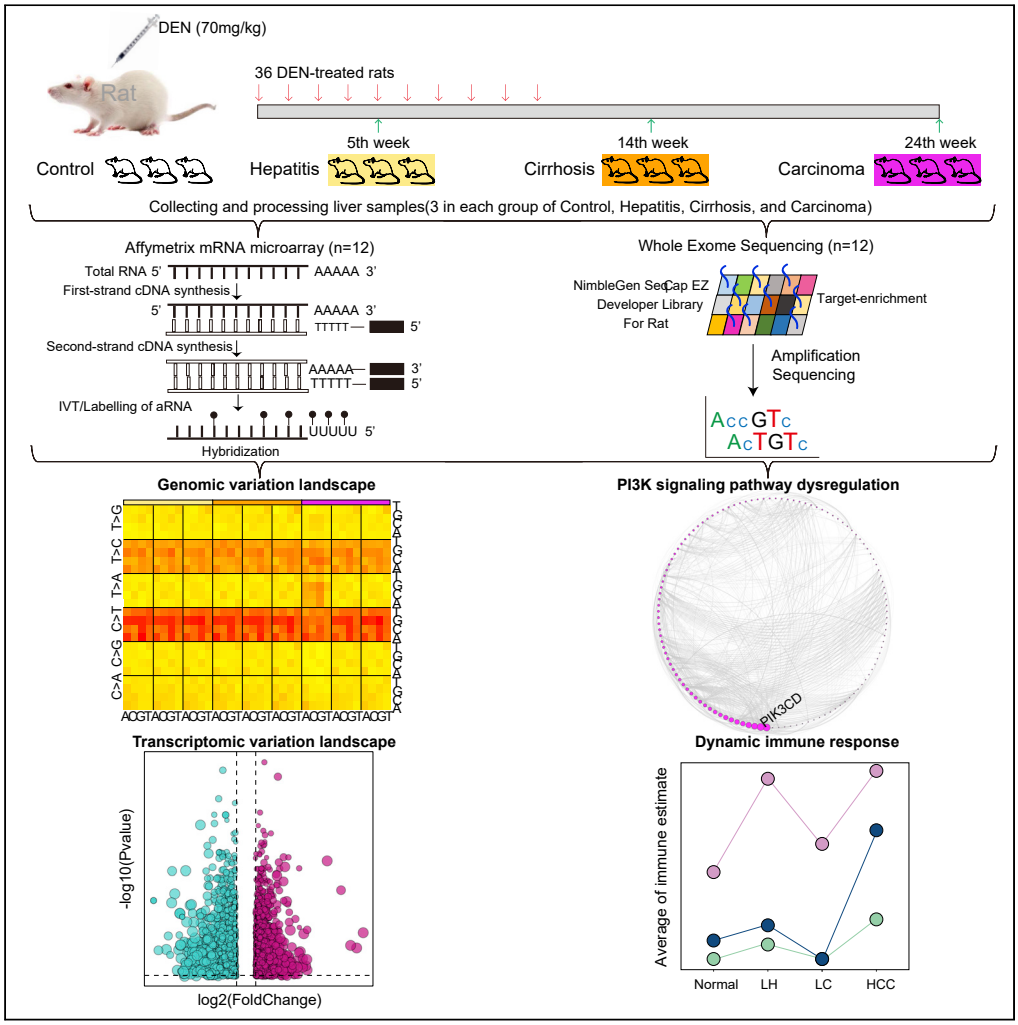
iScience

CellPress
OPEN ACCESS

Article

The Mutational and Transcriptional Landscapes of Hepatocarcinogenesis in a Rat Model



Zhiao Chen, Shengli Li, Mengting Shen, ..., Wenming Cong, Leng Han, Xianghuo He

leng.han@uth.tmc.edu (L.H.)
xhhe@fudan.edu.cn (X.H.)

HIGHLIGHTS

Genome instability and cancer-related signaling aberrance occurred in liver hepatitis

The rivalry between tumor-suppressive and oncogenic strengths peaked in liver cirrhosis

PI3K-Akt plays a key role in the process of hepatocarcinogenesis

Hepatocarcinogenesis witnessed dynamic changes of monocytes and natural killer cells

Chen et al., iScience 23, 101690
November 20, 2020
https://doi.org/10.1016/j.isci.2020.101690

iScience

CellPress
OPEN ACCESS

Article

The Mutational and Transcriptional Landscapes of Hepatocarcinogenesis in a Rat Model

Zhiao Chen,[1,6] Shengli Li,[2,5,6] Mengting Shen,[1,6] Xinyuan Lu,[3] Chunyang Bao,[1] Di Chen,[1] Jie Ding,[1] Qifeng Wang,[1] Shenglin Huang,[1] Wenming Cong,[3] Leng Han,[2,*] and Xianghuo He[1,4,7,*]

SUMMARY

Hepatocellular carcinoma (HCC) initiation is characterized by stepwise accumulation of molecular alterations, during which the early events are largely unknown. Here, we presented a comprehensive genomic and transcriptomic landscape at stages of hepatitis, cirrhosis, and HCC by using a diethylnitrosamine-induced rat HCC model. We observed the early occurrence of gene instability and aberrant cancer associated signaling pathways in liver hepatitis. We further characterized the progressive molecular changes during hepatocarcinogenesis, wherein the intense rivalry between tumor-suppressive and oncogenic strengths occurred in cirrhosis stage. Despite the significant pathological difference, mutation signatures and expression landscape are highly similar between hepatitis and cirrhosis stages. Furthermore, we identified PI3K-Akt signaling pathway as a key pathway in the process of hepatocarcinogenesis through integrative analysis, and PIK3CD is a potential biomarker indicating HCC recurrence. The dynamic immune response during hepatocarcinogenesis, such as continuous decline of monocytes, suggests an immunological intervention strategy beyond chemoprevention for liver cancer.

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the most common malignant tumors and is the second leading cause of cancer-related deaths worldwide (Villanueva, 2019). Major risk factors for HCC include infection with hepatitis B virus or hepatitis C virus, alcoholic liver disease, and most probably non-alcoholic fatty liver disease (El-Serag and Rudolph, 2007; Villanueva, 2019). HCC largely occurs in an established background of chronic liver disease and cirrhosis but rarely develops in healthy liver during normal aging. Patients with chronic liver disease have sustained hepatic inflammation, cirrhosis, and aberrant hepatocyte regeneration. These abnormalities can cause cirrhosis, which is present in 80–90% of patients with HCC. HCC can arise in patients who have chronic liver disease but do not have established cirrhosis, although the HCC risk remains low at this stage. The cancer risk increases sharply in response to chronic liver injury at the cirrhosis stage. A series of genetic and epigenetic events occurred during this process and culminate in the formation of preneoplastic lesions. On this foundation, additional molecular alterations in the cirrhosis tissue provide liver cells with sustained proliferative, invasive, and survival advantages and complete the transition to HCC(El-Serag, 2012; Yu et al., 2013).

Although many approaches can provide new insights into the complex molecular pathogenesis of HCC, i.e., identification of novel oncogenic pathways and cancer driver genes (Fujimoto et al., 2012; Guichard et al., 2012; Li et al., 2019), sequential analysis of the development of tumors in humans is very difficult. It is still a challenge to study the development of HCC, especially those early events, e.g., from chronic liver disease to cirrhosis stage (Sottoriva et al., 2015; Tao et al., 2011). For instance, most analyses have failed to identify gene profiles or the genes that play key roles during the entire carcinogenic process of tumor development. Animal models of hepatocarcinogenesis in mice and rats expanded our understanding of molecular events driving HCC and identification of key signaling pathways involved in this process. Although there is not yet a single model that fully recapitulates all aspects of human liver cancer, the availability of multiple chemical and genetic models allows us for selecting the appropriate one to investigate the dynamic changes in hepatitis and cirrhosis stages to better prevent and diagnose cancers (CavigLia and Schwabe, 2014). Chronic administration of diethylnitrosamine (DEN) causes liver tumors with a sequential progression of hepatitis, cirrhosis and tumor formation. Comparing to the 34% of induction rate for mice,

[1]Fudan University Shanghai Cancer Center and Institutes of Biomedical Sciences, Shanghai Medical College, Fudan University, Shanghai 200032, China

[2]Department of Biochemistry and Molecular Biology, McGovern Medical School at The University of Texas Health Science Center at Houston, Houston, TX 77030, USA

[3]Department of Pathology, Eastern Hepatobiliary Surgery Hospital, Second Military Medical University, Shanghai 200438, China

[4]Key Laboratory of Breast Cancer in Shanghai, Fudan University Shanghai Cancer Center, Fudan University, Shanghai 200032, China

[5]Institute of Translational Medicine, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 201620, China

[6]These authors contributed equally

[7]Lead Contact

*Correspondence: leng.han@uth.tmc.edu (L.H.), xhhe@fudan.edu.cn (X.H.)

https://doi.org/10.1016/j.isci.2020.101690

recent studies demonstrated that Sprague-Dawley (SD) rats treated with 10 weekly doses of DEN (70 mg/kg body weight, intraperitoneally) develop HCC with up to 80% (CavigLia and Schwabe, 2014). This DEN-induced rat HCC model closely mimics human liver cancer development, which typically involve sequential stimuli for initiation and progression. DEN is administered to cause DNA damage, that can induce hepatocellular injury, compensatory proliferation, inflammation (1∼8 weeks after first DEN treatment), followed by cirrhosis (8∼16 weeks). Ultimately, it leads to the amplification of initiated cells and clonal expansion to form tumors (16∼22 weeks) (Magee and Barnes, 1956). Therefore, this experimental model of HCC can reproduce the frequent alterations of the hepatic environment. The genetic and transcriptomic landscape of this rat model remains to be investigated in detail, particularly at hepatitis and cirrhosis stages.

It is well accepted that chronic inflammation is a key driver of disease progression in the prototypical etiologies of HCC (Llovet et al., 2016). As a central player in immunoregulation, the liver ensures organ and systemic protection while maintaining immunotolerance (Jenne and Kubes, 2013). A dysregulated immune system plays a key role in the development of HCC. Persistent up-regulation of inflammation signals leads to changes in the number and/or function of immune cells, altered immunological, survival and proliferation signals, and subsequently, the induction of tumorigenesis (Notas et al., 2009; Robinson et al., 2016) Moreover, the deregulation of innate and adaptive immune systems also enable the liver fail to detect and eliminate the transformed cells (Ringelhan et al., 2018). However, the roles of immunotolerance mechanisms and immune cell subsets involved in the progression of chronic liver disease are poorly understood. Thus, understanding the immunological network of the liver and alterations to the microenvironment in DEN-induced rat HCC model is also vital for liver cancer research.

RESULTS

DEN-Induced Hepatocarcinogenesis in Rat

We used a well established protocol to generate a genotoxic hepatocarcinogenesis model wherein 36 male SD rats were administered 10 once-weekly doses of DEN (intraperitoneal injection, 70 mg/kg body weight) for 10 weeks (Figure 1A). To capture developing stages of inflammation, cirrhosis, and liver tumor, we sacrificed 12 rats at fifth weeks, 14th weeks, and 24th weeks, respectively. At the fifth week after the first DEN administration, inflammatory infiltration occurred in 9 of 12 rats, which represents the liver hepatitis (LH) stage (Figure S1). Pathological histology in livers from this group showed hepatocyte swelling, expansion of some portal areas, and inflammatory cell infiltrating in portal duct areas. At the 14th week, liver cirrhosis (LC) occurred in 10 of 12 rats, which represent the stage of LC stage (Figure S1). Results of pathological sections showed fibrous tissue hyperplasia, fibrous bands, and significant portal-to-portal and portal-to-central bridging with pseudolobular formation in livers from this group. At the 24th week, 7 of 12 rats developed multiple, macroscopically identifiable tumors, and among which 6 developed HCC, which represent the stage of HCC (Figure S1). Tumor nodule from livers in this group showed tumor cells arranged irregularly, small cell change with high cell density and cytological atypia. As for the sample collection, 3 rats of 9 with the occurrence of inflammatory infiltration (at week 5), 3 rats of 10 with the occurrence of cirrhosis (at week 14), and 3 rats of 6 with the occurrence of HCC (at week 24) were chosen, and one pathologic tissue sample from each of the rat, which have similar histological character in each group, was collected to perform the whole-exome sequencing (WES). Three histologically confirmed HCC from 3 rats were collected to perform WES. As for the control group, we used the liver tissues from the control rats sacrificed at the fifth week to perform WES and gene expression microarray profiling (Figure 1B).

Genomic and Transcriptomic Variations in Liver Cancer Developing Stages

To understand the molecular basis of hepatocarcinogenesis, we investigated the genomic and transcriptomic variations in the rat model of DEN-induced HCC. Our analysis revealed that 88.4% of point mutations are non-synonymous, with no detectable bias in distribution of missense, intorn/intergenic and other non-synonymous mutations across different stages (Figure 2A and Table S1). As expected, the HCC stage carried the highest mutation burden (52.3 missense mutations per megabase) (Figure 2A). In addition, an attenuate mutation load was observed in LC samples. By analyzing all 96 possible substitutions based on trinucleotide context, we revealed highly consistent mutation signature across all DEN-treated rat samples (Figure S2A). Among these SNV mutations, C>T transitions was the most frequent SNV, following by T>C transitions. Furthermore, we applied a non-negative matrix factorization method to analyze the 96 substitution patterns (see Methods) and identified mutational signatures in all stages (Figure 2B). Two signatures were identified in HCC samples, wherein Signature 1 was characterized by the prominence of C>T transitions at NpCpG trinucleotides, while Signature 2 was characterized by T>A transversions at NpTpG

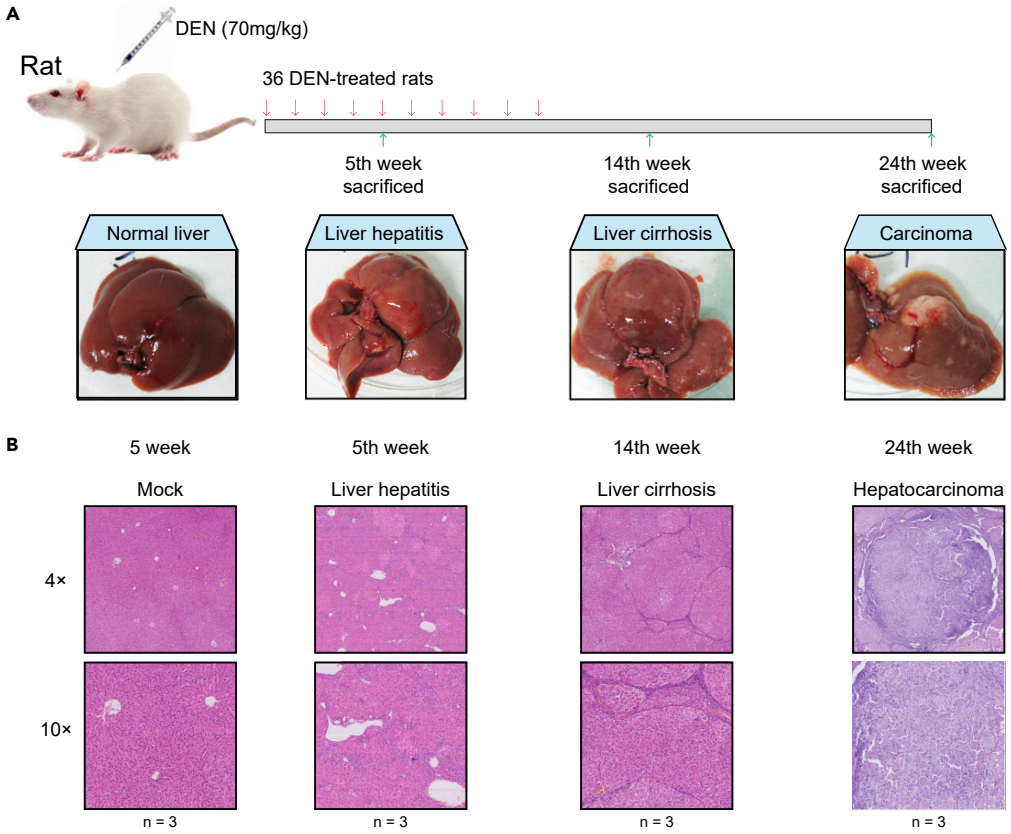


Figure 1. DEN-Induced Hepatocarcinogenesis in Rat Models

(A) Schematic diagram of the timeline for DEN-induced liver cancer rat model. Cohorts of SD male rats were either injected intraperitoneally with a dosage of 70 mg/kg DEN or distilled water once per week for 10 weeks. Liver samples were collected from cohorts of LH, LC, and HCC after administration of DEN for in 5, 14, and 24 weeks, respectively.

(B) Representative photomicrographs of serial sections of liver tissue from DEN-treated rats. H&E staining demonstrates tissue morphology. Original magnification, ×40 or ×100, as indicated.

See also Figure S1.

trinucleotides, and T>C transitions at NpTpC trinucleotides. The signatures matched the previously identified signature 12, 16, 3, and 5 for human cancers (Figure S2B), which are described in the Catalog of Somatic Mutations in Cancer database (Tate et al., 2019). Specifically, the signature 12 and 16 were exclusively found in liver cancer, while signature 3 and 5 were identified in most cancer types (Alexandrov et al., 2013). To seek evidence for important genes of hepatocarcinoma in our rat model, we explored several existing human liver cancer cohorts. These mutated genes in our DEN-induced HCC rat captured about half of those in TCGA Liver HCC (Totoki et al., 2017), Schulze et al. (Schulze et al., 2015), Ahn et al. (Ahn et al., 2014) or Harding et al. (Harding et al., 2019) cohort (Figure S2C). Additionally, we identified 38 recurrently mutated genes (mutation frequency >5% in at least two cohorts), including the tumor suppressor genes *TP53,* and oncogene *CTNNB1* in WNT pathway (Figure S2D). We checked the mutation status of specific genes in additional DEN treatment rat sample, such as Dnah5 and Fras1 (Figure S3). The corresponding mutation points were found in additional DEN treatment rat samples. We further built the phylogenetic tree by using R package ape (Paradis et al., 2004). In particular, pair-wide distances between samples were calculated as the number of genomic loci where the sample genotypes differ as described by previous study (Connor et al., 2018). As expected, samples in the same groups were closer (Figure S4A). After developing LH, rat livers were more genetically heterogeneous as they develop to cirrhosis/HCC. Furthermore, we found that genomic variants in aberrant growth and genomic instability occurred from early stage of hepatocarcinogenesis, such as "mitotic spindle" and "TGF beta signaling" (Figure S4B).

We further characterized transcriptomic changes in each stage by comparing gene expression with those normal control samples. We detected 855, 966, and 2,601 differentially expressed genes in LH, LC, and

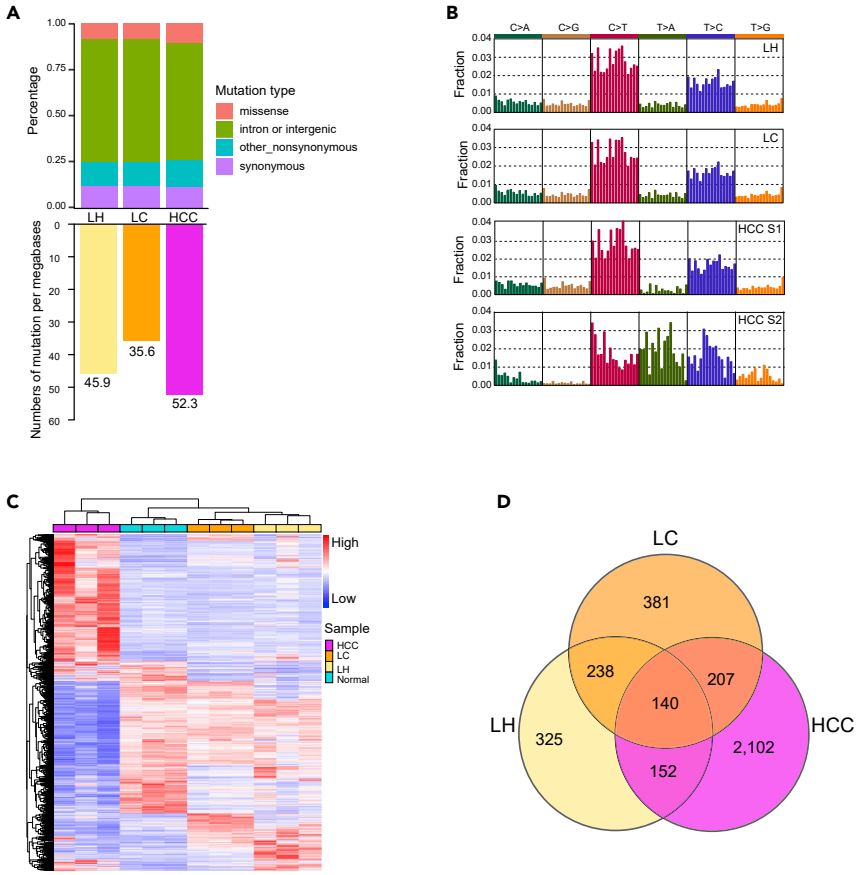


Figure 2. Genomic and Transcriptomic Variations in Each Stage of Liver Cancer Development

(A) Upper bar plots show the percentages of different mutation types in each stage, and lower bar plots are mutation burden (missense mutation) in different stages.

(B) Patterns of mutation signatures in each stage.

(C) Hierarchical clustering of significantly differentially expressed genes in DEN-treated rats from samples of hepatitis, cirrhosis, and hepatocellular carcinomas. Red color indicates higher expression level compared with the control, while blue shows lower expression.

(D) Venny plot presents overlaps of differentially expressed genes among LH, LC, and HCC stages.

See also Figures S2 and S5, and Table S2.

HCC stage, respectively (Figures 2C and S5 and Table S2). Only a few dysregulated genes were shared by all stages, while the majority were HCC-specific (80.8%, 2,102) (Figure 2D). In addition, most HCC-specific dysregulated genes contributed to cell proliferation such as "MYC targets" and "G2M checkpoints", whereas stages-shared dysregulated genes were involved in liver metabolism such as "bile acid metabolism" and "glycolysis" (Figure S5).

Characterization of Changing Genome and Transcriptome during Hepatocarcinogenesis

Our DEN-induced model represents the progressive stages of liver cancer, which enables us to capture the molecular changes from stage to stage. In order to characterize the mutagenesis during hepatocarcinogenesis, mutated genes were divided into 7 groups according to their mutated stages, wherein HCC stage witnessed a burst of mutagenesis with 1,195 mutated genes (Figure 3A and Table S3). Notably, the mutations of quite a number of genes (461 genes) occurred from the LH stage, which were in enriched in "WNT beta-catenin signaling", "apical junction" and "coagulation" (Figure 3B). Moreover, 91 genes were found to be mutated from LC mutations, which were involved in more cancer-related hallmarks such as "IL2 STAT5 signaling", "WNT beta-catenin signaling" and "TGF beta signaling". As expected, the burst of transcriptional dysregulation happened in HCC stage (Figure 3C and Table S4 and S5). Interestingly, the progress from hepatitis to cirrhosis stage showed less transcriptional dysregulation, compared to those from

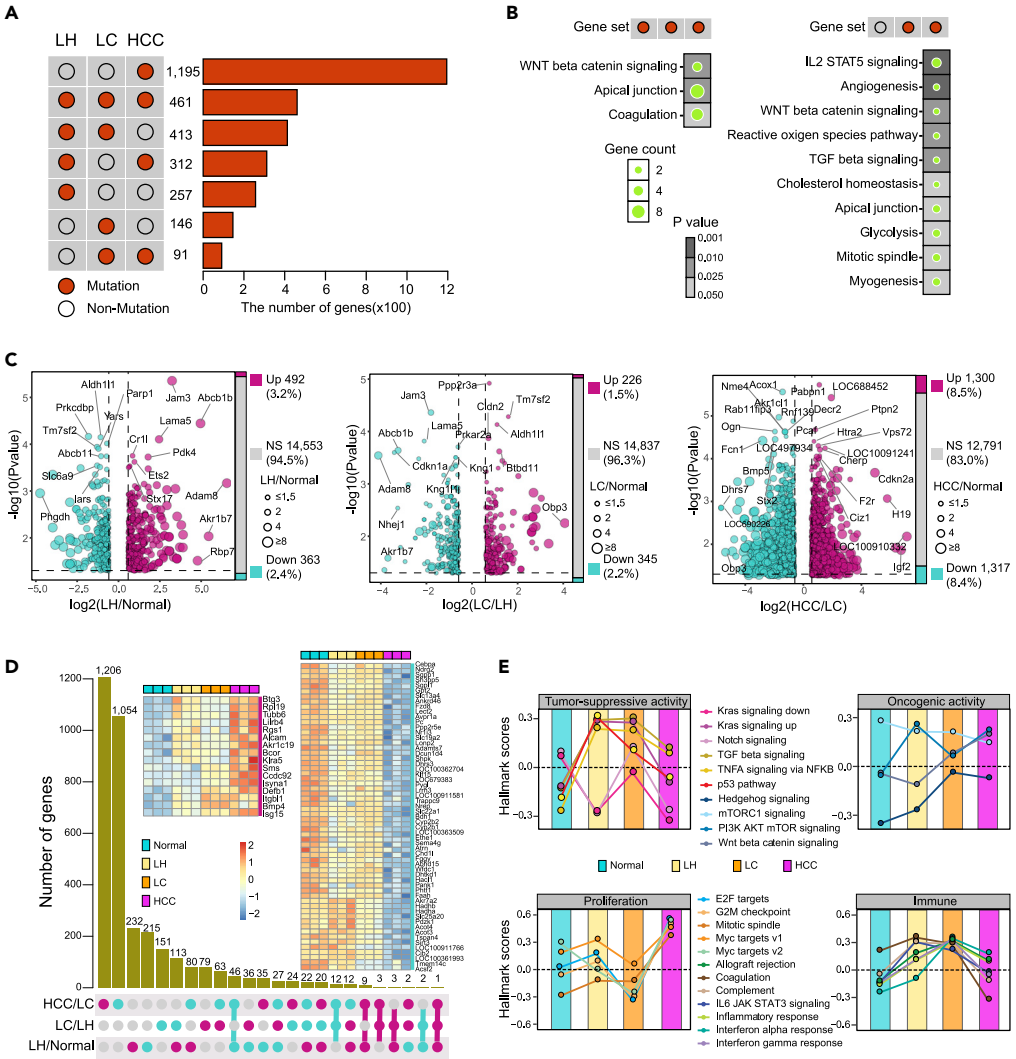


Figure 3. Characterization of Changing Genome and Transcriptome during Hepatocarcinogenesis

(A) The number of genes in different mutation patterns.

(B) Significantly enriched hallmarks for genes that have mutations in all stages and last two stages, respectively.

(C) Differentially expressed genes in each stage compared to its prior stage.

(D) Gene counts in various expression changing patterns. Left heatmap shows continuous up-regulation genes, and right heatmap shows continuous down-regulation genes.

(E) The dynamic changes of hallmarks during DEN-induced liver carcinogenesis.

See also Table S3 and Table S4.

hepatitis and from cirrhosis to liver cancer, which was implicated by the closer clustering and similar expression pattern as well (Figures 2C and S6A). The expression similarity between LH and LC was also observed in these human samples (Figure S6B). In addition to the transcriptional dysregulation burst from LC to HCC stage, we found 16 and 60 genes showed continuous up-regulation and down-regulation across the progressive stages, respectively (Figure 3D). To further characterize the biological changes across the progressive stages during hepatocarcinogenesis, we calculated the scores of key hallmarks in each stage (see Methods). The tumor-suppressive activities showed sharp decrease from LC to HCC stage, whereas oncogenic activities displayed up-regulation or maintained high levels (Figure 3E). However, some tumor-suppressive pathways showed the highest activity in LC stage such as Notch signaling pathway. Moreover, prior to the increase in HCC stage, PI3K-AKT-mTOR signaling pathway displayed sharp down-regulation from LH to LC stage. Intriguingly, cell proliferation exhibited a decline in LC and then increased in liver cancer. The immune-related hallmarks continued up-regulation until LC, then declined in liver cancer. In sum,

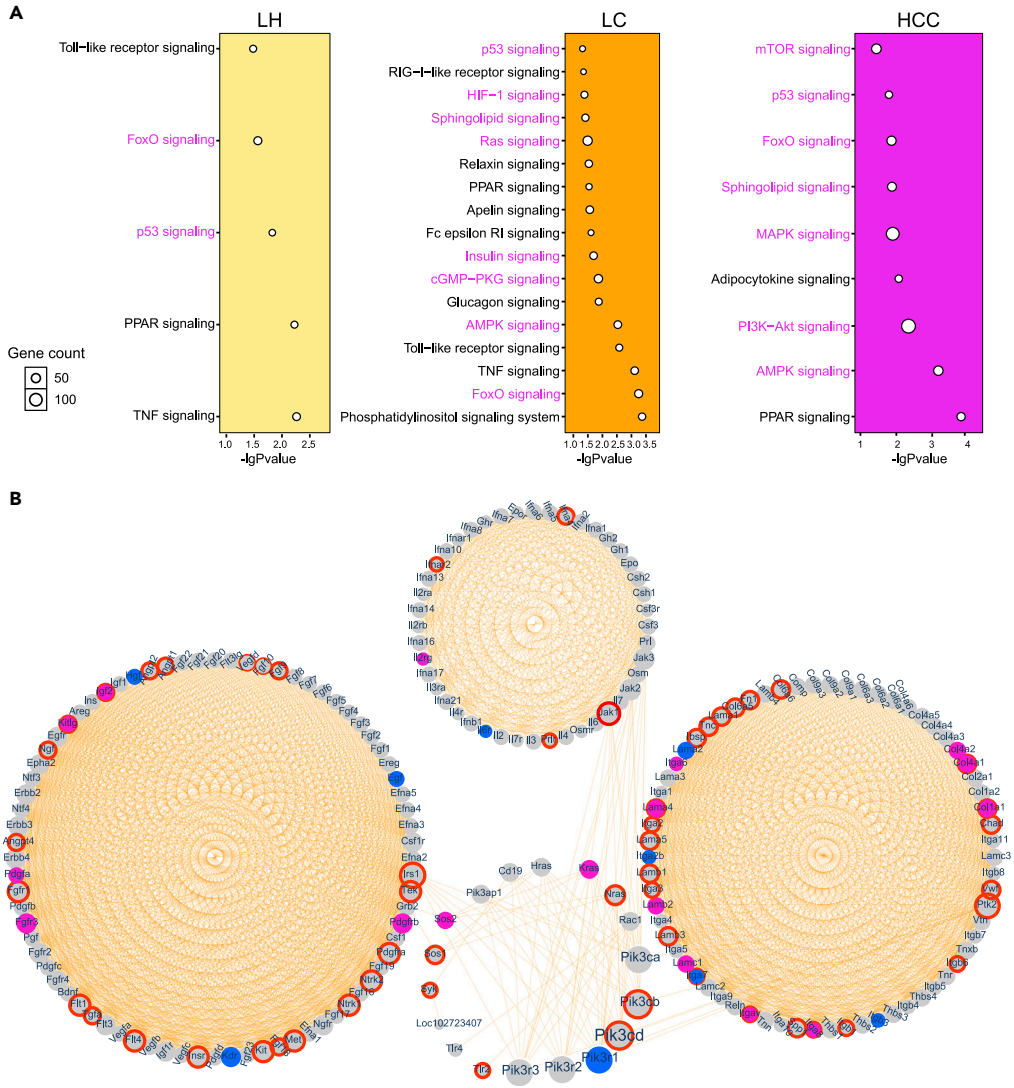


Figure 4. Integration Analysis Identified the Key Affected Signaling Pathways Involved in Hepatocarcinogenesis

(A) The significantly enriched signaling pathways by dysregulated genes in LH, LC, and HCC stages.

(B) The dysregulation of PI3K signaling pathway. Nodes with red borders indicate genes with mutation in HCC stage. Nodes in pink represent higher expression in HCC, while blue ones represent genes with lower expression. The size of nodes show the closeness of genes in the network.

the decrease of cell proliferation, high-level tumor-suppressive activities and active immune response in LC suggested that drastic rivalry between anti-tumor and oncogenic strength occurred in cirrhosis. Presumably, enhancing anti-tumor strength such as Notch signaling pathway or suppressing oncogenic factors such as PI3K-AKT-mTOR signaling pathway in LC would be helpful for preventing the transition from cirrhosis to liver cancer.

Identification of the Key Signaling Pathways Involved in Hepatocarcinogenesis

To further elucidate the regulation of pathways in the progression from LH to cancer, we performed KEGG pathway analysis of dysregulated genes that showed differential expression or mutation in transformed tissues compared to normal tissues. The majority of dysregulated pathways in LH stage were related to inflammation, such as "Toll-like receptor signaling", "PPAR signaling", and "TNF signaling" (Figure 4A). Most disordered pathways in HCC stage were tumor-related, including "AMPK signaling", "PI3K-Akt signaling", "MAPK signaling", and "p53 signaling". As the transitional stage from LH to HCC, the LC stage

showed a mixture of tangled inflammation and tumor-related signaling pathways (Figure 4A). Interestingly, some tumor-related signaling pathways, including "FoxO signaling" and "p53 signaling" were perturbed from the initiation of liver cancer, suggesting a potential molecular mechanism for hepatocarcinogenesis. Of note, the "PI3K-Akt signaling" pathway harbored the most dysregulated genes in HCC, that 26 and 50 genes were observed to differentially expressed and mutated, respectively (Figure 4B). In the centered PI3K, Pik3r1 was down-regulated, and Pik3cd and Pik3cb harbored mutations, which were a missense mutation in Pik3cd (GenBank: NM_001108978, c.T2084C, p.F695S) and an intron mutation in Pik3cb, respectively. Our results suggested that the PI3K pathway might play important roles in the development of liver cancer.

Biological and Clinicopathological Significance of PIK3CD in HCC

Through the integrative analysis of mutation and expression data, we identified an important oncogenic network, PI3K-Akt signaling pathway, wherein PI3K plays central roles. To determine the effect of PI3K-Akt activities on liver tumor growth, we performed loss-of-functional studies in HCC cell lines. The results showed that the knockdown of the PIK3CB did not affect the cell proliferation and colony formation of HCC cells in HepG2/C3A and Huh 7 cell lines (Figures S7A–S7F). In addition, the knockdown of PIK3R1 was observed to markedly promote HCC cell proliferation in Huh 7 cells but not in HepG2/C3A cells (Figures S7G–S7J). Notably, the knockdown of the PIK3CD significantly inhibited the proliferation and colony formation of HCC cells in HepG2/C3A and PLC/PRF/5 cell lines (Figures 5A and 5B).

To further determine the clinicopathological significance of PIK3CD in HCC, we investigated the protein level of PIK3CD by using immunohistochemical staining in a tissue array that included 118 paired HCC and adjacent non-tumor tissues. The results were graded from 0 to 3 depending on the average number of PIK3CD-positive cells in the aggregated field (Figure 5C). The results showed that PIK3CD was primarily localized in the cytoplasm. Notably, high PIK3CD expression was found in primary HCC samples compared with adjacent non-tumor tissues ($p = 0.00069$), thus indicating that an increase in PIK3CD expression is a frequent event in HCC (Figures 5D and 5E). Furthermore, we evaluated the relationship between PIK3CD expression and early recurrence (less than 3 years) by analyzing tumor liver tissue of early stage HCC patients undergoing curative ablation. We found that high expression of PIK3CD was significantly associated with early recurrence (Figure 5F). These results suggest that PIK3CD may represent useful biomarkers for the risk of HCC recurrence and may be used to identify patients who should be closely monitored after curative HCC ablation.

Dynamic Immune Response during Hepatocarcinogenesis

Despite the molecular lesions, the immunological microenvironment has been shown to play crucial roles in the development of liver cancer (Makarova-Rusher et al., 2015), and we observed immunology reprogramming in our hepatocarcinogenesis rat model (Figures 3E and 4A). To maximize the utility of our hepatocarcinogenesis model in understanding extensive molecular dynamics during the development of liver cancer and their clinical implications, we then analyzed the compositions of different immune cells in each sample across all stages based on gene expression profile (See Methods). We observed composition changes of immune cell types in different stages (Figure 6A). Some of the cell types keep stable abundance, such as memory B cells, whereas some showed notable changes among different stages, such as resting dendritic cells and monocytes. Moreover, memory B cells and monocytes are the major parts of immune composition in normal samples, while HCC samples primarily comprise memory B cells and resting dendritic cells. In particular, we found a consecutive decrease of monocytes (Figure 6B), and a significant decrease of M0 macrophages from LC to HCC stage was observed (Figure 6C). In addition, an increase of M1 macrophages, and resting dendritic cells during the progress of hepatocarcinogenesis (Figures 6D and 6E), which might indicate that macrophage activation erupts during the transition from LC to HCC. These results suggest that initiation of liver cancer might stimulate the differentiation of monocytes into activated macrophages and dendritic cells. Furthermore, we found a co-regulation of activated mast cells, CD8 T cells, and follicular helper T cells, wherein the decline of cell abundance in LC stage was followed by an increase in HCC (Figure 6F). The decrease of these immune cells in LC stage might indicate a repaired microenvironment or immune evasion in LC stage. Additionally, we also observed a shift in the immune status of nature killer cells (Figure 6G). Specifically, an abundant decline in activated natural killer (NK) cells was accompanied by an increase in resting NK cells in LH stage, which was followed by a sharp increase of activated NK cells and a concurrent decline of resting NK cells in the LC stage (Figure 6G). To validate the performance of cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT) on

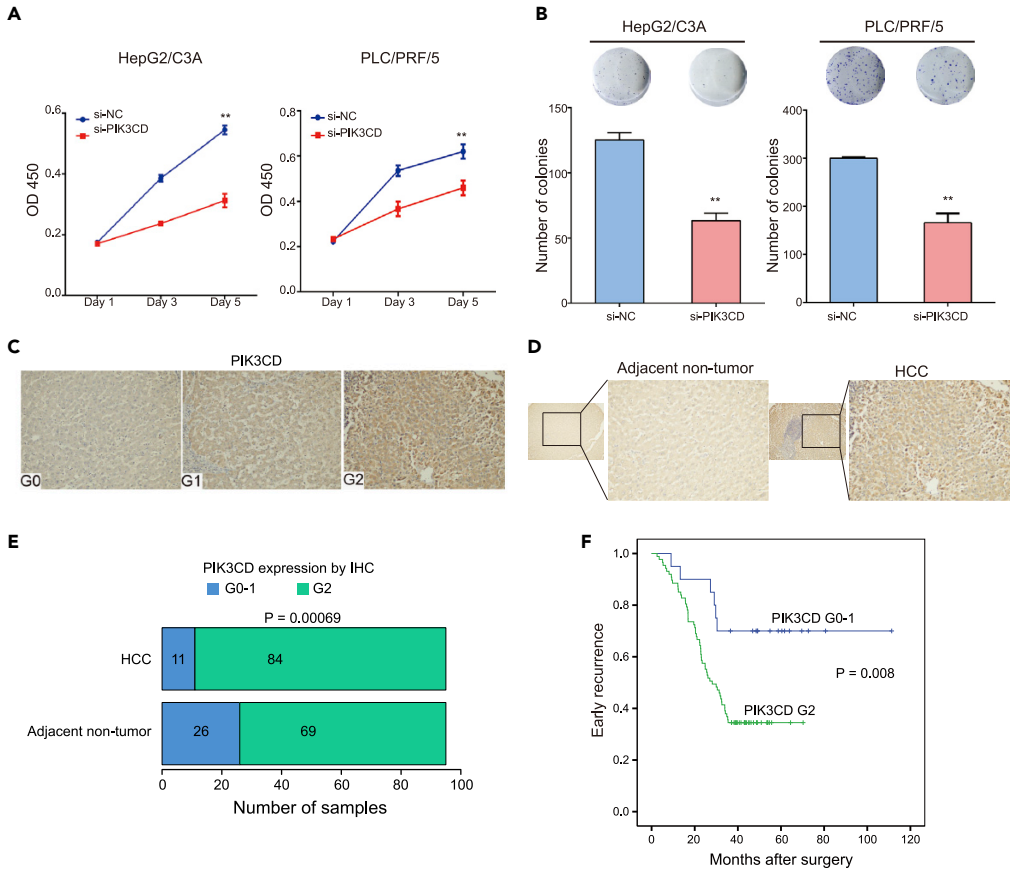


Figure 5. Biological and Clinicopathological Significance of PIK3CD in HCC

(A) Representative CCK-8 assays showing proliferation of HepG2/C3A and PLC/PRF/5 cells after PIK3CD knockdown.

(B) Colony formation assay showing the effect of PIK3CD knockdown on proliferation of HepG2/C3A and PLC/PRF/5 cells.

(C) Representative images of PIK3CD immunohistochemistry of liver biopsies from HCC patients.

(D) Representative case of PIK3CD expression is shown for HCC and adjacent non-tumor tissues. Original magnification, ×400.

(E) IHC analysis of PIK3CD in 118 paired HCC and adjacent non-tumor tissues. The p-value was determined by McNemar test.

(F) Kaplan-Meier analysis of the correlation between PIK3CD expression and recurrence-free survival of 95 HCC patients. Log rank tests were used to determine statistical significance. (Data are represented as mean ± SEM, **p < 0.01).

See also Figure S7.

tumor-infiltrating leukocytes (TILs) cell subsets, we performed immunohistochemistry (IHC) to evaluate the expression of CD14 and CD56 on TILs. We found difference in the density of CD14 and CD56 positive cells among the four developments stages. Consistent with the immune gene expression evolution, CD14 positive monocytes/macrophages showed an increase during the progress of hepatocarcinogenesis. Additionally, we also observed an abundant decline of CD56 positive NK cells in LC stage, which was followed by a sharp increase of NK cells in HCC stage (Figure S8). The shift of NK status mirrors the dynamic response of immune system to the progress of hepatocarcinogenesis. Taken together, our analysis revealed dynamic immune configurations during the development of liver cancer. The differentiation of monocytes into dendritic cells and macrophages during hepatocarcinogenesis suggest an immune response mechanism for the initiation liver cancer.

DISCUSSION

HCC is an inflammation-driven disease, 90% of which develop due to underlying chronic liver inflammation and the induction of subsequent cirrhosis. The liver has the ability to repair itself after acute damage. In chronic hepatitis stage, constant cell death, compensatory regeneration, and activation of other cells,

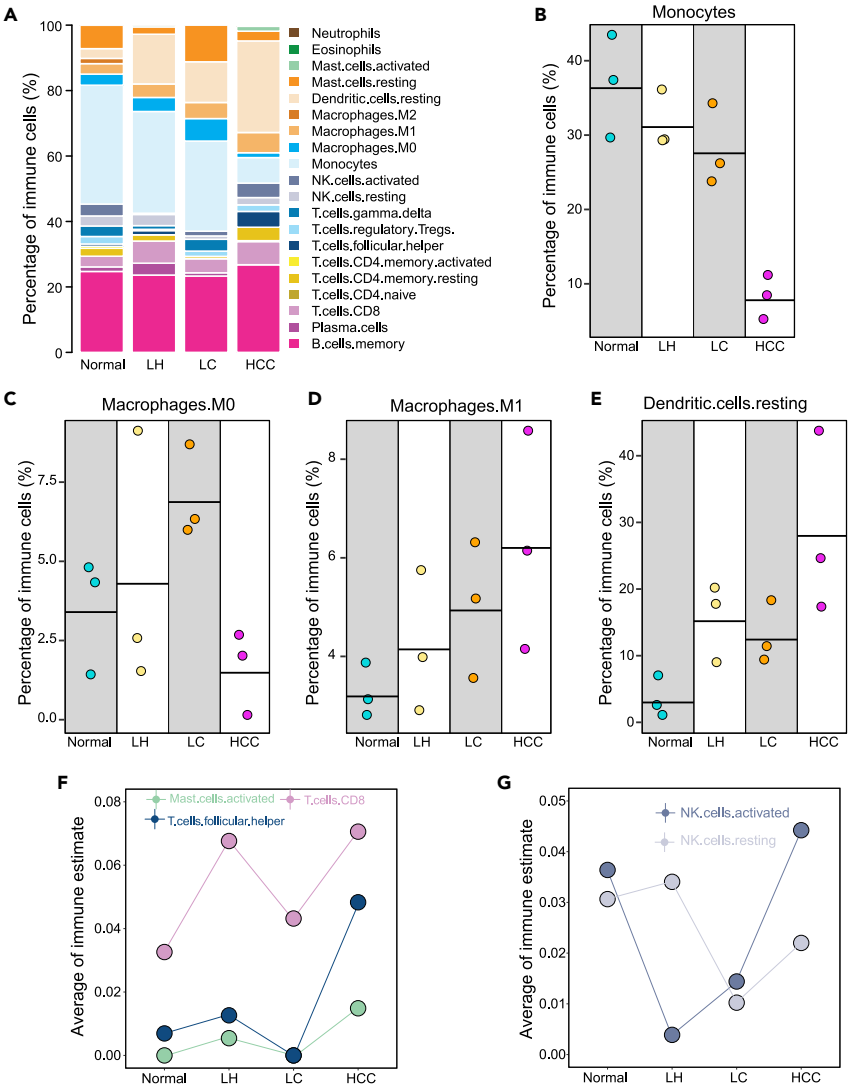


Figure 6. Dynamic Immune Response during Hepatocarcinogenesis

(A) The abundance estimation of different immune cell types in LH, LC, and HCC stages.

(B–E) (B) Monocytes show continuous decrease during liver cancer development. M0 macrophages (C), M1 Macrophages (D), and resting dendritic cells (E) show increase during hepatocarcinogenesis.

(F) Activated mast cells, CD8 T cells, and follicular helper T cells show co-regulation during hepatocarcinogenesis.

(G) The shift of immune status for NK cells during hepatocarcinogenesis.

See also Figure S8.

promote LC and tumorigenesis. Connor et al. compared mutation signatures between liver tumors for DEN-treated and untreated mice, and human HCCs by using mouse model of DEN-initiated HCC (Connor et al., 2018). They provided insights into the mutational landscape of DEN-initiated HCC mouse model, which facilitated the better understanding of this disease model. However, dynamic molecular changes during the progressive development from normal liver to HCC still remain vague. In the present study, we characterized the dynamic molecular changes underlying hepatocarcinogenesis from our rat model established for stepwise carcinogenesis stages of hepatocarcinoma, including hepatitis, cirrhosis, and liver cancer.

During the process of hepatocarcinogenesis in our rat model, chronic inflammation preceded cirrhosis and ultimately led to carcinoma. Interestingly, the most frequent SNVs in the hepatitis and cirrhosis stages were NCG>T (or the complement, CGN>A). NCG>T is a signature of genomic instability in cancer and normal

cells (Alexandrov et al., 2013; Welch et al., 2012), thus indicating that genomic instability occurred in the stage of LH. Genomic instability and mutation are clearly enabling characteristics that are causally associated with the acquisition of hallmark capabilities, which endow cancer cells with genetic alterations that drive tumor progression (Hanahan and Weinberg, 2011). Therefore, this mutability observed in LH may be achieved through increased sensitivity to mutagenic agents. We identified two mutation signatures (Figure 2B) in DEN rat HCC samples, i.e. HCC S1 and HCC S2. The mutation profile of HCC S2 is quite similar to that of DEN mouse HCC as shown in Figure 4A of Connor et al. paper, which are both characterized with C>T, T>A and T>C. Compared to DEN mouse HCC, HCC S1 lacks T>A mutations, which may due to the heterogeneity of hepatocarcinogenesis. In addition, we compared cancer genes that presented in Connor et al. paper between DEN rat and mouse. We found 46 cancer genes that were mutated in DNE mouse but not DEN rat, such as Gata3 and Notch1 (Table S6). Comparison that include large cohorts of rat and mouse HCC samples will reveal more mutation differences.

p53 signaling pathway is a major tumor-suppressor pathway, the activation of this signaling pathway also indicated that the aberrant cell survival or proliferation signals and disorder of genome integrity had already occurred in the early stage. Our results suggest that the genomic instability in LH activated the p53 signaling pathway and initiated DNA repair, thereby maintaining or reducing the number of somatic mutations in the cirrhosis stage. One striking observation is that an oncogenic pathway, PI3K-Akt/FoxO signaling pathway, has been already activated in the hepatitis stage. Activation of the Akt pathway suppresses transforming growth factor β-induced apoptosis and growth-inhibitory activity of CCAAT/enhancer binding protein α. These results suggested this effect may promote tumor formation at the very early stage. Furthermore, integrative analysis of gene mutations and expression identified PI3K-Akt pathway as one of the oncogenic networks. An activation of PI3K-Akt signaling and impaired expression of phosphatase and tensin homolog (a negative regulator of Akt) have been reported in 40–60% of human HCC. Although the activation of oncogenic pathways in human HCC appears to be more heterogeneous compared with other cancer types, activation of PI3K-Akt signaling pathway may occur early and throughout in human hepatocarcinogenesis.

Increasing evidences suggest that a dysregulated immune system, including changes in the number or function of immune cells significantly contribute to the development of HCC. Our data show that the dynamic response of different immune cell types, reflecting the battle between immune surveillance and cancerous lesions, occurs at early stages of pre-malignant transformation and the initiation of carcinoma, which supports the hypothesis of immune surveillance in pre-cancerous lesions. It has been reported that macrophages play a critical role in hepatic progenitor cell (HPC) expansion to chronic liver injury and HPC-mediated hepatocellular regeneration (Boulter et al., 2012; Viebahn et al., 2010). Our study revealed decrease of monocytes and increase of macrophage subtypes during the progress of hepatocarcinogenesis, reflecting the activation of the hepatocellular regeneration and contribute to chronic inflammation. Tumor-related cytotoxic T cells, $CD4^+$ T cells, NK cells, and the cross talk between them have all been reported to be involved in the development of HCC (Cariani and Missale, 2019; Ringelhan et al., 2018). Co-regulation of immune cells from both innate and adaptive immunity showed significant decline of most LH-increased immune cells in LC groups, which suggests that T cell and NK cell exhaustion contribute to the escape of liver cancer cells from immune surveillance might occur in LC stage. Our observation proposed an immunological intervention in LC stage, but further study is needed for the precise immunotherapy to prevent liver cancer. Resting or immature dendritic cells induce tolerance in T lymphocytes (Abbas and Sharpe, 2005). The continuous increased resting dendritic cells (DCs) may shift the immune response toward tumor tolerance, resulting in the pre-malignant transformation.

In conclusion, we report a comprehensive genomic and transcriptomic analysis of DEN-induced rat HCC resulting from hepatitis, cirrhosis of hepatocytes, and their malignant transformation, in which the earliest molecular alterations affect proliferation and metabolism. Integrated analyses revealed several key signaling pathways and biological processes in the pathogenesis of HCC. These observations have a potential clinical impact to guide the identification of novel therapeutic targets and/or stratify patients with a high risk of malignant transformation who may require more aggressive treatment. In addition, models that encompass heterogeneity in carcinogen exposure may also be useful for the preclinical screening of cancer therapeutics.

Limitations of the Study

In this work, we comprehensively characterize the molecular changes during hepatocarcinogenesis by using a rat HCC model. One limitation of our study is the relatively small sample size of rat model. Tracing molecular changes during hepatocarcinogenesis in human is a huge challenge, and re-building rat HCC model is time consuming. Our proposed analytical strategy is worth testing in larger sample size.

Resource Availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Xianghuo He (xhhe@fudan.edu.cn).

Materials Availability

No unique reagents have been generated in this study and all the materials are commercially available.

Data and Code Availability

The raw DNA sequencing reads were deposited in Sequence Read Archive database under the accession code number BioPPRJNA589289. The microarray expression data was deposited in the Gene Expression Omnibus database under accession code number GSE141090.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101690.

ACKNOWLEDGMENTS

Our work was supported by grants from the National Natural Science Foundation of China (81972247, 81790252), the State Key Project on Infectious Diseases of China (2018ZX10723204) and the Natural Science Foundation of Shanghai (18ZR1407900). We thank Drs. Zhaoping Qiu, Yan Shen, and Ruopeng Zha for their help with animal treatment. We would like to thank Genergy Bio Technology (Shanghai, China) and Capital-Bio Corporation (Beijing, China) for the operation of Illumina library construction, Illumina sequencing and the Affymetrix gene-chips.

AUTHOR CONTRIBUTIONS

Z.C., S.L., S.H., L.H., and X.H. designed and carried out bioinformatics analyses; Z.C. and M.S. performed and analyzed the primary cell experiments; M.S., D.C., and X.L. performed IHC; M.S., Q.W., X.L., and W.C. analyzed IHC data; Z.C. and J.D. performed animal experiments; Z.C., S.L., L.H., and X.H. wrote the paper with comments from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 13, 2020
Revised: September 17, 2020
Accepted: October 13, 2020
Published: November 20, 2020

REFERENCES

Abbas, A.K., and Sharpe, A.H. (2005). Dendritic cells giveth and taketh away. Nat. Immunol. *6*, 227–228.

Ahn, S.M., Jang, S.J., Shim, J.H., Kim, D., Hong, S.M., Sung, C.O., Baek, D., Haq, F., Ansari, A.A., Lee, S.Y., et al. (2014). Genomic portrait of resectable hepatocellular carcinomas: implications of RB1 and FGF19 aberrations for patient stratification. Hepatology *60*, 1972–1982.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.L., et al. (2013). Signatures of mutational processes in human cancer. Nature *500*, 415–421.

Boulter, L., Govaere, O., Bird, T.G., Radulescu, S., Ramachandran, P., Pellicoro, A., Ridgway, R.A., Seo, S.S., Spee, B., Van Rooijen, N., et al. (2012).

Macrophage-derived Wnt opposes Notch signaling to specify hepatic progenitor cell fate in chronic liver disease. Nat. Med. *18*, 572–579.

Cariani, E., and Missale, G. (2019). Immune landscape of hepatocellular carcinoma microenvironment: implications for prognosis and therapeutic applications. Liver Int. *39*, 1608–1621.

CavigLia, J.M., and Schwabe, R.F. (2014). Experimental Hepatocarcinogenesis, Pathobiology of Human Disease: A Dynamic Encyclopedia of Disease Mechanisms (Elsevier Inc).

Connor, F., Rayner, T.F., Aitken, S.J., Feig, C., Lukk, M., Santoyo-Lopez, J., and Odom, D.T. (2018). Mutational landscape of a chemically-induced mouse model of liver cancer. J. Hepatol. *69*, 840–850.

El-Serag, H.B. (2012). Epidemiology of viral hepatitis and hepatocellular carcinoma. Gastroenterology *142*, 1264–1273.e1.

El-Serag, H.B., and Rudolph, K.L. (2007). Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology *132*, 2557–2576.

Fujimoto, A., Totoki, Y., Abe, T., Boroevich, K.A., Hosoda, F., Nguyen, H.H., Aoki, M., Hosono, N., Kubo, M., Miya, F., et al. (2012). Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nat. Genet. *44*, 760–764.

Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I.B., Calderaro, J., Bioulac-Sage, P., Letexier, M., Degos, F., et al. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat. Genet. *44*, 694–698.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Harding, J.J., Nandakumar, S., Armenia, J., Khalil, D.N., Albano, M., Ly, M., Shia, J., Hechtman, J.F., Kundra, R., Dika, I. El, et al. (2019). Prospective genotyping of hepatocellular carcinoma: clinical implications of next-generation sequencing for matching patients to targeted and immune therapies. Clin. Cancer Res. *25*, 2116–2126.

Jenne, C.N., and Kubes, P. (2013). Immune surveillance by the liver. Nat. Immunol. *14*, 996–1006.

Li, S., Hu, Z., Zhao, Y., Huang, S., and He, X. (2019). Transcriptome-wide analysis reveals the landscape of aberrant alternative splicing events in liver cancer. Hepatology *69*, 359–375.

Llovet, J.M., Zucman-rossi, J., Pikarsky, E., Sangro, B., Sherman, M., and Gores, G. (2016). Hpatocellular carcinoma. Nat. Rev. Dis. Primers *2*, 16018.

Magee, P.N., and Barnes, J.M. (1956). The production of malignant primary hepatic tumours in the rat by feeding dimethylnitrosamine. Br. J. Cancer *10*, 114–122.

Makarova-Rusher, O.V., Medina-Echeverz, J., Duffy, A.G., and Greten, T.F. (2015). The yin and yang of evasion and immune activation in HCC. J. Hepatol. *62*, 1420–1429.

Notas, G., Kisseleva, T., and Brenner, D. (2009). NK and NKT cells in liver injury and fibrosis. Clin. Immunol. *130*, 16–26.

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics *20*, 289–290.

Ringelhan, M., Pfister, D., O'Connor, T., Pikarsky, E., and Heikenwalder, M. (2018). The immunology of hepatocellular carcinoma. Nat. Immunol. *19*, 222–232.

Robinson, M.W., Harmon, C., and O'Farrelly, C. (2016). Liver immunology and its role in inflammation and homeostasis. Cell. Mol. Immunol. *13*, 267–276.

Schulze, K., Imbeaud, S., Letouzé, E., Alexandrov, L.B., Calderaro, J., Rebouissou, S., Couchy, G., Meiller, C., Shinde, J., Soysouvanh, F., et al. (2015). Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat. Genet. *47*, 505–511.

Sottoriva, A., Kang, H., Ma, Z., Graham, T.A., Salomon, M.P., Zhao, J., Marjoram, P., Siegmund, K., Press, M.F., Shibata, D., et al. (2015). A big bang model of human colorectal tumor growth. Nat. Genet. *47*, 209–216.

Tao, Y., Ruan, J., Yeh, S.-H., Lu, X., Wang, Y., Zhai, W., Cai, J., Ling, S., Gong, Q., Chong, Z., et al. (2011). Rapid growth of a hepatocellular carcinoma and the driving mutations revealed by cell-population genetic analysis of whole-genome data. Proc. Natl. Acad. Sci. U S A *108*, 12042–12047.

Tate, J.G., Bamford, S., Jubb, H.C., Sondka, Z., Beare, D.M., Bindal, N., Boutselakis, H., Cole, C.G., Creatore, C., Dawson, E., et al. (2019). COSMIC: the Catalogue of somatic mutations in cancer. Nucleic Acids Res. *47*, D941–D947.

Totoki, Y., Beroukhim, R., Lee, S., Heiman, D., Hsu, T.-K., Saggioro, F.P., Sadeghi, S., Liu, X., Stuart, J., Akbani, R., et al. (2017). Comprehensive and integrative genomic characterization of hepatocellular carcinoma. Cell *169*, 1327–1341.e23.

Viebahn, C.S., Benseler, V., Holz, L.E., Elsegood, C.L., Vo, M., Bertolino, P., Ganss, R., and Yeoh, G.C.T. (2010). Invading macrophages play a major role in the liver progenitor cell response to chronic liver injury. J. Hepatol. *53*, 500–507.

Villanueva, A. (2019). Hepatocellular carcinoma. N. Engl. J. Med. *380*, 1450–1462.

Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., et al. (2012). The origin and evolution of mutations in acute myeloid leukemia. Cell *150*, 264–278.

Yu, J., Shen, J., Sun, T.T., Zhang, X., and Wong, N. (2013). Obesity, insulin resistance, NASH and hepatocellular carcinoma. Semin. Cancer Biol. *23*, 483–491.

iScience, Volume *23*

Supplemental Information

The Mutational and Transcriptional Landscapes of Hepatocarcinogenesis in a Rat Model

Zhiao Chen, Shengli Li, Mengting Shen, Xinyuan Lu, Chunyang Bao, Di Chen, Jie Ding, Qifeng Wang, Shenglin Huang, Wenming Cong, Leng Han, and Xianghuo He

Transparent Methods

Generation of hepatocarcinoma rat models

Male Sprague-Dawley (SD) rats weighing 120 to 150 g at the beginning of the experiments were obtained from SLAC Laboratory Animal Co. Ltd. (Shanghai, PR China). The animals were acclimatized to standard laboratory conditions (temperature 22-25°C, relative humidity 50-60%, and 12-hour photoperiods (lights on 07:00-19:00)) and were housed in stainless steel wire-mesh cages (3 rats per cage). During the entire period of study, the rats were supplied with a semipurified basal diet and water ad libitum. All experiments followed the Guide for the Care and Use of Laboratory Animals.

Collecting rat samples

Briefly, 50 SD rats were randomly divided into two groups: the control (14 rats) and diethylnitrosamine (DEN, Sigma Chemical Co. St Louis, MO) groups (36 rats; 12 for the hepatitis group, 12 for the cirrhosis group, and 12 for the carcinoma group). After 1 week on a basal diet, rats belonging to the DEN group were injected intraperitoneally with a dosage of 70 mg/kg of DEN once per week for 10 weeks. Animals that belonged to the control group were injected with distilled water. The rats were sacrificed with 1.5% (g/ml) pentobarbital sodium (30 mg/kg) by intraperitoneal injection at different time points. At the end of the 5^{th} (hepatitis group), $13/14^{th}$ (cirrhosis group) and $23/24^{th}$ (carcinoma group) weeks after DEN-treatment, 12 rats were sacrificed, respectively. The entire liver of each rat was observed grossly and weighed. The samples from each rats were collected and divided into pieces. Some samples were snap-frozen directly in liquid nitrogen and stored at -80°C prior to DNA and RNA isolation for microarray analysis. The remaining pieces were preserved in 10% phosphate-buffered formalin. The liver tissue fixed in neutral formalin was embedded in paraffin, sectioned, and stained with haematoxylin and eosin (H&E). Histopathologic examinations of the liver sections were conducted by a pathologist.

DNA/RNA extraction

DNA was isolated using the Qiagen reagents and total RNA was isolated using TRIzol reagent (Invitrogen, Carlsbad, CA). Genomic DNA from tumour samples weighing up to 10 mg was extracted using the QIAamp DNA Micro Kit (Qiagen, Waltham, MA) following the manufacturer's instructions.

Illumina library construction and sequencing

To characterize the somatic mutation spectrum during hepatocarcinogenesis, we performed whole-exome sequencing (WES) on 12 liver tissues, including 3 from control group, 3 from liver hepatitis group (LH), 3 from liver cirrhosis group (LC), and 3 from HCC group. All methods for library construction were performed as described in the SeqCap EZ Library protocols (Roche NimbleGen, Madison, WI). Genomic libraries were prepared using the SepCap EZ Developer Library for rats per the manufacturer's instructions (Roche NimbleGen). Each exome was sequenced using the 100 bp paired-end protocol on an Illumina HiSeq 2000 DNA Analyzer to produce approximately 12 to 20 Gb of sequence per exome.

Processing DNA sequencing data

For each sample, low-quality bases were removed from the raw sequencing reads by using Trimmomatic (version 0.38)(Bolger et al., 2014). Then the trimmed reads were aligned to the reference rat genome (rn6) using BWA (version 0.7.17)(Li et al., 2010). After alignment, MarkDuplicates program (version 2.18.27) in Picard was employed to mark the duplicate reads. We then recalibrated the base quality scores and realigned reads around indels using GATK (version 3.8-0)(McKenna, Aaron, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, David Altshuler, Stacey Gabriel et al., 2010). Alignment and coverage metrics were then collected. Specifically, we sequenced an average of 102 million unique on-target reads per sample and the targeted exonic regions were sequenced to an average depth of 142x.

Identification and annotation of SNVs

SNVs were identified using the somatic variant detection program MuTect2 in GATK (version 3.8-0). Each sample was called against DNA taken from normal livers isolated from the control group. Rn6 served as the reference genome during the mutation calling process. Each set of variants was then used to generate a subset of variants that passed MuTect2 filters and had a minimum read depth of 10. The associated data were then called for variants using MuSE (version 1.0rc)(Fan et al., 2016), SomaticSniper (version 1.0.5.0)(Larson et al., 2012) and VarScan2 (version 2.3.9)(Wilson et al., 2012). Variants from each caller were then filtered for sites with a minimum quality of 20 and minimum depth of 10. Variants called by a minimum of two softwares were filtered for downstream analysis. Surviving variants were annotated using SnpEff (version 4.1)(Cingolani et al., 2012).

Mutation spectra analysis

SNVs in all samples were annotated on the basis of the 96 possible tri-nucleotide context substitutions (6 types of substitutions × 4 possible flanking 5' bases × 4 possible flanking 3' bases) and summed for each sample, thus creating a matrix of 96 substitutions. The *signeR* package, which employed an empirical Bayesian treated non-negative matrix factorization (NMF) approach for mutational signature discovery(Rosales et al., 2017), was applied to mutation counts to infer the number of operating mutational signatures and the mutational signatures (96 normalized weights per process) in each sample.

Gene expression analysis

Total RNAs from three samples of each group were extracted for the GeneChip analysis. Biotin-labelled cRNA samples were used for hybridization of Affymetrix GeneChip Rat Genome 230 2.0 arrays. The arrays were prepared according to the protocol supplied with the GeneChip Sample Clean up module (P/N 900371, Affymetrix Inc., Santa Clara, CA). The microarray expression data was deposited in the Gene Expression Omnibus (GEO) database under accession GSE141090.The probes from the same genes were merged on the basis of their average values. The differential analysis of gene expression between different stages was conducted by Wilcoxon rank sum test. Genes with fold-change > 1.5 and P value < 0.05 were considered to be statistically differentially expressed.

Cell lines and cell culture

Two liver cancer cell lines were used: HepG2/C3A (ATCC, ATCC Number: CRL-10741), and PLC/PRF/5 (ATCC, ATCC Number: CRL-8024). Cells were cultured in DMEM supplemented with 10% fetal bovine serum, 100 U/ml penicillin, and 100 μg/ml streptomycin. Cells were regularly certified as free of mycoplasma contamination.

Cell proliferation and colony formation assays

Cell proliferation was monitored by counting viable cells using the Cell Counting Kit (CCK)-8 (Dojindo, Kumamoto, Japan). Cells were seeded at 1000 cells/well in 96-well plates and incubated. A 10 μl volume of the CCK-8 solution was added to the triplicate wells and incubated for 2 hours. Subsequently, the absorbance at 450 nm was measured to calculate the number of vital cells in each well. For colony formation assays, cells were trypsinized, resuspended, seeded into 6-well plates at a concentration of 500 cells per well, and cultured at 37°C with 5% CO_2 for 2 weeks. The media were changed every 3-4 days. At the end of the incubation, the cells were fixed with 100% methanol and stained with 0.1% crystal violet. Megascopic cell colonies were counted by Image-

Pro Plus 5.0 (Media Cybernetics, Bethesda, MD).

Transient transfection and RNA interference

Cells were transfected using the Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instructions. The total amount of transfected plasmid DNA was kept constant between experimental conditions by the addition of an empty-vector plasmid. For small interfering RNA (siRNA) transfection, cells seeded in antibiotic-free medium at 30 to 50% confluence were transfected with 50 nM siRNA duplexes by using RNAiMAX (Invitrogen) according to the manufacturer's instructions.

Human Tissue microarray and immunohistochemistry

The tissue microarray was constructed as described previously (Jia et al., 2010). Core samples were obtained from representative regions of each tumour based on H&E staining. Duplicate 1-mm cores were taken from different areas of the same tissue block for each case (intratumoural tissue and peritumoural tissue). Serial sections (4-μm-thick) were placed on slides coated with 3-aminopropyltriethoxysilane. The immunohistochemistry analysis was carried out as described previously. The primary mAbs used were rabbit anti-human PIK3CD (1:100, ThermoFisher Scientific, Waltham, MA), rabbit anti-rat CD14 (1:4000, proteintech, Wuhan, Hubei, China) and rabbit anti-rat NCAM (CD56) (1:400, Cell Signaling Technology, Danvers, MA, USA). The PIK3CD immunostaining intensities was scored semi-quantitatively. All samples were anonymously and independently scored by two investigators. In cases of disagreement, the slides were re-examined, and a consensus was reached by the observers.

This study also involved a set of 118 HCC patients from the Eastern Hepatobiliary Surgery Hospital (Shanghai, PR China). The patients had undergone curative resection for HCC, and the isolated tumour tissues were prepared on tissue microarray slides. Patients with HCC had first undergone radical resection of HCC, suffered a relapse a few years later, and then underwent a second resection of HCC. The follow-up procedures and postoperative treatments were based on a uniform guideline and have been described previously (Sho et al., 2009). Tumour differentiation was graded using the Edmondson grading system. Clinical staging was performed according to the 6^{th} edition of the AJCC/UICC TNM classification system. The data were censored at the last follow-up for patients without relapse or upon death. Institutional review board approval was obtained, and each patient provided written informed consent.

Calculation of hallmark scores

The hallmark gene sets were retrieved from the Molecular Signature Database (MSigDB) (Liberzon et al., 2015). The hallmark score for each sample was calculated based on gene set variation analysis (Hänzelmann et al., 2013). In particular, we adopted the Gene Set Variation Analysis (GSVA) method to estimate the variation of hallmark activities over each sample in an unsupervised manner by employing expression profiles of genes annotated in the corresponding hallmarks.

Network construction

We first extracted the network connections of signalling pathways from the KEGG database (Kanehisa et al., 2017). Then dysregulated genes with differential expression or non-synonymous mutations were mapped to these networks. Moreover, the topology analysis and visualization of the networks were conducted by Cytoscape (version 3.4.0) (Paul Shannon et al., 2003).

Functional enrichment analysis

The list of unique genes that harbored nonsynonymous mutations or showed differential expression was extracted to determine the enrichment in hallmark gene sets. A hypergeometric test was performed to calculate the enrichment significance. The probability P was computed to assess the enrichment significance as following (Li et al., 2019):

$$\mathrm{P} = 1 - \mathrm{F}(x|N, K, M) = 1 - \sum_{t=0}^{x} \frac{\binom{K}{t}\binom{N-K}{M-t}}{\binom{N}{M}}$$

where N is the total number of all protein-coding genes annotated in GENCODE (v28), K represents the number of genes in the hallmark under investigation, M indicates the number of mutated or differential genes for analysis, and x is the number of genes shared between investigated hallmark and mutated or differential genes.

Estimation of immune cell abundance

The CIBERSORT algorithm (Newman et al., 2015) was adopted to estimate the abundance of different immune cell types in each sample. Specifically, the predefined LM22 signature in CIBERSORT method was applied for immune cell deconvolution from gene expression. The LM22 signature is a leukocyte gene signature which includes 547 genes that could distinguish 22 types of hematopoietic cells.

Ethics approval and consent to participate

All experimental methods comply with the Helsinki Declaration. Liver cancer cell lines used in

this study were obtained by ATCC and Shanghai Second Military Medical University (Shanghai, China). All experiments involving animals followed the Guide for the Care and Use of Laboratory Animals. All work relating to human tissues was performed previously at Shanghai Second Military Medical University under an institutional review board-approval protocol, and each patient provided written informed consent.

Statistical analysis

The experiments were repeated at least three times, and the results are expressed as the means and the standard error of the mean (SEM). Student's t tests (two-tailed) and two-way analysis of variance (ANOVA) were used to compare the means of two or more samples unless otherwise indicated. The chi-squared ($\chi 2$) test was used to evaluate the different expression of PIK3CD between HCC and adjacent non-tumour tissues. Significant pathways were identified using the hypergeometric distribution model (hypergeom). The P-values for each analysis were adjusted using the Benjamini-Hochberg method for multiple-testing correction. The cumulative recurrence and survival rates were determined using the Kaplan-Meier method (log-rank test). All data was visualized in R using summary statistics and basic plotting functions before statistical testing, and variance was comparable in all cases in which the Wilcoxon rank-sum test was used. All reported P-values were two tailed, and differences were considered to be statistically significant when the P-value was less than 0.05.

Supplemental Tables

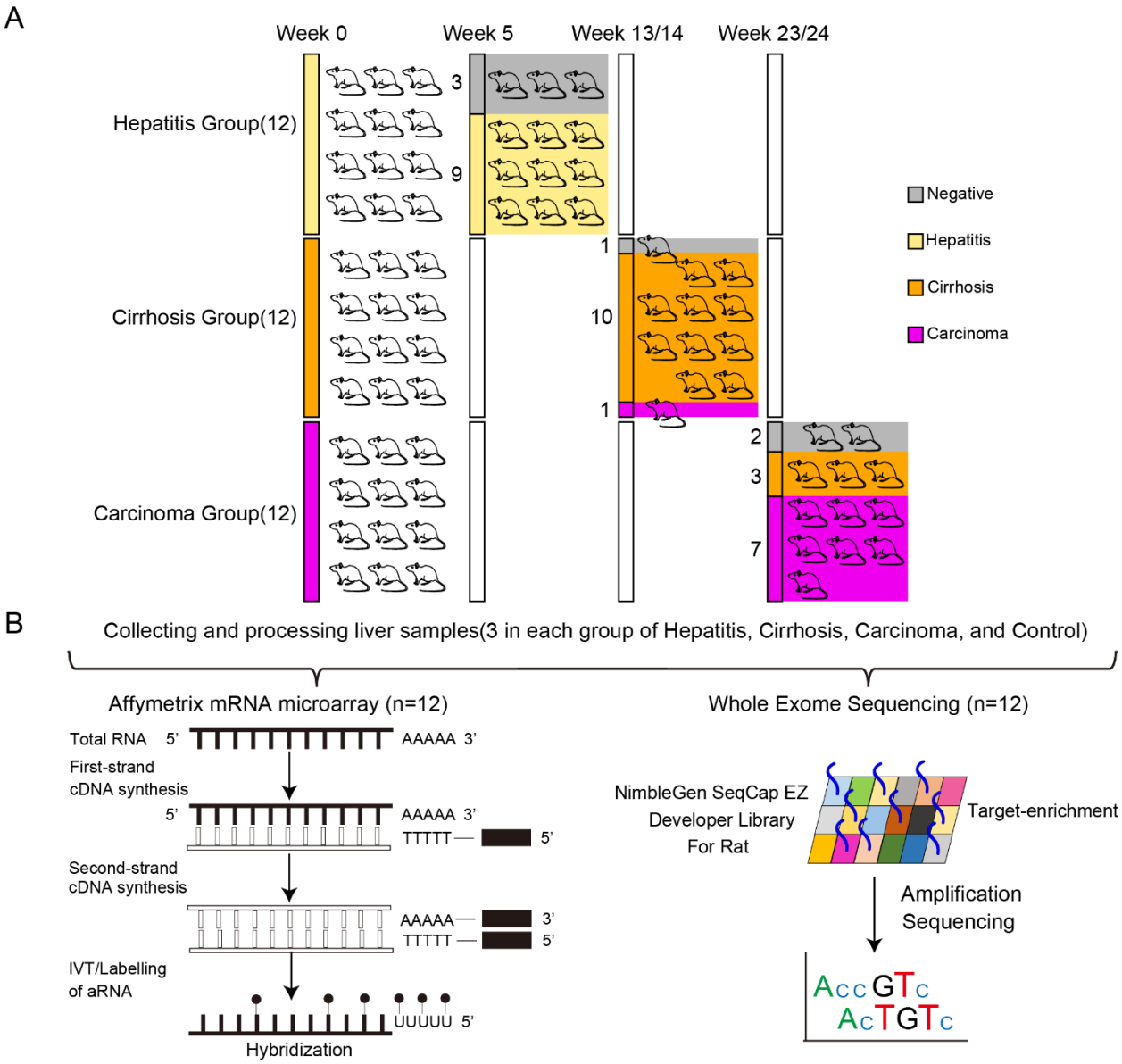
Table S6. The mutation status of 77 cancer genes in DEN rat and mouse HCC. Related to Figure 2.

Gene	DEN rat	DEN mouse
Cdc73	Mutated	Mutated
Gata3	Non-Mutated	Mutated
Notch1	Non-Mutated	Mutated
Tsc1	Non-Mutated	Mutated
Nfe2l2	Mutated	Mutated
Asxl1	Mutated	Mutated
Gnas	Non-Mutated	Mutated
Gm20721	Non-Mutated	Mutated
Pik3ca	Mutated	Mutated
Fbxw7	Mutated	Mutated
Notch2	Non-Mutated	Mutated
Tet2	Non-Mutated	Mutated
Fubp1	Non-Mutated	Mutated
Pax5	Mutated	Mutated
Klf4	Non-Mutated	Mutated
Cdkn2a	Non-Mutated	Mutated
Jak1	Mutated	Mutated
Arid1a	Mutated	Mutated
Kmt2c	Non-Mutated	Mutated
Pdgfra	Mutated	Mutated
Kit	Mutated	Mutated
Crlf2	Non-Mutated	Mutated
Card11	Non-Mutated	Mutated
Brca2	Mutated	Mutated
Met	Mutated	Mutated
Braf	Non-Mutated	Mutated

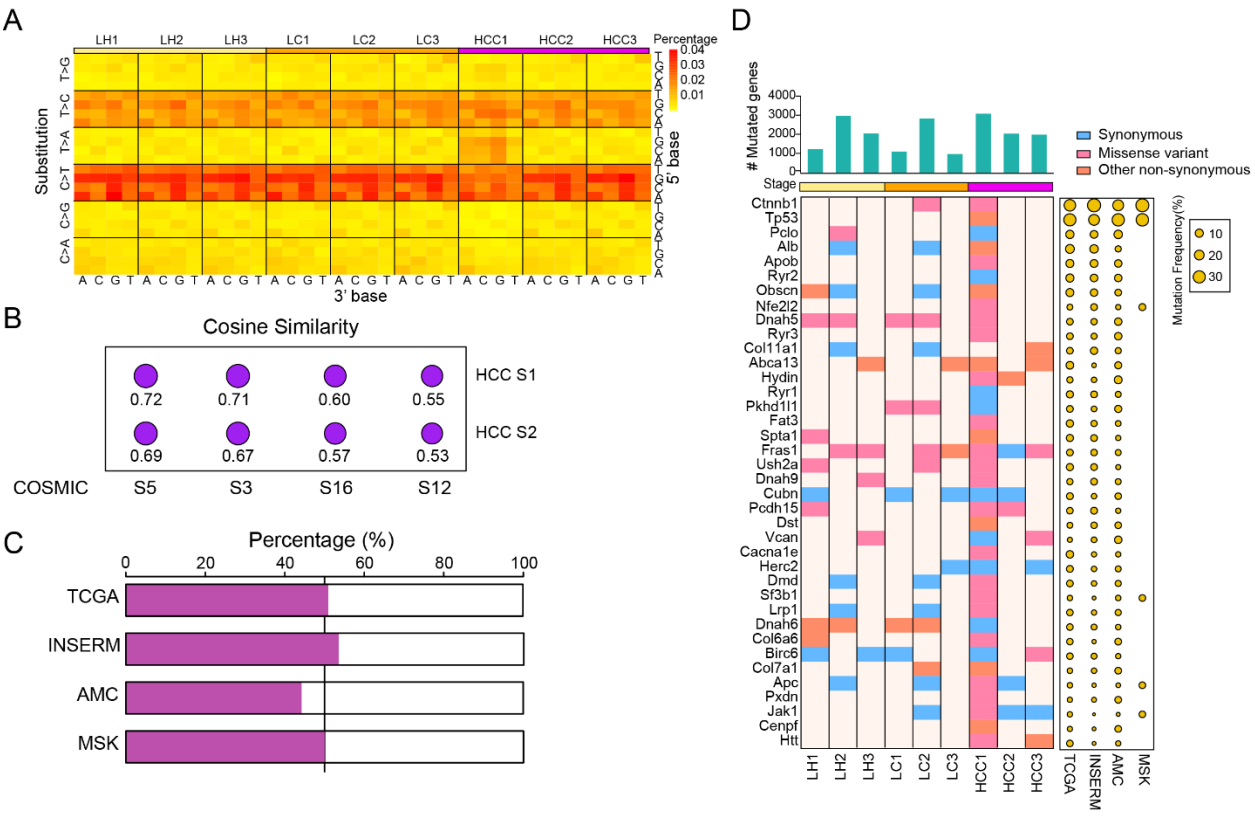
Ezh2	Non-Mutated	Mutated
Ret	Mutated	Mutated
Gata2	Mutated	Mutated
Kras	Non-Mutated	Mutated
Cic	Non-Mutated	Mutated
Cebpa	Non-Mutated	Mutated
Idh2	Non-Mutated	Mutated
Fgfr2	Non-Mutated	Mutated
Hras	Non-Mutated	Mutated
Jak3	Non-Mutated	Mutated
Cyld	Non-Mutated	Mutated
Cdh1	Mutated	Mutated
Dnmt1	Non-Mutated	Mutated
Smarca4	Non-Mutated	Mutated
Cbl	Non-Mutated	Mutated
Atm	Mutated	Mutated
Map2k1	Non-Mutated	Mutated
Setd2	Mutated	Mutated
Mlh1	Mutated	Mutated
Myd88	Non-Mutated	Mutated
Ctnnb1	Mutated	Mutated
Egfr	Mutated	Mutated
Ncor1	Mutated	Mutated
Nf1	Mutated	Mutated
Rnf43	Non-Mutated	Mutated
Erbb2	Non-Mutated	Mutated
Brca1	Mutated	Mutated
Tshr	Non-Mutated	Mutated
Map3k1	Non-Mutated	Mutated
Pbrm1	Non-Mutated	Mutated

Rb1	Mutated	Mutated
Ep300	Non-Mutated	Mutated
Arid2	Non-Mutated	Mutated
Kmt2d	Non-Mutated	Mutated
Crebbp	Mutated	Mutated
Runx1	Non-Mutated	Mutated
Arid1b	Mutated	Mutated
Ppp2r1a	Non-Mutated	Mutated
Traf7	Non-Mutated	Mutated
Axin1	Non-Mutated	Mutated
U2af1	Mutated	Mutated
Alk	Non-Mutated	Mutated
Msh6	Mutated	Mutated
Apc	Mutated	Mutated
Csf1r	Non-Mutated	Mutated
Smad2	Non-Mutated	Mutated
Setbp1	Mutated	Mutated
Men1	Non-Mutated	Mutated
Gnaq	Mutated	Mutated
Jak2	Non-Mutated	Mutated
Pten	Non-Mutated	Mutated

Supplemental Figures



Supplementary Figure S1. Collecting and processing rat liver samples. (**A**) Breakdown of different histological findings in each treatment group. (**B**) Genomic DNAs and transcriptomic RNAs from these samples were used for whole exome sequencing and microarray, respectively. Related to Figure 1.



Supplementary Figure S2. Mutational profiles for each stage in DEN-induced rat model of hepatocarcinogenesis. (**A**) Stacked heatmaps of mutation spectra for the LH, LC and HCC stages. Substitutions are shown on the left of each heatmap, while the 5'- and 3'-flanking base displayed down and right, respectively. (**B**) The cosine similarities between HCC mutation signatures and COSMIC signature 12, 16, 3, and 5. (**C**) Barplots show the percentage of mutated genes identified in TCGA, INSERM, AMC, and MSK cohorts. (**D**) The mutation landscape of recurrently mutated genes (>5% mutation frequency in at least two cohorts) in LH, LC and HCC stages. Barplots in upper panel show the number of mutated genes in each sample. Points in left panel indicate the mutation frequency of corresponding genes in TCGA, INSERM, AMC, and MSK datasets. Related to Figure 2.

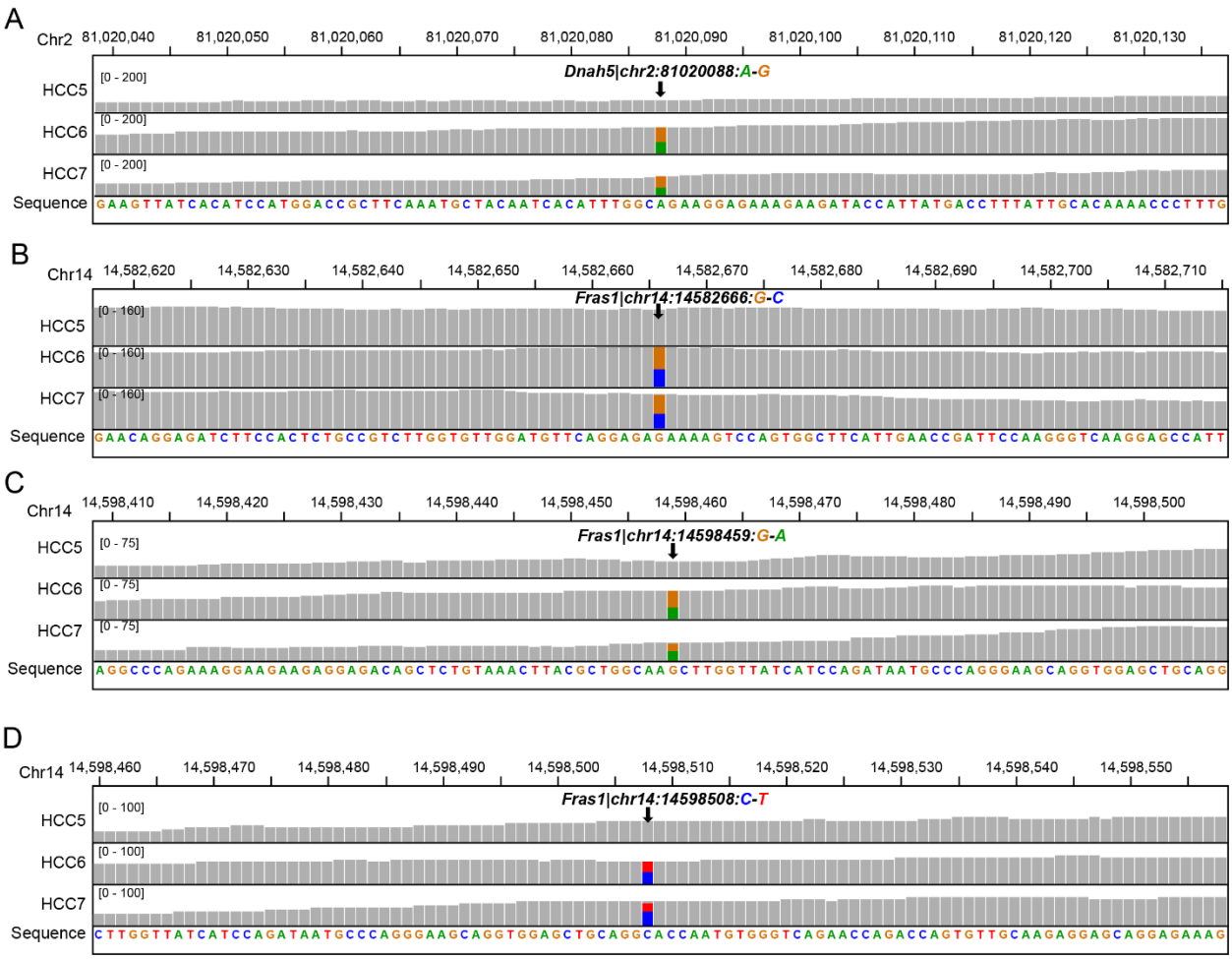
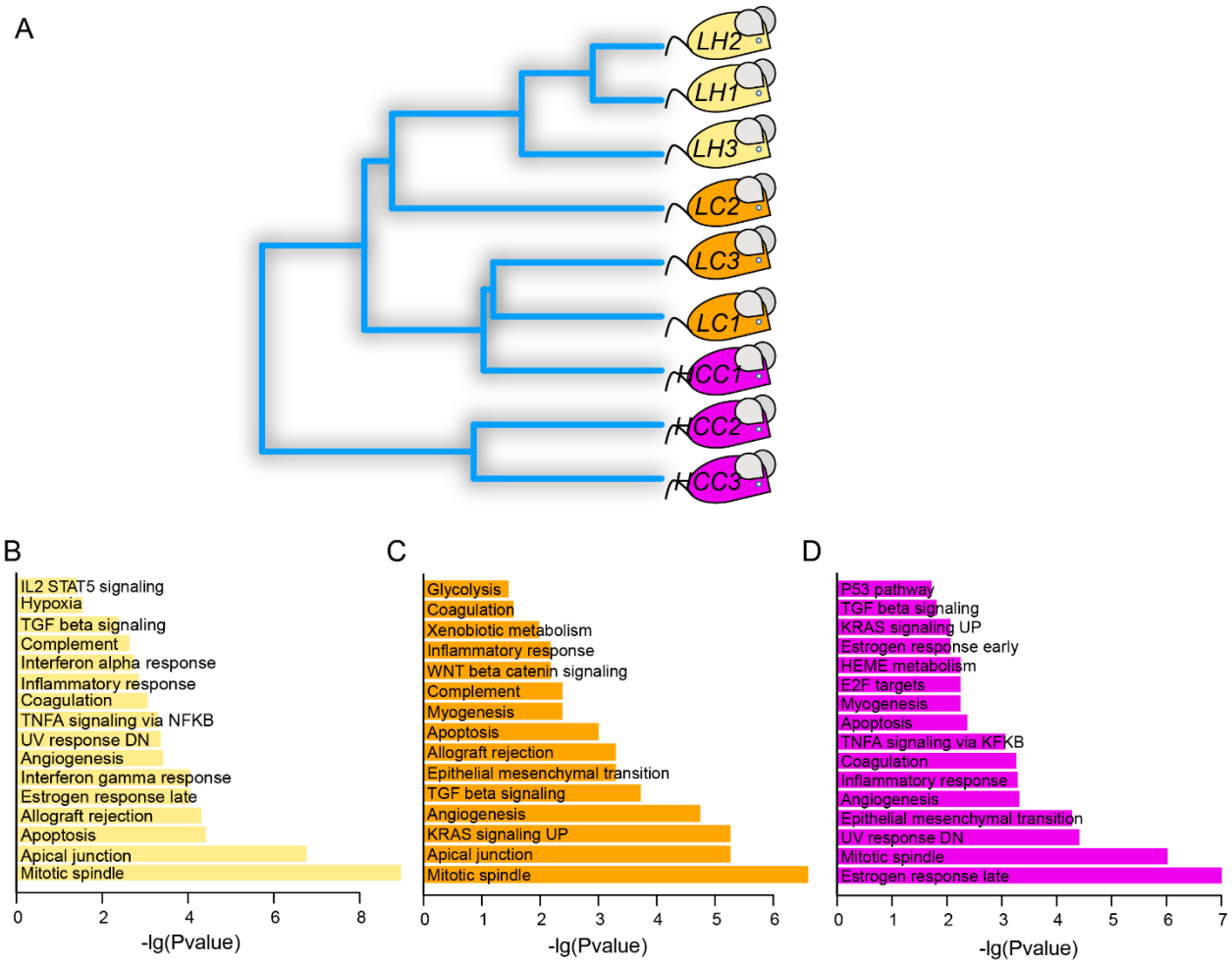


Figure S3. IGV visualization of mutation points of specific genes, i.e. Dnah5|chr2:81020088:A-G (**A**), Fras1|chr14:14582666:G-C (**B**), Fras1|chr14:14598459:G-A (**C**), Fras1|chr14:14598508:C-T (**D**). Bars indicate reads coverage mapped in the corresponding loci. Related to Figure 2.



Supplementary Figure S4. The mutational relationship and enrichment analysis of mutated genes among different stages. (**A**) Phylogenetic analysis of DEN treatment rat samples. The functional enrichment analysis of genes wiht non-synonymous mutations in LH stage (**B**), LC stage (**C**) and HCC stage (**D**). LH, liver hepatitis; LC, liver cirrhosis; HCC, hepatocellular carcinoma. Related to Figure 2.

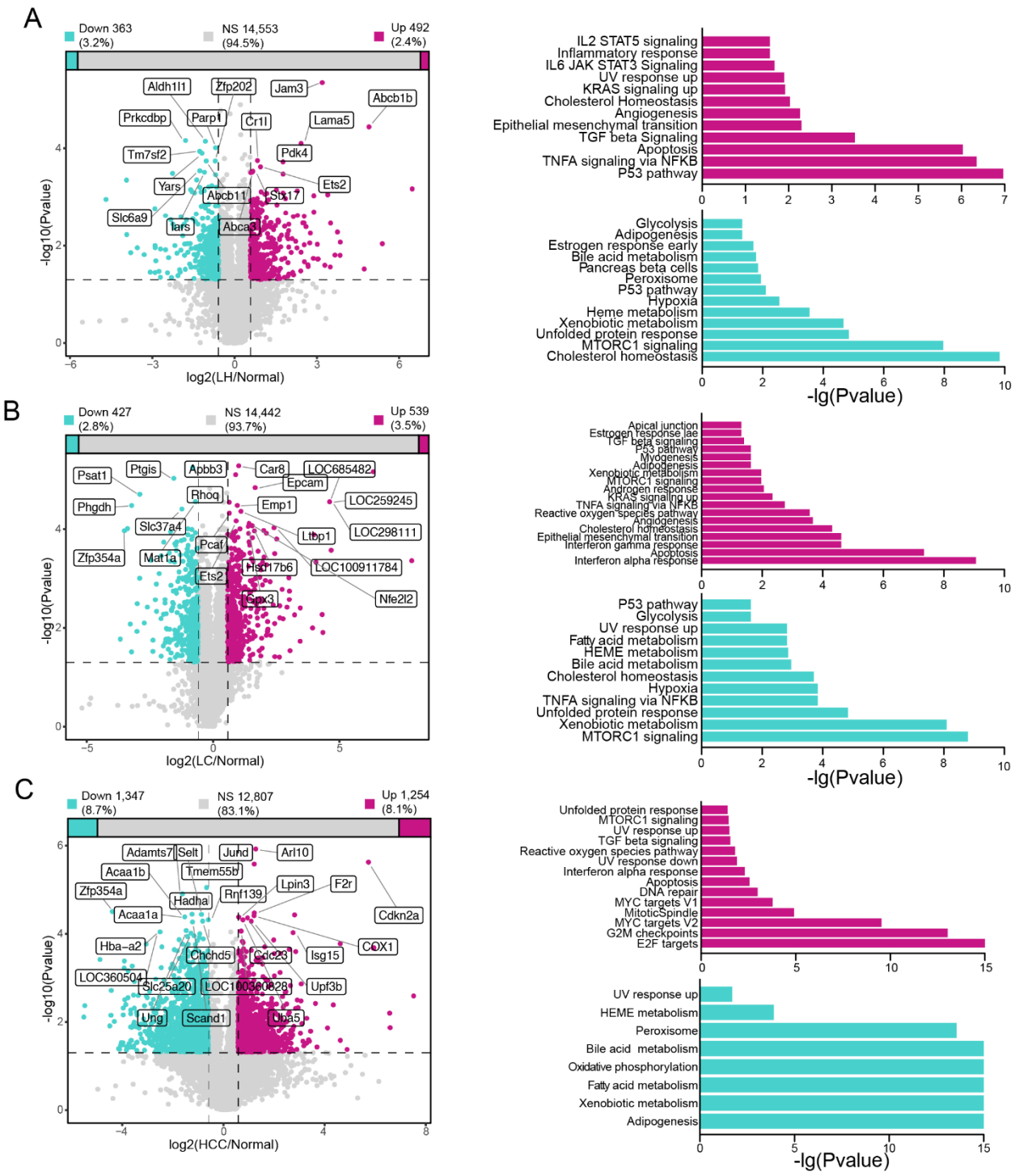


Figure S5. The significance and functional enrichment analysis of differentially expressed genes in LH stage (**A**), LC stage (**B**), and HCC stage (**C**). Related to Figure 2.

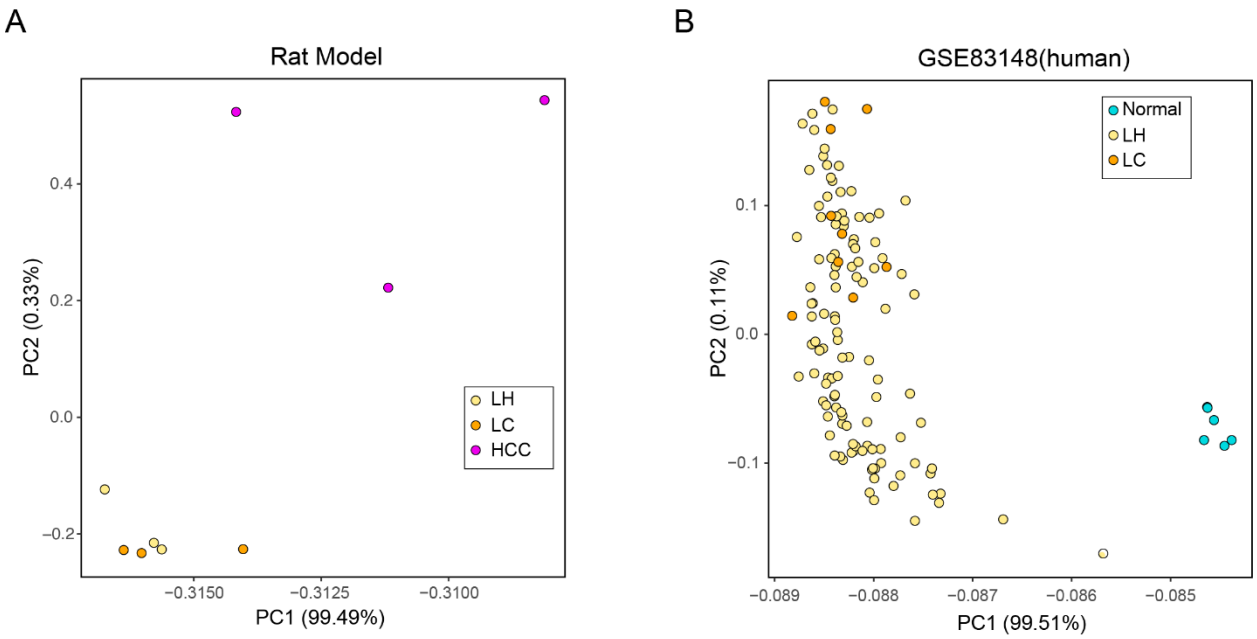


Figure S6. PCA analysis of gene expression among different stages. (**A**) PCA analysis of samples in LH, LC and HCC stages from DEN-induced rats. (**B**) PCA analysis of samples in normal, LH and LC stages from a human cohort (GSE83148). Related to Figure 3.

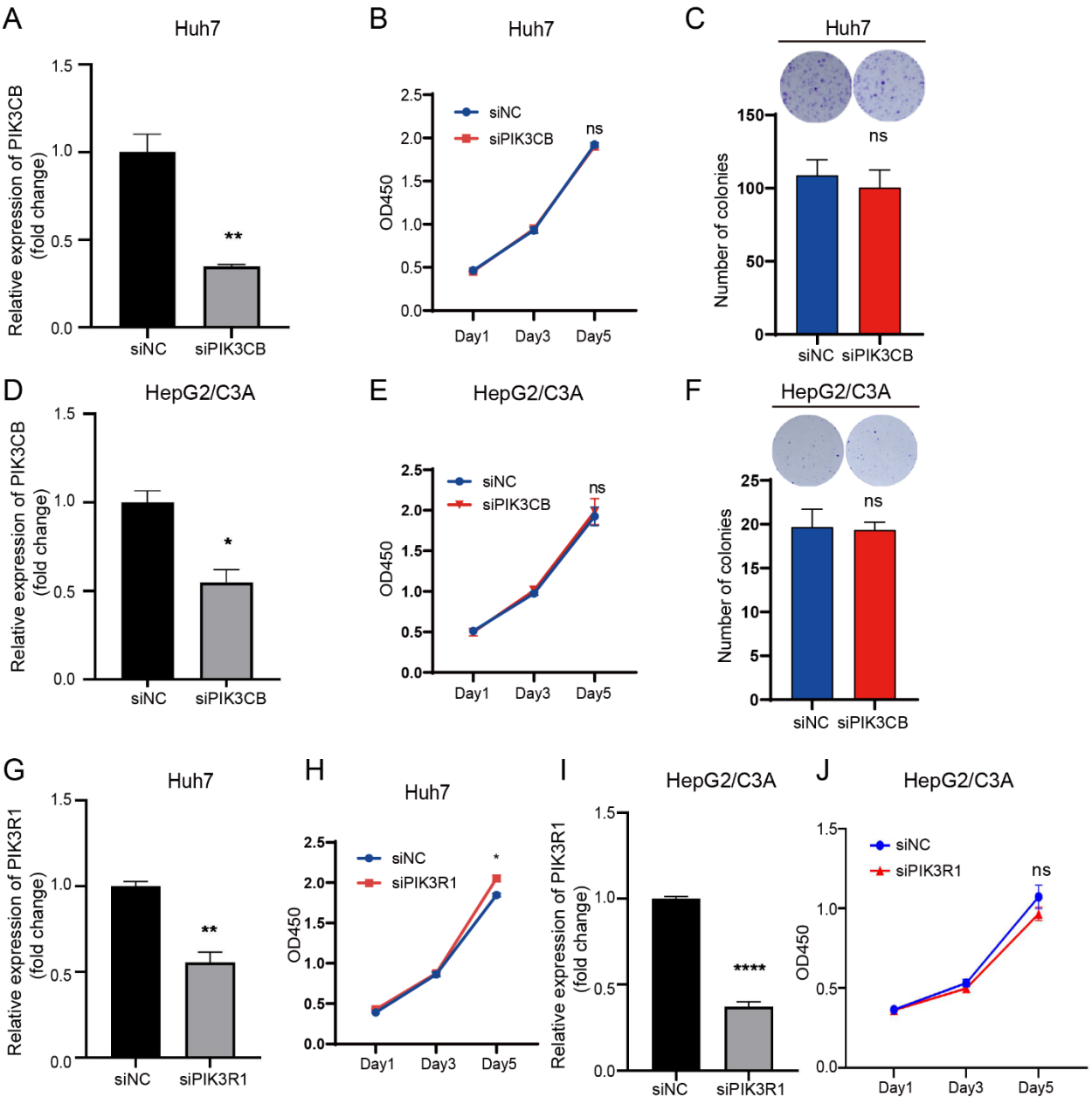


Figure S7. Biological function of PIK3CB and PIK3R1 in HCC. (**A**) The PIK3CB mRNA level in Huh 7 cells transfected with small interfering RNA (siRNA) against PIK3CB. (**B**) Representative CCK-8 assays showing proliferation of Huh 7 cells after PIK3CB knockdown. (**C**) Colony formation assay showing the effect of PIK3CB knockdown on proliferation of Huh 7 cells. (**D**) The PIK3CB mRNA level in HepG2/C3A cells transfected with small interfering RNA (siRNA) against PIK3CB. (**E**) Representative CCK-8 assays showing proliferation of HepG2/C3A cells after PIK3CB knockdown. (**F**) Colony formation assay showing the effect of PIK3CB knockdown on proliferation of HepG2/C3A cells. (**G**) The PIK3R1 mRNA level in Huh 7 cells transfected with small interfering RNA (siRNA) against PIK3R1. (**H**) Representative CCK-8 assays showing proliferation of Huh 7 cells after PIK3R1 knockdown. (**I**) The PIK3R1 mRNA level in HepG2/C3A cells transfected with small interfering RNA (siRNA) against PIK3R1. (**J**) Representative CCK-8 assays showing proliferation of Huh 7 cells after PIK3CB knockdown. Related to Figure 5.

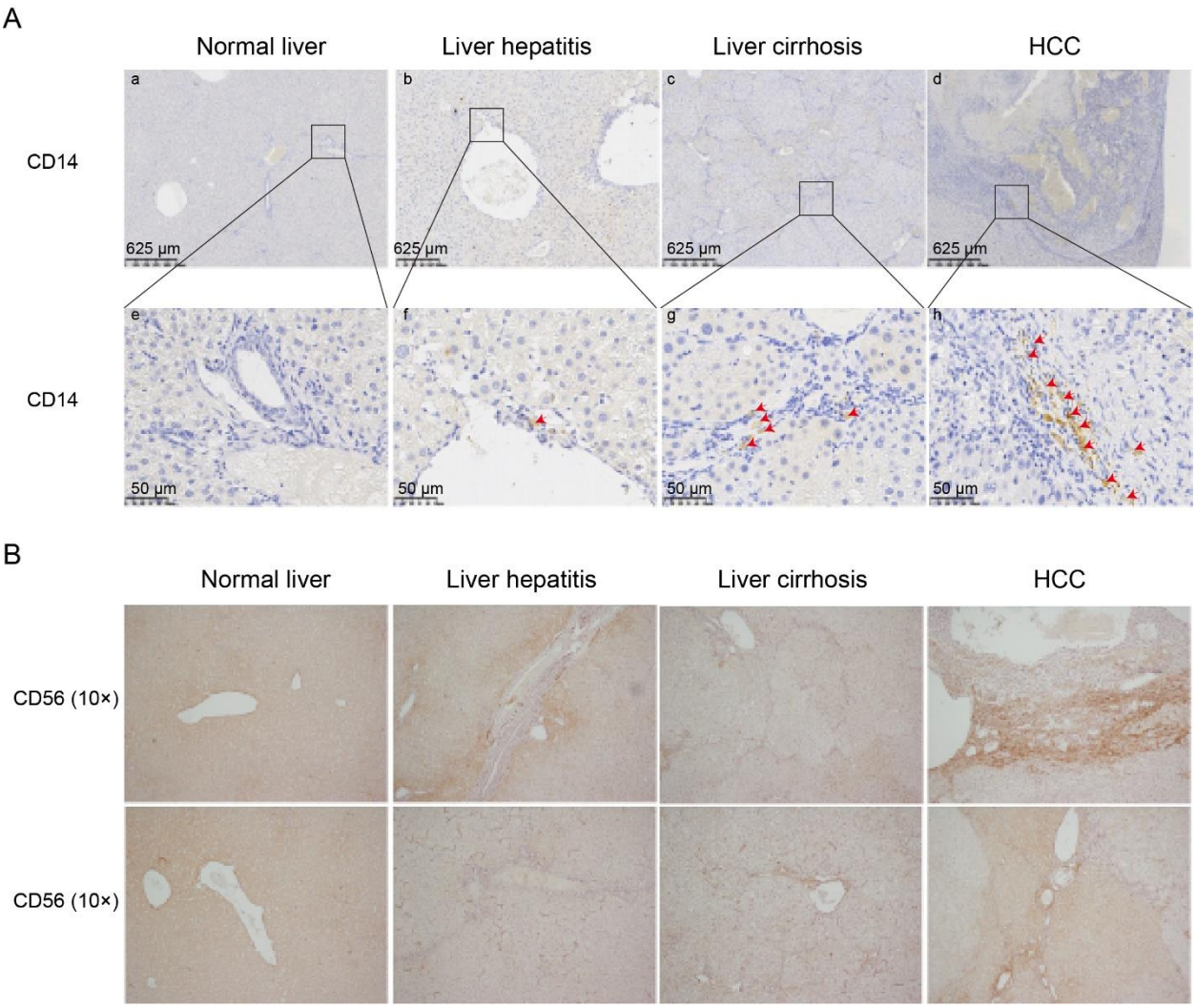


Figure S8. IHC sections of liver tissues. (**A**) Representative image of CD14 expression are shown for liver tissues among the four stages. (**B**) Representative image of CD56 expression are shown for liver tissues among the four stages. Red arrows indicate CD14 positive monocytes/macrophages. Magnification: a, b, c, and d, × 10; e, f, g, and h, × 40; B, × 10. Related to Figure 6.

Supplemental References

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 6, 80–92.

Fan, Y., Xi, L., Hughes, D.S.T., Zhang, Jianjun, Zhang, Jianhua, Futreal, P.A., Wheeler, D.A., Wang, W., 2016. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. Genome Biol. 17, 178.

Hänzelmann, S., Castelo, R., Guinney, J., 2013. GSVA: Gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics 14.

Jia, J.-B., Wang, W.-Q., Sun, H.-C., Zhu, X.-D., Liu, L., Zhuang, P.-Y., Zhang, J.-B., Zhang, W., Xu, H.-X., Kong, L.-Q., et al., 2010. High Expression of Macrophage Colony-Stimulating Factor-1 Receptor in Peritumoral Liver Tissue Is Associated with Poor Outcome in Hepatocellular Carcinoma After Curative Resection. Oncologist 15, 732–743.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., Morishima, K., 2017. KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 45, D353–D361.

Larson, D.E., Harris, C.C., Chen, K., Koboldt, D.C., Abbott, T.E., Dooling, D.J., Ley, T.J., Mardis, E.R., Wilson, R.K., Ding, L., 2012. Somaticsniper: Identification of somatic point mutations in whole genome sequencing data. Bioinformatics 28, 311–317.

Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595.

Li, S., Hu, Z., Zhao, Y., Huang, S., He, X., 2019. Transcriptome-Wide Analysis Reveals the Landscape of Aberrant Alternative Splicing Events in Liver Cancer. Hepatology 69, 359–375.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., Tamayo, P., 2015. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst. 1, 417–425.

McKenna, Aaron, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, David Altshuler, Stacey Gabriel, M.D., DePristo, and M.A., McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., et al., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 254–260.

Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., Alizadeh, A.A., 2015. Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods 12, 453–457.

Paul Shannon, 1, Andrew Markiel, 1, Owen Ozier, 2 Nitin S. Baliga, 1 Jonathan T. Wang, 2 Daniel Ramage, 2, Nada Amin, 2, Benno Schwikowski, 1, 5 and Trey Ideker2, 3, 4, 5, 2003. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. 13, 6.

Rosales, R.A., Drummond, R.D., Valieris, R., Dias-Neto, E., Da Silva, I.T., 2017. signeR: An empirical Bayesian approach to mutational signature discovery. Bioinformatics 33, 8–16.

Sho, M., Gao, Q., Yamato, I., Wang, X.-Y., Qiu, S.-J., Xiao, Y.-S., Nakajima, Y., Fan, J., Zhou, J., Li, B.-Z., et al., 2009. Overexpression of PD-L1 Significantly Associates with Tumor Aggressiveness and Postoperative Recurrence in Human Hepatocellular Carcinoma. Clin. Cancer Res. 15, 971–979.

Wilson, R.K., Mardis, E.R., McLellan, M.D., Koboldt, D.C., Shen, D., Zhang, Q., Ding, L., Larson, D.E., Lin, L., Miller, C.A., 2012. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 22, 568–576.