Am J Cancer Res 2020;10(10):3157-3178
www.ajcr.us /ISSN:2156-6976/ajcr0119638

Original Article

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

Yao Yang[1], Xiaoyu Deng[1], Xinjian Chen[2], Shihan Chen[3], Liang Song[1], Meng Meng[1], Qi Han[4], Saber Imani[5], Shuhui Li[6], Zhaoyang Zhong[7], Xiaohui Li[1], Youcai Deng[1]

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

Table 1. Characteristics of the GEO datasets

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

mechanisms of HCC development and shed light on the identification of druggable epigenetic targets for the prevention and treatment of HCC.

A newly reported epigenetic modification referred to as active enhancers (AEs), marked by posttranslational modifications of both H3K27ac and H3K4me1, are defined as gene-distal 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

(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, ENO1, 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

the Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org/) training set [36].

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: GSE112221 [4] and 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 indispensable 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,

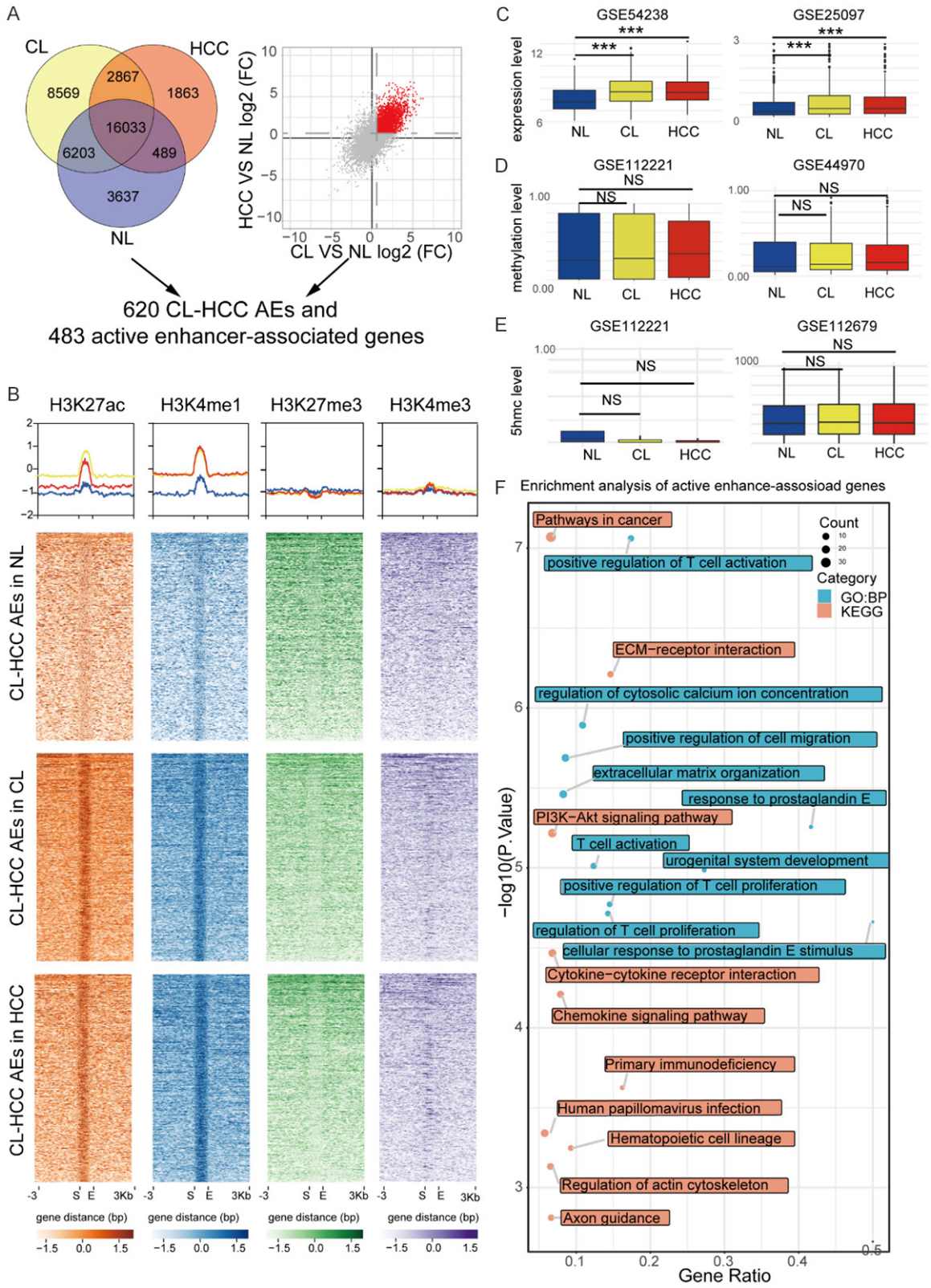


Figure 1. Identified CL-HCC AEs and their characteristics. (A) Identified CL-HCC AEs. First, common AEs were acquired from both patients with cirrhosis and those with HCC. Then, AE-associated genes that were upregulated in both CL and HCC were identified. Venn diagrams (left) depicting the unique and 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.

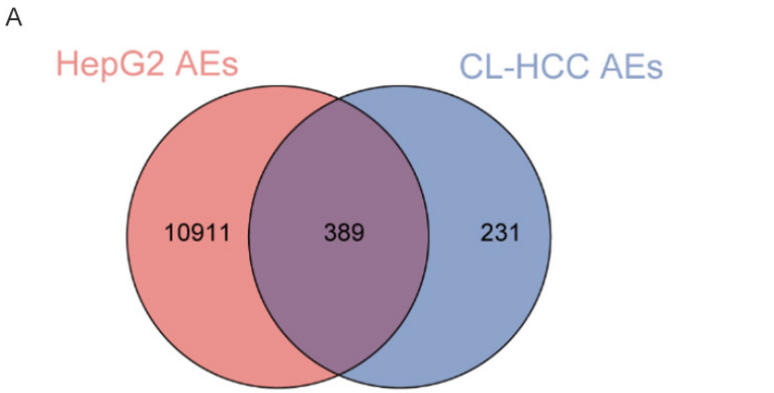
Potential diagnostic ability of CL-HCC AE-associated genes for HCC

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**).

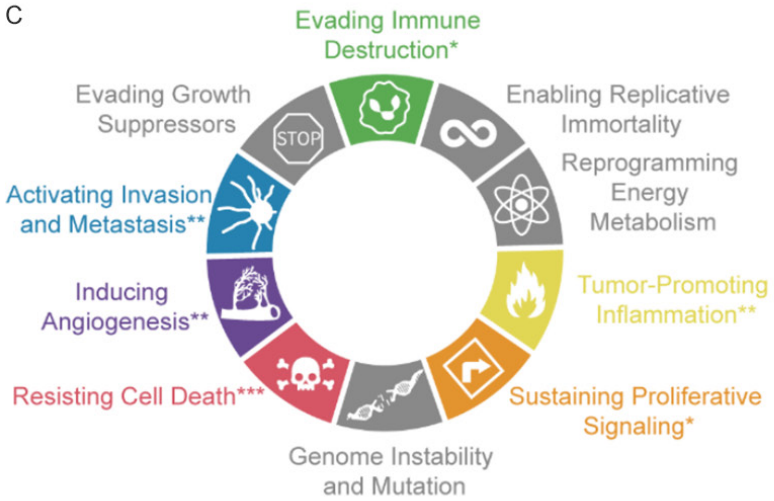
These results suggest that this 5-gene LR model of CL-HCC AE-associated genes could act as a potential diagnostic biomarker of HCC.

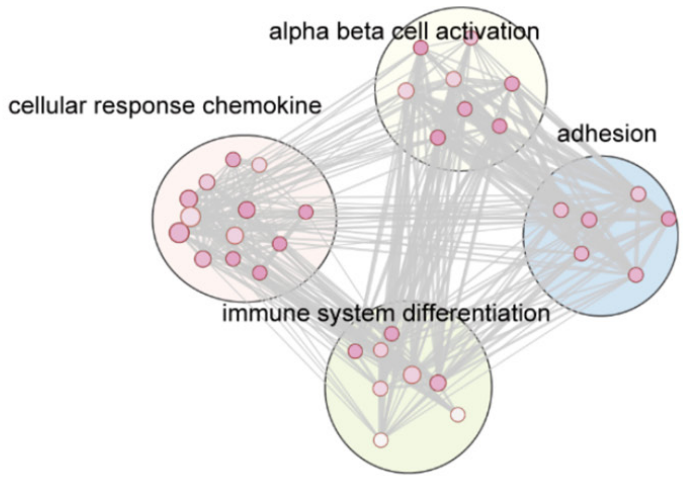
MS identification based on CL-HCC AE-associated genes

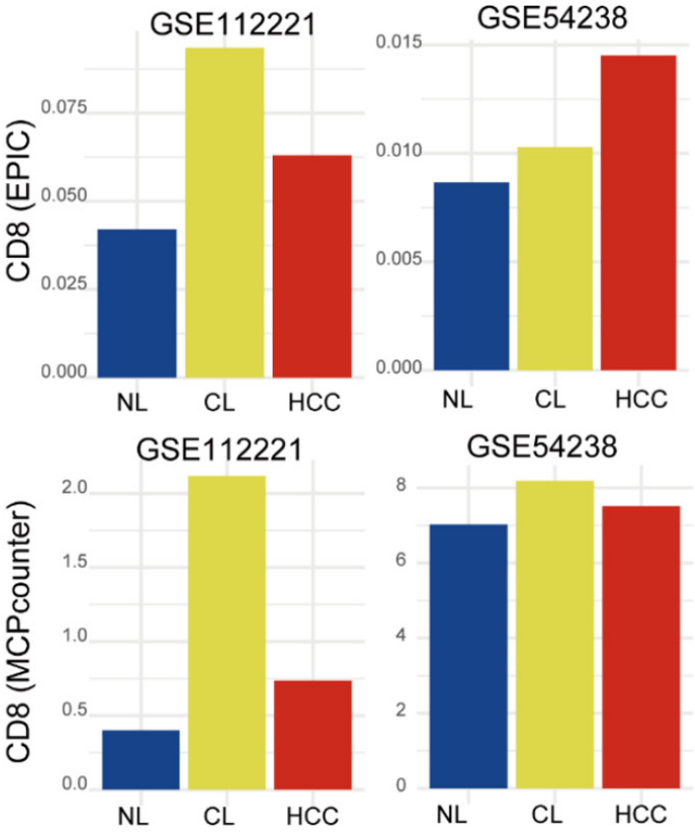
To better understand the role of CL-HCC AE-associated genes in HCC heterogeneity, 370 HCC patients were clustered based on the expression of these genes by executing consensus NMF, and we identified 3 HCC MSs (**Figure 4A**) [32]. The mRNA expression levels of the CL-HCC AE-associated genes were significantly 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


A
HepG2 AEs
CL-HCC AEs
10911
389
231


B
387
11
chr4 87,900,000
88,000,000
SPP1


C
Evading Immune Destruction*
Enabling Replicative Immortality
Reprogramming Energy Metabolism
Tumor-Promoting Inflammation**
Sustaining Proliferative Signaling*
Genome Instability and Mutation
Resisting Cell Death***
Inducing Angiogenesis**
Activating Invasion and Metastasis**
Evading Growth Suppressors
STOP


E
alpha beta cell activation
cellular response chemokine
adhesion
immune system differentiation


F
GSE112221
GSE54238
CD8 (EPIC)
0.075
0.050
0.025
0.000
0.015
0.010
0.005
NL
CL
HCC
CD8 (MCPcounter)
2.0
1.5
1.0
0.5
0.0
8
6
4
2
0

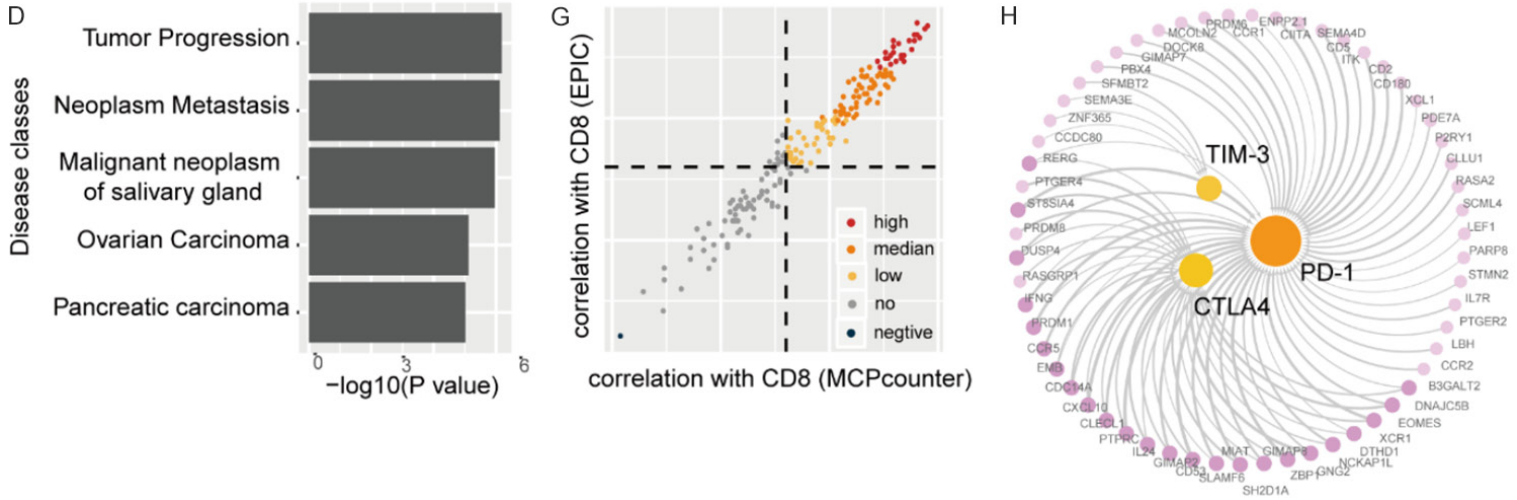


Figure 2. Different roles of CL-HCC AE-associated genes from CL to HCC. A. Venn diagrams depicting the overlapping and nonoverlapping AEs in HepG2 cells (left, red) and CL-HCC AEs (right, blue). A total of 389 AEs overlapped with those in HepG2 cells, whereas 231 did not. B. Hi-C interaction frequency (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.

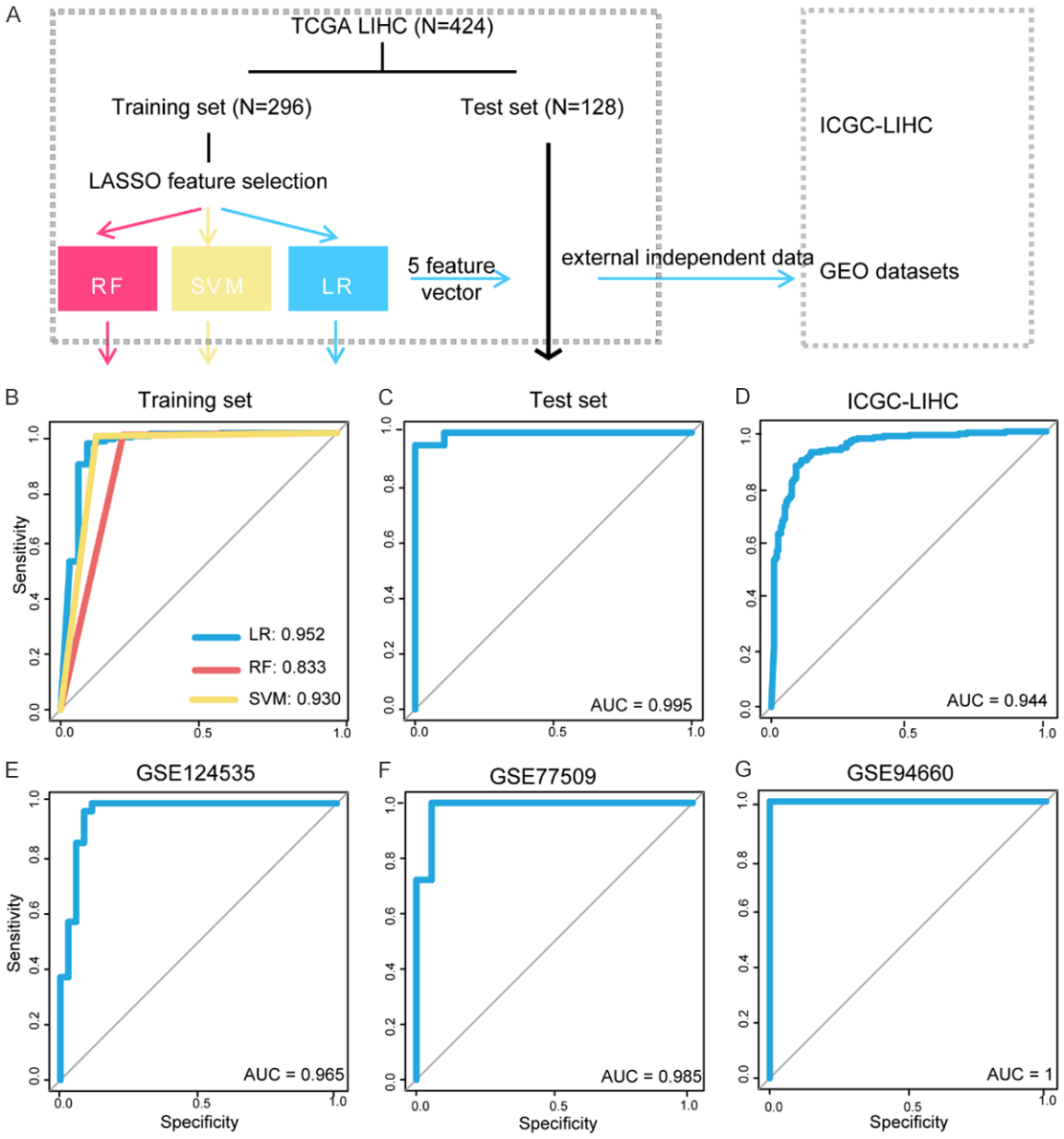


Figure 3. Potential diagnostic ability of CL-HCC AE-associated genes in HCC and other cancers. (A) Flowchart describing the schematic overview of the design. First, the TCGA-LIHC dataset (n = 424) was divided into training (n = 296) and test (n = 128) sets. Second, the LASSO algorithm was used to reduce 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 GSE94600 (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

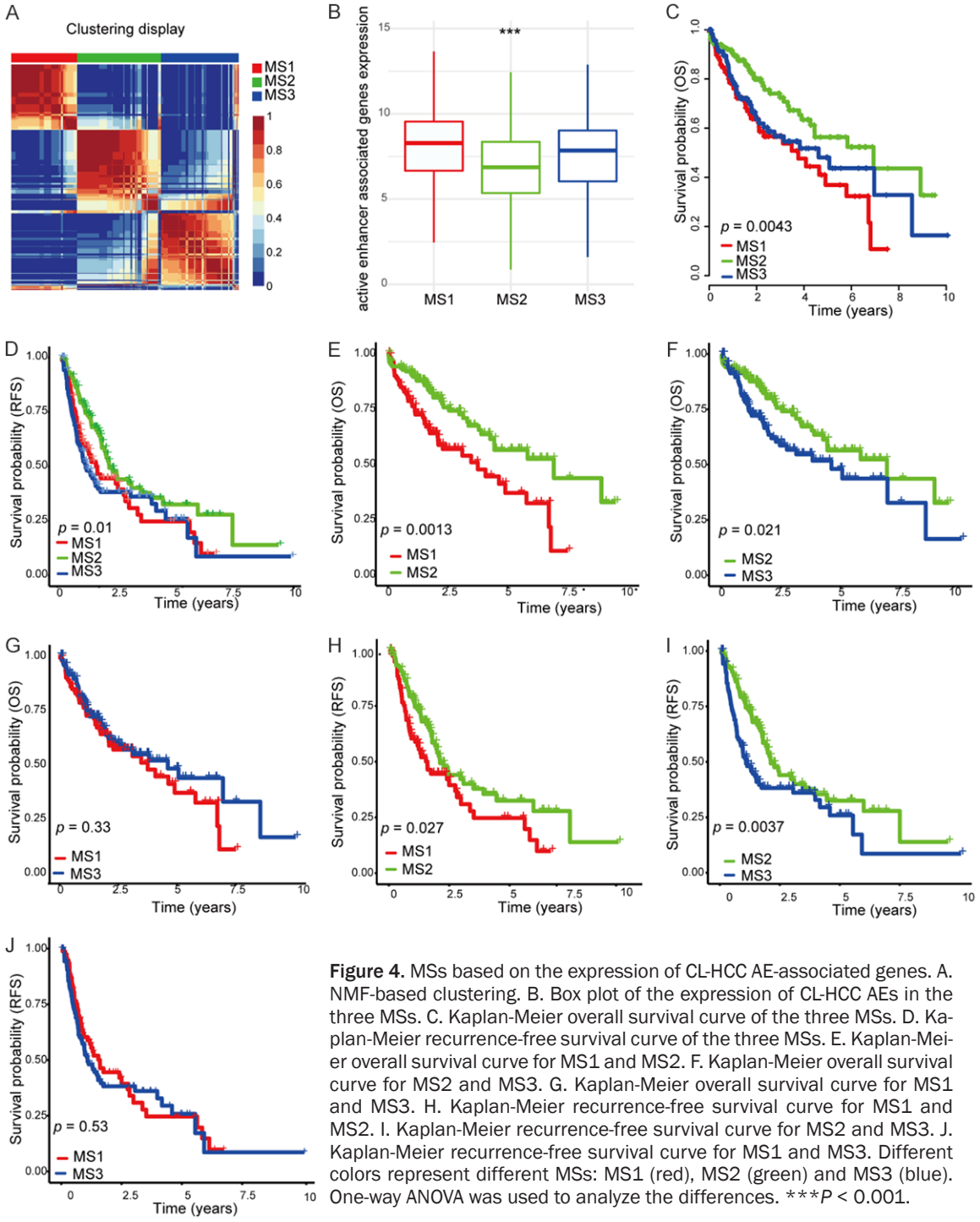


Figure 4. MSs based on the expression of CL-HCC AE-associated genes. A. NMF-based clustering. B. Box plot of the expression of CL-HCC AEs in the three MSs. C. Kaplan-Meier overall survival curve of the three MSs. D. Kaplan-Meier recurrence-free survival curve of the three MSs. E. Kaplan-Meier overall 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.

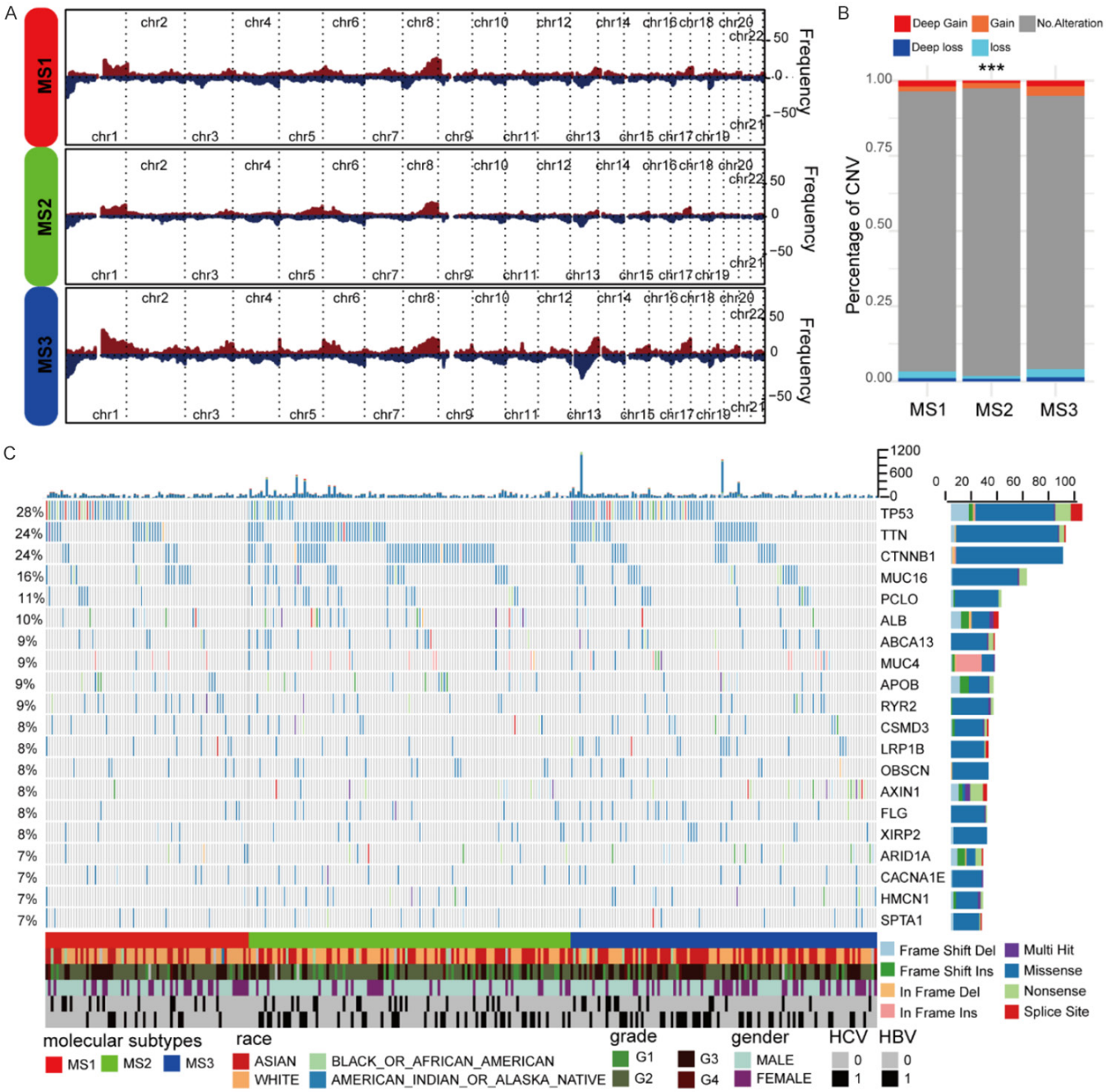
These results suggest that the MSs established by the CL-HCC AE-associated genes have different responses to immunotherapies, chemotherapy drugs and molecular targeted drugs, which could be helpful for precision therapy and immunotherapy for HCC patients.

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


A
MS1
MS2
MS3
Frequency
B
Deep Gain
Gain
No.Alteration
Deep loss
loss
***
Percentage of CNV
MS1
MS2
MS3
C
TP53
TTN
CTNNB1
MUC16
PCLO
ALB
ABCA13
MUC4
APOB
RYR2
CSMD3
LRP1B
OBSCN
AXIN1
FLG
XIRP2
ARID1A
CACNA1E
HMCN1
SPTA1
Frame Shift Del
Frame Shift Ins
In Frame Del
In Frame Ins
Multi Hit
Missense
Nonsense
Splice Site
molecular subtypes
MS1
MS2
MS3
race
ASIAN
BLACK_OR_AFRICAN_AMERICAN
WHITE
AMERICAN_INDIAN_OR_ALASKA_NATIVE
grade
G1
G2
G3
G4
gender
MALE
FEMALE
HCV
HBV

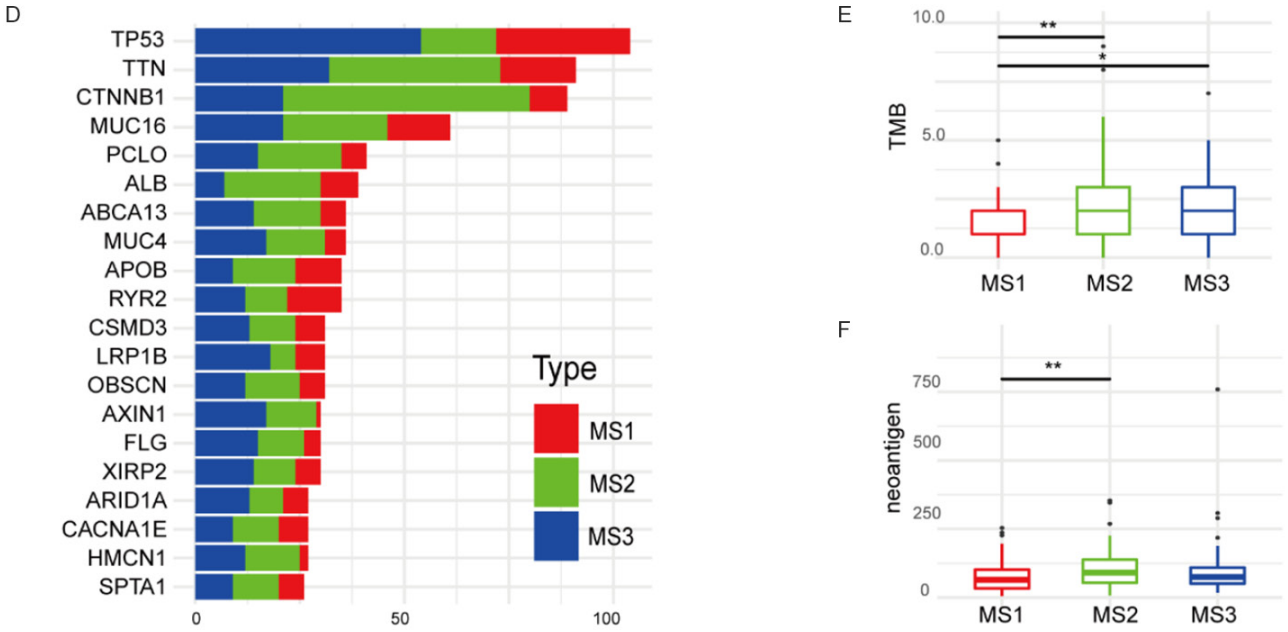
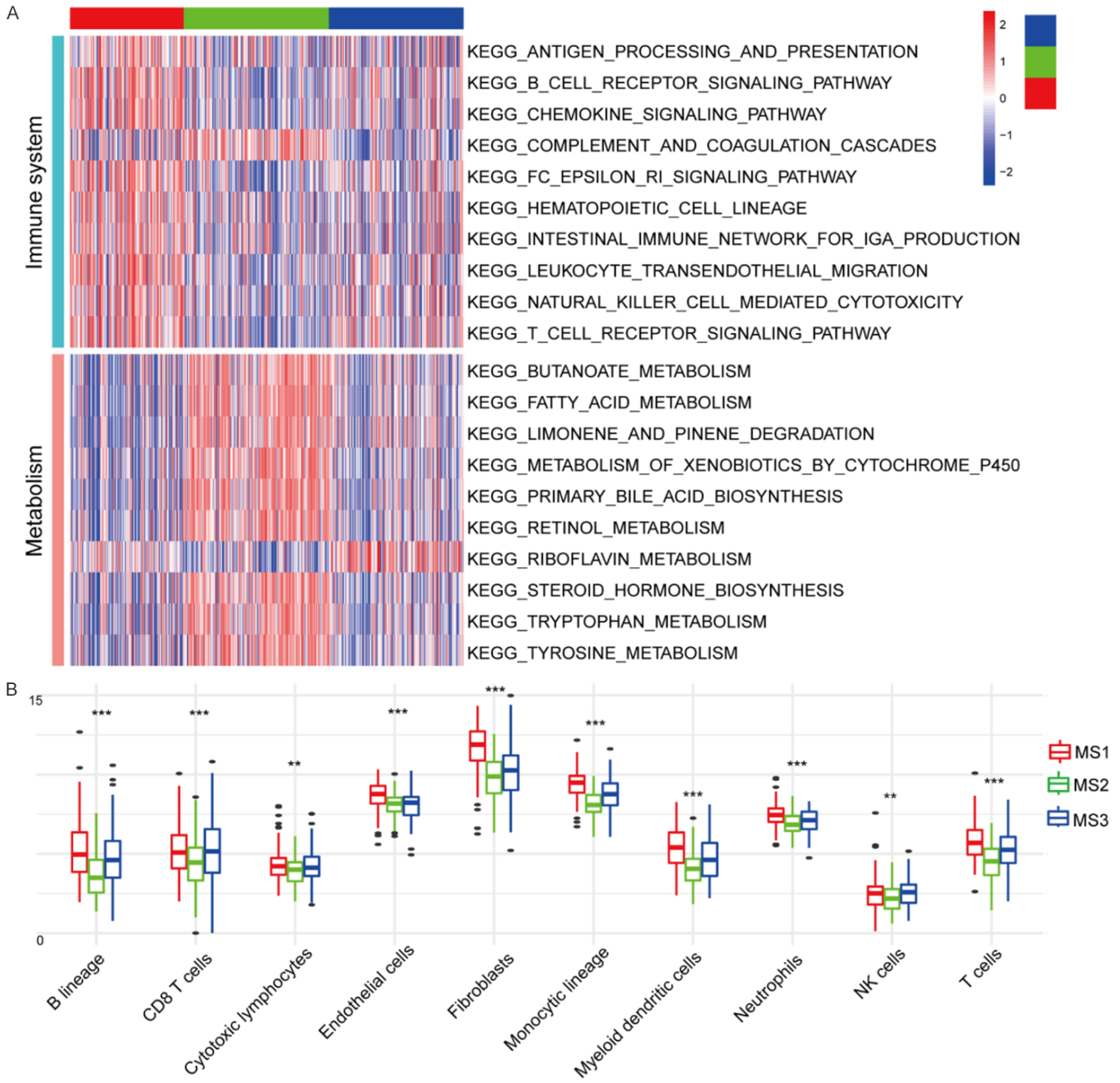


Figure 5. CNV and mutation spectrum of the MSs. (A) Composite copy number profiles for MS1, MS2 and MS3, with gains in red and losses in blue. (B) CNV distribution of all genes among the three MS; each color indicates a different type of CNV. (C) Profiles of the top 20 significant mutations across the 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



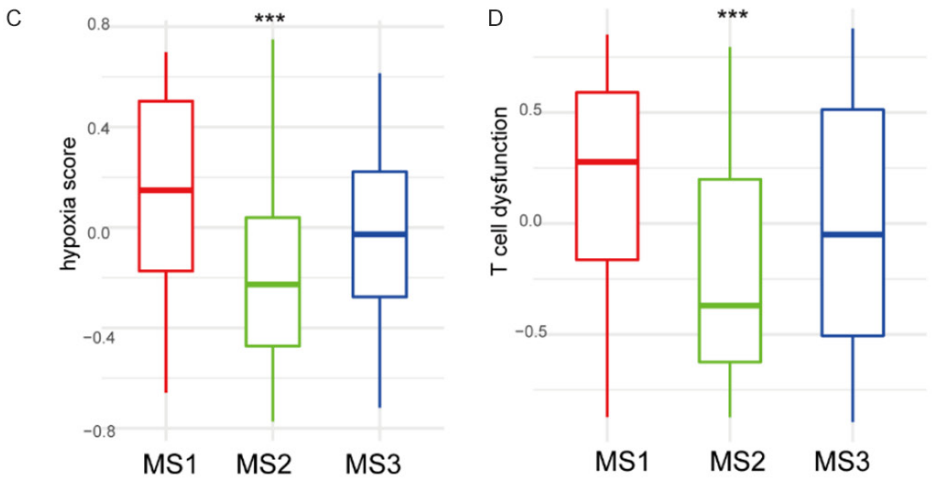


Figure 6. Characterization and tumor microenvironment of the MSs. (A) Heatmap of the enrichment level calculated by single sample gene set enrichment analysis for the top 10 significantly different KEGG immune-related pathways and metabolism-related pathways among the three MSs derived from GSEA. (B) 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. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

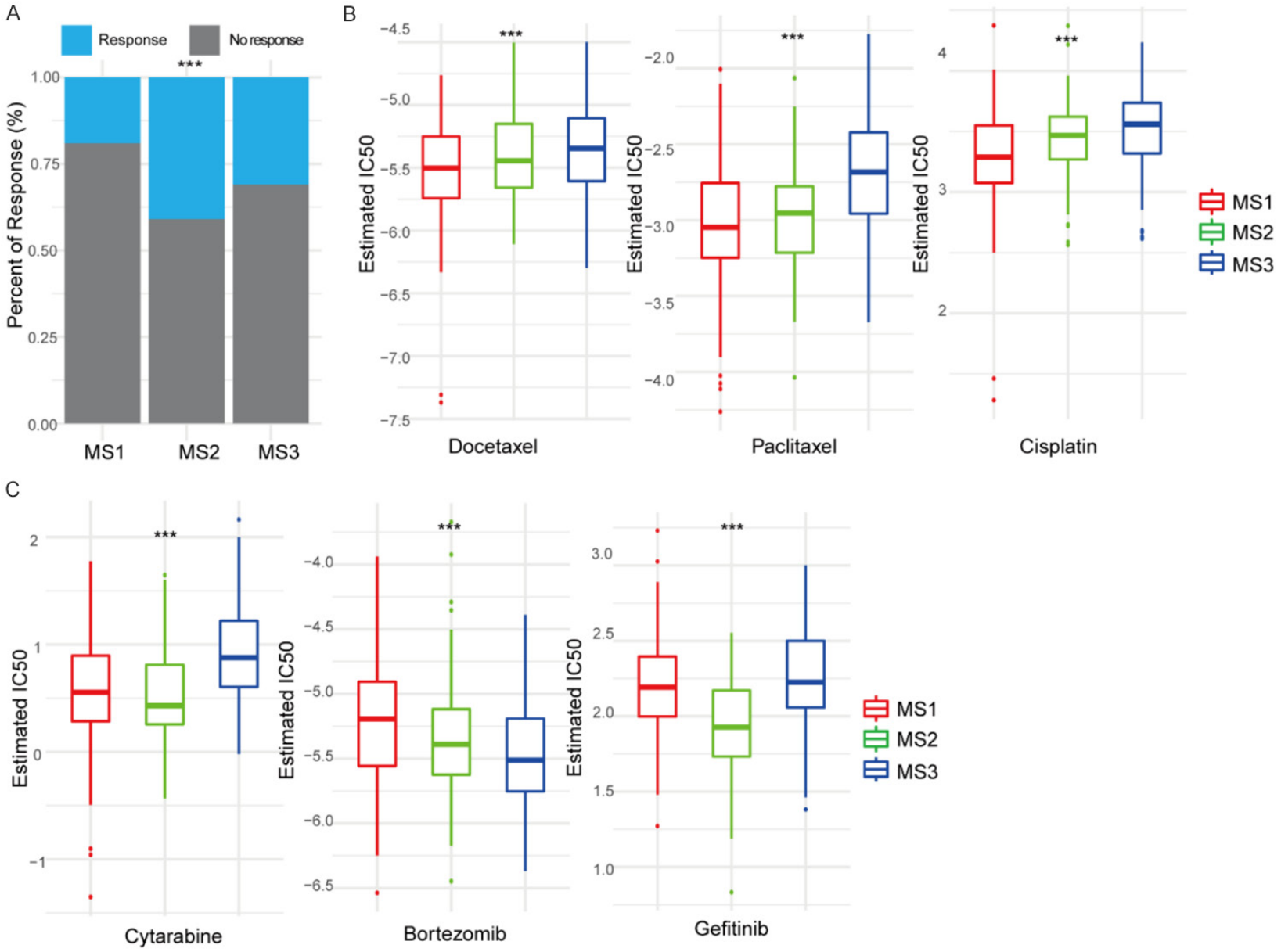

A
Response
No response
***
Percent of Response (%)
1.00
0.75
0.50
0.25
0.00
MS1
MS2
MS3
B
Estimated IC50
Docetaxel
Paclitaxel
Cisplatin
MS1
MS2
MS3
C
Cytarabine
Bortezomib
Gefitinib

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.

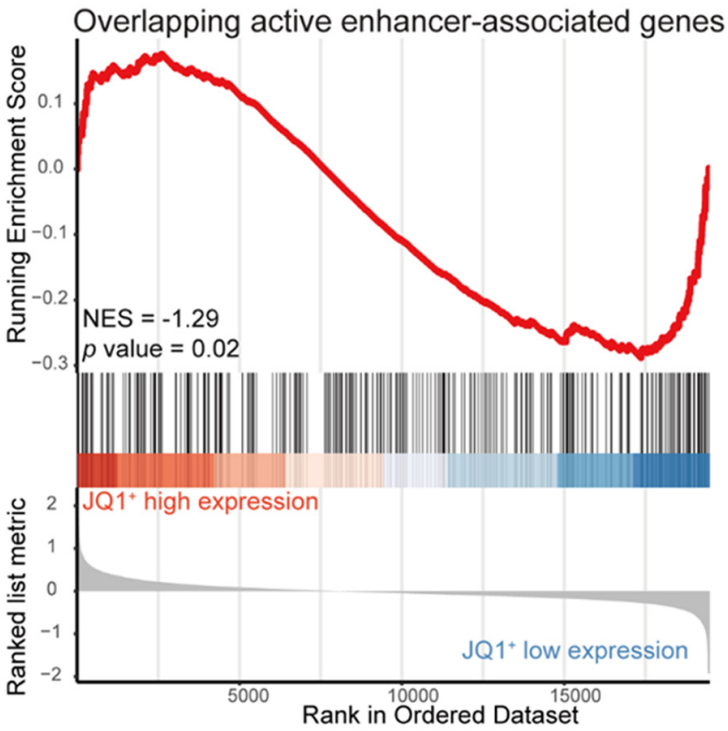


Figure 8. GSEA of hepatocyte-intrinsic CL-HCC AE-associated genes in HepG2 cells treated with the BET inhibitor JQ1. GSEA plots, normalized enrichment scores (NESs) and *p* values are shown for the hepatocyte-intrinsic CL-HCC AE-associated gene sets in HepG2 cells treated with JQ1.

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-intrinsic tumorigenesis and the tumor immune response. 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 diagnostic biomarkers [55]. By using CL-HCC AE-associated gene expression profiles, we established a five-gene LR model that showed reliable diagnostic capability for HCC, with high sensitivity and specificity. All five genes, namely, THBS4, OLFML2B, CDKN3, GABRE and HDAC11, have been reported to be independent diagnostic biomarkers for cancer [56-60]. Therefore, CL-HCC AE-associated genes could be used to search for new diagnostic biomarkers to provide a new strategy for HCC diagnosis and to forecast the incidence of HCC at the stage of cirrhosis.

Although many studies on HCC subtype classifications based on gene expression or epigenetics have been proposed in recent years, there has not yet been a consensus [61-63]. In this study, we identified HCC MSs based on CL-HCC AE-associated genes. We identified three MSs of HCC (MS1, MS2, and MS3) and then explored their differences in prognosis, CNV and mutation, functional pathways, TME cell subtypes, immunotherapy responses and putative drug responses. The results showed that MS1 displayed the highest expression of the CL-HCC AE-associated genes, increased hypoxia and maximum immune cell infiltration but serious T cell dysfunction and a poor prognosis. MS2 showed the lowest expression of the CL-HCC AE-associated genes, the lowest inflammatory conditions and preserved the default metabolic program of NL. This subtype showed good immunotherapeutic sensitivity and the best prognosis. MS3, with median expression levels of the CL-HCC AE-associated genes, had most features falling somewhere between those of MS1 and MS2. MS3 showed a worse prognosis than MS2 but not MS1. However, MS3 was highly sensitive to all chemotherapy and molecular targeted drugs

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 downregulated 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); lpsh008@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

[1] Balogh J, Victor D 3rd, Asham EH, Burroughs SG, Boktour M, Saharia A, Li X, Ghobrial RM and Monsour HP Jr. Hepatocellular carcinoma: a review. J Hepatocell Carcinoma 2016; 3: 41-53.

[2] Forner A, Reig M and Bruix J. Hepatocellular carcinoma. Lancet 2018; 391: 1301-1314.

[3] Tian Y and Ou JH. Genetic and epigenetic alterations in hepatitis B virus-associated hepatocellular carcinoma. Virol Sin 2015; 30: 85-91.
[4] Hlady RA, Sathyanarayan A, Thompson JJ, Zhou D, Wu Q, Pham K, Lee JH, Liu C and Robertson KD. Integrating the epigenome to identify drivers of hepatocellular carcinoma. Hepatology 2019; 69: 639-652.
[5] Nishida N and Kudo M. Alteration of epigenetic profile in human hepatocellular carcinoma and its clinical implications. Liver Cancer 2014; 3: 417-427.
[6] Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA and Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature 2011; 470: 279-283.
[7] Kimura H. Histone modifications for human epigenome analysis. J Hum Genet 2013; 58: 439-445.
[8] Murakawa Y, Yoshihara M, Kawaji H, Nishikawa M, Zayed H, Suzuki H, Fantom C and Hayashizaki Y. Enhanced identification of transcriptional enhancers provides mechanistic insights into diseases. Trends Genet 2016; 32: 76-88.
[9] Chen H, Li C, Peng X, Zhou Z and Weinstein JN; Cancer Genome Atlas Research Network, Liang H. A pan-cancer analysis of enhancer expression in nearly 9000 patient samples. Cell 2018; 173: 386-399, e312.
[10] Patten DK, Corleone G, Gyorffy B, Perone Y, Slaven N, Barozzi I, Erdos E, Saiakhova A, Goddard K, Vingiani A, Shousha S, Pongor LS, Hadjiminas DJ, Schiavon G, Barry P, Palmieri C, Coombes RC, Scacheri P, Pruneri G and Magnani L. Enhancer mapping uncovers phenotypic heterogeneity and evolution in patients with luminal breast cancer. Nat Med 2018; 24: 1469-1480.
[11] Goldman MJ, Craft B, Hastie M, Repecka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, Zhu J and Haussler D. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol 2020; 38: 675-678.
[12] Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G and Noushmehr H. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res 2016; 44: e71.
[13] Zang C, Schones DE, Zeng C, Cui K, Zhao K and Peng W. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics 2009; 25: 1952-1958.
[14] Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F and Manke T. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 2016; 44: W160-165.
[15] Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, Li D, Choudhary MNK, Li Y, Hu M, Hardison R, Wang T and Yue F. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biol 2018; 19: 151.
[16] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW and Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016; 44: W90-97.
[17] Pinero J, Ramirez-Anguita JM, Sauch-Pitarch J, Ronzano F, Centeno E, Sanz F and Furlong LI. The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Res 2020; 48: D845-D855.
[18] Zhang D, Huo D, Xie H, Wu L, Zhang J, Liu L, Jin Q and Chen X. CHG: a systematically integrated database of cancer hallmark genes. Front Genet 2020; 11: 29.
[19] Deist TM, Dankers F, Valdes G, Wijsman R, Hsu IC, Oberije C, Lustberg T, van Soest J, Hoebers F, Jochems A, El Naqa I, Wee L, Morin O, Raleigh DR, Bots W, Kaanders JH, Belderbos J, Kwint M, Solberg T, Monshouwer R, Bussink J, Dekker A and Lambin P. Erratum: "machine learning algorithms for outcome prediction in (chemo)radiotherapy: an empirical comparison of classifiers" [Med Phys 45 (7), 3449-3459 (2018)]. Med Phys 2019; 46: 1080-1087.
[20] Friedman J, Hastie T and Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010; 33: 1-22.
[21] Alderden J, Pepper GA, Wilson A, Whitney JD, Richardson S, Butcher R, Jo Y and Cummins MR. Predicting pressure injury in critical care patients: a machine-learning model. Am J Crit Care 2018; 27: 461-468.
[22] Cinelli M, Sun Y, Best K, Heather JM, Reich-Zeliger S, Shifrut E, Friedman N, Shawe-Taylor J and Chain B. Feature selection using a one dimensional naive Bayes' classifier increases the accuracy of support vector machine classification of CDR3 repertoires. Bioinformatics 2017; 33: 951-955.
[23] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC and Muller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011; 12: 77.
[24] Racle J, de Jonge K, Baumgaertner P, Speiser DE and Gfeller D. Simultaneous enumeration

of cancer and immune cell types from bulk tumor gene expression data. Elife 2017; 6: e26476.
[25] Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautes-Fridman C, Fridman WH and de Reynies A. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016; 17: 218.
[26] Mayakonda A, Lin DC, Assenov Y, Plass C and Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res 2018; 28: 1747-1756.
[27] Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, Schrock A, Campbell B, Shlien A, Chmielecki J, Huang F, He Y, Sun J, Tabori U, Kennedy M, Lieber DS, Roels S, White J, Otto GA, Ross JS, Garraway L, Miller VA, Stephens PJ and Frampton GM. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med 2017; 9: 34.
[28] Hakimi AA, Reznik E, Lee CH, Creighton CJ, Brannon AR, Luna A, Aksoy BA, Liu EM, Shen R, Lee W, Chen Y, Stirdivant SM, Russo P, Chen YB, Tickoo SK, Reuter VE, Cheng EH, Sander C and Hsieh JJ. An integrated metabolic atlas of clear cell renal cell carcinoma. Cancer Cell 2016; 29: 104-116.
[29] Hanzelmann S, Castelo R and Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013; 14: 7.
[30] Ye Y, Hu Q, Chen H, Liang K, Yuan Y, Xiang Y, Ruan H, Zhang Z, Song A, Zhang H, Liu L, Diao L, Lou Y, Zhou B, Wang L, Zhou S, Gao J, Jonasch E, Lin SH, Xia Y, Lin C, Yang L, Mills GB, Liang H and Han L. Characterization of hypoxia-associated molecular features to aid hypoxia-targeted therapy. Nat Metab 2019; 1: 431-444.
[31] Liu F, Qin L, Liao Z, Song J, Yuan C, Liu Y, Wang Y, Xu H, Zhang Q, Pei Y, Zhang H, Pan Y, Chen X, Zhang Z, Zhang W and Zhang B. Microenvironment characterization and multi-omics signatures related to prognosis and immunotherapy response of hepatocellular carcinoma. Exp Hematol Oncol 2020; 9: 10.
[32] Xu T, Le TD, Liu L, Su N, Wang R, Sun B, Colaprico A, Bontempi G and Li J. CancerSubtypes: an R/Bioconductor package for molecular cancer subtype identification, validation and visualization. Bioinformatics 2017; 33: 3131-3133.
[33] Vasudevan P and Murugesan T. Cancer subtype discovery using prognosis-enhanced neural network classifier in multigenomic data. Technol Cancer Res Treat 2018; 17: 1533033818790509.
[34] Yu G, Wang LG, Han Y and He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287.
[35] Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P and Liu XS. Large-scale public data reuse to model immunotherapy response and resistance. Genome Med 2020; 12: 21.
[36] Geeleher P, Cox N and Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One 2014; 9: e107468.
[37] Calo E and Wysocka J. Modification of enhancer chromatin: what, how, and why? Mol Cell 2013; 49: 825-837.
[38] Yuan SX, Wang J, Yang F, Tao QF, Zhang J, Wang LL, Yang Y, Liu H, Wang ZG, Xu QG, Fan J, Liu L, Sun SH and Zhou WP. Long noncoding RNA DANCR increases stemness features of hepatocellular carcinoma by derepression of CTNNB1. Hepatology 2016; 63: 499-511.
[39] Tung EK, Mak CK, Fatima S, Lo RC, Zhao H, Zhang C, Dai H, Poon RT, Yuen MF, Lai CL, Li JJ, Luk JM and Ng IO. Clinicopathological and prognostic significance of serum and tissue Dickkopf-1 levels in human hepatocellular carcinoma. Liver Int 2011; 31: 1494-1504.
[40] Revill K, Wang T, Lachenmayer A, Kojima K, Harrington A, Li J, Hoshida Y, Llovet JM and Powers S. Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma. Gastroenterology 2013; 145: 1424-1435, e1421-1425.
[41] Cai J, Chen L, Zhang Z, Zhang X, Lu X, Liu W, Shi G, Ge Y, Gao P, Yang Y, Ke A, Xiao L, Dong R, Zhu Y, Yang X, Wang J, Zhu T, Yang D, Huang X, Sui C, Qiu S, Shen F, Sun H, Zhou W, Zhou J, Nie J, Zeng C, Stroup EK, Zhang X, Chiu BC, Lau WY, He C, Wang H, Zhang W and Fan J. Genome-wide mapping of 5-hydroxymethylcytosines in circulating cell-free DNA as a non-invasive approach for early detection of hepatocellular carcinoma. Gut 2019; 68: 2195-2205.
[42] Rinaldi L, Nevola R, Franci G, Perrella A, Corvino G, Marrone A, Berretta M, Morone MV, Galdiero M, Giordano M, Adinolfi LE and Sasso FC. Risk of hepatocellular carcinoma after HCV clearance by direct-acting antivirals treatment predictive factors and role of epigenetics. Cancers (Basel) 2020; 12: 1351.
[43] Moeini A, Torrecilla S, Tovar V, Montironi C, Andreu-Oller C, Peix J, Higuera M, Pfister D, Ramadori P, Pinyol R, Sole M, Heikenwalder M, Friedman SL, Sia D and Llovet JM. An immune gene expression signature associated with development of human hepatocellular carcinoma

identifies mice that respond to chemopreventive agents. Gastroenterology 2019; 157: 1383-1397, e1311.
[44] Maros ME, Capper D, Jones DTW, Hovestadt V, von Deimling A, Pfister SM, Benner A, Zucknick M and Sill M. Machine learning workflows to estimate class probabilities for precision cancer diagnostics on DNA methylation microarray data. Nat Protoc 2020; 15: 479-512.
[45] Chae YK, Davis AA, Raparia K, Agte S, Pan A, Mohindra N, Villaflor V and Giles F. Association of tumor mutational burden with DNA repair mutations and response to anti-PD-1/PD-L1 therapy in non-small-cell lung cancer. Clin Lung Cancer 2019; 20: 88-96 e86.
[46] Fancello L, Gandini S, Pelicci PG and Mazzarella L. Tumor mutational burden quantification from targeted gene panels: major advancements and challenges. J Immunother Cancer 2019; 7: 183.
[47] Chae YK, Anker JF, Oh MS, Bais P, Namburi S, Agte S, Giles FJ and Chuang JH. Mutations in DNA repair genes are associated with increased neoantigen burden and a distinct immunophenotype in lung squamous cell carcinoma. Sci Rep 2019; 9: 3235.
[48] Tang H, Qiao J and Fu YX. Immunotherapy and tumor microenvironment. Cancer Lett 2016; 370: 85-90.
[49] Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, Ramaswamy S, Futreal PA, Haber DA, Stratton MR, Benes C, McDermott U and Garnett MJ. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res 2013; 41: D955-961.
[50] Avila MA, Berasain C, Sangro B and Prieto J. New therapies for hepatocellular carcinoma. Oncogene 2006; 25: 3866-3884.
[51] Lee JE, Park YK, Park S, Jang Y, Waring N, Dey A, Ozato K, Lai B, Peng W and Ge K. Brd4 binds to active enhancers to control cell identity gene induction in adipogenesis and myogenesis. Nat Commun 2017; 8: 2217.
[52] Picaud S, Wells C, Felletar I, Brotherton D, Martin S, Savitsky P, Diez-Dacal B, Philpott M, Bountra C, Lingard H, Fedorov O, Muller S, Brennan PE, Knapp S and Filippakopoulos P. RVX-208, an inhibitor of BET transcriptional regulators with selectivity for the second bromodomain. Proc Natl Acad Sci U S A 2013; 110: 19754-19759.
[53] Flavahan WA, Gaskell E and Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. Science 2017; 357: eaal2380.
[54] Feinberg AP, Koldobskiy MA and Gondor A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. Nat Rev Genet 2016; 17: 284-299.
[55] Bertoletti A, Kennedy PTF and Durantel D. HBV infection and HCC: the 'dangerous liaisons'. Gut 2018; 67: 787-788.
[56] Chen CD, Wang CL, Yu CJ, Chien KY, Chen YT, Chen MC, Chang YS, Wu CC and Yu JS. Targeted proteomics pipeline reveals potential biomarkers for the diagnosis of metastatic lung cancer in pleural effusion. J Proteome Res 2014; 13: 2818-2829.
[57] Liu J, Liu Z, Zhang X, Gong T and Yao D. Correction to: Bioinformatic exploration of OLFML2B overexpression in gastric cancer base on multiple analyzing tools. BMC Cancer 2019; 19: 759.
[58] Tu H, Wu M, Huang W and Wang L. Screening of potential biomarkers and their predictive value in early stage non-small cell lung cancer: a bioinformatics analysis. Transl Lung Cancer Res 2019; 8: 797-807.
[59] Moller M, Strand SH, Mundbjerg K, Liang G, Gill I, Haldrup C, Borre M, Hoyer S, Orntoft TF and Sorensen KD. Heterogeneous patterns of DNA methylation-based field effects in histologically normal prostate tissue from cancer patients. Sci Rep 2017; 7: 40636.
[60] Yu Z, Wang R, Chen F, Wang J and Huang X. Five novel oncogenic signatures could be utilized as afp-related diagnostic biomarkers for hepatocellular carcinoma based on next-generation sequencing. Dig Dis Sci 2018; 63: 945-957.
[61] Huang X, Yang C, Wang J, Sun T and Xiong H. Integrative analysis of DNA methylation and gene expression reveals distinct hepatocellular carcinoma subtypes with therapeutic implications. Aging (Albany NY) 2020; 12: 4970-4995.
[62] Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY, Villanueva A, Newell P, Ikeda K, Hashimoto M, Watanabe G, Gabriel S, Friedman SL, Kumada H, Llovet JM and Golub TR. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. Cancer Res 2009; 69: 7385-7392.
[63] Chiang DY, Villanueva A, Hoshida Y, Peix J, Newell P, Minguez B, LeBlanc AC, Donovan DJ, Thung SN, Sole M, Tovar V, Alsinet C, Ramos AH, Barretina J, Roayaie S, Schwartz M, Waxman S, Bruix J, Mazzaferro V, Ligon AH, Najfeld V, Friedman SL, Sellers WR, Meyerson M and Llovet JM. Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. Cancer Res 2008; 68: 6779-6788.
[64] Hegde PS and Chen DS. Top 10 challenges in cancer immunotherapy. Immunity 2020; 52: 17-35.
[65] Zhu AX, Finn RS, Edeline J, Cattan S, Ogasawara S, Palmer D, Verslype C, Zagonel V, Fartoux L, Vogel A, Sarker D, Verset G, Chan SL,

Knox J, Daniele B, Webber AL, Ebbinghaus SW, Ma J, Siegel AB, Cheng AL and Kudo M; KEYNOTE-224 investigators. Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. Lancet Oncol 2018; 19: 940-952.

[66] Charoentong P, Angelova M, Efremova M, Gallasch R, Hackl H, Galon J and Trajanoski Z. Bioinformatics for cancer immunology and immunotherapy. Cancer Immunol Immunother 2012; 61: 1885-1903.

[67] Austin JW, Lu P, Majumder P, Ahmed R and Boss JM. STAT3, STAT4, NFATc1, and CTCF regulate PD-1 through multiple novel regulatory regions in murine T cells. J Immunol 2014; 192: 4876-4886.

[68] Palsson-McDermott EM, Dyck L, Zaslona Z, Menon D, McGettrick AF, Mills KHG and O'Neill LA. Pyruvate kinase M2 is required for the expression of the immune checkpoint PD-L1 in immune cells and tumors. Front Immunol 2017; 8: 1300.

[69] De Plaen E, Lurquin C, Van Pel A, Mariame B, Szikora JP, Wolfel T, Sibille C, Chomez P and Boon T. Immunogenic (tum-) variants of mouse tumor P815: cloning of the gene of tum- antigen P91A and identification of the tum- mutation. Proc Natl Acad Sci U S A 1988; 85: 2274-2278.

[70] Roe JS, Mercan F, Rivera K, Pappin DJ and Vakoc CR. BET bromodomain inhibition suppresses the function of hematopoietic transcription factors in acute myeloid leukemia. Mol Cell 2015; 58: 1028-1039.

[71] Kanno T, Kanno Y, LeRoy G, Campos E, Sun HW, Brooks SR, Vahedi G, Heightman TD, Garcia BA, Reinberg D, Siebenlist U, O'Shea JJ and Ozato K. BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. Nat Struct Mol Biol 2014; 21: 1047-1057.

[72] Jiang G, Deng W, Liu Y and Wang C. General mechanism of JQ1 in inhibiting various types of cancer. Mol Med Rep 2020; 21: 1021-1034.

[73] Ding N, Hah N, Yu RT, Sherman MH, Benner C, Leblanc M, He M, Liddle C, Downes M and Evans RM. BRD4 is a novel therapeutic target for liver fibrosis. Proc Natl Acad Sci U S A 2015; 112: 15713-15718.

[74] Kagoya Y, Nakatsugawa M, Yamashita Y, Ochi T, Guo T, Anczurowski M, Saso K, Butler MO, Arrowsmith CH and Hirano N. BET bromodomain inhibition enhances T cell persistence and function in adoptive immunotherapy models. J Clin Invest 2016; 126: 3479-3494.

[75] Jiang Y, Sun A, Zhao Y, Ying W, Sun H, Yang X, Xing B, Sun W, Ren L, Hu B, Li C, Zhang L, Qin G, Zhang M, Chen N, Zhang M, Huang Y, Zhou J, Zhao Y, Liu M, Zhu X, Qiu Y, Sun Y, Huang C, Yan M, Wang M, Liu W, Tian F, Xu H, Zhou J, Wu Z, Shi T, Zhu W, Qin J, Xie L, Fan J, Qian X and He F; Chinese Human Proteome Project (CNHPP) Consortium. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma. Nature 2019; 567: 257-261.

[76] Yang Y, Chen L, Gu J, Zhang H, Yuan J, Lian Q, Lv G, Wang S, Wu Y, Yang YT, Wang D, Liu Y, Tang J, Luo G, Li Y, Hu L, Sun X, Wang D, Guo M, Xi Q, Xi J, Wang H, Zhang MQ and Lu ZJ. Recurrently deregulated lncRNAs in hepatocellular carcinoma. Nat Commun 2017; 8: 14421.

[77] Yoo S, Wang W, Wang Q, Fiel MI, Lee E, Hiotis SP and Zhu J. A pilot systematic genomic comparison of recurrence risks of hepatitis B virus-associated hepatocellular carcinoma with low- and high-degree liver fibrosis. BMC Med 2017; 15: 214.