



Spatial Multiomics Analysis Reveals Only Minor Genetic and Epigenetic Changes in Human Liver Cancer Stem-Like Cells Compared With Other Tumor Parenchymal Cells

Dan Liu^{1,2†}, Hong Li^{3*†}, Hui Dong^{4†}, Mincheng Qu^{1,2}, Liguang Yang³, Lina Chen³, Yixue Li³, Hongyang Wang^{1,5*} and Yufei He^{2*}

OPEN ACCESS

Edited by:

Suman S Thakur,
Centre for Cellular and Molecular
Biology (CCMB), India

Reviewed by:

Alexey Ruzov,
University of Nottingham,
United Kingdom
Yue J Wang,
Florida State University, United States

*Correspondence:

Hong Li
lihong01@sibs.ac.cn
Hongyang Wang
hywangk@vip.sina.com
Yufei He
yfe@sibcb.ac.cn

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Molecular and Cellular Oncology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 07 November 2021

Accepted: 17 January 2022

Published: 09 February 2022

Citation:

Liu D, Li H, Dong H, Qu M, Yang L,
Chen L, Li Y, Wang H and He Y (2022)
Spatial Multiomics Analysis Reveals
Only Minor Genetic and Epigenetic
Changes in Human Liver Cancer Stem-
Like Cells Compared With Other
Tumor Parenchymal Cells.
Front. Cell Dev. Biol. 10:810687.
doi: 10.3389/fcell.2022.810687

¹Model Animal Research Center, Medical School, Nanjing University, Nanjing, China, ²Molecular Pathology Laboratory, National Center for Liver Cancer, Eastern Hepatobiliary Surgery Hospital, Shanghai, China, ³CAS Key Laboratory of Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China, ⁴Department of Pathology, Eastern Hepatobiliary Surgery Hospital, Shanghai, China, ⁵National Center for Liver Cancer and International Cooperation Laboratory on Signal Transduction, Eastern Hepatobiliary Surgery Institute/Hospital, Shanghai, China

Cancer stem cells (CSCs) usually account for a very small tumor cell population but play pivotal roles in human cancer development and recurrence. A fundamental question in cancer biology is what genetic and epigenetic changes occur in CSCs. Here we show that the *in-situ* global levels of DNA cytosine modifications, including 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC) and 5-formylcytosine (5fC), are similar between liver cancer stem-like (LCSL) cells and paratumor liver cells of liver cancer patients. We then developed a robust method combining immunohistochemistry, laser capture microdissection and genome sequencing with ultra-low-input cells (CIL-seq) to study the detailed genetic and DNA methylation changes in human LCSL cells. We first used clinical samples of mixed hepatocellular carcinoma-cholangiocarcinoma (HCC-CCA) with stem cell features to investigate human LCSL cells. The CIL-seq analysis of HCC-CCA and HCC patients showed that LCSL cells had strong spatial genetic and epigenetic heterogeneity. More interestingly, although the LCSL cells had some potential key changes in their genome, they had substantially fewer somatic single nucleotide variants (SNVs), copy number alterations (CNAs) and differentially methylated regions than other tumor parenchymal cells. The cluster analysis of SNVs, CNAs, DNA methylation patterns and spatial transcriptomes all clearly showed that the LCSL cells were clustered with the paratumor liver cells. Thus, spatial multiomics analysis showed that LCSL cells had only minor genetic and epigenetic changes compared with other tumor parenchymal cells. Targeting key changes in CSCs, not just changes in bulk tumor cells, should be more effective for human cancer therapy.

Abbreviations: AFP, alpha fetal protein; ALB, albumin; CK19, cytokeratin 19; CNA, copy number alteration; CSC, Cancer stem cell; DMR, differentially methylated region; EpCAM, epithelial cell adhesion molecule; 5fC, 5-formylcytosine; FFPE, formalin-fixed, paraffin-embedded; GO, gene ontology; HCC-CCA, hepatocellular carcinoma-cholangiocarcinoma; H&E, hematoxylin and eosin; 5hmC, 5-hydroxymethylcytosine; IHC, immunohistochemistry; LCM, laser capture microdissection; LCSL, liver cancer stem-like; 5mC, 5-methylcytosine; OCT, optimal cutting temperature; PcG, polycomb group; PDR, paratumor ductular reaction; PL, paratumor liver; PLC, primary liver cancer; SALL4, Sal-like protein 4; SNV, single nucleotide variant; ST, spatial transcriptome; UMI, unique molecular identifier; WES, whole exome sequencing; WGBS, whole genome bisulfite sequencing; WGS, whole-genome sequencing.

Keywords: DNA methylation, cytosine modification, CIL-seq, exome sequencing, spatial transcriptome

INTRODUCTION

Human cancer development is a complex evolutionary process (Greaves and Maley, 2012; Black and McGranahan, 2021). To date, it is universally acknowledged that cancer cells possess many significant genetic and epigenetic changes caused by multiple factors compared with normal cells. Some of these changes are essential for tumor evolution (Black and McGranahan, 2021). Many cancers are thought to originate from cancer stem cells (CSCs), which pose a high risk of therapy resistance and cancer relapse (Greaves and Maley, 2012; Black and McGranahan, 2021). Understanding the genetic and epigenetic changes in human CSCs should shed light on a better understanding of the developmental and evolutionary trajectory of a tumor and the design of better cancer therapeutic approaches. However, it is always difficult to study human CSCs, as they generally account for a very small proportion of tumor specimens from clinical patients. Using cell culture and animal models, including the xenotransplantation approach, to investigate the properties of CSCs has inherent technical and conceptual limitations (Batlle and Clevers, 2017). Consequently, the genetic and epigenetic changes that occur in human CSCs are still poorly understood.

Liver cancer is one of the most common cancer types worldwide. In the current study, we developed a robust spatial multiomics method for genome sequencing with ultra-low-input cells, to uncover the genetic and DNA methylation changes related to human CSCs with different spatial locations. We first selected specimens with a stem cell phenotype from mixed hepatocellular cholangiocarcinoma (HCC-CCA) patients to investigate liver cancer stem-like (LCSL) cells. We also investigated LCSL cells in other liver cancer samples using spatial multiomics analysis.

MATERIALS AND METHODS

Patients and Collection of Clinical Specimens

Formalin-Fixed, Paraffin-Embedded (FFPE) Specimens

For mixed HCC-CCA cases collection, pathological sections of 12,603 primary liver cancer (PLC) patients were investigated who received surgical resection from 2017 to 2019 at the Eastern Hepatobiliary Surgery Hospital (EHBH) in Shanghai, China, and finally total 278 HCC-CCA cases were found out and further investigated for LCSL cells. In accordance with the number of HCC-CCA cases, 265 HCC patients who underwent surgical resection from 2009 to 2020 at EHBH were randomly selected and investigated for LCSL cells (Supplementary Table S1).

Fresh Frozen Specimens

Fresh PLC tissues and matched paratumor liver tissues were obtained from 58 patients (35 cases in cohort 1 and 23 cases in

cohort 2) who underwent surgical resection at EHBH (Supplementary Table S1). Each tissue specimen was embedded in optimal cutting temperature (OCT) compound medium, immediately frozen in an isopentane slurry made with liquid nitrogen, and finally stored at -80°C until further processing. Each tissue was embedded within 30 min for frozen sectioning after surgical removal. All histological specimens were evaluated by at least two experienced pathologists.

The archives of all patients were collected by the EHBH archive system. Informed consent was obtained from the patients, and all procedures were approved by the ethical committee of EHBH.

Immunohistochemistry (IHC)

For IHC staining of frozen sections, briefly, OCT-embedded frozen tissues were cut into $8\mu\text{m}$ thick sections, and endogenous peroxidases were inactivated by incubation in 0.3% H_2O_2 . Nonspecific signals were blocked using 5% bovine serum albumin. The primary antibodies used were as follows: anti-EpCAM (1:500, ab7504; Abcam), anti-OV6 (1:200, MAB2020; R&D Systems) and anti-GPC3 (1:100, ab207080; Abcam). HRP-conjugated antibodies were used as the secondary antibodies. Diaminobenzidine colorimetric reagent solution was used for staining followed by hematoxylin counterstaining. The slides were finally scanned, and representative images were illustrated.

For IHC staining of FFPE sections, the tissues were fixed overnight in 4% PFA, embedded in paraffin and cut into $4\mu\text{m}$ serial consecutive sections. After deparaffinization, antigen retrieval was performed with sodium citrate buffer (10mM sodium citrate, 0.05% Tween-20, pH 6) (for H&E, omit this step, subsequently by hematoxylin and eosin standard protocols). The following steps were the same as those in the aforementioned frozen section. The primary antibodies used here were as follows: anti-SALL4 (1:100, sc-101147; Santa Cruz Biotechnology), anti-CK19 (1:100, ab9221; Abcam), and anti-Ki-67 (1:1000, ab15580; Abcam).

For 5mC, 5hmC or 5fC immunostaining, DNA was denatured with 2N HCl for 15 min at room temperature after antigen retrieval treatment, followed by neutralization with 100mM Tris-HCl (pH 8.0) for 10 min at room temperature. The primary antibodies used were as follows: anti-5mC (1:5000, Cat#: 28692; Cell Signaling Technology), anti-5hmC (1:10000, Cat#: 39769; Active Motif); and anti-5fC (1:1000, Cat#: 61227; Active Motif). For EpCAM and 5hmC double-staining, primary anti-5hmC antibody (rabbit) was first incubated for 16 h at 4°C , followed by alkaline phosphatase-based streptavidin/biotinylated link system to visualize dark purple cell nuclei. Anti-EpCAM antibody (mouse) was then incubated for 2 h at room temperature, followed by detection using a horseradish peroxidase/AEC system to visualize the red cell membrane. Specifically, the slides were sealed with glycerol mounting medium and without hematoxylin counterstaining. We could not obtain good results for EpCAM and 5fC double staining,

possibly because of the low 5fC content in cells and thus usually weak staining for 5fC.

CIL-Seq IHC and Laser Capture Microdissection (LCM)

OCT-embedded frozen tissues were cut into 10 μ m thick sections and mounted on PEN Membrane Glass Slides (Leica). After IHC of EpCAM (*see above*), tissues were dehydrated through rising ethanol concentrations (50, 75, 85, 95, 100 and 100% ethanol, 60 s each) in 50 ml sterile centrifuge tubes. When membrane slides dried completely, the LCM procedure was initiated by a Leica LMD 7000 laser microdissection system according to the manufacturer's protocol. Two pathologists confirmed the different tumor or paratumor components independently. All captured samples were collected in 0.2 ml PCR tubes. Finally, the membrane slides were cleared with xylene and sealed for long-term preservation. Certified RNase/DNase-free materials were used whenever available.

Genetic Libraries

Ultralow DNA from LCM samples was extracted by a Quick-DNA Microprep Plus Kit (Zymo Research) before single-cell multiple displacement amplification (MDA) using a REPLI-g Single Cell Kit (Qiagen) for trace DNA amplification. After obtaining high yields of high-quality whole-genome amplified DNA, whole-exome capture and library construction were carried out by an Agilent SureSelect Human All Exon V6 kit (Agilent) according to the manufacturer's instructions. Whole-genome libraries were constructed by the QIAseq FX Single Cell DNA Library kit (Qiagen).

Whole-Genome Bisulfite Sequencing (WGBS) Libraries

Tissue samples from LCM were digested for up to 4 h with 1 mg/ml proteinase K according to the EZ DNA Methylation-Direct Kit (Zymo Research). After thorough digestion, LCM cells were immediately sent to the Pico Methyl-Seq Library Prep Kit (Zymo Research) to accommodate ultralow DNA input according to the manufacturer's instructions with minor adjustment: bisulfite conversion time was extended to 90 min to ensure complete conversion; for the sample M1BULK, the library only needed six cycles in **Section 4**, while other libraries needed 11 cycles.

Sequencing

All libraries were sequenced with the Illumina Xten platform to generate 150 bp paired-end reads.

Whole-Genome Sequencing (WGS) and Whole-Exome Sequencing (WES) Data Analysis

Data Quality and Filtering

Raw reads were filtered with Trimmomatic to remove adaptor sequences and low-quality bases (Bolger et al., 2014). The remaining reads were aligned to the human genome (hg38) by the BWA algorithm (Li and Durbin, 2009). After mapping, the duplicate reads were removed. Germline variants were called using GATK (Auwera et al., 2013).

Somatic Mutations

Taking bulk immune infiltrating cells as a control, somatic mutations were identified from BAM files using muTect and varScan (Koboldt et al., 2012; Cibulskis et al., 2013). High-quality somatic single nucleotide variants (SNVs) were obtained by applying the following filters: 1) Removal of potential germline variants that are recorded in dbSNP or variants with allele frequency larger than 1% in population (Sherry et al., 2001; Genomes Project Consortium et al., 2015); 2) Removal of variants located within 10 bp distance, which are more possible to be errors introduced during library preparation; 3) Removal of mutations that are identified by only one algorithm; 4) Depth of the mutation sites should be larger than 10 in sample j and control sample; 5) There are at least two reads supporting the alternative allele and the allele frequency is larger than 5% in sample j , and no reads have the alternative allele in control sample; 6) Since there are more samples in P2 than in P7, only mutations that occur in more than one of P2 samples are analyzed to improve quality. This filter is not suitable for P7.

Functional consequences of somatic SNVs were annotated by ANNOVAR (Wang et al., 2010). Somatic SNVs of all samples were converted to a 0–1 matrix (M), in which 1 indicates having a mutation. To reduce the effects of allele dropout, when site i had low coverage ($\leq 5\times$) in the sample, M_{ij} was adjusted based only on other samples of the same pathology group. $M_{ij} = 0.9$ if more samples carry the mutation or $M_{ij} = 0.1$ if more samples are wild type. Euclidean distances between paired samples were calculated from M and hierarchical clustering was performed to cluster samples.

Copy Number Alterations

Copy number alterations (CNAs) were identified from WES data by CNVkit (Talevich et al., 2016). Samples covering $<50\%$ target regions were removed from the CNA analysis. Deletions were ignored to avoid the effects of uncovered regions. Segments and genes with $\log_2(\text{CN}/2) \geq 1$ were regarded as amplifications. To decrease false positives of amplified genes, we only analyzed genes that were amplified in $\geq 5\%$ TCGA liver cancer patients (<https://www.cancer.gov/tcga>) and genes with three more amplifications in tumor samples than in paratumor samples. Euclidean distances were calculated from amplified genes, and hierarchical clustering was used to plot the clustering tree.

WGBS Data Analysis

Data Quality and Filtering

Raw reads were filtered with Trimmomatic to remove adaptor sequences and low-quality bases (Bolger et al., 2014). The remaining reads were aligned to the human genome (hg38) by Bismark (Krueger and Andrews, 2011). Potential PCR duplicates were removed. The bisulfite conversion rate was estimated as the ratio between the number of methylated non-CpGs and total non-CpG sites.

Data Processing

The R package methylKit was used to generate and analyze the methylation matrix (Akalin et al., 2012). Bases whose coverage was larger than $500\times$ or lower than $3\times$ were discarded. The

genome was split into 300-bp windows with 300-bp step size for fresh-frozen samples. For FFPE samples, the genome was split into 1 Mb windows with a 1 Mb step size. Windows that were covered by less than three bases were discarded. Similarity between paired samples was measured by Pearson correlation coefficients. Hierarchical clustering was used to explore the clusters of samples. For differential methylation analysis between two groups, the chi-square test was used to compare windows covered by at least two samples in each group. Differentially methylated regions (DMRs) were windows that satisfied Q -value < 0.05 and absolute difference > 0.25 . DMRs were annotated with genomic contexts and 15 chromatin states of embryonic stem cell line H9 (downloaded from Roadmap Epigenomics Project) (Roadmap Epigenomics Consortium et al., 2015). Annotation enrichment analysis was performed by Fisher's exact test. Annotations with Q -values < 0.01 were regarded as significantly enriched.

Gene Ontology (GO) Enrichment Analysis

The Metascape web-based tool (Zhou et al., 2019) was used for GO analysis of the genes of interest. The Metascape analysis was performed using the default settings. A value of $p < 0.05$ was considered statistically significant.

DNA Modification Analysis

For statistical analysis of 5mC, 5hmC and 5fC staining, we set a standard assessment rule: within one patient, the strongest stained areas were marked with three scores. Specifically, strong positive, score = 3; medium positive, score = 2; weak positive, score = 1; no signal, score = 0. Five randomly selected equal fields (400 \times) of IHC images of each cell type were analyzed.

Spatial transcriptome Analysis

ST Experiments

Spatial transcriptome (ST) experiments were performed according to the user guide of Visium Spatial Gene Expression Reagent Kits (10 \times Genomics). Briefly, HCC tissues were gently washed with cold PBS and filled with OCT in a proper mold and then snap frozen in chilled isopentane. Cryosections were mounted onto a spatially barcoded array of ST with 10- μ m thickness, and serial adjacent cryosections were mounted onto regular glass slides for IHC staining of EpCAM or other markers with 8- μ m thickness. For processing, the tissue was fixed for 30 min with prechilled methanol at -20°C , followed by H&E staining. Slides were finally taken on a Leica SCN400 F whole-slide scanner at 40 \times resolution. After capturing ideal tissue morphology information and ensuring that RNA was not degraded ($\text{RIN} \geq 7$), tissue permeabilization and reverse transcription were immediately conducted by a Visium Spatial Tissue Optimization Kit (10 \times Genomics). Finally, the ST library was prepared with second strand synthesis and denaturation and sequenced by NovaSeq 6000 (Illumina). Each of the spots printed onto the array is 50 μ m in diameter and 100 μ m from center to center, covering an area of $6.5 \times 6.5 \text{ mm}^2$ (ref. (Stahl et al., 2016)).

Data Processing and Cluster Analysis

Raw sequenced ST reads were processed using Space Ranger analysis software (version 1.0.0, 10 \times Genomics) mapped to

the GRCH38 genome assembly following the standardized analysis rules. Unique molecular identifier counts in each spot were normalized and scaled by the median number transcript count across all spots. Cluster analysis was based on the K-Means algorithm in Space Ranger software.

Statistical Analyses

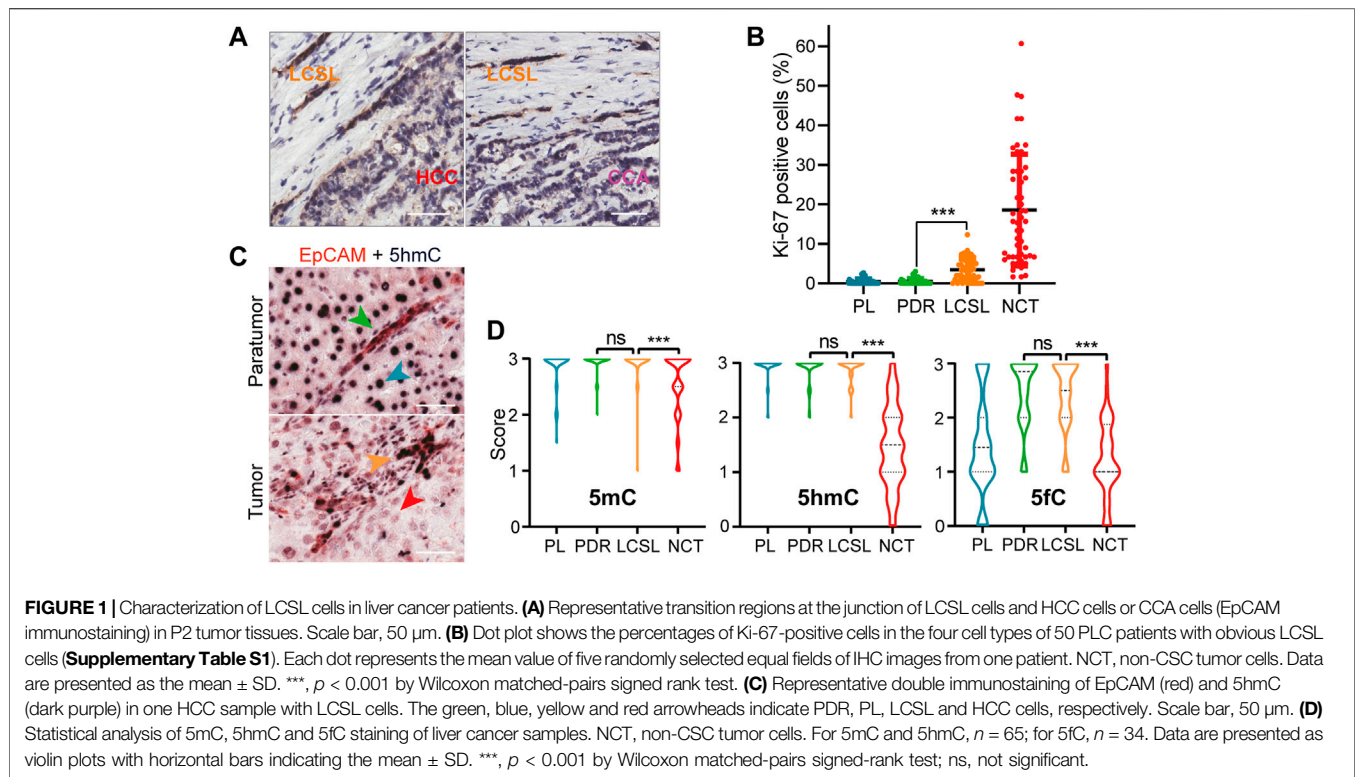
Statistical analysis and data visualization were carried out using the R/Biocoductor software packages (<http://www.bioconductor.org>) or GraphPad Prism 8 (GraphPad Software, La Jolla, CA, USA). All p -values less than 0.05 were considered statistically significant.

RESULTS

Characteristics of LCSL Cells in Liver Cancer Patients

Mixed hepatocellular cholangiocarcinoma (HCC-CCA) is a rare type (less than 5%) of primary liver cancer (PLC) with mixed differentiation (**Supplementary Figure S1A**), and its diagnosis depends on postoperative histopathology (Brunt et al., 2018). Some cancer cells in mixed HCC-CCAs have stem cell characteristics (Brunt et al., 2018; Munoz-Garrido and Rodrigues, 2019). The proportion of CSCs or LCSL cells in mixed HCC-CCAs is usually much higher than that in other types of tumors. Therefore, this type of PLC can provide us with a perfect tumor object directly from clinical patients to help investigate human CSCs or LCSL cells and compare the similarities and differences between LCSL cells and HCC or CCA cells within the same liver cancer sample. We collected fresh liver cancer specimens that were resected and diagnosed with PLC before surgery. According to the clinicopathological diagnosis, one mixed HCC-CCA case (patient P2, **Supplementary Table S1**) among a total of 35 cases (**Supplementary Table S1**) collected was identified. This mixed HCC-CCA specimen is the stem-cell feature type, which is the exact type of specimen that we aimed to research. Pathological histocytology (**Supplementary Figure S1B**) showed three distinct types of parenchymal cells coexisting in the tumor tissues: HCC cells, CCA cells, and LCSL cells. LCSL cells express liver stem cell markers, including epithelial cell adhesion molecule (EpCAM), cytokeratin 19 (CK19), OV6 and Sal-like protein 4 (SALL4), and account for approximately 10% of the overall tumor tissues (**Supplementary Figure S2**). Some LCSL cells grew in clonal clusters, and the cell morphology within each clone was not distinctly different under the microscope (**Supplementary Figure S1B**). Thereafter, we selected EpCAM, which was used to label liver stem cells (Huch et al., 2015; Aizarani et al., 2019) or liver CSCs (Yamashita et al., 2009; Yamashita et al., 2013; Matsumoto et al., 2017), as a marker molecule for LCSL cells or other liver stem-like cells in this study.

By integrating EpCAM staining and tissue/cell morphology, the liver parenchymal cells in the paratumor tissue of P2 could be further subdivided into two types: the most common type was EpCAM-negative cells, which are ordinary paratumor liver (PL) cells, and the other type was EpCAM-positive cells, which correspond to paratumor ductular reaction (PDR) cells (**Supplementary Figure S1B**). In PDR cells, each clone is generally believed to contain a portion of liver stem cells (Sato et al., 2019). Therefore, EpCAM



immunostaining combined with tissue/cell morphology features could clearly distinguish the two types of liver parenchymal cells in paratumor tissues and the three types of parenchymal cells in tumor tissues. At the junction between LCSL cells and HCC cells or CCA cells, some transitional areas could be clearly observed (**Figure 1A**), supporting that HCC and CCA cells may be derived from LCSL cells.

LCSL cells and PDR cells had similar characteristic markers, such as EpCAM, CK19, OV6, and SALL4 (**Supplementary Figure S1B**). Therefore, LCSL cells closely resembled PDR cells in both histomorphology and key markers of liver stem cells. We investigated the pathological sections of 12,603 PLC patients who received surgical resection from 2017 to 2019 at our hospital. Among them, 278 cases (2.2%) were pathologically diagnosed as HCC-CCA and 58 cases of them (**Supplementary Table S1**) could be observed such LCSL cells as in patient P2. We also investigated 265 HCC patients and found that 35 cases of HCC (**Supplementary Table S1**) could be observed such LCSL cells, whereas the number of LCSL clones was much less than that in HCC-CCA. Notably, however, LCSL cells had significantly more Ki-67-positive staining than PDR cells (**Figure 1B** and **Supplementary Figure S1B**), indicating that LCSL cells have greater proliferative activity.

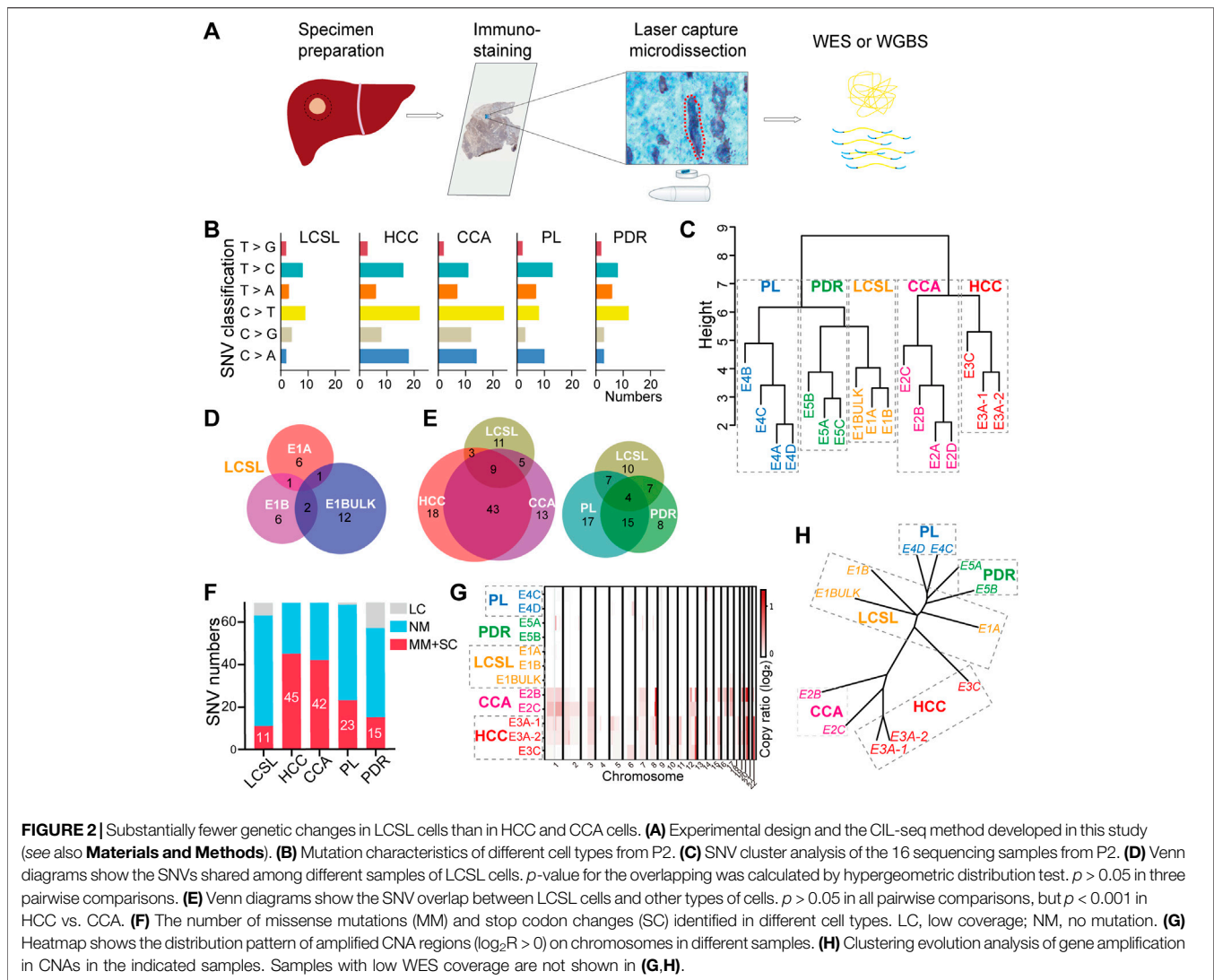
Only Minor Global DNA Cytosine Modification Changes Occur in Human LCSL Cells Compared With Other Tumor Parenchymal Cells

An important epigenetic change during cancer development is the global demethylation of DNA. We and others have found that

5-methylcytosine (5mC) can be oxidized to the demethylation intermediates 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) in mammalian cells (He et al., 2011; Xu and Bochtler, 2020). Recently, we found that in the early stage of HCC, the global content of 5hmC and 5fC was decreased (Liu et al., 2019), but it is not clear whether there is such a change in LCSL cells. We immunostained the FFPE section samples of P2 tissues for 5mC, 5hmC and 5fC. Surprisingly, the results showed that most LCSL cells were different from other tumor parenchymal cells which had obviously decreased 5hmC and 5fC staining (**Supplementary Figure S3**). Examination of 65 specimens with obvious LCSL cells (**Supplementary Table S1**), including 30 mixed HCC-CCAs and 35 HCC patients, showed similar results: there was little difference in 5mC, 5hmC or 5fC staining between LCSL cells and PDR cells (**Figures 1C,D**), suggesting that LCSL cells do not undergo extensive DNA demethylation. These results suggest that only minor global DNA cytosine modification changes occur in human LCSL cells.

Development of CIL-Seq Method

To explore the detailed genetic and epigenetic differences between LCSL cells and other EpCAM-stained cells in both tumor and paratumor tissues, we developed a spatial multiomics method combining immunohistochemistry (IHC), laser capture microdissection (LCM) and genome sequencing with ultra-low-input cells (CIL-seq) (**Figure 2A**, Materials and Methods). First, to better obtain LCSL cells for further study, we carried out EpCAM IHC staining on frozen sections. Second, we used LCM to capture needed cells more accurately and visually in multiple spatial regions, which is helpful for the investigation of



tumor biology and heterogeneity. Then, to test whether the IHC process damages the integrity of genomic DNA, we selected approximately 40 pure HCC or PL cells from specimen P1 (**Supplementary Figure S4A** and **Supplementary Table S1**) for whole genome sequencing (WGS) library analysis based on single-cell technology. Simultaneous sequencing of several same-type cells could reduce the inherent single-base false positives caused by whole-genome amplification of single cells as much as possible. Last, the WGS results showed satisfactory coverage; at a $15\times$ mean sequencing depth, WGS covered approximately 78–80% of the genome at $\geq 1\times$ coverage (**Supplementary Table S2**), confirming the integrity of genomic DNA.

Whole-Exome Sequencing (WES) Analysis of LCSL Cells Reveals Only Minor Genetic Changes

We performed a WES analysis of specimen P2 based on CIL-seq. According to the aforementioned EpCAM staining and

tissue/cell morphology, we isolated five types of cell samples (**Supplementary Figures S4A–C**) from frozen sections, including three components in the tumor: LCSL cells (three samples), CCA cells (four samples) and HCC cells (two samples), and two components in the paratumor: PDR cells (three samples) and PL cells (four samples). In addition, we used bulk immune infiltrating cells (~ 500 cells) from the paratumor tissue (**Supplementary Figure S4B**) in this patient as the normal genomic reference. Taken together, WES libraries were prepared from a total of 17 samples and sequenced. One of the 17 samples, E3A, was sequenced repeatedly, and the results showed good sequencing concordance (**Supplementary Figure S4D**). At least two samples of each type had good coverage (average target depth $>62\times$) (**Supplementary Table S2**).

We utilized two algorithms to identify somatic single nucleotide variants (SNVs) and removed false positives by commonly used filters (**Materials and Methods**). Consequently, we detected 125 SNVs in all samples, and

Sanger sequencing further confirmed the accuracy of SNV detection (**Supplementary Table S3**). They were predominantly C to T transitions (**Figure 2B**), similar to those previously reported (Marquardt et al., 2015; Blokzijl et al., 2016; Zhu et al., 2019). Through cluster analysis of the 125 SNVs, it was clearly seen that different samples from the same cell types could be clustered together. Intriguingly, the LCSL samples were clustered with the paratumor samples instead of the HCC or CCA samples (**Figure 2C**).

Most SNVs were detected in HCC or CCA samples (73/125 and 70/125, respectively, **Supplementary Table S3**), and the number of SNVs was consistent with that of previous findings in HCC or mixed HCC-CCA samples (Schulze et al., 2015; Lin et al., 2017; Wang et al., 2018; Xue et al., 2019). Notably, only minor SNVs (28 in total) were detected in LCSL cells. There were considerably more independent SNVs among LCSL samples (**Figure 2D**, **Supplementary Figure S5A** and **Supplementary Table S3**), suggesting that LCSL cells harbor strong genetic heterogeneity. HCC and CCA cells shared 52 SNVs (**Figure 2E** and **Supplementary Table S3**), indicating a monoclonal origin, which is in agreement with results from previous reports on the monoclonal origin of mixed HCC-CCA (Wang et al., 2018; Xue et al., 2019). More than half of the mutations in LCSL cells (17/28) coincided with those in HCC and CCA cells (**Figure 2E** and **Supplementary Table S3**), supporting the same origin of LCSL cells and HCC or CCA cells. Previous reports showed extensive SNVs in normal or cirrhotic liver cells (Brunner et al., 2019; Zhu et al., 2019). Consistent with these findings, we found many SNVs (43 in total, **Supplementary Table S3**) in PL cells.

Of the 69 genes with missense mutations or stop codon changes, only a small portion of these SNVs were detected in LCSL cells (**Figure 2F** and **Supplementary Table S3**). We found that three genes (*DHX9*, *ERBB4*, *FAM46D*) were among the 299 reported driver genes (Bailey et al., 2018), and one (*ZFH4*) was among the 161 driver genes of HCC (Schulze et al., 2015). These driver gene mutations were detected only in HCC or CCA cells (**Supplementary Table S3**). Taken together, these results suggest that crucial gene mutations that may promote tumor growth have occurred in HCC and CCA cells; however, substantially fewer mutations occurred in LCSL cells.

Only Few Key Genetic Changes in LCSL Cells Compared With Other Tumor Parenchymal Cells

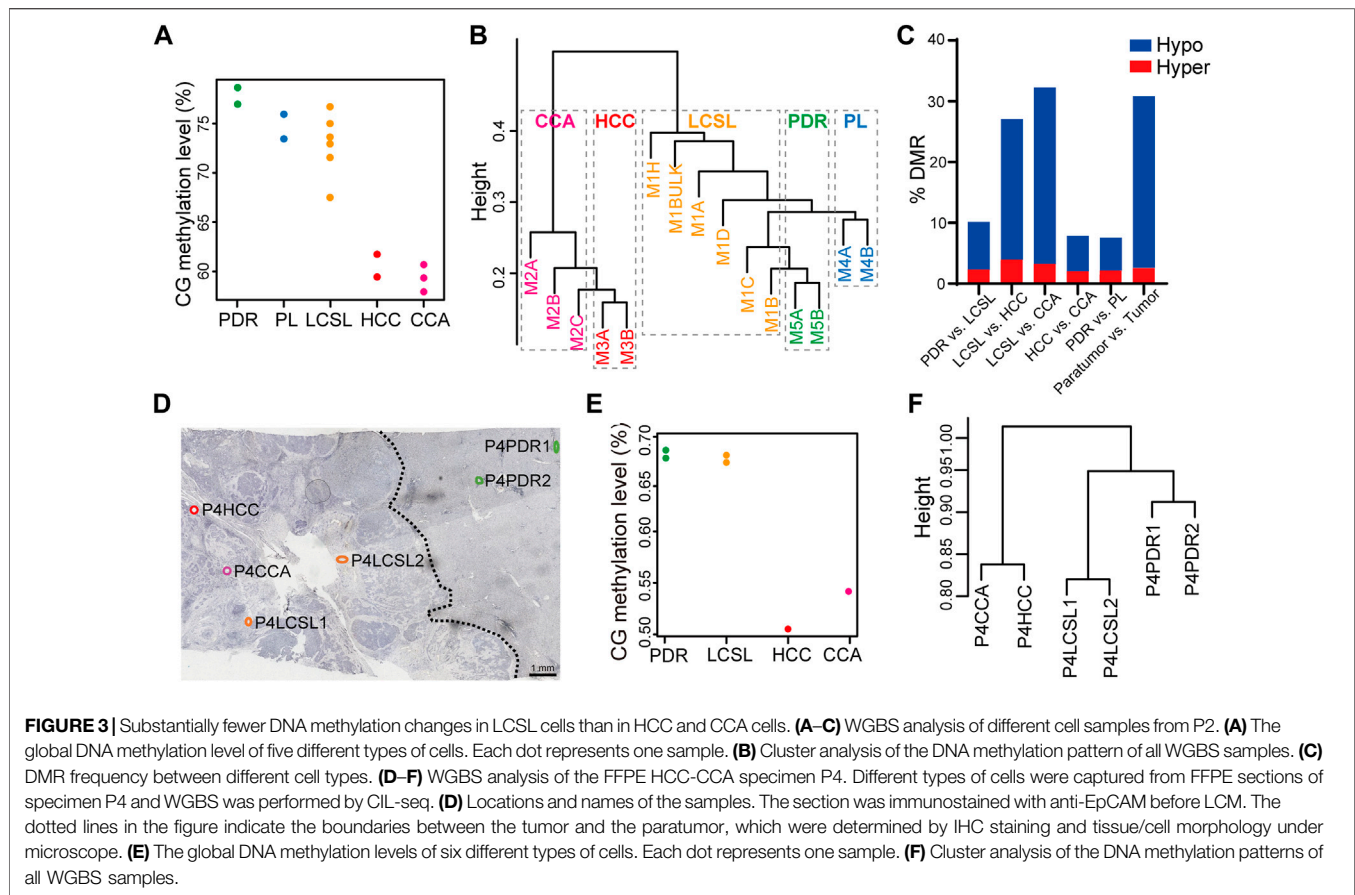
We next analyzed the copy number alteration (CNA) using the WES data. Our analysis focused only on the amplified loci because of the low coverage of some samples. We found many CNA amplification regions in the tumor, but most of them were centralized in HCC and CCA cells, including the amplification of chromosome 8q, which is often detected in liver cancer (Zhai et al., 2017; Cancer Genome Atlas Research Team et al., 2017; Xue et al., 2019), while only very few CNA regions were detected in LCSL cells (**Figure 2G** and **Supplementary Table S4**). We also analyzed the amplified genes. Overall, consistent with the CNA amplification regions, at the gene level, LCSL

cells also had apparently fewer amplification genes than HCC and CCA cells (**Supplementary Figure S5B** and **Supplementary Table S4**). A phylogenetic tree constructed using the amplified genes clearly showed that similar to SNVs, LCSL cells were more similar to paratumor cells than to HCC and CCA cells (**Figure 2H**). Interestingly, some of these amplified genes are oncogenes, such as *MYC*, which was also amplified in one sample of LCSL cells (**Supplementary Figure S6** and **Supplementary Table S4**). Activation of *MYC* is thought to be related to the formation of CSCs in the liver (Chow et al., 2012; Marquardt et al., 2015). Taken together, WES data showed that LCSL cells have only a few SNVs and CNAs compared to those in HCC and CCA cells, but LCSL cells might possess few crucial changes, such as the amplification of the oncogene *MYC*, which may be important for the proliferation of LCSL cells during cancer development and evolution.

Whole-Genome Bisulfite Sequencing (WGBS) Analysis of LCSL Cells Reveals Only Minor Epigenetic Changes

Epigenetic aberrations play an important role in cancer development (Flavahan et al., 2017; Vicente-Dueñas et al., 2018; Rozenblatt-Rosen et al., 2020; Black and McGranahan, 2021), and our WES data showed only a few striking genetic changes in LCSL cells. Thus, we are curious about what epigenetic changes are hidden in LCSL cells. Considering the actual situation that the input DNA amount of a pure cluster of LCSL cells captured by LCM was very low (only ~100 pg), we mainly examined the DNA methylation landscape of these samples. DNA methylation is the most essential DNA modification and is closely associated with other epigenetic contents (REC et al., 2015). We performed WGBS analysis of captured different cell clones (**Supplementary Figure S4**) from specimen P2 using CIL-seq. The bisulfite conversion efficiency of most samples was greater than 98%, and CpG coverage in most samples (**Supplementary Table S2**) was higher than that reported for single-cell WGBS sequencing (Smallwood et al., 2014; Gravina et al., 2016; Linker et al., 2019).

In general, HCC and CCA cells have undergone broad DNA demethylation compared with paratumor cells (**Figure 3A**). However, the global methylation level of LCSL cells changed much less (**Figure 3A**), indicating a closer relationship to paratumor cells, consistent with the above *in situ* global DNA cytosine modification results (**Figure 1D**). In addition, there was extensive DNA methylation heterogeneity in tumor cells (**Supplementary Figure S7A**), as reported (Mazor et al., 2016; Lin et al., 2017; Zhang et al., 2019; Marusyk et al., 2020). Notably, there was also strong DNA methylation heterogeneity among LCSL samples (**Supplementary Figure S7B**). The cluster analysis of the DNA methylation pattern (**Figure 3B**) clearly illustrated that HCC and CCA cells were clustered together, and different samples in the paratumor were clustered together. Intriguingly, LCSL cells were clustered with paratumor cells. Thus, these results suggest that, for DNA methylation, LCSL cells were also more similar to paratumor cells than to HCC and CCA



cells, consistent with the WES analysis shown above (Figures 2C,H).

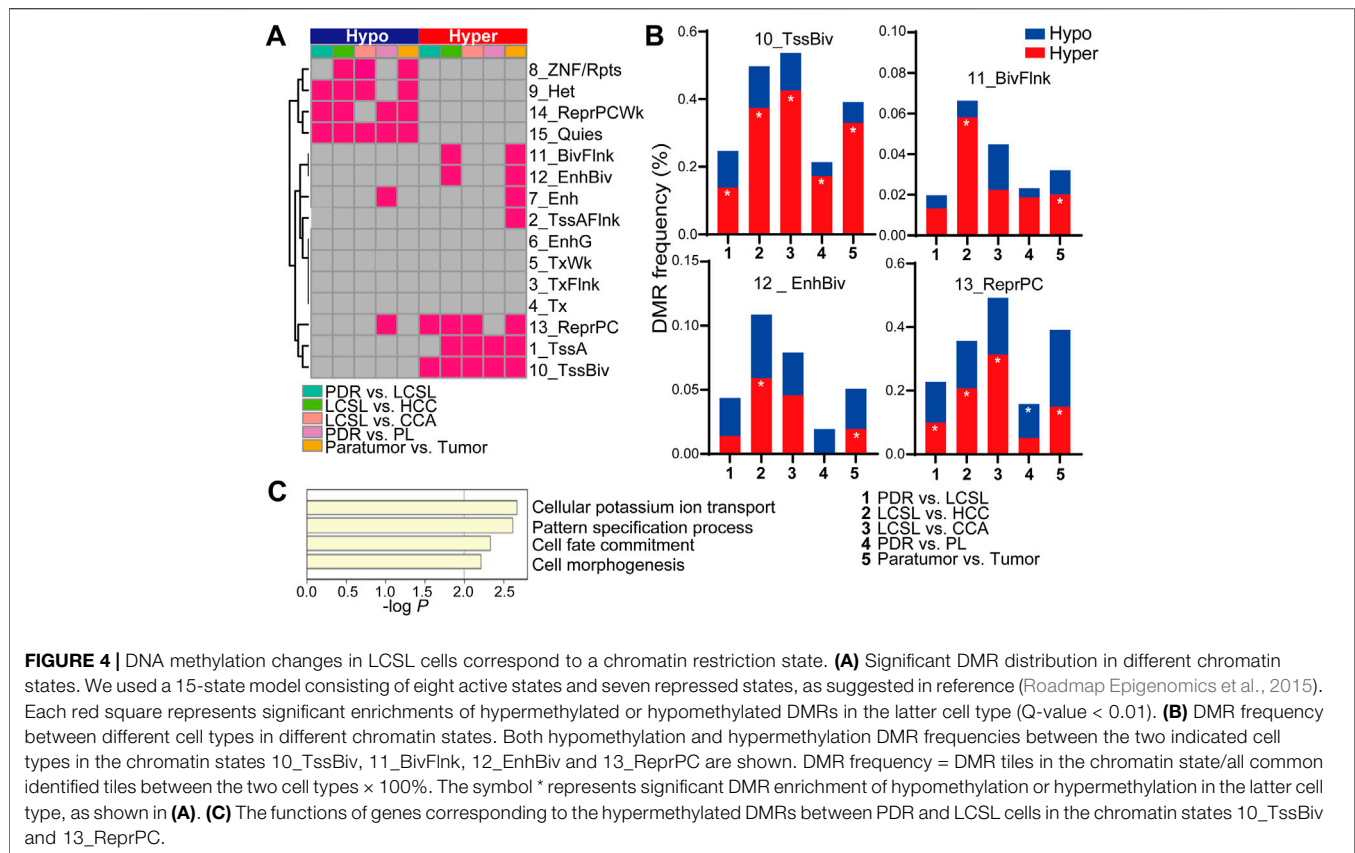
By comparing the differentially methylated regions (DMRs) of distinct sample types, we found that LCSL cells displayed limited variance with paratumor cells but had greater variance with HCC and CCA cells, which is consistent with the global DNA methylation pattern (Figure 3B). Specifically, 10.1% DMRs were found between LCSL cells and PDR cells, while 27.1% or 32.3% DMRs were found between LCSL cells and HCC or CCA cells (Figure 3C and Supplementary Table S5).

Formalin-fixed, paraffin-embedded (FFPE) tissue resources are far richer than frozen tissues. We chose an FFPE HCC-CCA specimen (P4, Supplementary Table S1) and conducted WGBS analysis based on the CIL-seq method (Figure 3F). Although the quality of the sequencing data was much lower than that of frozen tissues (Supplementary Table S2), we still observed that the overall methylation level and DNA methylation pattern of LCSL cells were similar to those of PDR cells (Figures 3E,F), consistent with the results of patient P2. Taken together, these results suggest that in the global pattern of DNA methylation, very large methylation changes occurred in HCC and CCA cells compared with paratumor cells, whereas there were fewer differences in DNA methylation patterns between LCSL cells and paratumor cells.

DNA Methylation Changes in LCSL Cells Correspond to a Chromatin Restriction State

We next examined the distribution region of DMRs in the genome of P2 samples. Overall, there were significantly fewer regions with prominent DNA methylation changes in LCSL cells than in HCC and CCA cells (Supplementary Table S6). According to the 15-state model of chromatin (REC et al., 2015), the significant methylation changes in HCC and CCA cells were closely related to the methylation level of the 15 chromatin states in normal cells. Chromatin states with significant demethylation occurred on those initially exhibiting a hypermethylated state in normal cells (REC et al., 2015), including the 9_Het and 15_Quies chromatin states, and chromatin states with significant hypermethylation occurred on those initially exhibiting a hypomethylated state in normal cells (REC et al., 2015), including the 1_TssA, 10_TssBiv, 11_BivFlnk and 12_EnhBiv states (Figures 4A,B and Supplementary Figure S7C). However, substantially fewer methylation changes occurred in these chromatin states in LCSL cells, which is consistent with the minor change in the global methylation level in LCSL cells (Figure 3A).

Remarkably, we found that a few crucial chromatin regions in LCSL cells had the same DNA methylation tendency as in HCC and CCA cells. Significant



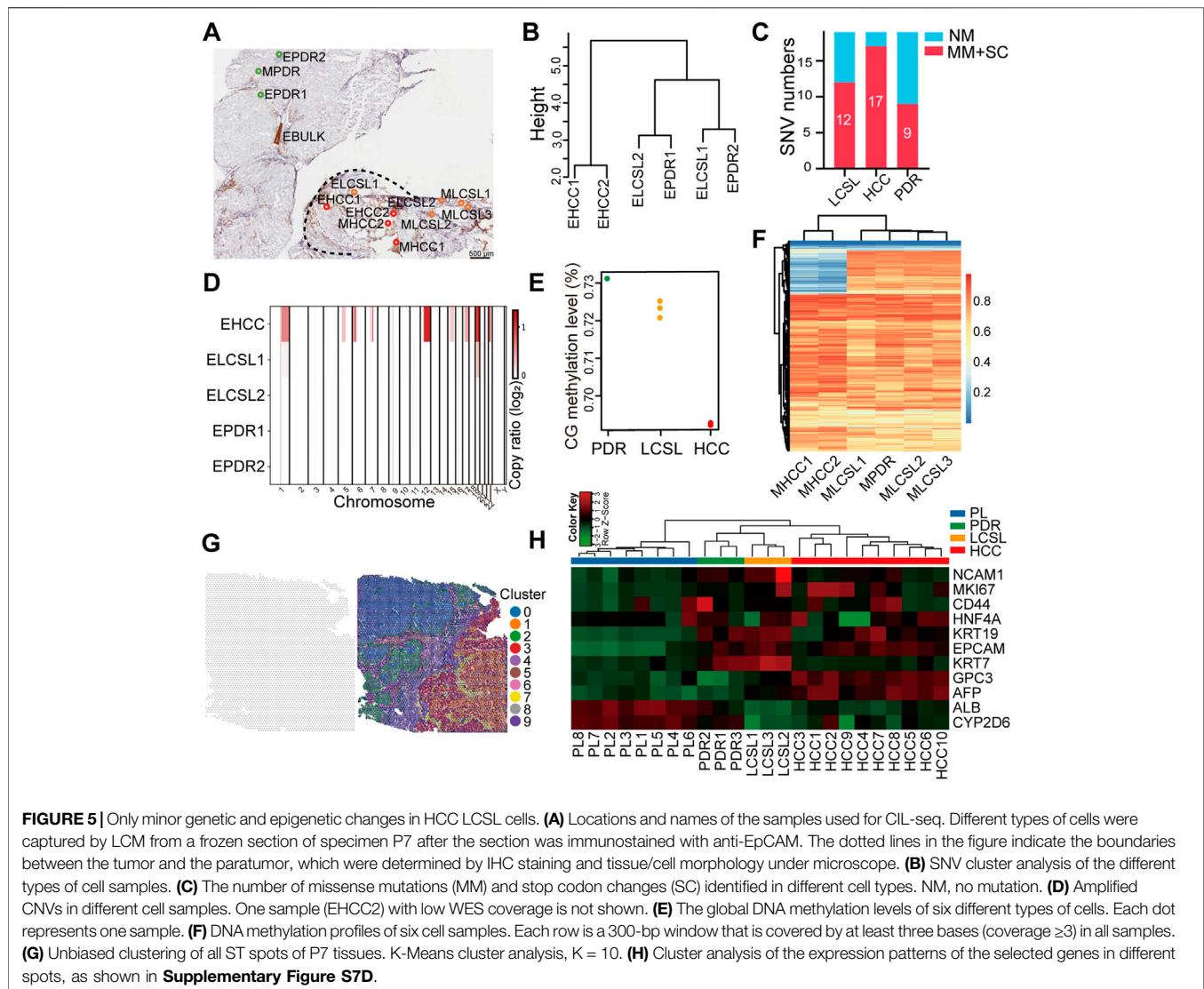
hypermethylation changes in repressive Polycomb group (PcG) protein-marked regions (including the chromatin states 10_TssBiv, 11_BivFlnk, 12_EnhBiv and 13_ReprPC) occurred in HCC and/or CCA cells (Figures 4A,B), which is consistent with previous findings that PcG-marked genes in both embryonic and adult progenitor systems have a higher chance of becoming hypermethylated in cancers (Easwaran et al., 2012). In addition, we also found that two of these PcG-marked regions had already undergone significant changes in LCSL cells compared with PDR cells, including in the bivalent state 10_TssBiv and the chromatin state 13_ReprPC (Figure 4B). The functions of those genes corresponding to these hypermethylated regions in LCSL cells are mainly involved in cell differentiation and fate determination (Figure 4C and Supplementary Table S7), and more than half of them (20/32) also correspond to bivalent genes in the human fetal liver (Yan et al., 2016), such as *APC2* (Supplementary Table S7), a homologue of the *APC* tumor suppressor (van Es et al., 1999). Therefore, hypermethylation changes in these regions in LCSL cells might result in a restrictive chromatin state that can block a differentiation program (Flavahan et al., 2017). Taken together, these results suggest that although the DNA methylation changes in LCSL cells were much less than those in HCC and CCA cells, the changes in LCSL cells might also result in a restrictive chromatin state, which might block or inhibit a normal differentiation program of

these cells, resulting in the persistent abnormal proliferation of the cells.

Only Minor Genetic and Epigenetic Changes in LCSL Cells Compared With Other Tumor Parenchymal Cells in HCC

HCC is a major type of PLC. We collected frozen tissues and prepared sections from 23 PLC cases (Supplementary Table S1). One of them (patient P7) was pathologically diagnosed with HCC and had obvious EpCAM-positive LCSL cells. We performed WES and WGBS analysis of this case with CIL-seq (Figure 5A and Supplementary Table S2). The results showed that, except for the genetic and epigenetic heterogeneity in the LCSL cells, the genetic difference (SNV and CNV) or DNA methylation pattern difference between LCSL cells and PDR cells was small (Figures 5B–F), consistent with the analysis results of patients P2 and P4. Interestingly, among the minor genetic changes in LCSL cells, we found a stop gain mutation in the driver gene *MGA* (Bailey et al., 2018), which also appeared in an HCC sample (Supplementary Table S3). *MGA* is significantly mutated in lung adenocarcinoma and participates in the negative regulation of *MYC* (Llabata et al., 2020). Thus, *MGA* mutation in certain LCSL cells may confer them a proliferative and evolutionary advantage.

Finally, we aimed to determine whether the RNA expression profiles of LCSL cells are similar to those of paratumor liver cells,



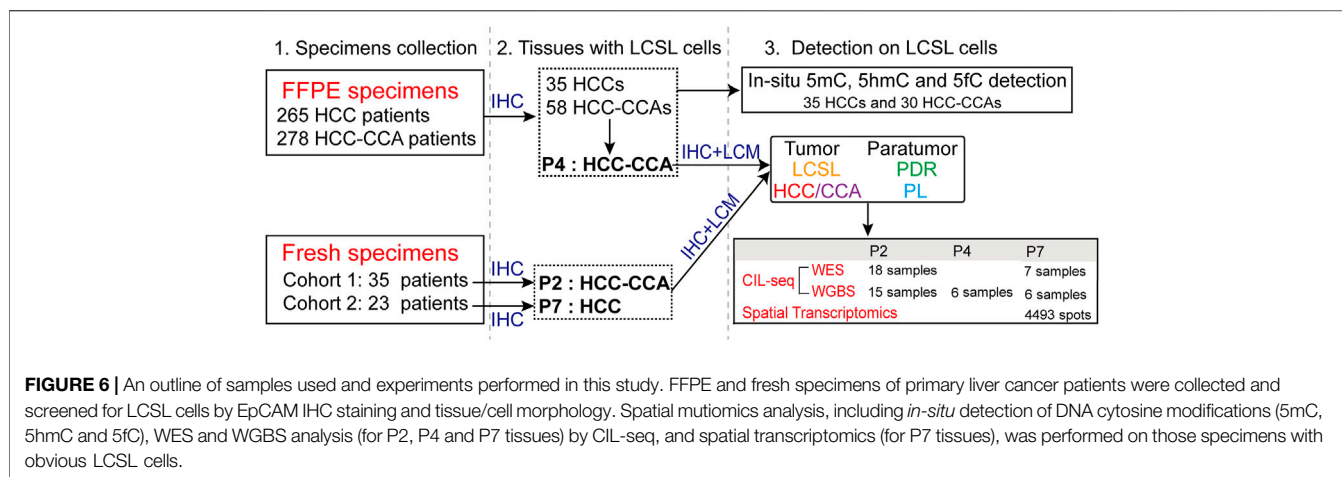
since the genetic and epigenetic changes in LCSL cells are not significant. We performed a spatial transcriptomic (ST) analysis of P7 frozen tissues. We generated 4493 spot transcriptomes at a medium depth of 14,455 UMIS/spot and 4064 genes/spot (**Supplementary Figure S8A**), which can clearly distinguish tumor and paratumor cells (**Supplementary Figure S8B**). In general, there were large RNA expression differences between the tumor and paratumor cells (**Supplementary Figure S8C**). Further cluster analysis showed that the tumor cells had extensive heterogeneity (**Figure 5G**). Because of the small proportion of LCSL cells, we manually selected all LCSL cell spots in the tumor tissues and PDR cell spots in the paratumor tissues (**Supplementary Figures S8D,E**). In addition, we randomly selected some spots of tumor or paratumor parenchymal cells. Cluster analysis of the whole expression profiles showed that LCSL cells and PDR cells were clustered together (**Supplementary Figure S8F**). The cluster analysis of the expression profiles of 11 liver-related genes of these spots showed similar results (**Figure 5H**), indicating that there was

only a small difference in RNA expression patterns between LCSL cells and PDR cells, consistent with the results of the genetic and epigenetic changes.

DISCUSSION

In the present study, using spatial multiomics analysis (**Figure 6**), we provide evidence that only a few genetic and epigenetic changes occur in human LCSL cells compared with other tumor parenchymal cells. Our data showed that HCC and CCA cells have undergone considerable genetic and epigenetic changes that we usually see in human cancers, but only minor genetic and epigenetic changes have occurred in LCSL cells, although some of the changes may be important for CSC formation.

In the past few years, there have been many studies on the genetics and DNA methylation of various human cancers, including various types of liver cancer, and found a wide



range of changes in tumor tissues. However, these studies were mainly focused on local bulk tumor tissues, and there were few studies and little information about CSCs directly from patients. Because of the possible great genetic and epigenetic differences between CSCs and other tumor parenchymal cells, the detection of genetic and epigenetic changes in whole tumor tissue could not reflect what has happened in CSCs. The sequencing signals from bulk tumor samples would be dominated by major tumor parenchymal cells, rendering rare CSCs undetectable. Therefore, treatments that target most of the significant genetic and epigenetic changes (and hence the changes in RNA level and protein level) in these tumors may not be effective for CSCs because CSCs may not possess those changes.

The LCSL cells in the patients in this study harbored CSC characteristics, although they closely resembled PDR cells in cell morphology, the expression of some key markers of liver stem cells (**Supplementary Figure S1B**), and the global levels of DNA cytosine modifications (**Figure 1D**). The CSC characteristics of LCSL cells include 1) the expression of some commonly recognized liver stem cell markers, 2) a higher proliferative capacity than PDR cells (**Figure 1B**) and 3) histopathological (**Figure 1A**) and genetic evidence (**Figure 2H**) suggesting that LCSL cells may be the common origin of HCC and CCA cells. In addition, a very small number of key genetic or epigenetic changes, for example, *MYC* amplification (**Supplementary Figure S6**), were identified in these LCSL cells, but not in PDR cells. Taken together, these characteristics of LCSL cells clearly indicate that LCSL cells have characteristics of both tumor cells and stem cells and are different from paratumor PDR cells.

The proportion of CSCs or LCSL cells in some mixed HCC-CCAs is usually much higher than that in other types of tumors; therefore, such mixed HCC-CCAs can provide us with an ideal tumor object to help investigate human CSCs. At present, mixed HCC-CCA can only be diagnosed by pathological examination after surgery, and the proportion of mixed HCC-CCA in PLC is very low, so it is not easy to collect many samples of mixed HCC-CCA for frozen sections. However, DNA methylation data obtained by the CIL-seq

method from FFPE samples (**Figures 3D–F**) could also provide useful information to support our conclusion. In addition, *in situ* detection of the global level of cytosine modifications (**Figure 1D**) and the results obtained from HCC tissues (**Figure 5**) all support our conclusion that only minor genetic and epigenetic changes occurred in LCSL cells.

The epigenetic changes in LCSL cells of P2 patient are interesting, although the changes are minor compared with those in HCC cells or CCA cells. Since there are only minor genetic changes in LCSL cells, we propose that epigenetic changes in LCSL cells may be important for cancer initiation and evolution. Through restricting changes in gene activity, chromatin structures and regulators may increase the heights of energy walls between cell states and resist changes in cell identity (Yuan et al., 2019). Our results showed that a few crucial chromatin regions in LCSL cells, such as repressive PcG protein-marked regions, had a tendency of hypermethylation. Hypermethylation changes in these regions might result in a restrictive chromatin state in LCSL cells, which can block a differentiation program (Yuan et al., 2019). Thus, epigenetic changes in LCSL cells might block or inhibit the normal differentiation program of these cells, resulting in persistent abnormal proliferation and evolution of the cells.

In summary, our findings revealed the important genetic and epigenetic properties and changes in the LCSL cells of liver cancer patients. In particular, we found only minor genetic and epigenetic changes in LCSL cells compared with other tumor parenchymal cells. Further studies are needed to determine whether such properties and changes are common in CSCs in patients with other types of cancer. Identifying few but key genetic and epigenetic changes in both cancer stem-like cells (or CSCs) and other tumor parenchymal cells of each specific cancer patient, such as *MYC* amplification in patient P2 or *MGA* mutation in patient P7 in this study, could help design more effective methods to treat human cancer in the future.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The raw sequencing data were deposited in the

National Omics Data Encyclopedia under accession number OEP000991 (<https://www.biosino.org/node/project/detail/OEP000991>).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethical committee of the Eastern Hepatobiliary Surgery Hospital (EHBH) in Shanghai, China. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YH and HW designed and supervised the research work. YH wrote the manuscript with assistance from HL and DL. HL performed the bioinformatics analyses with assistance from LY, LC, and YL. DL developed techniques and performed experiments with assistance from MQ and HD.

REFERENCES

- Aizarani, N., Saviano, A., SagarMaily, L., Durand, S., Herman, J. S., et al. (2019). A Human Liver Cell Atlas Reveals Heterogeneity and Epithelial Progenitors. *Nature* 572 (7768), 199–204. doi:10.1038/s41586-019-1373-2
- Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., et al. (2012). methylKit: A Comprehensive R Package for the Analysis of Genome-wide DNA Methylation Profiles. *Genome Biol.* 13 (10), R87. doi:10.1186/gb-2012-13-10-r87
- Auweru, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinformatics* 43 (1110), 11.10.1–11.10.33. doi:10.1002/0471250953.bi1110s43
- Bailey, M. H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., et al. (2018). Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* 173 (2), 371–385.e18. doi:10.1016/j.cell.2018.02
- Battle, E., and Clevers, H. (2017). Cancer Stem Cells Revisited. *Nat. Med.* 23 (10), 1124–1134. doi:10.1038/nm.4409
- Black, J. R. M., and McGranahan, N. (2021). Genetic and Non-genetic Clonal Diversity in Cancer Evolution. *Nat. Rev. Cancer* 21 (6), 379–392. doi:10.1038/s41568-021-00336-2
- Blokzijl, F., de Ligt, J., Jager, M., Sasselli, V., Roerink, S., Sasaki, N., et al. (2016). Tissue-specific Mutation Accumulation in Human Adult Stem Cells During Life. *Nature* 538 (7624), 260–264. doi:10.1038/nature19768
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170
- Brunner, S. F., Roberts, N. D., Wylie, L. A., Moore, L., Aitken, S. J., Davies, S. E., et al. (2019). Somatic Mutations and Clonal Dynamics in Healthy and Cirrhotic Human Liver. *Nature* 574 (7779), 538–542. doi:10.1038/s41586-019-1670-9
- Brunt, E., Aishima, S., Clavien, P.-A., Fowler, K., Goodman, Z., Gores, G., et al. (2018). cHCC-CCA: Consensus Terminology for Primary Liver Carcinomas with Both Hepatocytic and Cholangiocytic Differentiation. *Hepatology* 68 (1), 113–126. doi:10.1002/hep.29789
- Cancer Genome Atlas Research Network (2017). Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* 169 (7), 1327–1341.e23. doi:10.1016/j.cell.2017.05.046

FUNDING

This work was supported by the National Natural Science Foundation of China (81972644 to YH, 91859205, 81830054 and 81988101 to HW, 31771472 to HL) and the National Science and Technology Major Project of China (2018ZX10723204 to YH). HL was supported by the CAS Youth Innovation Promotion Association, SA-SIBS Scholarship Program.

ACKNOWLEDGMENTS

We thank L. Guo for assistance with patient specimen acquisition and D. Cao for experimental help.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2022.810687/full#supplementary-material>

- Chow, E. K.-H., Fan, L.-L., Chen, X., and Bishop, J. M. (2012). Oncogene-specific Formation of Chemoresistant Murine Hepatic Cancer Stem Cells. *Hepatology* 56 (4), 1331–1341. doi:10.1002/hep.25776
- Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., et al. (2013). Sensitive Detection of Somatic Point Mutations in Impure and Heterogeneous Cancer Samples. *Nat. Biotechnol.* 31 (3), 213–219. doi:10.1038/nbt.2514
- Easwaran, H., Johnstone, S. E., Van Neste, L., Ohm, J., Mosbrugger, T., Wang, Q., et al. (2012). A DNA Hypermethylation Module for the Stem/progenitor Cell Signature of Cancer. *Genome Res.* 22 (5), 837–849. doi:10.1101/gr.131169.111
- Flavahan, W. A., Gaskell, E., and Bernstein, B. E. (2017). Epigenetic Plasticity and the Hallmarks of Cancer. *Science*, 357 (6348), eaal2380. doi:10.1126/science.aal2380
- Genomes Project ConsortiumAuton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., et al. (2015). A Global Reference for Human Genetic Variation. *Nature* 526 (7571), 68–74. doi:10.1038/nature15393
- Gravina, S., Dong, X., Yu, B., and Vijg, J. (2016). Single-cell Genome-wide Bisulfite Sequencing Uncovers Extensive Heterogeneity in the Mouse Liver Methylome. *Genome Biol.* 17 (1), 150. doi:10.1186/s13059-016-1011-3
- Greaves, M., and Maley, C. C. (2012). Clonal Evolution in Cancer. *Nature* 481 (7381), 306–313. doi:10.1038/nature10762
- He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., et al. (2011). Tet-mediated Formation of 5-carboxylcytosine and its Excision by TDG in Mammalian DNA. *Science* 333 (6047), 1303–1307. doi:10.1126/science.1210944
- Huch, M., Gehart, H., van Boxtel, R., Hamer, K., Blokzijl, F., Verstegen, M. M. A., et al. (2015). Long-term Culture of Genome-Stable Bipotent Stem Cells from Adult Human Liver. *Cell* 160 (1-2), 299–312. doi:10.1016/j.cell.2014.11.050
- Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., et al. (2012). VarScan 2: Somatic Mutation and Copy Number Alteration Discovery in Cancer by Exome Sequencing. *Genome Res.* 22 (3), 568–576. doi:10.1101/gr.129684.111
- Krueger, F., and Andrews, S. R. (2011). Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications. *Bioinformatics* 27 (11), 1571–1572. doi:10.1093/bioinformatics/btr167
- Li, H., and Durbin, R. (2009). Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* 25 (14), 1754–1760. doi:10.1093/bioinformatics/btp324

- Lin, D.-C., Mayakonda, A., Dinh, H. Q., Huang, P., Lin, L., Liu, X., et al. (2017). Genomic and Epigenomic Heterogeneity of Hepatocellular Carcinoma. *Cancer Res.* 77 (9), 2255–2265. doi:10.1158/0008-5472.CAN-16-2822
- Linker, S. M., Urban, L., Clark, S. J., Chhatriwala, M., Amatya, S., McCarthy, D. J., et al. (2019). Combined Single-Cell Profiling of Expression and DNA Methylation Reveals Splicing Regulation and Heterogeneity. *Genome Biol.* 20 (1), 30. doi:10.1186/s13059-019-1644-0
- Liu, J., Jiang, J., Mo, J., Liu, D., Cao, D., Wang, H., et al. (2019). Global DNA 5-Hydroxymethylcytosine and 5-Formylcytosine Contents Are Decreased in the Early Stage of Hepatocellular Carcinoma. *Hepatology* 69 (1), 196–208. doi:10.1002/hep.30146
- Llabata, P., Mitsuishi, Y., Choi, P. S., Cai, D., Francis, J. M., Torres-Diz, M., et al. (2020). Multi-Omics Analysis Identifies MGA as a Negative Regulator of the MYC Pathway in Lung Adenocarcinoma. *Mol. Cancer Res.* 18 (4), 574–584. doi:10.1158/1541-7786.MCR-19-0657
- Marquardt, J. U., Andersen, J. B., and Thorgeirsson, S. S. (2015). Functional and Genetic Deconstruction of the Cellular Origin in Liver Cancer. *Nat. Rev. Cancer* 15 (11), 653–667. doi:10.1038/nrc4017
- Marusyk, A., Janiszewska, M., and Polyak, K. (2020). Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. *Cancer cell* 37 (4), 471–484. doi:10.1016/j.ccell.2020.03.007
- Matsumoto, T., Takai, A., Eso, Y., Kinoshita, K., Manabe, T., Seno, H., et al. (2017). Proliferating EpCAM-Positive Ductal Cells in the Inflamed Liver Give Rise to Hepatocellular Carcinoma. *Cancer Res.* 77 (22), 6131–6143. doi:10.1158/0008-5472.CAN-17-1800
- Mazor, T., Pankov, A., Song, J. S., and Costello, J. F. (2016). Intratumoral Heterogeneity of the Epigenome. *Cancer Cell* 29 (4), 440–451. doi:10.1016/j.ccell.2016.03.009
- Munoz-Garrido, P., and Rodrigues, P. M. (2019). The Jigsaw of Dual Hepatocellular-Intrahepatic Cholangiocarcinoma Tumours. *Nat. Rev. Gastroenterol. Hepatol.* 16 (11), 653–655. doi:10.1038/s41575-019-0185-z
- Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., et al. (2015). Integrative Analysis of 111 Reference Human Epigenomes. *Nature* 518 (7539), 317–330. doi:10.1038/nature14248
- Rozenblatt-Rosen, O., Regev, A., Oberdoerffer, P., Nawy, T., Hupalowska, A., Rood, J. E., et al. (2020). The Human Tumor Atlas Network: Charting Tumor Transitions Across Space and Time at Single-Cell Resolution. *Cell* 181 (2), 236–249. doi:10.1016/j.cell.2020.03.053
- Sato, K., Marzioni, M., Meng, F., Francis, H., Glaser, S., and Alpini, G. (2019). Ductular Reaction in Liver Diseases: Pathological Mechanisms and Translational Significances. *Hepatology* 69 (1), 420–430. doi:10.1002/hep.30150
- Schulze, K., Imbeaud, S., Letouze, E., Alexandrov, L. B., Calderaro, J., Rebouissou, S., et al. (2015). Exome Sequencing of Hepatocellular Carcinomas Identifies New Mutational Signatures and Potential Therapeutic Targets. *Nat. Genet.* 47 (5), 505–511. doi:10.1038/ng.3252
- Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., et al. (2001). dbSNP: The NCBI Database of Genetic Variation. *Nucleic Acids Res.* 29 (1), 308–311. doi:10.1093/nar/29.1.308
- Smallwood, S. A., Lee, H. J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., et al. (2014). Single-cell Genome-wide Bisulfite Sequencing for Assessing Epigenetic Heterogeneity. *Nat. Methods* 11 (8), 817–820. doi:10.1038/nmeth.3035
- Stahl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., et al. (2016). Visualization and Analysis of Gene Expression in Tissue Sections by Spatial Transcriptomics. *Science* 353 (6294), 78–82. doi:10.1126/science.aaf2403
- Talevich, E., Shain, A. H., Botton, T., and Bastian, B. C. (2016). CNVkit: Genome-wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *Plos Comput. Biol.* 12 (4), e1004873. doi:10.1371/journal.pcbi.1004873
- van Es, J. H., Kirkpatrick, C., van de Wetering, M., Molenaar, M., Miles, A., Kuipers, J., et al. (1999). Identification of APC2, a Homologue of the Adenomatous Polyposis Coli Tumour Suppressor. *Curr. Biol.* 9 (2), 105–S2. doi:10.1016/s0960-9822(99)80024-4
- Vicente-Dueñas, C., Hauer, J., Cobaleda, C., Borkhardt, A., and Sánchez-García, I. (2018). Epigenetic Priming in Cancer Initiation. *Trends Cancer* 4 (6), 408–417. doi:10.1016/j.trecan.2018.04.007
- Wang, A., Wu, L., Lin, J., Han, L., Bian, J., Wu, Y., et al. (2018). Whole-exome Sequencing Reveals the Origin and Evolution of Hepato-Cholangiocarcinoma. *Nat. Commun.* 9 (1), 894. doi:10.1038/s41467-018-03276-y
- Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data. *Nucleic Acids Res.* 38 (16), e164. doi:10.1093/nar/gkq603
- Xu, G.-L., and Bochtler, M. (2020). Reversal of Nucleobase Methylation by Dioxygenases. *Nat. Chem. Biol.* 16 (11), 1160–1169. doi:10.1038/s41589-020-00675-5
- Xue, R., Chen, L., Zhang, C., Fujita, M., Li, R., Yan, S.-M., et al. (2019). Genomic and Transcriptomic Profiling of Combined Hepatocellular and Intrahepatic Cholangiocarcinoma Reveals Distinct Molecular Subtypes. *Cancer cell* 35 (6), 932–947.e8. doi:10.1016/j.ccell.2019.04.007
- Yamashita, T., Honda, M., Nakamoto, Y., Baba, M., Nio, K., Hara, Y., et al. (2013). Discrete Nature of EpCAM+ and CD90+ cancer Stem Cells in Human Hepatocellular Carcinoma. *Hepatology* 57 (4), 1484–1497. doi:10.1002/hep.26168
- Yamashita, T., Ji, J., Budhu, A., Forgues, M., Yang, W., Wang, H. Y., et al. (2009). EpCAM-positive Hepatocellular Carcinoma Cells Are Tumor-Initiating Cells with Stem/progenitor Cell Features. *Gastroenterology* 136 (3), 1012–1024. doi:10.1053/j.gastro.2008.12.004
- Yan, L., Guo, H., Hu, B., Li, R., Yong, J., Zhao, Y., et al. (2016). Epigenomic Landscape of Human Fetal Brain, Heart, and Liver. *J. Biol. Chem.* 291 (9), 4386–4398. doi:10.1074/jbc.M115.672931
- Yuan, S., Norgard, R. J., and Stanger, B. Z. (2019). Cellular Plasticity in Cancer. *Cancer Discov.* 9 (7), 837–851. doi:10.1158/2159-8290.CD-19-0015
- Zhai, W., Lim, T. K.-H., Zhang, T., Phang, S.-T., Tiang, Z., Guan, P., et al. (2017). The Spatial Organization of Intra-tumour Heterogeneity and Evolutionary Trajectories of Metastases in Hepatocellular Carcinoma. *Nat. Commun.* 8, 4565. doi:10.1038/ncomms14565
- Zhang, Q., Lou, Y., Yang, J., Wang, J., Feng, J., Zhao, Y., et al. (2019). Integrated Multiomic Analysis Reveals Comprehensive Tumour Heterogeneity and Novel Immunophenotypic Classification in Hepatocellular Carcinomas. *Gut* 68 (11), 2019–2031. doi:10.1136/gutjnl-2019-318912
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6
- Zhu, M., Lu, T., Jia, Y., Luo, X., Gopal, P., Li, L., et al. (2019). Somatic Mutations Increase Hepatic Clonal Fitness and Regeneration in Chronic Liver Disease. *Cell* 177 (3), 608–621.e12. doi:10.1016/j.cell.2019.03.026

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

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

Copyright © 2022 Liu, Li, Dong, Qu, Yang, Chen, Li, Wang and He. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.