



Published in final edited form as:

Hepatology. 2015 June ; 61(6): 1945–1956. doi:10.1002/hep.27732.

DNA Methylation-Based Prognosis and Epdrivers in Hepatocellular Carcinoma

Augusto Villanueva^{1,2,3,*}, Anna Portela^{4,*}, Sergi Sayols^{4,5}, Carlo Battiston⁶, Yujin Hoshida¹, Jesús Méndez-González⁴, Sandrine Imbeaud^{7,8}, Eric Letouzé⁹, Virginia Hernandez-Gea³, Helena Cornella³, Roser Pinyol³, Manel Solé³, Josep Fuster³, Jessica Zucman-Rossi^{7,8}, Vincenzo Mazzaferro⁶, Manel Esteller^{4,10,11,}, Josep M. Llovet^{1,3,11,**} on behalf of the HEPROMIC Consortium**

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

²Division of Hematology and Medical Oncology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY

³Barcelona-Clínic Liver Cancer Group (Liver Cancer Translational Research Laboratory, Liver Unit, Pathology Department, Surgery Department), Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), CIBEREHD, Hospital Clínic de Barcelona, Universitat de Barcelona (UB), Barcelona, Spain

⁴Cancer Epigenetics and Biology Program, Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Spain

⁵Institute of Molecular Biology, Mainz, Germany

⁶Gastrointestinal Surgery and Liver Transplantation Unit, National Cancer Institute, Milan, Italy

⁷Inserm, UMR-1162, Génomique fonctionnelle des tumeurs solides, IUH, Paris, France

⁸Université Paris Descartes; Université Paris Diderot, Université Paris 13, Labex Immunooncology, Sorbonne Paris Cité, Faculté de Médecine, Paris, France

⁹Programme Cartes d'Identité des Tumeurs, Ligue Nationale Contre le Cancer, Paris, France

¹⁰Department of Physiological Sciences II, School of Medicine, University of Barcelona, Barcelona, Spain

¹¹Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain.

Abstract

Epigenetic deregulation has emerged as a driver in human malignancies. There is no clear understanding of the epigenetic alterations in hepatocellular carcinoma (HCC) and of the potential role of DNA methylation markers as prognostic biomarkers. Analysis of tumor tissue from 304 patients with HCC treated with surgical resection allowed us to generate a methylation-based

*These authors equally contributed to this study.

Author names in bold designate shared co-first authorship.

**Shared senior authorship.

prognostic signature using a training-validation scheme. Methylome profiling was done with the Illumina HumanMethylation450 array (Illumina, Inc., San Diego, CA), which covers 96% of known cytosine-phosphate-guanine (CpG) islands and 485,000 CpG, and transcriptome profiling was performed with Affymetrix Human Genome U219 Plate (Affymetrix, Inc., Santa Clara, CA) and miRNA Chip 2.0. Random survival forests enabled us to generate a methylation signature based on 36 methylation probes. We computed a risk score of mortality for each individual that accurately discriminated patient survival both in the training (221 patients; 47% hepatitis C-related HCC) and validation sets (n = 83; 47% alcohol-related HCC). This signature correlated with known predictors of poor outcome and retained independent prognostic capacity of survival along with multinodularity and platelet count. The subset of patients identified by this signature was enriched in the molecular subclass of proliferation with progenitor cell features. The study confirmed a high prevalence of genes known to be deregulated by aberrant methylation in HCC (e.g., *Ras association [RalGDS/AF-6] domain family member 1*, *insulin-like growth factor 2*, and *adenomatous polyposis coli*) and other solid tumors (e.g., *NOTCH3*) and describes potential candidate epidrivers (e.g., *septin 9* and *ephrin B2*).

Conclusions: A validated signature of 36 DNA methylation markers accurately predicts poor survival in patients with HCC. Patients with this methylation profile harbor messenger RNA-based signatures indicating tumors with progenitor cell features.

Liver cancer represents a major health problem, being the second-leading cause of cancer death worldwide.¹ The disease burden of this malignancy continues to grow and it is a leading cause of disability.² Surgical resection, liver transplantation, and local ablation are the recommended treatment options for early hepatocellular carcinoma (HCC), but can only be applied to approximately 30% of patients in the West. Current clinical practice guidelines recommend resection in patients with single tumors and well-preserved liver function.^{3,4} Even in these cases, tumor recurrence occurs in up to 70% of patients at 5 years, and no adjuvant therapy is presently available. Recent studies demonstrate how messenger RNA (mRNA)-based gene signatures obtained from HCC resection specimens and biopsies improve prognostic performance of conventional clinical and pathological variables.^{5–7} Direct translation of these prognostic signatures into clinical decision making has not yet occurred. Further understanding of tumor biology in HCC is needed to optimize prognostic accuracy and improve trial design and clinical management.

DNA methylation regulates cell differentiation and participates in tumorigenesis.⁸ Global loss of DNA methylation is a hallmark of human cancer, also characterized by selective hypermethylation confined to gene promoters. Solid evidence indicates that epigenetic marks could be used as prognostic and predictive biomarkers in oncology.⁹ Analysis of methylomes enabled classification of colorectal carcinoma patients based on their prognosis.¹⁰ In HCC, there is no clear understanding of the methylome and epidrivers, and few studies have comprehensively evaluated methylation biomarkers using high-throughput platforms.¹¹ Herein, using genome-wide methylation profiling, we introduce and validate a 36-probe methylation signature able to accurately predict survival in HCC patients, as well as describe the landscape of aberrant methylation of key potential tumors suppressors and oncogenes in this cancer.

Materials and Methods

Human Samples and Molecular Profiling.

Initially, the study included samples from 331 surgically resected HCC and 19 nontumor tissues, including 9 cirrhosis and 10 normal livers. The training set (Heptomic data set, $n = 248$; flow chart in Supporting Fig. 1) were samples obtained from two institutions of the HCC Genomic Consortium: IRCCS Istituto Nazionale Tumori (Milan, Italy; $n = 217$) and Hospital Clínic (Barcelona, Spain; $n = 31$). All samples included in this study were fresh-frozen. For RNA and DNA extraction, we used the Qiagen RNeasy Mini (500 ng of total RNA at a concentration of 100 ng/ μ L; Qiagen, Hilden, Germany) and Invitrogen ChargeSwitch genomic DNA Mini Tissue (1 μ g of total DNA at a concentration of 100 ng/ μ L; Invitrogen, Carlsbad, CA) kits, respectively. Median sample storage time from collection to DNA/RNA extraction was 7 years. RNA profiling was conducted on 228 HCC and 168 nontumor liver adjacent cirrhotic tissues using the Affymetrix Human Genome U219 Array Plate (Affymetrix, Inc., Santa Clara, CA), which is able to interrogate more than 20,000 mapped genes. After hybridization, only one tumor sample was discarded for transcriptome analysis owing to poor quality.

Methylome profiling was performed on 248 samples with the Illumina Infinium HumanMethylation450 BeadChip array (Illumina, Inc., San Diego, CA) that interrogates more than 485,000 cytosine-phosphate-guanine (CpG) sites covering 96% of known CpG islands.¹² Of the 248 HCC samples from the training set, 221 HCC qualified for final analyses after quality filtering (i.e., less than 5% of the probes incorrectly interrogated; $P < 0.01$). In 205 patients, there was information on both transcriptome and methylation profiling. For validation of the methylation-based signature, we analyzed data from a cohort of 83 HCC patients treated with resection in two French institutions (Bordeaux and Créteil hospitals), profiled with the same technology as the training set. The institutional review boards of the participating centers approved the study.

Pyrosequencing Validation.

DNA methylation was evaluated with a pyrosequencing assay in a subset of samples previously analyzed by the Illumina Infinium Human Methylation450 array. A minimum of 500 ng of DNA was converted using the EZ DNA Methylation-Gold (Zymo Research Corporation, Irvine, CA) bisulfite conversion kit, following the manufacturer's recommendations. Specific sets of primers for polymerase chain reaction (PCR) amplification and sequencing were designed using specific software (PyroMark assay design, version 2.0.01.15). These are summarized in Supporting Table 1. Primer sequences were designed, when possible, to hybridize with CpG-free sites to ensure methylation-independent amplification. PCR was performed under standard conditions with biotinylated primers, and the PyroMark Vacuum Prep Tool (Biotage AB, Uppsala, Sweden) was used to prepare single-stranded PCR products, according to the manufacturer's instructions. PCR products were observed at 2% agarose gels before pyrosequencing. Reactions were performed in a PyroMark Q96 System (version 2.0.6; Qiagen) using appropriate reagents and protocols.

Data Analysis.

To study differential methylation between HCC and normal liver tissue, probes containing single-nucleotide polymorphisms (SNPs) or located on sex chromosomes were eliminated, leaving 434,728 probes for analysis. Only those with a high signal quality ($P < 0.01$) were considered. Probes hypomethylated ($B < 0.33$) in at least 90% of normal liver samples and hypermethylated ($B > 0.5$) in at least 5% of tumors (31,052 CpG sites) and probes hypermethylated ($B > 0.5$) in at least 90% of normal liver and hypomethylated ($B < 0.33$) in at least 5% of tumors were selected (68,086 CpG sites). The beta value (B) is used to estimate the methylation level of the CpG locus using the ratio of intensities between methylated and unmethylated alleles. For prognosis purposes, analyses focused only on CpG islands located in TSS1500, TSS200, 5' UTR (untranslated region), and 1stExon ($n = 84,448$) shown to be differentially methylated in tumors versus normal liver according to predefined criteria ($n = 11,307$).¹³ This enabled us to select for those methylation markers primarily deregulated in cancer tissues.

In order to discover potential candidate HCC epidrivers, an F score directly proportional to intergroup variability (normal liver versus HCC) and inversely proportional to normal sample intragroup variability was calculated for all probes located in promoters and CpG islands. The first 500 ranked probes were grouped per gene. Genes with more than five hits were further studied as epidrivers. Previously reported epigenetically deregulated genes were also analyzed, considering the mean of all probes located in the TSS200 region, which shows the best correlation with expression. For those samples without TSS200 probes in the array, TSS1500 were used instead, as the closest marker available in the array.

We generate a methylome signature able to predict risk of death for each individual measured by a mortality index (MI). The prediction signature was generated using the random survival forest (RSF) method developed for variable selection.¹³ RSF is a method for prediction and variable selection for right-censored survival and competing risk data by growing survival trees to estimate a cumulative hazard function (CHF), which derives from each tree of the RSF. Input variables were the filtered 11,307 CpG sites. First, probes were randomly split in six sets of 2,000 or less probes to fit a model each using the RSF method by growing 1,000 survival trees. Variable importance (VIMP) scores were computed for all the probes used to grow the trees, and the 350 most informative from each model were selected to fit a new model using the same RSF method. Once the model was built, we removed redundant and “noisy” variables by generating incremental RSF models adding one more variable, ranked by the VIMP calculated in the previous model, at each step, getting an estimate of the error as a measure of how much misclassification increases, or decreases, for a new test case if a given variable was not available for that case. The final set of selected variables (36 probes) was used to build the model having the lowest error rate using only the two most significant decimals of the obtained error rate. Thus, the methylation signature was based upon those 36 probes. The predictive accuracy of the 36-signature was examined by bootstrap cross-validation.¹³ This allows estimating the prediction error curves of different survival models on 100 bootstrap samples and comparing the performance of the different models (i.e., Cox's and RSF) with a reference (estimation for the whole group without splitting based on the prediction variables; Supporting Fig. 2).

From the last RSF model, we generated a DNA methylation-based mortality risk score (i.e., MI; range, 36–360) for every individual computed as a sum of the CHF for each patient evaluated at a set of distinct time points weighed by the number of individuals at risk at the different time points. High MI correlates with higher risk of death. Using this index, we were able to split our training set into two risk groups based on percentiles, selecting the patients with the 20% highest levels (high risk [MI >230] and nonhigh risk [MI <230]) and test the variables selected as a prediction signature by confronting the high- and low-risk groups. To test the significance of the MI index, the same MI value used to split the training set was also used for the validation set. In other words, for the validation, we did not use percentiles and split patients using the actual same value of the MI generated in the training set.

Processing of transcriptome data (i.e., normalization, background correction, and filtering) was conducted as previously reported.⁵ Prediction of liver cancer mRNA-based signatures in the training set was performed using the nearest template prediction method, as implemented in the specific module of Gene Pattern software.¹⁴ All mRNA signatures analyzed were already reported, being deposited in the Molecular Signature Database (www.broadinstitute.org/gsea/msigdb). To provide biological insight on samples with high mortality risk score (>230; percentile, 80), we used the Gene Set Enrichment Analysis (GSEA), as implemented in Gene Pattern. Gene ontology by PANTHER, INTERPRO, and KEGG pathways enrichment analysis was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID; v6.7).

Outcome analyses were framed within current guidelines from the Progress Partnership for prognostic factor research.¹⁵ Kaplan-Meier's plots and Cox's regression were used to assess association with outcome of the methylation signature and clinical variables. Overall survival was defined as the time between surgical resection and death of any cause or last follow-up, whereas cancer-related survival (validation set) was the time between resection to HCC-related death, as previously defined.¹⁶ Time to recurrence is the time between resection and the radiological evidence of first tumor recurrence. Several clinical variables previously reported as outcome predictors in HCC were included in the analysis: gender; age; etiology of liver disease; tumor size; serum albumin; serum bilirubin; platelet count; serum alpha-fetoprotein (AFP); Barcelona Clinic Liver Cancer (BCLC) stage¹⁷; microvascular invasion; satellites; and degree of differentiation. Variables with a *P* value less than 0.05 in the univariate analysis (log-rank test) were separately evaluated in a multivariate Cox's model aimed to identify independent predictors of survival and recurrence. Missing values were less than 10% for each variable included in the Cox's model. Analyses were performed using the GenePattern analytical toolkit (www.broad.mit.edu/cancer/software/genepattern/) and the R statistical package (www.r-project.org).

Results

Methylation Signature Predicts Survival in HCC.

RSFs enabled us to generate a methylation-based signature that accurately discriminated patients based on their survival (training set; *n* = 221). Table 1 summarizes the clinical characteristics of this cohort, that mostly included males (172; 78%), with a median of

66 years of age, viral-related HCC (hepatitis C virus [HCV]: 101 [47%]; hepatitis B virus [HBV]: 44 [20%]), and a median tumor size of 3.5 cm. Patients had predominantly uninodular disease (166; 75%) and no microvascular invasion (142; 65%) or satellites (158; 71%), being the majority at early clinical stages (BCLC O/A: 191 [87%]) with a median follow-up of 48.5 months. Using the array data, we generated a methylation-based signature formed by 36 unique probes (Supporting Table 2) that enable us to provide a DNA-methylation-based risk score of mortality (MI, ranging from 36 to 360; Fig. 1A). As predicted, these 36 probes were hypermethylated in tumor, compared to nontumor, tissue (Supporting Fig. 3). To ensure reproducibility of the methylation status of these probes, we performed technical validation and determine methylation levels by pyrosequencing in a subset of 10 probes in 20 HCCs. There was a significant and high concordance on the methylation status of these probes using both methods (Supporting Fig. 4).

The MI is a metric based on the weighed hazard of death for an individual across all trees (see Materials and Methods for details). This DNA-methylome-based score of mortality risk significantly correlated with known clinical and pathological predictors of survival, such as satellites ($P = 0.02$), multinodularity ($P = 0.002$), vascular invasion ($P < 0.001$), BCLC staging ($P = 0.001$), AFP ($P = 0.001$), bilirubin ($P = 0.01$), platelet count ($P = 0.009$), and albumin ($P = 0.03$; Fig. 1B-I). Concordantly, univariate analysis showed that DNA-methylation MI significantly correlated with risk of death ($P < 0.001$; Fig. 2A). Median survival was significantly lower for patients with DNA-methylation MI >230 —which represents 20% of the cohort—than for those with MI <230 (13 vs. 79.7 months, respectively; $P < 0.001$; Fig. 2B). Multivariate Cox's modeling confirmed DNA-methylation MI as an independent predictor of survival (hazard ratio [HR]: 13.35; 95% confidence interval [CI]: 7.94–22.42; $P < 0.001$) along with multinodularity (HR, 1.65; 95% CI: 1.03–2.62; $P < 0.001$) and platelet count (HR, 1.58; 95% CI: 1.01–2.41; $P = 0.03$; Table 2). Similarly, DNA-methylation MI score predicted tumor recurrence (Fig. 2C) and Cox's modeling showed that the signature was also an independent predictor of overall recurrence (HR, 5.8; 95% CI: 3.1–11; $P < 0.001$) along with multinodularity (HR, 1.8; 95% CI: 1.1–3.0; $P = 0.007$; Table 2).

We next sought to validate the prognostic performance of our methylation-based MI in an independent data set. The validation cohort also included 83 HCC patients treated by surgical resection from different French institutions. Unlike the training set, the main etiology of these patients was related to alcohol intake (47%; 39 of 83), presented with more aggressive disease (microvascular invasion 54%, satellites 42%, and tumors >3 cm in 83%) and had shorter median follow-up time (35 months). Patients with a high DNA-methylation-based MI had a significantly lower cancer-related survival ($P = 0.01$; Fig. 2D). Even when arbitrary categorizing MI by percentiles, there is a direct relation between MI and poor outcome, both in training and validation sets (Supporting Fig. 5). Overall, these data demonstrate how epigenetic deregulation correlates with patient outcome in surgically resected HCC and introduces a first-in-class DNA-validated, methylation-based signature.

Integrated Prognostic Classification With Transcriptome-Based Predictors.

In order to further characterize the subset of patients identified by the DNA-methylation signature, we integrate whole-genome mRNA and methylation data of 205 HCCs. DNA-methylation-based high MI was significantly enriched in tumors that harbored mRNA signatures capturing progenitor cell features, such as epithelial cell adhesion molecule (EpCAM)¹⁸ ($P = 0.009$) and S2 (a reported molecular class of potential progenitor cell origin¹⁹; $P = 0.006$; Fig. 3). This was further confirmed when we performed GSEA on the samples with the highest MI (Supporting Table 3). In addition, we found a significant enrichment of Gene Ontologies related to chromosomal dynamics and RNA processing in these samples (false discovery rate: <0.05). Conversely, tumors with DNA-methylation-based low MI were enriched in gene sets associated with metabolism and homeostasis (Supporting Table 3).

Noncoding mRNA expression profile and DNA methylation data were tested in 205 samples. Patients with a high DNA-methylation MI had a differential noncoding mRNA expression pattern comprising five microRNAs (miRNAs) and one small nucleolar RNA: miR-3193 ($P < 0.001$); miR-27b-star ($P = 0.002$); miR-422a ($P = 0.02$); miR-378c ($P = 0.03$); miR-938 ($P = 0.03$); and SNORD126 (annotated as miR-1201 in the array; $P = 0.03$). All six noncoding RNAs were significantly down-regulated in patients with DNA-methylation high MI score (Supporting Fig. 6).

Aberrant Methylation in Human HCC.

Similarly to other malignancies, there was a remarkable predominance of hypomethylated probes in HCC, compared to normal liver (68%; Fig. 4A). Hypomethylated probes were mainly located in the intergenic (39.9%) and body regions (34.5%), whereas hypermethylated probes were predominantly located in promoter areas (50.5%). Regarding CpG island relation, hypomethylated probes were mainly located in open sea regions (63.55%), whereas hypermethylated probes tended to locate in CpG islands (63.9%) and shores (24.8%; Fig. 4A).

To identify potential methylation markers of malignant transformation, we searched for probes that could accurately differentiate normal liver tissue and HCC, among the 11,307 above-described probes. Analyses focused on probes with low intragroup variability (tumor and nontumor), but high intergroup variability, as defined by an F score (mean- β differences between groups/mean- β differences within groups). The top-100-ranked probes using this score were able to accurately differentiate normal liver and HCC tissues (Fig. 4B; Supporting Table 4). These 100 probes were all hypermethylated in HCC and corresponded to 70 different genes. Gene ontology analyses showed enrichment in entries related to gene expression regulation (such as transcription factors [$P = 2.9E^{-6}$]) and different signaling cascades, such as insulin growth factor (IGF; $P = 2.1E^{-3}$), phosphatidylinositol 3-kinase (PI3K; $P = 3E^{-3}$), transforming growth factor beta (TGF- β ; $P = 5.8E^{-3}$), and cadherin ($P = 3.5E^{-2}$), among others. When compared to a recently published analysis in 24 HCCs,²⁰ we found 43 common probes (Supporting Fig. 7A; Supporting Table 5) and validated their ability to distinguish HCC from nontumor tissue (Supporting Fig. 7B). We next sought to explore the landscape of known and novel genes candidates as potential epdrivers in HCC.

First, we validated other previously reported aberrantly DNA methylated genes in HCC such as hypermethylation of RASSF1 (82%), APC (78%) or NEFH (43%)²¹ and hypomethylation of *IGF2* (51%), among others (Table 3). Then, we identified that known drivers in other tumors show aberrant promoter methylation in HCC, such as *NOTCH3*,²² *nuclear receptor binding SET domain protein 1 (NSD1)*,²² and *zinc finger of the cerebellum 1 (ZIC1)*. Finally, in order to describe novel candidate potential epdrivers not previously described in HCC, we selected the top 500 *F*-scored probes that were clustered in their target genes (Supporting Table 6) and identified genes involved in TGF- β , or fibroblast growth factor (FGF) signaling. Those genes with the highest number of deregulated hits between HCC and normal liver were highlighted as the most appealing candidates (Table 3). Some of these candidates were also validated with pyrosequencing (Supporting Fig. 8). Among them *SEPT9*, a tumor suppressor described in colon and ovarian cancer,²³ and *ephrin-B2* ligand (*EFNB2*), of which hypermethylation was reported in patients with acute leukemia.²⁴ These new candidate epdrivers were also found significantly deregulated in the validation cohort of 83 HCCs (Fig. 5).

Discussion

The understanding of molecular pathogenesis of HCC and gene-based prognostic prediction has improved during the last decade. By using next-generation sequencing and SNP array analysis, the main structural alterations and drivers in HCC have been uncovered. Similarly, transcriptome analysis allowed for establishing molecular subclasses of this cancer, although with no direct implications in the management of patients thus far.¹¹ Nonetheless, there still is a clear lack of understanding of the role of epigenetics in HCC and whether parameters related to methylation have clinical and outcome relevance. Few studies have addressed the role of epigenetics in hepatocarcinogenesis,^{20,25,26} and there is even less information regarding the relevance of DNA-methylation-based signatures as prognostic biomarkers of this prevalent and lethal cancer. Herein, by using a genome-wide approach, we present a first-in-class 36-probe DNA-methylation signature able to characterize a molecularly aggressive HCC subtype. In addition, we describe a landscape of aberrantly methylated promoters of genes potentially involved in hepatocarcinogenesis.

To develop the signature, we profiled a total of 304 HCCs (221 in training and 83 in validation sets) using high-density methylation arrays in a training-validation approach. Previous studies have looked at epigenetic prognostic markers in HCC, but using lower-density arrays, less samples, and without independent validation.^{20,26,27} Similarly to other solid tumors,⁸ our study demonstrates the capability of methylation data to accurately classify HCC patients based on their outcome. Unlike previous prediction methods applied in prognostic studies in HCC,⁵ the design of the 36-probe signature provides a quantitative death risk score (i.e., MI) to each new patient, eliminating signal redundancy. Correlation with other prognostic clinical parameters in the training set reinforces its prognostic performance. Of note, despite remarkable etiology differences between the training (mostly hepatitis C) and validation (mainly alcohol-related) sets, the signature confirmed its ability for outcome prediction. In addition, the DNA-methylation signature defines high risk of mortality in 20% of patients in the cohort with best outcome (i.e., training set), compared to 50% of patients in the validation cohort, proved to have more-aggressive tumors, a

feature that suggests the ability of the DNA-methylation signature to identify patients with aggressive tumors across different stages.

Frequently, prognosis research studies lack the high standards required in other fields of medicine, such as therapeutic trials and genetic epidemiology.²⁸ This is paradigmatic in prognostic gene signatures considering the high discrepancy between reported ones and those that are finally implemented in clinical practice. In fact, currently, guidelines in HCC management do not include any in the routine care of HCC patients.^{3,4} Our study is entirely consonant with the prognostic factor research scheme of the PROGRESS Partnership,²⁸ an initiative that builds upon previous recommendations for reporting prognostic biomarkers in oncology (REMARK statement²⁹). Nonetheless, in order to confirm the clinical strength of the novel information obtained with the DNA-methylation signature, an external independent validation would be required to conform to recent guidelines in HCC.³ In terms of its clinical implementation, prospective validation of the signature in cohort studies will provide solid evidence to facilitate its inclusion in clinical practice guidelines. Preliminary studies suggest the feasibility of DNA-methylation profiling from biopsy specimens or even plasma (i.e., circulating cell-free DNA).³⁰

Integrative analysis with transcriptome data enabled us to obtain additional insight into molecular subclasses in HCC. Considering that DNA methylation provides one of the layers of epigenetic control of tissue specification and differentiation,³¹ it is remarkable that patients within the high mortality risk group, based on the methylation signature (defined by a MI higher than 230), were significantly enriched in mRNA proliferation subclasses capturing progenitor cell origin (EpCAM¹⁸ and S2¹⁹). Interestingly, we found that a set of noncoding RNA significantly down-regulated in patients within the high-risk methylation group. Of note, miR-27b has been characterized as a regulatory hub in lipid metabolism in liver,³² being also involved in peroxisome proliferator-activated receptor (PPAR) signaling.³³ The role of PPAR receptors in HCC remains controversial,³⁴ but in retrospective studies the use of PPARG agonists was associated with a decrease in liver cancer.³⁵

Approximately 140 genetic drivers have been described in oncogenesis,³⁶ among which epidrivers are the most unknown category. We describe a landscape of DNA-methylation aberrations in 221 HCC patients using a last-generation technology. Our findings point to IGF, PI3K, TGF- β , and WNT signaling as the pathways clearly deregulated by DNA methylation in HCC. We also explored the role of methylation in WNT signaling in a subset of 47 samples for which we had *catenin (cadherin-associated protein) beta 1 (CTNNB1)* mutation status data. Interestingly, we found significant hypermethylation of the four WNT pathway genes included among the top 500 *F*-scored probes in *CTNNB1*-mutated samples (12 of 47; 25%), compared to wild type ($P < 0.05$; Supporting Fig. 9). Whether epigenetic changes contribute to further deregulate WNT signaling in addition to *CTNNB1* mutations deserves additional investigations. We confirm the high prevalence of DNA-methylation aberrations of promoters of described tumor suppressors (*RSSFA1*, *APC*, *NEFH*)²¹ or potential oncogenes (*IGF2*) in HCC. On the other hand, we identified aberrant methylation in HCC of epidrivers described in other neoplasms, such as *NOTCH3* in acute leukemias,³⁷ *NSD1* in glioblastoma,³⁸ and *ZIC1* in colorectal cancer.³⁹ Finally, we are pointing to

novel candidates that, despite being involved in carcinogenesis, have not been described as deregulated by DNA methylation, such as *SEPT9*, a tumor suppressor described in colon and ovarian cancer, and ephrin-B2 ligand *EFNB2*, of which hypermethylation was reported in patients with acute leukemia, and others involved in TGF- β receptor-signaling (*homeobox A9*, *forkhead box G1*, and *runt-related transcription factor 3*) and FGF-signaling pathways (*FGF8* and *FGF6*). This novel information provides a complementary portrait of epigenetic changes in HCC.

In summary, we describe a novel prognostic biomarker based on promoter DNA methylation changes in HCC that identifies patients at high risk of death and provides complementary epigenetic characterization of this tumor. Early detection of “poor biology” HCCs could have a major impact in the decision making and allocation of resources, such as in the adjuvant setting or complex therapies, such as resection or transplantation. In addition, potential candidate epdrivers have been described in HCC, which would require functional studies for biological confirmation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement:

The authors thank Loreto Boix and Jordi Bruix for their input and critical analysis of the manuscript. The authors also thank Paulette Bioulac-Sage, Charles Balabaud, Jean Saric, and Christophe Laurent (CHU Bordeaux); Jeanne Tran Van Nhieu, Daniel Cherqui, and Daniel Azoulay (CHU Henri Mondor, Créteil); and the tumor bank of CHU Bordeaux and CHU Henri Mondor for contributing to the French tissue collection.

The study is supported by the European Commission Framework Programme 7 (Heptromic; proposal no.: 259744). J.M.L. is supported by grants from the European Commission Framework Programme 7 (Heptromic; proposal no.: 259744), The Samuel Waxman Cancer Research Foundation, the Spanish National Health Institute (J.M.L.: SAF-2013–41027), and the Asociación Española para el Estudio del Cáncer (AECC). M.E. is funded by Cellex Foundation, Botin Foundation, and the Health and Science Departments of the Catalan Government (Generalitat de Catalunya). Y.H. is supported by the National Institute of Diabetes and Digestive and Kidney Diseases (R01-DK099558). This work was supported by the INCa within the ICGC project, the Ligue Nationale Contre le Cancer (“ Carte d’identité des tumeurs” program), and the Réseau national CRB Foie. CIBERehd is funded by Instituto de Salud Carlos III.

Submitted microarray and methylation data to a publicly available database (accession nos.: GSE56588 and GSE63898).

Potential conflict of interest:

Dr. Battiston is on the speakers’ bureau for Bayer. Dr. Mazzaferro is on the speakers’ bureau for and received grants from Bayer. He is on the speakers’ bureau for BTG and Ipsen. Dr. Llovet consults for and received grants from Bayer, Bristol-Myers Squibb, and Boehringer Ingelheim. He consults for Lilly, Blueprint, Celsion, Novartis, and GlaxoSmithKline.

Abbreviations:

AFP	alpha-fetoprotein
APC	adenomatous polyposis coli
BCLC	Barcelona Clinic Liver Cancer

CHF	cumulative hazard function
CI	confidence interval
CpG	cytosine-phosphate-guanine
CTNNB1	catenin (cadherin-associated protein) beta 1
EFNB2	ephrin-B2
EpCAM	epithelial cell adhesion molecule
FGF	fibroblast growth factor
GSEA	Gene Set Enrichment Analysis
HBV	hepatitis B virus
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
HR	hazard ratio
IGF	insulin growth factor
MI	mortality index
miRNAs	microRNAs
mRNA	messenger RNA
NEFH	neurofilament, heavy polypeptide
NSD1	nuclear receptor binding SET domain protein 1
PCR	polymerase chain reaction
PPAR	peroxisome proliferator-activated receptor
PI3K	phosphatidylinositol 3-kinase
RASSF1	Ras association (RalGDS/AF-6) domain family member 1
RSF	random survival forests
SEPT9	septin 9
SNPs	single-nucleotide polymorphisms
TGF-β	transforming growth factor beta
UTR	untranslated region
VIMP	variable importance
ZIC1	zinc finger of the cerebellum 1

References

1. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380:2095–2128. [PubMed: 23245604]
2. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2013;380:2197–2223.
3. European Association for the Study Of The Liver, European Organisation for Research and Treatment of Cancer. EASL-EORTC Clinical Practice Guidelines: management of hepatocellular carcinoma. *J Hepatol* 2012;56:908–943. [PubMed: 22424438]
4. Bruix J, Sherman M; American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma: an update. *HEPATOLOGY* 2011;53:1020–1022. [PubMed: 21374666]
5. 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;140:1501–1502.e2. [PubMed: 21320499]
6. 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]
7. 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]
8. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol* 2010;28:1057–1068. [PubMed: 20944598]
9. Heyn H, Esteller M. DNA methylation profiling in the clinic: applications and challenges. *Nat Rev Genet* 2012;13:679–692. [PubMed: 22945394]
10. Hinoue T, Weisenberger DJ, Lange CPE, Shen H, Byun HM, Van Den Berg D, et al. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome Res* 2012;22:271–282. [PubMed: 21659424]
11. Villanueva A, Hoshida Y, Toffanin S, Lachenmayer A, Alsinet C, Savic R, et al. New strategies in hepatocellular carcinoma: genomic prognostic markers. *Clin Cancer Res* 2010;16:4688–4694. [PubMed: 20713493]
12. Sandoval J, Heyn HA, Moran S, Serra-Musach J, Pujana MA, Bibikova M, Esteller M. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* 2011;6:692–702. [PubMed: 21593595]
13. Ishwaran H, Kogalur UB. Consistency of random survival forests. *Stat Probab Lett* 2010;80:1056–1064. [PubMed: 20582150]
14. Hoshida Y. Nearest template prediction: a single-sample-based flexible class prediction with confidence assessment. *PLoS One* 2010;5:e15543.
15. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. ; PROGRESS Group. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 2013;10:e1001380.
16. Nault JC, de Reyniès A, Villanueva A, Calderaro J, Rebouissou S, Couchy G, et al. A hepatocellular carcinoma 5-gene score associated with survival of patients following liver resection. *Gastroenterology* 2013;145:176–187. [PubMed: 23567350]
17. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. *Lancet* 2012; 379:1245–1255. [PubMed: 22353262]
18. Yamashita T, Ji J, Budhu A, Forgues M, Yang W, Wang HY, et al. EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. *Gastroenterology* 2009;136:1012–1024. [PubMed: 19150350]
19. Hoshida Y, Nijman SMB, Kobayashi M, Chan JA, Brunet JP, Chiang DY, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res* 2009;69:7385–7392. [PubMed: 19723656]

20. Song MA, 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.
21. 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;145:1424–1435.e1–25. [PubMed: 24012984]
22. Garraway LA, Lander ES. Lessons from the cancer genome. *Cell* 2013; 153:17–37. [PubMed: 23540688]
23. Gyparaki MT, Basdra EK, Papavassiliou AG. DNA methylation biomarkers as diagnostic and prognostic tools in colorectal cancer. *J Mol Med* 2013;91:1249–1256. [PubMed: 24057814]
24. Kuang SQ, Bai H, Fang ZH, Lopez G, Yang H, Tong W, et al. Aberrant DNA methylation and epigenetic inactivation of Eph receptor tyrosine kinases and ephrin ligands in acute lymphoblastic leukemia. *Blood* 2010;115:2412–2419. [PubMed: 20061560]
25. Toffanin S, Hoshida Y, Lachenmayer A, Villanueva A, Cabellos L, Minguez B, et al. MicroRNA-based classification of hepatocellular carcinoma and oncogenic role of miR-517a. *Gastroenterology* 2011;140:1618–1628.e16. [PubMed: 21324318]
26. Mah WC, Thurnherr T, Chow PKH, Chung AY, Ooi LL, Toh HC, et al. Methylation profiles reveal distinct subgroup of hepatocellular carcinoma patients with poor prognosis. *PLoS One* 2014;9:e104158.
27. Shen J, Wang S, Zhang YJ, Wu HC, Kibriya MG, Jasmine F, et al. Exploring genome-wide DNA methylation profiles altered in hepatocellular carcinoma using Infinium HumanMethylation 450 BeadChips. *Epigenetics* 2013;8:34–43. [PubMed: 23208076]
28. Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. ; PROGRESS Group. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* 2013;346:e5595. [PubMed: 23386360]
29. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM; Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst* 2005;97:1180–1184. [PubMed: 16106022]
30. Li X, Zhou F, Jiang C, Wang Y, Lu Y, Yang F, et al. Identification of a DNA methylome profile of esophageal squamous cell carcinoma and potential plasma epigenetic biomarkers for early diagnosis. *PLoS One* 2014;9:e103162.
31. Esteller M Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* 2007;8:286–298. [PubMed: 17339880]
32. Vickers KC, Shoucri BM, Levin MG, Wu H, Pearson DS, OseiHwedieh D, et al. MicroRNA-27b is a regulatory hub in lipid metabolism and is altered in dyslipidemia. *HEPATOLOGY* 2013;57:533–542. [PubMed: 22777896]
33. Lee JJ, Drakaki A, Iliopoulos D, Struhl K. MiR-27b targets PPARc to inhibit growth, tumor progression and the inflammatory response in neuroblastoma cells. *Oncogene* 2012;31:3818–3825. [PubMed: 22120719]
34. Tanaka N, Moriya K, Kiyosawa K, Koike K, Gonzalez FJ, Aoyama T. PPARalpha activation is essential for HCV core protein-induced hepatic steatosis and hepatocellular carcinoma in mice. *J Clin Invest* 2008;118:683–694. [PubMed: 18188449]
35. Chang CH, Lin JW, Wu LC, Lai MS, Chuang LM, Chan KA. Association of thiazolidinediones with liver cancer and colorectal cancer in type 2 diabetes mellitus. *HEPATOLOGY* 2012;55:1462–1472. [PubMed: 22135104]
36. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science* 2013;339:1546–1558. [PubMed: 23539594]
37. Cui H, Kong Y, Xu M, Zhang H. Notch3 functions as a tumor suppressor by controlling cellular senescence. *Cancer Res* 2013;73:3451–3459. [PubMed: 23610446]
38. Berdasco M, Ropero S, Setien F, Fraga MF, Lapunzina P, Losson R, et al. Epigenetic inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1 in human neuroblastoma and glioma. *Proc Natl Acad Sci U S A* 2009; 106: 21830–21835. [PubMed: 20018718]

39. Gan L, Chen S, Zhong J, Wang X, Lam EK, Liu X, et al. ZIC1 is downregulated through promoter hypermethylation, and functions as a tumor suppressor gene in colorectal cancer. *PLoS One* 2011;6:e16916.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

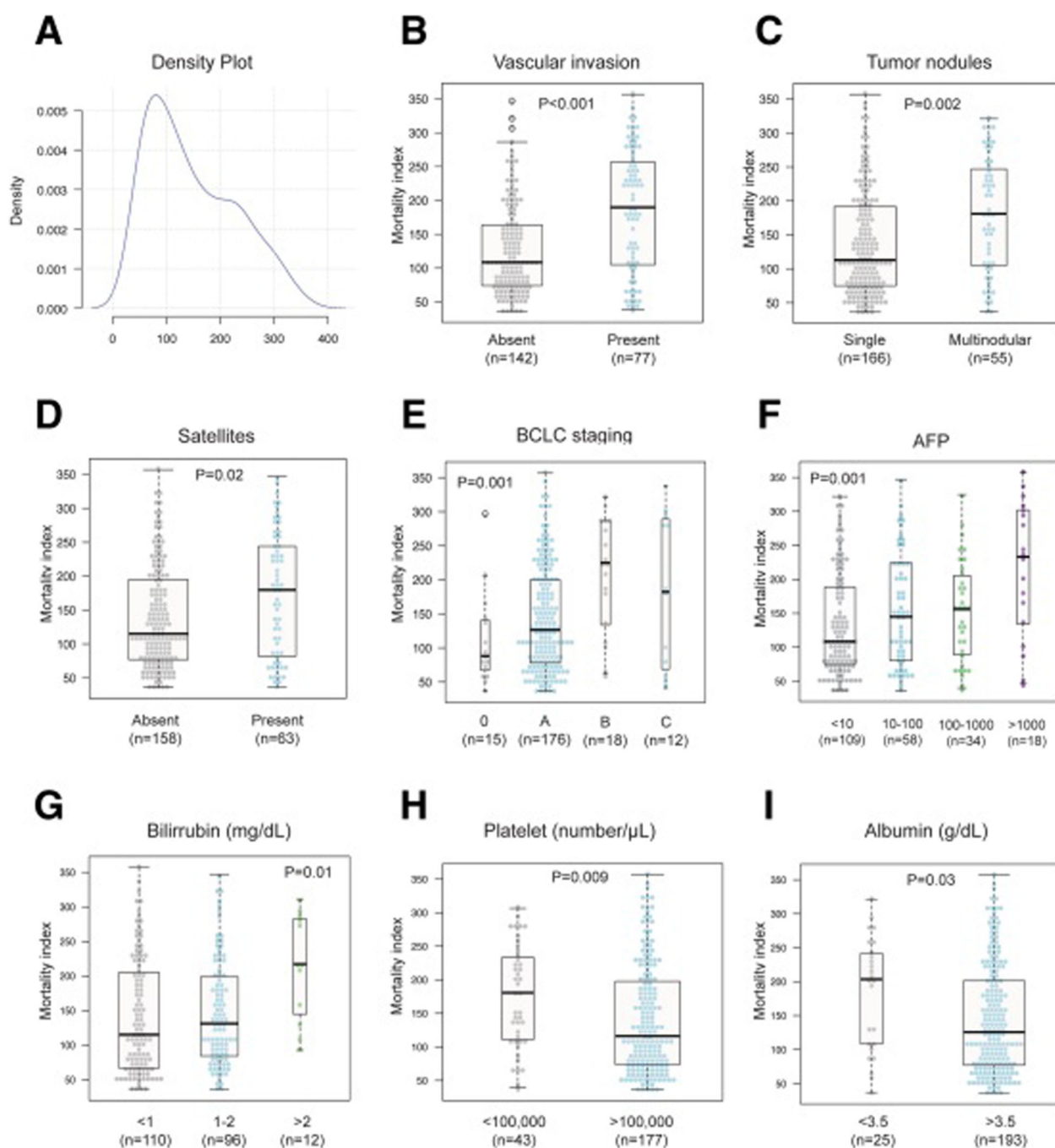


Fig. 1. Distribution of DNA-methylation-derived MI in the training set (Heptromic) as per density plot (A). Box plots (with overimposed individual values) for different clinical and pathological variables known to be prognostic in HCC and the mortality index (B-I).

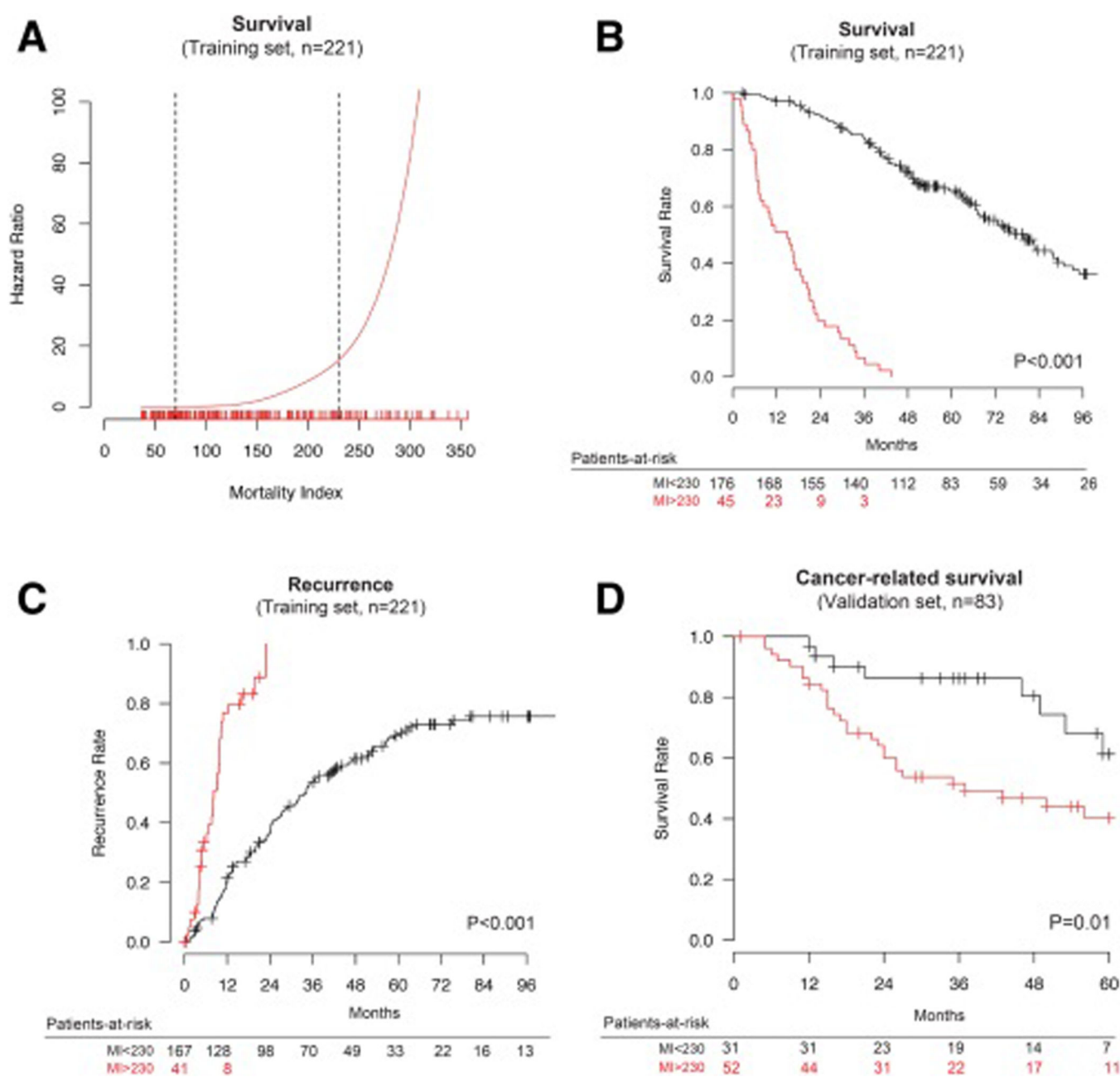


Fig. 2.
Predicted HR for survival based on DNA-methylation-derived MI values (A). Kaplan-Meier's plots for outcome analysis in Heptronic (n = 221, training; B-C) and French data sets (n = 83, validation; D).

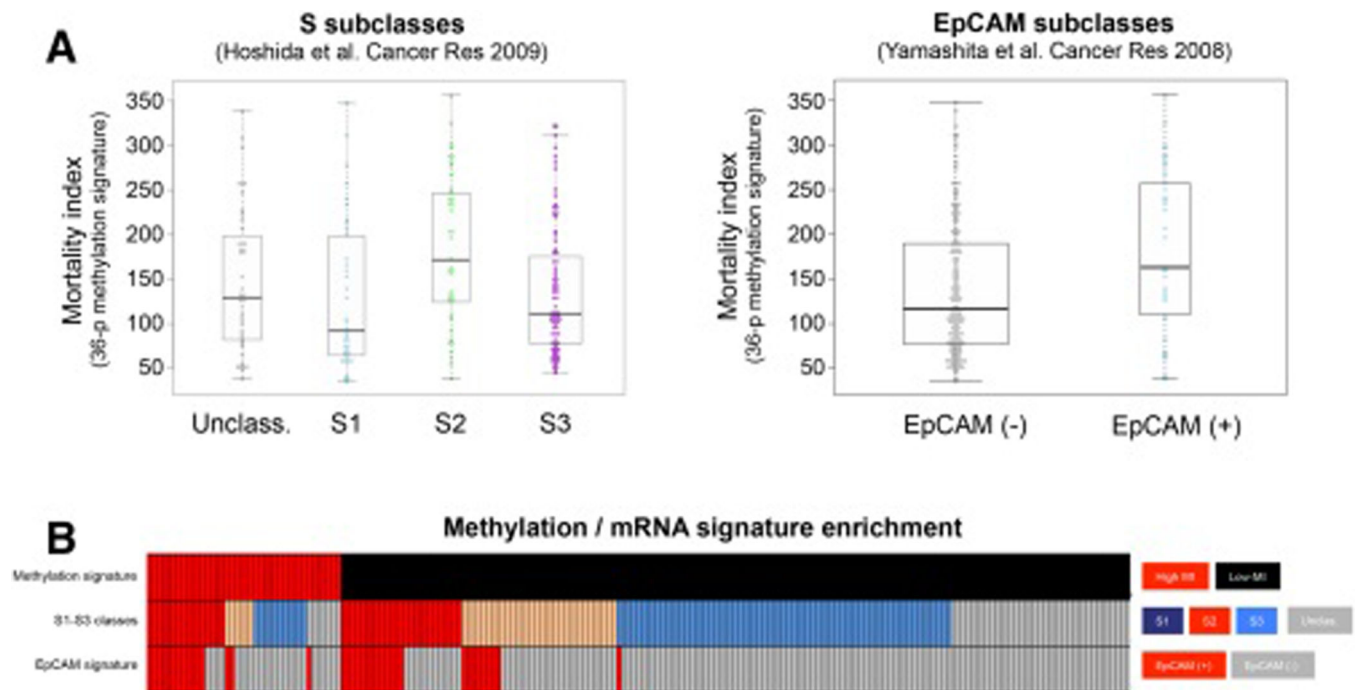


Fig. 3. Box plots (with overimposed individual values) of MI according to S and EpCAM mRNA-based molecular subclasses (A). Patients in the S2 and EpCAM classes (reported initially as identifying tumors with a progenitor cell origin) have significantly higher levels of DNA-methylation-derived MI based on predictions from the methylation signatures. (B) Bottom panel represents enrichment of the 205 samples with methylation and mRNA data based on the levels of risk from the MI score. Top 20% patients with the highest DNA-methylation-derived MI score are significantly enriched in S2 and EpCAM classes.

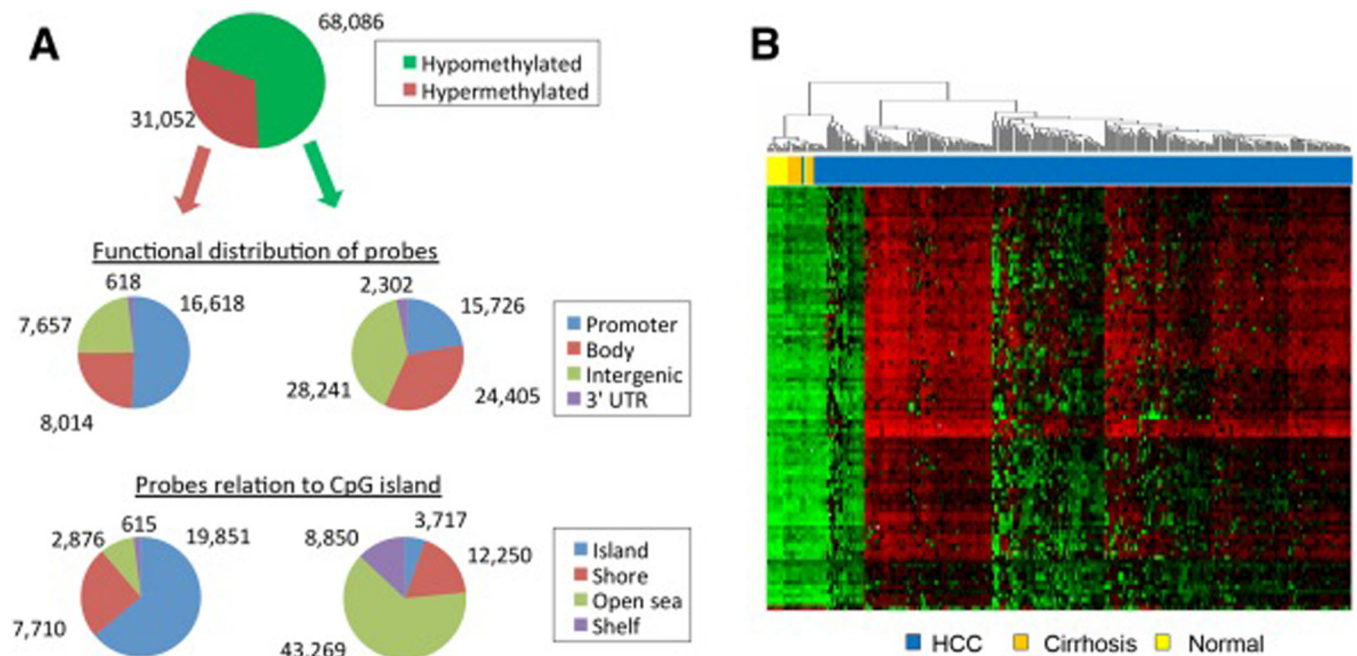


Fig. 4. Genome-wide DNA methylation profiling of HCC patients. Genomic distribution of the differentially methylated probes between normal liver and HCC samples according to: (1) functional genomic distribution (promoter, body, 3' UTR and intergenic) and (2) CpG content and neighborhood context (island, shore, shelf, and open sea) (A). Unsupervised hierarchical clustering of 10 normal liver (yellow), 9 cirrhotic liver (orange), and 221 HCC samples (blue) using the top 100 differentially methylated probes according to *F* score (B).

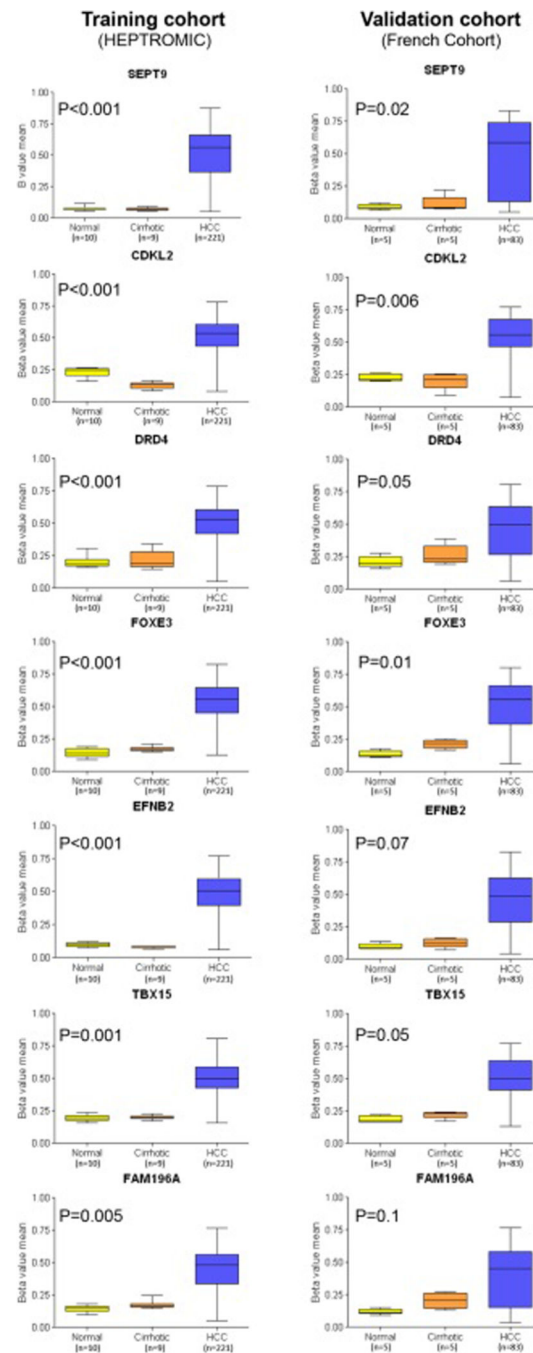


Fig. 5. (Left panels) Box plots represent demethylation status of new candidate epigenetic drivers generated from the training cohort (Heptromic). Right panels show their values on the French cohort (validation). *P* values computed comparing HCC versus normal liver. Abbreviations: CDKL2, cyclin-dependent kinase-like 2; DRD4, dopamine receptor D4; FAM196A, family with sequence similarity 196, member A; FOXE3, forkhead box protein E3; TBX15, T-box 15.

Table 1.
Clinical Characteristics of HCC Patients in the Training and Validation Set

Variable	Training (n = 221)	Validation (n = 83)
Median age	66	66
Gender, male (%)	172 (78)	68 (81)
Etiology (%)		
Hepatitis C	101 (47)	9 (11)
Hepatitis B	44 (20)	17 (20)
Alcohol	35 (16)	39 (47)
Others	37 (17)	NA
Child-Pugh score (%)		
A	216 (96)	74 (92)
B	4 (2)	9 (8)
Tumor size, cm (%)		
<2	26 (11)	5 (6)
2–3	73 (33)	9 (11)
>3	122 (55)	69 (83)
Multiple nodules (%)		
Absent	166 (75)	69 (84)
Present	55 (25)	16 (16)
Microvascular invasion (%)		
Absent	142 (64)	38 (46)
Present	77 (35)	44 (54)
Satellites (%)		
Absent	158 (71)	48 (58)
Present	63 (29)	35 (42)
BCLC early stage, 0-A (%)	191 (87)	59 (73)
Degree of tumor differentiation (%)		
Well	34 (18)	28 (34)
Moderately	105 (57)	35 (43)
Poor	45 (24)	19 (23)
Bilirubin, >1 mg/dL (%)	92 (42)	10 (19) [†]
Albumin, <3.5 g/L	25 (11)	NA
Platelet count, <100,000/mm ³ (%)	43 (19)	NA
AFP, >100 mg/dL (%)	51 (23)	31 (41) [†]
Events (%)		
Recurrence	151 (69)	NA
Death	139 (62)	36 (43) [*]
Median follow-up, months	48.5	35

^{*} Cancer-related death in the validation set, as defined in Nault et al.¹⁶

[†] Missing values: bilirubin (n = 32); AFP (n = 7).

Abbreviation: NA, not applicable.

Table 2.
Uni- and Multivariate Survival Analyses of HCC Patients^a

	Univariate Analysis	Multivariate Analysis (Cox's Regression)		
Variable	P Value	HR	CI (95% Low-High Limits)	P Value
DNA methylation MI	<0.001	13.35	7.94–22.42	<0.001
Tumor size (>35 mm)	<0.001			
Vascular invasion	<0.001			
BCLC stage B or C	<0.001			
Multinodularity	<0.001	1.65	1.03–2.62	0.03
Albumin levels (<3.5 gr/L)	0.001			
186-gene signature ⁶	0.001			
Gender	0.005			
Platelet count (<100,000/mm ³)	0.006	1.56	1.01–2.41	0.04
AFP levels (>100 mg/dL)	0.02			
Bilirubin (>1 mg/dL)	0.02			
Satellites	0.03			
Etiology (HCV)	0.04			

^a
n = 221.

Table 3.

Candidate Tumor Suppressors and Oncogenes With Aberrant Methylation of Promoters in HCC *

	Gene	Chromosome Location	% Hypermethylated		P Value
			HCC Samples (n = 221)	Normal Liver (n = 10)	
Genes reported aberrantly methylated in HCC	<i>RASSF1</i>	3	82	10	<0.001
	<i>IGF2</i>	11	51	100	0.001
	<i>APC</i>	5	78	0	<0.001
	<i>NKX6-2</i>	10	36	0	0.01
	<i>SFRP5</i>	10	7	0	ns
	<i>NEFH</i>	22	43	0	0.006
	<i>RASSF5</i>	1	6	0	ns
Genes reported as drivers in other cancers	<i>NOTCH3</i>	19	40	0	0.007
	<i>NSD1</i>	5	80	0	<0.001
	<i>ZIC1</i>	3	58	0	<0.001
Candidate potentia novel epidrivers in HCC	<i>SEPT9</i>	17	61	0	<0.001
	<i>CDKL2</i>	4	59	0	<0.001
	<i>DRD4</i>	11	57	0	<0.001
	<i>FOXE3</i>	1	57	0	<0.001
	<i>EFNB2</i>	13	53	0	<0.001
	<i>TBX15</i>	1	51	0	0.001
	<i>FAM196A</i>	10	44	0	0.005

* Hepatocellular carcinoma cohort (n = 221).

Abbreviations: NKX6-2, NK6 homeobox 2; SFRP5, secreted frizzled-related protein 5; CDKL2, cyclin-dependent kinase-like 2; DRD4, dopamine receptor D4; FOXE4, forkhead box protein E3; TBX15, T-box 15; FAM196A, family with sequence similarity 196, member A; ns, not significant.