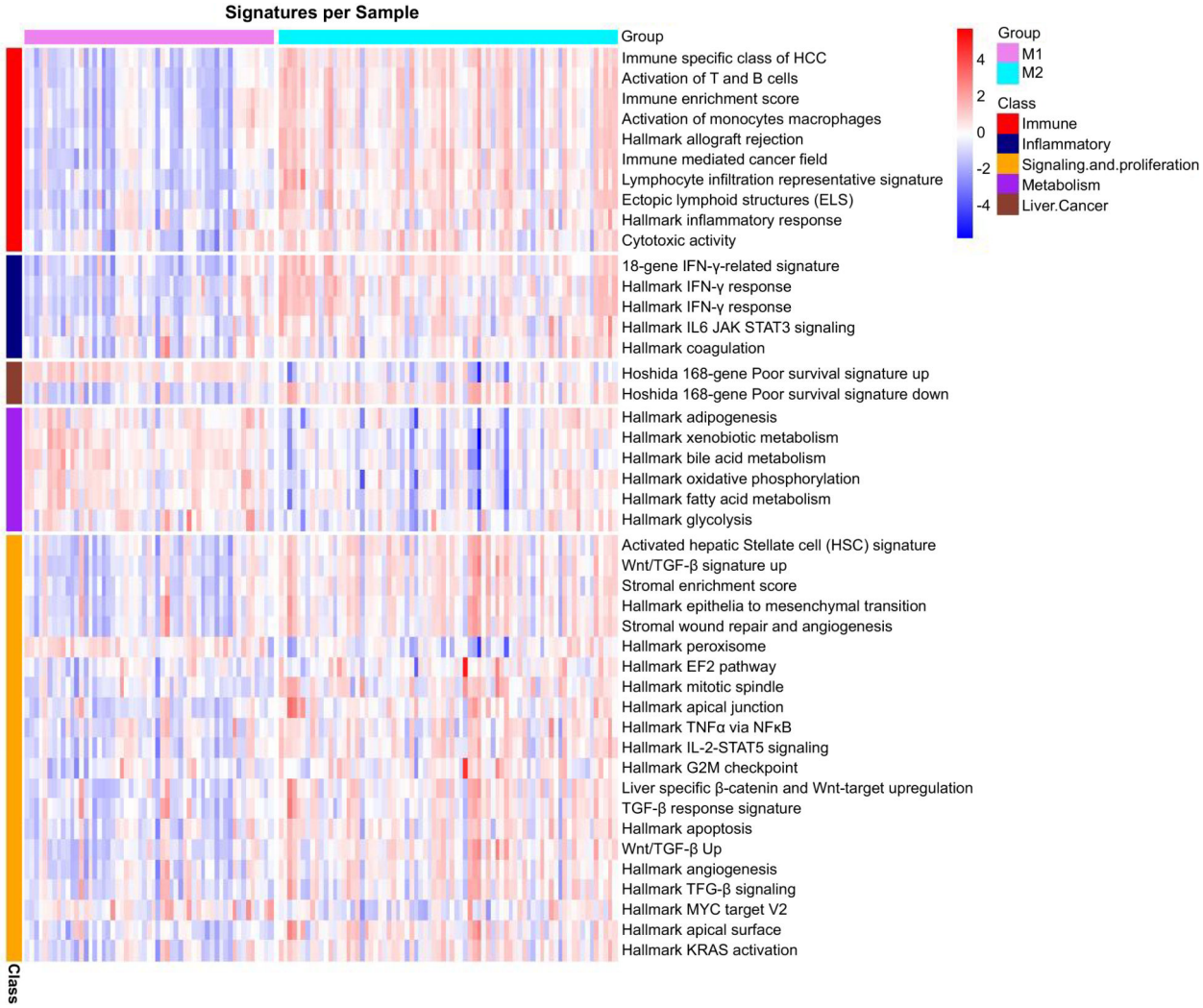


Figure 6. Single sample gene set enrichment analysis, differential enrichment of M1 and M2.

Differentially enriched signatures in the Hallmark database and signatures associated with the immune-mediated cancer field. M2 showed significant enrichment for immune, inflammatory, liver cancer, signaling and proliferation signatures. In comparison, M1 showed significant enrichment for signatures linked to metabolic pathways.

Table 1.

Demographic and clinical characteristics of HCC patients with matched cirrhotic tissue

Characteristic	Overall, N = 130	Statistic	M1, N = 55	M2, N = 75	p-value[1]
Sex					0.074
Female	27 (21%)	n (%)	7 (13%)	20 (28%)	
Male	99 (79%)	n (%)	47 (87%)	52 (72%)	
Etiology					**<0.001**
Alcohol	18 (15%)	n (%)	13 (25%)	5 (6.9%)	
HBV	30 (24%)	n (%)	17 (33%)	13 (18%)	
HCV	55 (45%)	n (%)	11 (22%)	44 (61%)	
Others	20 (16%)	n (%)	10 (20%)	10 (14%)	
Age	66 (61, 72)	median (IQR)	65 (59, 71)	66 (62, 72)	0.2
Size	4.0 (2.8, 7.0)	median (IQR)	5.0 (3.0, 8.0)	3.5 (2.7, 5.0)	**0.024**
BCLC					0.10
BCLC 0	6 (4.8%)	n (%)	3 (5.7%)	3 (4.2%)	
BCLC A	94 (75%)	n (%)	42 (79%)	52 (72%)	
BCLC B	17 (14%)	n (%)	3 (5.7%)	14 (19%)	
BCLC C	8 (6.4%)	n (%)	5 (9.4%)	3 (4.2%)	
Degree of tumor differentiation					0.6
Well	15 (15%)	n (%)	5 (12%)	10 (16%)	
Moderate	61 (59%)	n (%)	24 (57%)	37 (61%)	
Poor	27 (26%)	n (%)	13 (31%)	14 (23%)	
Child-Pugh score					>0.9
A	124 (99%)	n (%)	53 (100%)	71 (99%)	
B	1 (0.8%)	n (%)	0 (0%)	1 (1.4%)	
Microvascular invasion					0.2
Absent	76 (61%)	n (%)	27 (52%)	49 (68%)	
Macro	8 (6.5%)	n (%)	5 (9.6%)	3 (4.2%)	
Micro	40 (32%)	n (%)	20 (38%)	20 (28%)	
Satellite nodules	36 (28%)	n (%)	16 (30%)	20 (27%)	>0.9
Bilirubin (mg/dL)	1.00 (0.70, 1.30)	median (IQR)	0.80 (0.60, 1.10)	1.20 (0.80, 1.50)	**<0.001**
Albumin (g/L)	4.10 (3.70, 4.40)	median (IQR)	4.20 (3.80, 4.50)	4.00 (3.60, 4.40)	0.065
Platelets (100,000/mm^3)	147 (102, 189)	median (IQR)	178 (145, 226)	122 (91, 167)	**<0.001**
AFP (mg/dL)	12 (4, 158)	median (IQR)	12 (3, 268)	13 (5, 65)	0.4

[1] Statistical tests performed: chi-square test of independence; Wilcoxon rank-sum test; Fisher's exact test

Table 2.

Univariate and multivariate outcomes analysis (overall survival)

	Univariate analysis		Multivariate analysis	
Variable	**HR (95% CI)**	**P-value**	**HR (95% CI)**	**P-value**
Sex (male)	0.65 (0.39 – 1.1)	0.11		
Age	1.01 (0.97 – 1.02)	0.82		
Number of nodules (multiple)	2.24 (1.39 – 3.59)	<0.001	1.73 (1.05 – 2.88)	0.03
Tumor size (cm)	1.03 (0.97 – 1.08)	0.26		
Vascular invasion	1.88 (1.19 – 2.98)	0.006	1.82 (1.12 – 2.94)	0.01
Satellite nodules	1.34 (0.80 – 2.23)	0.25		
Bilirubin (mg/dL)	1.52 (0.98 – 2.36)	0.06		
Albumin (g/dL)	0.52 (0.34 – 0.79)	0.002	0.56 (0.36 – 0.89)	0.01
Platelet count	0.99 (0.99 – 1.01)	0.24		
AFP (mg/dL)	1.03 (0.97 – 1.10)	0.22		
DNA methylation class (M2)	1.92 (1.16 – 3.16)	0.01	1.64 (0.95 – 2.84)	0.07
Hoshida 186 gene signature	1.65 (1.03 – 2.66)	0.03	1.07 (0.64 – 1.81)	0.77

CI, confidence interval; HR, hazard ratio