Original Article

Landscape of active enhancers developed de novo in cirrhosis and conserved in hepatocellular carcinoma

Yao Yang, Xiaoyu Deng, Xinjian Chen, Shihan Chen, Liang Song, Meng Meng, Qi Han, Saber Imani, Shuhui Li, Zhaoyang Zhong, Xiaohui Li, Youcai Deng

1Institute of Materia Medica, College of Pharmacy, Army Medical University (Third Military Medical University), Chongqing 400038, China; 2Department of Cardiovascularology, Airforce Hospital of Southern Theater Command, Guangdong, Guangdong 510062, China; 3Department of Hepatobiliary Surgery, The First Affiliated Hospital of Army Medical University (Third Military Medical University), Chongqing 400038, China; 4The General Hospital of Tibet Military Region, Lhasa, Tibet 850000, China; 5Department of Oncology, The Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan 646000, China; 6Department of Clinical Biochemistry, Faculty of Pharmacy and Laboratory Medicine, Army Medical University (Third Military Medical University), Chongqing 400038, China; 7Cancer Center, Daping Hospital and Research Institute of Surgery, Army Medical University (Third Military Medical University), Chongqing 400042, China

Received August 5, 2020; Accepted September 22, 2020; Epub October 1, 2020; Published October 15, 2020

Abstract: Hepatocellular carcinoma (HCC) patients always have a background of cirrhosis. Aberrant epigenetic changes in cirrhosis provide a conductive environment for HCC tumorigenesis. Active enhancers (AEs) are essential for epigenetic regulation and play an important role in cell development and the progression of many diseases. However, the role of AEs in the progression from cirrhosis to HCC remains unclear. We systemically constructed a landscape of AEs that developed de novo in cirrhosis and were conserved in HCC, referred to as CL-HCC AEs. We observed significant upregulation of these CL-HCC AE-associated genes in cirrhosis and HCC, with no other epigenetic changes. Enrichment analysis of these CL-HCC AE-associated genes revealed enrichment in both hepatocyte-intrinsic tumorigenesis and tumor immune response, which might contribute to HCC tumorigenesis. Analysis of the diagnostic ability of these CL-HCC AE-associated genes revealed enrichment in both hepatocyte-intrinsic tumorigenesis and tumor immune response, which might contribute to HCC tumorigenesis. Analysis of the diagnostic ability of these CL-HCC AE-associated genes provided a five-gene (THBS4, OLFML2B, CDKN3, GABRE, and HDAC11) diagnostic biomarker for HCC. Molecular subtype (MS) identification based on the CL-HCC AE-associated genes identified 3 MSs. Samples representing the 3 MSs showed differences in CL-HCC AE-associated gene expression levels, prognosis, copy number variation (CNV)/mutation frequencies, functional pathways, tumor microenvironment (TME) cell subtypes, immunotherapy responses and putative drug responses. We also found that the BET bromodomain inhibitor JQ1 downregulated the expression of CL-HCC AE-associated genes. Collectively, our results suggest that CL-HCC AEs and their associated genes contribute to HCC tumorigenesis and evolution, and could be used to distinguish the different landscapes of HCC and help explore the mechanism, classification, prediction, and precision therapy of HCC.

Keywords: Hepatocellular carcinoma, active enhancer, cirrhosis, classification, biomarkers, immunotherapy, immune dysfunction, JQ1

Introduction

Hepatocellular carcinoma (HCC) is one of the most common cancers with high mortality worldwide. A large number of studies on the mechanisms of HCC initiation and development have helped to improve the diagnosis, treatment and prognosis of HCC, but the goal is still far from reach [1]. HCC arises from chronic liver disease, fibrosis, and cirrhosis in 70-80% of patients [2], and recent studies have demonstrated that early and progressive epigenetic changes in cirrhosis are a critical determinant of HCC tumorigenesis [3-5]. However, the lack of information regarding changes in the epigenome from normal liver (NL) to cirrhosis or HCC has limited our further knowledge of HCC tumorigenesis. Therefore, an in-depth exploration of epigenetic changes from NL to cirrhosis and HCC may provide new insights into the
Landscape of active enhancers in HCC development

A newly reported epigenetic modification referred to as active enhancers (AEs), marked by posttranslational modifications of both H3K27ac and H3K4me1, are defined as gene-distant cis-regulatory sequences that are capable of inducing strong expression of their target genes [6, 7]. Several lines of evidence have shown that AEs are not only required for cell development, but also participate in cancer initiation and development [8-10]. Robertson et al. recently broadly profiled epigenetic regulation during HCC progression and identified driver events linked to epigenetic deregulation during the initiation, progression and prognosis of HCC [4]. This study provided data on posttranslational modifications of H3K27ac and H3K4me1 that helped define the evolutionary profiles of AEs from NL to cirrhosis and HCC [4]. Therefore, a systematic and deep review of the AE profiles of HCC initiation and progression will help us further understand the mechanisms of HCC development and provide new theoretical evidence for HCC diagnosis, treatment and prognosis.

In this study, we focused on the role of AEs that formed de novo in cirrhosis and were conserved in HCC, referred to as CL-HCC AEs. The data demonstrated that these AEs are major factors in the aberrant expression of their target genes and induce HCC tumorigenesis from cirrhosis in different ways. Furthermore, the target genes of these AEs play an important role in the diagnosis, molecular typing and precise treatment of HCC. Finally, we discuss potential treatments that could reduce or eliminate the aberrant expression of their target genes.

Materials and methods

Data collection

We gathered ten cohorts from the Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) (Table 1). mRNA expression data and corresponding clinical information from the Cancer Genome Atlas (TCGA) pan-cancer cohort were retrieved from the UCSC Xena browser (https://xenabrowser.net/datapages/) [11]. Gene somatic mutation data of the TCGA-LIHC cohort were retrieved using TCGAbiolinks [12]. DNA copy data of the TCGA-LIHC cohort were retrieved from Firehose (https://gdac.broadinstitute.org/).

ChIP-Seq and Hi-C data processing

ChIP-seq reads were mapped to hg19. The SICER package was used to call the peaks [13]. The R package ChIPseeker (version 1.18.0) was used for peak annotation and comparison. AEs were identified by overlapping H3K27ac peaks and H3K4me1 peaks. DeepTools was used to visualize the ChIP-seq data [14]. 3D Genome Browser (http://promoter.bx.psu.edu/hi-c/) was used to visualize the Hi-C data [15].

Enrichment analysis

EnrichR (https://amp.pharm.mssm.edu/Enrichr/) was used for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) enrichment analyses [16]. DisGeNET

### Table 1. Characteristics of the GEO datasets

<table>
<thead>
<tr>
<th>GEO accession</th>
<th>Platform</th>
<th>Data type</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE112221 [4]</td>
<td>GPL16791</td>
<td>RNA-seq, ChIP-seq, DNA methylation and 5-hydroxymethylation</td>
<td>4 HCC, 4 CL and 2 NL tissues</td>
</tr>
<tr>
<td>GSE54238 [38]</td>
<td>GPL16955</td>
<td>Gene chip</td>
<td>26 HCC, 10 CL and 10 NL tissues</td>
</tr>
<tr>
<td>GSE25097 [39]</td>
<td>GPL10687</td>
<td>Gene chip</td>
<td>268 HCC, 243 adjacent nontumor, 40 CL and 6 NL tissues</td>
</tr>
<tr>
<td>GSE44970 [40]</td>
<td>GPL8490</td>
<td>DNA methylation chip</td>
<td>20 HCC, 8 CL and 8 NL tissues</td>
</tr>
<tr>
<td>GSE112679 [41]</td>
<td>GPL18573</td>
<td>DNA 5-hydroxymethylation</td>
<td>1204 HCC, 392 CL, and 958 NL tissues</td>
</tr>
<tr>
<td>GSE124535 [75]</td>
<td>GPL20795</td>
<td>RNA-seq</td>
<td>35 HCC and 35 nontumor tissues</td>
</tr>
<tr>
<td>GSE77509 [76]</td>
<td>GPL16791</td>
<td>RNA-seq</td>
<td>19 HCC and 19 nontumor tissues</td>
</tr>
<tr>
<td>GSE94660 [77]</td>
<td>GPL16791</td>
<td>RNA-seq</td>
<td>21 HCC and 21 nontumor tissues</td>
</tr>
<tr>
<td>GSE51143 [52]</td>
<td>GPL6244</td>
<td>Gene chip</td>
<td>HepG2 cells treated with DMSO or JQ1</td>
</tr>
</tbody>
</table>
Landscape of active enhancers in HCC development

(https://www.disgenet.org/) was used for disease enrichment analysis [17]. Gene sets of 10 cancer hallmarks were downloaded from the CHG database (http://bio-bigdata.hrbmu.edu.cn/CHG/nav_help.html), and a hypergeometric distribution was used to calculate the enrichment degree [18].

**Establishment of the diagnostic model**

First, the R package caret (6.0-84) was used to divide samples into a training set and a test set [19]. Next, the R package glmnet (version 2.0-18) was used to run the LASSO algorithm, which reduces the dimensions and selects features of the training set [20]. Then, random forest (RF), support vector machine (SVM) and logistic regression (LR) were used to build classifiers based on the results of the LASSO algorithm via the R packages glmnet (version 2.0-18), randomForest (4.6-14) and e1071 (1.7-3), respectively [20-22]. The R package pROC (version 1.15.3) was used to display the receiver operating characteristic (ROC) curves and calculate the area under the curve (AUC) [23]. In the end, the best-performing classifier that had the highest AUC was utilized to examine the test set, external independent HCC datasets and TCGA pan-cancer dataset.

**Estimation of immune cell infiltration**

We used Microenvironment Cell Oopulation-counter (MCPCounter) and Estimate the Proportion of Immune and Cancer cells (EPIC) to compute the immune cell infiltration using gene expression data from the GSE112221 and TCGA-LIHC datasets. The R packages MCPCounter (version 1.1.0) and EPIC (version 1.1.5) were used to robustly quantify the immune cells [24, 25].

**Copy number variation (CNV) and mutation analyses**

Gene Pattern modules 2.0 (GISTIC) was used to investigate significant amplification or deletion events (CNVs) in the regions of the genome associated with HCC. The R package maftools was used to analyze mutations based on the TCGA-LIHC Mutect2 pipeline [26]. The tumor mutational burden (TMB) was calculated as the number of mutations per Mb in the genome for each patient based on the TCGA-LIHC Mutect2 pipeline [27]. Predicted neoantigens for each patient were downloaded from The Cancer Imaging Archive (TCIA) dataset (https://tcia.at/home) [28].

**Gene set variation analysis (GSVA)**

The R package GSVA, a nonparametric and unsupervised gene set enrichment method, was used to estimate the score of certain pathways or signatures for single HCC patients based on transcriptome data [29]. The KEGG pathway signatures (c2.cp.kegg.v7.0.symbols.gmt) were downloaded from the Molecular Signatures Database (MSigDB; https://www.gsea-msigdb.org/gsea/msigdb/index.jsp). The hypoxia signature (ACOT7, ADM, ALDOA, CDKN3, EN01, LDHA, MIF, MRPS17, NDRG1, P4HA1, PGAM1, SLC2A1, TPI1, TUBB6 and VEGFA) and T cell dysfunction signature (TGFB1, CD274, CTLA4, IL10, PDCD1, CD276, HAVCR2, TNFRSF9, LAG3, TIGIT, and ICOS) were obtained from previously published studies [30, 31].

**Identification of HCC molecular subtype (MS)**

All 425 CL-HCC AE-associated genes were subjected to nonnegative matrix factorization (NMF) clustering using the R package CancerSubtypes (version 1.8.0) [32]. Before performing NMF, we calculated the expression of the genes associated with overall survival (OS) by univariate Cox analysis, and the genes with significant prognostic values (P < 0.005) were used for sample clustering. Then, we used the sum of squared error (SSE) to evaluate the best number of clusters, and the samples were classified into 3 MSs [33].

**Gene set enrichment analysis (GSEA)**

The R package clusterProfiler (3.10.1) was used to perform GSEA [34]. P < 0.05 was considered statistically significant.

**Immunotherapy and drug responsiveness**

The Tumor Immune Dysfunction and Exclusion (TIDE) tool (http://tide.dfci.harvard.edu/) was used to predict immunotherapy responsiveness [35]. The R package pRRophetic (version 0.5) was used to predict drug sensitivity, which estimated the half-maximal inhibitory concentration (IC50) for each sample by ridge regression, after which the prediction accuracy was evaluated by 10-fold cross-validation based on
Statistical analysis

All computational and statistical analyses were performed using R (version 3.5.1). An unpaired Student’s t-test or the Wilcoxon test was used to compare the differences between 2 groups. One-way ANOVA was used to analyze the differences among 3 or more groups. The chi-square test was used to analyze contingency table variables. The Kaplan-Meier method was used to analyze the differences in prognosis, and then the log-rank test was used to test the significance of prognosis. P < 0.05 was considered statistically significant.

Results

Identified CL-HCC AEs and their target genes

To investigate the implications of AE alterations in cirrhosis and hepatocarcinogenesis, we compared the AE landscapes between NL, cirrhotic liver (CL) and HCC tissue. Based on the histone modifications of H3K27ac and H3K4me1 and gene expression, we identified 620 AEs associated with 483 genes that occur in cirrhosis and are maintained in HCC (Figure 1A). These CL-HCC AE regions showed stronger modification of H3K27ac and H3K4me1 and weaker modification of H3K27me3 (an inactive enhancer mark) and H3K4me3 (a primed enhancer mark, prior to activation) compared with other regions. In addition, the H3K27ac and H3K4me1 signals were stronger in CL and HCC than in NL (Figure 1B) [37]. For the mRNA expression of AE-associated genes, the mRNA expression of CL-HCC AE-associated genes was higher in CL and HCC than in NL in two independent datasets (GSE54238 [38] and GSE25097 [39]) (Figure 1C).

To determine whether the expression of these CL-HCC AE-associated genes is related to DNA methylation or DNA mutations, we analyzed the DNA methylation levels by DNA methylation or 5-hydroxymethylcytosine (5 hmC) using additional independent datasets. The data revealed that neither the DNA methylation nor 5 hmC levels of the CL-HCC AE-associated genes showed significant differences among NL, CL and HCC (DNA methylation: GSE44970 [40]; 5 hmC: GSE112221 [4] and GSE112679 [41]) (Figure 1D and 1E).

Enrichment analysis of the CL-HCC AE-associated genes showed that these genes are mainly involved in tumor-associated pathways (e.g., pathways in cancer) and immune-related pathways (e.g., positive regulation of T cell activation), which are two main conditions needed for the development of HCC (Figure 1F) [42].

All the evidence above demonstrates that the abnormal mRNA expression of these CL-HCC AE-associated genes in CL and HCC is mainly caused by CL-HCC AEs and may play an indispensible role in HCC tumorigenesis.

The role of CL-HCC AEs and their target genes from CL to HCC

Because bulk ChIP-seq data of tumor tissues contain information from multiple cell types, including tumor cells and other stromal or immune cells in the tumor microenvironment (TME), we further explored the role of CL-HCC AEs and CL-HCC AE-associated genes in different cell types. We divided these CL-HCC AEs and their associated genes into two subtypes according to the absence or presence of the CL-HCC AEs in HepG2 cell AEs, an HCC cell line (Figure 2A).

In total, 389 CL-HCC AEs overlapped with AEs in HepG2 cells, indicating their specificity to hepatoma carcinoma cells but not as hepatocyte-intrinsic CL-HCC AEs. The Hi-C data showed that the hepatocyte-intrinsic CL-HCC AEs were closely correlated with their associated genes in HepG2 cells (Figure 2B). Further analysis revealed that the hepatocyte-intrinsic CL-HCC AE-associated genes were significantly involved in cancer hallmarks (Figure 2C). The hepatocyte-intrinsic CL-HCC AE-associated genes were also enriched in cancer among all the disease classes and cancer development (Figure 2D).

In contrast, 231 CL-HCC AEs did not overlap with HepG2 cell AEs, which might be related to the non-tumor cells in the HCC TME; we refer to these AEs as TME-related CL-HCC AEs. GO biological process (BP) analysis of the TME-related CL-HCC AE-associated genes revealed enrichment of immune-related biological processes,
Landscape of active enhancers in HCC development

Figure 1. Identified CL-HCC AEs and their characteristics. (A) Identified CL-HCC AEs. First, common AEs were acquired from both patients with cirrhosis and those with HCC. Then, AE-associated genes that were upregulated in both CL and HCC were identified. Venn diagrams (left) depicting the unique and overlapping AEs in NL (blue), CL (yellow), and HCC (red). Scatter plot (right) shows the fold change of genes in CL and HCC compared with NL. AEs with a
fold change (FC) > 1.5 in both CL and HCC are shown (red plots). (B) Epigenetic characteristics of CL-HCC AEs. Line chart shows the average H3K27ac, H3K4me1, H3K27me3 and H3K4me3 signals in the CL-HCC AE region with 3 kb surrounding DNA in NL (blue), cirrhotic (yellow), and HCC (red) tissues. Heatmap showing the H3K27ac (orange), H3K4me1 (blue), H3K27me3 (green) and H3K4me3 (purple) signals in the CL-HCC AE region with 3 kb surrounding DNA in NL (top), CL (middle) and HCC (bottom) tissues. Box plots of the expression (C), methylation level (D), and 5 hmC level (E) of CL-HCC AE-associated genes in NL (blue), CL (yellow) and HCC (red) tissues based on different datasets. (F) KEGG (orange) and GO BP (light blue) enrichment analyses of the AE-associated genes. An unpaired Student’s t-test was used to assess the difference. NS, no significance P > 0.05; ***P < 0.001.

such as immune system differentiation, cellular response to chemokines, and adhesion (Figure 2E).

To further explore changes in the immune system from NL to HCC and the relationship of TME-related CL-HCC AE-associated genes, we next analyzed the difference in immune cell infiltration among NL, CL and HCC samples based on two different algorithms, EPIC [24] and MCPcounter [25], from two independent datasets. The results showed that CD8+ T cell infiltration was significantly increased in CL and HCC samples compared with NL samples, consistent with previous studies (Figure 2F) [43]. Further analysis of most of the TME-related CL-HCC AE-associated genes revealed a correlation to CD8+ T cell infiltration based on the GSE112221 dataset (Figure 2G). Most of these genes are co-expressed with PD-1, CTLA4 and TIM-3, which are immune checkpoints during CD8+ T cell exhaustion (Figure 2H).

The above findings demonstrated that CL-HCC AEs play an intricate and important role in hepatocarcinogenesis from CL to HCC. Some oncogenes might be activated by hepatocyte-intrinsic CL-HCC AEs during the CL stage and continue to promote robust transcription in HCC, initiating malignant transformation and sustaining cancer cell growth. In addition, TME-related CL-HCC AEs might participate in changes in the TME, especially those involved in CD8+ T cell infiltration and exhaustion. Overall, these combined effects of CL-HCC AEs ultimately participate in HCC oncogenesis.

Potential diagnostic ability of CL-HCC AE-associated genes for HCC

CL-HCC AE-associated genes were abnormally upregulated before hepatocarcinogenesis, and CL-HCC AEs were identified from HCC patients with a background of cirrhosis and were paired, suggesting their diagnostic capability in HCC. To assess the potential clinical utility of CL-HCC AE-associated genes, we first randomly divided the TCGA-LIHC samples into training and test sets. We next reduced the dimensions of the training set using the LASSO algorithm and then applied three machine learning algorithms, RF, SVM and LR [44], to evaluate their diagnostic ability based on AUC through five-fold cross-validation. Lastly, we examined the diagnostic ability of the best machine learning model with external independent data on HCC (Figure 3A). As expected, we found that the expression of CL-HCC AE-associated genes could accurately classify normal and HCC tissues (RF AUC = 0.883, SVM AUC = 0.930, LR AUC = 0.951, Figure 3B). Then, the best-performing algorithm, LR, which includes 5 genes (THBS4, OLFML2B, CDKN3, GABRE, and HDAC11), achieved a high AUC on the test set (0.998) (Figure 3C). The LR model based on the 5 CL-HCC AE-associated genes could predict the occurrence of HCC (Figures 3D-G).

These results suggest that this 5-gene LR model of CL-HCC AE-associated genes could act as a potential diagnostic biomarker of HCC.

MS identification based on CL-HCC AE-associated genes

To better understand the role of CL-HCC AE-associated genes in HCC heterogeneity, 370 HCC patients were clustered based on the expression of these genes by executing consensus NMF, and we identified 3 HCC MSs (Figure 4A) [32]. The mRNA expression levels of the CL-HCC AE-associated genes were significantly different among the different MSs, as follows: subclass MS1 showed the highest mRNA expression of the CL-HCC AE-associated genes and MS2 showed the lowest (Figure 4B). Furthermore, we analyzed the prognostic capacity of HCC based on the 3 HCC MSs, and a significant prognostic difference was observed in the OS (log-rank test P = 0.004) and recurrence-free survival (RFS) (log-rank test P = 0.01) (Figure 4C and 4D). Further
Landscape of active enhancers in HCC development

A. HepG2 AEs and CL-HCC AEs

B. chr4 87,900,000 - SPP1 88,000,000

C. Diagram showing various cancer-related activities:
   - Evading Immune Destruction*
   - Evading Growth Suppressors
   - Activating Invasion and Metastasis**
   - Inducing Angiogenesis***
   - Resisting Cell Death***
   - Sustaining Proliferative Signaling*
   - Enabling Replicative Immortality
   - Reprogramming Energy Metabolism
   - Tumor-Promoting Inflammation**
   - Genome Instability and Mutation

D. Network of cellular response chemokine, adhesion, and immune system differentiation

E. Graph showing CD8 (EPIC) levels for different samples (NL, CL, HCC) for GSE112221 and GSE54238.
Figure 2. Different roles of CL-HCC AE-associated genes from CL to HCC. A. Venn diagrams depicting the overlapping and nonoverlapping AEs in HepG2 cells (left, red) and CL-HCC AEs (right, blue). A total of 389 AEs overlapped with those in HepG2 cells, whereas 231 did not. B. Hi-C interaction frequency (normalized interaction counts) and genomic coordinates comprising a portion of a topological domain in HepG2 cells. C. Enrichment analysis of hepatocyte-intrinsic CL-HCC AE-associated genes and the 10 hallmarks of cancer. A hypergeometric distribution was used to calculate the enrichment degree. *P < 0.05, **P < 0.01, ***P < 0.001, gray color indicates no significance. D. Enrichment analysis of hepatocyte-intrinsic CL-HCC AEs in the disease class. E. Enrichment map for the GO BP enrichment results of the TME-related CL-HCC AE-associated genes. Each node represents a GO BP term. F. CD8 T cell infiltration in NL (blue), CL (yellow) and HCC (red) tissues in the GSE112221 (left) and GSE54238 (right) datasets based on EPIC (top) and MCPcounter (bottom). G. Correlation of the expression of TME-related CL-HCC AE-associated genes and CD8 T cell infiltration in the GSE112221 dataset. H. Correlation of the expression of TME-related CL-HCC AE-associated genes and CD8 T cell exhaustion markers (PD-1, CTLA4 and TIM-3). Genes with a correlation coefficient > 0.7 are shown in the figure.
Landscape of active enhancers in HCC development

Figure 3. Potential diagnostic ability of CL-HCC AE-associated genes in HCC and other cancers. (A) Flowchart describing the schematic overview of the design. First, the TCGA-LIHC dataset (n = 424) was divided into training (n = 296) and test (n = 128) sets. Second, the LASSO algorithm was used to reduce the dimensions of the training set. Then, classifiers were built through five-fold cross-validation within the training set. The best-performing classifier (LR model) was used to examine the test set, external independent HCC datasets. (B) ROC curves of the training set based on three machine learning algorithms. ROC curves of the test set (C), ICGC (D), GSE124535 (E), GSE77509 (F), and GSE94660 (G).

comparisons among the three groups revealed significantly different OS and RFS outcomes among MS1-MS2 and MS3-MS2, but there was no significant difference in the outcomes between MS1 and MS3 (Figure 4E-J).

DNA copy number and mutation spectrum among different MSs

To further identify the differences between the above-established MSs, we determined the numbers of CNVs (i.e., gains and losses) for
Landscape of active enhancers in HCC development

Figure 4. MSs based on the expression of CL-HCC AE-associated genes. A. NMF-based clustering. B. Box plot of the expression of CL-HCC AEs in the three MSs. C. Kaplan-Meier overall survival curve of the three MSs. D. Kaplan-Meier recurrence-free survival curve of the three MSs. E. Kaplan-Meier overall survival curve for MS1 and MS2. F. Kaplan-Meier overall survival curve for MS2 and MS3. G. Kaplan-Meier overall survival curve for MS1 and MS3. H. Kaplan-Meier recurrence-free survival curve for MS1 and MS2. I. Kaplan-Meier recurrence-free survival curve for MS2 and MS3. J. Kaplan-Meier recurrence-free survival curve for MS1 and MS3. Different colors represent different MSs: MS1 (red), MS2 (green) and MS3 (blue). One-way ANOVA was used to analyze the differences. ***$P < 0.001$.

each sample (Figure 5A) and found significant differences among the three subtypes. MS1 and MS3 had more CNV gains and losses than MS2 by the chi-square test (Figure 5B). The mutation spectrum of the top 20 genes with the highest mutation rate in various MSs showed significant differences by the chi-square test ($P < 0.0001$) (Figure 5C). Mutation spectrum analysis showed a significantly lower mutation frequency of TP53 in MS2, but TTN, CTNNB1, MUC16 and PCLO showed a higher mutation frequency in MS2 than in MS1 and MS3 (Figure 5D).

Previous studies have revealed that DNA mutations are significantly associated with both pre-
dicted TMB and neoantigens, which has emerged as a promising biomarker for the immunotherapy response in cancer patients and are capable of inducing tumor-specific T cell recognition, respectively [45-47]. In the current study, the data revealed that both predicted TMB and neoantigens were higher in MS2 than in MS1 (Figure 5E and 5F).

All these data suggest that the MSs have different CNVs and mutation landscapes and may have different responses to immunotherapy.

Characterization of the MSs involved in different functional pathways and TME constituents

We next characterized the functional pathway differences among the 3 MSs by ssGSEA. ssGSEA of the KEGG pathways showed that many pathways related to the immune response and metabolism were different among the above 3 MSs (Figure 6A). Alterations in numerous immune response-associated pathways prompted us to then explore differences in the TME constituents among the 3 MSs by characterizing the ratio of various immune cell subsets as well as stromal-related cell subtypes in the TME [25]. The data showed that 10 immune- and stromal-related cell subtypes, including T cells, CD8 T cells, cytotoxic lymphocytes, NK cells, B lineage cells, monocyte lineage cells, myeloid dendritic cells, neutrophils, endothelial cells, and fibroblasts, had more infiltration in MS1 than in MS2 and MS3 (Figure 6B). Although many immune cells infiltrated in MS1 and MS3, MS1 and MS3 were associated with a higher hypoxia state of the TME than MS2, which often limits anticancer immunity (Figure 6C). Predictably, many immune cells will infiltrate in MS1 and MS3, but more T cells may become dysfunctional in MS1 and MS3 (Figure 6D).

Differential immunotherapy and putative drug responses according to MS

Differences in the TMB, neoantigens and TME constituents of the subtypes prompted us to further investigate differences in the TME constituents and the likelihood of responding to immuno-oncology therapy [48]. The results showed higher sensitivity to immunotherapies for MS2 than MS1 and MS3 (Figure 7A) [35]. For further research into precision therapy for HCC patients, we evaluated the putative drug responses [49]. As chemotherapy and molecular targeted therapy are important steps in the comprehensive treatment of HCC patients, we aimed to assess the response of the 3 MSs to three classic chemotherapy drugs (docetaxel, paclitaxel and cisplatin) and three classic molecular targeted drugs (gefitinib, cytarabine and bortezomib) [50]. We observed a significant difference in the estimated IC50 values among the 3 MSs for the 3 chemotherapy drugs and 3 molecular targeted drugs (Figure 7B and 7C). MS3 was more sensitive to commonly administered chemotherapies and molecular targeted therapies, except for bortezomib. MS1 showed decreased sensitivity to chemotherapy treatment but increased sensitivity to molecular targeted therapies.

These results suggest that the MSs established by the CL-HCC AE-associated genes have different responses to immunotherapies, chemotherapy drugs and molecular targeted drugs, which could be helpful for precision therapy and immunotherapy for HCC patients.

The mRNA levels of hepatocyte-intrinsic CL-HCC AE-associated genes were significantly downregulated in HepG2 cells after treatment with the BET bromodomain inhibitor JQ1

According to the above data, the high expression of CL-HCC AE-associated genes tends to play a role in promoting cancer and suppressing immunity. To identify drugs that could downregulate the expression of these CL-HCC AE-associated genes, we further analyzed the mRNA expression levels of the hepatocyte-intrinsic CL-HCC AE-associated genes in HepG2 cells treated with 0.5 µM JQ1, a BET bromodomain inhibitor that could reduce the frequency of AEs and the expression of their associated genes [51, 52]. GSEA showed that these genes were significantly downregulated after JQ1 treatment (Figure 8).

Discussion

Epigenetic alterations in AEs and their malfunction have been recognized as driving causes of tumorigenesis and progression [53, 54]. Previous studies have shown that the de novo acquisition of AEs correlates with cancer-related pathways and is associated with prognosis [4]. In the current study, by performing an integrated analysis of the histone posttranslational
Landscape of active enhancers in HCC development
Figure 5. CNV and mutation spectrum of the MSs. (A) Composite copy number profiles for MS1, MS2 and MS3, with gains in red and losses in blue. (B) CNV distribution of all genes among the three MS; each color indicates a different type of CNV. (C) Profiles of the top 20 significant mutations across the MSs. (D) Distribution pattern of the top 20 significant mutations among the three MSs. Comparison of TMB (E) and predicted neoantigens (F) among the three MSs. Different colors represent different MSs: MS1 (red), MS2 (green) and MS3 (blue). The chi-square test was used to analyze the differences in CNVs and mutations among the MSs. An unpaired Student’s t-test was used to assess the differences in TMB, and the Wilcoxon test was used to assess the differences in predicted neoantigens. *P < 0.05, **P < 0.01, ***P < 0.001.
Landscape of active enhancers in HCC development

A

Immune system

- KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION
- KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY
- KEGG_CHEMOKINE_SIGNALING_PATHWAY
- KEGG_COMPLEMENT_AND_COAGULATION_CASCADES
- KEGG_FC_EPSILON_RI_SIGNALING_PATHWAY
- KEGG_HEMATOPOIETIC_CELL_LINEAGE
- KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION
- KEGG_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION
- KEGG_NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY
- KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY

Metabolism

- KEGG_BUTANOATE_METABOLISM
- KEGG_FATTY_ACID_METABOLISM
- KEGG_LIMONENE_AND_PINENE_DEGRADATION
- KEGG_METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450
- KEGG_PRIMARY_BILE_ACID_BIOSYNTHESIS
- KEGG RETINOL_METABOLISM
- KEGG_RIBOFLAVIN_METABOLISM
- KEGG_STEROID_HORMONE_BIOSYNTHESIS
- KEGG_TRYPTOPHAN_METABOLISM
- KEGG_TYROSINE_METABOLISM

B

- MS1
- MS2
- MS3

- B lineage
- CD8 T cells
- Cytotoxic lymphocytes
- Endothelial cells
- Fibroblasts
- Monocytic lineage
- Myeloid dendritic cells
- Neutrophils
- NK cells
- T cells
Figure 6. Characterization and tumor microenvironment of the MSs. (A) Heatmap of the enrichment level calculated by single sample gene set enrichment analysis for the top 10 significantly different KEGG immune-related pathways and metabolism-related pathways among the three MSs derived from GSEA. (B) Box plot shows different immune cell enrichment levels among the three MSs. Comparison of hypoxia scores (C) and T cell dysfunction (D) (one-way ANOVA) among the three MSs. Different colors represent different MSs: MS1 (red), MS2 (green) and MS3 (blue). One-way ANOVA was used to analyze the differences. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. 
Landscape of active enhancers in HCC development

A

Percent of Response (%)

MS1  MS2  MS3

0.00  0.25  0.50  0.75  1.00

Response  No response

B

Estimated IC50

Docetaxel  Paclitaxel  Cisplatin

C

Estimated IC50

Cytarabine  Bortezomib  Gefitinib

MS1  MS2  MS3

Red  Green  Blue

***
Landscape of active enhancers in HCC development

Figure 7. Immunotherapy response and putative drug response of the MSs. Comparison of the predicted response to immunotherapy (A) (Fisher’s exact test), classic chemotherapy drugs (B) and classic molecular targeted drugs (C). One-way ANOVA was used to analyze the differences. *P < 0.05, **P < 0.01, ***P < 0.001.

Figure 8. GSEA of hepatocyte-intrinsic CL-HCC AE-associated genes in HepG2 cells treated with the BET inhibitor JQ1. GSEA plots, normalized enrichment scores (NESs) and p values are shown for the hepatocyte-intrinsic CL-HCC AE-associated gene sets in HepG2 cells treated with JQ1.

modifications of H3K27ac and H3K4me1 from NL to CL and HCC. We systemically studied the function and potential clinical effects of CL-HCC AEs that are absent in normal liver tissue and are de novo acquired in cirrhosis and sustained during HCC development. By comparing with CL-HCC AEs and HepG2 AEs, we divided the CL-HCC AEs into hepatocyte-intrinsic CL-HCC AEs and TME-related CL-HCC AEs. We believe that CL-HCC AEs can direct the activation of their associated genes and contribute to HCC development by regulating both hepatocyte-intrinsictumorogenesis and the tumorimmuneresponse. ROC curve analysis and MS identification demonstrated that the CL-HCC AE-associated genes show potential diagnostic and precision therapy power for HCC.

HCC patients are often diagnosed at advanced stages because of the lack of sufficient diagnos-
Landscape of active enhancers in HCC development

except for bortezomib. Therefore, our data suggest that MS identification based on CL-HCC AE-associated genes could be used to guide the therapeutic strategy (i.e., chemotherapy, molecular targeted therapy and immunotherapy) for HCC patients.

Immunotherapy is considered a promising approach for HCC patients; however, only a small proportion of HCC patients benefit from it [64, 65]. Therefore, it is urgent to develop a more accurate MS identification to determine the efficiency prediction of immunotherapy [66, 67]. A previous study showed that both H3K4me1 and H3K27ac, marked AEs, increase PD-1 expression through the combined stimulation of TCR and cytokines [67]. Our data showed that many CL-HCC AE-associated genes are positively correlated with PD-1 expression; predictably, the high expression of these genes was determined to be involved in CD8+ T cell dysfunction in HCC. In our study, although the TME of MS1 patients contained abundant CD8+ T cells and NK cells, the simultaneous presence of a large number of immunosuppressive cells and a hypoxic status promoted an immunosuppressive microenvironment [68]. The combination of these effects ultimately led to a poor response to immunotherapy in MS1 patients. MS2 patients had lower expression of the CL-HCC AE-associated genes and showed higher sensitivity to immunotherapy. This might be because the TME of MS2 patients predicts higher neoantigen exposure that could induce better antitumor immune responses, better T cell function, and less inflammation [69]. As such, our data provide another option for predicting the immune therapy response of HCC patients based on CL-HCC AE-associated gene-directed MS identification.

Bromodomain containing 4 (BRD4) is enriched in AEs and controls the induction of gene expression [51, 70]. JQ1 is a noteworthy anticancer drug that inhibits BET bromodomains and reduces the development of AEs as well as the expression of their associated genes [71, 72]. Ronald M et al. found that JQ1 can protect and reverse the fibrotic response in carbon tetrachloride-induced fibrosis in mouse models [73]. JQ1 can also enhance both T cell persistence and function when combined with immunotherapy [74]. However, the mechanistic linkage of JQ1 and the immunotherapy response is still not clear. We found that CL-HCC AE-associated genes were significantly down-regulated in JQ1-treated HepG2 cells. Therefore, our data provide evidence for the potential of applying JQ1 for HCC treatment, which might show dual roles in both the direct inhibition of cancer cell growth and an improvement in the immunotherapy response.

In conclusion, we comprehensively explored CL-HCC AEs and their target genes in HCC. We demonstrate that alterations in CL-HCC AEs and their target genes play critical roles in HCC initiation, development and the therapeutic response. CL-HCC AE-associated genes could be specific for identifying biomarkers, classification and therapeutic predictions of HCC. In general, we provide new perspectives that CL-HCC AE-associated genes can distinguish the different landscapes of HCC and help to explore the mechanism, classification and precision therapy of HCC.

Acknowledgements

This study was supported by grants from the National Natural Science Foundation of China (No. 81520108029 to X.L.; No. 81922068 and 81874313 to Y.C.D.; No. 81703521 to Y.F.D.) and the Natural Science Foundation of Tibet (XZ2019ZRG-135 to QH).

Disclosure of conflict of interest

None.

Address correspondence to: Youcai Deng and Xiaohui Li, Institute of Materia Medica, College of Pharmacy, Army Medical University (Third Military Medical University), 30# Gaotanyan Road, Shapingba District, Chongqing 400038, China. E-mail: youcai.deng@tmmu.edu.cn (YCD); lps008@aliyun.com (XHL); Zhaoyang Zhong, Cancer Center, Daping Hospital and Research Institute of Surgery, Army Medical University (Third Military Medical University), 10# Daping Changjiangzhilu Road, Yuzhong District, Chongqing 400042, China. E-mail: zhongzhaoyang08@hotmail.com

References


Landscape of active enhancers in HCC development


Landscape of active enhancers in HCC development


Landscape of active enhancers in HCC development


