Integrating the Epigenome and Transcriptome of Hepatocellular Carcinoma to Identify Systematic Enhancer Aberrations and Establish an Aberrant Enhancer-Related Prognostic Signature

Peng Huang[1], Bin Zhang[1], Junsheng Zhao[1] and Ming D. Li[1,2]*

[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

Recently, emerging evidence has indicated that aberrant enhancers, especially super-enhancers, play pivotal roles in the transcriptional reprogramming of multiple cancers, including hepatocellular carcinoma (HCC). In this study, we performed integrative analyses of ChIP-seq, RNA-seq, and whole-genome bisulfite sequencing (WGBS) data to identify intergenic differentially expressed enhancers (DEEs) and genic differentially methylated enhancers (DMEs), along with their associated differentially expressed genes (DEE/DME-DEGs), both of which were also identified in independent cohorts and further confirmed by HiC data. Functional enrichment and prognostic model construction were conducted to explore the functions and clinical significance of the identified enhancer aberrations. We identified a total of 2,051 aberrant enhancer-associated DEGs (AE-DEGs), which were highly concurrent in multiple HCC datasets. The enrichment results indicated the significant overrepresentations of crucial biological processes and pathways implicated in cancer among these AE-DEGs. A six AE-DEG-based prognostic signature, whose ability to predict the overall survival of HCC was superior to that of both clinical phenotypes and previously published similar prognostic signatures, was established and validated in TCGA-LIHC and ICGC-LIRI cohorts, respectively. In summary, our integrative analysis depicted a landscape of aberrant enhancers and associated transcriptional dysregulation in HCC and established an aberrant enhancer-derived prognostic signature with excellent predictive accuracy, which might be beneficial for the future development of epigenetic therapy for HCC.

Keywords: enhancer, super-enhancer, DNA methylation, histone modification, prognostic model, eRNA, RNA-seq, hepatocellular carcinoma

INTRODUCTION

Liver cancer is the sixth most common malignant tumor and the third leading cause of cancer-related deaths, accounting for approximately 700,000 deaths annually worldwide and poses a severe health threat and economic burden to the world (Likhitsup and Parikh, 2020; Sung et al., 2021). This is especially true in China, which has the largest HCC risk population (HBV carriers) throughout the world. The latest epidemiological report showed that primary liver cancer is the fourth most common tumor in China (Feng et al., 2019) and the vast majority of liver cancers are HCCs. Since there are usually no evident symptoms in the early developmental state of liver cancer, patients are often diagnosed in the late stage of liver cancer, resulting in an extremely high probability of death (Grandhi et al., 2016). Although the survival duration of early- and intermediate-stage HCCs has improved over the past decades, the prognosis for advanced-stage HCC patients has remained poor, with no significant improvement. Even with the survival benefits of several first- and second-line therapeutic options available for patients with advanced HCC, such as sorafenib and lenvatinib, the median survival time of intermediate to advanced HCC is only 1–2 years (Marrero et al., 2018). Clinical studies of immune checkpoint inhibitors have yielded promising survival benefits, although the suppressive milieu and tumor immunosurveillance escape mechanisms in the liver still dampen the effectiveness of immunotherapy (Nakano et al., 2020). Hence, there is an urgent need to explore the underlying genetic and epigenetic mechanisms implicated in hepatocarcinogenesis to identify potential targets/biomarkers for the diagnosis, treatment and prognosis of HCC.

Cancer is a complex disease involving both genetic mutations and epigenetic aberrations. By definition, epigenetics refers to heritable states of gene activities that do not involve alteration of DNA sequence itself. Epigenetic changes such as DNA hypermethylation or hypomethylation, dysregulation of histone modification patterns, chromatin remodeling, and aberrant expression of noncoding RNAs are demonstrated to be involved in the initiation and progression of HCC (Wahid et al., 2017). Unlike genetic mutations, epigenetic alterations are reversible and various drugs targeting epigenetic regulators have exhibited viable therapeutic potential for solid tumors in both preclinical and clinical studies (Cheng et al., 2019). A better understanding of the epigenetic mechanisms underlying hepatocarcinogenesis will facilitate the discovery of new targets and biomarkers for HCC therapy.

Like most malignancies, HCC is also characterized by widespread abnormal gene expression. Enhancers are distal, noncoding genomic regulatory elements with multiple transcription factor binding sites that interact with promoters to enhance the transcription of target genes. Nucleosomes in the neighborhood of active enhancers usually contain histones with iconic posttranslational modifications, such as H3 lysine monomethylation (H3K4me1) and H3 lysine acetylation (H3K27ac) at their amino termini (Shlyueva et al., 2014). Super-enhancers are large clusters of enhancers that synergistically promote gene transcription (Herranz et al.,

2014). Emerging evidence shows that cancer cells can acquire super-enhancers in the vicinity of key oncogenes, such as *MYC* and *TAL1*, during the development of cancer (Hnisz et al., 2013; Herranz et al., 2014; Mansour et al., 2014). Moreover, pancancer studies of TCGA data also showed wide-spread aberrant super-enhancer activities in cancers (Chen et al., 2018a; Chen and Liang, 2020).

In HCC, Wong et al. demonstrated that the super-enhancer landscape and components of the *trans*-acting super-enhancer complex, composed of *CDK7*, *BRD4*, *EP300*, and *MED1*, were significantly altered (Tsang et al., 2019). Additionally, Deng et al. reported an aberrant landscape of active enhancers developed in cirrhosis and conserved in hepatocarcinogenesis (Yang et al., 2020). However, those two studies lacked a comprehensive collection of enhancers in the liver, reliable identification of enhancer target genes, and replication of enhancer aberrations in independent cohorts.

In the present study, through the integration of transcriptome and epigenome data, we aimed to: 1) manually curate a comprehensive catalog of enhancers in the liver; 2) systematically identify and replicate enhancer aberrations and associated target genes in HCC; and 3) explore the function and prognostic significance of identified aberrant enhancers.

MATERIALS AND METHODS

Patient Data and Tissues Collection

Paired tumor tissues and adjacent non-tumor tissues used in this study were collected from 33 HCC patients who underwent hepatectomy at the First Affiliated Hospital, Zhejiang University School of Medicine. Board-certified pathologists reviewed each specimen to confirm that all frozen sections were histologically consistent with tumor or non-tumor tissues. This study was approved by the Institutional Review Board of The First Affiliated Hospital. Written informed consent was obtained from each participant.

High-Throughput Sequencing and Computational Preprocessing

DNA methylation and gene expression of 33 pairs of tumour and adjacent tissues were assessed by whole-genome bisulfite sequencing (WGBS) and mRNA-seq on the Illumina X Ten platform with standard procedures. After quality control, clean WGBS reads were aligned with the reference genome (hg38) using Bismark (v. 0.16.1) (Krueger and Andrews, 2011) with default parameters. The harvested count data for each strand were combined for methylation level estimation. Differentially methylated loci (DML) and differentially methylated regions (DMRs) were detected with customized R scripts like our previous WGBS study (Huang et al., 2021). For RNAseq data, clean reads that passed quality control were aligned with the hg38 genome, and the reference transcriptome was downloaded from GENCODE (v. 29) (Harrow et al., 2012) with STAR (v. 2.5.2a) (Dobin et al., 2013). Estimated raw count gene expression from STAR was imported into DESeq2 (Love et al., 2014) for

differential expression analysis. STAR generated alignment BAM files were utilized as input for enhancer RNA (eRNA) expression quantification via bedtools (v. 2.27.1) (Quinlan and Hall, 2010). More details about high-throughput sequencing and bioinformatic preprocessing can be found in the Supplementary Methods.

Curation of a Comprehensive Catalog of Enhancers in Liver

Eleven histone ChIP-seq liver relevant samples were collected from the public domain. Specifically, bed files containing the pseudo-replicated peaks identified from six H3K4me1- and H3K27ac-based ChIPseq profiled samples (i.e., one HepG2, one Hepatocyte, and four normal adult liver tissue samples) were downloaded from ENCODE (Consortium, 2004). For each ENCODE sample, regions with overlapped H3K27ac and H3K4me1 peaks were annotated as active enhancers, and regions with only H3K4me1 peaks were considered as primed enhancers. In one case of an adult liver ENCODE sample without H3K27ac profiling data, H3K4me1 peaks were included as enhancers (primed or active). Four types of histone (including H3K4me1 and H3K27ac) ChIP-seq profiling-based ChromHMM state annotation files of five adult liver tissue samples (i.e., one normal liver sample and two tumor and matched adjacent cirrhosis samples from two HCC patients) were retrieved from the recent integrative epigenomic study on HCC (Hlady et al., 2019). Specifically, regions whose ChromHMM states were annotated as "poised enhancer" (refers to regions with only H3K4me1 peaks) were included as primed enhancers, and regions annotated as "active enhancer" (refer to regions with both H3K4me1 and H3K27ac) were collected as active enhancers. All (active or primed) enhancers from each sample were merged together via bedtools (Quinlan and Hall, 2010). Afterwards, enhancers with a lenth of <50 bp or overlapped with any promoter (upstream 1,500 bp to downstream 500 bp from TSS) were excluded from further analysis. The concurrence of each merged enhancer was estimated as the number of ChIP-seq samples in which the merged enhancer was annotated as a primed or active enhancer. In other words, a merged enhancer with higher concurrence represents a more highly conserved and reliable enhancer among those 11 liver-related ChIP-seq samples.

Identification of Intergenic Differentially Expressed Enhancers and Associated Differentially Expressed Genes

The collected enhancers in the liver were divided into two groups, namely, intergenic enhancers and genic enhancers, according to their genomic locations. For intergenic enhancers, the read count-based expression levels of eRNAs were estimated via the "coverage" module of bedtools (Quinlan and Hall, 2010). A paired t-test was applied to the normalized expression ($\log_2$ transformed fragment per million, $\log_2$ FPM) of each eRNA to identify significant differentially expressed eRNA ($|\log_2$ fold change of FPM$| > 0.5$ and BH-FDR < 0.05). Intergenic enhancers with significant differential expression of

eRNA were defined as intergenic differentially expressed enhancers (intergenic DEEs). Nearby (TSS located $\pm$ 1 Mb from the center of corresponding intergenic DEEs) differentially expressed genes (DEG) ($|\log_2$ fold change (LFC)$| > 0.5$ and BH-FDR < 0.05) displayed a significant correlation (Spearman Rho ≥ 0.7 and Bonferroni-corrected p-value < 0.01) with eRNA expression were identified as intergenic DEE-associated DEGs (intergenic DEE-DEGs).

Replication of Intergenic DEE-DEGs

For independent replication of intergenic DEE-DEGs, four HCC RNA-seq datasets were downloaded from the GEO: GSE77314 (paired tumor and adjacent nontumor tissue samples from 50 HCC patients) (Liu et al., 2016), GSE124535 (paired tumor and adjacent nontumor tissue samples from 35 HCC patients) (Jiang et al., 2019), GSE148355 (62 tumor and 47 adjacent nontumor samples) (Yoon et al., 2021), and GSE77509 (paired tumor and adjacent nontumor samples from 20 HCC patients) (Yang et al., 2017). The same protocols in the discovery cohort were applied to detect intergenic DEEs and associated DEE-DEGs in these four datasets. Afterward, identified intergenic DEEs and DEE-DEGs from each dataset were compared with those from the discovery cohort to calculate the concurrence of each intergenic DEE and DEE-DEG. Specifically, the concurrence of each DEE was calculated as one plus the number of GEO datasets in which the DEE was successfully replicated, while the concurrence of each DEE-DEG was calculated as one plus the number of GEO datasets in which the corresponding DEE and DEG were significant and the correlation between them was also significant.

Assessment of the Roles of Epigenetic Modification Aberrations in Intergenic DEEs

Intergenic DEEs that overlapped with at least one DMR and displayed significant methylation-eRNA Spearman correlation (BH-FDR < 0.05) were defined as methylation-associated DEEs, and corresponding DEE-DEGs were classified as methylation-associated DEE-DEGs. Meanwhile, we further investigated the dysregulation of histone posttranslational modification (PTM) modifiers and their potential implications in those identified intergenic enhancer aberrations. Top differentially expressed histone PTM modifiers ($|LFC| > 1$ and BH-FDR $< 5\%$) in the discovery cohort were screened out for subsequent coexpression analyses to determine the ratios of DEEs and DEE-DEGs that significantly correlated ($|Spearman$ correlation coefficient$| > 0.5$ and BH-FDR $< 5\%$) with the mRNA expression of those histone PTM modifiers.

Identification of Genic Differentially Methylated Enhancers and Associated Differentially Expressed Genes

Reliable genic enhancers (concurrence among the 11 ChIP-seq samples ≥ 2) that overlapped (length$_{overlap} \geq 200$ bp, length$_{overlap}$/length$_{enhancer} \geq 0.3$, and with at least 5 CpGs) with at least one

DMR were identified as potential genic differentially methylated enhancers (genic DMEs). For each potential genic DME, their associated DEGs were screened *via* the Spearman correlation test. Nearby (distance of enhancer to TSS $\leq \pm 1$ Mb) DEGs ($|LFC| > 0.5$ and BH-FDR < 0.05) that show significant correlation ($|Rho| \geq 0.5$ and FDR ≤ 0.01) between gene expression and DNA methylation level were identified as genic DEE-associated DEGs (genic DME-DEGs). Genic DME candidates with at least one associated DEG were identified as genic DMEs.

Replication of Genic DMEs and DME-DEGs

The normalized gene expression and DNA methylation level matrix of TCGA-LIHC were retrieved *via* the RTCGA R package (Kosinski and Biecek, 2015). For each genic DEE-DEG pair identified in the discovery cohort, we examined the significance of differential methylation, differential expression, and Spearman correlation between DNA methylation and gene expression in TCGA-LIHC. A genic DEE-DEG pair was considered as "successful replication" only when there was simultaneous significant differential methylation, differential expression, and a significant correlation between methylation and expression in TCGA-LIHC. Considering the platform limitation of the 450 k methylation array in covering enhancer CpG, we classified all replication failures of DEE-DEGs into two groups: 1) "type I failure" refers to replication failure due to the lack of CpG for corresponding genic DMEs in TCGA-LIHC, and 2) "type II failure" refers to replication failure except type I failure. The raw replication rate of genic DEE-DEGs was calculated as the ratio of genic DEE-DEGs that achieved successful replication, while the platform-adjusted replication rate was defined as: $\text{Count}_{\text{Successful replication}}/(\text{Count}_{\text{Successful replication}} + \text{Count}_{\text{Type II failure}})*100$.

Functional Enrichment of Aberrant Enhancer-Associated Differentially Expressed Genes

AE-DEGs were defined as the union of those identified intergenic DEE-DEGs and genic DME-DEGs. Pathway/gene ontology (GO) enrichment analyses of upregulated and downregulated AE-DEGs were performed *via* the online web tool Metascape (Zhou et al., 2019). In addition, 10 cancer hallmark gene sets were downloaded from the Cancer Hallmark Gene (CHG) database (Zhang et al., 2020a). The enrichment degrees of AE-DEGs for cancer hallmarks were evaluated through a hypergeometric test followed by BH-FDR multitest correction in R.

Bioinformatic Confirmation of AE-DEGs Using Public Hi-C Data

The bed files containing topologically associated domains (TADs) and chromatin loops of Hi-C-profiled HepG2 and one normal adult liver tissue sample were downloaded from the 3D Genome Browser (http://3dgenome.fsm. northwestern.edu/) (Wang et al., 2018). Each pair of AE

and AE-DEG was examined to determine whether both the enhancer and its associated DEG were located in the same TAD or located in the two elements of a chromatin loop, respectively.

Establishment of an AE-Derived Prognostic Model

Clinical phenotype data, including overall survival (OS) time and status, were retrieved from the integrated TCGA pancancer clinical data resource (Liu et al., 2018a). Univariate Cox proportional hazards regression analysis was conducted to screen for AE-DEGs associated with the OS of HCC patients in TCGA-LIHC via the function "coxph" in the R package "survival" (Therneau, 2020). AE-DEGs with univariate Cox *p*-value < 0.05 were incorprated into the least absolute shrinkage and selection operator (LASSO) regression model by using the glmnet package (Friedman et al., 2010) for identification of the most prominent survival-associated AE-DEGs in TCGA-LIHC. Afterward, the multivariate proportional hazards Cox regression model was employed to establish a gene signature for predicting the OS of HCC patients. Multivariate Cox regression-derived coefficients (β) were used to calculate the risk score as follows: risk score = (β_{gene1} * normalized expression level of gene1 + β_{gene2} * normalized expression level of gene2 + . . . + β_{geneN} * normalized expression of geneN) (Lossos et al., 2004). Based on the optimal cutoff of risk score determined by minimizing log-rank test *p*-value, HCC patients were divided into high- and low-risk groups, whose differences in OS probability across time were visualized through a Kaplan-Meier survival curve by using the function "ggsurvplot" in the survminer package (Alboukadel Kassambara, 2021). The prognostic performance of the risk score was evaluated by time-dependent receiver operating characteristic (ROC) curve analysis via the function "survivalROC" in the survivalROC package (Patrick, 2013). The independent prognostic role of the identified gene signature in TCGA-LIHC was assessed by building a multivariate Cox regression model including the risk group, age, gender, and pathologic tumor-node-metastasis (TNM) stage of each patient. All factors that passed through the multivariate Cox regression model were utilized for the construction of a predictive nomogram *via* the rms package (RMS, 2021). Calibration plots and time-dependent ROC curves were applied to assess the predictive performance of the established nomogram.

Validtion of the AE-Derived Prognostic Model

Regarding the independent validation of the prognostic signature, clinical phenotypes and gene expression data of the International Cancer Genome Consortium Liver Cancer-RIKEN (LIRI-JP) were downloaded from the ICGC website. Multivariate Cox regression-derived coefficients from TCGA-LIHC were used to calculate the corresponding risk score for each patient in ICGC-LIRI. Similarly, ICGC-LIRI patients
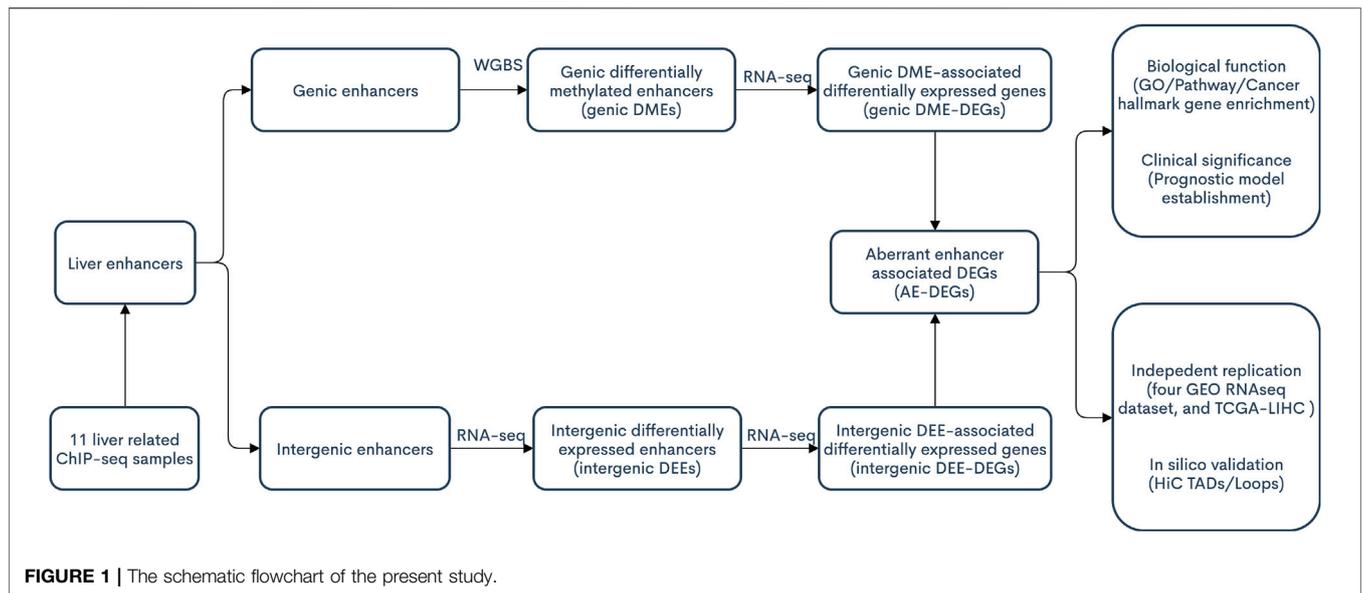
FIGURE 1 | The schematic flowchart of the present study.
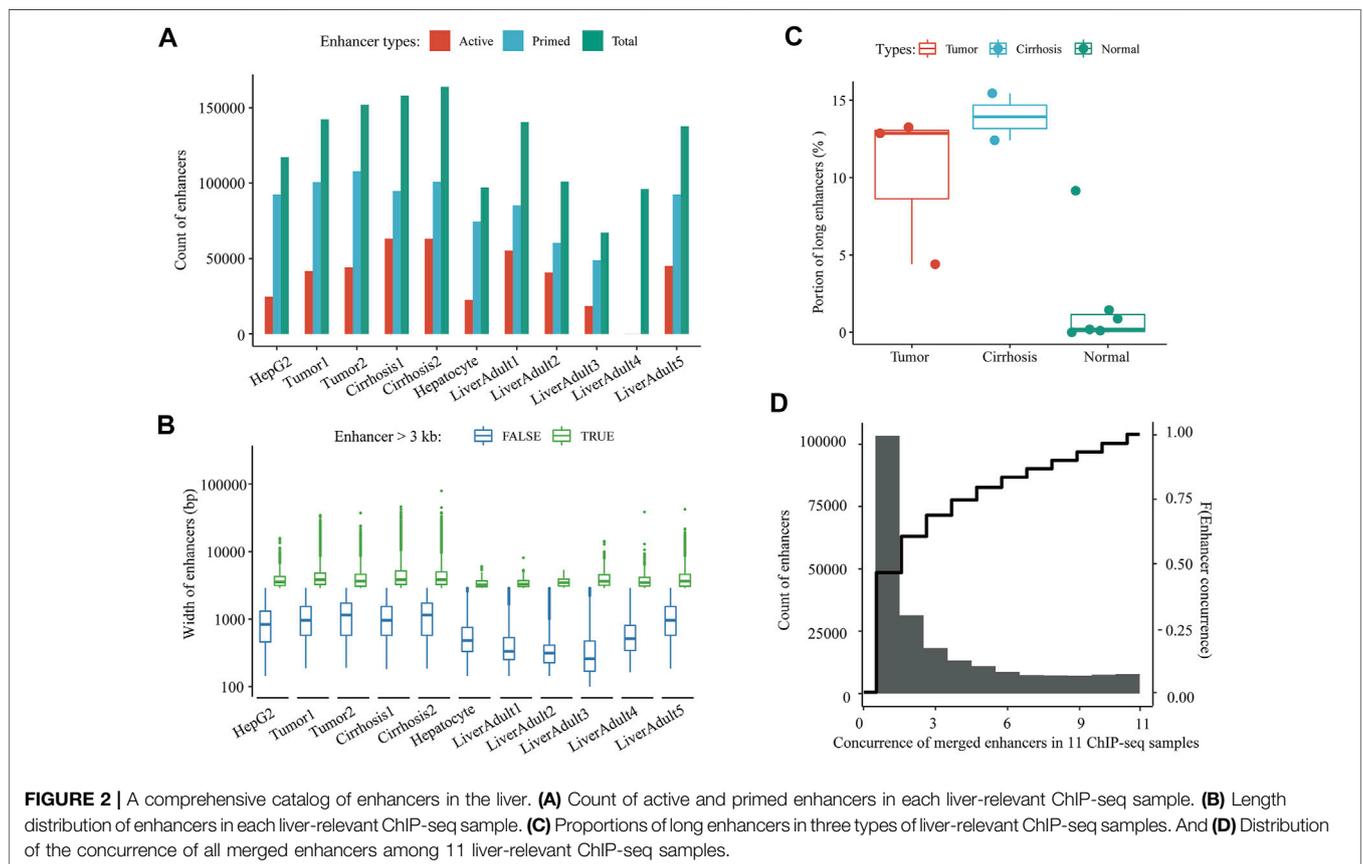
FIGURE 2 | A comprehensive catalog of enhancers in the liver. **(A)** Count of active and primed enhancers in each liver-relevant ChIP-seq sample. **(B)** Length distribution of enhancers in each liver-relevant ChIP-seq sample. **(C)** Proportions of long enhancers in three types of liver-relevant ChIP-seq samples. And **(D)** Distribution of the concurrence of all merged enhancers among 11 liver-relevant ChIP-seq samples.

were divided into high- and low-risk groups according to the cutoff determined by minimizing the log-rank test p-value. Comparison of the difference in OS probability, evaluation of

predictive performance, and assessment of predictive independence were performed with identical procedures employed for TCGA-LIHC.

TABLE 1 | Characteristics and enhancer identification strategies applied for 11 liver-relevant ChIP-seq samples.

Sample name	Source	H3K4me1	H3K27ac	Enhancer identification strategy
HepG2	ENCODE (Consortium, 2004)	√	√	H3K4me1 only (primed enhancer) + H3K4me1 and H3K27ac (active enhancer)
Hepatocyte		√	√	
LiverAdult 1		√	√	
LiverAdult 2		√	√	
LiverAdult 3		√	√	
LiverAdult 4		√	X	H3K4me1 (primed or active enhancer)
LiverAdult 5	The integrative epigenomic HCC study (Hlady et al., 2019)	√	√	H3K4me1 only (primed enhancer) + H3K4me1 and H3K27ac (active enhancer)
Cirrhosis 1		√	√	
Cirrhosis 2		√	√	
Tumor 1		√	√	
Tumor 2		√	√	
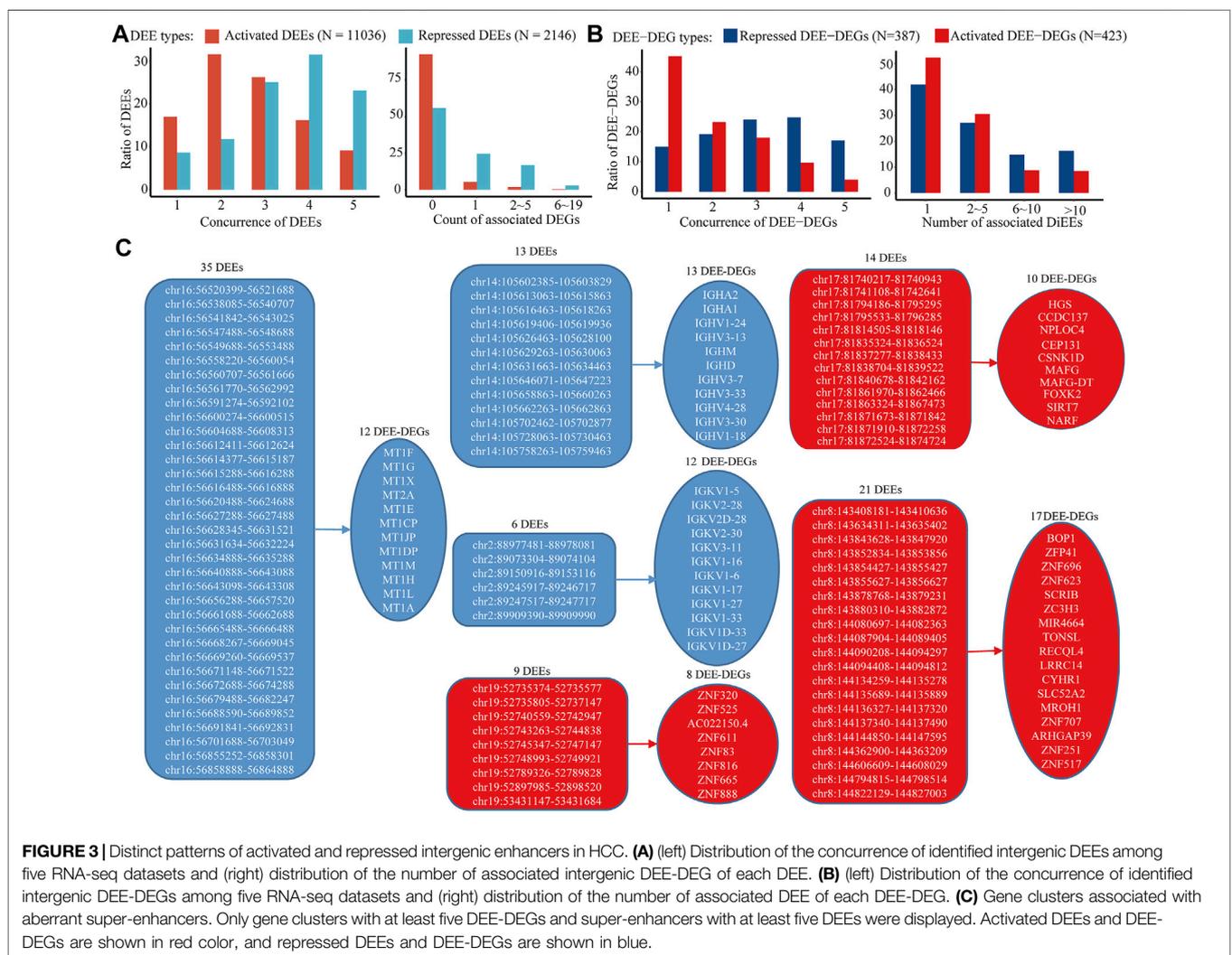
FIGURE 3 | Distinct patterns of activated and repressed intergenic enhancers in HCC. **(A)** (left) Distribution of the concurrence of identified intergenic DEEs among five RNA-seq datasets and (right) distribution of the number of associated intergenic DEE-DEG of each DEE. **(B)** (left) Distribution of the concurrence of identified intergenic DEE-DEGs among five RNA-seq datasets and (right) distribution of the number of associated DEE of each DEE-DEG. **(C)** Gene clusters associated with aberrant super-enhancers. Only gene clusters with at least five DEE-DEGs and super-enhancers with at least five DEEs were displayed. Activated DEEs and DEE-DEGs are shown in red color, and repressed DEEs and DEE-DEGs are shown in blue.

RESULTS

Comprehensive Collection of Enhancers in the Liver

The procedures of this study are shown in the schematic flowchart (**Figure 1**). Among 11 liver-relevant ChIP-seq profiled samples, we identified numerous enhancers whose counts ranged from 96,124 to 163,953 (**Figure 2A**; Table S1.1). On average, there were 124,838 enhancers in each sample, among which approximately one third (32.08%) were active enhancers with both H3K4me1 and H3K27ac peak signals (**Table 1**; **Supplementary Table S1.1**). Interestingly, there was a higher proportion of long enhancers (length > 3 kb) in nonnormal samples (i.e., tumoral and cirrhosis samples), and the median widths of classical enhancers (length ≤ 3 kb) were also higher than those of normal samples (**Figures 2B,C**; **Supplementary Table S1.2**). Specifically, the mean percentages of long enhancers (>3 kb) in tumor samples, adjacent cirrhosis samples, and normal liver samples were 10.17, 12.93, and 1.97%, respectively (**Figure 2C**). After combining these enormous enhancers, we obtained a comprehensive catalog of 223,007 unique enhancers in the liver. Over one half (53.64%) of them were concurrent enhancers that consistently existed in at least two samples (**Figure 2D**; **Supplementary Table S1.3**). In addition, genomic location-based annotation showed that approximately 29.84% (66,551/223,007) of those 223,007 enhancers were located in intergenic regions (**Supplementary Table S1.3**).

Activated and Repressed Intergenic Enhancers Show Different Patterns in Concurrence and Transcriptional Regulation in HCC

In the discovery cohort, 23,601 of the 66,551 collected intergenic enhancers displayed active transcription of eRNA, and 13,182 of them were identified as intergenic DEEs, including 11,036 activated DEEs and 2,146 repressed DEEs (**Figure 3A**). Through bioinformatic inference for target genes, 842 activated DEEs and 951 repressed DEEs were found to be correlated with 423 upregulated DEE-DEGs, and 387 downregulated DEE-DEGs, respectively (**Figure 3B**; **Supplementary Table S2**). Although the number of activated DEEs was over fivefold that of repressed DEEs (11036 *vs.* 2,146), each repressed DEE was found to be simultaneously associated with more DEE-DEGs (**Figure 3A**). Specifically, 19.67% of repressed DEEs displayed high correlations with multiple DEE-DEGs, while only 2.28% of activated DEEs showed this pattern (**Figure 3A**). Moreover, 387 downregulated and 423 upregulated DEE-DEGs also showed differences in terms of the number of associated DEEs. Compared with upregulated DEE-DEG, each downregulated DEE-DEG tended to be simultaneously regulated by more DEEs (**Figure 3B**). Taken together, we found a higher portion of potential transcriptional master regulators among the repressed DEEs, and more downregulated DEE-DEGs were simultaneously associated with aberrant super-enhancers that were consisted

of multiple adjacent synergistic enhancers (**Figure 3C**). For example, in 16q13, a cluster of 35 repressed intergenic DEEs was identified as potential regulators of the metallothionein (MT) family (i.e., each of the 35 DEEs was significantly correlated with the expressions of all 12 metallothionein genes) (**Figure 3C**; **Table 2**). A literature searching revealed that nine of those 12 MT genes were previously implicated in HCC (**Table 2**). Besides, in chromosome 17, we also identified a super-enhancer whose activation was correlated with upregulation of 10 DEGs including nine previously-reported oncogenes in HCC or other cancers (**Table 2**). Beyond these, we also identified another four gene clusters likely regulated by super-enhancers on chromosomes 14, 2, 19, and 8 (**Figure 3C**).

Moreover, the identified intergenic DEEs and DEE-DEGs were overall highly replicated in four independent GEO datasets (**Table 3**). A total of 83.03% of activated DEEs, and 91.33% of repressed DEEs were observed in at least one GEO dataset (i.e., concurrence ≥ 2) (**Figure 3A**). Furthermore, 54.85% of upregulated DEE-DEGs and 85.01% of downregulated DEE-DEGs were identified in at least one GEO dataset (**Figure 3B**). Compared with activated DEEs and upregulated DEE-DEGs, those repressed DEEs and downregulated DEE-DEGs were more likely to be conserved in multiple GEO datasets (i.e., concurrence higher than three) (**Figures 3A,B**). For instance, 54.52% of repressed DEEs and 41.86% of downregulated DEE-DEGs were consistently replicated in more than three GEO datasets, while only 25.34% activated DEEs and 13.71% upregulated DEE-DEGs were also observed in three or more GEO datasets (**Figures 3A,B**).

Potential Roles of Epigenetic Modification Aberrations in Identified Aberrant Intergenic Enhancers

Through integration with matched WGBS data in the discovery cohort, the differential expression of 10.61% of the activated DEEs and 11.14% of the repressed DEEs was significantly correlated with regional differential DNA methylation, especially hypomethylation, in corresponding enhancers (**Figure 4A**). Nevertheless, these differential methylation-associated DEEs correlated with 34.04% of the upregulated DEE-DEGs and 37.21% of the downregulated DEE-DEGs (**Figure 4B**), suggesting that those methylation-associated DEEs were more likely to be transcriptional master regulators that targeted multiple genes. In addition to DNA methylation, there was also substantial dysregulation of histone modification in HCC. Three histone methyltransferases (EZH2, EHMT2, and SMYD3), two demethylases (KDM5B and KDM6B), and two deacetylases (HDAC11 and HDAC9) were differentially expressed in both the discovery and four GEO datasets (**Figure 4C**; **Supplementary Table S2.2**). Coexpression tests showed that many DEEs and DEE-DEGs were significantly correlated with the differential expression of those seven histone modification regulators, especially EZH2, EHMT2, and SMYD3 (**Figures 4D,E**). Notably, 75.89% of the upregulated DEE-DEGs and 63.57% of the downregulated DEE-DEGs displayed significant coexpression with EZH2 and SMYD3,

TABLE 2 | Summary information about two representative sets of DEE-DEGs regulated by intergenic DEE clusters.

Enhancer cluster	Average LFC.DEE	Avergae LogFDR.DEE	Gene name	LFC DEG	FDR DEG	Average rho	Average concurrence	Implicated cancers
chr16:56520399–56864888 (~344 kb, 35 DEEs)	–3.3	6.62	MT1A	**–2.08**	**1.45E-06**	**0.76**	**2.56**	HCC (Ning et al., 2021)
			MT1CP	–2.01	1.20E-03	0.82	3.51	Unknown
			MT1DP	**–1.37**	**1.64E-03**	**0.81**	**3.71**	HCC (Yu et al., 2014a) and others (Gai et al., 2020)
			MT1E	**–3.58**	**2.35E-16**	**0.88**	**3.72**	HCC (Liu et al., 2020) and others (Hur et al., 2016)
			MT1F	**–3.70**	**3.24E-17**	**0.91**	**3.75**	HCC (Lu et al., 2003) and others (Lin et al., 2017a)
			MT1G	**–3.43**	**6.06E-10**	**0.90**	**3.67**	HCC (Wang et al., 2019) and others (Fu et al., 2013)
			MT1H	**–2.13**	**1.48E-03**	**0.86**	**2.43**	HCC (Zheng et al., 2017)
			MT1JP	**–5.27**	**1.36E-20**	**0.82**	**3.64**	HCC (Wu et al., 2020a) and others (Zhang et al., 2018; Yu et al., 2019)
			MT1L	–2.58	5.28E-10	0.86	2.86	Unknown
			MT1M	**–3.65**	**2.33E-12**	**0.90**	**2.78**	HCC (Fu et al., 2017) and others (Li et al., 2021)
			MT1X	**–3.55**	**1.47E-17**	**0.88**	**3.86**	HCC (Liu et al., 2018b)
			MT2A	–3.35	4.85E-24	0.91	3.83	Breast cancer (Kim et al., 2011) and others (Pan et al., 2013)
chr17: 81740217–81874724 (~134 kb, 14 DEEs)	1.05	4.53	HGS	**1.04**	**1.95E-20**	**0.87**	**1.22**	HCC (Canal et al., 2015)
			CCDC137	0.95	3.37E-16	0.80	1.78	Unknown
			NPLOC4	0.82	3.53E-11	0.79	1.22	Bladder cancer (Lu et al., 2019) and others (Skrott et al., 2017; Pan et al., 2021)
			CEP131	**1.42**	**2.39E-19**	**0.79**	**1.78**	HCC (Liu et al., 2017) and others (Kim et al., 2019; Wang et al., 2020)
			CSNK1D	0.55	8.36E-08	0.76	1.22	Breast cancer (Bar et al., 2018) and others (Peer et al., 2021)
			MAFG	**0.98**	**3.10E-07**	**0.71**	**1.00**	HCC (Liu et al., 2018c)
			MAFG-DT	**3.03**	**3.42E-28**	**0.73**	**1.33**	HCC (Ouyang et al., 2019) and others (Cui et al., 2018; Li et al., 2019; Sui et al., 2019; Qu and Liu, 2020; Xiao et al., 2020)
			FOXK2	**0.68**	**3.22E-07**	**0.73**	**1.67**	HCC (Lin et al., 2017b) and others (Shan et al., 2016; Nestal de Moraes et al., 2019)
			SIRT7	**0.61**	**8.43E-06**	**0.72**	**1.00**	HCC (Zhao et al., 2019) and others (Yu et al., 2014b; Zhang et al., 2015)
			NARF	1.07	9.12E-15	0.74	2.00	Glioblastoma (Anderson et al., 2010)

Notes: Enhancer cluster: the cluster of intergenic DEEs that were simultaneously associated with the corresponding cluster of genes; average LFC.DEE: the arithmetic mean of $\log_2$ fold change of the FPM of all DEEs in the enhancer cluster; average LogFDR.DEE: the arithmetic mean of the $-\log_{10}$FDR of the differential expression test of all DEEs in the enhancer cluster; average rho: the arithmetic mean of Spearman correlation coefficients of all DEE-DEG pairs between corresponding DEGs and DEEs in the enhancer cluster; average concurrence: the arithmetic mean of the concurrence of all DEE-DEG pairs between corresponding DEGs and DEEs in the enhancer cluster; implicated cancers: results of literature searching (only molecular mechanism studies) to determine the relevance between DEE-DEG and cancers (genes implicated in HCC were highlighted with a bold font).

TABLE 3 | Characteristics of five RNA-seq datasets used in the present study.

Dataset	No. of tumor tissues	No. of adjacent tissues	Additional data type	Reference (PMID)
Discovery cohort	33	33	WGBS	—
GSE77314	50	50	—	27119355 (Liu et al., 2016)
GSE124535	35	35	—	30814741 (Jiang et al., 2019)
GSE148355	62	47	—	33772139 (Yoon et al., 2021)
GSE77509	20	20	—	28194035 (Yang et al., 2017)

FIGURE 4 | DNA methylation and histone PTM modifer-associated intergenic DEEs and DEE-DEGs. **(A)** Percentage of DNA methylation-associated intergenic DEEs. **(B)** Percentage of DNA methylation-associated intergenic DEE-DEGs. **(C)** Significant differential expression of seven histone modification regulators (three histone methyltransferases, two histone demethylases, and three histone deacetylases) in the discovery cohort. **(D)** Proportion of histone modification regulator-associated intergenic DEEs. **(E)** Percentage of histone modification regulator-associated intergenic DEE-DEGs.

respectively, which were much higher than the corresponding percentages for significantly correlated DEEs (26.26 and 35.41%, respectively) (**Figures 4D,E**), suggesting that histone modification-associated DEEs are also more likely to be transcriptional master regulators.

Aberrant Genic Enhancers-Associated With DNA Methylation Alterations

Among those collected genic enhancers, 1,119 DMEs and their associated DME-DEGs were identified through the integration of WGBS, ChIP-seq, and RNA-seq data. Overall, hypomethylated and hypermethylated DMEs displayed similar transcriptional regulation patterns (i.e., they tended to be correlated with equal number of DME-DEGs) (**Figure 5A**). In total, there were 1,442 genic DEE-DEGs, including 120 hypermethylated upregulated DEE-DEGs (HyperUp), 168 hypermethylated downregulated DEE-DEGs (HyperDown), 517 hypomethylated upregulated DEE-DEGs (HypoUp), and 637 hypomethylated downregulated DEE-DEGs (HypoDown) (**Figure 5B**; **Supplementary Table S3.1**). Approximately half (52.50%) of the identified DEE-DEGs exhibited a nonclassical positive correlation between DNA methylation and gene expression. The results of independent replication of those 1,442 DME-DEGs in TCGA-LIHC showed that the raw

replication rates of the four types of DEE-DEGs (i.e., HyperUp, HyperDown, HypoUp, and HypoDown) were 47.50, 42.86, 28.63, and 14.44%, respectively (**Figure 5C**). Since the 450 k methylation array barely covered CpGs located in the gene body and intergenic regions, which were primarily hypomethylated, it was not surprising to observe much lower raw replication rates and higher type I failure ratios for the HypoUp and HypoDown groups. In contrast, their platform-adjusted replication rates reached 67.86, 59.02, 65.78, and 50.00% (**Figure 5D**), which were comparable to each other.

Intergration of Aberrant Enhancer-Associated Transcriptional Dysregulation and Sucessfully in Silico Verification Based on HiC Data

After combining 1,442 genic DME-DEGs with the 810 intergenic DEE-DEGs, we obtained a set of 2,051 aberrant enhancer-associated DEGs (AE-DEGs), which was composed of 1,092 upregulated AE-DEGs and 959 downregulated AE-DEGs (**Figure 6A**; **Supplementary Table S3.2**). Pathway/biological process enrichment analyses demonstrated that 1,092 activated AE-DEGs were overrepresented for genes implicated in the cell cycle, nuclear division, DNA repair, and DNA replication

FIGURE 5 | Identification and validation of genic DMEs and associated DME-DEGs in HCC. **(A)** Distribution density of the number of associated DME-DEGs of hypermethylated and hypomethylated genic DME. **(B)** Count of four types of genic DME-DEGs. "HyperUp" refers to hypermethylated enhancer-associated upregulated DME-DEGs; "HyperDown" refers to hypomethylated enhancer-associated downregulated DME-DEGs; "HypoUp" refers to hypermethylated enhancer-associated upregulated DME-DEGs, and "HypoDown" refers to hypomethylated enhancer-associated downregulated DME-DEGs. **(C)** Distribution of the three types of replication results of genic DME-DEGs. "Successful replication" refers to the successful replication of genic DME-DEGs for correlated differential methylation and differential expression in TCGA-LIHC; "type I failure" refers to replication failure due to lack of CpG for the corresponding genic DMEs in TCGA-LIHC; and "type II failure" refers to replication failures except type I failure. **(D)** Platform-adjusted replication rates of four types of genic DME-DEGs in TCGA-LIHC. Platform-adjusted replication rates were calculated as $(Count_{Successful\ Replication} + Count_{Successful\ Replication} + Count_{Type\ I\ Failure}) * 100\%$.

(**Figure 6B**), while 959 repressed AE-DEGs were enriched for genes involved in monocarboxylic acid metabolism, adaptive immune response, biological oxidation, and cytochrome P450 epoxygenase pathway (**Figure 6C**). Moreover, hypergeometric test revealed that AE-DEGs showed significant enrichment for genes related to four cancer hallmarks including genome instability and mutation (FDR = 6.3e−9), reprogramming energy metabolism (FDR = 1.2e−3), resisting cell death (FDR = 7.3e−3), and evading immune destruction (FDR = 4.2e−2) (**Figure 6D**).

In addition, AE-DEGs were further verificated according to HiC-produced TADs and chromatin loops. The results showed that 63.24 and 71.62% of AE-DEGs were in the same TAD, in which their corresponding enhancers were located, in the HiC profiled HepG2 and normal adult liver tissue sample (**Figure 6E**; **Supplementary Table S3.3**). Moreover, the TAD validation results in two HiC samples were highly consistent. Specifically, 56.95 and 80.35% of AE-DEGs were successfully supported by TADs in both samples and either sample, respectively (**Figure 6E**;

FIGURE 6 | Biological functions and in silico verification of AE-DEGs. **(A)** Venn diagram displaying the overlap between genic DME-DEGs and intergenic DEE-DEGs. The union of them were defined as aberrant enhancer-associated DEGs (AE-DEGs). **(B)** and **(C)** Top ten overrepresented pathways/GO terms of activated AE-DEGs and repressed AE-DEGs, respectively. **(D)** Enrichment of AE-DEGs for ten cancer hallmarks. "*"refers to hypergeometric test FDR < 0.05; "***"refers to FDR < 1e-2; and "****"refers to FDR < 1e-3. **(E)** Percentage of AE-DEGs that were successfully validated by TADs in HiC-profiled HepG2 and normal liver samples. "HepG2 or liver" refers to successful validation in HepG2 or the liver sample; "HepG2 and liver" refers to successful validation in both HepG2 and the liver sample.

Supplementary Table S3.3). Furthermore, 12 and one AE-DEG were confirmed by chromatin interaction loops in HepG2 and normal liver sample (**Supplementary Table S3.4**), respectively.

Construction of a Six AE-DEGs -Based Prognostic Model

Through univariate Cox regression, LASSO, and multivariate Cox regression model filtering, 2051 AE-DEGs were eventually filtered to six genes to build a prognostic model for OS in TCGA-LIHC. These six AE-DEGs included procollagen-

lysine,2-oxoglutarate 5-dioxygenase 2 (PLOD2), homeobox D9 (HOXD9), BOP1 ribosomal biogenesis factor, which is also known as Block of Proliferation (BOP1), Ras-related protein Rab-26 (RAB26), killer cell lectin-like receptor K1 (KLRK1), and Ral guanine nucleotide dissociation stimulator like 4 (RGL4) (**Figure 7A**). A prognostic risk score was calculated for each patient as follows: the risk score = (0.424 * expression of PLOD2) + (0.109 * expression of HOXD9) + (0.184 * expression of BOP1) + (−0.134 * expression of RAB26) + (−0.185 * expression of KLRK1) + (−0.0547 * expression of RGL4). An optimal cutoff at 7.37 was applied to divide all

FIGURE 7 | Construction of a six AE-DEG-based prognostic model for HCC in TCGA-LIHC. **(A)** The expression heatmap of six AE-DEGs constituted the identified prognostic model for OS of HCC in TCGA-LIHC. Multivariate Cox regression derived coefficients used for the calculation of risk score are given in parentheses. Patients were ranked according to corresponding calculated risk scores. **(B)** Distribution of the calculated risk scores of HCC patients in TCGA-LIHC. **(C)** Kaplan-Meier analysis of the six AE-DEG-based prognostic signature in TCGA-LIHC. **(D)** Distribution of duration and survival status of HCC patients in TCGA-LIHC. **(E)** Box plots display the comparison of survival times between high- and low-risk HCC patients and the comparison of risk scores between alive and deceased HCC patients in TCGA-LIHC. Wilcox p-values were calculated and displayed with each boxplot. **(F)** Time-dependent ROC analyses of the six AE-DEG-based prognostic signature in TCGA-LIHC. **(G)** Forest plot of the multivariate Cox regression analysis in TCGA-LIHC.

patients into high-risk (N = 100) and low-risk (N = 265) groups (**Figure 7B**). Kaplan-Meier analysis revealed significant differences in OS probability across time between high-risk and low-risk groups ($p < 2.0e{-}16$) (**Figure 7C**). Wilcox rank-sum exact tests illuminated significantly less OS duration in high-risk patients ($p = 1.2e{-}9$), and lower risk scores among alive patients ($p = 4.5e{-}9$) (**Figures 7D,E**). The areas under the time-dependent ROC curves (AUCs) for 1-, 3-, and 5-years OS were estimated to be 0.783, 0.797, and 0.715, respectively (**Figure 7F**). A multivariate Cox regression model constructed using both age, gender, and pathologic TNM stage demonstrated that TNM stage ($p < 0.001$, HR = 2.16) and risk group ($p < 0.001$, HR = 4.42) were both independent prognostic biomarkers for OS of HCC patients in TCGA-LIHC (**Figure 7G**).

Subsequently, a predictive nomogram was built by combining the risk score and TNM stage for accurate prediction of overall survival probability in 1, 3, and 5 years (**Figure 8A**). The calibration plots for internal validation of the nomogram

showed high consistency between the predicted OS outcomes and actual observations (**Figure 8B**). Time-dependent ROC curves revealed the best predictive performance of the nomogram, with AUCs of 0.796, 0.830, and 0.773 for 1-year, 3-years, and 5-years OS, respectively (**Figure 8C**).

Consistent Validation of the Six AE-DEGs-Based Prognostic Model in ICGC-LIRI

Univariate Cox regression revealed that all six AE-DEGs that constituted the identified prognostic signature were significant OS-related biomarkers in ICGC-LIRI cohort (**Figure 9A**). Risk scores were calculated for each ICGC-LIRI patient by using the coefficients estimated from TCGA-LIHC. Similarly, 198 ICGC-LIRI patients were divided into high-risk (N = 65) and low-risk (N = 135) groups according to the corresponding optimal cutoff (**Figure 9B**).

FIGURE 8 | Nomogram for the prediction of overall survival of HCC in TCGA-LIHC. (A) A prognostic nomogram for predicting the probabilities of 1-year, 3-years, and 5-years overall survival of HCC patients in TCGA-LIHC. (B) Calibration plots for evaluation of the predictive performance of the constructed nomogram. (C) Time-dependent ROC curves displayed the comparisons of AUCs among diverse prognostic models.

FIGURE 9 | Validation of the six AE-DEG-based prognostic model for OS of HCC in ICGC-LIRI cohort. **(A)** The expression heatmap of six AE-DEGs constituted the identified prognostic model for overall survival of HCC in ICGC-LIRI cohort. Patients were ranked according to their risk scores. **(B)** Distribution of the calculated risk scores of HCC patients in ICGC-LIRI. **(C)** Kaplan-Meier analysis of the 6-gene prognostic signature in ICGC-LIRI. **(D)** Distribution of duration and survival status of HCC patients in ICGC-LIRI. **(E)** Boxplots display the comparison of survival time between high- and low-risk HCC patients and the comparison of risk score between alive and deceased HCC patients in ICGC-LIRI. **(F)** Time-dependent ROC analysis of the six AE-DEG-based prognostic signature in ICGC-LIRI. **(G)** Forest plot of the multivariate Cox regression analysis in ICGC-LIRI.

Kaplan-Meier analysis revealed significant differences in OS probability across time between high-risk and low-risk group ICGC-LIRI patients ($p = 7.0e-9$) (**Figure 9C**). Wilcox rank-sum exact tests illuminated significantly less OS duration in high-risk patients ($p = 0.0061$), and lower risk scores among alive patients ($p = 6.5e-6$) (**Figures 9D,E**). The AUCs for 1-, 3-, and 5-years OS were estimated as 0.795, 0.756, and 0.800 (**Figure 9F**), respectively. A multivariate Cox regression model constructed using both age, gender, and pathologic TNM stage also confirmed that the risk group ($p < 0.001$, HR = 5.03) was an independent prognostic biomarker for OS of HCC patients in ICGC-LIRI (**Figure 9G**).

The Six AE-DEG-Based Prognostic Model Display Superb Predictive Performance for OS of HCC Patients

Furthermore, the good predictive performance of our identified AE-DEG-based signature was assessed through comparisons with seven established similar prognostic models (Long et al., 2019; Zhang et al., 2020b; Ouyang et al., 2020; Tang et al., 2020; Zhu et al., 2020; He et al., 2021; Wang et al., 2021). Among all signatures, the hypoxia-related gene-based signature and our AE-DEG-based signature were the only two models in which all AUCs were higher than 0.7, which is a well-accepted criterion for high predictive accuracy. Moreover, our model's average AUCs in the discovery and validation cohorts were both higher than those of the hypoxia-related model (0.765 and 0.784 *vs.* 0.723 and 0.763) (**Table 4**). Overall, our prognostic signature had more predictive power than others.

DISCUSSION

In the present study, integration of ChIP-seq and RNA-seq data revealed substantial intergenic DEEs and associated DEE-DEGs in HCC. Compared with activated DEEs and DEE-DEGs, the repressed DEEs and DEE-DEGs displayed higher consistency in multiple HCC

TABLE 4 | Comparison of the predictive performance of our AE-DEG-based signature with seven previously established prognostic signatures in HCC.

Signature name	AUCs for OS in discovery			AUCs for OS in validation		
	1-year	3-years	5-years	1-year	3-years	5-years
Methylation-driven gene based signature 1 (Long et al., 2019)	0.6885	0.6563	0.6548	0.6397	0.6644	0.5942
Methylation-driven gene based signature 2 (He et al., 2021)	0.742	0.661	—	0.695	0.655	—
Angiogenic gene based signature (Zhu et al., 2020)	0.74	0.66	0.66	0.78	0.74	—
EMT related gene based signature (Wang et al., 2021)	0.824	0.798	0.800	0.688	0.674	0.876
Ferroptosis and iron-metabolism related gene based signature (Tang et al., 2020)	0.77	0.71	0.64	0.67	0.73	—
Differentially expressed gene signature (Ouyang et al., 2020)	0.77	0.73	0.72	0.63	0.68	0.65
Hypoxia related gene based signature (Zhang et al., 2020b)	0.78	0.70	0.70	0.75	0.77	0.77
Our AE-DEG based signature	0.783	0.797	0.715	0.795	0.756	0.800

cohorts. Remarkably, 162 of those 387 intergenic DEE-DEGs were concurrent in at least four of the five HCC cohorts that were analyzed in this study. Functional enrichment analysis by Metascape (Zhou et al., 2019) revealed that half of those highly concurrent genes were liver-specific. Enrichment of repressed DEGs for liver-specific genes was previously reported in HCC (Lian et al., 2018). A highly plausible mechanism underlying this phenomenon might be cell dedifferentiation. Cell dedifferentiation is a process that implicates the epigenetic reprogramming of gene activity to transform cells into a less differentiated state like their parent cell type. In the development of HCC, stepwise dedifferentiation is a certain event that exhibits loss of hepatic functions and morphology and gain of hepatic progenitor markers (Chao et al., 2020). Moreover, the well-known demethylation agent 5-azacytidine (5-AZA) displayed potential for usage in dedifferentiation therapy in HCC cell lines and cell-derived xenograft (Gailhouste et al., 2018). In addition, it has been shown that upon loss of the mouse Igκ gene's downstream enhancers, E3′ and Ed, the mature B cells unexpectedly undergo reversible retrograde differentiation (Zhou et al., 2013). Hence, our finds about conservative enhancer repression associated suppression of liver-specific genes might shed new light on epigenetic mechanisms underlying the dedifferentiation that occurs in hepatocarcinogenesis and provide potential targets for dedifferentiation-targeted therapy of HCC.

Notably, highly conserved intergenic DEE-DEGs with counts less than 200 unexpectedly included the majority of all MT genes in the genome. MTs are small cysteine-rich proteins that play pivotal roles in metal homeostasis and protection against heavy metal-related cytotoxicity, DNA damage, and oxidative stress (Coyle et al., 2002). Dysregulation of MTs is ubiquitous in most malignancies, and emerging evidence shows that MTs are implicated in tumor formation, progression, and drug resistance (Si and Lang, 2018; Merlos Rodrigo et al., 2020). As mentioned earlier, nine identified DEE-associated differentially expressed MT isoforms were reported to be involved in liver cancer. Specifically, the upregulation of *MT1A* mediated the attenuation of malignant behaviors of *CT23* knockdown in HCC cells (Ning et al., 2021). *MT1DP* is a pivotal anticancer long noncoding RNA (lncRNA), whose suppression mediates the vital carcinogenetic roles of *RUNX2* and *YAP* in HCC (Yu et al., 2014a). *MT1E* was newly identified as a novel tumor suppressor for HCC that could induce apoptosis and suppress cell growth and metastasis (Liu et al., 2020). Exogenous expression of *MT1F* displayed a strong inhibitive effect on the growth of HepG2 cells

(Lu et al., 2003). *MT1G* was uncovered as a tumor suppressor in HCC by inducing the transcriptional activity of p53 through direct interaction and supply of appropriate zinc ions to p53 (Wang et al., 2019). *MT1H* functions as a tumor suppressor that suppresses the proliferation and invasion of HCC cells by inhibiting the Wnt/β-catenin pathway (Zheng et al., 2017). Overexpression of the lncRNA *MT1JP* remarkably inhibited the proliferation and enhanced apoptosis, which might be mediated by regulating the expression of *AKT* (Wu et al., 2020a). Similarly, *MT1M* also showed a tumor-suppressive ability to suppress cell viability, migration, and invasion and activate apoptosis *in vitro* (Fu et al., 2017). *MT1X* was demonstrated to be a tumor suppressor that suppresses tumor growth and metastasis *in vivo* and induces cell cycle arrest and apoptosis by repressing the NF-κB signaling pathway in HCC (Liu et al., 2018b). The roles of *MT1CP*, *MT1L*, and *MT2A* in HCC are still unknown, while *MT2A* could promote breast cancer invasiveness and might play a suppressive role in gastric cancer through inhibition of the NK-κB signaling pathway (Kim et al., 2011; Pan et al., 2013).

Besides, there were also several sets of upregulated genes associated with activated super-enhancers in HCC. On chromosome 17, a group of 10 activated DEE-DEGs was found to be associated with increased enhancer activity for 14 intergenic DEEs. It is intriguing that six (*HGS*, *CEP131*, *MAFG*, *MAFG-DT*, *FOXK2*, and *SIRT7*) of them have already been discovered as proto-oncogenes in HCC (Canal et al., 2015; Lin et al., 2017b; Liu et al., 2017; Liu et al., 2018c; Ouyang et al., 2019; Zhao et al., 2019). In particular, the lncRNA *MAFG-DT*, which is likewise known as *MAFG-AS1*, was also recently shown to play oncogenic roles in multiple tumors in addition to HCC, including colorectal cancer (Cui et al., 2018), breast carcinoma (Li et al., 2019), bladder urothelial carcinoma (Xiao et al., 2020), esophageal squamous cell carcinoma (Qu and Liu, 2020), and lung adenocarcinoma (Sui et al., 2019). *NPLOC4*, also known as *NPL4*, is uncharacterized in HCC but has been revealed as an important oncogene in bladder cancer (Lu et al., 2019) and a critical target of the anticancer drug disulfiram (Skrott et al., 2017; Pan et al., 2021). Similarly, *CSNK1D* has recently been identified as a novel drug target in Hedgehog/GLI-driven cancers (Peer et al., 2021), and silencing of *CSNK1D* attenuates the migration and metastasis of triple-negative breast cancer cells (Bar et al., 2018). As an E3 ubiquitin

ligase, *NARF* was identified as a positive regulator of cell growth in glioblastoma (Anderson et al., 2010). *CCDC137* has not been characterized in any cancer, but its depletion via HIV could cause cell cycle arrest (Zhang and Bieniasz, 2020). Taken together, the elevated activity of the super-enhancer, which is composed of a cluster of 14 synergistic enhancers located on chromosome 17, was demonstrated to be associated with the activation of several critical oncogenes implicated in HCC and/or other cancers. Therefore, inhibition of this activated super-enhancer might be a promising therapy for HCC.

Our integrative transcriptomic analyses discovered massive concurrent DEEs in HCC, which might be caused by either genetic mutations or epigenetic aberrations. However, those DEEs, especially repressed DEEs, were ubiquitous and conserved in multiple HCC cohorts, which suggests a higher possibility of epigenetic aberration-relevant underlying mechanisms. Indeed, our investigation revealed that considerable DEEs and DEE-DEGs were linked to DNA methylation and histone modification. Notably, there were strong associations between the activation of three histone methyltransferases (*EZH2*, *EHMT2*, and *SMYD3*) and enhancer aberrations. This was consistent with the previous findings that mutations and expression changes of epigenetic modifiers are common events leading to an aggressive gene expression and poor clinical outcomes in HCC (Bayo et al., 2019). *EZH2*, *EHMT2*, and *SMYD3* are vital epigenetic regulators that could be targeted for cancer therapy (Cheng et al., 2019). Unlike those of *EZH2* (Gao et al., 2014; Liu et al., 2015; Zhuang et al., 2016; Chen et al., 2018b), the roles of *EHMT2* and *SMYD3* in mediating transcriptional regulation in carcinogenesis are still not fully characterized in HCC (Zhang et al., 2021a; Guo et al., 2021). Our findings serve as a proof-of-concept that activation of histone methyltransferases, such as *EZH2*, *EHMT2*, and *SMYD3* might promote hepatocarcinogenesis by inducing enhancer aberration of crucial cancer-related genes.

To better assess the clinical outcomes of HCC patients, in this study, we applied machine learning approaches to explore the prognostic significance of AE-DEGs in HCC and established a prognostic model based on a panel of six AE-DEGs, including *PLOD2*, *HOXD9*, *BOP1*, *RAB26*, *KLRK1*, and *RGL4*. Our identified AE-DEG-based signature outperformed clinical characteristics such as the TNM stage and seven previously established similar prognostic models in terms of predictive accuracy, suggesting that those six AE-DEGs might play important roles in HCC. *PLOD2* encodes a key enzyme mediating the formation of the stabilized collagen cross-links, which are considered as the "highway" for cancer cell migration and invasion (Provenzano et al., 2006). The roles of *PLOD2* in breast cancer, sarcoma, bladder cancer, and renal cell carcinoma were thoroughly discussed in a previous review (Du et al., 2017). *PLOD2* was first demonstrated as a prognostic marker for HCC in 2011 (Noda et al., 2012), while the function and mechanism of *PLOD2* activation in HCC have not been thoroughly explored. *HOXD9* and *BOP1* were both uncovered as the oncogenic promoters of epithelial-mesenchymal transition (EMT) in HCC (Chung et al., 2011; Lv et al., 2015), which was in line with their unfavorable prognostic contribution in our identified prognostic signature. On the other hand, *RAB26* was novel in HCC but was newly identified as a suppressor of the migration and invasion of breast cancer cells (Liu et al., 2021). The roles of *KLRK1* and *RGL4* have not been investigated in any malignancies but have been identified as prognostic factors in lung adenocarcinoma (Sun et al., 2020; Zhang et al., 2021b). In summary, previous studies revealed pivotal cancer-related functions of *PLOD2*, *HOXD9*, *BOP1*, and *RAB26*, manifesting our findings of their AE-associated dysregulation and prognostic significance in OS of HCC, and suggesting the possibility that *PLOD2*, *RAB26*, *KLRK1*, and *RGL4* play essential roles in the progression and survival of HCC, although further experimental investigations are warranted.

Our study uncovered systematic enhancer aberrations with important functions and excellent prognostic significance in HCC. There are still several potential limitations. First of all, RNA-seq is still commonly used in the literature (Chen et al., 2018a; Wu et al., 2020b; Chen and Liang, 2020) but is not one of the best choices for the comprehensive detection of eRNA; for example, GRO-seq would be a better approach (Danko et al., 2015; Zhang et al., 2020c). Second, aberrant genic enhancers might be only partially captured by identifying genic DMEs, especially considering the relatively low ratio of DNA methylation-associated DEEs in total intergenic DEEs. Moreover, although our identified AE-DEGs were successfully replicated in independent cohorts and confirmed by TADs from HiC, further validation of enhancer-mediated transcriptional regulation of particular genes *via* experimental technologies such as CRISPR, like in previous enhancer related studies (Chen et al., 2018a; Xiong et al., 2019), was lacking in our present study and will be part of our ongoing works.

CONCLUSION

In conclusion, our integrative analysis of the epigenome and transcriptome depicted and verified a systematic landscape of aberrant enhancers and 2051 associated DEGs, including many well-known cancer-related genes, in HCC. These findings provide new insight into the roles of epigenetic aberration induced aberrant enhancers in the progression of HCC. Furthermore, our established prognostic signature based on six AE-DEGs displayed superior predictive performance over previous models for predicting the long-term and short-term OS of HCC patients.

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 below: NCBI SRA database [accession: PRJNA762641].

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The institutional review board of the First Affiliated Hospital, Zhejiang University School of Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

PH, BZ, JZ, and ML contributed to the conception and design of the study. PH organized the database and performed the statistical analysis. PH wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.827657/full#supplementary-material

REFERENCES

Alboukadel Kassambara, M. K. (2021). *Przemyslaw Biecek, and Scheipl Fabian Survminer: Survival Analysis and Visualization.*

Anderson, T. W., Wright, C., and Brooks, W. S. (2010). The E3 Ubiquitin Ligase NARF Promotes colony Formation *In Vitro* and Exhibits Enhanced Expression Levels in Glioblastoma Multiforme *In Vivo. Am. J. Undergraduate Res.* 9, 23–30. doi:10.33697/ajur.2010.017

Bar, I., Merhi, A., Larbanoix, L., Constant, M., Haussy, S., Laurent, S., et al. (2018). Silencing of Casein Kinase 1 delta Reduces Migration and Metastasis of Triple Negative Breast Cancer Cells. *Oncotarget* 9 (56), 30821–30836. doi:10.18632/oncotarget.25738

Bayo, J., Fiore, E. J., Dominguez, L. M., Real, A., Malvicini, M., Rizzo, M., et al. (2019). A Comprehensive Study of Epigenetic Alterations in Hepatocellular Carcinoma Identifies Potential Therapeutic Targets. *J. Hepatol.* 71 (1), 78–90. doi:10.1016/j.jhep.2019.03.007

Canal, F., Anthony, E., Lescure, A., Del Nery, E., Camonis, J., Perez, F., et al. (2015). A Kinome siRNA Screen Identifies HGS as a Potential Target for Liver Cancers with Oncogenic Mutations in CTNNB1. *BMC Cancer* 15, 1020. doi:10.1186/s12885-015-2037-8

Chao, J., Zhao, S., and Sun, H. (2020). Dedifferentiation of Hepatocellular Carcinoma: Molecular Mechanisms and Therapeutic Implications. *Am. J. Transl Res.* 12 (5), 2099–2109.

Chen, H., Li, C., Peng, X., Zhou, Z., Weinstein, J. N., Caesar-Johnson, S. J., et al. (2018). A Pan-Cancer Analysis of Enhancer Expression in Nearly 9000 Patient Samples. *Cell* 173 (2), 386–e12. e312. doi:10.1016/j.cell.2018.03.027

Chen, H., and Liang, H. (2020). A High-Resolution Map of Human Enhancer RNA Loci Characterizes Super-enhancer Activities in Cancer. *Cancer Cell* 38 (5), 701–715. e705. doi:10.1016/j.ccell.2020.08.020

Chen, S., Pu, J., Bai, J., Yin, Y., Wu, K., Wang, J., et al. (2018). EZH2 Promotes Hepatocellular Carcinoma Progression through Modulating miR-22/galectin-9 axis. *J. Exp. Clin. Cancer Res.* 37 (1), 3. doi:10.1186/s13046-017-0670-6

Cheng, Y., He, C., Wang, M., Ma, X., Mo, F., Yang, S., et al. (2019). Targeting Epigenetic Regulators for Cancer Therapy: Mechanisms and Advances in Clinical Trials. *Sig Transduct Target. Ther.* 4, 62. doi:10.1038/s41392-019-0095-0

Chung, K.-Y., Cheng, I. K.-C., Ching, A. K.-K., Chu, J.-H., Lai, P. B.-S., and Wong, N. (2011). Block of Proliferation 1 (BOP1) Plays an Oncogenic Role in Hepatocellular Carcinoma by Promoting Epithelial-To-Mesenchymal Transition. *Hepatology* 54 (1), 307–318. doi:10.1002/hep.24372

Consortium, E. P. (2004). The ENCODE (ENCyclopedia of DNA Elements) Project. *Science* 306 (5696), 636–640. doi:10.1126/science.1105136

Coyle, P., Philcox, J. C., Carey, L. C., and Rofe, A. M. (2002). Metallothionein: the Multipurpose Protein. *Cell Mol. Life Sci. (Cmls)* 59 (4), 627–647. doi:10.1007/s00018-002-8454-2

Cui, S., Yang, X., Zhang, L., Zhao, Y., and Yan, W. (2018). LncRNA MAFG-AS1 Promotes the Progression of Colorectal Cancer by Sponging miR-147b and Activation of NDUFA4. *Biochem. Biophysical Res. Commun.* 506 (1), 251–258. doi:10.1016/j.bbrc.2018.10.112

Danko, C. G., Hyland, S. L., Core, L. J., Martins, A. L., Waters, C. T., Lee, H. W., et al. (2015). Identification of Active Transcriptional Regulatory Elements from GRO-Seq Data. *Nat. Methods* 12 (5), 433–438. doi:10.1038/nmeth.3329

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* 29 (1), 15–21. doi:10.1093/bioinformatics/bts635

Du, H., Pang, M., Hou, X., Yuan, S., and Sun, L. (2017). PLOD2 in Cancer Research. *Biomed. Pharmacother.* 90, 670–676. doi:10.1016/j.biopha.2017.04.023

Feng, R.-M., Zong, Y.-N., Cao, S.-M., and Xu, R.-H. (2019). Current Cancer Situation in China: Good or Bad News from the 2018 Global Cancer Statistics? *Cancer Commun.* 39 (1), 22. doi:10.1186/s40880-019-0368-6

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* 33 (1), 1–22. doi:10.18637/jss.v033.i01

Fu, C.-L., Pan, B., Pan, J.-H., and Gan, M.-F. (2017). Metallothionein 1M Suppresses Tumorigenesis in Hepatocellular Carcinoma. *Oncotarget* 8 (20), 33037–33046. doi:10.18632/oncotarget.16521

Fu, J., Lv, H., Guan, H., Ma, X., Ji, M., He, N., et al. (2013). Metallothionein 1G Functions as a Tumor Suppressor in Thyroid Cancer through Modulating the PI3K/Akt Signaling Pathway. *BMC Cancer* 13, 462. doi:10.1186/1471-2407-13-462

Gai, C., Liu, C., Wu, X., Yu, M., Zheng, J., Zhang, W., et al. (2020). MT1DP Loaded by Folate-Modified Liposomes Sensitizes Erastin-Induced Ferroptosis via Regulating miR-365a-3p/NRF2 axis in Non-small Cell Lung Cancer Cells. *Cell Death Dis* 11 (9), 751. doi:10.1038/s41419-020-02939-3

Gailhouste, L., Liew, L. C., Yasukawa, K., Hatada, I., Tanaka, Y., Nakagama, H., et al. (2018). Differentiation Therapy by Epigenetic Reconditioning Exerts Antitumor Effects on Liver Cancer Cells. *Mol. Ther.* 26 (7), 1840–1854. doi:10.1016/j.ymthe.2018.04.018

Gao, S.-B., Zheng, Q.-F., Xu, B., Pan, C.-B., Li, K.-L., Zhao, Y., et al. (2014). EZH2 Represses Target Genes through H3K27-dependent and H3K27-independent Mechanisms in Hepatocellular Carcinoma. *Mol. Cancer Res.* 12 (10), 1388–1397. doi:10.1158/1541-7786.mcr-14-0034

Grandhi, M. S., Kim, A. K., Ronnekleiv-Kelly, S. M., Kamel, I. R., Ghasebeh, M. A., and Pawlik, T. M. (2016). Hepatocellular Carcinoma: From Diagnosis to Treatment. *Surg. Oncol.* 25 (2), 74–85. doi:10.1016/j.suronc.2016.03.002

Guo, Y., Zhao, Y.-R., Liu, H., Xin, Y., Yu, J.-Z., Zang, Y.-J., et al. (2021). EHMT2 Promotes the Pathogenesis of Hepatocellular Carcinoma by Epigenetically Silencing APC Expression. *Cell Biosci* 11 (1), 152. doi:10.1186/s13578-021-00663-9

Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., et al. (2012). GENCODE: the Reference Human Genome Annotation for the ENCODE Project. *Genome Res.* 22 (9), 1760–1774. doi:10.1101/gr.135350.111

He, D., Liao, S., Cai, L., Huang, W., Xie, X., and You, M. (2021). Integrated Analysis of Methylation-Driven Genes and Pretreatment Prognostic Factors in Patients with Hepatocellular Carcinoma. *BMC Cancer* 21 (1), 599. doi:10.1186/s12885-021-08314-5

Herranz, D., Ambesi-Impiombato, A., Palomero, T., Schnell, S. A., Belver, L., Wendorff, A. A., et al. (2014). A NOTCH1-Driven MYC Enhancer Promotes

T Cell Development, Transformation and Acute Lymphoblastic Leukemia. *Nat. Med.* 20 (10), 1130–1137. doi:10.1038/nm.3665

Hlady, R. A., Sathyanarayan, A., Thompson, J. J., Zhou, D., Wu, Q., Pham, K., et al. (2019). Integrating the Epigenome to Identify Drivers of Hepatocellular Carcinoma. *Hepatology* 69 (2), 639–652. doi:10.1002/hep.30211

Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-André, V., Sigova, A. A., et al. (2013). Super-enhancers in the Control of Cell Identity and Disease. *Cell* 155 (4), 934–947. doi:10.1016/j.cell.2013.09.053

Huang, P., Xu, M., Han, H., Zhao, X., Li, M. D., 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

Hur, H., Ryu, H.-H., Li, C.-H., Kim, I. Y., Jang, W.-Y., and Jung, S. (2016). Metallothinein 1E Enhances Glioma Invasion through Modulation Matrix Metalloproteinases-2 and 9 in U87MG Mouse Brain Tumor Model. *J. Korean Neurosurg. Soc.* 59 (6), 551–558. doi:10.3340/jkns.2016.59.6.551

Jiang, Y., Sun, A., Sun, A., Zhao, Y., Ying, W., Sun, H., et al. (2019). Proteomics Identifies New Therapeutic Targets of Early-Stage Hepatocellular Carcinoma. *Nature* 567 (7747), 257–261. doi:10.1038/s41586-019-0987-8

Kim, D. H., Ahn, J. S., Han, H. J., Kim, H.-M., Hwang, J., Lee, K. H., et al. (2019). Cep131 Overexpression Promotes Centrosome Amplification and colon Cancer Progression by Regulating Plk4 Stability. *Cel Death Dis* 10 (8), 570. doi:10.1038/s41419-019-1778-8

Kim, H. G., Kim, J. Y., Han, E. H., Hwang, Y. P., Choi, J. H., Park, B. H., et al. (2011). Metallothionein-2A Overexpression Increases the Expression of Matrix Metalloproteinase-9 and Invasion of Breast Cancer Cells. *FEBS Lett.* 585 (2), 421–428. doi:10.1016/j.febslet.2010.12.030

Kosinski, M., and Biecek, P. (2015). *RTCGA:The Cancer Genome Atlas Data Integration*. R bioconductor package. Available at: https://rtcga.github.io

Krueger, F., and Andrews, S. R. (2011). Bismark: a Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications. *Bioinformatics* 27 (11), 1571–1572. doi:10.1093/bioinformatics/btr167

Li, D., Peng, W., Wu, B., Liu, H., Zhang, R., Zhou, R., et al. (2021). Metallothionein MT1M Suppresses Carcinogenesis of Esophageal Carcinoma Cells through Inhibition of the Epithelial-Mesenchymal Transition and the SOD1/PI3K Axis. *Mol. Cell* 44 (4), 267–278. doi:10.14348/molcells.2021.2179

Li, H., Zhang, G. Y., Pan, C. H., Zhang, X. Y., and Su, X. Y. (2019). LncRNA MAFG-AS1 Promotes the Aggressiveness of Breast Carcinoma through Regulating miR-339-5p/MMP15. *Eur. Rev. Med. Pharmacol. Sci.* 23 (7), 2838–2846. doi:10.26355/eurrev_201904_17561

Lian, Q., Wang, S., Zhang, G., Wang, D., Luo, G., Tang, J., et al. (2018). HCCDB: A Database of Hepatocellular Carcinoma Expression Atlas. *Genomics, Proteomics & Bioinformatics* 16 (4), 269–275. doi:10.1016/j.gpb.2018.07.003

Likhitsup, A., and Parikh, N. D. (2020). Economic Implications of Hepatocellular Carcinoma Surveillance and Treatment: A Guide for Clinicians. *Pharmacoeconomics* 38 (1), 5–24. doi:10.1007/s40273-019-00839-9

Lin, M.-F., Yang, Y.-F., Peng, Z.-P., Zhang, M.-F., Liang, J.-Y., Chen, W., et al. (2017). FOXK2, Regulted by miR-1271-5p, Promotes Cell Growth and Indicates Unfavorable Prognosis in Hepatocellular Carcinoma. *Int. J. Biochem. Cel Biol.* 88, 155–161. doi:10.1016/j.biocel.2017.05.019

Lin, Z., Lai, S., He, X., Zhuo, W., Wang, L., Si, J., et al. (2017). Decreased Long Non-coding RNA MTM Contributes to Gastric Cancer Cell Migration and Invasion via Modulating MT1F. *Oncotarget* 8 (57), 97371–97383. doi:10.18632/oncotarget.22126

Liu, G., Hou, G., Li, L., Li, Y., Zhou, W., and Liu, L. (2016). Potential Diagnostic and Prognostic Marker Dimethylglycine Dehydrogenase (DMGDH) Suppresses Hepatocellular Carcinoma Metastasis *In Vitro* and *In Vivo*. *Oncotarget* 7 (22), 32607–32616. doi:10.18632/oncotarget.8927

Liu, H., Liu, Y., Liu, W., Zhang, W., and Xu, J. (2015). EZH2-mediated Loss of miR-622 Determines CXCR4 Activation in Hepatocellular Carcinoma. *Nat. Commun.* 6, 8494. doi:10.1038/ncomms9494

Liu, H., Zhou, Y., Qiu, H., Zhuang, R., Han, Y., Liu, X., et al. (2021). Rab26 Suppresses Migration and Invasion of Breast Cancer Cells through Mediating Autophagic Degradation of Phosphorylated Src. *Cel Death Dis* 12 (4), 284. doi:10.1038/s41419-021-03561-7

Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* 173 (2), 400–e11. e411. doi:10.1016/j.cell.2018.02.052

Liu, Q., Lu, F., and Chen, Z. (2020). Identification of MT1E as a Novel Tumor Suppressor in Hepatocellular Carcinoma. *Pathol. - Res. Pract.* 216 (11), 153213. doi:10.1016/j.prp.2020.153213

Liu, T., Yang, H., Fan, W., Tu, J., Li, T. W. H., Wang, J., et al. (2018). Mechanisms of MAFG Dysregulation in Cholestatic Liver Injury and Development of Liver Cancer. *Gastroenterology* 155 (2), 557–571. e514. doi:10.1053/j.gastro.2018.04.032

Liu, X.-H., Yang, Y.-F., Fang, H.-Y., Wang, X.-H., Zhang, M.-F., and Wu, D.-C. (2017). CEP131 Indicates Poor Prognosis and Promotes Cell Proliferation and Migration in Hepatocellular Carcinoma. *Int. J. Biochem. Cel Biol.* 90, 1–8. doi:10.1016/j.biocel.2017.07.001

Liu, Z., Ye, Q., Wu, L., Gao, F., Xie, H., Zhou, L., et al. (2018). Metallothionein 1 Family Profiling Identifies MT1X as a Tumor Suppressor Involved in the Progression and Metastastatic Capacity of Hepatocellular Carcinoma. *Mol. Carcinogenesis* 57 (11), 1435–1444. doi:10.1002/mc.22846

Long, J., Chen, P., Lin, J., Bai, Y., Yang, X., Bian, J., et al. (2019). DNA Methylation-Driven Genes for Constructing Diagnostic, Prognostic, and Recurrence Models for Hepatocellular Carcinoma. *Theranostics* 9 (24), 7251–7267. doi:10.7150/thno.31155

Lossos, I. S., Czerwinski, D. K., Alizadeh, A. A., Wechser, M. A., Tibshirani, R., Botstein, D., et al. (2004). Prediction of Survival in Diffuse Large-B-Cell Lymphoma Based on the Expression of Six Genes. *N. Engl. J. Med.* 350 (18), 1828–1837. doi:10.1056/nejmoa032520

Love, M. I., Huber, W., and Anders, S. (2014). Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8

Lu, B.-S., Yin, Y.-W., Zhang, Y.-P., Guo, P.-Y., Li, W., and Liu, K.-L. (2019). Upregulation of NPL4 Promotes Bladder Cancer Cell Proliferation by Inhibiting DXO Destabilization of Cyclin D1 mRNA. *Cancer Cel Int* 19, 149. doi:10.1186/s12935-019-0874-2

Lu, D. D., Chen, Y. C., Zhang, X. R., Cao, X. R., Jiang, H. Y., and Yao, L. (2003). The Relationship between metallothionein-1F (MT1F) Gene and Hepatocellular Carcinoma. *Yale J. Biol. Med.* 76 (2), 55–62.

Lv, X., Li, L., Lv, L., Qu, X., Jin, S., Li, K., et al. (2015). HOXD9 Promotes Epithelial-Mesenchymal Transition and Cancer Metastasis by ZEB1 Regulation in Hepatocellular Carcinoma. *J. Exp. Clin. Cancer Res.* 34, 133. doi:10.1186/s13046-015-0245-3

Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., et al. (2014). An Oncogenic Super-enhancer Formed through Somatic Mutation of a Noncoding Intergenic Element. *Science* 346 (6215), 1373–1377. doi:10.1126/science.1259037

Marrero, J. A., Kulik, L. M., Sirlin, C. B., Zhu, A. X., Finn, R. S., Abecassis, M. M., et al. (2018). Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. *Hepatology* 68 (2), 723–750. doi:10.1002/hep.29913

Merlos Rodrigo, M. A., Jimenez Jimemez, A. M., Haddad, Y., Bodoor, K., Adam, P., Krizkova, S., et al. (2020). Metallothionein Isoforms as Double Agents - Their Roles in Carcinogenesis, Cancer Progression and Chemoresistance. *Drug Resist. Updates* 52, 100691. doi:10.1016/j.drup.2020.100691

Nakano, S., Eso, Y., Okada, H., Takai, A., Takahashi, K., and Seno, H. (2020). Recent Advances in Immunotherapy for Hepatocellular Carcinoma. *Cancers (Basel)* 12 (4), 775. doi:10.3390/cancers12040775

Nestal de Moraes, G., Carneiro, L. D. T., Maia, R. C., Lam, E. W., and Sharrocks, A. D. (2019). FOXK2 Transcription Factor and its Emerging Roles in Cancer. *Cancers (Basel)* 11 (3). doi:10.3390/cancers11030393

Ning, W., Huang, M., Wu, S., Wang, H., Yao, J., Ge, Y., et al. (2021). CT23 Knockdown Attenuating Malignant Behaviors of Hepatocellular Carcinoma Cell Is Associated with Upregulation of Metallothionein 1. *Cell Biol Int* 45 (6), 1231–1245. doi:10.1002/cbin.11567

Noda, T., Yamamoto, H., Takemasa, I., Yamada, D., Uemura, M., Wada, H., et al. (2012). PLOD2 Induced under Hypoxia Is a Novel Prognostic Factor for Hepatocellular Carcinoma after Curative Resection. *Liver Int.* 32 (1), 110–118. doi:10.1111/j.1478-3231.2011.02619.x

Ouyang, G., Yi, B., Pan, G., and Chen, X. (2020). A Robust Twelve-Gene Signature for Prognosis Prediction of Hepatocellular Carcinoma. *Cancer Cel Int* 20, 207. doi:10.1186/s12935-020-01294-9

Ouyang, H., Zhang, L., Xie, Z., and Ma, S. (2019). Long Noncoding RNA MAFG-AS1 Promotes Proliferation, Migration and Invasion of Hepatocellular

Carcinoma Cells through Downregulation of miR-6852. *Exp. Ther. Med.* 18 (4), 2547–2553. doi:10.3892/etm.2019.7850

Pan, M., Zheng, Q., Yu, Y., Ai, H., Xie, Y., Zeng, X., et al. (2021). Seesaw Conformations of Npl4 in the Human P97 Complex and the Inhibitory Mechanism of a Disulfiram Derivative. *Nat. Commun.* 12 (1), 121. doi:10.1038/s41467-020-20359-x

Pan, Y., Huang, J., Xing, R., Yin, X., Cui, J., Li, W., et al. (2013). Metallothionein 2A Inhibits NF-Kb Pathway Activation and Predicts Clinical Outcome Segregated with TNM Stage in Gastric Cancer Patients Following Radical Resection. *J. Transl Med.* 11, 173. doi:10.1186/1479-5876-11-173

Patrick, J. (2013). *Heagerty aPS-C: survivalROC: Time-dependent ROC Curve Estimation from Censored Survival Data*.

Peer, E., Aichberger, S. K., Vilotic, F., Gruber, W., Parigger, T., Grund-Gröschke, S., et al. (2021). Casein Kinase 1D Encodes a Novel Drug Target in Hedgehog-GLI Driven Cancers and Tumor-Initiating Cells Resistant to SMO Inhibition. *bioRxiv* 13 (16), 4227. doi:10.3390/cancers13164227

Provenzano, P. P., Eliceiri, K. W., Campbell, J. M., Inman, D. R., White, J. G., and Keely, P. J. (2006). Collagen Reorganization at the Tumor-Stromal Interface Facilitates Local Invasion. *BMC Med.* 4 (1), 38. doi:10.1186/1741-7015-4-38

Qu, Y., and Liu, J. (2020). lncRNA MAFG-AS1 Contributes to Esophageal Squamous-Cell Carcinoma Progression via Regulating miR143/LASP1. *Ott* 13, 8359–8370. doi:10.2147/ott.s258157

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* 26 (6), 841–842. doi:10.1093/bioinformatics/btq033

Rms (2021). *Rms: Regression Modeling Strategies*. R package.

Shan, L., Zhou, X., Liu, X., Wang, Y., Su, D., Hou, Y., et al. (2016). FOXK2 Elicits Massive Transcription Repression and Suppresses the Hypoxic Response and Breast Cancer Carcinogenesis. *Cancer Cell* 30 (5), 708–722. doi:10.1016/j.ccell.2016.09.010

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional Enhancers: from Properties to Genome-wide Predictions. *Nat. Rev. Genet.* 15 (4), 272–286. doi:10.1038/nrg3682

Si, M., and Lang, J. (2018). The Roles of Metallothioneins in Carcinogenesis. *J. Hematol. Oncol.* 11 (1), 107. doi:10.1186/s13045-018-0645-x

Skrott, Z., Mistrik, M., Andersen, K. K., Friis, S., Majera, D., Gursky, J., et al. (2017). Alcohol-abuse Drug Disulfiram Targets Cancer via P97 Segregase Adaptor NPL4. *Nature* 552 (7684), 194–199. doi:10.1038/nature25016

Sui, Y., Lin, G., Zheng, Y., and Huang, W. (2019). LncRNA MAFG-AS1 Boosts the Proliferation of Lung Adenocarcinoma Cells via Regulating miR-744-5p/MAFG axis. *Eur. J. Pharmacol.* 859, 172465. doi:10.1016/j.ejphar.2019.172465

Sun, Y., Zhang, Y., Ren, S., Li, X., Yang, P., Zhu, J., et al. (2020). Low Expression of RGL4 Is Associated with a Poor Prognosis and Immune Infiltration in Lung Adenocarcinoma Patients. *Int. Immunopharmacology* 83, 106454. doi:10.1016/j.intimp.2020.106454

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Tang, B., Zhu, J., Li, J., Fan, K., Gao, Y., Cheng, S., et al. (2020). The Ferroptosis and Iron-Metabolism Signature Robustly Predicts Clinical Diagnosis, Prognosis and Immune Microenvironment for Hepatocellular Carcinoma. *Cell Commun Signal* 18 (1), 174. doi:10.1186/s12964-020-00663-1

Therneau, T. M. (2020). *A Package for Survival Analysis in R*.

Tsang, F. H., Law, C. T., Tang, T. C., Cheng, C. L., Chin, D. W., Tam, W. V., et al. (2019). Aberrant Super-enhancer Landscape in Human Hepatocellular Carcinoma. *Hepatology* 69 (6), 2502–2517. doi:10.1002/hep.30544

Wahid, B., Ali, A., Rafique, S., and Idrees, M. (2017). New Insights into the Epigenetics of Hepatocellular Carcinoma. *Biomed. Res. Int.* 2017, 1609575. doi:10.1155/2017/1609575

Wang, J., Yang, X., Han, S., and Zhang, L. (2020). CEP131 Knockdown Inhibits Cell Proliferation by Inhibiting the ERK and AKT Signaling Pathways in Non-small Cell Lung Cancer. *Oncol. Lett.* 19 (4), 3145–3152. doi:10.3892/ol.2020.11411

Wang, X., Xing, Z., Xu, H., Yang, H., and Xing, T. (2021). Development and Validation of Epithelial Mesenchymal Transition-Related Prognostic Model for Hepatocellular Carcinoma. *Aging* 13 (10), 13822–13845. doi:10.18632/aging.202976

Wang, Y., Song, F., Zhang, B., Zhang, L., Xu, J., Kuang, D., et al. (2018). The 3D Genome Browser: a Web-Based Browser for Visualizing 3D Genome Organization and Long-Range Chromatin Interactions. *Genome Biol.* 19 (1), 151. doi:10.1186/s13059-018-1519-9

Wang, Y., Wang, G., Tan, X., Ke, K., Zhao, B., Cheng, N., et al. (2019). MT1G Serves as a Tumor Suppressor in Hepatocellular Carcinoma by Interacting with P53. *Oncogenesis* 8 (12), 67. doi:10.1038/s41389-019-0176-5

Wu, J. H., Xu, K., Liu, J. H., Du, L. L., Li, X. S., Su, Y. M., et al. (2020). LncRNA MT1JP Inhibits the Malignant Progression of Hepatocellular Carcinoma through Regulating AKT. *Eur. Rev. Med. Pharmacol. Sci.* 24 (12), 6647–6656. doi:10.26355/eurrev_202006_21651

Wu, Y., Yang, Y., Gu, H., Tao, B., Zhang, E., Wei, J., et al. (2020). Multi-omics Analysis Reveals the Functional Transcription and Potential Translation of Enhancers. *Int. J. Cancer* 147 (8), 2210–2224. doi:10.1002/ijc.33132

Xiao, M., Liu, J., Xiang, L., Zhao, K., He, D., Zeng, Q., et al. (2020). MAFG-AS1 Promotes Tumor Progression via Regulation of the HuR/PTBP1 axis in Bladder Urothelial Carcinoma. *Clin. Transl Med.* 10 (8), e241. doi:10.1002/ctm2.241

Xiong, L., Wu, F., Wu, Q., Xu, L., Cheung, O. K., Kang, W., et al. (2019). Aberrant Enhancer Hypomethylation Contributes to Hepatic Carcinogenesis through Global Transcriptional Reprogramming. *Nat. Commun.* 10 (1), 335. doi:10.1038/s41467-018-08245-z

Yang, Y., Deng, X., Chen, X., Chen, S., Song, L., Meng, M., et al. (2020). Landscape of Active Enhancers Developed De Novo in Cirrhosis and Conserved in Hepatocellular Carcinoma. *Am. J. Cancer Res.* 10 (10), 3157–3178.

Yang, Y., Chen, L., Gu, J., Zhang, H., Yuan, J., Lian, Q., et al. (2017). Recurrently Deregulated lncRNAs in Hepatocellular Carcinoma. *Nat. Commun.* 8, 14421. doi:10.1038/ncomms14421

Yoon, S.-H., Choi, S.-W., Nam, S. W., Lee, K. B., and Nam, J.-W. (2021). Preoperative Immune Landscape Predisposes Adverse Outcomes in Hepatocellular Carcinoma Patients with Liver Transplantation. *Npj Precis. Onc.* 5 (1), 27. doi:10.1038/s41698-021-00167-2

Yu, H., Wang, S., Zhu, H., and Rao, D. (2019). LncRNA MT1JP Functions as a Tumor Suppressor via Regulating miR-214-3p Expression in Bladder Cancer. *J. Cel Physiol* 1, 1. doi:10.1002/jcp.28274

Yu, H., Ye, W., Wu, J., Meng, X., Liu, R.-y., Ying, X., et al. (2014). Overexpression of Sirt7 Exhibits Oncogenic Property and Serves as a Prognostic Factor in Colorectal Cancer. *Clin. Cancer Res.* 20 (13), 3434–3445. doi:10.1158/1078-0432.ccr-13-2952

Yu, W., Qiao, Y., Tang, X., Ma, L., Wang, Y., Zhang, X., et al. (2014). Tumor Suppressor Long Non-coding RNA, MT1DP Is Negatively Regulated by YAP and Runx2 to Inhibit FoxA1 in Liver Cancer Cells. *Cell Signal.* 26 (12), 2961–2968. doi:10.1016/j.cellsig.2014.09.011

Zhang, B., Tang, B., Gao, J., Li, J., Kong, L., and Qin, L. (2020). A Hypoxia-Related Signature for Clinically Predicting Diagnosis, Prognosis and Immune Microenvironment of Hepatocellular Carcinoma Patients. *J. Transl Med.* 18 (1), 342. doi:10.1186/s12967-020-02492-9

Zhang, D., Huo, D., Xie, H., Wu, L., Zhang, J., Liu, L., et al. (2020). CHG: A Systematically Integrated Database of Cancer Hallmark Genes. *Front. Genet.* 11, 29. doi:10.3389/fgene.2020.00029

Zhang, F., and Bieniasz, P. D. (2020). HIV-1 Vpr Induces Cell Cycle Arrest and Enhances Viral Gene Expression by Depleting CCDC137. *Elife* 9, e55806. doi:10.7554/eLife.55806

Zhang, G., Li, S., Lu, J., Ge, Y., Wang, Q., Ma, G., et al. (2018). LncRNA MT1JP Functions as a ceRNA in Regulating FBXW7 through Competitively Binding to miR-92a-3p in Gastric Cancer. *Mol. Cancer* 17 (1), 87. doi:10.1186/s12943-018-0829-6

Zhang, H., Zheng, Z., Zhang, R., Yan, Y., Peng, Y., Ye, H., et al. (2021). SMYD3 Promotes Hepatocellular Carcinoma Progression by Methylating S1PR1 Promoters. *Cel Death Dis* 12 (8), 731. doi:10.1038/s41419-021-04009-8

Zhang, S., Chen, P., Huang, Z., Hu, X., Chen, M., Hu, S., et al. (2015). Sirt7 Promotes Gastric Cancer Growth and Inhibits Apoptosis by Epigenetically Inhibiting miR-34a. *Sci. Rep.* 5, 9787. doi:10.1038/srep09787

Zhang, Y., Chen, Z., and Gao, G. (2021). *KLRK1 as a Prognostic Biomarker for Lung Adenocarcinoma Cancer*.

Zhang, Y., Qian, M., Tang, F., Huang, Q., Wang, W., Li, Y., et al. (2020). Identification and Analysis of P53-Regulated Enhancers in Hepatic Carcinoma. *Front. Bioeng. Biotechnol.* 8, 668. doi:10.3389/fbioe.2020.00668

Zhao, J., Wozniak, A., Adams, A., Cox, J., Vittal, A., Voss, J., et al. (2019). SIRT7 Regulates Hepatocellular Carcinoma Response to Therapy by Altering the P53-dependent Cell Death Pathway. *J. Exp. Clin. Cancer Res.* 38 (1), 252. doi:10.1186/s13046-019-1246-4

Zheng, Y., Jiang, L., Hu, Y., Xiao, C., Xu, N., Zhou, J., et al. (2017). Metallothionein 1H (MT1H) Functions as a Tumor Suppressor in Hepatocellular Carcinoma through Regulating Wnt/β-Catenin Signaling Pathway. *BMC Cancer* 17 (1), 161. doi:10.1186/s12885-017-3139-2

Zhou, X., Xiang, Y., Ding, X., and Garrard, W. T. (2013). Loss of an Igκ Gene Enhancer in Mature B Cells Results in Rapid Gene Silencing and Partial Reversible Dedifferentiation. *Mol. Cel Biol* 33 (10), 2091–2101. doi:10.1128/mcb.01569-12

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6

Zhu, J., Tang, B., Li, J., Shi, Y., Chen, M., Lv, X., et al. (2020). Identification and Validation of the Angiogenic Genes for Constructing Diagnostic, Prognostic, and Recurrence Models for Hepatocellular Carcinoma. *Aging* 12 (9), 7848–7873. doi:10.18632/aging.103107

Zhuang, C., Wang, P., Huang, D., Xu, L., Wang, X., Wang, L., et al. (2016). A Double-Negative Feedback Loop between EZH2 and miR-26a Regulates Tumor Cell Growth in Hepatocellular Carcinoma. *Int. J. Oncol.* 48 (3), 1195–1204. doi:10.3892/ijo.2016.3336

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.