SETD1A drives stemness by reprograming epigenetic landscape in hepatocellular carcinoma stem cells

Jianxu Chen\textsuperscript{1,2,3#}, Zhijie Xu\textsuperscript{1,2#}, Hongbin Huang\textsuperscript{1,2#}, Yao Tang\textsuperscript{1,2}, Hong Shan\textsuperscript{1,4*}, Fei Xiao\textsuperscript{1,2*}

\textsuperscript{1}Guangdong Provincial Engineering Research Center of Molecular Imaging, the Fifth Affiliated Hospital, Sun Yat-Sen University, Zhuhai 519000, China.

\textsuperscript{2}Department of Infectious Diseases, the Fifth Affiliated Hospital, Sun Yat-Sen University, Zhuhai, Guangdong Province, China.

\textsuperscript{3}Department of Gastroenterology, the Fifth Affiliated Hospital, Sun Yat-Sen University, Zhuhai, Guangdong Province, China.

\textsuperscript{4}Center for Interventional Medicine, the Fifth Affiliated Hospital, Sun Yat-Sen University, Zhuhai, Guangdong Province, China.

\#These authors contributed equally to this work.

*Corresponding authors: Hong Shan, Guangdong Provincial Engineering Research Center of Molecular Imaging, the Fifth Affiliated Hospital, Sun Yat-Sen University, 52 Mei Hua East Road, Zhuhai 51900, Guangdong Province, China. Tel/Fax: 86-756-2528527. E-mail: shanhong@mail.sysu.edu.cn;

Fei Xiao, Guangdong Provincial Engineering Research Center of Molecular Imaging, the Fifth Affiliated Hospital, Sun Yat-Sen University, 52 Mei Hua East Road, Zhuhai 51900, Guangdong Province, China. Tel/Fax: 86-765-2528527. E-mail: xiaof35@sysu.edu.cn.
Abstract
Cancer stem cells (CSCs) are responsible for tumor progression and recurrence. However, the mechanisms regulating hepatocellular carcinoma (HCC) stemness remain unclear. Applying a genome-scale CRISPR knockout screen, we identify that the H3K4 methyltransferase SETD1A and other members of Trithorax group proteins drive cancer stemness in HCC. SETD1A is positively correlated with poor clinical outcome in HCC patients. Combination of SETD1A and serum AFP significantly improves the accuracy of predicting HCC relapse. Mechanistically, SETD1A mediates transcriptional activation of various histone-modifying enzymes, facilitates deposition of H3K4me3 and H3K27me3 and activates oncogenic enhancers and super-enhancers, leading to activation of oncogenes and inactivation of tumor suppressor genes simultaneously in liver CSCs. In addition, SETD1A cooperates with PABPC1 to regulate H3K4me3 modification on oncogenes. Our data pinpoint SETD1A as a key epigenetic regulator driving HCC stemness and progression, highlighting the potential of SETD1A as a candidate target for HCC intervention and therapy.

Keywords: hepatocellular carcinoma; cancer stem cells; CRISPR screening; SETD1A; histone modification; PABPC1.

Introduction
Hepatocellular carcinoma (HCC) is one of the leading causes of cancer-related death and displays high level of heterogeneity, which limits the efficacy of clinical treatment. The heterogeneity of HCC has been reported to be driven by cancer stem cells (CSCs), which are a subpopulation of tumor cells with self-renewal capacity and
differentiation potential (1). Since John Dick has isolated leukemia stem cells in 1994, accumulating evidence have shown that CSCs are responsible for cancer drug resistance, relapse, and metastasis. Therefore, CSCs represent promising targets for cancer therapy (2). Understanding determinants and regulatory mechanism of CSCs self-renewal is important in developing CSC-targeted therapy.

Cell surface markers CD24, CD133, CD90, EpCAM, and CD13 are frequently used to identify and isolate liver CSCs (3). However, liver CSCs expressing different markers have discrete characteristics and tumor-initiating capacity. For example, EpCAM+ liver CSCs have features of epithelial cells and are associated with rapid growth, while CD90+ liver CSCs present features of vascular endothelial cells and with high incidence of distant organ metastasis(4). Compared to CD24+ liver CSCs and CD133+ liver CSCs, CD24+CD133+ liver CSCs have a stronger tumor-initiating capacity, As few as ten CD24+CD133+ liver CSCs could initiate tumorigenicity in NOD/SCID mice, patients with CD24+CD133+ liver CSCs have worse clinical outcome, revealing the critical role of CD24+CD133+ CSCs in HCC tumorigenesis (5).

High-throughput CRISPR screening is a powerful tool for discovering key regulators in cancer development, progression and resistance, with the advantage of low-noise, minimal off-target effects and consistently high efficiency over RNA interfere (RNAi) based genetic screening (6). Using kinome CRISPR knockout library, kinases CDK7, CDK12 and CDC7 have been identified to regulate HCC progression (7). Here, we established a genome-wide CRISPR knockout screening in CD24+CD133+ liver CSCs and identified the H3K4 methyltransferase SETD1A as a critical diver for HCC stemness and progression, providing a therapeutic target for HCC treatment.
Results

A genome-wide CRISPR screening for factors contributing to HCC stemness

We used pooled genome-wide CRISPR/Cas9 knockout library (GeCKO v2) consisting of 58,028 gRNAs targeting 19,009 genes (three gRNAs per gene) to investigate genes contributing to stemness of CD24+CD133+ HCC in PLC cell (Fig. 1A and Supplemental Fig. 1A). We ranked sgRNAs that were enriched in CD24+CD133+ non-CSCs compared with non-sorted library cell population using MAGeCK robust ranking algorithm (RRA) (Fig. 1B). We identified 504 genes contributing to HCC stemness, including ABCG1, CDK16, and MYCT1, which have been reported to drive HCC stemness (Supplemental Fig. 1B and Supplemental Fig. 1C) (8-10).

Interestingly, we found several top hits belonging to Trithorax group (TrxG) and Polycomb group (PcG) proteins family, including KMT2A (also known as MLL/MLL1), KMT2D (also known as MLL2/MLL4), SETD1A, SMARCD1, SMARCE1, PCGF5, and CBX4 (Supplemental Fig. 1D). TrxG and PcG proteins activate oncogenes transcription and inhibit transcription of tumor suppressors through regulating histone modification and higher-order chromatin structure(11). PcG protein CBX4 has been reported to promote HCC stemness and increase sorafenib resistance in advanced HCC (12). Another PcG protein PCGF5 is also associated with CSCs expansion (13). TrxG proteins SMARCD1 and SMARCE1 have been reported to promote tumor progression (14, 15). KMT2A and KMT2D are the targets of mutant P53R249S which is the most common missense mutation of p53 (16). In addition, KMT2A is essential for HGF-MET signaling induced HCC
metastasis (17). Gene ontology (GO) enrichment analysis of the 504 hits indicated that GO terms of H3K4 methylation, post-translation protein modification and SWI/SNF complex were prominently enriched, revealing these epigenetic regulators, especially TrxG proteins, play a core role in liver CSCs expansion. Other prominently enriched GO terms included WNT pathway (18) and non-canonical NF-κB pathway (19), which also have been proved to promote HCC stemness (Fig. 1C). Collectively, our CRISPR screening identified TrxG and PcG proteins as key factors contributing to HCC stemness.

**SETD1A promotes the HCC stemness in vitro and in vivo**

Histone methyltransferase SETD1A, which is the catalytic subunit of the SET1/COMPASS complex containing WDR5, ASH2L, RBBP5, DPY30, SETD1B, HCF1, WDR82 and CFP1, was among the top-ranking genes in the screen analysis. gRNAs targeting SETD1A were significantly enriched in non-CSCs (Fig. 1D). It has been reported that SETD1A is associated with cell differentiation, development and tumor progression(20). A previous study using hepatocyte-like cells differentiated from human induced pluripotency stem cells (iPSC) showed that SETD1A incorporated with CUDR to trigger the malignant transformation of hepatocyte cells. Overexpression of SETD1A alone failed to promote malignant transformation (21). However, our data revealed SETD1A as an independent regulator of HCC stemness. The discrepancy between our results and the previous finding could be explained by the iPSCs-derived hepatocyte-like cells used in the previous study, which couldn’t fully display functional characteristics of hepatic stem/progenitor cells (3, 22). To further investigate the role of SETD1A in HCC stemness, we first used ATAC-seq assay to determine its chromatin states in liver CSCs and found its promoter region
was more accessible in CD24⁺CD133⁺ liver CSCs than in CD24⁻CD133⁻ non-CSCs (Fig. 1E). qRT-PCR assay also showed that \( \text{SETD1A} \) was significantly upregulated in CD24⁺CD133⁺ HCC CSCs compared to CD24⁺CD133⁻ non-CSCs, confirming the result of ATAC-seq assay (Fig. 1F). These results demonstrated \( \text{SETD1A} \) was transcriptionally activated in liver CSCs.

To validate the function of \( \text{SETD1A} \) in HCC stemness, we established stable HCC cell lines of PLC, Huh-7 and Hep3B with shRNA targeting \( \text{SETD1A} \), and confirmed its knockdown by western blotting and qRT-PCR assay (Supplemental Fig. 1E and Supplemental Fig. 1F). Notably, \( \text{SETD1A} \) knockdown significantly decreased the proportion of CD24⁺CD133⁺ liver CSCs (Fig. 1G, Supplemental Fig. 1G). Meanwhile, we observed that \( \text{SETD1A} \) knockdown significantly reduced the proportion of EpCAM⁺ liver CSCs (Fig. 1H). Spheroid-formation assays revealed \( \text{SETD1A} \) knockdown effectively inhibited spheroid formation (Fig. 1I). Moreover, we assessed the tumorigenicity of serial dilutions of \( \text{SETD1A} \) knockdown HCC cells in NOD/SCID mice. As shown in Fig. 1J and 1K, \( \text{SETD1A} \) knockdown significantly reduced tumor initiating capacity and CSCs frequency. Kalpan-Meier survival assay showed that the mice transplanted with \( \text{SETD1A} \) knockdown HCC cells had significantly longer survival time than those transplanted with scramble control HCC cells (Fig. 1L). Taken together, we demonstrated that \( \text{SETD1A} \) promotes HCC stemness \textit{in vitro} and \textit{in vivo}. In addition, we found \( \text{SETD1A} \) significantly promoted HCC cell proliferation, invasion, migration, sorafenib resistance and induced EMT (Supplemental Fig 2), suggesting \( \text{SETD1A} \) promotes HCC progression.

\textbf{SETD1A is upregulated in HCC tissues and associated with poor clinical}
outcome in HCC patients

To investigate the clinical significance of SETD1A in HCC, we determined SETD1A expression in HCC tissues using the TCGA database and Oncomine database. We found SETD1A was significantly upregulated in HCC tissues compared to their matched adjacent normal liver tissues (Fig. 2A and 2B). Its expression was associated with advanced stages of HCC (Fig. 2C). Then, we performed Kalpan-Meier survival assay using TCGA and Kaplan-Meier Plotter databases and found the overall survival and disease-free survival of patients with high SETD1A expression was significantly shorter than those with low SETD1A expression (Fig. 2D-2F). Immunohistochemical analysis of 90 matched pairs of HCC and adjacent normal liver tissues using immunohistochemical staining (IHC) confirmed high expression of SETD1A in HCC tissue (Fig. 2G-2H). SETD1A expression was positively correlated with tumor size, tumor encapsulation and tumor recurrence (Table 1), which was further validated by Univariate and Multivariate Cox proportional hazard regression analyses (Table 2). Higher SETD1A expression was associated with shorter survival times and higher recurrence rate (Fig. 2I-2J). Using Receiver operator characteristic (ROC) analysis, we showed that SETD1A exhibited better performance than serum APF in prediction of HCC relapse (23). Combination of SETD1A and serum AFP significantly improved the accuracy in prediction of HCC relapse (Fig. 2K-2L). Taken together, these findings indicated that high SETD1A expression is associated with poor outcome and relapse in HCC patients.

SETD1A promotes HCC stemness and progression through directly transcriptional activate histone-modifying enzymes.

The SETD1A-generated H3K4me3 modification marks the promoters of actively transcribed genes. To further determine the regulatory mechanism of SETD1A driving
HCC stemness, we used CUT&Tag to investigate the genomic distributions of SETD1A and H3K4me3 in CD24^+CD133^+ liver CSCs. H3K27me3 modification marks the promoters of transcriptionally silent genes. To determine the transcriptional activation of genes, we also investigate the genomic distribution of H3K27me3. CUT&Tag assay showed that SETD1A and H3K4me3 co-occupied the promoters of some known drivers of HCC stemness, such as PRMT6, BMI1, SOX9, ZIC2, ANGPTL4, PDK4 and IRAK1 (24), suggesting SETD1A activates the transcription of these genes. Notably, we found that SETD1A forms a positive feedback loop with itself through binding to its own promoter (Fig. 3A and Supplemental Fig. 3A).

Importantly, the GO analysis of the top 3000 SETD1A-regulated genes showed that chromatin modifying enzymes and histone modification were the prominently enriched (Fig. 3B). Alterations of chromatin modification are linked to dysregulated expression of genes which are critical for tumorigenesis and development (25). SETD1A and H3K4me3 co-occupied the promoters of lots of histone methyltransferases (HMTs), histone demethylases, and histone acetyltransferases (HATs), including KMT2A, KMT2D, H3K27me3 methyltransferase EZH2, H3K9 methyltransferase SETDB1, H3K9me3 and H3K36me3 demethylase KDM4A, H3K9me3 methyltransferase SUV39H2 and SUV39H1, H3K79 methyltransferase DOT1L, H3K20me1/me2 methyltransferase KMT5C, lysine acetyltransferase KAT6A, KAT5, KAT7 and KAT8 (Fig. 3C and Supplemental Fig. 3B). In addition, the promoters of these genes aren’t bound by H3K27me3, suggesting these genes are transcriptional activation in liver CSCs. Among them, EZH2, SETD1B, KDM4A,
DOTL1, SUV39H1, KAT7, KAT8, and KAT5 have been reported to promote HCC stemness, HCC growth, EMT, and metastasis (26). These results revealed that SETD1A promotes the HCC stemness through depositing H3K4me3 on the promoters of various histone modifiers to promote their transcription, highlighting the core role of SETD1A in epigenetic regulation of HCC stemness.

**SETD1A knockdown remolds the chromatin modification states of H3K4me3 and H3K27me3.**

We found that SETD1A and H3K4me3 co-occupied the promoters of EZH2 (Fig. 3C). Knockdown of SETD1A inhibited EZH2 expression (Supplemental Fig. 3C), suggesting SETD1A transcriptionally activates EZH2. In addition, EZH2 catalyzes H3K27me3 and promotes HCC stemness and progression (27). These findings suggested that SETD1A regulates H3K27me3 modification. Therefore, it’s our speculation that SETD1A may promote HCC stemness through depositing H3K4me3 on the promoters of oncogenes and H3K27me3 on the promoters the tumor suppressor genes. Next, we determined the genome-wide profiles of H3K4me3 and H3K27me3 in CD24⁺CD133⁺ liver CSCs and SETD1A knockdown CD24⁺CD133⁺ liver CSCs. A total of 10,464 H3K4me3-marked genes and 5,419 H3K27me3-marked genes were detected in CD24⁺CD133⁺ HCC CSCs. Meanwhile, a total of 10,380 H3K4me3-marked genes and 6,172 H3K27me3-marked genes were detected in SETD1A knockdown CD24⁺CD133⁺ liver CSCs (Supplemental Fig. 3D). In H3K4me3-marked genes, SETD1A knockdown resulted in 15 (~0.15%) genes gaining H3K27me3, 610 (~5.8%) bivalent genes and 355 (~3.4%) genes losing H3K4me3 mark. In
H3K27me3-marked genes, SETD1A knockdown resulted in 17 (~0.3%) genes gaining H3K4me3 and 482 (~8.9%) bivalent genes. In 1,800 bivalent genes, SETD1A knockdown resulted in 79 genes (4.4%) losing H3K27me3 and 72 genes (~4.0%) losing H3K4me3 (Fig. 3D and 3E). On the pathway level, we determined the impact of SETD1A knockdown on the effect of H3K4me3 and H3K27me3 modification dynamics in HCC stemness using GO analysis. GO terms associated with protein kinase activity, autophagy, MAPK signaling and eIF3 complex were significantly enriched among the genes gaining H3K27me3 and losing H3K4me3 (Fig. 3F). Analysis of protein–protein interactions (PPI) analysis of these genes revealed four functional modules including MAPK signaling pathway, autophagy, eIF3 family and protein kinases (Fig. 3G and 3J), all of which have been reported to play a crucial role in HCC stemness as well as cancer progression(28). To determine the function of the genes losing H3K27me3 and gaining H3K4me3 upon SETD1A knockdown, we performed GO and PPI network analysis and found that these genes were involved in collagen biosynthesis, differentiation, and negative regulation of proliferation (Fig. 3H, 3I and 3K). Then we used qRT-PCR and western blot to verify above findings and found SETD1A knockdown significantly inhibited genes associated with MAPK pathway, autophagy, and protein kinase activity, such as RAG1, ATG7, MAPK8, MTM1, CCNL1, and EFNA5. Meanwhile, SETD1A knockdown significantly upregulated genes associated with negative regulation of proliferation, differentiation and collagen biosynthesis, such as MSX2, ZFPM2, FGF9, PAX5, COL1A1, TLL1, and TLL2 (Supplemental Fig. 4A). Western blotting also showed
SETD1A knockdown promoted Collagen 1 expression (Supplemental Fig. 4B). Taken together, these results suggested that SETD1A increases promoter activity of oncogenes and inhibiting promoter activity of tumor suppressor genes to drive HCC stemness.

**SETD1A drives HCC stemness through administering H3K27ac deposition.**

Enhancer malfunction drives the aberrant regulation of oncogenes in cancer. Considering H3K27ac that marks active enhancers is directly blocked by H3K27me3 and have a synergistic effect with H3K4me3 on tumor progression(29, 30). Thus, we determined the H3K27ac-enriched enhancer regions in CD24⁺CD133⁺ CSCs and SETD1A knockdown CD24⁺CD133⁺ CSCs, and found that SETD1A knockdown resulted in the change of H3K27ac distribution. We observed loss of H3K27ac mark in the enhancers of 442 genes, including lncRNA MALAT1, oncogenes ZFX and EPS8, as well as gain of H3K27ac mark in the enhancers of 134 genes including the tumor suppressor PPP1R12B (31, 32) (Fig. 4A). Gene set enrichment analysis (GSEA) of the genes associated with loss of H3K27ac revealed that SETD1A knockdown resulted in loss of SETD1A binding sites signature, histone methyltransferase complex signature, liver cancer growth associated, doxorubicin resistance and MYC target genes, all of which have been reported to be associated with CSCs stemness, therapeutic resistance and tumor progression (33) (Fig. 4B). In addition, GO analysis of the genes associated with loss of H3K27ac were prominently enriched in signaling pathways related to CSCs stemness, tumorigenesis and tumor growth (Fig. 4C). All together, these findings suggested that SETD1A increases activity of oncogenic
enhancers to promote HCC stemness and progression.

Super-enhancers (SEs) that recruit a great number of transcription factors and cofactors to confer strong transcriptional regulation have been proved to play an essential role in CSCs self-renewal and tumor progression through increasing transcriptional activity of oncogenes (34). To determine whether SETD1A regulates SEs, we annotated 583 SEs and 749 SEs in CD24⁺CD133⁺ liver CSCs and SETD1A knockdown CD24⁺CD133⁺ liver CSCs, respectively, and observed that SE-associated genes, such as FNDC3B, PTP4A1, PBX1 (35) and ELF3 (28), promote CSCs self-renewal and tumor progression in CD24⁺CD133⁺ liver CSCs (Fig. 4D). SETD1A knockdown changed the landscape of SEs, resulting in activation of the transcription of SE-driven tumor suppressors, such as ST3GAL4, AKAP12, PTPN1, AQP9 and ANKRD11 (Fig. 4E). Overall, these results suggested that SETD1A promotes HCC stemness through increasing oncogenic activity of enhancers and SEs.

**SETD1A drives HCC stemness via interacting with PABPC1 to regulate H3K4me3 modification.**

Next, we used co-immunoprecipitation (co-IP) followed by mass spectrometry (MS) to identify SETD1A-interacting proteins. Notably, we found that RNA-binding protein polyadenylate-binding protein cytoplasmic 1 (PABPC1) interacted with SETD1A and its core complex proteins (Fig. 5A and 5B). Cytoplasmic PABPC1 has been reported to promote the progression of different types of tumors. Mechanistically, PABPC1 mainly is regulated by non-coding RNAs, such as circPTK2 in bladder cancer (36) and lcnRNA SNHG14 in HCC (37). However, the
role of nuclear PABPC1 hasn’t been investigated. Previous studies showed that PABPC1 interacts with AGO2 and eukaryotic initiation factor 4G (eIFG4) in cytoplasm to regulate mRNA translation and HCC proliferation (38, 39). But its role in promoting HCC stemness hasn’t been studied yet. We showed that PABPC1 was upregulated in HCC tissues (Fig. 5C). HCC patients with high PABPC1 expression had shorter survival time than patients with low PABPC1 expression (Fig. 5D). CUT&Tag assay showed that SETD1A directly binds to the promoter sequences of PABPC1 in CD24+CD133+ liver CSCs. ATAC-seq data analysis revealed that PABPC1 promoter is more accessible in CD24+CD133+ liver CSCs than in non-CD24+CD133+ HCC cells (Fig. 5E), suggesting that SETD1A not only interacted with PABPC1, but also regulated the transcriptional activity of PABPC1. To determine whether SETD1A promotes HCC stemness and progression through regulating PABPC1, we overexpressed PABPC1 in SETD1A knockdown HCC cells (Supplemental Fig. 4C). FACS and sphere formation assays showed that PABPC1 overexpression partly reversed the effect of SETD1A knockdown on the percent of CD24+CD133+ population and sphere number (Fig. 5F and 5G). Cell proliferation, and transwell invasion and migration assays showed PABPC1 overexpression partly reversed the effect of SETD1A knockdown on the HCC proliferation, migration, and invasion (Fig. 5H-5J). These findings revealed that SETD1A regulates HCC stemness and progression partly via PABPC1. We then used CUT&Tag assay to determine the genomic distributions of SETD1A, H3K4me3 and PABPC1 in CD24+CD133+ liver CSCs and found that 54% (3019/5552) PABPC1 targets were co-occupied by
SETD1A. 65% (3604/5552) PABPC1 targets were co-occupied by H3K4me3. 81% SETD1A targets were co-occupied by H3K4me3. 45% (2499/5552) PABPC1 targets were co-occupied by H3K4me3 and SETD1A. (Fig. 5K). In addition, GO analysis showed that PABPC1 targets were enriched for GO terms including protein phosphorylation, gene transcription, apoptosis, cell cycle, cell migration, TGFβ signaling, all of which are involved in HCC stemness and progression (Fig. 5L). GO analysis showed that co-occupied targets of SETD1A and PABPC1 as well as PABPC1, SETD1A and H3K4me3 also enriched for protein phosphorylation, gene transcription, apoptosis, cell cycle, cell migration, and TGFβ signaling, indicating that SETD1A interacts with PABPC1, which is essential for the transcriptional activation of SETD1A-regulated oncogenes, to activate the co-regulated targets of SETD1A to promote HCC stemness and progression (Fig. 5M and 5N). Taken together, our data show that SETD1A cooperates with PABPC1 to regulate H3K4me3 modification on the promoters of oncogenes to drive HCC stemness.

Discussion

In this study, we performed a genome-scale CRISPR knockout screening on CD24⁺CD133⁺ liver CSCs for determinants of HCC stemness and found that the most enriched were the genes associated with H3K4 methylation, chromatin remodeling, cell cycle, WNT and non-canonical NF-κB pathway, especially the genes regulating H3K4 methylation, including the H3K4 methyltransferase KMT2A, KMT2D and SETD1A. Among the above three H3K4 methyltransferases, SETD1A mainly
catalyzes H3K4me3 at gene promoters and exhibits the most dramatic effect on global H3K4me3 and gene expression (40). Previous studies have indicated that increased H3K4me3 modification is associated with tumor progression and also a poor prognosis in patients with cancer(41). SETD1A has been found to promote tumors progression in various cancers, such as lung cancer, breast cancer, gastric cancer, colorectal cancer and leukemia through regulating TGFβ, Wnt and Hippo/YAP signaling pathways(42). However, its role in CSCs remains unknown. Therefore, we focused on investigating the function of SETD1A in promoting HCC stemness. We showed that SETD1A expression is upregulated in HCC tissues and positively correlated with poor clinical outcome in HCC patients, exhibiting good performance in predicting HCC relapse. SETD1A knockdown inhibited HCC stemness.

To further investigate the mechanism underlying SETD1A driving HCC stemness, we determined direct targets of SETD1A and its impact on histone modification profiling in liver CSCs. As a tagmentation-based epigenomic profiling method, CUT&Tag has distinct advantages compared to ChIP-seq, such as easy-to-handle, compatibility with reduced cell inputs, low costs and high-quality data, therefore becoming the best fit in CSCs study(43). Applying the CUT&Tag technique, we identified that SETD1A directly targets various histone-modifying enzymes in CD24⁺CD133⁺ liver CSCs, some of which have been reported to promote HCC stemness and progression. Histone modification profiling showed that SETD1A-generated H3K4me3 modifies the promoters of oncogenes involved in activation of WNT, MAPK, EGFR and c-MYB pathways, autophagy and protein kinase activity. In addition, we found that
SETD1A promoted EZH2 transcription through directly binds to its promoter, leading to reprogramming of H3K27me3-modified genomic regions and inhibition of transcription of tumor suppressors genes related to cell proliferation suppression, differentiation and collagen biosynthesis. Meanwhile. We found PABPC1 interacted with SETD1A/COMPASS to recruit SETD1A to the promoters of genes associated with protein phosphorylation, apoptosis, cell cycle, cell migration, and TGFβ signaling. These findings suggest that SETD1A serves as a master regulator for HCC stemness, representing an ideal target for HCC therapy. Accumulating evidence has shown that epigenetic dysregulation contributes to aberrant transcriptional programs that promote tumor stemness and progression. Histone modification plays a central role in epigenetic regulation(44). Interestingly, we found that SETD1A directly binds to the promoters of various histone-modifying enzymes which mediate H3K4me1, H3K4me3, H3K27me3, H3K9me3, H3K36me3, H3K79me3, and H4K16ac modification. Our study reveals that SETD1A drives HCC stemness through epigenetic regulation. These histone modifications and their crosstalk determine the transcriptional outcomes to promote HCC stemness and progression. Thus, it’s worth further studies on SETD1A-mediated histone crosstalk to better understand the mechanism. Our CRISPR screening also revealed that multiple members of TrxG proteins including KMT2A, KMT2D, SETD1A, SMARCE1 and SMARCD1, and PcG proteins including CBX4 and PCGF5 orchestrate HCC stemness, suggesting that TrxG and PcG proteins might cooperate or crosstalk to promote HCC stemness. TrxG
and PcG proteins regulate cellular memory, cell fate determination, tumorigenesis and tumor development via regulating histone modification, chromatin accessibility and chromatin compaction(11). However, there are scarce reports regarding the role of TrxG and PcG proteins in HCC stemness. Among more than 40 family members of TrxG proteins, only BRG1 has been reported to promote HCC stemness through regulating lncRNA lncFZD6(45). Similarly, among PcG proteins family members, BMI1 is the only one that has been reported to promote HCC stemness (46). Our study highlights the role of TrxG and PcG proteins in HCC stemness. Since range of small molecules targeting TrxG and PcG proteins have gained success in preclinical development, such as WDR5-MLL interaction antagonist OICR-9429, MENIN inhibitor MI-403 and EZH2 inhibitor EPZ-6438(47), our data provide the rationale to test these TrxG- or PcG-targeted drugs in the treatment of HCC.

Aberrant landscape of enhancer and SEs results in abnormal transcriptional program, leading to tumorigenesis(48). SETD1A has been found to mediate long-range interactions between enhancer and promoter. Here, we demonstrate that SETD1A promotes activity of oncogenic enhancers and SEs in liver CSCs. It has been recently reported that H3K4me3 reader PHF23 interacts with SIN3-HDAC complex to mediate a synergistic action of H3K4me3 and H3K27ac on inhibition of the deacetylation activity of SIN3-HDAC complex, resulting in activation of tumor suppressor genes(30). This recent finding indicates that H3K4me3 methyltransferase, H3K27ac acetyltransferase or H3K27ac deacetyltransferase might form a complex with H3K4me3 reader to mediate a synergistic effect of H3K4me3 and H3K27ac on
regulation of gene transcription. The detailed mechanism underlying the regulatory effect of SETD1A on activity of oncogenic enhancers and SEs needs further investigation.

In summary, we identified the H3K4 methyltransferase SETD1A as an oncogenic regulator driving HCC stemness through epigenetic modification and PABPC1 to promote oncogene transcription, providing an attractive therapeutic target for the treatment of HCC.

Methods

Cell lines and cell culture
HCC cell lines PLC, Huh7 and Hep3B were purchased from Cell Bank of the Chinese Academy of Sciences (Shanghai, China). HCC cells were cultured in DMEM high glucose (HyClone) supplemented with 10% fetal bovine serum (FBS, WISENT), 50 U/ml penicillin and 50 mg/ml streptomycin (HyClone). All the cell lines were authenticated by Short Tandem Repeat (STR) profiling and proved to be Mycoplasma-free by Myco-Blue Mycoplasma Detector (Vazyme, Nanjing, China).

Antibodies and reagents
Antibodies used were APC-conjugated CD133 antibody (cat. 130-113-184) from Miltenyi Biotec, PE-conjugated anti-CD24 (cat. 555428) from BD Biosciences, FITC-conjugated anti-EpCAM (cat. 60136FI) from Stemcell Technologies, anti-SETD1A (cat. A300-289A) from Bethyl Laboratories, anti-PABPC1 (cat. A14872), anti-WDR5 (cat. A3259), anti-CXXC1 (cat. A13423), and anti-ASH2L (cat. A4892) from
Abclonal, anti-H3K27ac (cat. ab4729) and anti-H3K27me3 (cat. ab6002) from Abcam, Anti-β-Actin (cat. AC004) from Abclonal Technology, anti-E-cadherin (cat. 3195S), anti-Vimentin (cat. 5741S) and anti-H3K4me3 (cat. C42D8) from Cell Signaling Technology, and Alexa Fluor Plus 488-conjugated Goat anti-Rabbit IgG (cat. A32731) and Alexa Fluor 594-conjugated Donkey anti-Rabbit IgG (cat. A-21207) secondary antibodies from ThermoFisher Scientific. The detailed information for antibodies was shown in Supplemental Table 3. DAPI (cat. D9542) was from Sigma. B27 (cat. A3582801) and N2 supplements (cat. 17502001) were from Gibco. bFGF (cat. 157AA) was from Novoprotein and EGF (cat. 236-EG) was from R&D Systems.

**Genome-wide CRISPR/Cas9 knockout library screen.**

Human CRISPR Knockout Pooled Library (GeCKO V2, #1000000049) was purchased from Addgene. The libraries were amplified using Endura cells (cat. 60242, Lucigen). The workflow of CRISPR/Cas9 pooled screen was shown in Fig. 1A. PLC cells were infected with lentiviral particles packaging Cas9 protein at an MOI less than 0.7. The blasticidin selected Cas9-expressing PLC cells were infected with pooled lentiviral CRISPR library at an MOI of 0.3 (1000× coverage) to ensure single-copy sgRNA integration in each cell. A pool of knockout cells was created after 7 days of selection with 2.5 μg/ml puromycin. CD24⁺CD133⁺ HCC CSCs and CD24⁻CD133⁻ non-HCC CSCs were sorted by FACS. Genomic DNA of cells were extracted using Quick-DNA Microprep Plus Kit (cat. D4074, Zymo Research) according to the manufacturer’s instructions. Sequencing libraries were constructed and sequenced by
Genewiz company. The sequencing data were analyzed by MAGeCKFlute(49).

**Establishment of SETD1A knockdown cell lines**

shRNA targeting SETD1A was cloned into pLKO.1-puro vector. The sequence was shown in Supplemental Table 1. To generate stable SETD1A knockdown cell lines, lentiviral particles were generated by co-transfection of pLKO.1-shSETD1A, the packaging plasmid psPAX2 and the envelope plasmid pMD2.G into 293T cells using the calcium phosphate transfection method and harvesting of the supernatant 48 h and 72h after transfection to infect HCC cell lines. Then, the cells were selected by puromycin for 24 h. The effect of SETD1A knockdown was evaluated by qRT-PCR and western blotting analysis.

**RNA extraction and quantitative real-time PCR (qRT-PCR)**

Total RNA was extracted using NucleoZOL (MN) according to the manufacturer’s instructions. cDNA was synthesized using HiScript III All-in-one RT SuperMix Perfect for qPCR kit (cat. R333-01, Vazyme). TB Green Premix Ex Taq (Tli RNase H Plus) (cat. RR420A, Takara) was used for qRT-PCR. To ensure the authenticity of the results, the experiments were performed in triplicate. Changes in mRNA expression was calculated based on comparison of the cycle threshold value after normalization to β-actin expression. The primers used in this study are listed in Supplementary Table 2.

**Western blotting**

Cell or tissue proteins were extracted by RIPA buffer (cat. P00013C, Beyotime) supplemented with cOmplete protease inhibitor cocktail (cat. 4693116001, Sigma)
according to the manufacturer’s instructions. The protein concentration was
determined by the Pierce BCA Protein Assay kit (cat. 23225, Thermo). 20 µg protein
was loaded on the SDS-PAGE gel and transferred to PVDF membrane, followed by
immunodetection of proteins. β-Actin was used as a loading control.

**Cell proliferation and transwell migration and invasion assay**

A live real-time IncuCyte ZOOM (Incucyte S3, ESSEN Bioscience) was used to
determine cell proliferation as previously described (50). In the transwell migration
and invasion assays, cells were suspended by serum-free medium and gently added in
a transwell chamber (cat. 3422, Corning) with or without Matrigel (cat. 356234,
Corning). After incubating at 37°C for 24 h, cells at the lower surface of the transwell
chamber were fixed, stained with 0.1% crystal violet (Solarbio) and counted at a
magnification of 200×.

**Sphere formation assay**

Cancer cells were seeded in the Ultra-Low Attachment 6-well plates (cat. 3471,
Corning) at a density of 1,500 cells/well and cultured in DMEM high glucose medium
supplemented with 1× B27, 1 × N2, 20ng/ml EGF and 20ng/ml bFGF for 2 weeks.
Spheres were counted at a magnification of 40×.

**Immunohistochemistry**

Tissues microarrays (TMA) with patients’ survival information were purchased from
Shanghai Outdo Biotech Company (cat. HLivH180Su16, Shanghai, China). The TMA
were treated with citrate buffer with 95°C. Endogenous peroxidase were inactivated
using 3% H₂O₂. The TMA were blocked using goat serum for 1h at room temperature
and then incubated with anti-SETD1A antibody overnight at 4°C. The TMA were then incubated with HRP-conjugated secondary antibody and detected using DAB. The nuclei were stained with hematoxulin.

**Mice model**

4-week-old female NOD/SCID mice were purchased from the Shanghai Model Organisms Center, Inc. Mice were housed in IVC systems in SPF grade animal room and fed standard laboratory diet with water and food. All the animal experiments were approved by the Animal Ethics Committee of the Fifth Affiliated Hospital of Sun Yat-Sen University (Approval No. 2020101401). And were performed in compliance with the National Institutes of Health Guidelines.

1×10^4 or 5×10^4 cancer cells mixed with Matrigel (cat. 356231, Corning) were injected subcutaneously into a NOD/SCID mouse. Each experimental group contained 8 mice. Tumor size was measured every three days. The animals were euthanized when tumor size reached 1.0-1.5 cm in diameter. The tumor initiating frequency was calculated using Extreme Limiting Dilution Analysis (http://bioinf.wehi.edu.au/software/elda/).

**CUT&Tag**

CD24^+^CD133^+^ liver CSC and CD24^−^CD133^−^ non-CSC cells were sorted by the BD Aria Fusion Cell Sorter (BD Biosciences). CUT&Tag was performed using the Hyperactive In-Situ ChIP Library Prep Kit (cat. TD901-01, Vazyme) according to the manufacturer’s instructions. Briefly, a total of 100,000 cells were collected and washed using Wash buffer supplemented with the protease inhibitors cocktail. Then cells bound to ConA magnetic beads were suspended in the antibody dilution buffer
and incubated with primary antibody at room temperature for 2 h or 4°C overnight. ConA-bound cells were then washed to remove unbound primary antibody, resuspended in Dig-wash buffer containing seconding antibody and incubated at room temperature for 2 h. Samples were washed using Dig-wash buffer for 3 times and resuspended in Dig-300 buffer containing Hyperactive pG-Tn5/pA Transposon and incubated at room temperature for 2 h. Samples were washed with Dig-300 buffer and resuspended in Tagmentation buffer for 1 h at 37°C. To stop tagmentation, 10µl 0.5M EDTA, 3µl 10% SDS and 2.5µl 20mg/ml Proteinase K were added to samples and incubated for 1 h at 55°C. The phenol-chloroform extraction method was used to extract DNA. To amplify libraries, PCR was performed using the following cycling conditions: 72°C, 3min; 98°C, 30s; 15 cycles of 98°C for 15 s, 60°C for 30 s, 72°C for 30 s and 72°C for 5 min. The PCR amplified sequencing library was further purified using 1.2 × AMPure XP beads (cat. A63881, Beckman). The library quality control was prepared using Agilent 2200. DNA sequencing was performed using Illumina NovaSeq 6000.

**ATAC-seq**

ATAC-seq was performed using TruePrep DNA Library Prep Kit V2 for Illumina (cat. TD501, Vazyme). Briefly, a total of 50,000 cells were collected and washed by PBS buffer. Cells were centrifuged at 500g for 5 min at room temperature, resuspended in cold Lysis buffer (10mM Tris-HCl, pH 7.4, 10mM NaCl, 3mM MgCl2, 0.1% IGEPAL CA-630) and incubated on ice for 10 min. Nuclear pellets were then collected by centrifugation for 5 min at 500 g at 4 °C and resuspended in transposition
reaction mix and incubated for 30 min at 37℃. DNA was extracted using 2× Agencourt AMPure XP beads. ATAC-seq libraries were amplified using Phanta HS Super-Fidelity DNA Polymerase using the following the program: 72℃ for 3 min; 98℃ for 30 s; 15 cycles of 98℃ for 15 s, 60℃ for 30 s, 72℃ for 30 s and 72℃ for 5 min. The PCR amplified sequencing library was purified and sequenced as previously described.

**CUT&Tag and ATAC-seq analysis**

For CUT&Tag and ATAC-seq, the reads were aligned to hg19 using the Bowtie2 (version 2.3.4.3) and run MACS2 to call peaks with a q-value of < 0.05 by using the MACS2 (version 2.1.2) with default parameters, a genome size of 2.7e9 bp and the appropriate input control sample. Through the annotatePeaks.pl in Homer (version 4.11), we further annotated the peaks with their related genes and distance to the closest transcription start sites (TSS).

For H3K27ac pairwise comparison, reads from CUT&Tag were counted using featureCounts (version 2.0.1). To screen the differentially expressed genes (DEGs) between the groups, the datasets were analyzed using the R package DESeq2 (version 1.30.1) with the default DESeq2 setting. The values for statistical significance were set as adjusted P value ≤ 0.05 and |Fold change| ≥ 1. Volcano maps were drawn using the R package ggplot2 (version 3.3.3). To functionally annotate DEGs, visualization and annotation of GO terms was utilized by Metascape (http://metascape.org/gp/index.html#/main/step1). GSEA was performed using the GSEA software (version 4.1.0) from the Broad Institute. The default weighted
enrichment method was applied for enrichment analysis. The random combination was set for 1000 times.

**Identifying Super-enhancers**

To identify super-enhancers, which were defined as regions of CUT&Tag enrichment for H3K27ac, we used the Rank Ordering of Super-Enhancers (ROSE) algorithm (https://bitbucket.org/young_computation/rose). The enhancer peaks of H3K27ac were stitched together if they were located within 12.5 kb distance of each other; peaks within 2.5 kb from a RefSeq transcription start site were excluded. To distinguish the super-enhancers from the typical enhancers, the point along with X-axis at which a line with a slope of 1 was tangent to the curve was found by scaling the data such that the x and y axis were from 0-1. Enhancers above this point were defined as super-enhancers, while enhancers below that point were typical enhancers. Enhancers were then assigned to the transcript whose TSS was nearest the center of the enhancer.

**Co-immunoprecipitation (CoIP)**

PLC cells transduced with FLAG-tagged PABPC1 vector or empty vector were lysed using NP40 buffer (150 mM NaCl, 1.5 mM MgCl2, 0.5% NP40, 50 mM Tris-HCl at pH8.0) supplemented with protease inhibitor cocktail (Sigma). Cell lysates were incubated with anti-FLAG antibody (cat. F1804, Sigma). Protein-antibody complex were conjugated to Protein A Magnetic Beads (MCE) by incubation at 4°C overnight. The beads were washed with IP washing buffer for 3 times. Proteins were dissolved in 1 × SDS loading buffer using boiling. The interacted proteins were identified using
mass spectrometry.

**Statistical analysis**

Statistical data were analyzed by SPSS version 13.0 (Chicago, IL, USA). Overall survival (OS) and disease-free survival (DFS) curves were plotted by Kaplan-Meier survival analysis and log-rank test. Two-way Student’s t-test or ANOVA test were used to determine the differences among groups. P values less than 0.05 were considered statistically significant.

**Data availability**

The data that support the findings of this study are available from the corresponding author upon reasonable request. All of the sequence data have been deposited at NCBI GEO (PRJNA991165).

**Acknowledgement**

This work was supported by National Natural Science Foundation of China (81870411 and 82172241), Natural Science Foundation of Guangdong Province, China 2019A1515010138, Science and Technology Planning Project of Guangdong Province China 2019A030317017, The Innovation Research Team for Basic and Clinical Studies on Chronic Liver Diseases of 2018 High-Level Health Teams of Zhuhai. We thank the Research Core Facility of Guangdong Provincial Engineering Research Center of Molecular Imaging, the Fifth Affiliated Hospital, Sun Yat-Sen University.

**Conflict of interest:** The authors declare no conflict of interest.
Author contributions

Conceptualization: F.X., and H.S.


Bioinformation analysis: Z.J.X.

Writing: F.X., and J.X.C.

References


40. Li Y, Schulz VP, Deng C, Li G, Shen Y, Tusi BK, et al. Setd1a and NURF


Figure and figure legends:
Fig. 1. SETD1A promotes HCC stemness in vitro and in vivo. A. Schematic outline of CRISPR/Cas9 knockout screen. B. A scatterplot of the gRNA distribution from the GeCKO screen. The TrxG proteins are labeled in red and the PcG proteins are labeled in blue. C. GO enrichment analysis of the hits promoting liver CSCs expansion. D. Enrichment of the gRNAs targeting SETD1A in the CD24⁺CD133⁻ non-CSCs. E. ATAC-seq analysis of the accessibility of SETD1A locus in the CD24⁺CD133⁺ CSCs.
and CD24+CD133- non-CSCs. **F.** qRT-PCR analysis of SETD1A expression in CD24+CD133+ CSCs and CD24+CD133- non-CSCs (n=3). Comparison of the proportion of CD24+CD133+ CSCs (**G**) as well as EpCAM+ CSCs (**H**) in scramble control and SETD1A knockdown HCC cells using flow cytometry (n=3). **I.** The spheroid-formation assays showing the role of SETD1A in HCC stemness *in vitro* (n=3). Scale bar represents 500 μm. **J.** The images of tumors formation in NOD/SCID mice injected subcutaneously with the scramble control and shSETD1A HCC cells. **K.** Extreme limiting dilution analysis for comparing the scramble control group and shSETD1A group. **L.** Overall survival curves of mice transplanted with the scramble control and shSETD1A HCC cells. Data are presented as mean ± SEM. Statistical analysis was performed by unpaired two-tailed Student’s t-test. **p < 0.01, and ***p < 0.001.**
Fig. 2. High-level expression of SETD1A is associated with poor outcome in HCC patients. Analysis of relative expression of SETD1A in HCC samples and their matched normal samples using the TCGA (A) and Oncomine (B) database. C. Relative expression of SETD1A at different HCC stages. (D-E) Analysis of overall survival curves of patients with high and low SETD1A expression levels using the
TCGA (D) and Kaplan-Meier Plotter (E) databases. F. Analysis of disease-free survival curves in HCC patients with high or low SETD1A expression using the TCGA database. G. IHC analysis of SETD1A expression in human HCC TMA. Scale bar represents 60 μm. H. Analysis of relative expression of SETD1A in HCC samples and their matched normal samples using IHC. I. Analysis of overall survival curves of HCC patients with high and low SETD1A expression levels. J. Analysis of recurrence curves of HCC patients with high and low SETD1A expression levels. K-L. ROC analysis for evaluating the association between SETD1A and HCC recurrence. **p < 0.01, and ***p < 0.001.

Fig. 3. SETD1A promotes HCC stemness and progression through regulating histone modification. A. Representative SETD1A, H3K4me3, and H3K27me3 CUT&Tag profiles in CD24⁺CD133⁺ CSCs at BMI1, ZIC2, SETD1A, and PDK4 locus. B. GO enrichment analysis of SETD1A-regulated genes in CD24⁺CD133⁺ CSCs. C. Representative SETD1A, H3K4me3, and H3K27me3 CUT&Tag profiles in CD24⁺CD133⁺ CSCs at EZ.
H2, SETD1B, KDM4A, KAT5, DOTIL and KMT2D locus. D. The changes of H3K4me3-marked genes (indicated as K4me3), H3K37me3-marked genes (indicated as K27me3) and bivalent genes (indicated as K4/K27me3) resulted by SETD1A knockdown. F. GO enrichment analysis of the genes losing H3K4me3 and gaining H3K27me3 resulted by SETD1A knockdown. G. PPI network analysis of the
regulators losing H3K4me3 and gaining H3K27me3 upon SETD1A knockdown. **H.**

GO enrichment analysis for the genes gaining H3K4me3 and losing H3K27me3 upon SETD1A knockdown. **I.** PPI network analysis of the regulators gaining H3K4me3 and losing H3K27me3 upon SETD1A knockdown. **J.** Representative H3K27me3 CUT&Tag profiles in the control and SETD1A knockdown CD24+CD133+ CSCs at KITLG and EPHA3 locus. The expression of KITLG and EPHA3 is shown on the right. **K.** Representative H3K4me3 CUT&Tag profiles in control and SETD1A knockdown CD24+CD133+ CSCs at NGFR and WNT7A locus. Their expression of NGFR and WNT7A is shown on the right (n=3). Data are presented as mean ± SEM. Statistical analysis was performed by unpaired two-tailed Student’s t-test. *p < 0.05, **p < 0.01, and ***p < 0.001.
Fig. 4. SETD1A promotes the activity of oncogenic enhancers and SEs. A. Volcano plots illustrating distribution of the enhancer-associated genes in the SETD1A knockdown versus control CD24⁺CD133⁺ CSCs. Genes losing H3K27ac are marked in blue. Genes gaining H3K27ac are marked in red. B. Leading-edge analysis of the enrichment of genes losing H3K27ac in the SETD1A knockdown versus control CD24⁺CD133⁺ CSCs based on GSEA. C. GO enrichment analysis of the
genes losing H3K27ac in SETD1A knockdown versus control in CD24+CD133+ CSCs. **D.** Distribution of H3K27ac signal across enhancer in the control CD24+CD133+ CSCs and SETD1A knockdown CD24+CD133+ CSCs. Prominent genes associated with SEs are highlighted with their respective SE ranks and roles in tumor initiation and progression. **E.** Representative H3K27ac CUT&Tag profiles in the control and SETD1A knockdown CD24+CD133+ CSCs at FNDC3B, PTP4A1, ELF3, AKAP12, ST3GAL4, and PTPN1 locus (Top). The expression of FNDC3B, PTP4A1, ELF3, AKAP12, ST3GAL4, and PTPN1 is shown on the bottom (n=3). Data are presented as mean ± SEM. Statistical analysis was performed by unpaired two-tailed Student’s t-test. **p < 0.01, and ***p < 0.001.
Fig. 5. SETD1A cooperates with PABPC1 to promote HCC stemness and progression. A. Mass spectrometry identification of PABPC1 as an interaction protein of SETD1A. B. Western blots of endogenous co-IPs for PABPC1, SETD1A, CXXC1, ASH2L and WDR5 in PLC and Huh7 cells. C. Analysis of relative expression of PABPC1 in HCC samples and their matched normal samples using the TCGA database. D. Analysis of overall survival curves of patients with high and low...
PABPC1 expression levels using the TCGA database. E. ATAC-seq analysis of the accessibility of PABPC1 locus in the CD24⁺CD133⁺ CSCs and CD24⁻CD133⁻ non-CSCs. And the representative SETD1A CUT&Tag profiles in CD24⁺CD133⁺ CSCs at PABPC1 locus. F. Comparation of the proportion of CD24⁺CD133⁺ CSCs in SETD1A knockdown and PABPC1 expressing SETD1A knockdown HCC cells using flow cytometry (n=3). G. The spheroid-formation assays showing the effect of PABPC1 overexpression on the stemness of SETD1A knockdown HCC cells (n=3). Scale bar represents 500 μm. H. Cell proliferation assay for the effect of PABPC1 overexpression on the cell proliferation of SETD1A knockdown HCC cells (n=3). I-J. Transwell assay with/without Matrigel assessing the effect of SETD1A knockdown PABPC1 overexpression on the migration and invasion of SETD1A knockdown HCC cells (n=3). Scale bar represents 200 μm. K. Venn diagram showing the extent of overlap for PABPC1, SETD1A, H3K4me3-bound regions in liver CSCs. L. GO analysis for biological processes of PABPC1 target genes. M. GO analysis for biological processes of PABPC1 and SETD1A co-regulation genes. N. GO analysis for biological processes of PABPC1, H3K4me3, and SETD1A co-regulation genes. Data are presented as mean ± SEM. Statistical analysis was performed by unpaired two-tailed Student’s t-test. *p < 0.05, **p < 0.01, and ***p < 0.001.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>SETD1A expression</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low expression</td>
<td>High expression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age(years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤50</td>
<td>40(44.4%)</td>
<td>22(55.0%)</td>
<td>18(45.0%)</td>
</tr>
<tr>
<td>&gt;50</td>
<td>50(55.6%)</td>
<td>27(54.0%)</td>
<td>23(46.0%)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>10(11.1%)</td>
<td>6(60.0%)</td>
<td>4(40.0%)</td>
</tr>
<tr>
<td>female</td>
<td>80(88.9%)</td>
<td>43(53.8%)</td>
<td>37(46.3%)</td>
</tr>
<tr>
<td>Pathological Grades</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>3(3.3%)</td>
<td>1(33.3%)</td>
<td>2(66.7%)</td>
</tr>
<tr>
<td>II</td>
<td>54(60%)</td>
<td>33(61.1%)</td>
<td>21(38.9%)</td>
</tr>
<tr>
<td>III</td>
<td>33(36.7%)</td>
<td>15(45.5%)</td>
<td>18(54.5%)</td>
</tr>
<tr>
<td>AJCC Grades</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>63(70%)</td>
<td>38(60.3%)</td>
<td>25(39.7%)</td>
</tr>
<tr>
<td>II</td>
<td>25(27.8%)</td>
<td>10(40.0%)</td>
<td>15(60.0%)</td>
</tr>
<tr>
<td>III</td>
<td>2(2.2%)</td>
<td>1(50.0%)</td>
<td>1(50.0%)</td>
</tr>
<tr>
<td>Tumor size, cm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤5 cm</td>
<td>55(61.1%)</td>
<td>37(67.3%)</td>
<td>18(32.7%)</td>
</tr>
<tr>
<td>&gt;5 cm</td>
<td>35(38.9%)</td>
<td>12(34.3%)</td>
<td>23(65.7%)</td>
</tr>
<tr>
<td>Cirrhosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>9(10%)</td>
<td>6(66.7%)</td>
<td>3(33.3%)</td>
</tr>
<tr>
<td>yes</td>
<td>81(90%)</td>
<td>43(53.1%)</td>
<td>38(46.9%)</td>
</tr>
<tr>
<td>Cirrhosis nodule</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>solitary</td>
<td>9(10%)</td>
<td>5(55.6%)</td>
<td>4(44.4%)</td>
</tr>
<tr>
<td>multiple</td>
<td>81(90%)</td>
<td>44(54.3%)</td>
<td>37(45.7%)</td>
</tr>
<tr>
<td>Tumor number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>solitary</td>
<td>79(87.8%)</td>
<td>45(57.0%)</td>
<td>34(43.0%)</td>
</tr>
<tr>
<td>multiple</td>
<td>11(12.2%)</td>
<td>4(36.4%)</td>
<td>7(63.6%)</td>
</tr>
<tr>
<td>Tumor recurrence</td>
<td>complete ALT normal</td>
<td>complete ALT increase</td>
<td>complete TB normal</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>no</td>
<td>41(45.6%)</td>
<td>28(68.3%)</td>
<td>13(31.7%)</td>
</tr>
<tr>
<td>yes</td>
<td>49(54.4%)</td>
<td>21(42.9%)</td>
<td>28(57.1%)</td>
</tr>
<tr>
<td>ALT normal</td>
<td>53(58.9%)</td>
<td>25(47.2%)</td>
<td>28(52.8%)</td>
</tr>
<tr>
<td>ALT increase</td>
<td>37(41.1%)</td>
<td>24(64.9%)</td>
<td>13(35.1%)</td>
</tr>
<tr>
<td>TB normal</td>
<td>76(25.9%)</td>
<td>43(56.6%)</td>
<td>33(43.4%)</td>
</tr>
<tr>
<td>TB increase</td>
<td>14(74.1%)</td>
<td>6(42.9%)</td>
<td>8(57.1%)</td>
</tr>
<tr>
<td>AFP &lt;20</td>
<td>36(40%)</td>
<td>17(47.2%)</td>
<td>19(52.8%)</td>
</tr>
<tr>
<td>AFP ≥20</td>
<td>54(60%)</td>
<td>32(59.3%)</td>
<td>22(40.7%)</td>
</tr>
<tr>
<td>HBsAg negative</td>
<td>19(21.1%)</td>
<td>9(47.4%)</td>
<td>10(52.6%)</td>
</tr>
<tr>
<td>HBsAg positive</td>
<td>71(78.9%)</td>
<td>40(56.3%)</td>
<td>31(43.7%)</td>
</tr>
<tr>
<td>HBcAb negative</td>
<td>7(7.8%)</td>
<td>4(57.1%)</td>
<td>3(42.9%)</td>
</tr>
<tr>
<td>HBcAb positive</td>
<td>83(92.2%)</td>
<td>45(54.2%)</td>
<td>38(45.8%)</td>
</tr>
<tr>
<td>GTT normal</td>
<td>44(48.9%)</td>
<td>27(61.4%)</td>
<td>17(38.6%)</td>
</tr>
<tr>
<td>GTT decrease</td>
<td>46(51.1%)</td>
<td>22(47.8%)</td>
<td>24(52.2%)</td>
</tr>
</tbody>
</table>

ALT, Alanine aminotransferase; HBsAg, hepatitis B surface antigen; HBcAb, hepatitis B core antigen AFP, alpha-fetoprotein; TB, total bilirubin; GTT, γ-glutamyl transpeptidase.
Table 2. Univariate and multivariate analyses of clinicopathologic variables and SETD1A in HCC.

<table>
<thead>
<tr>
<th>Variables</th>
<th>SETD1A Univariate P value</th>
<th>SETD1A Multivariate 95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (≤50/&gt;50y)</td>
<td>0.925</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>0.709</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathological Grades</td>
<td>0.360</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AJCC Grades</td>
<td>0.123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor size (≤5/&gt;5 cm)</td>
<td><strong>0.003</strong> *</td>
<td>1.624-11.077</td>
<td><strong>0.003</strong> *</td>
</tr>
<tr>
<td>Cirrhosis (no/yes)</td>
<td>0.442</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cirrhosis nodule (solitary/multiple)</td>
<td>0.944</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor number (solitary/multiple)</td>
<td>0.208</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Encapsulation (complete/none)</td>
<td><strong>0.003</strong> *</td>
<td>1.592-10.584</td>
<td><strong>0.003</strong> *</td>
</tr>
<tr>
<td>Tumor recurrence (no/yes)</td>
<td><strong>0.017</strong> *</td>
<td></td>
<td>0.086</td>
</tr>
<tr>
<td>ALT (normal/increase)</td>
<td>0.099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TB (normal/increase)</td>
<td>0.347</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AFP (&lt;20/&gt;20)</td>
<td>0.263</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBsAg (negative/positive)</td>
<td>0.487</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBcAb (negative/positive)</td>
<td>0.881</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTT (normal/decrease)</td>
<td>0.199</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI, confidence interval; ALT, Alanine aminotransferase; HBsAg, hepatitis B surface antigen; HBcAb, hepatitis B core antigen; AFP, alpha-fetoprotein; TB, total bilirubin; GTT, γ-glutamyl transpeptidase.