frontiers in Oncology

ORIGINAL RESEARCH
published: 10 November 2021
doi: 10.3389/fonc.2021.769390

Integrative Analysis of Epigenome and Transcriptome Data Reveals Aberrantly Methylated Promoters and Enhancers in Hepatocellular Carcinoma

Peng Huang[1], Mengxiang Xu[1], Haijun Han[1], Xinyi Zhao[1], Ming D. Li[1,2] and Zhongli Yang[1]**

[1] State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2] Research Center for Air Pollution and Health, Zhejiang University, Hangzhou, China

DNA methylation is a key transcription regulator, whose aberration was ubiquitous and important in most cancers including hepatocellular carcinoma (HCC). Whole-genome bisulfite sequencing (WGBS) was conducted for comparison of DNA methylation in tumor and adjacent tissues from 33 HCC patients, accompanying RNA-seq to determine differentially methylated region-associated, differentially expressed genes (DMR-DEGs), which were independently replicated in the TCGA-LIHC cohort and experimentally validated *via* 5-aza-2-deoxycytidine (5-azadC) demethylation. A total of 9,867,700 CpG sites showed significantly differential methylation in HCC. Integrations of mRNA-seq, histone ChIP-seq, and WGBS data identified 611 high-confidence DMR-DEGs. Enrichment analysis demonstrated activation of multiple molecular pathways related to cell cycle and DNA repair, accompanying repression of several critical metabolism pathways such as tyrosine and monocarboxylic acid metabolism. In TCGA-LIHC, we replicated about 53% of identified DMR-DEGs and highlighted the prognostic significance of combinations of methylation and expression of nine DMR-DEGs, which were more efficient prognostic biomarkers than considering either type of data alone. Finally, we validated 22/23 (95.7%) DMR-DEGs in 5-azadC-treated LO2 and/or HepG2 cells. In conclusion, integration of epigenome and transcriptome data depicted activation of multiple pivotal cell cycle-related pathways and repression of several metabolic pathways triggered by aberrant DNA methylation of promoters and enhancers in HCC.

Keywords: epigenetics, DNA methylation, promoter, enhancer, WGBS, RNA-Seq, hepatocellular carcinoma, ChIP-seq

OPEN ACCESS

Edited by:
Shicheng Guo,
University of Wisconsin-Madison, United States

Reviewed by:
Baosen Jia,
Cornell University, United States
Liang Xu,
Tongji University School of Medicine, China
Bowen Lin,
Tongji University, China
Feiyang Ma,
University of California, Los Angeles, United States

***Correspondence:**
Ming D. Li
ml2km@zju.edu.cn
Zhongli Yang
zy3p@zju.edu.cn

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *02 September 2021*
Accepted: *14 October 2021*
Published: *10 November 2021*

Citation:
Huang P, Xu M, Han H, Zhao X, Li MD and Yang Z (2021) Integrative Analysis of Epigenome and Transcriptome Data Reveals Aberrantly Methylated Promoters and Enhancers in Hepatocellular Carcinoma. Front. Oncol. 11:769390. doi: 10.3389/fonc.2021.769390

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the most common malignancies and a growing burden in global health (1, 2), especially in China, which has the highest incidence of HCC due to the high prevalence of hepatitis B virus (HBV) infection (3). Even after decades of research, the 5-year survival rate for liver cancer remains very low, generally less than 5% (4). Therefore, further research on the pathogenesis of HCC and the development of effective diagnosis and prognosis biomarkers is urgently needed.

HCC is a complex disease in which both genetic mutations and epigenetic alterations have been implicated (5). DNA methylation is a critical epigenetic regulator whose aberrations are ubiquitous in many cancers (6). Aberrant DNA methylation has been investigated with several techniques in HCC, including the Human Methylation27 BeadChips (7), Methylation450 BeadChip (5, 8–10), and targeted bisulfite sequencing (11). Alterations in DNA methylation, including both global hypomethylation and regional specific hypermethylation, frequently occur in HCC and related preneoplastic conditions. Hypermethylation events in HCC occur predominantly in promoter-associated CpG islands (CGIs) and correlate with attenuated gene expression (12, 13). Therefore, in HCC, there were recurrent hypermethylated promoter-associated repressions of well-known tumor suppressor genes such as *RASSF1A*, *RUNX3*, *SOCS1*, *HHIP*, *SFRP2*, *APC*, *CDKN1A*, *CDKN2B*, and *CDH* (14–20). However, hypermethylation alterations of CGIs located in the gene bodies of oncogenes were consistently associated with their transcriptional activation (21). Most DNA hypomethylation events in HCC occur in repeat DNA sequences, intergenic regions, and regions outside CGIs (22). Hypomethylation was associated with increased genome instability in HCC (23, 24). Besides, hypomethylation in transcriptional regulatory elements could activate pivotal oncogenic genes like CCAAT/enhancer-binding protein-beta (C/EBPβ) (25). Overall, the transcriptional dysregulations perturbed by abnormal DNA methylation are still not thoroughly clear in HCC.

Furthermore, the commonly used array techniques for studying DNA methylation alterations in HCC lack a good coverage in non-coding regions such as enhancers, which have been implicated as playing pivotal regulatory roles in cancer initiation and development (26, 27). In contrast, whole-genome bisulfite sequencing (WGBS) provides comprehensive single-base-pair resolution-based methylome profiling of more than 90% (>26 million) of all CpGs in the human genome (28). To overcome the limitation of array technologies, WGBS was recently applied to epigenomic profiling of HCC in two studies (24, 25) but with a relatively small sample (generally less than five). Specifically, Dr. Shibata and his colleagues illuminated the interplay between DNA methylation and genetic aberrations by integrating WGBS data and whole-genome shotgun sequencing data (24). On the other side, WGBS was applied to perform global enhancer methylation profiling of three HCC tumors, in which aberrant enhancer hypomethylation of C/EBPβ was discovered and validated as causally linked to C/EBPβ overexpression, thereby contributing to hepatocarcinogenesis through global transcriptional reprogramming (25).

In the present study, we performed WGBS of tumor tissues and paired adjacent tissues from 33 patients for a systematical investigation of the DNA methylation abruption, especially in promoter and enhancer regions, and its associated genes and pathways dysregulated in HCC (**Supplementary Figure S1**). We also aimed to replicate our findings of methylation aberration-associated genes and explore their clinical significances in the TCGA-LIHC cohort to identify potentially effective prognosis biomarkers for HCC.

MATERIALS AND METHODS

Patient Collection

This study was approved by the Institutional Review Board of The First Affiliated Hospital. All tissue samples used in the current study were obtained from patients with HCC who underwent a partial hepatectomy at the First Affiliated Hospital, Zhejiang University School of Medicine. Each specimen was reviewed by a board-certified pathologist to confirm that the frozen section was histologically consistent with tumor or non-tumor tissues. Written informed consent was obtained from each patient.

WGBS and RNA Sequencing

Paired tumor and adjacent non-tumor tissue samples from 33 HCC patients were subjected to WGBS on the Illumina X Ten platform with the procedures described in our previous WGBS paper (29). Briefly, a 200-ng genomic DNA sample was sheared to about 300-bp fragments by sonication. Then, DNA fragments were subjected to end-repair, addition of adenosine to the 3′ end, and TruSeq adaptor ligation (Illumina, San Diego, CA, USA). Bisulfite conversion was implemented *via* the EZ DNA Methylation kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's protocol. After that, bisulfite-converted DNA was enriched through several cycles of PCR amplification using the KAPA HiFi HotStart uracil DNA polymerases (Kapa Biosystems, Boston, MA, USA). The PCR conditions were set as 45 s at 98°C followed by 10 cycles at 98°C for 15 s, 65°C for 30 s, 72°C for 30 s, ending with 72°C for 1 min. The quality of each WGBS library was assessed by Qubit 2.0 (Life Tech, Carlsbad, CA, USA) and an Agilent 2100 Bioanalyzer. Finally, 150-bp pair-end sequencing was conducted on the Illumina X Ten sequencing platform. High-throughput mRNA-seq was performed for each WGBS sample. Similarly, all 66 RNA samples with high quality (RIN ≥ 7) were applied to the Illumina X Ten platform for sequencing. Specifically, total RNA was extracted and purified using the RNeasy Micro Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. The quality of RNA was assessed ***via*** an Agilent 2100 Bioanalyzer. Libraries for poly(A)+ RNA were prepared according to the Illumina standard protocol. Constructed libraries were sequenced on HiSeq X Ten platform by WuXi AppTec (Wuxi, Jiangsu China).

Computational Preprocessing of the Next-Generation Sequencing Data

For raw reads from RNA-seq, Cutadapter (30) (v.1.12) and Trimmomatic (31) (v. 0.33) were applied for adapter removal and trimming of low-quality sequences, followed by FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) for a quality check. After quality control, clean reads were submitted as the input of Kallisto (32) (v.0.44) for abundance quantification of transcripts based on a gene model download from the GENCODE (v. 29) (33). Normalized expression (transcripts per million [TPM] reads) of each gene was summarized from the transcript level *via* the R package tximport (v. 1.6.0) (34). Non-expressed and low-expressed genes (defined as those with TPM < 0.01 among more than half of the total 66 samples) were excluded from downstream analysis. Clean reads were also aligned against the reference genome by STAR (v. 2.5.2a) (35), and the resulting BAM files were utilized as input for enhancer RNA expression quantification *via* bedtools (v.2.27.1) (36). Besides, the gene count outputs of STAR were used as inputs of DESeq2 (v.1.18.1) (37) for identification of differential expressed genes (FDR < 5% and $|\text{Log}_2\text{FoldChange}| > 0.5$). The clean WGBS reads that passed preprocessing were aligned with the hg38 reference genome using Bismark (v. 0.16.1) (38) with default parameters. The harvested count information from each strand was then combined. As recommended by the R package DSS (v.2.26.0) developer (39), the smoothing approach was adopted for estimation of smoothed methylation level for all 28.9 million CpGs with default parameters.

Identification of DML and Differentially Methylated Regions

In order to identify overall significant differential methylation between all tumor and non-tumor samples, a combined Baumgartner–Weiβ–Schindler (BWS) test (40) was applied to carry out age-adjusted differentially methylated loci (DML) detection *via* the R package BWStest (v.0.2.2). We divided all the 33 HCC patients into three age groups: "young" (age < 55 years; n = 10), "medium" (55 < age ≤ 65; n = 13), and "old" (age > 65; n = 10). A single BWS test was performed for each age group on every CpG to obtain two individual BWS *p*-values (p_{left} and p_{right}). Afterward, three one-sided *p*-values were combined as statistic T_{left} (or T_{right}) = $-2^* \Sigma\log_{10}(p_{\text{left}}$ [or p_{right}]), and a new statistic T was defined as $\max(T_{\text{left}}, T_{\text{right}})$ (40). The empirical distribution of the T statistics of combined BWS test was determined by 2.0×10^8 time permutations. At last, an overall empirical *p*-value was estimated as the combined BWS *p*-values for each CpG. CpG with a combined BWS *p*-value < 1.0×10^{-5} was identified as DML for subsequent differentially methylated loci (DMR) calling.

Tumor-associated DMRs were determined by R script with the following two steps: 1) DML were combined into pre-DMRs if the distance between neighbor CpGs was < 200 bp; and 2) all CpGs located between the start and the end of each pre-DMR were included as a final DMR. The arithmetic mean of T statistics for all CpGs in each DMR was calculated for estimating the empirical combined BWS *p*-value for each DMR. Group-level methylation was estimated as the arithmetic average of DNA methylation of all CpGs in each corresponding DMR.

Annotation of DML and DMRs

Annotation of the genomic location of each identified DML and DMR was realized by using the function "findOverlaps" in the R package GenomicRanges (v.1.30.3) (41). Specifically, location annotations were determined according to overlaps (>1 bp) between the range of each DML/DMR and all known genomic regions. These genomic region annotations were defined on the basis of the gene model downloaded from the GENCODE, which included promoter (upstream 1,500 bp and downstream 500 bp from the TSS of each transcript), exon, intron, 5′-UTR, 3′-UTR, and intergenic region.

Identification of Promoter- and Enhancer-Like DMRs

Functional annotation was performed for all identified DMRs to search for potentially active promoter- and enhancer-like DMRs. Specifically, liver-active promoter/enhancers were obtained according to candidate regulatory element annotation of eight liver-relevant histone ChIP-seq samples, which included five tissue samples (i.e., two tumor and adjacent samples from two HCC patients and one normal liver sample) from a recent integrative epigenomic HCC study (42) and three liver-relevant samples (i.e., one adult liver tissue, one hepatocyte, and one HepG2 cell sample) from the ENCODE database (43). The regulatory element annotation of these five liver tissue samples was based on 10-state chromHMM annotations, whereas those three ENCODE samples were based on five-state candidate regulatory element annotations. Hence, active liver promoters were composed of regions annotated as "activeTSS" or "activePromoter" (refer to regions with both H3K4me3 and H3K27ac peaks) in any of those five tissue samples and "promoter-like cRE" (refer to regions with both H3K4me3 and DNase peaks) in any of those three ENCODE samples. Similarly, active liver enhancers included regions annotated as "activeEnhancer" (regions with both H3K4me1 and H3K27ac peaks) in any of those five tissue samples and "enhancer-like cRE" (regions with both H3K27ac and DNase peaks) in any of those three ENCODE samples. Finally, DMRs that overlapped with at least one active promoter/enhancer in any sample were identified as promoter-like or enhancer-like DMRs. Besides, promoter-like DMRs also included DMRs that were annotated as promoters in the genomic location annotation. For identified promoter/enhancer-like DMRs, their promoter/enhancer-like activity scores were calculated as the number of ChIP-seq liver samples in which they were annotated as promoter/enhancer-like regulatory elements. Moreover, as a supplement of annotated enhancers in the public domain, we estimated enhancer RNA (eRNA) expression, which was reported to be a reliable indicator of enhancer activity (44, 45), for all intergenic DMRs *via* bedtools (36) for identification of potential novel enhancers with active enhancer expression (count of reads ≥ 3 in at least one third of tumor or adjacent HCC samples) among our HCC samples.

Identification of DMR-DEGs

To investigate the effect of the identified tumor-associated DMRs in transcriptional regulation, we performed integrative genomic analysis by integrating the paired DNA methylomic and transcriptomic data from our original HCC cohort and histone ChIP-seq data from the public domain. To ameliorate potential false-positive findings, we focused only on DMRs ($|\Delta_{methylation}| \geq 0.15$) in candidate active regulatory elements, i.e., active promoters and enhancers in the liver. As mentioned earlier, DMRs that overlapped with active promoters or enhancers were identified as promoter/enhancer-like DMRs. For promoter-like DMRs, nearby genes (TSS ≤ 2 kb away from the DMR start or end site) were tested for correlation between mRNA expression and methylation of DMR. Only differentially expressed genes (FDR < 5% and $|\text{Log}_2$ FoldChange$| > 0.5$) with a BH-corrected Spearman correlation *p*-value < 0.05 were designated promoter-like DMR-associated genes.

As for genic enhancer-like DMRs, nearby genes whose distances from DMR were < 100 kb were examined for Spearman correlation, and differentially expressed genes with a BH-corrected correlation *p*-value < 0.05 were identified as enhancer-like DMR-associated genes. Regarding those identified intergenic enhancer-like DMRs, neighbor genes within ± 0.5 Mb of those intergenic enhancer-like DMRs were screened for differentially expressed genes with significant DMR-eRNA–gene triple correlation (simultaneous significant Spearman correlation between DMR methylation and eRNA expression, between eRNA and gene expression, and between DMR methylation and gene expression). Additionally, enhancer-like DMR-associated target genes were ruled out of genes that were also identified as promoter-like DMR-associated genes. Those target genes associated with promoter-like DMRs, genic enhancer-like DMRs, or intergenic enhancer-like DMRs were defined as DMR-DEGs.

Pathway Enrichment Analysis

Genic- and intergenic enhancer-like DMR-associated genes were first combined as enhancer-like DMR-DEGs, along with promoter-like DMR-DEGs which were then used as inputs for enrichment analysis *via* Metascape (v.3.5) (46).

Identification of High-Confidence DMR-DEGs

Strict screening procedures were imposed to identify DMR-DEGs with high confidence and a low possibility of false-positive results, which would be more applicable to downstream validation. For promoter-like DMRs and genic enhancer-like DMRs, we conducted another genomic location and functional annotation with a much stricter criterion. Specifically, only DMRs overlapping with at least 80% of a promoter region, an active promoter, or an active enhancer were defined as high-confidence promoter/enhancer-like DMRs. Afterward, all DEGs associated with those high-confidence promoter/enhancer-like DMRs were identified as high-confidence DMR-DEGs. As for intergenic enhancer-like DMRs, only target DEGs that were highly positively co-expressed (Spearman correlation coefficient $\rho \geq 0.7$) with corresponding eRNAs passed the screening and were included as high-confidence DMR-DEGs.

Independent Replication and Clinical Significance of DMR-DEGs in the TCGA-LIHC Cohort

In silico replication was conducted for each high-confidence DMR–DEG pair using methylomic and transcriptomic data of the TCGA-LIHC cohort. The clinical phenotypes, DNA methylation, and gene expression dataset of the TCGA-LIHC cohort were downloaded *via* the RTCGA (v.1.22.0) R package (47). The 450 k array methylation data were updated to hg38 from hg19 *via* the UCSC genome liftover tool (genome.ucsc.deu/cgi-bin/hgLiftOver). For each high-confidence DMR–gene pair, the availability of all CpG sites in the DMR and active expression of the corresponding target gene were verified in the LIHC-TCGA. If no CpGs existed or the gene was not expressed, the replication was said to have failed for that DMR–gene pair. Differential DNA methylation, differential gene expression, and the correlation between DNA methylation and gene expression were evaluated *via* Wilcox test and Spearman correlation, respectively. A DMR–gene pair was considered to have been replicated in TCGA-LIHC only when there were significant differential methylation, differential gene expression, and consistent (same sign) significant correlation. About the calculation of replication rates of our identified DMR-DEGs in the TCGA-LIHC, it might be unfair to replicate our promoter/enhancer-like DMR-DEGs directly in the TCGA-LIHC cohort, whose DNA methylation levels were profiled by 450k array, especially for DMRs in enhancers, which were rarely covered by 450k array. Thus, DMR-DEGs that failed to replicate were divided into two groups: Group I, replication failure was because no CpG of the DMR was covered in the 450k array; Group II, at least one CpG was available in the 450k array but still failed to replicate the correlated differential methylation and differential expression. Platform-adjusted replicated rates were calculated as: $\text{Count}_{\text{Replicated}}/(\text{Count}_{\text{Type II failure}} + \text{Count}_{\text{Replicated}})$.

In the end, the clinical significance of those high-confidence DMR-DEGs was investigated by way of survival analysis and tumor-stage association analysis. Briefly, the overall survival time (OS) and progression-free survival (PFS) time of TCGA-LIHC samples were retrieved from the integrated TCGA pan-cancer clinical data resource (48). Survival analyses for methylation and gene expression were performed on the basis of the univariate Cox proportional hazards regression model. The Kaplan–Meier method was used to create the survival plots, and the log-rank test was employed to compare the difference between survival curves. The optimal cutoffs of DNA methylation, gene expression, and combination of methylation and expression were determined by minimizing the p-values of log-rank tests. The differences in methylation and gene expression in various tumor stages were compared using ANOVA in R (v.3.5).

In Vitro DNA Methylation-Unmasking Treatment

For expression detection, 5×10^5 L02 or HepG2 cells were seeded in 12-well plates and allowed to reach 80%–90% confluence. Then, freshly prepared 5-aza-2-deoxycytidine (5-azadC; decitabine) solution was added to the medium at a final concentration of

100 μM. Cells were allowed to grow for 72 h at 37°C with 5% CO2 and then harvested for RNA extraction and qRT-PCR quantification. The cDNA was reverse-transcribed using the iScript™ cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA) according to the manufacturer's protocol. The real-time PCR was conducted with SYBR Premix Ex Taq (TaKaRa, Kyoto, Japan). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as an internal control for amplification of mRNAs. The comparative C_t method was used to calculate the relative mRNA expression. There were three replicates for each experimental condition (2 cell lines × 2 treatment concentrations).

RESULTS

Age-Dependent Global Hypomethylation in HCC

By WGBS, we obtained the methylation profile of a total of 28,978,826 CpG sites with an average sequencing depth of 12.76 × in 33 pairs of HCC tumor and adjacent non-tumor tissue samples (**Figure 1A**, **Supplementary Figure S2A** and **Supplementary Table S1**). To rescue some CpGs with relatively low depth, smoothed methylation amounts were obtained for all CpGs. We observed a significant methylation difference between tumoral and adjacent tissues and a negative correlation (ρ = -0.49; p = 0.0038) between the average extent of DNA methylation and chronological age in the tumor tissues (**Figure 1B**). This negative correlation also was significant in most genomic regions, including exons, introns, and intergenic regions (**Supplementary Figures S2B–J**). In the PCA plot of all CpGs among all samples, HCCs were clearly separated from paired adjacent tissues, whereas there existed considerable heterogeneities among tumor samples (**Figure 1C**).

Through the combined BWS test, we identified a total of 9,867,700 significant DML between paired HCCs and non-cancerous tissues, including 157,320 hypermethylated DML (hyper-DML) and 9,710,380 hypomethylated DML (hypo-DML). The genomic location annotation of those DML showed that hyper-DML were depleted in the intergenic and intron regions and enriched in other regions, particularly in

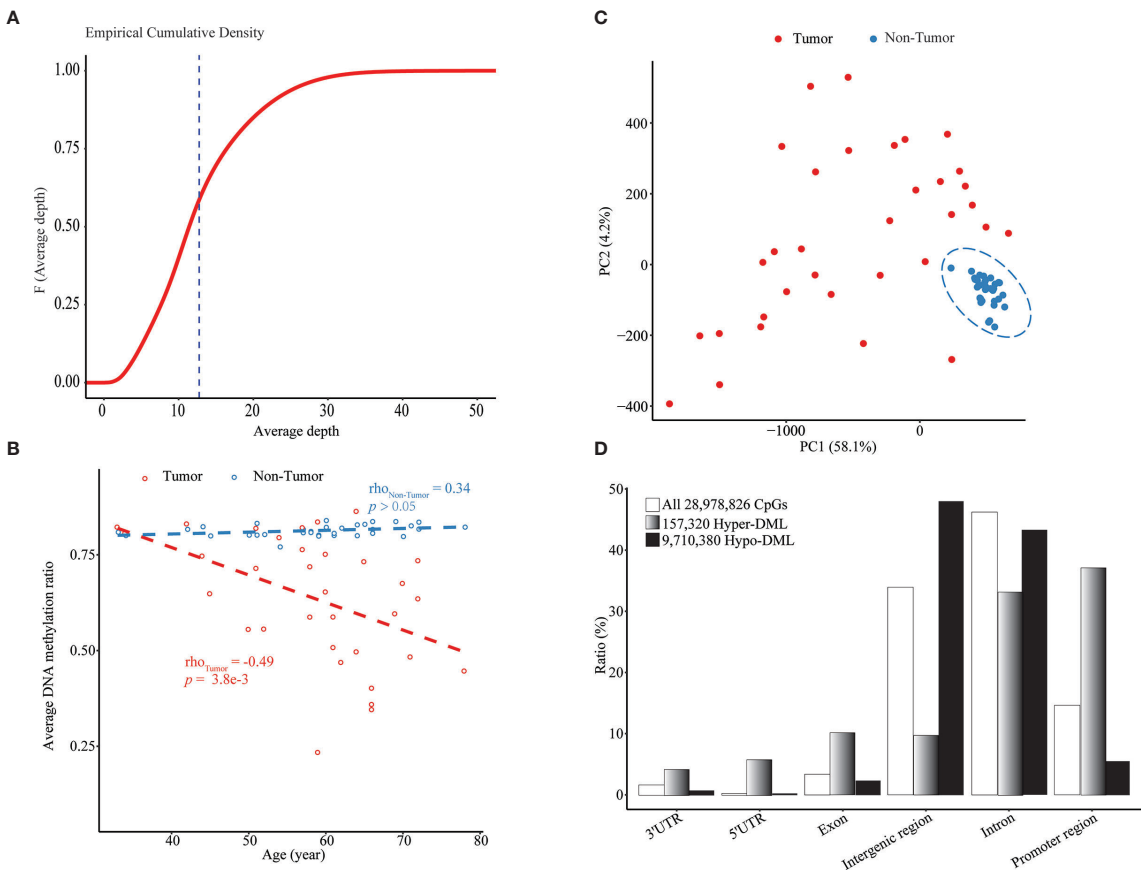

FIGURE 1 | Age-dependent global hypomethylation in HCC. **(A)** Empirical cumulative density plot of average depth of all 28.9 million CpGs profiled by WGBS in all 66 samples. **(B)** Median methylation level of all CpGs in each sample vs. chronological age. **(C)** A PCA plot of all tumor and non-tumor samples based on smoothed methylation levels of all CpGs. **(D)** Genomic location distribution and enrichment of hypomethylated and hypermethylated DML. Upstream 1,500 bp to downstream 500 bp from TSS of each transcript was defined as "promoter region".

promoters (defined as -1,500 ~ +500 bp from a TSS) (**Figure 1D**). As for hypo-DML, the pattern was in the opposite direction (**Figure 1D**). Moreover, 47.94% and 43.44% of those hypo-DML were located in the intergenic and intron regions, respectively (**Figure 1D**), which usually were missed in previous arrays or target sequencing-based methylation studies.

Aberrantly Methylated Promoters and Enhancers in HCC

After DMR calling from those DML, we identified 608,279 DMRs composed of 6,924 hypermethylated DMRs (hyper-DMRs) and 601,355 hypomethylated DMRs (hypo-DMRs). Promoter annotation of these DMRs revealed 2,882 promoter-like hyper-DMRs and 44,611 promoter-like hypo-DMRs. Of them, 1,569 promoter-like hyper-DMRs and 9,285 promoter-like hypo-DMRs exhibited active promoter-associated histone peaks (H3K4me3) in at least one of those eight liver-related ChIP-seq samples (**Figures 2A, B**). From their corresponding enhancer activities in tumor and non-tumor liver ChIP-seq samples, we found 3,232 enhancer-like hyper-DMRs and 20,568 enhancer-like hypo-DMRs (**Figures 2C, D**). Overall, 61.54% of those 6,924 hyper-DMRs were annotated as promoter or enhancer-like regulatory elements (**Supplementary Figure S3A**), while only 4.17% of those 601,355 hypo-DMRs were annotated as promoter or enhancer-like regulatory elements (**Supplementary Figure S3B**). This indicated that hypermethylation events were much less than hypomethylation events but functionally more critical for transcriptional regulation in HCC.

Those promoter-/enhancer-like DMRs were inferred to be "repressed in tumor" (**Figure 2E** and **Supplementary Figure S3C**) or "activated in tumor" (**Figure 2F**; **Supplementary Figure S3D**) according to comparison of their promoter/enhancer activities in tumor and non-tumor samples. In line with the classical negative regulatory relationship between DNA methylation and gene expression, a majority (82.47% and 85.02%, respectively) of those promoter- and enhancer-like hyper-DMRs were recognized as being

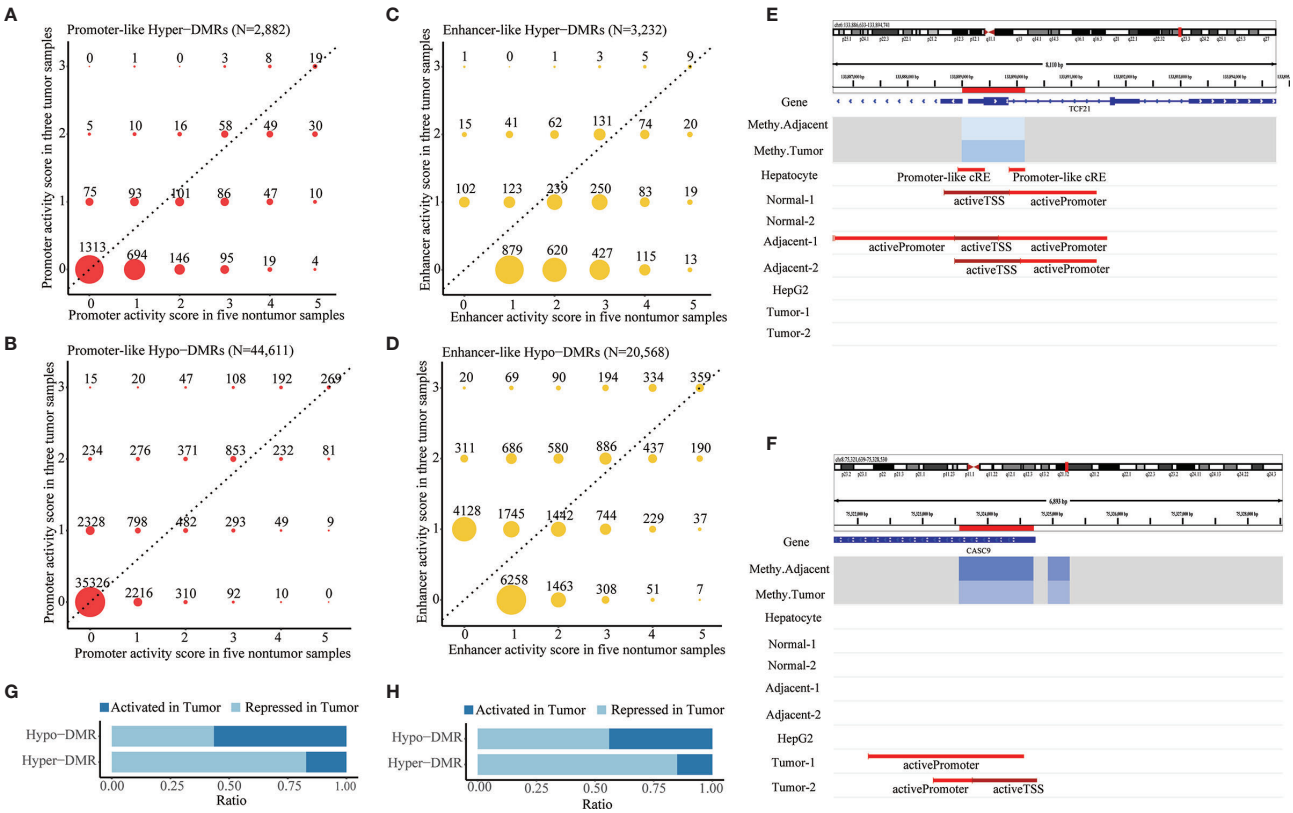

FIGURE 2 | Identification of aberrant DNA methylation in promoter- and enhancer-like regulatory elements. **(A,B)** Distribution of promoter activity of hypermethylated and hypomethylated promoter-like DMRs in tumor *vs*. non-tumor liver ChIP-seq samples. **(C,D)** Distribution of enhancer activity of hypermethylated and hypomethylated enhancer-like DMRs in tumor vs. non-tumor liver ChIP-seq samples. Promoter/enhancer activity scores in tumor and non-tumor reflect counts of tumor and non-tumor liver ChIP-seq samples that show active promoter-/enhancer-related histone peaks in corresponding DMRs. Promoter/enhancers that show greater activity (dots in upper area of dotted line) in tumor condition would be defined as "activated (promoter/enhancer) in tumor" and vice versa. **(E,F)** Paradigm of repressed and activated promoter associated with differential methylation in HCC. Location of DMRs highlighted in red in the chromosome. "Methy.Adjacent" and "Methy.Tumor" refer to average methylation of DMR among adjacent and tumor WGBS samples, respectively. "Promoter-Like cRE" refers to a genomic region with both H3K4me3 and DNase peak; "activePromoter" and "activeTSS" refer to similar regions with both H3K4me3 and H3K27ac peak. **(G,H)** Summarized ratio of "activated (promoter/enhancer) in tumor" of hypermethylated and hypomethylated promoter/enhancer-like DMRs.

repressed promoters or enhancers in HCC (**Figures 2G, H**) whereas about half (44.01% and 56.03%, respectively) of those promoter-like or enhancer-like hypo-DMRs were inferred to be repressed promoters or enhancers in HCC (**Figures 2G, H**), indicating the potential presence of substantial non-classical positive regulation between DNA methylation and histone modification.

The genomic location annotation for all 6,924 hyper-DMRs and 601,355 hypo-DMRs revealed that there were 641 intergenic hyper-DMRs and 250,932 intergenic hypo-DMRs (**Figure S3E**). Quantification of eRNA expression for all intergenic DMRs revealed 36,651 hypo-DMRs and 309 hyper-DMRs with active eRNA expression in our HCC samples. They were recognized as candidate active enhancers in the liver given that eRNA expression is implicated to be a reliable indicator of enhancer activity. The majority of those intergenic active enhancer candidates appeared to be novel. Specifically, 2,378 of them (6.4%) were annotated as active enhancers in at least one of the eight ChIP-seq liver samples, whereas only 6,622 of all 283,631 intergenic DMRs (2.3%) were annotated as active enhancers, an indication of a nearly three-fold enrichment of known enhancers among our intergenic active enhancer candidates. After correlation analysis between DNA methylation and eRNA expression, 4,833 intergenic hypo-DMRs and 23 intergenic hyper-DMRs exhibited a significant negative methylation-eRNA correlation, whereas 2,126 hypo-DMRs and 52 hyper-DMRs displayed non-classical positive methylation-enhancer regulation(**Figure S3F**). Only these 7,034 intergenic DMRs with both active eRNA expression and significant methylation–eRNA correlation were defined as intergenic enhancer-like DMRs for downstream analyses.

Genes and Pathways Deregulated by Aberrant DNA Methylation in HCC

Aberrant methylation associated genes were determined by integrating DNA methylation with transcriptomic data for our identified promoter-like DMRs, genic enhancer-like DMRs, and intergenic enhancer-like DMRs. Specifically, we found a total of 1,323 potential target genes (i.e., promoter-like DMR-DEGs) for all promoter-like DMRs (**Figure 3A** and **Supplementary Table S2.1.1**). As for genic enhancer-like DMRs, we determined 1,751 genes to be their potential targets (i.e., genic enhancer-like DMR-DEGs) (**Figure 3A** and **Supplementary Table S2.1.3**). Regarding intergenic enhancer-like DMRs, there were 562 genes (i.e., intergenic enhancer-like DMR-DEGs) that passed the methylation-eRNA–gene triple correlation examination; i.e., these 562 DMR-DEGs displayed simultaneous significant

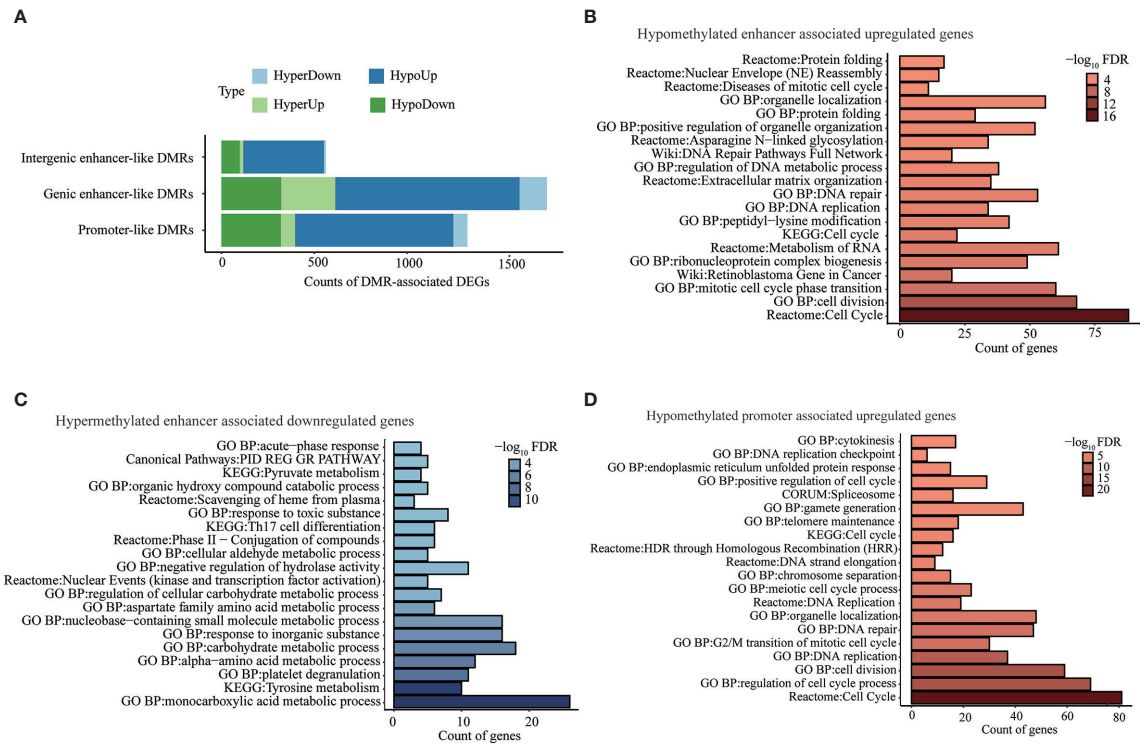

FIGURE 3 | Genes and pathways deregulated by aberrant DNA methylation in promoter- and enhancer-like regulatory elements. **(A)** Counts of genes classically or non-classically associated with identified promoter-, genic enhancer-, and novel intergenic enhancer-like DMRs. "HyperDown" refers to hypermethylated and downregulated gene, "HyperUp" refers to hypermethylated and upregulated gene, "HypoDown" refers to hypomethylated and downregulated gene, and "HypoUp" refers to hypomethylated and upregulated gene. **(B)** Top 20 pathways/process that significantly enriched in activated genes associated with hypomethylated enhancer-like DMRs. **(C)** Top 20 pathways/processes that significantly enriched in repressed genes associated with hypermethylated enhancer-like DMRs. **(D)** Top 20 pathways/processes that significantly enriched in activated genes associated with hypomethylated promoter-like DMRs.

Spearman correlations between DMR methylation and eRNA expression, between eRNA and gene expression, and between DMR methylation and gene expression (**Figure 3A** and **Supplementary Table S2.1.5**). Overall, 70.0%, 65.0%, and 79.0% of those identified promoter-like DMR-DEGs, genic enhancer-like DMR-DEGs, and intergenic enhancer-like DMR-DEGs appeared to be negatively regulated by DNA methylation.

The pathway enrichment analysis of negatively correlated DMR-DEGs (i.e., HyperDown and HypoUp) demonstrated deregulated DNA methylations in enhancer-induced activation of genes implicated in the cell cycle, retinoblastoma gene in cancer, DNA replication, and DNA repair (**Figure 3B**) accompanied by the repression of genes in various critical metabolism pathways, including monocarboxylic acid metabolic process, tyrosine metabolism, and carbohydrate metabolic process (**Figure 3C**). Similarly, hypomethylated promoters activated genes implicated in the cell cycle, DNA replication, and DNA repair (**Figure 3D**). On the other hand, 75 genes repressed by promoter hypermethylation failed to be enriched in any pathways, possibly because of the small number of genes, but many of them, such as ST8SIA6-AS1 (49) and GRHL2 (50), were reported to play a suppressor role in multiple cancers including HCC (**Table 2**). Pathway enrichment analysis of both negatively and positively methylation-correlated DMR-DEGs displayed similar overrepresented pathways (**Supplementary Table S2.2**).

In Silico Replication and Clinical Significance Investigation of High-Confidence DMR-DEGs

We identified a set of 611 DMR-DEGs with high confidence through strict screening for genes whose associated DMRs overlapped completely with annotated promoters or enhancers from ChIP-seq (**Table 1**). Specifically, we discovered 171 high-confidence promoter-like DMR-DEGs (**Supplementary Table S2.1.2**), 338 high-confidence genic enhancer-like DMR-DEGs (**Supplementary Table S2.1.4**), and 102 high-confidence intergenic enhancer-like DMR-DEGs (**Supplementary Table S2.1.6**). Most of the differential DNA methylation in the promoter and genic enhancer regions (70.18% and 73.96%) exhibited a negative correlation with expression of the target gene, but a considerable proportion (41.18%) of those intergenic enhancers showed hypomethylation-associated gene repression.

Literature searching of the identified 56 top differentially expressed high-confidence DMR-DEGs in our HCC sample indicated that 22 of them were implicated in HCC carcinogenesis and the other 15 genes were involved in other types of cancers (**Tables 2**, **3**). Subsequently, 139/661 high-confidence DMR-DEGs were replicated in the TCGA-LIHC cohort, which consisted of 63 promoter-like DMR-DEGs, 67 genic enhancer-like DMR-DEGs, and 9 intergenic enhancer-like DMR-DEGs (**Table 1**). Given the limited availability of WGBS-profiled CpGs in the 450k array, particularly the CpGs in non-coding regions, the raw replication rate of DMR-DEGs was modest except for the promoter-like DMR-associated DEGs. Nevertheless, when the platform effect was adjusted, we achieved a considerably higher replication rate (66.32%, 60.36%, and 52.94%) for the above three groups of DMR-DEGs in the TCGA-LIHC cohort (**Table 1**).

To explore further the clinical significance of our identified 661 potential methylation drivers, we carried out separate survival analyses and tumor stage-association tests based on DNA methylation and transcription data. At the expression level, survival analysis showed that 108 and 82 of those 661

TABLE 1 | Distribution and replication of 611 high-confident DMR-DEGs.

Type of DMR-DEGs	Count in discovery	Count of type I failure	Count of type II failure	Count passed replication	Replication rate (%)	Adjusted replication rate (%)
Promoter-like DMR-DEGs						
HyperDown	12	4	1	7	58.33	87.50
HypoUp	108	61	15	32	29.63	68.09
HyperUp	21	1	10	10	47.62	50.00
HypoDown	30	10	6	14	46.67	70.00
Total	171	76	32	63	36.84	66.32
Genic enhancer-like DMR-DEGs						
HyperDown	14	11	1	2	14.29	66.67
HypoUp	236	155	34	47	19.92	58.02
HyperUp	34	19	7	8	23.53	53.33
HypoDown	54	42	2	10	18.52	83.33
Total	338	227	44	67	19.82	60.36
Intergenic enhancer-like DMR-DEGs						
HyperDown	6	6	0	0	0.00	/
HypoUp	54	45	6	3	5.56	33.33
HyperUp	0	0	0	0	/	/
HypoDown	42	34	2	6	14.29	75.00
Total	102	85	8	9	8.82	52.94

*HyperDown, hypermethylation-associated downregulated gene; HypoUp, hypomethylation-associated upregulated gene; HyperUp, hypermethylation associated upregulated gene; HypoDown, hypomethylation associated with downregulated gene; Type I replication failure, no CpG available in 450k for corresponding DMR; Type II replication failure, at least one CpG available but no significant differential methylation-associated differential gene expression; replication rate = Count $_{passed\ replication}$/Count $_{in\ discovery}$ * 100; Adjusted replication rate = Count $_{passed\ replication}$/(Count $_{in\ discovery}$ - Count $_{type\ I\ failure}$) * 100.*

TABLE 2 | Top differentially expressed promoter- and genic enhancer-like DMR-DEGs and implicated cancers.

DMR location	Δ_{methy}	Gene name	LFC	Dist (bp)	Rho	Rho.padj	TCGA-LIHC replication	Implicated cancer
Promoter-likeDMR-DEGs								
chr8:101492028-101494873	0.30	GRHL2	-3.36	0	-0.52	9.70E-06	Replicated	HCC (50) and others (51, 52)
chr14:21022886-21023796	0.29	TPPP2	-2.93	-313	-0.79	<2.00E-16	Replicated	HCC (53)
chr6:160122738-160123492	0.25	SLC22A1	-2.54	949	-0.73	<2.00E-16	Type I failure	Unknown
chr4:52051223-52052036	0.26	SPATA18	-2.51	0	-0.56	1.90E-06	Replicated	Breast (54) and others (55, 56)
chr14:21022886-21023796.1	0.29	AL161668.4	-2.49	-376	-0.79	<2.00E-16	Type II failure	Unknown
chr1:58575901-58577030	0.32	TACSTD2	-2.31	-743	-0.39	0.0012	Replicated	Cholangiocarcinoma (57) and others (58)
chr8:85465316-85466086	0.17	CA2	-2.14	1464	-0.44	0.00056	Type I failure	HCC (59) and others (60)
chr7:128030940-128032690	0.37	LRRC4	-2.08	0	-0.53	7.20E-06	Replicated	Glioma (61) and Ovarian cancer (62)
chr6:133889000-133890152	0.19	TCF21	-1.81	0	-0.39	0.0015	Replicated	HCC (63) and others (64, 65)
chr1:118983368-118990519	0.31	TBX15	-1.12	0	-0.4	0.00098	Replicated	Renal cell carcinoma (66)
chr19:28606496-28607360	-0.25	AC079466.1	9.26	0	-0.58	6.30E-07	Type I failure	Unknown
chr7:153409385-153414141	-0.32	LINC01287	8.83	0	-0.53	7.70E-06	Type I failure	HCC (67) and others (68, 69)
chr13:64076307-64079199	-0.35	LINC00355	6.60	263	-0.55	3.00E-06	Type I failure	Gastric (70) and others (71)
chr10:17386395-17387833	-0.32	ST8SIA6-AS1	6.23	0	-0.65	3.10E-09	Type I failure	HCC (49) and others (72)
chrX:133987205-133987743	-0.24	GPC3	6.20	1310	-0.66	1.30E-09	Type I failure	HCC (73) and others (74)
chr22:45286491-45287226	-0.30	UPK3A	5.86	1509	-0.39	0.0012	Type I failure	Unknown
chr1:26862894-26863264	-0.26	SFN	5.30	0	-0.69	<2.00E-16	Replicated	HCC (75) and others (76)
chr15:23565016-23566853	-0.35	MKRN3	4.40	0	-0.71	1.80E-11	Replicated	Lung cancer (77)
chr1:43359350-43360396	-0.16	CDC20	4.18	395	-0.65	1.40E-08	Replicated	HCC (78) and others (79)
chr13:100088021-100088317	-0.27	AL355338.1	4.09	-531	-0.42	0.001	Type I failure	Unknown
Genic enhancer-like DMR-DEGs								
chr9:133368447-133369064	0.22	ADAMTS13	-3.27	-45294	-0.61	1.20E-06	Type I failure	Unknown
chr1:8004016-8004614	0.27	AL034417.4	-3.08	12882	-0.5	0.00012	Type I failure	Unknown
chr7:26365323-26365835	0.18	AC004540.2	-2.73	-10866	-0.66	7.20E-09	Type I failure	Unknown
chr19:3428124-3428792	0.22	SMIM24	-2.47	-51750	-0.41	0.0024	Type I failure	Unknown
chr5:172871061-172871507	0.29	DUSP1	-2.15	99866	-0.42	0.0022	Type I failure	HCC (80) and others (81)
chr11:66717972-66718255	0.18	SPTBN2	-1.99	-10971	-0.59	1.80E-06	Type I failure	Unknown
chr1:25686093-25686578	0.21	MAN1C1	-1.70	68625	-0.39	0.0012	Type I failure	Renal cell carcinoma (82)
chr16:2036508-2038495	0.27	SLC9A3R2	-1.26	11152	-0.57	1.40E-05	Replicated	Unknown
chr2:3535603-3535962	0.26	ADI1	-1.21	15867	-0.48	5.40E-05	Type I failure	HCC (83) and others (84)
chr17:4222301-4222630	0.18	CYB5D2	-1.20	79133	-0.46	0.00011	Replicated	Breast cancer (85) and others (86)
chr3:124858896-124859415	-0.15	MUC13	6.72	-94404	-0.52	1.70E-05	Type I failure	HCC (87) and others (88)
chr5:147773067-147773789	-0.21	SPINK1	5.45	-57997	-0.45	0.00051	Type I failure	HCC (89) and others (90)
chr20:43652556-43653516	-0.30	MYBL2	4.35	-13503	-0.59	1.00E-06	Type I failure	HCC (91) and others (92)
chr17:44905503-44905963	-0.21	C1QL1	4.22	-62108	-0.47	0.00081	Type I failure	Lung adenocarcinoma (93)
chr15:40100406-40101049	-0.21	BUB1B	4.16	-59974	-0.43	0.0016	Type I failure	Glioblastoma (94) and others (95)
chr6:44094402-44095522	-0.19	AL109615.3	4.08	19750	-0.42	0.002	Type II failure	Breast cancer (96)
chr1:44676848-44677387	-0.19	KIF2C	3.95	-62431	-0.66	1.10E-08	Type I failure	HCC (97) and others (98)
chr9:98862188-98863173	-0.24	COL15A1	3.81	-80006	-0.47	3.00E-04	Type I failure	Unknown
chr10:5284186-5284850	-0.18	LINC02561	3.80	12950	-0.64	4.10E-08	Type I failure	Unknown
chr5:176545906-176546537	-0.23	GPRIN1	3.23	-63596	-0.59	1.40E-06	Type I failure	Unknown

LFC, estimated log_2 transformation of fold change of gene expression between tumor and non-tumor (baseline) by DESeq2; Dist, distance between the DMR and associated target genes; Rho, the Spearman correlation coefficient between DNA methylation and gene expression; Rho.padj, BH-adjusted p-value of the Spearman correlation test; Type I failure, no CpG available in 450k for corresponding DMR; Type II failure, at least one CpG available but no significant differential methylation-associated differential gene expression.

genes were significantly associated with overall survival (OS) and progression-free survival (PFS) (**Figure 4A**), with a high overlap (72 genes) between the two sets of genes. The tumor stage-association test confirmed the significant association between the expression of 140 of the 661 methylation drivers and tumor progression, and 48 of them also were associated with survival (**Figure 4A** and **Supplementary Table S3.1**). As for methylation analyses, most of the 661 high confidence DMR-DEGs, except those 139 DMR-DEGs replicated in the TCGA-LIHC cohort, were excluded because of the low coverage in the 450k methylation array. Among those 139 genes, 27, 30, and 18 were highlighted as OS, PFS, and tumor stage-associated methylation biomarkers (**Figure 4B** and **Supplementary Table S3.2**). In the same manner, there were extensive overlaps among the three sets of clinically relevant genes (**Figure 4B**). Integration of DNA methylation and gene expression data highlighted six OS-associated, eight PFS-associated, and eight tumor stage-associated DMR-DEGs whose transcription and DNA

TABLE 3 | Top differentially expressed intergenic enhancer-like DMR-DEGs and implicated cancers.

DMR location	Δ_{methy}	Gene name	LFC	Dist (bp)	Rho.Me	Rho.eG	Rho.MG	TCGA-LIHC replication	Implicated cancer
chr11:1667908-1668094	0.23	FAM99A	-2.68	2311	-0.49	0.92	-0.45	Type I failure	HCC (99)
chr15:74758829-74759080	0.19	CYP1A2	-4.56	9985	-0.52	0.86	-0.55	Type I failure	HCC (100)
chr18:31601037-31601741	0.20	TTR	-2.35	44027	-0.47	0.72	-0.74	Type I failure	Lung cancer (101)
chr4:154603041-154603623	0.17	FGG	-1.35	-9344	-0.52	0.72	-0.75	Type I failure	HCC (102) and others (103)
chr4:154603041-154603623	0.17	FGB	-1.56	40085	-0.52	0.74	-0.73	Type I failure	Renal cell carcinoma (104)
chr5:147837794-147838222	-0.19	SPINK1	5.45	6008	-0.64	0.82	-0.64	Type I failure	HCC (105) and others (90)
chrX:109740935-109741180	-0.19	ACSL4	3.76	7532	-0.44	0.7	-0.4	Type I failure	HCC (106) and others (107)
chr14:19242162-19242285	-0.23	DUXAP10	3.70	-95445	-0.46	0.83	-0.39	Type I failure	HCC (108) and others (109)
chr4:49513136-49516430	-0.27	AC119751.4	3.43	-63403	-0.57	0.72	-0.65	Type II failure	Unknown
chr18:6683286-6684199	-0.34	ARHGAP28	2.37	-45519	-0.63	0.71	-0.43	Type I failure	Unknown
chr8:122448922-122449306	-0.17	SMILR	2.16	20371	-0.36	0.76	-0.4	Type I failure	Unknown
chr8:144145816-144146462	-0.21	TSSK5P	2.15	2152	-0.68	0.73	-0.66	Type I failure	Unknown
chr1:109309171-109309440	-0.18	SORT1	2.06	-88511	-0.5	0.74	-0.54	Type I failure	Lung cancer and others (110)
chr8:122448922-122449306	-0.17	AC108136.1	1.93	-40509	-0.36	0.71	-0.4	Type I failure	Unknown
chr8:144347135-144347612	-0.28	TONSL	1.72	-96832	-0.43	0.71	-0.61	Type I failure	Gastric cancer (111) and others (112)

LFC, estimated log_2 transformation of fold change of gene expression between tumor and non-tumor (baseline) by DESeq2; Dist, distance between the DMR and associated target gene; Rho.Me, the Spearman correlation coefficient between DNA methylation and enhancer RNA expression; Rho.eG, the Spearman correlation coefficient between enhancer RNA expression and associated target gene expression; Rho.MG, the Spearman correlation coefficient between DNA methylation and associated target gene expression; Type I failure, No CpG available in 450k for corresponding DMR; Type II failure, at least one CpG available but no significant differential methylation-associated differential gene expression.

methylation were both plausible biomarkers (**Figure 4C**). Interestingly, five of those six genes whose expression and methylation were both significant OS biomarkers were more powerful (lower log-rank p-values) prognostic biomarkers when considering both expression and methylation data together (**Figure 4D** and **Supplementary Table S3.3**). Similarly, seven of those eight genes whose expression and methylation were both significant PFS biomarkers show higher significance (lower log-rank p-values) when considering both expression and methylation together (**Figure 4E** and **Supplementary Table S3.3**). In addition, four genes (CDC20, UCK2, HEATR6, and SLC9A3R2) were concordantly associated with both survival duration and tumor progression (**Figure 4F** and **Supplementary Tables S3.1, 3.2**).

Successful *In Vitro* Demethylation Treatment-Based Validation of DMR-DEGs

For the sake of validation, 15 top differentially expressed STRING protein–protein interaction (PPI) network hub genes of the overrepresented pathways of DMR-DEGs, including cell cycle, DNA repair, and metabolic pathways, and nine genes significantly associated with OS, PFS, and/or tumor stages were selected for in vitro DNA methylation unmasking validation in LO2 and HepG2 cells. After 5-azadC treatment, 20/23 (87.0%) and 15/23 (65.2%) of those selected genes (one gene, named CDC20, belonged to both the pathway hub gene and the clinically relevant gene) showed significant upregulation in LO2 and HepG2, respectively. For instance, four of the five hub genes in the cell cycle and all six hub genes related to the DNA repair pathway showed apparent upregulation in LO2, and a majority of them also presented upregulation in HepG2 after methylation unmasking (**Figures 5A, B**). Likewise, we obtained similar high validation rates for the remaining pathways (**Supplementary Figure S4**). Furthermore, after 5-azadC treatment in LO2 and HepG2, the qRT-PCR results showed significant upregulation of all seven survival-associated genes and five of the six tumor stage-associated genes (**Figures 5C, D**).

DISCUSSION

In recent years, large-scale genome-wide DNA methylome studies using methylation array and next-generation sequencing technologies have reshaped our understanding of epigenetic aberrations' vital roles in tumor formation and maintenance. Owing to the technical merits of WGBS, we identified about 9.8 million differentially methylated CpGs and more than 600,000 regional differential methylations, most of which were located in the intergenic and intronic regions, which could rarely be discovered by the array or target sequencing platform. To the best of our knowledge, this work is the third, but the largest, WGBS-based DNA methylation study in HCC. Considering the small samples of the previous two WGBS studies with sample sizes of eight (five tumor samples and three non-tumor control samples) in one study (24) and of five (two tumor samples and three non-tumor control samples) in the other study (25), they might lack sufficient power to detect methylomic aberrations comprehensively, given the heterogeneity of tumor tissues.

It is well established that HCC is a complex disease contributed to by a disrupted genome that harbors numerous genetic mutations and epigenetic aberrations during the development and maintenance of liver carcinogenesis. Applying an integration of multi-omics data to gain a deeper understanding of the hepatocarcinogenesis mechanisms underlying HCC has become increasingly popular. For instance, the integration of multiple epigenomics data that included DNA methylation, DNA hydromethylation, and four types of histone ChIP-seq data

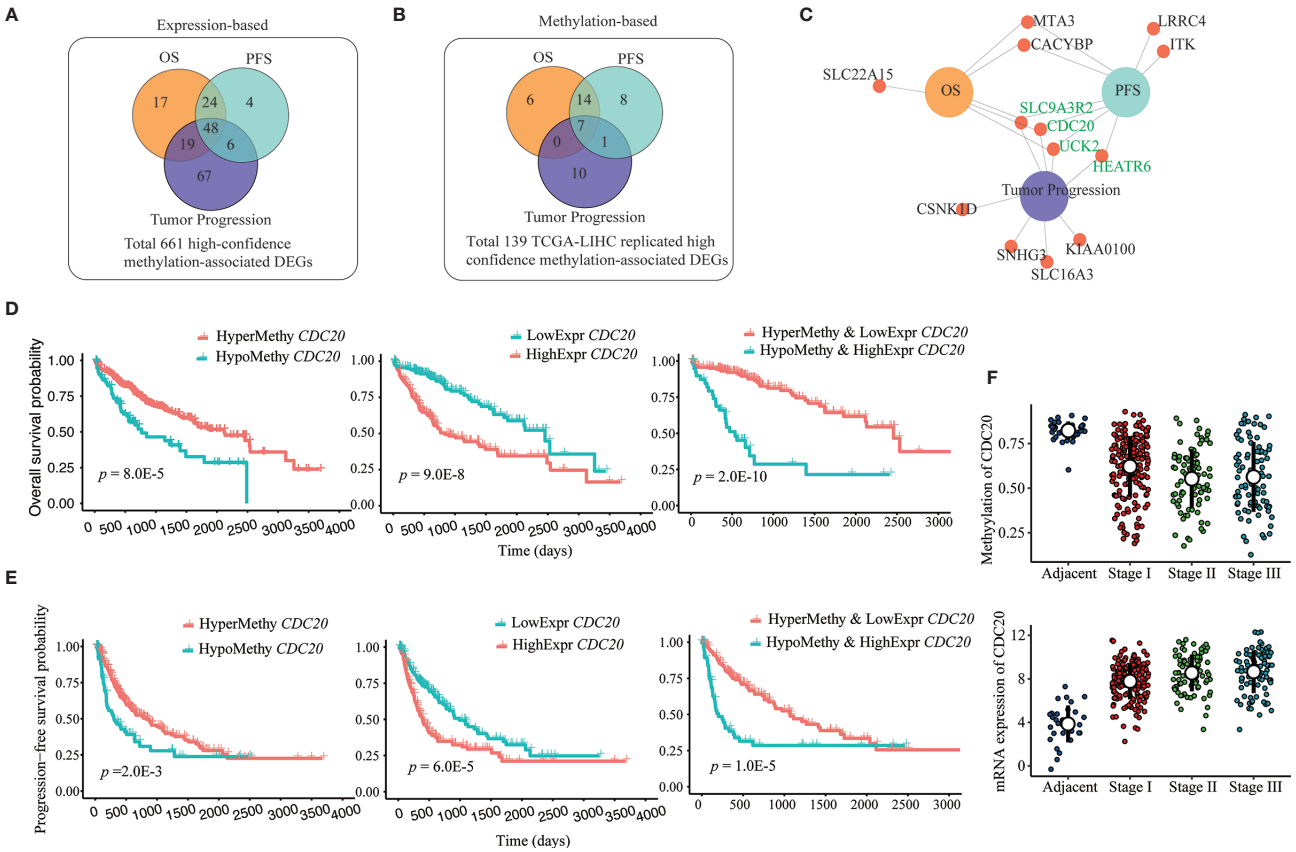

FIGURE 4 | Prognosis and tumor progression-associated biomarkers among DMR-DEGs in HCC. **(A, B)** Overlap between identified survival (OS and PFS) and tumor progression (tumor stage)-associated expression and methylation-based biomarkers among identified 661 high-confidence DMR-DEGs. **(C)** DMR-DEGs, whose expression and methylation were significantly associated with survival and/or tumor progression in HCC. Four genes that associated with both survival and tumor stages are highlighted in green. **(D)** Combination of both promoter hypomethylation and gene upregulation of CDC20 identified more efficient OS biomarkers. **(E)** Combination of both promoter hypomethylation and gene upregulation of CDC20 identified more powerful PFS biomarkers. **(F)** Progressively hypomethylated promoter and increased expression of CDC20 in HCC.

identified novel tumor-suppressor genes for HCC (42). Besides integration with genomic mutations or other epigenomic data, DNA methylomic data most commonly were integrated with genome-wide transcriptome profiling for identification of potential methylation-associated tumor suppressors or oncogenes (113). In our DMR-DEG identification procedure, DNA methylation was systematically integrated with genome-wide gene expression profiling, intergenic eRNA expression, and histone ChIP-seq data. The combination of histone ChIP-seq peak signals and active eRNA expression with DNA methylation profiling contributed to identifying differentially methylated transcription regulatory elements effectively and credibly, followed by integration of gene expression to identify significantly correlated nearby genes as candidate DMR-DEGs for downstream replication and validation. Thus, our strategy would be more powerful and reliable to identify epigenetic drivers with high confidence, in contrast to a simple integrative analysis of the DNA methylome and transcriptome.

Through integrating WGBS-based DNA methylation profiling and RNA-seq based transcriptomic data from paired tumor and adjacent tissues of 33 HCC patients, along with the integration of liver histone ChIP-seq data from the public domain, we identified 661 differential methylated promoter/enhancer-associated target genes and replicated 139 of them in the TCGA-LIHC cohort, which is a high, platform-adjusted, independent replication rate. Moreover, the set of high-confidence DMR-DEGs contains a high proportion of previously experimentally validated HCC driver genes, many other cancer-relevant genes, and some uncharacterized genes with considerable biological function, for instance, cell division cycle 20 (CDC20), a critical coactivator of the cellular division essential complex—anaphase-promoting complex/cyclosome (APC/C), whose overexpression has been associated with the development of a multitude of cancers such as those of the prostate (79) and liver (78). Silencing of CDC20 introduced effective antitumor activity into the orthotopic liver tumor model (114). Although CDC20 is prevalently overexpressed in HCCs (115), the underlying mechanism still was obscure. In our HCC sample, we identified significantly correlated the hypomethylated promoter and transcriptional activation of CDC20, which was replicated successfully in the TCGA-LIHC

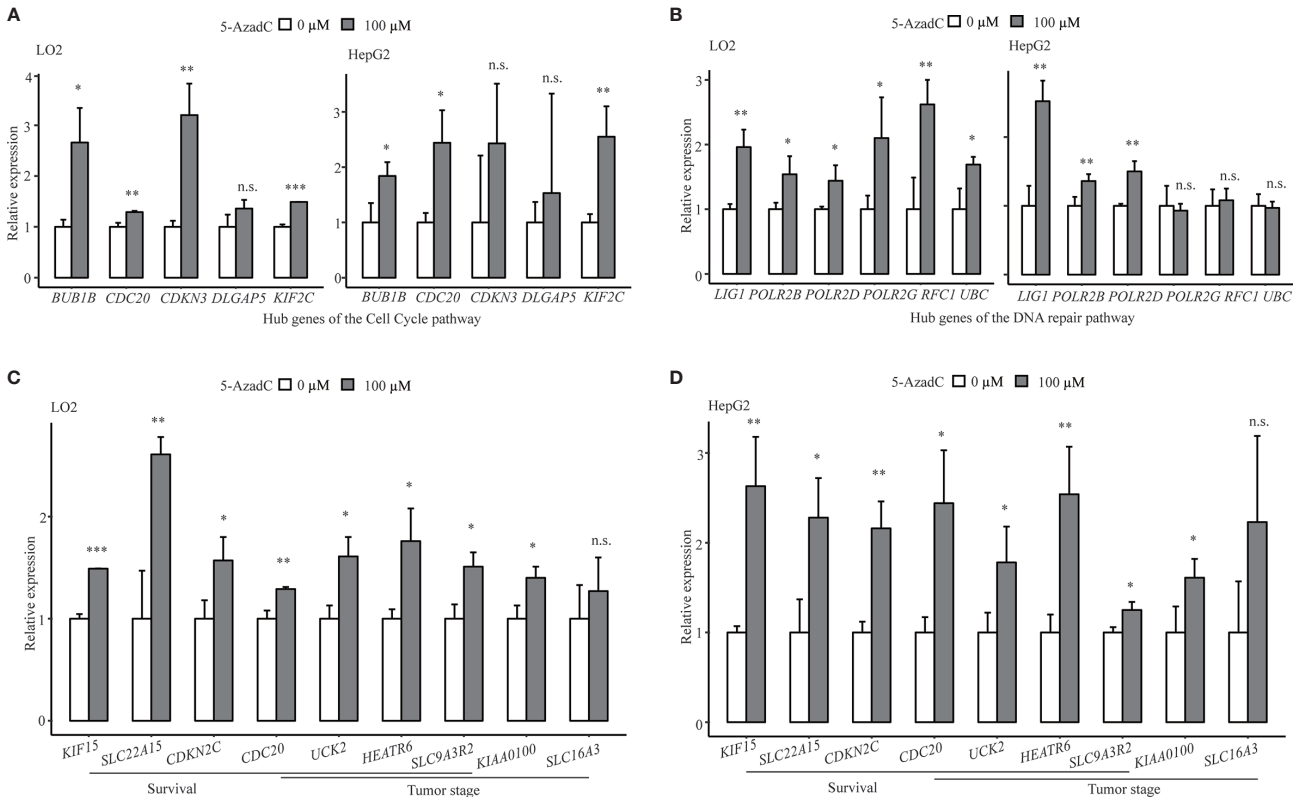

FIGURE 5 | Successful *in vitro* validation of methylation-mediated transcriptional regulation of key pathway hub genes and clinical associated genes. **(A)** Demethylation validation results of five hub genes of cell cycle pathway in LO2 and HepG2 cells. **(B)** Demethylation validation results of selected hub genes from DNA repair pathway in LO2 and HepG2 cells. **(C, D)** Demethylation validation results of selected survival (OS or PFS) and tumor stage-associated genes in LO2 and HepG2 cells. n.s., non-significant; *p < 0.05, **p < 0.01, and ***p < 0.001.

dataset and validated in methylation-unmasked LO2 and HepG2. Hence, hypomethylated promoter-associated activation might represent a plausible mechanism underlying the widespread dysregulation of CDC20 in HCC. Besides, other known HCC-related genes such as TPPP2 (53), TCF21 (63), GRHL2 (50), and CA2 (59) also were found to be negatively regulated by aberrant promoter DNA methylation in our study. Moreover, the spermatogenesis-associated protein 18 (SPATA18) is a p53-inducible protein involved in the mitochondrial quality-control process, whose dysregulation is associated with cancer. Unlike CDC20, the role of SPATA18 is uncharacterized in HCC, although it also showed concurrent transcriptional repression (115). However, SPATA18 was reported to suppress growth of murine intestinal tumor (116) and human breast cancer (54) *via* mitochondrial quality control. Therefore, our integrative epigenomic analysis might shed new light on the epigenetic-mediated roles of novel genes such as SPATA18 in the process of liver carcinogenesis. Furthermore, accumulating evidence indicates the significance of aberrant enhancer-mediated transcriptional dysregulation in the formation and maintenance of multiple tumors (117, 118), including HCC (25). In the present study, we identified abundant hypomethylated enhancer-associated activated HCC-related genes such as MUC13 (119), SPINK1 (105), and KIF2C (97), plus hypermethylated enhancer-associated repression of known HCC suppressors like DUSP1 (80) and ADI1 (83). Additionally, we found aberrant enhancer-associated dysregulation of genes whose functions are uncharacterized in cancer but harbor possibly essential biological functions. For example, the liver-specific long non-coding (Lnc) gene, FAM99A, was characterized only a few months ago as a powerful regulator of metastasis of HCC (99).

It is well known that DNA methylation modulates gene transcription in negative regulation, especially promoter hypermethylation-induced silencing of tumor suppre sors, which is a hallmark of most cancers. However, there were new studies suggesting that the effect of DNA methylation of CpG islands in gene bodies on transcriptional regulation is different (21, 120). Furthermore, some transcription factors such as CEBPB (121) and RXRA (122) have been reported to prefer methylated CpGs in their binding sites, suggesting a positive correlation between promoter/enhancer methylation and gene transcription. A newly published prostate cancer study also reported extensive, robust associations between DNA hypermethylation and gene upregulation (123), indicating the diversity of epigenetic regulation. In our findings, 181 of those 661 high-confidence DMR-DEGs displayed a positive correlation

between methylation and gene expression, and 48 of these 181 genes represented the same non-classical association in the TCGA-LIHC cohort. These 48 genes also contained several noted HCC-relevant genes, such as CELSR3 (124) and PCK1 (125), as well as some genes involved in other cancers, such as NTF3 in breast cancer (126) and TIMD4 in B-cell lymphoma (127).

Our integrative analysis advances the understanding of the disordered methylome of HCC, although there still are several potential limitations. We implemented the first relatively large-scale WGBS-based global DNA methylome profiling of paired tumor and adjacent non-tumor tissues from 33 HCC patients, which covered almost all gene body and intergenic CpG islands that could barely be estimated by the 450k methylation array or target sequencing. However, the average depth of WGBS samples in our study was medium because of the significant cost of WGBS. Besides, we identified a total of 661 high-confidence differentially methylated promoter/enhancer-associated DEGs and achieved a high ratio of successful replication in the TCGA-LIHC cohort after platform limitation adjustment, whereas the considerable DMR-DEGs suffered from lack of replication. Further replication in a larger independent cohort with WGBS-based DNA methylation profiling is greatly needed, which might be a promising method of discovering more epigenetic drivers associated with aberrant methylation in the gene body and intergenic regions. In addition, further validation of methylation-mediated regulation of particular genes *via* technologies like CRISPR, like previous studies (25, 26), were lacking in our present study but would be part of our ongoing works. Besides, it would be better if histone modification, DNA methylation, and gene expression were performed in the same samples, which would provide a more accurate functional annotation of identified DMRs.

Collectively, our integrative analysis of epigenome and transcriptome of HCC convincingly proved the powerful potential of WGBS in uncovering the global DNA methylation aberrations in HCC, especially numerous enhancers in the intron and intergenic regions. Specifically, we identified a group of 661 DMR-DEGs with high confidence, and they were substantially replicated in an independent cohort and validated by *in vitro* methylation unmasking experiments. Intriguingly, those genes reflected a high percentage of known HCC or other cancer-relevant vital genes. These findings depicted activated pathways such as those for the cell cycle and DNA repair and repressed key metabolic pathways induced by aberrant DNA methylation of promoters and enhancers in HCC. Beyond those results, our perfectly matched methylome and transcriptome sequencing data from relatively large-scale paired tumoral and adjacent non-tumoral tissues also provide a valuable resource for follow-up studies in HCC, in which WGBS-based methylome data were insufficient.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at the NCBI SRA database [accession: PRJNA762641].

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of the First Affiliated Hospital of Zhejiang University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

Conceptualization, PH, MDL, and ZY. Data curation, PH, MX, HH, XZ, MDL, and ZY. Formal analysis, PH, HH, and XZ. Funding acquisition, MDL. Investigation, MX and HH. Methodology, PH, MX, XZ, and MDL. Project administration, MDL and ZY. Resources, MDL and ZY. Software, PH and XZ. Supervision, MDL and ZY. Validation, PH, MX, HH, and ZY. Visualization, PH. Writing—original draft, PH. Writing—review and editing, PH, MDL, and ZY. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported in part by the China Precision Medicine Initiative (2016YFC0906300), Research Center for Air Pollution and Health of Zhejiang University, and the Independent Task of State Key Laboratory for Diagnosis and Treatment of Infectious Diseases.

ACKNOWLEDGMENTS

We thank Dr. David L. Bronson for his excellent editing of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.769390/full#supplementary-material

REFERENCES

1. Bruix J, Gores GJ, Mazzaferro V. Hepatocellular Carcinoma: Clinical Frontiers and Perspectives. *Gut* (2014) 63(5):844–55. doi: 10.1136/gutjnl-2013-306627
2. Villanueva A. Hepatocellular Carcinoma. *N Engl J Med* (2019) 380 (15):1450–62. doi: 10.1056/NEJMra1713263
3. Yan YP, Su HX, Ji ZH, Shao ZJ, Pu ZS. Epidemiology of Hepatitis B Virus Infection in China: Current Status and Challenges. *J Clin Trans Hepatol* (2014) 2(1):15–22. doi: 10.14218/JCTH.2013.00030

4. Siegel R, Naishadham D, Jemal A. Cancer Statistics, 2013. *CA: Cancer J Clin* (2013) 63(1):11–30. doi: 10.3322/caac.21166
5. Jee BA, Choi JH, Rhee H, Yoon S, Kwon SM, Nahm JH, et al. Dynamics of Genomic, Epigenomic, and Transcriptomic Aberrations During Stepwise Hepatocarcinogenesis. *Cancer Res* (2019) 79(21):5500–12. doi: 10.1158/0008-5472.CAN-19-0991
6. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-Wide Association Studies for Common Human Diseases. *Nat Rev Genet* (2011) 12(8):529–41. doi: 10.1038/nrg3000
7. Huidobro C, Torano EG, Fernandez AF, Urdinguio RG, Rodriguez RM, Ferrero C, et al. A DNA Methylation Signature Associated With the Epigenetic Repression of Glycine N-Methyltransferase in Human Hepatocellular Carcinoma. *J Mol Med* (2013) 91(8):939–50. doi: 10.1007/s00109-013-1010-8
8. 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(1):34–43. doi: 10.4161/epi.23062
9. Fan X, Li Y, Yi X, Chen G, Jin S, Dai Y, et al. Epigenome-Wide DNA Methylation Profiling of Portal Vein Tumor Thrombosis (PVTT) Tissues in Hepatocellular Carcinoma Patients. *Neoplasia* (2020) 22(11):630–43. doi: 10.1016/j.neo.2020.09.007
10. Zhang C, Ge S, Wang J, Jing X, Li H, Mei S, et al. Epigenomic Profiling of DNA Methylation for Hepatocellular Carcinoma Diagnosis and Prognosis Prediction. *J Gastroenterol Hepatol* (2019) 34(10):1869–77. doi: 10.1111/jgh.14694
11. Gao F, Liang H, Lu H, Wang J, Xia M, Yuan Z, et al. Global Analysis of DNA Methylation in Hepatocellular Carcinoma by a Liquid Hybridization Capture-Based Bisulfite Sequencing Approach. *Clin Epigenet* (2015) 7 (1):1–11. doi: 10.1186/s13148-015-0121-1
12. Nishida N, Kudo M, Nagasaka T, Ikai I, Goel A. Characteristic Patterns of Altered DNA Methylation Predict Emergence of Human Hepatocellular Carcinoma. *Hepatology* (2012) 56(3):994–1003. doi: 10.1002/hep.25706
13. 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(6):1424–35.e1425. doi: 10.1053/j.gastro.2013.08.055
14. Toh TB, Lim JJ, Chow EK-H. Epigenetics of Hepatocellular Carcinoma. *Clin Trans Med* (2019) 8(1):13. doi: 10.1186/s40169-019-0230-0
15. Nishida N, Goel A. Genetic and Epigenetic Signatures in Human Hepatocellular Carcinoma: A Systematic Review. *Curr Genomics* (2011) 12(2):130–7. doi: 10.2174/138920211795564359
16. Fukai K, Yokosuka O, Imazeki F, Tada M, Mikata R, Miyazaki M, et al. Methylation Status of P14arf, P15ink4b, and P16ink4a Genes in Human Hepatocellular Carcinoma. *Liver Int* (2005) 25(6):1209–16. doi: 10.1111/j.1478-3231.2005.01162.x
17. Wei Y, Van Nhieu JT, Prigent S, Srivatanakul P, Tiollais P, Buendia M-A. Altered Expression of E-Cadherin in Hepatocellular Carcinoma: Correlations With Genetic Alterations, β-Catenin Expression, and Clinical Features. *Hepatology* (2002) 36(3):692–701. doi: 10.1053/jhep.2002.35342
18. Yeo W, Wong N, Wong WL, Lai PB, Zhong S, Johnson PJ. High Frequency of Promoter Hypermethylation of RASSF1A in Tumor and Plasma of Patients With Hepatocellular Carcinoma. *Liver Int* (2005) 25(2):266–72. doi: 10.1111/j.1478-3231.2005.01084.x
19. Mori T, Nomoto S, Koshikawa K, Fujii T, Sakai M, Nishikawa Y, et al. Decreased Expression and Frequent Allelic Inactivation of the RUNX3 Gene at 1p36 in Human Hepatocellular Carcinoma. *Liver Int* (2005) 25(2):380–8. doi: 10.1111/j.1478-3231.2005.1059.x
20. Okochi O, Hibi K, Sakai M, Inoue S, Takeda S, Kaneko T, et al. Methylation-Mediated Silencing of SOCS-1 Gene in Hepatocellular Carcinoma Derived From Cirrhosis. *Clin Cancer Res* (2003) 9(14):5295–8.
21. Arechederra M, Daian F, Yim A, Bazai SK, Richelme S, Dono R, et al. Hypermethylation of Gene Body Cpg Islands Predicts High Dosage of Functional Oncogenes in Liver Cancer. *Nat Commun* (2018) 9(1):3164. doi: 10.1038/s41467-018-06482-w
22. 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(6):1945–56. doi: 10.1002/hep.27732
23. Calvisi DF, Ladu S, Gorden A, Farina M, Lee J-S, Conner EA, et al. Mechanistic and Prognostic Significance of Aberrant Methylation in the Molecular Pathogenesis of Human Hepatocellular Carcinoma. *J Clin Invest* (2007) 117(9):2713–22. doi: 10.1172/JCI31457
24. Hama N, Totoki Y, Miura F, Tatsuno K, Saito-Adachi M, Nakamura H, et al. Epigenetic Landscape Influences the Liver Cancer Genome Architecture. *Nat Commun* (2018) 9(1):1643. doi: 10.1038/s41467-018-03999-y
25. Xiong L, Wu F, Wu Q, Xu L, Cheung OK, Kang W, et al. Aberrant Enhancer Hypomethylation Contributes to Hepatic Carcinogenesis Through Global Transcriptional Reprogramming. *Nat Commun* (2019) 10(1):335. doi: 10.1038/s41467-018-08245-z
26. Chen H, Li C, Peng X, Zhou Z, Weinstein JNCancer Genome Atlas Research N, et al. A Pan-Cancer Analysis of Enhancer Expression in Nearly 9000 Patient Samples. *Cell* (2018) 173(2):386–399 e312. doi: 10.1016/j.cell.2018.03.027
27. Zhang Z, Lee JH, Ruan H, Ye Y, Krakowiak J, Hu Q, et al. Transcriptional Landscape and Clinical Utility of Enhancer Rnas for Erna-Targeted Therapy in Cancer. *Nat Commun* (2019) 10(1):4562. doi: 10.1038/s41467-019-12543-5
28. Chatterjee A, Rodger EJ, Morison IM, Eccles MR, Stockwell PA. Tools and Strategies for Analysis of Genome-Wide and Gene-Specific DNA Methylation Patterns. *Methods Mol Biol* (2017) 1537:249–77. doi: 10.1007/978-1-4939-6685-1_15
29. Wang M, Zhao J, Wang Y, Mao Y, Zhao X, Huang P, et al. Genome-Wide DNA Methylation Analysis Reveals Significant Impact of Long-Term Ambient Air Pollution Exposure on Biological Functions Related to Mitochondria and Immune Response. *Environ Pollution (Barking Essex: 1987)* (2020) 264:114707. doi: 10.1016/j.envpol.2020.114707
30. Martin M. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnet J* (2011) 17(1):10–2. doi: 10.14806/ej.17.1.200
31. Bolger AM, Lohse M, Usadel B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* (2014) 30(15):2114–20. doi: 10.1093/bioinformatics/btu170
32. Bray NL, Pimentel H, Melsted P, Pachter L. Near-Optimal Probabilistic RNA-Seq Quantification. *Nat Biotechnol* (2016) 34(5):525–7. doi: 10.1038/nbt.3519
33. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: The Reference Human Genome Annotation for the ENCODE Project. *Genome Res* (2012) 22(9):1760–74. doi: 10.1101/gr.135350.111
34. Soneson C, Love MI, Robinson MD. Differential Analyses for RNA-Seq: Transcript-Level Estimates Improve Gene-Level Inferences. *F1000Res* (2015) 4:1521. doi: 10.12688/f1000research.7563.1
35. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* (2013) 29(1):15–21. doi: 10.1093/bioinformatics/bts635
36. Quinlan AR, Hall IM. Bedtools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* (2010) 26(6):841–2. doi: 10.1093/bioinformatics/btq033
37. Love MI, Huber W, Anders S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data With Deseq2. *Genome Biol* (2014) 15(12):550. doi: 10.1186/s13059-014-0550-8
38. Krueger F, Andrews SR. Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications. *Bioinformatics* (2011) 27(11):1571–2. doi: 10.1093/bioinformatics/btr167
39. Park Y, Wu H. Differential Methylation Analysis for BS-Seq Data Under General Experimental Design. *Bioinformatics* (2016) 32(10):1446–53. doi: 10.1093/bioinformatics/btw026
40. Huang H, Chen Z, Huang X. Age-Adjusted Nonparametric Detection of Differential DNA Methylation With Case-Control Designs. *BMC Bioinf* (2013) 14:86. doi: 10.1186/1471-2105-14-86
41. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, et al. Software for Computing and Annotating Genomic Ranges. *PloS Comput Biol* (2013) 9(8):e1003118. doi: 10.1371/journal.pcbi.1003118
42. Hlady RA, Sathyanarayan A, Thompson JJ, Zhou D, Wu Q, Pham K, et al. Integrating the Epigenome to Identify Drivers of Hepatocellular Carcinoma. *Hepatology* (2019) 69(2):639–52. doi: 10.1002/hep.30211
43. Consortium EP:. The ENCODE (Encyclopedia of DNA Elements) Project. *Science* (2004) 306(5696):636–40. doi: 10.1126/science.1105136

44. Zhu Y, Sun L, Chen Z, Whitaker JW, Wang T, Wang W. Predicting Enhancer Transcription and Activity From Chromatin Modifications. *Nucleic Acids Res* (2013) 41(22):10032–43. doi: 10.1093/nar/gkt826
45. Wu H, Nord AS, Akiyama JA, Shoukry M, Afzal V, Rubin EM, et al. Tissue-Specific RNA Expression Marks Distant-Acting Developmental Enhancers. *PloS Genet* (2014) 10(9):e1004610. doi: 10.1371/journal.pgen.1004610
46. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6
47. BP KM: RTCGA. *The Cancer Genome Atlas Data Integration. R Bioconductor Package* (2015). Available at: https://rtcga.github.io.RTCGA.
48. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* (2018) 173(2):400–416.e411. doi: 10.1016/j.cell.2018.02.052
49. Li Y, Jiang A. ST8SIA6-AS1 Promotes Hepatocellular Carcinoma by Absorbing MiR-5195-3p to Regulate HOXB6. *Cancer Biol Ther* (2020) 21 (7):647–55. doi: 10.1080/15384047.2020.1743150
50. Tanaka Y, Kanai F, Tada M, Tateishi R, Sanada M, Nannya Y, et al. Gain of GRHL2 Is Associated With Early Recurrence of Hepatocellular Carcinoma. *J Hepatol* (2008) 49(5):746–57. doi: 10.1016/j.jhep.2008.06.019
51. Xiang X, Deng Z, Zhuang X, Ju S, Mu J, Jiang H, et al. Grhl2 Determines the Epithelial Phenotype of Breast Cancers and Promotes Tumor Progression. *PloS One* (2012) 7(12):e50781. doi: 10.1371/journal.pone.0050781
52. Pan X, Zhang R, Xie C, Gan M, Yao S, Yao Y, et al. GRHL2 Suppresses Tumor Metastasis *via* Regulation of Transcriptional Activity of Rhog in Non-Small Cell Lung Cancer. *Am J Transl Res* (2017) 9(9):4217–26.
53. Du Y, Kong G, You X, Zhang S, Zhang T, Gao Y, et al. Elevation of Highly Up-Regulated in Liver Cancer (HULC) by Hepatitis B Virus X Protein Promotes Hepatoma Cell Proliferation *via* Down-Regulating P18. *J Biol Chem* (2012) 287(31):26302–11. doi: 10.1074/jbc.M112.342113
54. Gaowa S, Futamura M, Tsuneki M, Kamino H, Tajima JY, Mori R, et al. Possible Role of P53/Mieap-Regulated Mitochondrial Quality Control as a Tumor Suppressor in Human Breast Cancer. *Cancer Sci* (2018) 109 (12):3910–20. doi: 10.1111/cas.13824
55. Sano H, Futamura M, Gaowa S, Kamino H, Nakamura Y, Yamaguchi K, et al. P53/Mieap-Regulated Mitochondrial Quality Control Plays an Important Role as a Tumor Suppressor in Gastric and Esophageal Cancers. *Biochem Biophys Res Commun* (2020) 529(3):582–9. doi: 10.1016/j.bbrc.2020.05.168
56. Mussazhanova Z, Shimamura M, Kurashige T, Ito M, Nakashima M, Nagayama Y. Causative Role for Defective Expression of Mitochondria-Eating Protein in Accumulation of Mitochondria in Thyroid Oncocytic Cell Tumors. *Cancer Sci* (2020) 111(8):2814–23. doi: 10.1111/cas.14501
57. Sawanyawisuth K, Tantapotinan N, Wongkham C, Riggins GJ, Kraiklang R, Wongkham S, et al. Suppression of Trophoblast Cell Surface Antigen 2 Enhances Proliferation and Migration in Liver Fluke-Associated Cholangiocarcinoma. *Ann Hepatol* (2016) 15(1):71–81. doi: 10.5604/16652681.1184223
58. Nakanishi H, Taccioli C, Palatini J, Fernandez-Cymering C, Cui R, Kim T, et al. Loss of MiR-125b-1 Contributes to Head and Neck Cancer Development by Dysregulating TACSTD2 and MAPK Pathway. *Oncogene* (2014) 33(6):702–12. doi: 10.1038/onc.2013.13
59. Zhang C, Wang H, Chen Z, Zhuang L, Xu L, Ning Z, et al. Carbonic Anhydrase 2 Inhibits Epithelial-Mesenchymal Transition and Metastasis in Hepatocellular Carcinoma. *Carcinogenesis* (2018) 39(4):562–70. doi: 10.1093/carcin/bgx148
60. Angeli A, Carta F, Nocentini A, Winum JY, Zalubovskis R, Akdemir A, et al. Carbonic Anhydrase Inhibitors Targeting Metabolism and Tumor Microenvironment. *Metabolites* (2020) 10(10):412. doi: 10.3390/metabo10100412
61. Tang H, Wang Z, Liu X, Liu Q, Xu G, Li G, et al. LRRC4 Inhibits Glioma Cell Growth and Invasion Through a MiR-185-Dependent Pathway. *Curr Cancer Drug Targets* (2012) 12(8):1032–42. doi: 10.2174/156800912803251180
62. Zhao C, She X, Zhang Y, Liu C, Li P, Chen S, et al. LRRC4 Suppresses E-Cadherin-Dependent Collective Cell Invasion and Metastasis in Epithelial Ovarian Cancer. *Front Oncol* (2020) 10:144. doi: 10.3389/fonc.2020.00144
63. França MM, Ferraz-de-Souza B, Lerario AM, Fragoso MCBV, Lotfi CFP. POD-1/TCF21 Reduces SHP Expression, Affecting LRH-1 Regulation and Cell Cycle Balance in Adrenocortical and Hepatocarcinoma Tumor Cells. *BioMed Res Int* (2015) 2015:841784. doi: 10.1155/2015/841784
64. Duan HX, Li BW, Zhuang X, Wang LT, Cao Q, Tan LH, et al. TCF21 Inhibits Tumor-Associated Angiogenesis and Suppresses the Growth of Cholangiocarcinoma by Targeting PI3K/Akt and ERK Signaling. *Am J Physiol Gastrointest Liver Physiol* (2019) 316(6):G763–73. doi: 10.1152/ajpgi.00264.2018
65. Ao X, Ding W, Zhang Y, Ding D, Liu Y. TCF21: A Critical Transcription Factor in Health and Cancer. *J Mol Med* (2020) 98(8):1055–68. doi: 10.1007/s00109-020-01934-7
66. Zheng Z, Chen Z, Zhong Q, Zhu D, Xie Y, Shangguan W, et al. Circpvt1 Promotes Progression in Clear Cell Renal Cell Carcinoma by Sponging MiR-145-5p and Regulating TBX15 Expression. *Cancer Sci* (2021) 112(4):1443–56. doi: 10.1111/cas.14814
67. Mo Y, He L, Lai Z, Wan Z, Chen Q, Pan S, et al. LINC01287/MiR-298/STAT3 Feedback Loop Regulates Growth and the Epithelial-to-Mesenchymal Transition Phenotype in Hepatocellular Carcinoma Cells. *J Exp Clin Cancer Res* (2018) 37(1):149. doi: 10.1186/s13046-018-0831-2
68. Fu D, Ren Y, Wang C, Yu L, Yu R. LINC01287 Facilitates Proliferation, Migration, Invasion and EMT of Colon Cancer Cells *via* MiR-4500/MAP3K13 Pathway. *BMC Cancer* (2021) 21(1):782. doi: 10.1186/s12885-021-08528-7
69. Zhang J, Ma D, Kang H, Zhao J, Yang M. Long Noncoding RNA LINC01287 Promotes Proliferation and Inhibits Apoptosis of Lung Adenocarcinoma Cells *via* the MiR-3529-5p/RNASEH2A Axis Under the Competitive Endogenous RNA Pattern. *Environ Toxicol* (2021) 36(10):2093–104. doi: 10.1002/tox.23325
70. Zhao W, Jin Y, Wu P, Yang J, Chen Y, Yang Q, et al. LINC00355 Induces Gastric Cancer Proliferation and Invasion Through Promoting Ubiquitination of P53. *Cell Death Discov* (2020) 6:99. doi: 10.1038/s41420-020-00332-9
71. Lu S, Sun Z, Tang L, Chen L. LINC00355 Promotes Tumor Progression in HNSCC by Hindering Microrna-195-Mediated Suppression of HOXA10 Expression. *Mol Ther Nucleic Acids* (2020) 19:61–71. doi: 10.1016/j.omtn.2019.11.002
72. Luo ML, Li J, Shen L, Chu J, Guo Q, Liang G, et al. The Role of APAL/ST8SIA6-AS1 LncRNA in PLK1 Activation and Mitotic Catastrophe of Tumor Cells. *J Natl Cancer Inst* (2020) 112(4):356–68. doi: 10.1093/jnci/djz134
73. Jiang Z, Jiang X, Chen S, Lai Y, Wei X, Li B, et al. Anti-GPC3-CAR T Cells Suppress the Growth of Tumor Cells in Patient-Derived Xenografts of Hepatocellular Carcinoma. *Front Immunol* (2016) 7:690. doi: 10.3389/fimmu.2016.00690
74. Valsechi MC, Oliveira AB, Conceicao AL, Stuqui B, Candido NM, Provazzi PJ, et al. GPC3 Reduces Cell Proliferation in Renal Carcinoma Cell Lines. *BMC Cancer* (2014) 14:631. doi: 10.1186/1471-2407-14-631
75. Liu CC, Chang TC, Lin YT, Yu YL, Ko BS, Sung LY, et al. Paracrine Regulation of Matrix Metalloproteinases Contributes to Cancer Cell Invasion by Hepatocellular Carcinoma-Secreted 14-3-3sigma. *Oncotarget* (2016) 7(24):36988–99. doi: 10.18632/oncotarget.9234
76. Hu Y, Ma Y, Liu J, Cai Y, Zhang M, Fang X. LINC01128 Expedites Cervical Cancer Progression by Regulating MiR-383-5p/SFN Axis. *BMC Cancer* (2019) 19(1):1157. doi: 10.1186/s12885-019-6326-5
77. Li K, Zheng X, Tang H, Zang YS, Zeng C, Liu X, et al. E3 Ligase MKRN3 Is a Tumor Suppressor Regulating PABPC1 Ubiquitination in Non-Small Cell Lung Cancer. *J Exp Med* (2021) 218(8):e20210151. doi: 10.1084/jem.20210151
78. Shi M, Dai WQ, Jia RR, Zhang QH, Wei J, Wang YG, et al. APC(CDC20)-Mediated Degradation of PHD3 Stabilizes HIF-1a and Promotes Tumorigenesis in Hepatocellular Carcinoma. *Cancer Lett* (2021) 496:144–55. doi: 10.1016/j.canlet.2020.10.011
79. Wu F, Dai X, Gan W, Wan L, Li M, Mitsiades N, et al. Prostate Cancer-Associated Mutation in SPOP Impairs Its Ability to Target Cdc20 for Poly-Ubiquitination and Degradation. *Cancer Lett* (2017) 385:207–14. doi: 10.1016/j.canlet.2016.10.021
80. Hao PP, Li H, Lee MJ, Wang YP, Kim JH, Yu GR, et al. Disruption of a Regulatory Loop Between DUSP1 and P53 Contributes to Hepatocellular

Carcinoma Development and Progression. *J Hepatol* (2015) 62(6):1278–86. doi: 10.1016/j.jhep.2014.12.033
81. Moncho-Amor V, Ibanez de Caceres I, Bandres E, Martinez-Poveda B, Orgaz JL, Sanchez-Perez I, et al. DUSP1/MKP1 Promotes Angiogenesis, Invasion and Metastasis in Non-Small-Cell Lung Cancer. *Oncogene* (2011) 30(6):668–78. doi: 10.1038/onc.2010.449
82. Li H, Wang G, Yu Y, Jian W, Zhang D, Wang Y, et al. Alpha-1,2-Mannosidase MAN1C1 Inhibits Proliferation and Invasion of Clear Cell Renal Cell Carcinoma. *J Cancer* (2018) 9(24):4618–26. doi: 10.7150/jca.27673
83. Chu YD, Lai HY, Pai LM, Huang YH, Lin YH, Liang KH, et al. The Methionine Salvage Pathway-Involving ADI1 Inhibits Hepatoma Growth by Epigenetically Altering Genes Expression *via* Elevating s-Adenosylmethionine. *Cell Death Dis* (2019) 10(3):240. doi: 10.1038/s41419-019-1486-4
84. Oram SW, Ai J, Pagani GM, Hitchens MR, Stern JA, Eggener S, et al. Expression and Function of the Human Androgen-Responsive Gene ADI1 in Prostate Cancer. *Neoplasia* (2007) 9(8):643–51. doi: 10.1593/neo.07415
85. Ojo D, Rodriguez D, Wei F, Bane A, Tang D. Downregulation of CYB5D2 Is Associated With Breast Cancer Progression. *Sci Rep* (2019) 9(1):6624. doi: 10.1038/s41598-019-43006-y
86. Xie Y, Shen YT, Kapoor A, Ojo D, Wei F, De Melo J, et al. CYB5D2 Displays Tumor Suppression Activities Towards Cervical Cancer. *Biochim Biophys Acta* (2016) 1862(4):556–65. doi: 10.1016/j.bbadis.2015.12.013
87. Dai Y, Liu L, Zeng T, Liang JZ, Song Y, Chen K, et al. Overexpression of MUC13, a Poor Prognostic Predictor, Promotes Cell Growth by Activating Wnt Signaling in Hepatocellular Carcinoma. *Am J Pathol* (2018) 188:378–91. doi: 10.1016/j.ajpath.2017.10.016
88. Sheng YH, He Y, Hasnain SZ, Wang R, Tong H, Clarke DT, et al. MUC13 Protects Colorectal Cancer Cells From Death by Activating the NF-Kappab Pathway and Is a Potential Therapeutic Target. *Oncogene* (2017) 36(5):700–13. doi: 10.1038/onc.2016.241
89. Ying HY, Gong CJ, Feng Y, Jing DD, Lu LG. Serine Protease Inhibitor Kazal Type 1 (SPINK1) Downregulates E-Cadherin and Induces EMT of Hepatoma Cells to Promote Hepatocellular Carcinoma Metastasis *via* the MEK/ERK Signaling Pathway. *J Dig Dis* (2017) 18(6):349–58. doi: 10.1111/1751-2980.12486
90. Chen F, Long Q, Fu D, Zhu D, Ji Y, Han L, et al. Targeting SPINK1 in the Damaged Tumour Microenvironment Alleviates Therapeutic Resistance. *Nat Commun* (2018) 9(1):4315. doi: 10.1038/s41467-018-06860-4
91. Li ZB, Chu HT, Jia M, Li L. Long Noncoding RNA LINC01139 Promotes the Progression of Hepatocellular Carcinoma by Upregulating MYBL2 *via* Competitively Binding to MiR-30 Family. *Biochem Biophys Res Commun* (2020) 525(3):581–8. doi: 10.1016/j.bbrc.2020.02.116
92. Bayley R, Ward C, Garcia P. MYBL2 Amplification in Breast Cancer: Molecular Mechanisms and Therapeutic Potential. *Biochim Biophys Acta Rev Cancer* (2020) 1874(2):188407. doi: 10.1016/j.bbcan.2020.188407
93. Gao YJ, Chen F, Zhang LJ. C1q-Like 1 Is Frequently Up-Regulated in Lung Adenocarcinoma and Contributes to the Proliferation and Invasion of Tumor Cells. *J Chemother* (2021) 1–10. doi: 10.1080/1120009X.2021.1906035
94. Ma Q, Liu Y, Shang L, Yu J, Qu Q. The FOXM1/BUB1B Signaling Pathway Is Essential for the Tumorigenicity and Radioresistance of Glioblastoma. *Oncol Rep* (2017) 38(6):3367–75. doi: 10.3892/or.2017.6032
95. Fu X, Chen G, Cai ZD, Wang C, Liu ZZ, Lin ZY, et al. Overexpression of BUB1B Contributes to Progression of Prostate Cancer and Predicts Poor Outcome in Patients With Prostate Cancer. *Onco Targets Ther* (2016) 9:2211–20. doi: 10.2147/OTT.S101994
96. Zagorac S, de Giorgio A, Dabrowska A, Kalisz M, Casas-Vila N, Cathcart P, et al. SCIRT LncRNA Restrains Tumorigenesis by Opposing Transcriptional Programs of Tumor-Initiating Cells. *Cancer Res* (2021) 81(3):580–93. doi: 10.1158/0008-5472.CAN-20-2612
97. Wei S, Dai M, Zhang C, Teng K, Wang F, Li H, et al. KIF2C: A Novel Link Between Wnt/Beta-Catenin and Mtorc1 Signaling in the Pathogenesis of Hepatocellular Carcinoma. *Protein Cell* (2020) 2:788–809. doi: 10.1007/s13238-020-00766-y
98. Gan H, Lin L, Hu N, Yang Y, Gao Y, Pei Y, et al. KIF2C Exerts an Oncogenic Role in Nonsmall Cell Lung Cancer and Is Negatively Regulated by MiR-325-3p. *Cell Biochem Funct* (2019) 37(6):424–31. doi: 10.1002/cbf.3420
99. Zhao B, Ke K, Wang Y, Wang F, Shi Y, Zheng X, et al. HIF-1alpha and HDAC1 Mediated Regulation of FAM99A-Mir92a Signaling Contributes to Hypoxia Induced HCC Metastasis. *Signal Transduct Target Ther* (2020) 5(1):118. doi: 10.1038/s41392-020-00223-6
100. Ren J, Chen GG, Liu Y, Su X, Hu B, Leung BCS, et al. Cytochrome P450 1A2 Metabolizes 17β-Estradiol to Suppress Hepatocellular Carcinoma. *PloS One* (2016) 11(4):e0153863. doi: 10.1371/journal.pone.0153863
101. Lee CC, Ding X, Zhao T, Wu L, Perkins S, Du H, et al. Transthyretin Stimulates Tumor Growth Through Regulation of Tumor, Immune, and Endothelial Cells. *J Immunol* (2019) 202(3):991–1002. doi: 10.4049/jimmunol.1800736
102. Zhang X, Wang F, Huang Y, Ke K, Zhao B, Chen L, et al. FGG Promotes Migration and Invasion in Hepatocellular Carcinoma Cells Through Activating Epithelial to Mesenchymal Transition. *Cancer Manag Res* (2019) 11:1653–65. doi: 10.2147/CMAR.S188248
103. Peng HH, Wang JN, Xiao LF, Yan M, Chen SP, Wang L, et al. Elevated Serum FGG Levels Prognosticate and Promote the Disease Progression in Prostate Cancer. *Front Genet* (2021) 12:651647. doi: 10.3389/fgene.2021.651647
104. Chen Y, Zhu Y, Sheng Y, Xiao J, Xiao Y, Cheng N, et al. SIRT1 Downregulated FGB Expression to Inhibit RCC Tumorigenesis by Destabilizing STAT3. *Exp Cell Res* (2019) 382(2):111466. doi: 10.1016/j.yexcr.2019.06.011
105. Shek FH, Luo R, Lam BYH, Sung WK, Lam TW, Luk JM, et al. Serine Peptidase Inhibitor Kazal Type 1 (SPINK1) as Novel Downstream Effector of the Cadherin-17/Beta-Catenin Axis in Hepatocellular Carcinoma. *Cell Oncol* (2017) 40(5):443–56. doi: 10.1007/s13402-017-0332-x
106. Qin X, Zhang J, Lin Y, Sun XM, Zhang JN, Cheng ZQ. Identification of MiR-211-5p as a Tumor Suppressor by Targeting ACSL4 in Hepatocellular Carcinoma. *J Transl Med* (2020) 18(1):326. doi: 10.1186/s12967-020-02494-7
107. Wu X, Deng F, Li Y, Daniels G, Du X, Ren Q, et al. ACSL4 Promotes Prostate Cancer Growth, Invasion and Hormonal Resistance. *Oncotarget* (2015) 6(42):44849–63. doi: 10.18632/oncotarget.6438
108. Han K, Li C, Zhang X, Shang L. DUXAP10 Inhibition Attenuates the Proliferation and Metastasis of Hepatocellular Carcinoma Cells by Regulation of the Wnt/Beta-Catenin and PI3K/Akt Signaling Pathways. *Biosci Rep* (2019) 39(5):BSR20181457. doi: 10.1042/BSR20181457
109. Xu Y, Yu X, Wei C, Nie F, Huang M, Sun M. Over-Expression of Oncigenic Pesudogene DUXAP10 Promotes Cell Proliferation and Invasion by Regulating LATS1 and Beta-Catenin in Gastric Cancer. *J Exp Clin Cancer Res* (2018) 37(1):13. doi: 10.1186/s13046-018-0684-8
110. Liang M, Yao W, Shi B, Zhu X, Cai R, Yu Z, et al. Circular RNA Hsa_Circ_0110389 Promotes Gastric Cancer Progression Through Upregulating SORT1 *via* Sponging MiR-127-5p and MiR-136-5p. *Cell Death Dis* (2021) 12(7):639. doi: 10.1038/s41419-021-03903-5
111. Wang P, Yang X, Zhao L, Liu D, Liu J, Ding Y. A Novel Long Non-Coding RNA TONSL-AS1 Regulates Progression of Gastric Cancer *via* Activating TONSL. *Exp Cell Res* (2019) 382(1):111453. doi: 10.1016/j.yexcr.2019.05.034
112. Liu Y, Li L, Wang X, Wang P, Wang Z. LncRNA TONSL-AS1 Regulates MiR-490-3p/CDK1 to Affect Ovarian Epithelial Carcinoma Cell Proliferation. *J Ovarian Res* (2020) 13(1):60. doi: 10.1186/s13048-020-00657-0
113. Udali S, Guarini P, Ruzzenente A, Ferrarini A, Guglielmi A, Lotto V, et al. DNA Methylation and Gene Expression Profiles Show Novel Regulatory Pathways in Hepatocellular Carcinoma. *Clin Epigenet* (2015) 7:43. doi: 10.1186/s13148-015-0077-1
114. Liu M, Zhang Y, Liao Y, Chen Y, Pan Y, Tian H, et al. Evaluation of the Antitumor Efficacy of Rnai-Mediated Inhibition of CDC20 and Heparanase in an Orthotopic Liver Tumor Model. *Cancer Biother Radiopharm* (2015) 30(6):233–9. doi: 10.1089/cbr.2014.1799
115. Lian Q, Wang S, Zhang G, Wang D, Luo G, Tang J, et al. HCCDB: A Database of Hepatocellular Carcinoma Expression Atlas. *Genomics Proteomics Bioinf* (2018) 16(4):269–75. doi: 10.1016/j.gpb.2018.07.003
116. Tsuneki M, Nakamura Y, Kinjo T, Nakanishi R, Arakawa H. Mieap Suppresses Murine Intestinal Tumor *via* Its Mitochondrial Quality Control. *Sci Rep* (2015) 5:12472. doi: 10.1038/srep12472
117. Chapuy B, McKeown MR, Lin CY, Monti S, Roemer MG, Qi J, et al. Discovery and Characterization of Super-Enhancer-Associated

Dependencies in Diffuse Large B Cell Lymphoma. *Cancer Cell* (2013) 24 (6):777–90. doi: 10.1016/j.ccr.2013.11.003
118. Mack SC, Pajtler KW, Chavez L, Okonechnikov K, Bertrand KC, Wang X, et al. Therapeutic Targeting of Ependymoma as Informed by Oncogenic Enhancer Profiling. *Nature* (2018) 553(7686):101–5. doi: 10.1038/nature25169
119. Wu Y, Liu H, Ding H. GPC-3 in Hepatocellular Carcinoma: Current Perspectives. *J Hepatocell Carcinoma* (2016) 3:63–7. doi: 10.2147/JHC.S116513
120. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene Body Methylation can Alter Gene Expression and Is a Therapeutic Target in Cancer. *Cancer Cell* (2014) 26(4):577–90. doi: 10.1016/j.ccr.2014.07.028
121. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, et al. Impact of Cytosine Methylation on DNA Binding Specificities of Human Transcription Factors. *Science* (2017) 356(6337):eaaj2239. doi: 10.1126/science.aaj2239
122. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PW, Bauer C, et al. Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. *Cell* (2013) 152(5):1146–59. doi: 10.1016/j.cell.2013.02.004
123. Rauluseviciute I, Drablos F, Rye MB. DNA Hypermethylation Associated With Upregulated Gene Expression in Prostate Cancer Demonstrates the Diversity of Epigenetic Regulation. *BMC Med Genomics* (2020) 13(1):6. doi: 10.1186/s12920-020-0657-6
124. Xie Z, Dang Y, Wu H, He R, Ma J, Peng Z, et al. Effect of CELSR3 on the Cell Cycle and Apoptosis of Hepatocellular Carcinoma Cells. *J Cancer* (2020) 11 (10):2830–44. doi: 10.7150/jca.39328
125. Liu MX, Jin L, Sun SJ, Liu P, Feng X, Cheng ZL, et al. Metabolic Reprogramming by PCK1 Promotes TCA Cataplerosis, Oxidative Stress and Apoptosis in Liver Cancer Cells and Suppresses Hepatocellular Carcinoma. *Oncogene* (2018) 37(12):1637–53. doi: 10.1038/s41388-017-0070-6
126. Louie E, Chen XF, Coomes A, Ji K, Tsirka S, Chen EI. Neurotrophin-3 Modulates Breast Cancer Cells and the Microenvironment to Promote the Growth of Breast Cancer Brain Metastasis. *Oncogene* (2013) 32(35):4064–77. doi: 10.1038/onc.2012.417
127. Li Y, Zhang PY, Yang ZW, Ma F, Li FX. TIMD4 Exhibits Regulatory Capability on the Proliferation and Apoptosis of Diffuse Large B-Cell Lymphoma Cells *via* the Wnt/Beta-Catenin Pathway. *J Gene Med* (2020) 22(8):e3186. doi: 10.1002/jgm.3186

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.