Cellular Oncology (2026) 49:18
https://doi.org/10.1007/s13402-025-01157-3

RESEARCH

The epigenetic regulator SETDB1 as a key component of cancer stem cells and drug resistance in primary liver cancer

Maël Padelli[1,2] · Christophe Desterke[3] · Aurore Devocelle[1] · Georges Uzan[1] · Antoinette Lemoine[2,4] · Julien Giron-Michel[1]

Received: 22 September 2025 / Accepted: 25 December 2025
© The Author(s) 2025

Abstract

Background Hepatocellular carcinoma (HCC) is a highly heterogeneous malignancy with a poor prognosis and limited response to current therapies. Cancer stem cells (CSCs) contribute to this heterogeneity, driving tumor progression, immune evasion, and therapeutic resistance. Epigenetic regulators have emerged as pivotal modulators of CSC phenotypes. The histone methyltransferase SETDB1, a key stemness factor in both embryonic and adult stem cells, catalyzes H3K9 trimethylation (H3K9me3) and acts as a major oncogenic driver in a wide range of solid and hematological malignancies. Yet, its specific role in maintaining stemness and contributing to therapeutic resistance in HCC remains poorly defined.

Methods Integrative analyses of bulk and single-cell transcriptomic datasets from public HCC cohorts were performed. Tumors were stratified by SETDB1 expression, followed by pathway enrichment, immune deconvolution, and stemness scoring. Associations with clinical outcomes, immune phenotypes, drug resistance signatures, and chromatin-binding partners (DNMT3A, TRIM28, HDAC1) were also explored. The role of SETDB1 was then examined in vitro, where a stem-like phenotype was induced in HCC cell lines under hypoxic conditions, and patterns of SETDB1 expression and stemness marker levels were assessed following paclitaxel treatment.

Results SETDB1-low tumors were enriched in well-differentiated, immunologically active subclasses and were associated with increased predicted responsiveness to immune checkpoint blockade (ICB), sorafenib, and transarterial chemoembolization. Conversely, SETDB1-high tumors exhibited transcriptional features of dedifferentiation, elevated stemness scores, and enrichment in aggressive molecular subtypes. These tumors also displayed higher TP53 mutation rates, reduced immune infiltration, immune exclusion signatures, and activation of pathways linked to ferroptosis and ABC transporter dysfunction, consistent with resistance to ICB and conventional therapies. SETDB1 expression was also associated with hypoxia-related pathways, suggesting a role in maintaining CSC niches. In vitro, hypoxia-induced stemness coincided with increased SETDB1 levels, whereas paclitaxel treatment decreased both SETDB1 expression and stemness markers. A chromatin risk score incorporating SETDB1 and its partners predicted poor disease-free survival independently of clinical parameters.

Conclusions SETDB1 defines a stemness-enriched, immune-resistant HCC subtype associated with poor outcomes and therapeutic failure. Correlations between SETDB1 expression and stem-like features in vitro suggest a potential role in maintaining CSC phenotypes, supporting its relevance as a biomarker and candidate for epigenetic- and CSC-directed therapeutic strategies.

Keywords SETDB1 · Hepatocellular carcinoma · Cancer stem cells · Drug resistance · Epigenetics · Paclitaxel · Immunotherapy · Liver cancer

✉ Maël Padelli
mael.padelli@inserm.fr

[1] INSERM UMR-S-MD 1197, Ministère des Armées et Université Paris Saclay, Villejuif, France

[2] Department of Biochemistry and Oncogenetics, Paul Brousse Hospital, AP-HP, Villejuif, France

[3] INSERM UMR1310 et Université Paris-Saclay, Villejuif, France

[4] INSERM UMR1193 et Université Paris-Saclay, Villejuif, France

Published online: 06 January 2026

Abbreviations

AFP	Alpha-fetoprotein
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
CPTAC	Clinical Proteomic Tumor Analysis Consortium
CSCs	Cancer Stem Cells
DAPI	4',6-diamidino-2-phenylindole
DFS	Disease-Free Survival
DMEM	Dulbecco's Modified Eagle's Medium
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
GDC	Genomic Data Commons
GEO	Gene Expression Omnibus
GPC3	Glypican-3
GSEA	Gene Set Enrichment Analysis
HCC	Hepatocellular Carcinoma
iCCA	Intrahepatic Cholangiocarcinoma
IC50	Half-maximal Inhibitory Concentration
ITH	Intratumoral Heterogeneity
KEGG	Kyoto Encyclopedia of Genes and Genomes
KMTs	Histone Lysine Methyltransferases
LIHC	Liver Hepatocellular Carcinoma
MCC	Matthews Correlation Coefficient
MDR	Multi-Drug Resistance
MSigDB	Molecular Signatures Database
mRNAsi	mRNA-based Stemness Index
NES	Normalized Enrichment Score
PCA	Principal Component Analysis
PBS	Phosphate-Buffered Saline
PPI	Protein-Protein Interactions
PVDF	Polyvinylidene Fluoride
qRT-PCR	Real-time Quantitative Reverse Transcription Polymerase Chain Reaction
RIPA	Radioimmunoprecipitation Assay
ROC	Receiver Operating Characteristic
RSEM	RNA-Seq by Expectation Maximization
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SH	SETDB1 High
SL	SETDB1 Low
ssGSEA	Single Sample Gene Set Enrichment Analysis
TACE	Transarterial Chemoembolization
TCGA	The Cancer Genome Atlas
TIMER	Tumor Immune Estimation Resource
TKI	Tyrosine Kinase Inhibitor
TME	Tumor Microenvironment
TPM	Transcripts Per Million
t-SNE	t-distributed Stochastic Neighbor Embedding
UALCAN	University of Alabama at Birmingham CANcer data analysis Portal

1 Introduction

Primary liver cancer, predominantly hepatocellular carcinoma (HCC), is the third leading cause of cancer-related mortality [1]. Despite recent advancements in traditional treatments such as surgical resection, transarterial chemoembolization (TACE), chemotherapy, and immunotherapy, the 5-year survival rate for HCC remains dismal, often under 20%, largely due to drug resistance and disease recurrence [2].

A defining feature of HCC is its remarkable heterogeneity, both across patients (interpatient) and within individual tumors (intratumoral). Interpatient heterogeneity plays a central role in the stratification of patients for personalized medicine. However, intratumoral heterogeneity (ITH), which refers to the coexistence of genetically, epigenetically, and phenotypically distinct subclones within the same tumor, poses an even greater challenge to treatment success. ITH has been recognized as a major contributor to tumor progression, immune evasion, therapy resistance, and ultimately, poor clinical outcomes [3]. The most aggressive cancer cells display characteristics similar to progenitor cells, leading to their designation as cancer stem cells (CSCs). They exhibit significant self-renewal capacity, the ability to form metastases, and the potential to differentiate into various cancer cell lineages [4, 5], all of which grants them high tumorigenicity [6]. In HCC, CSCs have been identified using surface markers such as CD133, CD44, EPCAM, and CD24 [7]. CSCs are thought to be responsible for driving therapeutic resistance, especially to conventional treatments such as chemotherapy, targeted therapy, and immune checkpoint blockade. Moreover, CSCs often reside in protective niches shaped by the tumor microenvironment (TME), where they are shielded from immune attack and supported in their plasticity by hypoxia, stromal signals, and epigenetic modulation [8]. Given their role in cancer's persistence, targeting CSCs represents a highly promising, novel therapeutic approach for developing more effective and lasting treatments that can lead to long-term remission. However, no therapeutic strategy specifically employing this approach is currently used in the clinic, largely due to a limited understanding of the precise molecular targets that would effectively reduce their stem-like properties.

Alterations in different histone lysine methylations have been associated with tumorigenesis [9] and drug resistance [10]. Furthermore, blocking the enzymes responsible for these modifications, known as histone lysine methyltransferases (KMTs), can even reverse drug resistance [4]. However, the link between KMTs and CSCs remain largely unknown. The histone KMT SETDB1, a key stemness factor both in embryonic and adult stem cells [11, 12], has emerged as a predominant oncogene in a wide range of solid

cancers and hematological malignancies [13–15] promoting chemoresistance [10]. Recent analyses in HCC reveal that SETDB1 is overexpressed and strongly associated with advanced disease features, such as tumor microsatellite formation and metastasis. Conversely, its expression inversely correlates with immune cell infiltration [16, 17].

In this study, the expression patterns, regulatory mechanisms, and functional implications of SETDB1 in HCC were investigated using a combination of in vitro assays, in silico analyses, and public transcriptomic datasets (TCGA, GEO). The results showed that SETDB1 overexpression defines a high-risk HCC subtype characterized by stemness enrichment, aggressive behavior, and therapy resistance. These data position SETDB1 as a potential key component of CSCs, a critical biomarker for prognosis, and a promising candidate for epigenetic therapy targeting CSC-driven and immune-resistant liver tumors.

2 Material & methods

2.1 In vitro experiment

2.1.1 Human cell line culture and hypoxic treatment

Huh7 and HepG2 cell lines from ATCC were cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco, France) supplemented with 10% fetal bovine serum (FBS; Gibco, France) in a humidified incubator with 5% CO_2 and ambient oxygen (~21%) at 37 °C. For hypoxic treatment, cells were placed in a hypoxia workstation (Baker Ruskinn SCI-tive, Sanford, Maine, USA) and subjected to a gas mixture of 1% O_2, 5% CO_2, with the balance consisting of N_2, for durations ranging from 3 to 48 h.

2.1.2 RNA extraction and real-time quantitative RT-PCR

Total RNA was extracted using RNeasy® Mini Kit (Cat. No. 74104, Qiagen, Germany) according to the manufacturer's instructions. Complementary DNA (cDNA) was synthesized using SuperScript® III System (Cat. 18080044, ThermoFisher, Waltham, MA, USA). Real-time quantitative PCR was then performed using TaqMan Universal PCR Master Mix (Cat. 4318157, ThermoFisher) with TaqMan Gene Expression Assays. Each reaction consisted of 500 ng of starting cDNA and was run for 40 cycles. 18 S was used as the endogenous control for normalization. Relative gene expression was quantified by the $2^{-\Delta\Delta CT}$ method. The primers and probes used in this study are listed in Table S1. To ensure the reliability of the results, all experiments were performed in at least triplicate.

2.1.3 Flow cytometry

Suspensions of enzymatically detached cells were prepared in staining buffer (0.5% BSA in PBS). 10^5 cells were stained with conjugated antibodies directed against the cell markers CD24, EPCAM, CD44, and CD133. After rinsing in phosphate-buffered saline (PBS), 10,000 events were analyzed on an LSR Fortessa flow cytometer (BD Biosciences, Oxford, UK). Data were subsequently analyzed using FACS Diva software (BD Biosciences). Three technical replicates were used for each condition, and each experiment was performed at least three independent times.

2.1.4 Immunofluorescence staining

The intracellular localization of SETDB1 was examined in Huh7 cell line. For immunofluorescence staining, fixed Huh7 cells were permeabilized with 0.3% Triton-X-100 and then blocked with 0.1% Triton-X-100 + 1% BSA. Cells were incubated with anti-SETDB1 (MA5-15722, ThermoFisher, Plainville, MA, USA) overnight at 4 °C. After incubation with a fluorophore-conjugated secondary antibody (A-11001, ThermoFisher, Plainville, MA, USA), the glass coverslips were counterstained with DAPI and observed under a fluorescence microscope.

2.1.5 Western blotting

Protein extraction was performed using RIPA buffer (catalog P00013C, Beyotime, Shanghai, China) supplemented with cOmplete protease inhibitor cocktail (cat.4693116001, MilliporeSigma, Burlington, MA, USA) according to the manufacturer's instructions. Protein concentration was determined using the Pierce BCA Protein Assay kit (catalog 23225, Thermo Fisher Scientific, Waltham, MA, USA). A total of 20 µg of protein per sample was loaded onto an SDS-PAGE gel and subsequently transferred to a PVDF membrane. Immunodetection was performed with the following primary antibodies: SETDB1 (ab-107225, Abcam, 1:750), HIF1A (cat.610359, BD Biosciences, 1:1000), OCT4 (AF1759, Biotechne, 1/1000), NANOG (sc-134218, SantaCruz Biotechnology, 1/200), CD133 (130-092-395, Miltenyi Biotec, 1/200), and β-actin (sc-47778, SantaCruz Biotechnology, 1:250), which was used as a loading control.

2.1.6 Cell viability assay

Response to sorafenib chemotherapy was assessed by the CellTiter-Glo luminescent cell viability assay (Promega Corporation, Madison, WI, USA). Cells were plated at 3000 per well in 96-well microtiter plates and incubated for 48 h at 37 °C under normoxic or hypoxic conditions.

Subsequently, increasing concentrations of sorafenib (0 to 10 µM) were added to the wells, and the cultures were incubated for an additional 48 h. Cell viability was determined using the CellTiter-Glo luminescent cell viability kit from Promega Corporation (Madison, WI, USA) according to the manufacturer's instructions. The half-maximal inhibitory concentration (IC50), was calculated by nonlinear regression analysis using GraphPad Prism software (San Diego, CA, USA). The experiment was performed with at least three independent replicates.

2.1.7 Statistical analysis

GraphPad Prism (v.5.0) was used for statistical analysis, and $p<0.05$ was considered statistically significant. Wilcoxon rank sum test were employed in two-group comparisons to evaluate differentially expressed genes or microRNA (miRNAs).

2.2 Public transcriptome data collection and preprocessing

2.2.1 Bulk RNA-seq transcriptome analyses

Bulk RNA-sequencing data for liver hepatocellular carcinoma (LIHC) were retrieved from The Cancer Genome Atlas (TCGA) program [18]. Both normalized mRNA expression profiles (RNA-Seq by Expectation Maximization, RSEM), somatic mutation data, microRNA expression data, and comprehensive clinical annotations were obtained via the Genomic Data Commons (GDC) Firehose pipeline through cBioPortal (https://www.cbioportal.org/, accessed February 13, 2023) [19] or the LinkedOmics platform (http://www.linkedomics.org/) [20]. For analyses requiring raw count data, RNA-seq count matrices were normalized with edgeR Bioconductor R-package version 3.38.4 to account for library size and composition bias. Clinical data were integrated with expression matrices using the *dplyr* R package (v1.0.10). Only primary liver tumor samples with complete transcriptomic and clinical information were retained, yielding a final cohort of 364 patients for downstream analyses. Patients were stratified into two groups (SETDB1 high (SH) and SETDB1 low (SL)) based on the median SETDB1 expression ($n = 182$ per group) (Table S2). Differential expression analyses compared transcriptomic profiles between SH and SL tumors. Functional enrichment analyses were performed to identify pathways and gene sets associated with SETDB1 expression. Molecular and cellular phenotypes were computed using Gene Set Enrichment Analysis (GSEA) and single sample gene set enrichment analysis (ssGSEA) with predefined signatures from the Molecular Signatures Database (MSigDB), including KEGG (Kyoto Encyclopedia of Genes and Genomes), WikiPathways, Hallmark, and published HCC gene sets (Table S3). Enrichment scores for each tumor sample were generated with GenePattern (www.genepattern.org/, accessed on 30 April 2020). ssGSEA scores reflect coordinated up- or downregulation of gene sets within each sample [21], based on the Matthews correlation coefficient (MCC) for robust performance across all prediction categories [22]. In GSEA, significance was set at false discovery rate (FDR) < 0.25 and $p < 0.05$, and the most enriched pathways were selected based on normalized enrichment score (NES). The GSE104580 cohort was used as an independent validation dataset, in which patients were stratified into two groups based on SETDB1 expression levels (SH, $n = 74$; SL, $n = 73$).

2.2.2 Web tools used to analyse SETDB1 expression and correlation with stemness and immune microenvironment

SETDB1 mRNA expression in TCGA-LIHC tumors versus adjacent non-tumorous tissues was assessed using the University of Alabama at Birmingham CANcer data analysis Portal (UALCAN, http://ualcan.path.uab.edu/, accessed November 9, 2023). SETDB1 protein expression was further assessed in the Clinical Proteomic Tumor Analysis Consortium (CPTAC)-LIHC cohort via UALCAN. Group comparisons were performed using Wilcoxon rank-sum tests. The correlation between SETDB1 expression and the most frequent driver mutations reported in LIHC [23] was examined using the Tumor Immune Estimation Resource (TIMER 2.0) databases (http://timer.cistrome.org/, accessed November 9, 2023), with correlations evaluated by Spearman's test. SETDB1 expression in TACE and sorafenib responders versus non-responders was examined using HCC transcriptomic datasets GSE104580 and GSE109211 from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/, accessed on 7 November 2023) using the GEO2R tool. Stemness features were quantified using CancerStemnessOnline (http://bio-bigdata.hrbmu.edu.cn/CancerStemnessOnline, accessed on 7 November 2023), focusing on the mRNAsi (mRNA-based stemness index) and Stemness Index, where higher values indicate a more progenitor-like phenotype. Immune infiltration was estimated with xCell (https://xcell.ucsf.edu/, accessed on 7 November 2023), and differences in immune cell distributions between SETDB1 groups were compared. Correlations between SETDB1 and gene expression or stemness scores were assessed by Spearman's test, with $p < 0.05$ considered significant.

2.2.3 Single cell transcriptome analyses

Single-cell transcriptome analysis was performed using the HCC dataset GSE125449, which was downloaded from the NCBI website [24]. Bioinformatic analyses were conducted using the R software environment (version 4.0.2) and the Seurat R-package (version 3.2.2). In the preprocessing step, genes with positive expression in fewer than three cells were removed. Following canonical correlation-based batch correction, the merged dataset was normalized and scaled. Dimensionality reduction was performed using principal component analysis (PCA), followed by t-SNE for data visualization. The single-cell stemness score was calculated in Seurat by integrating the expression levels of the markers GDF15, ALDH1L1, GPC3, AFP, EPCAM, CD44, and CD24. Data visualization was generated using the FeaturePlot, DotPlot, and VlnPlot functions from the Seurat R package.

2.2.4 Prognosis risk score determination

Protein-protein interactions (PPI) for SETDB1 were determined using the STRING network server (version 11.5; confidence threshold > 0.7; Homo sapiens) on February 11th, 2023. A functional enrichment analysis on the resulting PPI network was performed using the Gene Ontology Molecular Function database.

For the TCGA liver cancer cohort, a SETDB1-related expression risk score was computed from partners with a univariate Cox p-value < 0.05 and an unfavorable expression profile. This score was calculated by summing the products of the Cox beta coefficients and the RNA-seq expression levels of these markers. The risk score threshold was determined using standardized log-rank statistics with the survminer R-package (version 0.4.9) and a loop of univariate Cox survival analysis was performed with the loopcolcox R-package (version 1.0.0).

A multivariable Cox model for disease-free survival (DFS) was built with the survival R-package (version 3.5-0) to include the SETDB1-related risk score. The model also integrated demographic parameters (age, gender, weight), liver cancer risk factors (hepatitis B, hepatitis C, NAFLD, alcohol consumption), and pathological classification (grades 1–4). The model was calibrated by bootstrapping with the rms R-package (version 6.4-1), and a corresponding nomogram was drawn with the regplot R-package (version 1.1).

Finally, expression analysis on TCGA RNA-sequencing data was performed via PCA with the prcomp R-base function. Expression heatmaps were generated with the pheatmap R-package (version 1.0.12), and unsupervised clustering was performed using Euclidean distances and the Ward.D2 method.

3 Results

3.1 SETDB1 expression is upregulated in HCC and correlates with genomic alterations, aggressive tumor features and poor prognosis

To explore the expression profile of SETDB1 in HCC, we analyzed transcriptomic and proteomic data using the UALCAN online platform. Differential expression analysis revealed a significant upregulation of SETDB1 mRNA in tumor tissues compared to adjacent normal liver tissues (Fig. 1A). In addition, proteomic data from the CPTAC module showed that SETDB1 protein levels were significantly elevated in HCC compared to non-tumoral liver samples ($p = 1.56E^{-41}$) (Fig. 1B).

Given the genomic complexity of HCC, which typically harbors 35 to 80 somatic mutations in coding regions per tumor, we assessed whether SETDB1 expression was associated with key driver mutations. Analysis of TCGA-LIHC data using the Timer2.0 web tool revealed a significant correlation between high SETDB1 expression and frequent mutations in TP53, RB1, and RPS6KA3, with Wilcoxon p-values of $4.6E^{-05}$, 0.0027, and 0.0024, respectively (Fig. 1C, D).

To further explore these associations, we stratified patients into SH and SL groups based on median SETDB1 expression. Mutational profiling using the Timer2.0 web tool showed that the SH group exhibited a higher proportion of patients with mutations in these key oncogenes (Fig. 1E). Notably, TP53 mutations were present in approximately 31% of cases and were especially enriched in the SH group. Meanwhile, mutations in RB1 and RPS6KA3 were each found in only 4% of the TCGA-LIHC cohort. These findings suggest that SETDB1 overexpression may be driven, at least in part, by genomic alterations, particularly TP53 mutations, and may play a role in the oncogenic reprogramming of HCC.

To investigate the biological consequences of SETDB1 overexpression, we conducted ssGSEA using curated gene signatures associated with poor prognosis in liver cancer. Our results in TCGA-LIHC showed that the SH group was significantly enriched in gene signatures linked to vascular invasion, early recurrence, and poor survival (Fig. 2A). In contrast, the SL group was associated with favorable survival-related gene sets.

Consistently, receiver operating characteristic (ROC) analysis revealed that gene signatures linked to poor prognosis were upregulated in the SH group, whereas

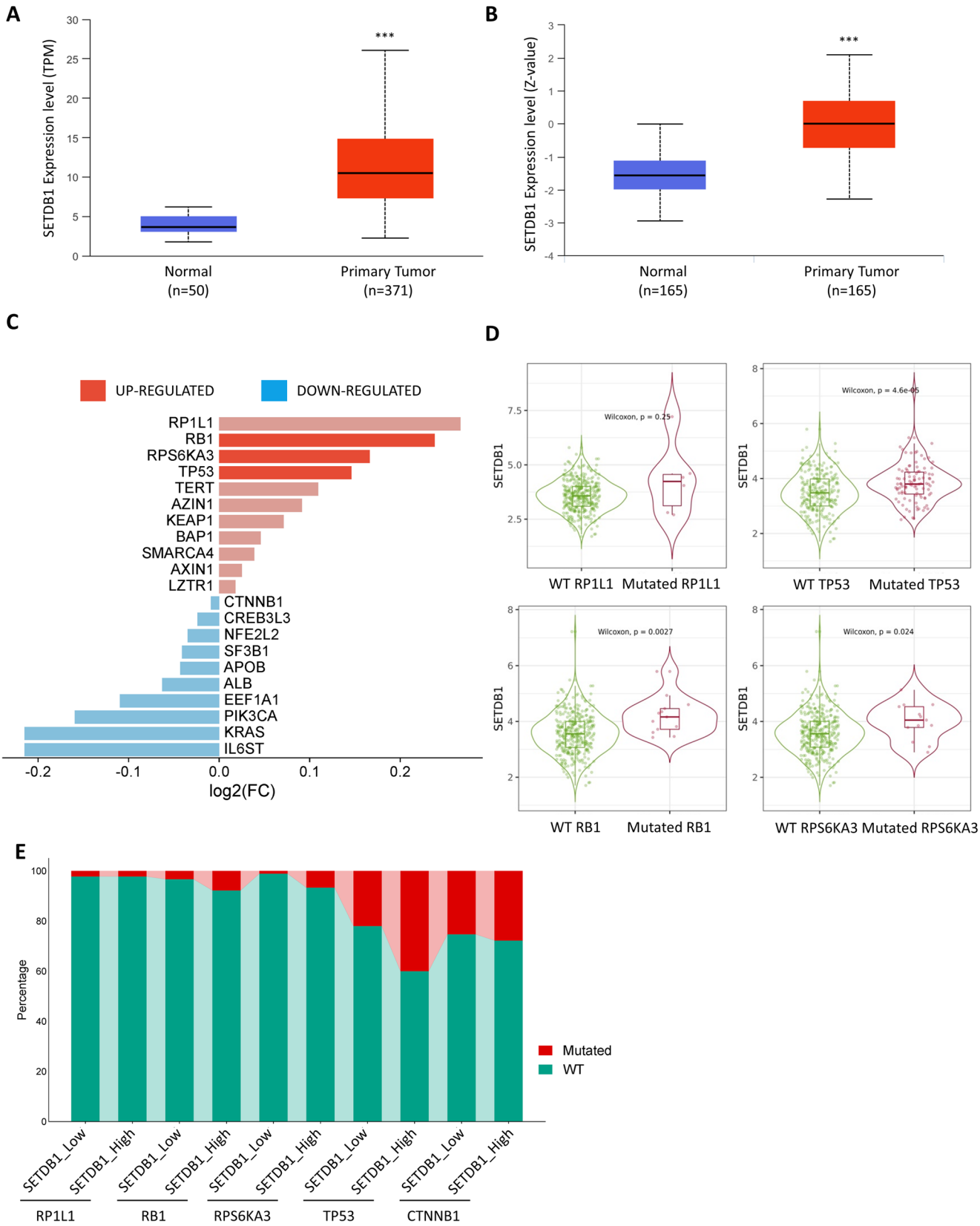

A
SETDB1 Expression level (TPM)
***
Normal (n=50)
Primary Tumor (n=371)
B
SETDB1 Expression level (Z-value)
***
Normal (n=165)
Primary Tumor (n=165)
C
UP-REGULATED
DOWN-REGULATED
RP1L1
RB1
RPS6KA3
TP53
TERT
AZIN1
KEAP1
BAP1
SMARCA4
AXIN1
LZTR1
CTNNB1
CREB3L3
NFE2L2
SF3B1
APOB
ALB
EEF1A1
PIK3CA
KRAS
IL6ST
log2(FC)
D
SETDB1
Wilcoxon, p = 0.25
WT RP1L1
Mutated RP1L1
Wilcoxon, p = 4.6e-05
WT TP53
Mutated TP53
Wilcoxon, p = 0.0027
WT RB1
Mutated RB1
Wilcoxon, p = 0.024
WT RPS6KA3
Mutated RPS6KA3
E
Percentage
Mutated
WT
SETDB1_Low
SETDB1_High
RP1L1
RB1
RPS6KA3
TP53
CTNNB1

Fig. 1 SETDB1 is overexpressed in primary tumors and correlates with specific gene mutations. **(A)** Analysis of relative SETDB1 gene expression (in Transcripts Per Million, TPM) in primary tumor samples ($n=371$) and adjacent normal tissues ($n=50$) from the TCGA database. Box plots illustrate the median, interquartile range (IQR), and minimum/maximum values (whiskers). SETDB1 expression is significantly higher in primary tumors. $P<0.001$. **(B)** SETDB1 protein expression levels (measured in Z-score) in normal ($n=165$) and primary tumor ($n=165$) tissues from CPTAC database. The protein levels are also significantly elevated in tumors. $P<0.001$. **(C)** Bar plot showing correlations between SETDB1 expression and mutation status of driver genes in LIHC. Up-regulation (red) or down-regulation (blue) is indicated by log2 fold change. Significant differences ($P<0.05$, Spearman's test) are shown with saturated colors; non-significant ($P>0.05$) with attenuated colors. **(D)** Violin plots illustrating SETDB1 expression levels in tumors stratified by the mutation status (wild-type vs. mutated) for four key genes: RP1L1, TP53, RB1, and RPS6KA3. Statistical comparisons were performed using a two-sided Wilcoxon rank-sum test. **(E)** A comparison of somatic mutation profiles between groups with high and low SETDB1 expression, defined by the median expression value. The stacked bar plot shows the percentage of mutated (red) and wild-type (green) samples for the top-ranked mutated genes in the TCGA-LIHC cohort

survival-associated signatures were downregulated (Fig. 2B). We also observed that the Coulouarn_temporal_TGFb1_signature, a hallmark of tumor progression and dedifferentiation, was specifically enriched in the SH group, reinforcing the association between SETDB1 expression and aggressive tumor phenotypes. Heatmaps (Fig. 2C) and scatter plots (Fig. 2D) illustrate the correlation between SETDB1 expression and Chen's eight-gene diagnostic model. SETDB1 expression positively correlates with the TPM values of five genes associated with poor prognosis: *TOP2A*, *CNNB1*, *MCM6*, *SSP1*, and *IGFBP3*, while it negatively correlates with *IGF1*, *ESR1*, and *F9*, which are considered indicative of a good prognosis [25]. These findings were confirmed in the independent validation cohort GSE104580 (Fig. S1A).

Overall, these data demonstrate that high SETDB1 expression is linked to both mutational burden and molecular programs associated with tumor aggressiveness and poor clinical outcomes in HCC. These results highlight the potential of SETDB1 as a prognostic biomarker and suggest its involvement in orchestrating transcriptional programs that drive HCC progression.

3.2 SETDB1 expression correlates with resistance to therapy

The predicted effects of SETDB1 expression on tumor biology do not fully account for the poorer survival outcomes observed in patients with high SETDB1 expression (SH group) within the TCGA-LIHC cohort. Consequently, we investigated whether this unfavorable prognosis could be linked to differences in therapeutic responses, specifically focusing on the impact of SETDB1 expression on sensitivity to immunotherapy, sorafenib, and TACE.

3.2.1 SETDB1 expression predicts response to immunotherapy

To characterize the immunological features of the SH and SL groups, we performed ssGSEA using established immune-related gene signatures. SL tumors were significantly enriched in signatures associated with immunogenic and inflamed tumor microenvironments, cytotoxic and antigen-presenting cell signatures, and the inflammatory signature (Fig. 3A). Conversely, the SH group showed enrichment in immune-excluded or non-inflamed tumor phenotypes, highlighting a less immunogenic state. These findings suggest that SETDB1-low tumors resemble "hot tumors" with active immune infiltration, whereas SETDB1-high tumors resemble "cold tumors" with an immunosuppressive microenvironment.

To further explore the immune contexture, we assessed immune infiltration scores using the xCell algorithm. As expected, the immune score was significantly higher in the SL group compared to the SH group (Fig. 3B), confirming a more active immune microenvironment. A detailed analysis of 29 immune cell subtypes revealed that the SL group displayed a greater abundance of antitumoral immune cells, notably CD8+T cells ($P<0.01$), CD4+effector memory T cells, NK cells, M1 macrophages, and dendritic cells, which are known to contribute to effective immune surveillance and tumor rejection. In contrast, the SH group exhibited enrichment of immunosuppressive cells such as Tregs, monocytes, and M2 macrophages, potentially contributing to immune evasion.

Prompted by this strong link between immune features and immunotherapy response, we investigated whether SETDB1 expression influences immunotherapeutic sensitivity. Using ssGSEA, we compared immunotherapy response-related gene signatures between the two groups. The SH group was enriched in gene sets associated with immunotherapy resistance, while the SL group displayed a strong enrichment for signatures predicting sensitivity to immune checkpoint blockade (Fig. 3C). The expression levels of 46 potentially targetable immune checkpoint genes, including the major ones like *PDCD1* (PD-1), *CD274* (PD-L1), and *CTLA4*, were markedly higher in the SL group, further supporting the hypothesis that SETDB1-low tumors are more likely to benefit from immune checkpoint blockade therapy.

To further explore genes linked to immune responsiveness, we examined the CUI_2023_IMI_Immunotherapy_sensitive signature. A heatmap (Fig. 3D) showed a negative correlation between SETDB1 expression and six key IMI

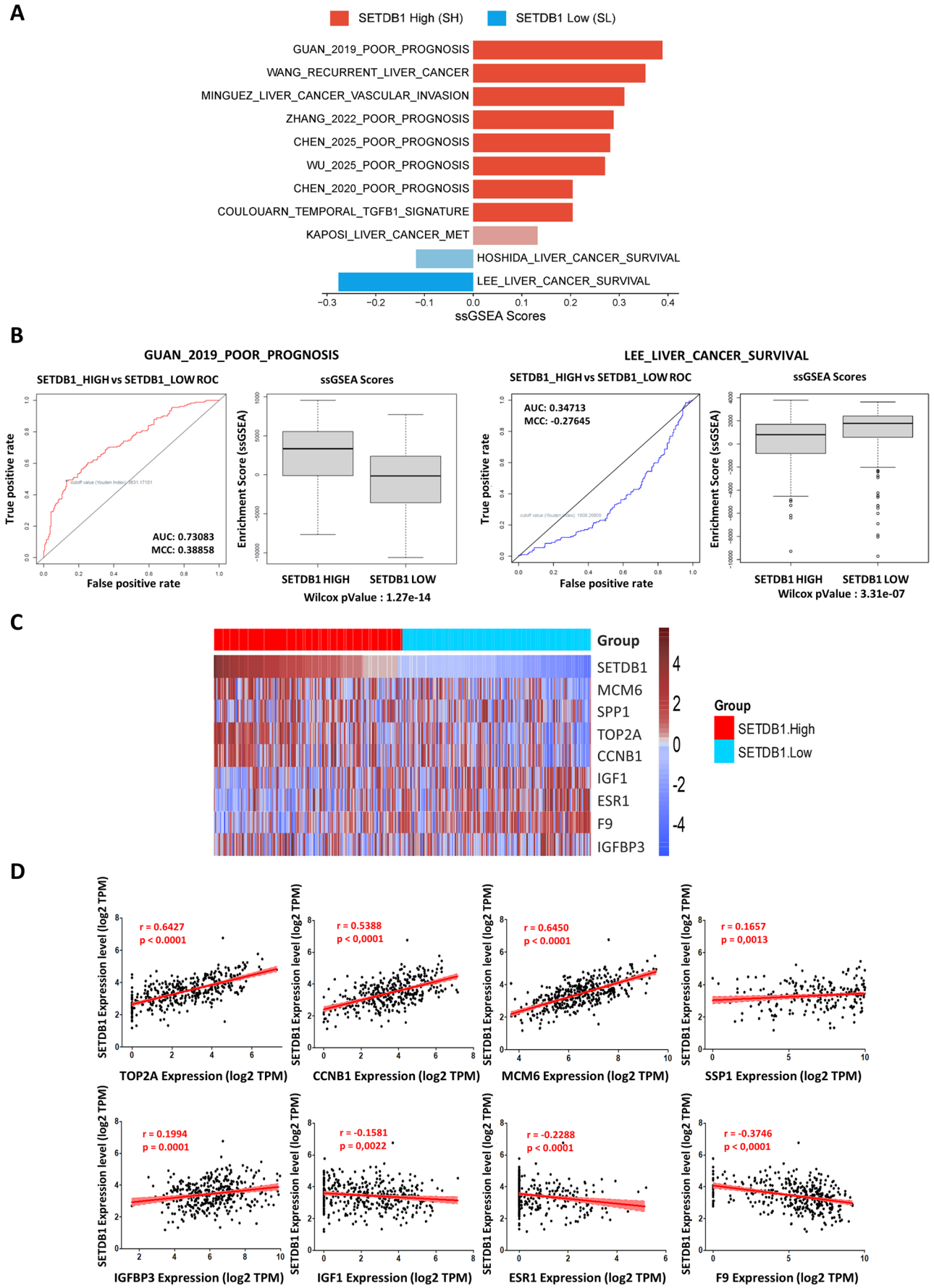

A
SETDB1 High (SH)
SETDB1 Low (SL)
GUAN_2019_POOR_PROGNOSIS
WANG_RECURRENT_LIVER_CANCER
MINGUEZ_LIVER_CANCER_VASCULAR_INVASION
ZHANG_2022_POOR_PROGNOSIS
CHEN_2025_POOR_PROGNOSIS
WU_2025_POOR_PROGNOSIS
CHEN_2020_POOR_PROGNOSIS
COULOUARN_TEMPORAL_TGFB1_SIGNATURE
KAPOSI_LIVER_CANCER_MET
HOSHIDA_LIVER_CANCER_SURVIVAL
LEE_LIVER_CANCER_SURVIVAL
ssGSEA Scores
B
GUAN_2019_POOR_PROGNOSIS
SETDB1_HIGH vs SETDB1_LOW ROC
True positive rate
False positive rate
AUC: 0.73083
MCC: 0.38858
ssGSEA Scores
Enrichment Score (ssGSEA)
SETDB1 HIGH
SETDB1 LOW
Wilcox pValue : 1.27e-14
LEE_LIVER_CANCER_SURVIVAL
SETDB1_HIGH vs SETDB1_LOW ROC
AUC: 0.34713
MCC: -0.27645
True positive rate
False positive rate
ssGSEA Scores
Enrichment Score (ssGSEA)
SETDB1 HIGH
SETDB1 LOW
Wilcox pValue : 3.31e-07
C
Group
SETDB1
MCM6
SPP1
TOP2A
CCNB1
IGF1
ESR1
F9
IGFBP3
4
2
0
-2
-4
Group
SETDB1.High
SETDB1.Low
D
SETDB1 Expression level (log2 TPM)
r = 0.6427
p < 0.0001
TOP2A Expression (log2 TPM)
r = 0.5388
p < 0,0001
CCNB1 Expression (log2 TPM)
r = 0.6450
p < 0.0001
MCM6 Expression (log2 TPM)
r = 0.1657
p = 0,0013
SSP1 Expression (log2 TPM)
r = 0.1994
p = 0.0001
IGFBP3 Expression (log2 TPM)
r = -0.1581
p = 0,0022
IGF1 Expression (log2 TPM)
r = -0.2288
p < 0.0001
ESR1 Expression (log2 TPM)
r = -0.3746
p < 0,0001
F9 Expression (log2 TPM)

Fig. 2 SETDB1 expression is associated with poor prognosis gene signatures. **(A)** Bar plot representing single-sample Gene Set Enrichment Analysis (ssGSEA) enrichment scores of gene signatures related to various adverse clinical outcomes. Positive scores (red bars) indicate an enrichment of the signature in the SETDB1 high expression (SH) group in the TCGA-LIHC cohort, while negative scores (blue bars) indicate enrichment in the SETDB1 low expression (SL) group. Gene sets with a significant enrichment ($P<0.05$, Wilcoxon test) are shown with saturated colors; non-significant sets ($P>0.05$) are shown with attenuated colors. **(B)** Receiver Operating Characteristic (ROC) curves and corresponding box plots illustrating the differential enrichment of two specific poor prognosis signatures between the SH and SL groups: "Guan_2019_Poor_Prognosis" and "Lee_Liver_Cancer_Survival". The Area Under the Curve (AUC) is provided for each ROC curve to assess the predictive power of the signatures. The Wilcoxon P-value for each comparison is also shown. **(C)** Heatmap showing the correlation between SETDB1 expression and the expression of genes from the "Chen_2020_Poor_Prognosis" signature across samples from the TCGA-LIHC cohort. Each row represents a gene from the signature, and each column represents a patient sample. The color intensity corresponds to the expression level (red for high, blue for low). **(D)** Scatter plots illustrating the Spearman's correlation between SETDB1 expression and the expression of individual genes from the "Chen_2020_Poor_Prognosis" signature. The Spearman's correlation coefficient (r) and the corresponding P-value are indicated on each plot

genes (*GZMH*, *CD8A*, *CD8B*, *CTSW*, *KLRK1*, and *IGJ*), all associated with robust cytotoxic T/NK cell activity and immunotherapy sensitivity [26]. Enrichment of this signature was higher in the SL group (Fig. 3E), consistent with a greater predicted benefit from immune-based treatments.

Collectively, these results highlight that SETDB1-low tumors exhibit a favorable immunological profile with enhanced infiltration of cytotoxic immune cells and upregulation of immune checkpoint molecules, correlating with a higher likelihood of response to immunotherapies. In contrast, SETDB1-high tumors exhibit immune exclusion and signatures of resistance, which may contribute to their poor prognosis. These results were consistently reproduced in the independent validation cohort GSE104580 (Fig. S1B, C).

3.2.2 SETDB1 overexpression is associated with TACE refractoriness

To investigate the association between SETDB1 expression and TACE response, we performed ssGSEA using curated gene signatures in the TCGA-LIHC cohort. Our findings suggest that the SH group exhibits features of TACE refractoriness, as they were significantly enriched in gene sets associated with TACE resistance (Fig. 3F). This association was further confirmed in the GSE104580 cohort (Fig. 3G), where we observed a higher expression of SETDB1 in TACE non-responders compared to responders. These results suggest that SETDB1 expression may be a valuable prognostic biomarker for TACE response in HCC patients, potentially influencing treatment decisions.

3.2.3 SETDB1 expression predicts Sorafenib response

We next evaluated the relationship between SETDB1 expression and sorafenib sensitivity, a common therapeutic agent and a tyrosine kinase inhibitor (TKI) used for HCC. Utilizing ssGSEA, we investigated the enrichment of sorafenib resistance-related signatures in both the SH and SL groups (Fig. 3H). The SH group exhibited significant enrichment in gene sets associated with sorafenib resistance, including those linked to multidrug resistance (MDR) [27] and ferroptosis [28], in both the TCGA-LIHC cohort and the GSE104580 validation cohort (Fig. S1D). This suggests that the SH group may exhibit increased resistance to sorafenib treatment due to altered gene expression involved in MDR and ferroptosis pathways. However, we did not observe a significant difference in SETDB1 expression between sorafenib responders and non-responders in the GSE109211 cohort (Fig. 3I). Despite this, a trend suggesting a possible association between SETDB1 expression and sorafenib response was observed, requiring further investigation.

In conclusion, our findings highlight the potential involvement of SETDB1 in mediating therapeutic resistance. These findings suggest that SETDB1 may serve as a valuable biomarker to predict treatment response and guide personalized therapeutic strategies for HCC patients.

3.3 SETDB1 expression defines distinct proliferative, metabolic, and immune transcriptional programs in TCGA-LIHC

ssGSEA and GSEA were performed using the "Hallmarks," "Kegg_Pathway," and "Wikipathways" gene set collections to interpret the transcriptomic differences between SH and SL groups.

In the SH group, both ssGSEA and GSEA consistently revealed enrichment for signatures related to cell proliferation, cell cycle progression, and DNA damage repair. Specifically, Hallmarks pathways such as E2F Targets, G2M Checkpoint, Mitotic Spindle, and MYC Targets (V1 and V2) were significantly upregulated (Fig. 4A), suggesting active mitotic processes and oncogenic transcriptional programs. Parallel Kegg_Pathway analysis showed significant enrichment for the cell cycle, DNA replication, and multiple DNA repair pathways (base excision repair, nucleotide excision repair, mismatch repair, and homologous recombination) in SH tumors (Fig. 4B), consistent with enhanced DNA repair capacity. Similarly, Wikipathways analysis revealed an overrepresentation of modules related to DNA damage response and transcriptional regulation (Fig. 4C). These results collectively suggest that high SETDB1 expression is associated with a proliferative, undifferentiated phenotype

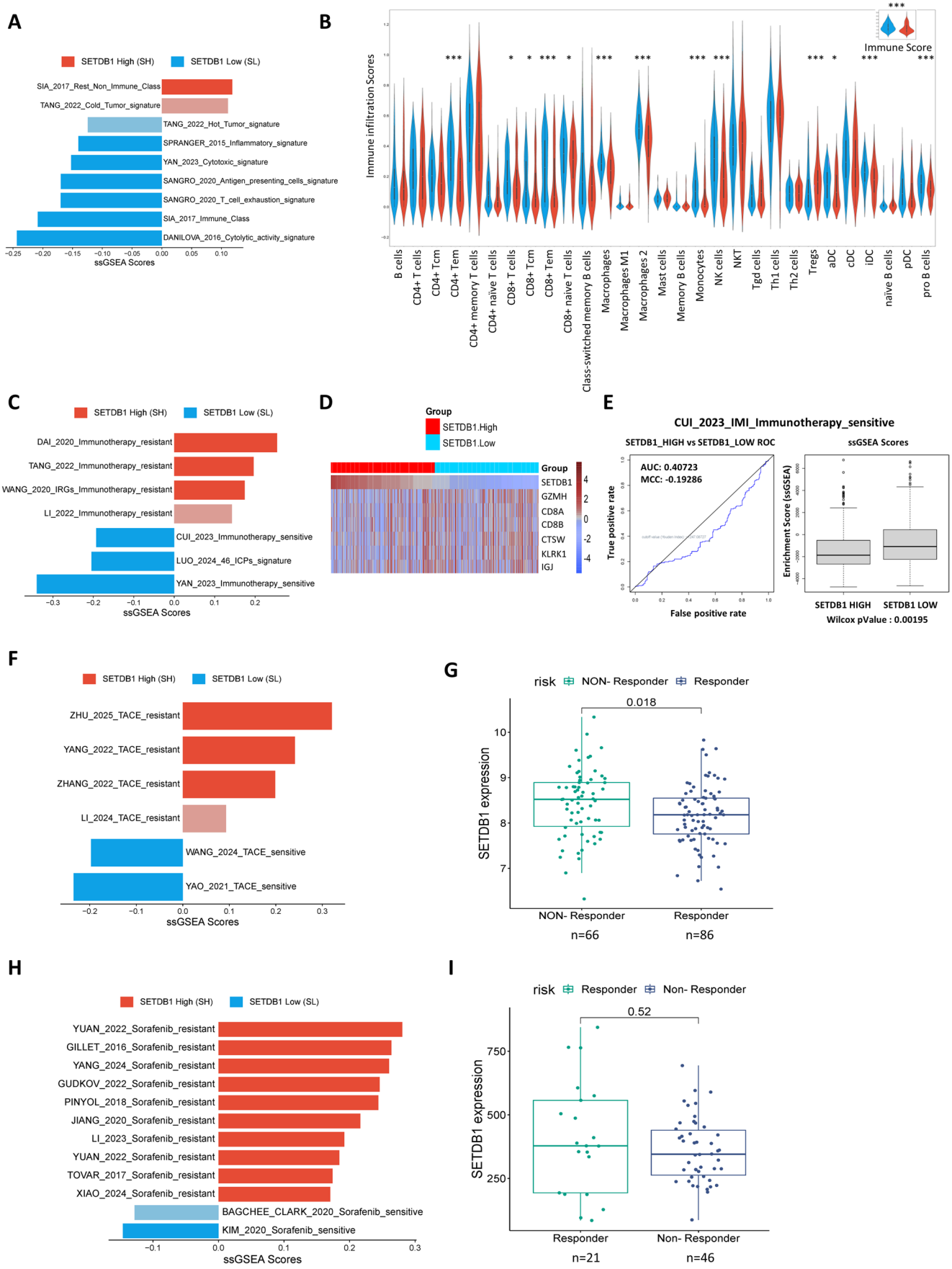

A
SETDB1 High (SH)
SETDB1 Low (SL)
SIA_2017_Rest_Non_Immune_Class
TANG_2022_Cold_Tumor_signature
TANG_2022_Hot_Tumor_signature
SPRANGER_2015_Inflammatory_signature
YAN_2023_Cytotoxic_signature
SANGRO_2020_Antigen_presenting_cells_signature
SANGRO_2020_T_cell_exhaustion_signature
SIA_2017_Immune_Class
DANILOVA_2016_Cytolytic_activity_signature
ssGSEA Scores
B
Immune infiltration Scores
Immune Score
B cells
CD4+ T cells
CD4+ Tcm
CD4+ Tem
CD4+ memory T cells
CD4+ naive T cells
CD8+ T cells
CD8+ Tcm
CD8+ Tem
CD8+ naive T cells
Class-switched memory B cells
Macrophages
Macrophages M1
Macrophages 2
Mast cells
Memory B cells
Monocytes
NK cells
NKT
Tgd cells
Th1 cells
Th2 cells
Tregs
aDC
cDC
iDC
naive B cells
pDC
pro B cells
C
SETDB1 High (SH)
SETDB1 Low (SL)
DAI_2020_Immunotherapy_resistant
TANG_2022_Immunotherapy_resistant
WANG_2020_IRGs_Immunotherapy_resistant
LI_2022_Immunotherapy_resistant
CUI_2023_Immunotherapy_sensitive
LUO_2024_46_ICPs_signature
YAN_2023_Immunotherapy_sensitive
ssGSEA Scores
D
Group
SETDB1.High
SETDB1.Low
Group
SETDB1
GZMH
CD8A
CD8B
CTSW
KLRK1
IGJ
E
CUI_2023_IMI_Immunotherapy_sensitive
SETDB1_HIGH vs SETDB1_LOW ROC
AUC: 0.40723
MCC: -0.19286
True positive rate
False positive rate
ssGSEA Scores
Enrichment Score (ssGSEA)
SETDB1 HIGH
SETDB1 LOW
Wilcox pValue : 0.00195
F
SETDB1 High (SH)
SETDB1 Low (SL)
ZHU_2025_TACE_resistant
YANG_2022_TACE_resistant
ZHANG_2022_TACE_resistant
LI_2024_TACE_resistant
WANG_2024_TACE_sensitive
YAO_2021_TACE_sensitive
ssGSEA Scores
G
risk
NON- Responder
Responder
0.018
SETDB1 expression
NON- Responder
Responder
n=66
n=86
H
SETDB1 High (SH)
SETDB1 Low (SL)
YUAN_2022_Sorafenib_resistant
GILLET_2016_Sorafenib_resistant
YANG_2024_Sorafenib_resistant
GUDKOV_2022_Sorafenib_resistant
PINYOL_2018_Sorafenib_resistant
JIANG_2020_Sorafenib_resistant
LI_2023_Sorafenib_resistant
YUAN_2022_Sorafenib_resistant
TOVAR_2017_Sorafenib_resistant
XIAO_2024_Sorafenib_resistant
BAGCHEE_CLARK_2020_Sorafenib_sensitive
KIM_2020_Sorafenib_sensitive
ssGSEA Scores
I
risk
Responder
Non- Responder
0.52
SETDB1 expression
Responder
Non- Responder
n=21
n=46

Fig. 3 SETDB1 expression predicts response to immunotherapy and systemic therapies in HCC. **(A)** Single-sample Gene Set Enrichment Analysis (ssGSEA) scores of gene signatures related to the immune response, compared between the SETDB1-high (SH) and SETDB1-low (SL) groups in the TCGA-LIHC cohort. Positive scores indicate enrichment in the SH group (red), while negative scores indicate enrichment in the SL group (blue). Saturated colors indicate significant gene sets ($P<0.05$, Wilcoxon test); attenuated colors are non-significant ($P>0.05$). **(B)** Violin plots showing the immune cell infiltration scores, as estimated by the XCell algorithm, for 29 immune cell types in SH (red) and SL (blue) groups. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ns = not significant. **(C)** ssGSEA scores of gene signatures related to sensitivity and resistance to immune checkpoint inhibitors (ICIs). Gene sets related to ICI resistance are enriched in the SH group, while those related to sensitivity are enriched in the SL group. **(D)** Heatmap illustrating the correlation between SETDB1 expression and genes from the Immune Index (IMI) signature, a prognostic model for HCC. **(E)** Receiver Operating Characteristic (ROC) curve and box plots evaluating the predictive performance of the IMI signature in distinguishing between SH and SL groups. The Area Under the Curve (AUC) and Wilcoxon P-value are provided. **(F)** ssGSEA scores of gene signatures associated with response to transarterial chemoembolization (TACE). The TACE-resistant signatures are enriched in the SH group. **(G)** Box plot comparing SETDB1 expression levels in TACE responders ($n=66$) and non-responders ($n=86$) from the GSE104580 cohort. The P-value from the statistical comparison is shown. **(H)** ssGSEA scores of gene signatures related to sensitivity and resistance to sorafenib. Sorafenib resistance-related signatures are enriched in the SH group. **(I)** Box plot comparing SETDB1 expression levels in sorafenib responders ($n=21$) and non-responders ($n=46$) from the GSE109211 cohort. The P-value from the statistical comparison is shown

in HCC, characterized by increased genomic instability and the upregulation of repair and checkpoint mechanisms.

In contrast, the SL group showed significant enrichment for immune-related and metabolic pathways. Immune processes included IL6/JAK/STAT3 signaling, Type I and II Interferon Responses, Toll-like Receptor signaling, and T cell receptor signaling, highlighting a more immunoreactive tumor microenvironment. Metabolic programs, particularly those linked to hepatocytic function (lipid metabolism, amino acid metabolism, and glucose/energy pathways), were also predominant. These findings suggest a metabolically active and well-differentiated hepatic state in the SL group.

In conclusion, the SH group appears to harbor aggressive and proliferative features with elevated DNA repair activity, whereas the SL group displays characteristics of differentiated hepatocytes with active metabolic and immune-competent properties.

ssGSEA and GSEA were performed using the "Hallmarks," "Kegg_Pathway," and "Wikipathways" gene set collections to characterize the biological processes associated with SETDB1 expression. These analyses were aimed at interpreting the transcriptomic differences between SH and SL groups.

In the SH group, consistent enrichment for signatures related to cell proliferation, cell cycle progression, and DNA damage repair was revealed by both ssGSEA and GSEA. Specifically, Hallmarks pathways such as E2F Targets, G2M Checkpoint, Mitotic Spindle, and MYC Targets (V1 and V2) were significantly upregulated (Fig. 4A), suggesting active mitotic processes and oncogenic transcriptional programs. In parallel, significant enrichment for the cell cycle, DNA replication, base excision repair, nucleotide excision repair, mismatch repair, and homologous recombination pathways was shown in SH tumors by Kegg_Pathway analysis (Fig. 4B), which is consistent with an enhanced DNA repair capacity. Similarly, an overrepresentation of modules related to DNA damage response and transcriptional regulation was revealed by Wikipathways analysis (Fig. 4C). These results collectively suggest that high SETDB1 expression is associated with increased genomic instability, necessitating the upregulation of repair and checkpoint mechanisms, and supports a proliferative, undifferentiated phenotype in HCC.

In contrast, significant enrichment for immune-related and metabolic pathways was shown in the SL group. The immune processes included IL6/JAK/STAT3 signaling, Type I and II Interferon Responses, Toll-like Receptor signaling, and T cell receptor signaling, highlighting a more immunoreactive tumor microenvironment. Metabolic programs, particularly those linked to hepatocytic function, were also predominant, including lipid metabolism, amino acid metabolism, and glucose/energy pathways. These findings suggest a metabolically active and well-differentiated hepatic state in the SL group.

In conclusion, the SH group appears to harbor aggressive and proliferative features with elevated DNA repair activity, whereas the SL group displays characteristics of differentiated hepatocytes, with active metabolic and immune-competent properties.

3.4 SETDB1 expression drives HCC molecular subclassification

To further investigate the role of SETDB1 in shaping HCC intratumoral heterogeneity, we performed a subclass-oriented transcriptomic analysis based on established molecular classification systems. The heatmap (Fig. 5A) illustrates enrichment scores from ssGSEA for various HCC molecular subclasses, as well as signatures of hepatocytic differentiation and metabolic activity, thus providing an integrative view of HCC molecular heterogeneity.

ssGSEA analysis (Fig. 5B, C) showed that the SL group was enriched for subclass signatures linked to well-differentiated, metabolically active HCC, including Hoshida's S3, Chiang's interferon subclass, and Désert's periportal subclass. These subclasses retain hepatocytic identity, display lower mitotic activity, and upregulate metabolic pathways such as the urea cycle, lipid metabolism, and bile acid

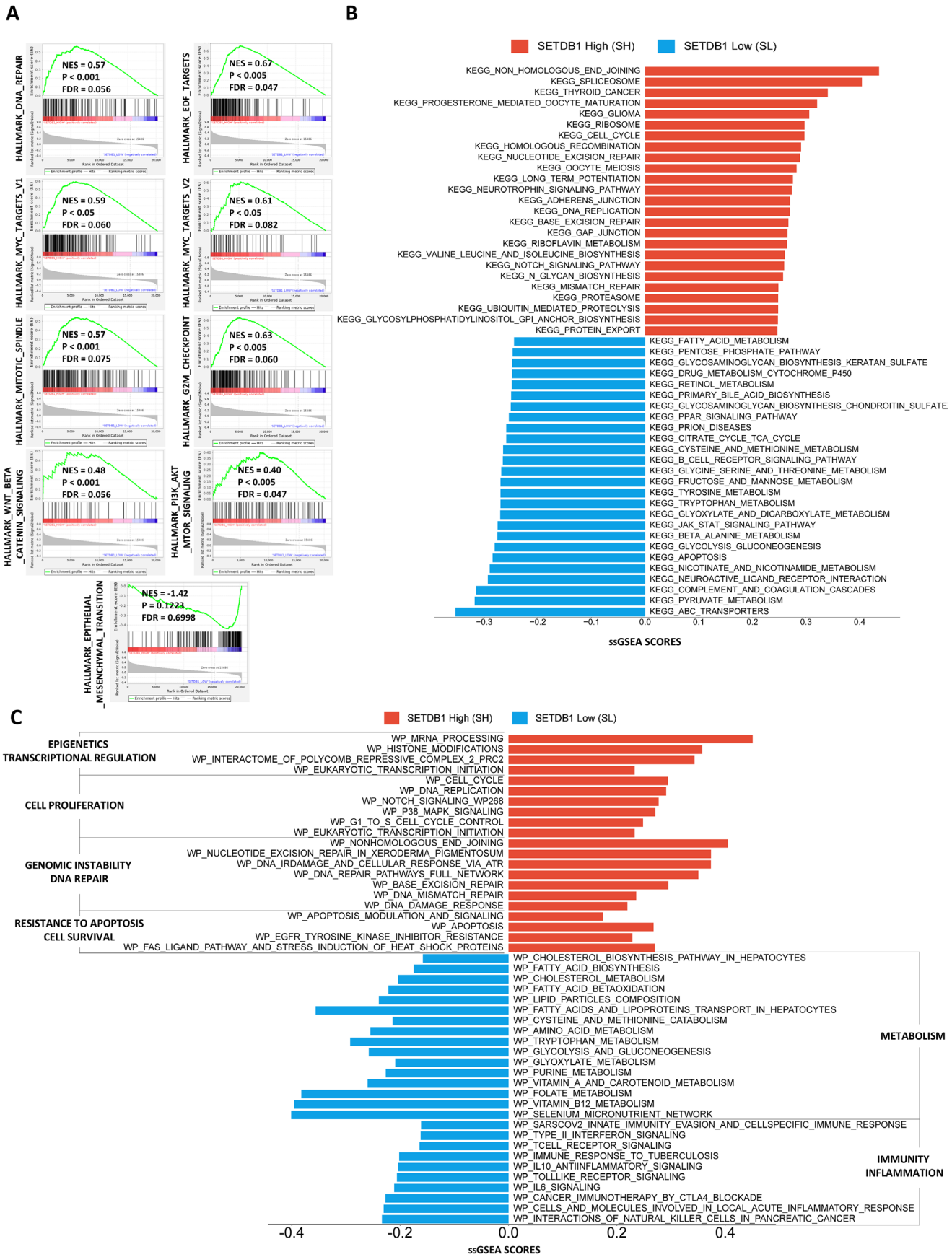

A
HALLMARK_DNA_REPAIR
NES = 0.57
P < 0.001
FDR = 0.056
HALLMARK_EDF_TARGETS
NES = 0.67
P < 0.005
FDR = 0.047
HALLMARK_MYC_TARGETS_V1
NES = 0.59
P < 0.05
FDR = 0.060
HALLMARK_MYC_TARGETS_V2
NES = 0.61
P < 0.05
FDR = 0.082
HALLMARK_MITOTIC_SPINDLE
NES = 0.57
P < 0.001
FDR = 0.075
HALLMARK_G2M_CHECKPOINT
NES = 0.63
P < 0.005
FDR = 0.060
HALLMARK_WNT_BETA_CATENIN_SIGNALING
NES = 0.48
P < 0.001
FDR = 0.056
HALLMARK_PI3K_AKT_MTOR_SIGNALING
NES = 0.40
P < 0.005
FDR = 0.047
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
NES = -1.42
P = 0.1223
FDR = 0.6998
B
SETDB1 High (SH)
SETDB1 Low (SL)
KEGG_NON_HOMOLOGOUS_END_JOINING
KEGG_SPLICEOSOME
KEGG_THYROID_CANCER
KEGG_PROGESTERONE_MEDIATED_OOCYTE_MATURATION
KEGG_GLIOMA
KEGG_RIBOSOME
KEGG_CELL_CYCLE
KEGG_HOMOLOGOUS_RECOMBINATION
KEGG_NUCLEOTIDE_EXCISION_REPAIR
KEGG_OOCYTE_MEIOSIS
KEGG_LONG_TERM_POTENTIATION
KEGG_NEUROTROPHIN_SIGNALING_PATHWAY
KEGG_ADHERENS_JUNCTION
KEGG_DNA_REPLICATION
KEGG_BASE_EXCISION_REPAIR
KEGG_GAP_JUNCTION
KEGG_RIBOFLAVIN_METABOLISM
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_BIOSYNTHESIS
KEGG_NOTCH_SIGNALING_PATHWAY
KEGG_N_GLYCAN_BIOSYNTHESIS
KEGG_MISMATCH_REPAIR
KEGG_PROTEASOME
KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS
KEGG_GLYCOSYLPHOSPHATIDYLINOSITOL_GPI_ANCHOR_BIOSYNTHESIS
KEGG_PROTEIN_EXPORT
KEGG_FATTY_ACID_METABOLISM
KEGG_PENTOSE_PHOSPHATE_PATHWAY
KEGG_GLYCOSAMINOGLYCAN_BIOSYNTHESIS_KERATAN_SULFATE
KEGG_DRUG_METABOLISM_CYTOCHROME_P450
KEGG_RETINOL_METABOLISM
KEGG_PRIMARY_BILE_ACID_BIOSYNTHESIS
KEGG_GLYCOSAMINOGLYCAN_BIOSYNTHESIS_CHONDROITIN_SULFATE
KEGG_PPAR_SIGNALING_PATHWAY
KEGG_PRION_DISEASES
KEGG_CITRATE_CYCLE_TCA_CYCLE
KEGG_CYSTEINE_AND_METHIONINE_METABOLISM
KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY
KEGG_GLYCINE_SERINE_AND_THREONINE_METABOLISM
KEGG_FRUCTOSE_AND_MANNOSE_METABOLISM
KEGG_TYROSINE_METABOLISM
KEGG_TRYPTOPHAN_METABOLISM
KEGG_GLYOXYLATE_AND_DICARBOXYLATE_METABOLISM
KEGG_JAK_STAT_SIGNALING_PATHWAY
KEGG_BETA_ALANINE_METABOLISM
KEGG_GLYCOLYSIS_GLUCONEOGENESIS
KEGG_APOPTOSIS
KEGG_NICOTINATE_AND_NICOTINAMIDE_METABOLISM
KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES
KEGG_PYRUVATE_METABOLISM
KEGG_ABC_TRANSPORTERS
−0.3 −0.2 −0.1 0.0 0.1 0.2 0.3 0.4
ssGSEA SCORES
C
SETDB1 High (SH)
SETDB1 Low (SL)
EPIGENETICS
TRANSCRIPTIONAL REGULATION
WP_MRNA_PROCESSING
WP_HISTONE_MODIFICATIONS
WP_INTERACTOME_OF_POLYCOMB_REPRESSIVE_COMPLEX_2_PRC2
WP_EUKARYOTIC_TRANSCRIPTION_INITIATION
CELL PROLIFERATION
WP_CELL_CYCLE
WP_DNA_REPLICATION
WP_NOTCH_SIGNALING_WP268
WP_P38_MAPK_SIGNALING
WP_G1_TO_S_CELL_CYCLE_CONTROL
WP_EUKARYOTIC_TRANSCRIPTION_INITIATION
GENOMIC INSTABILITY
DNA REPAIR
WP_NONHOMOLOGOUS_END_JOINING
WP_NUCLEOTIDE_EXCISION_REPAIR_IN_XERODERMA_PIGMENTOSUM
WP_DNA_IRDAMAGE_AND_CELLULAR_RESPONSE_VIA_ATR
WP_DNA_REPAIR_PATHWAYS_FULL_NETWORK
WP_BASE_EXCISION_REPAIR
WP_DNA_MISMATCH_REPAIR
WP_DNA_DAMAGE_RESPONSE
RESISTANCE TO APOPTOSIS
CELL SURVIVAL
WP_APOPTOSIS_MODULATION_AND_SIGNALING
WP_APOPTOSIS
WP_EGFR_TYROSINE_KINASE_INHIBITOR_RESISTANCE
WP_FAS_LIGAND_PATHWAY_AND_STRESS_INDUCTION_OF_HEAT_SHOCK_PROTEINS
METABOLISM
WP_CHOLESTEROL_BIOSYNTHESIS_PATHWAY_IN_HEPATOCYTES
WP_FATTY_ACID_BIOSYNTHESIS
WP_CHOLESTEROL_METABOLISM
WP_FATTY_ACID_BETAOXIDATION
WP_LIPID_PARTICLES_COMPOSITION
WP_FATTY_ACIDS_AND_LIPOPROTEINS_TRANSPORT_IN_HEPATOCYTES
WP_CYSTEINE_AND_METHIONINE_CATABOLISM
WP_AMINO_ACID_METABOLISM
WP_TRYPTOPHAN_METABOLISM
WP_GLYCOLYSIS_AND_GLUCONEOGENESIS
WP_GLYOXYLATE_METABOLISM
WP_PURINE_METABOLISM
WP_VITAMIN_A_AND_CAROTENOID_METABOLISM
WP_FOLATE_METABOLISM
WP_VITAMIN_B12_METABOLISM
WP_SELENIUM_MICRONUTRIENT_NETWORK
IMMUNITY
INFLAMMATION
WP_SARSCOV2_INNATE_IMMUNITY_EVASION_AND_CELLSPECIFIC_IMMUNE_RESPONSE
WP_TYPE_II_INTERFERON_SIGNALING
WP_TCELL_RECEPTOR_SIGNALING
WP_IMMUNE_RESPONSE_TO_TUBERCULOSIS
WP_IL10_ANTIINFLAMMATORY_SIGNALING
WP_TOLLLIKE_RECEPTOR_SIGNALING
WP_IL6_SIGNALING
WP_CANCER_IMMUNOTHERAPY_BY_CTLA4_BLOCKADE
WP_CELLS_AND_MOLECULES_INVOLVED_IN_LOCAL_ACUTE_INFLAMMATORY_RESPONSE
WP_INTERACTIONS_OF_NATURAL_KILLER_CELLS_IN_PANCREATIC_CANCER
-0.4 -0.2 0.0 0.2 0.4
ssGSEA SCORES

Fig. 4 SETDB1 expression is associated with distinct proliferative, metabolic, and immune transcriptional programs in HCC. **(A)** Gene Set Enrichment Analysis (GSEA) plots for six selected gene sets from the Hallmark collection, comparing gene expression profiles between the SETDB1-high (SH) and SETDB1-low (SL) groups. Each plot displays the Normalized Enrichment Score (NES), nominal P-value, and False Discovery Rate (FDR) q-value. Gene sets with an FDR q-value < 0.25 and a NES > 1 (or NES < -1) were considered significantly enriched. **(B)** Bar plot of ssGSEA scores for gene sets from the KEGG collection. Red bars represent pathways enriched in the SH group, while blue bars represent pathways enriched in the SL group. Only pathways with a significant enrichment ($P < 0.05$, Wilcoxon test) are shown. **(C)** Bar plot of ssGSEA scores for gene sets from the WikiPathways collection. The plot is organized by biological function categories, including Epigenetics, Cell Proliferation, Genomic Instability, DNA Repair, Resistance to Apoptosis, Cell Survival, Metabolism, Immunity, and Inflammation. Red bars indicate enrichment in the SH group and blue bars indicate enrichment in the SL group

biosynthesis, consistent with our previous ssGSEA results. The SL group also exhibited enrichment of immune-related pathways, including interferon-γ response, IL6/JAK/STAT3 signaling, and inflammatory signaling, in line with Chiang's interferon and Hoshida's S3 classes.

In contrast, the SH group was enriched in dedifferentiated and aggressive molecular subclasses, including Boyault's G1/G2 and G3, Chiang's proliferation subclass, Hoshida's S2, and Désert's STEM-type. Consistently, the results obtained in the GSE104580 cohort (Fig. S1E) confirmed the molecular subclassification patterns observed in the TCGA-LIHC cohort, although with slightly lower statistical significance.

Overall, the SH group is characterized by poor differentiation, TP53 mutations, chromosomal instability, and activation of oncogenic pathways such as MYC and TGF-beta signaling. In line with our previous observations, the SH group also showed strong activation of cell cycle, DNA repair, MYC targets, and chromatin remodeling programs, alongside a marked downregulation of metabolic pathways typical of Chiang's proliferation subclass. The SL group was enriched in well-differentiated tumors (grades 1–2) with preserved hepatocytic morphology, whereas the SH group predominated in grade 3 tumors, consistent with a dedifferentiated histopathological profile, higher mitotic activity, and more aggressive clinical behavior (Fig. 5D). In support of these findings, a heatmap (Fig. 5E) shows that the SH group exhibited higher expression of Alpha-fetoprotein (*AFP*), Glypican-3 (*GPC3*), and *EPCAM* (Fig. 5F), all markers associated with dedifferentiation, progenitor-like features, and aggressive behavior. This is further supported by enrichment in Désert's STEM-type signature.

3.5 SETDB1-high HCC tumors exhibit stemness-associated transcriptional programs, stem cell marker expression, and activation of pluripotency signaling pathways

Following the observation that SETDB1-high tumors are enriched in poorly differentiated and stem-like subclasses (Fig. 5), we investigated whether SETDB1 expression is functionally associated with cancer stemness programs in HCC. Using CancerStemnessOnline, we quantified mRNAsi and StemnessIndex and found both to be elevated in the SH group, indicating a stronger stem-like transcriptional identity (Fig. 6A). To extend these results, GSEA and ssGSEA with the PanglaoDB collection revealed a strong enrichment of pluripotent stem cell signatures in SH tumors, while hepatocyte signatures predominated in the SL group (Fig. 6B, C), highlighting the dichotomy between progenitor-like and hepatocytic differentiation states.

Of note, the SH group showed high expression of *KRT19*, a canonical marker of hepatic progenitors and cholangiocyte-like HCC associated with poor prognosis [29]. In line with this, the SH group exhibited strong enrichment of multiple stemness-related signatures (Fig. 6A and D). *SETDB1* expression positively correlated with mRNAsi, the PanglaoDB pluripotency score, and the Desert STEM score (Spearman $r = 0.4349$, $r = 0.6303$, and $r = 0.4139$, respectively, all $P < 0.0001$) (Fig. 6E). We also found that most genes in a custom HCC CSC signature, including *PROM1* (CD133), *CD24*, *EPCAM*, *KRT19*, *LGR5*, *THY1*, *ICAM1*, *SOX9*, *POU5F1*, *NANOG*, and *KLF4*, were positively correlated with *SETDB1* expression (Fig. 6F), reinforcing its association with progenitor-like states and tumor-initiating potential.

Mechanistically, ssGSEA revealed the activation of multiple developmental and stemness-related signaling pathways in SH tumors, including NOTCH, WNT, mTOR, and TGF-β (Fig. 6G, top). These findings were complemented by the enrichment of protein–protein interaction networks centered on pluripotency transcription factors (*POU5F1* (OCT4), *SOX2*, *NANOG*, *KLF4*, and *MYC*) (Fig. 6G, bottom), suggesting that SETDB1 may act upstream or in concert with these core stemness regulators. Analysis of the GSE104580 cohort (Fig. S1F–H) confirmed the trends in stemness-associated programs observed in TCGA-LIHC. Together, these results support a model in which SETDB1 drives dedifferentiation and stemness in HCC through coordinated transcriptional and signaling networks, fostering intratumoral plasticity and aggressive tumor behavior.

Given the role of non-coding RNAs in regulating differentiation and stemness, we analyzed miRNA profiles in SH and SL tumors. Differential expression analysis identified 5 miRNAs upregulated in the SH group and 13 in the

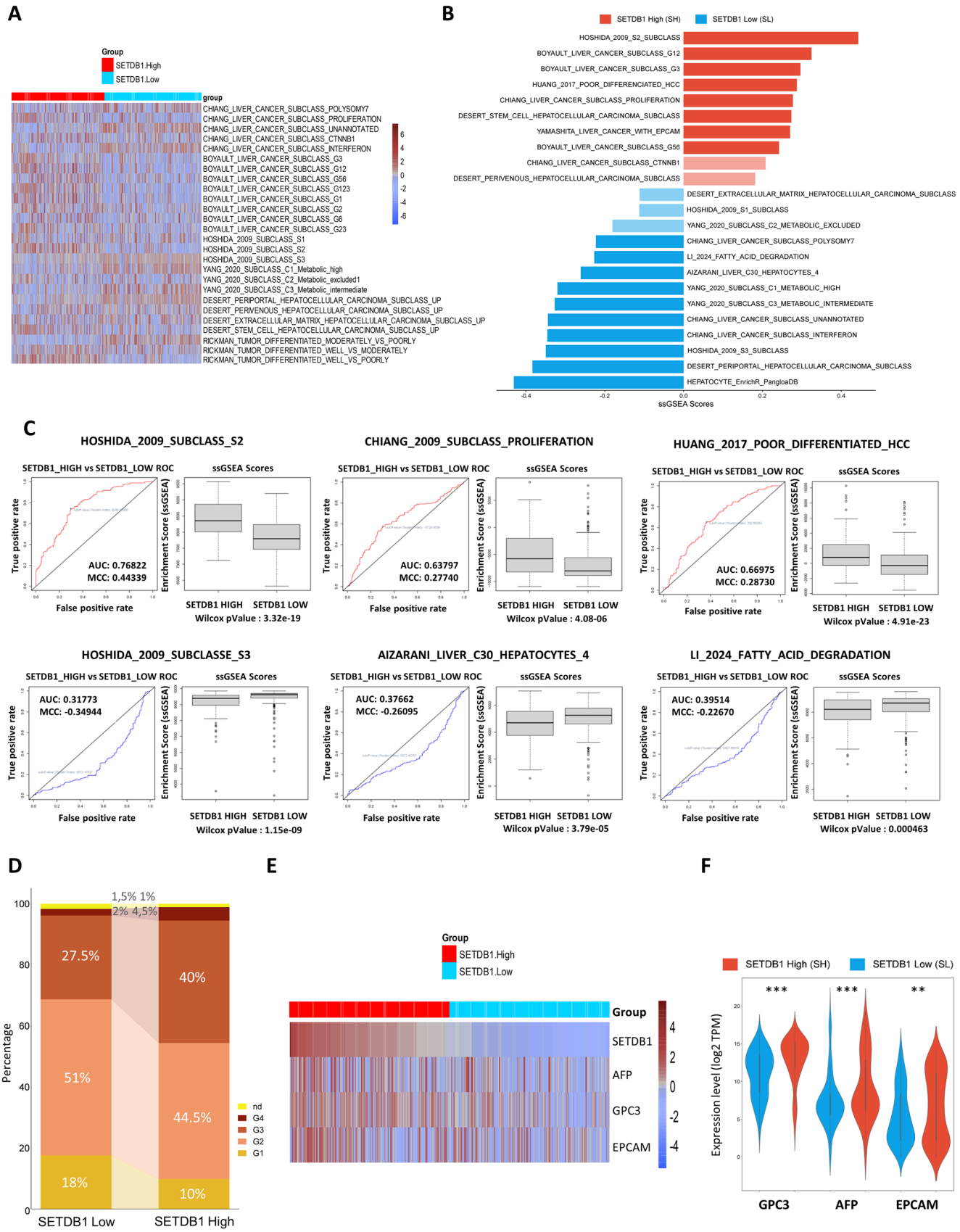

A
Group
SETDB1.High
SETDB1.Low
group
CHIANG_LIVER_CANCER_SUBCLASS_POLYSOMY7
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION
CHIANG_LIVER_CANCER_SUBCLASS_UNANNOTATED
CHIANG_LIVER_CANCER_SUBCLASS_CTNNB1
CHIANG_LIVER_CANCER_SUBCLASS_INTERFERON
BOYAULT_LIVER_CANCER_SUBCLASS_G3
BOYAULT_LIVER_CANCER_SUBCLASS_G12
BOYAULT_LIVER_CANCER_SUBCLASS_G56
BOYAULT_LIVER_CANCER_SUBCLASS_G123
BOYAULT_LIVER_CANCER_SUBCLASS_G1
BOYAULT_LIVER_CANCER_SUBCLASS_G2
BOYAULT_LIVER_CANCER_SUBCLASS_G6
BOYAULT_LIVER_CANCER_SUBCLASS_G23
HOSHIDA_2009_SUBCLASS_S1
HOSHIDA_2009_SUBCLASS_S2
HOSHIDA_2009_SUBCLASS_S3
YANG_2020_SUBCLASS_C1_Metabolic_high
YANG_2020_SUBCLASS_C2_Metabolic_excluded1
YANG_2020_SUBCLASS_C3_Metabolic_intermediate
DESERT_PERIPORTAL_HEPATOCELLULAR_CARCINOMA_SUBCLASS_UP
DESERT_PERIVENOUS_HEPATOCELLULAR_CARCINOMA_SUBCLASS_UP
DESERT_EXTRACELLULAR_MATRIX_HEPATOCELLULAR_CARCINOMA_SUBCLASS_UP
DESERT_STEM_CELL_HEPATOCELLULAR_CARCINOMA_SUBCLASS_UP
RICKMAN_TUMOR_DIFFERENTIATED_MODERATELY_VS_POORLY
RICKMAN_TUMOR_DIFFERENTIATED_WELL_VS_MODERATELY
RICKMAN_TUMOR_DIFFERENTIATED_WELL_VS_POORLY
6
4
2
0
-2
-4
-6
B
SETDB1 High (SH)
SETDB1 Low (SL)
HOSHIDA_2009_S2_SUBCLASS
BOYAULT_LIVER_CANCER_SUBCLASS_G12
BOYAULT_LIVER_CANCER_SUBCLASS_G3
HUANG_2017_POOR_DIFFERENCIATED_HCC
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION
DESERT_STEM_CELL_HEPATOCELLULAR_CARCINOMA_SUBCLASS
YAMASHITA_LIVER_CANCER_WITH_EPCAM
BOYAULT_LIVER_CANCER_SUBCLASS_G56
CHIANG_LIVER_CANCER_SUBCLASS_CTNNB1
DESERT_PERIVENOUS_HEPATOCELLULAR_CARCINOMA_SUBCLASS
DESERT_EXTRACELLULAR_MATRIX_HEPATOCELLULAR_CARCINOMA_SUBCLASS
HOSHIDA_2009_S1_SUBCLASS
YANG_2020_SUBCLASS_C2_METABOLIC_EXCLUDED
CHIANG_LIVER_CANCER_SUBCLASS_POLYSOMY7
LI_2024_FATTY_ACID_DEGRADATION
AIZARANI_LIVER_C30_HEPATOCYTES_4
YANG_2020_SUBCLASS_C1_METABOLIC_HIGH
YANG_2020_SUBCLASS_C3_METABOLIC_INTERMEDIATE
CHIANG_LIVER_CANCER_SUBCLASS_UNANNOTATED
CHIANG_LIVER_CANCER_SUBCLASS_INTERFERON
HOSHIDA_2009_S3_SUBCLASS
DESERT_PERIPORTAL_HEPATOCELLULAR_CARCINOMA_SUBCLASS
HEPATOCYTE_EnrichR_PangloaDB
-0.4
-0.2
0.0
0.2
0.4
ssGSEA Scores
C
HOSHIDA_2009_SUBCLASS_S2
SETDB1_HIGH vs SETDB1_LOW ROC
True positive rate
False positive rate
AUC: 0.76822
MCC: 0.44339
ssGSEA Scores
Enrichment Score (ssGSEA)
SETDB1 HIGH SETDB1 LOW
Wilcox pValue : 3.32e-19
CHIANG_2009_SUBCLASS_PROLIFERATION
AUC: 0.63797
MCC: 0.27740
Wilcox pValue : 4.08-06
HUANG_2017_POOR_DIFFERENTIATED_HCC
AUC: 0.66975
MCC: 0.28730
Wilcox pValue : 4.91e-23
HOSHIDA_2009_SUBCLASSE_S3
AUC: 0.31773
MCC: -0.34944
Wilcox pValue : 1.15e-09
AIZARANI_LIVER_C30_HEPATOCYTES_4
AUC: 0.37662
MCC: -0.26095
Wilcox pValue : 3.79e-05
LI_2024_FATTY_ACID_DEGRADATION
AUC: 0.39514
MCC: -0.22670
Wilcox pValue : 0.000463
D
Percentage
1,5% 1%
2% 4,5%
27.5%
40%
51%
44.5%
18%
10%
nd
G4
G3
G2
G1
SETDB1 Low
SETDB1 High
E
Group
SETDB1.High
SETDB1.Low
Group
SETDB1
AFP
GPC3
EPCAM
4
2
0
-2
-4
F
SETDB1 High (SH)
SETDB1 Low (SL)
***
***
**
Expression level (log2 TPM)
GPC3
AFP
EPCAM

Fig. 5 SETDB1 expression is associated with aggressive molecular subtypes and poor differentiation in HCC. **(A)** Heatmap of single-sample Gene Set Enrichment Analysis (ssGSEA) enrichment scores for gene signatures associated with various molecular subclasses and differentiation states of HCC. The heatmap shows that tumors in the SETDB1-high group are primarily enriched for signatures related to aggressive subtypes and poorly differentiated profiles. **(B)** Bar plot of ssGSEA enrichment scores, showing the segregation of SETDB1-high (SH) and SETDB1-low (SL) tumors based on a selection of previously defined HCC molecular subclass signatures in the TCGA-LIHC cohort. Red bars indicate enrichment in the SH group, and blue bars indicate enrichment in the SL group. Saturated colors indicate significant gene sets ($P<0.05$, Wilcoxon test); attenuated colors are non-significant ($P>0.05$). **(C)** Receiver Operating Characteristic (ROC) curves and corresponding box plots illustrating the performance of representative signatures in distinguishing between the SH and SL groups. The Area Under the Curve (AUC) and Wilcoxon P-value are provided for each comparison. **(D)** Stacked bar plot showing the distribution of tumor differentiation grades (G1–G4) in the SH and SL groups. This demonstrates a higher proportion of poorly differentiated tumors (G3 and G4) in the SH group. **(E)** Heatmap illustrating the correlation between SETDB1 expression and the expression of three specific markers of poor differentiation and aggressiveness in HCC: AFP, GPC3, and EPCAM. **(F)** Violin plots comparing the expression levels of AFP, GPC3, and EPCAM in the SETDB1-high and SETDB1-low groups. Statistical significance was assessed using the Wilcoxon rank-sum test

SL group (Fig. 6H). The two groups exhibited contrasting miRNA profiles, with SL tumors enriched in miRNAs linked to differentiation and stemness inhibition (e.g., hsa-miR-612, hsa-miR-365, hsa-miR-126, hsa-miR-34a, hsa-miR-145, hsa-miR-29 family, hsa-miR-144, hsa-miR-203, hsa-miR-451, hsa-miR-122) [26, 30–32], whereas SH tumors showed enrichment of miRNAs promoting stemness (e.g., hsa-miR-196b, miR-429, miR-410, hsa-miR-3682, hsa-miR-181 family, hsa-miR-454) [25, 33, 34]. These findings further reinforce the association between high SETDB1 expression, stemness, and aggressive tumor behavior in HCC.

3.6 SETDB1 expression is predominantly localized in malignant tumor cells and correlates with stemness at the single-cell level

Single-cell transcriptomes from 9 HCC and 10 iCCA samples in dataset GSE125449 [24] were investigated using the Seurat R package (version 4.3.0) [35] in the R software environment (version 4.2.1). After prefiltration (genes expressed in over 100 cells and genes with minimal expression in at least 3 cells), a total of 9946 single-cell transcriptomes were normalized, scaled, and subjected to dimension reduction by tSNE method (Fig. 7A).

Expression of SETDB1 was investigated by tSNE (Fig. 7B) and volcano plot (Fig. 7C) visualization after stratification by LIHC subgroups and the cellular types present in tumors and their microenvironment. We observed that SETDB1 is expressed in different cell types in both subtypes of liver cancer (Fig. 7C, D). A more frequent expression of SETDB1 was found in malignant cells from both cancer types, with a slightly higher tendency in iCCA-derived cells compared to HCC ones. These results indicate that SETDB1 expression is primarily found in malignant tumor cells and is less prevalent in microenvironmental cellular subpopulations.

Single-cell transcriptomic analysis revealed distinct patterns of stemness marker gene expression across tumor and stromal cell populations in HCC and iCCA. Stemness-associated genes such as GDF15, ALDH1L1, GPC3, AFP, EPCAM, CD44, and CD24 were predominantly expressed in tumor cells, with varying levels across other cell types including CAFs, TECs, and TAMs (Fig. 7E). Consistently, the stem cell score (sc_score1), computed from the expression of these markers, was significantly higher in tumor cells compared to stromal and immune populations, indicating enrichment of stem-like transcriptional programs (Fig. 7F). Given this enrichment, we next investigated whether SETDB1 expression was associated with stemness at the single-cell level. Across all cell populations, SETDB1 expression showed only a weak association with stem cell score ($R^2 = 0.02$), reflecting *substantial* cellular heterogeneity and the absence of a global linear relationship (Fig. 7G). Importantly, when focusing specifically on HCC tumor cells, SETDB1-positive cells exhibited significantly higher stem cell scores compared to SETDB1-negative counterparts (Fig. 7H, Wilcoxon $p=0.044$), supporting the association between SETDB1 expression and enhanced stem-like properties in HCC liver cancer cells.

3.7 Hypoxia increases SETDB1 expression and stemness, while paclitaxel treatment leads to their suppression

Hypoxia is a well-known inducer of stemness in LIHC, playing a crucial role in driving metabolic shifts, cancer progression, and therapeutic resistance. Given this, we assessed the relationship between SETDB1 expression and hypoxia. We performed ssGSEA using curated hypoxia-related signatures (Buffa and Ragnum, validated across multiple cancer types [36]) in the TCGA-LIHC cohort. Our analysis showed that SH tumors displayed pronounced hypoxic features (Fig. 8A). Consistently, *SETDB1* expression correlated positively with *HIF1A* and *EPAS1* (HIF-2α) expression (Fig. 8B), reinforcing its association with hypoxia in HCC.

To validate these findings in vitro, we exposed Huh7 and HepG2 cells to hypoxic conditions (1% O2) for 3–48 h. Western blot revealed a marked upregulation of SETDB1 and HIF-1α protein expression in Huh7 cells after 6 h of hypoxia (Fig. 8C). Moreover, immunofluorescence showed weak nuclear SETDB1 staining in normoxia, which became

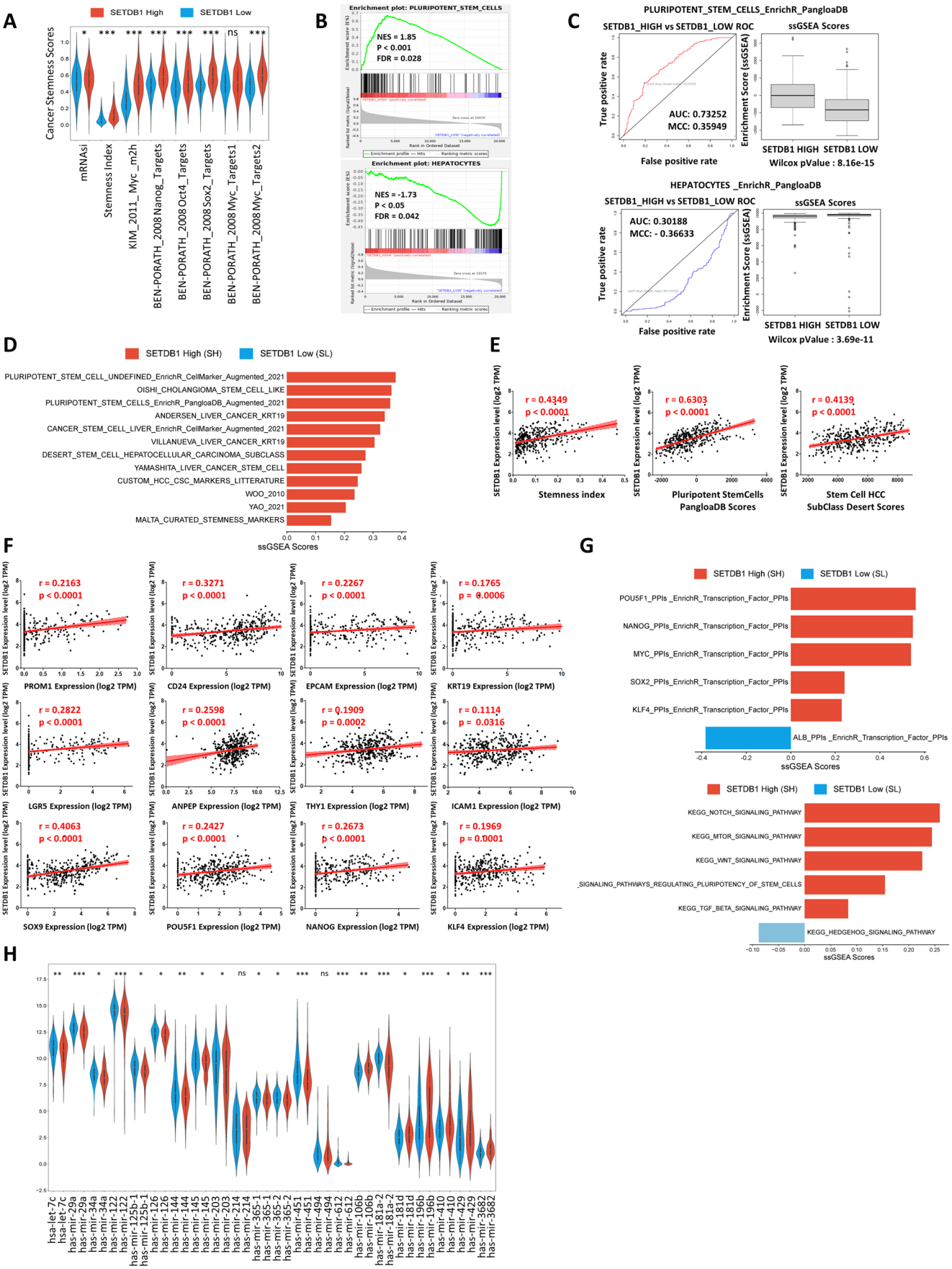

A
SETDB1 High
SETDB1 Low
Cancer Stemness Scores
mRNAsi
Stemness Index
KIM_2011_Myc_m2h
BEN-PORATH_2008 Nanog_Targets
BEN-PORATH_2008 Oct4_Targets
BEN-PORATH_2008 Sox2_Targets
BEN-PORATH_2008 Myc_Targets1
BEN-PORATH_2008 Myc_Targets2
B
Enrichment plot: PLURIPOTENT_STEM_CELLS
NES = 1.85
P < 0.001
FDR = 0.028
Enrichment plot: HEPATOCYTES
NES = -1.73
P < 0.05
FDR = 0.042
C
PLURIPOTENT_STEM_CELLS_EnrichR_PangloaDB
SETDB1_HIGH vs SETDB1_LOW ROC
AUC: 0.73252
MCC: 0.35949
True positive rate
False positive rate
ssGSEA Scores
Enrichment Score (ssGSEA)
SETDB1 HIGH SETDB1 LOW
Wilcox pValue : 8.16e-15
HEPATOCYTES _EnrichR_PangloaDB
SETDB1_HIGH vs SETDB1_LOW ROC
AUC: 0.30188
MCC: - 0.36633
Wilcox pValue : 3.69e-11
D
SETDB1 High (SH)
SETDB1 Low (SL)
PLURIPOTENT_STEM_CELL_UNDEFINED_EnrichR_CellMarker_Augmented_2021
OISHI_CHOLANGIOMA_STEM_CELL_LIKE
PLURIPOTENT_STEM_CELLS_EnrichR_PangloaDB_Augmented_2021
ANDERSEN_LIVER_CANCER_KRT19
CANCER_STEM_CELL_LIVER_EnrichR_CellMarker_Augmented_2021
VILLANUEVA_LIVER_CANCER_KRT19
DESERT_STEM_CELL_HEPATOCELLULAR_CARCINOMA_SUBCLASS
YAMASHITA_LIVER_CANCER_STEM_CELL
CUSTOM_HCC_CSC_MARKERS_LITTERATURE
WOO_2010
YAO_2021
MALTA_CURATED_STEMNESS_MARKERS
ssGSEA Scores
E
SETDB1 Expression level (log2 TPM)
r = 0.4349
p < 0.0001
Stemness index
r = 0.6303
p < 0.0001
Pluripotent StemCells PangloaDB Scores
r = 0.4139
p < 0.0001
Stem Cell HCC SubClass Desert Scores
F
r = 0.2163
p < 0.0001
PROM1 Expression (log2 TPM)
r = 0.3271
p < 0.0001
CD24 Expression (log2 TPM)
r = 0.2267
p < 0.0001
EPCAM Expression (log2 TPM)
r = 0.1765
p = 0.0006
KRT19 Expression (log2 TPM)
r = 0.2822
p < 0.0001
LGR5 Expression (log2 TPM)
r = 0.2598
p < 0.0001
ANPEP Expression (log2 TPM)
r = 0.1909
p = 0.0002
THY1 Expression (log2 TPM)
r = 0.1114
p = 0.0316
ICAM1 Expression (log2 TPM)
r = 0.4063
p < 0.0001
SOX9 Expression (log2 TPM)
r = 0.2427
p < 0.0001
POU5F1 Expression (log2 TPM)
r = 0.2673
p < 0.0001
NANOG Expression (log2 TPM)
r = 0.1969
p = 0.0001
KLF4 Expression (log2 TPM)
G
SETDB1 High (SH)
SETDB1 Low (SL)
POU5F1_PPIs _EnrichR_Transcription_Factor_PPIs
NANOG_PPIs_EnrichR_Transcription_Factor_PPIs
MYC_PPIs_EnrichR_Transcription_Factor_PPIs
SOX2_PPIs_EnrichR_Transcription_Factor_PPIs
KLF4_PPIs_EnrichR_Transcription_Factor_PPIs
ALB_PPIs _EnrichR_Transcription_Factor_PPIs
ssGSEA Scores
KEGG_NOTCH_SIGNALING_PATHWAY
KEGG_MTOR_SIGNALING_PATHWAY
KEGG_WNT_SIGNALING_PATHWAY
_SIGNALING_PATHWAYS_REGULATING_PLURIPOTENCY_OF_STEM_CELLS
KEGG_TGF_BETA_SIGNALING_PATHWAY
KEGG_HEDGEHOG_SIGNALING_PATHWAY
ssGSEA Scores
H
hsa-let-7c
hsa-let-7c
has-mir-29a
has-mir-29a
has-mir-34a
has-mir-34a
has-mir-122
has-mir-122
has-mir-125b-1
has-mir-125b-1
has-mir-126
has-mir-126
has-mir-144
has-mir-144
has-mir-145
has-mir-145
has-mir-203
has-mir-203
has-mir-214
has-mir-214
has-mir-365-1
has-mir-365-1
has-mir-365-2
has-mir-365-2
has-mir-451
has-mir-451
has-mir-494
has-mir-494
has-mir-612
has-mir-612
has-mir-106b
has-mir-106b
has-mir-181a-2
has-mir-181a-2
has-mir-181d
has-mir-181d
has-mir-196b
has-mir-196b
has-mir-410
has-mir-410
has-mir-429
has-mir-429
has-mir-3682
has-mir-3682

Fig. 6 SETDB1 expression is associated with stemness features and a poorly differentiated phenotype in the TCGA-LIHC cohort. **(A)** Violin plots comparing cancer stemness indices (mRNAsi, Stemness Index) between the SETDB1-high (SH) and SETDB1-low (SL) groups. The indices were calculated using the one-class logistic regression model. Statistical significance is indicated: *$p<0.05$, **$p<0.01$, ***$p<0.001$, ns, no significance. **(B)** Gene Set Enrichment Analysis (GSEA) plots demonstrating the enrichment of a pluripotent stem cell signature in the SH group and a hepatocyte signature in the SL group. The Normalized Enrichment Score (NES) and False Discovery Rate (FDR) q-value are shown. **(C)** Receiver Operating Characteristic (ROC) curves and corresponding box plots illustrating the performance of the pluripotent stem cell and hepatocyte signatures in distinguishing between SH and SL groups. The Area Under the Curve (AUC) and Wilcoxon P-value for each comparison are provided. **(D)** Bar plot of single-sample GSEA (ssGSEA) scores showing the enrichment of various stemness-related signatures in the SH (red) and SL (blue) groups. Significant enrichment ($P<0.05$, Wilcoxon test) is indicated by saturated colors. **(E)** Scatter plots illustrating the Spearman's correlation between SETDB1 expression and three different stemness scores (mRNAsi, PanglaoDB pluripotency score, and Désert STEM score). The correlation coefficient (r) and P-value are shown for each plot. **(F)** Scatter plots showing the Spearman's correlation between SETDB1 expression and the expression of ten genes from a custom HCC cancer stem cell (CSC) gene signature. The correlation coefficient (r) and P-value are provided. **(G)** Bar plot of ssGSEA scores for developmental and stemness-related signaling pathways (top) and protein-protein interaction (PPI) modules (bottom). This panel highlights the enrichment of key developmental pathways and pluripotency-related transcription factor modules in the SH group. **(H)** Violin plots comparing the differential expression of selected microRNAs between the SH and SL groups

markedly stronger after 6 h of hypoxia (Fig. 8D), further reinforcing the role of hypoxia in modulating SETDB1 expression. Building on these results, qPCR demonstrated elevated *SETDB1* mRNA along with pluripotency markers (*NANOG*, *POU5F1*, *SOX2*, *KLF4*, *CD24*, and *PROM1*) in Huh7 cells after 24 h of hypoxia (Fig. 8E). Flow cytometry analysis confirmed increased stem-like markers (CD133, CD24, and CD44) in Huh7 cells after 3 days of hypoxia (Fig. 8F). To explore the functional implications of these stemness traits, we assessed the sensitivity of Huh7 cells to Sorafenib. Sorafenib sensitivity assays showed an increased IC50 from 4.88 µM (normoxia) to 5.47 µM (3 days hypoxia) (Fig. 8G), suggesting enhanced therapeutic resistance.

Furthermore, to evaluate the relationship of SETDB1 levels and stemness gene expression, we used paclitaxel (taxol), a described SETDB1 inhibitor [37]. Western blot revealed a marked downregulation of SETDB1 upon paclitaxel treatment at both 50 nM and 500 nM (Fig. 8H). This decrease in SETDB1 inhibition was correlated with reduced mRNA levels of three key stemness transcription factors (*NANOG*, *SOX2*, *and POU5F1*) (Fig. 8I). At the protein level, NANOG, and to a lesser extent OCT4, was also decreased under paclitaxel at 50 nM and 500 nM. Experiments performed with HepG2 cells were concordant with those in Huh7cells, showing that hypoxia increased both SETDB1 and stemness markers, whereas paclitaxel treatment was associated with their reduction (Fig. S2A-D).

Collectively, these results support a strong association between SETDB1 and stemness traits under hypoxia, including elevated expression of stem cell markers and reduced drug sensitivity. Conversely, lower SETDB1 levels were observed alongside a less pronounced stem-like phenotype, indicating a close correlation between SETDB1 and stemness features.

3.8 SETDB1 expression and its chromatin binding partners are associated with poor disease-free survival in liver cancer

To investigate the prognostic role of SETDB1 and its partners in liver cancer, their PPI network was analyzed using the STRING server. The resulting network (Fig. 9A) contained eleven molecules, including SETDB1, and its functional enrichment revealed a significant enrichment for chromatin binding (FDR $p=1.8\times10^{-5}$).

Prognostic analyses identified three markers with a significant positive association with poor disease-free survival (DFS) (univariate $p<0.05$): DNMT3A (hazard ratio = 1.23, $p=0.015$), TRIM28 (hazard ratio = 1.30, $p=0.024$) and HDAC1 (hazard ratio = 1.32, $p=0.043$). The SETDB1 hazard ratio was 1.28 ($p=0.074$) (Fig. 9B, Table S4). These four prognostic markers are all involved in chromatin binding.

A risk score was then computed for the TCGA liver cancer cohort using the expression of these four genes. Patients were stratified into high- and low-risk groups for DFS using an optimal threshold (cutpoint: 5.8). A significant stratification of patients based on this risk score was confirmed by Kaplan-Meier survival analysis and a log-rank test (univariate $p=4\times10^{-4}$) (Fig. 9C).

A clear separation of the high- and low-risk groups was also demonstrated by PCA based on the expression of the four binders (Fig. 9D). Interestingly, samples were also stratified by pathological grade in this same PCA (Fig. 9E). Furthermore, high-risk samples with high-grade values were successfully aggregated by unsupervised clustering of these four chromatin binders' expression profiles (Fig. 9F). No association was observed between the clustering and other clinical risk factors (e.g., vascular invasion, hepatitis B or C, NAFLD, or alcohol consumption) (Fig. 9G).

To assess the independence of the SETDB1-related risk score, a multivariable Cox model for DFS was constructed. The risk score, demographic parameters (age, gender, weight), grade, and risk factors were included in this model. The model was significant (log-rank $p=0.001$, Fig. 10A), and the SETDB1-related risk score was confirmed as an independent unfavorable prognostic factor for DFS (multivariable hazard ratio = 1.78, $p=0.029$).

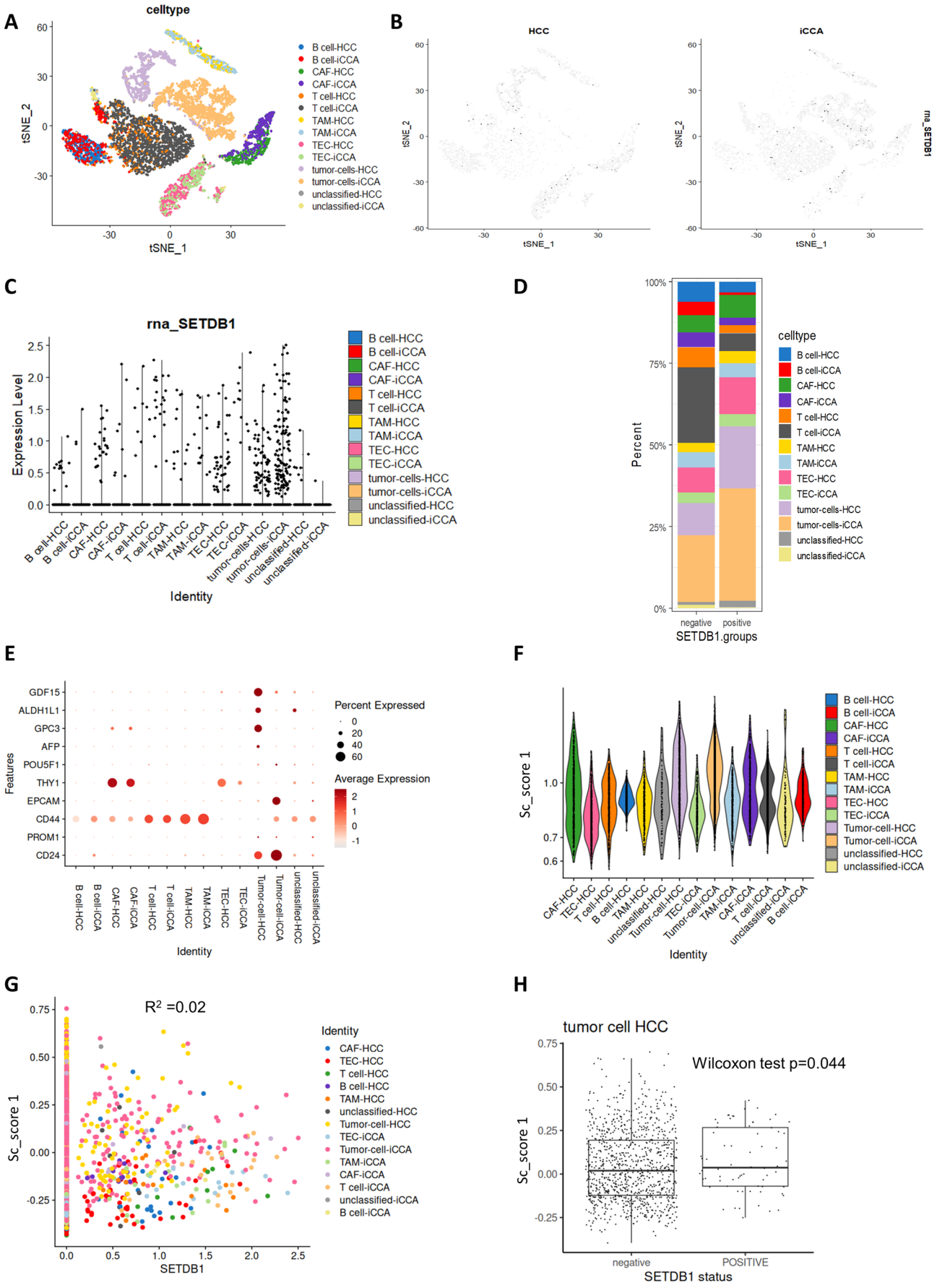

A
celltype
tSNE_2
tSNE_1
B cell-HCC
B cell-iCCA
CAF-HCC
CAF-iCCA
T cell-HCC
T cell-iCCA
TAM-HCC
TAM-iCCA
TEC-HCC
TEC-iCCA
tumor-cells-HCC
tumor-cells-iCCA
unclassified-HCC
unclassified-iCCA
B
HCC
iCCA
rna_SETDB1
C
rna_SETDB1
Expression Level
Identity
D
Percent
SETDB1.groups
negative
positive
E
Features
GDF15
ALDH1L1
GPC3
AFP
POU5F1
THY1
EPCAM
CD44
PROM1
CD24
Percent Expressed
Average Expression
Identity
F
Sc_score 1
Identity
G
R² =0.02
Sc_score 1
SETDB1
Identity
H
tumor cell HCC
Wilcoxon test p=0.044
Sc_score 1
negative
POSITIVE
SETDB1 status

Fig. 7 SETDB1 expression is predominantly localized in malignant tumor cells and correlates with stemness at the single-cell level in liver cancer. **(A)** t-SNE dimensional reduction plot of single-cell transcriptomes from liver cancer samples, showing different cell types from HCC (hepatocellular carcinoma) and iCCA (intrahepatic cholangiocarcinoma). Each cluster represents a distinct cell type, as indicated by the color legend. **(B)** t-SNE plots highlighting the expression levels of SETDB1 in HCC and iCCA samples. Darker points correspond to higher SETDB1 expression. This visualization shows that SETDB1 expression is more prominent in tumor cells from both cancer types, with a slightly higher tendency in iCCA tumor cells compared to HCC tumor cells. **(C)** Dot plot showing the expression levels of rna_SETDB1 across different cell types and cancer subtypes. Each dot represents a single cell, and the expression level is plotted on the Y-axis. This figure indicates that SETDB1 is expressed in various cell types but is most frequently expressed in malignant cells (tumor-cells-HCC and tumor-cells-iCCA). **(D)** Stacked bar plot comparing the percentage of different cell types in groups with high (positive) and low (negative) SETDB1 expression. This visualization confirms that the "positive" SETDB1 group is primarily composed of tumor cells. **(E)** Dot plot showing the expression of stemness marker genes (GDF15, ALDH1L1, GPC3, AFP, EPCAM, CD44, and CD24) across diverse cell populations, including tumor cells, cancer-associated fibroblasts (CAFs), tumor endothelial cells (TECs), tumor-associated macrophages (TAMs), B cells, T cells, and unclassified populations. Dot size represents the percentage of cells expressing each gene, while color intensity reflects the average expression level. **(F)** Violin plots depicting the distribution of stem cell score (sc_score1) across different cell identities, showing enrichment in tumor cell populations. **(G)** Scatter plot illustrating the positive correlation between SETDB1 expression and stem cell score (sc_score1) in HCC tumor cells. Each point represents an individual cell, with color indicating cell identity. **(H)** Box plot comparing stem cell scores (STEM/CELL_score) between SETDB1-positive and SETDB1-negative HCC tumor cells. SETDB1-positive cells exhibit significantly higher stemness scores (Wilcoxon test, $p = 0.044$)

The model's reliability was then validated through bootstrapping (500 iterations) to predict the probability of DFS at two and four years (Fig. 10B, C). Both calibrations showed that the model has good predictive power. The nomogram highlighted the significant weight of the SETDB1 risk score in predicting patient outcomes (Fig. 10D). These findings suggest that SETDB1 and its chromatin binding partners (DNMT3A, HDAC1, and TRIM28) are deregulated during liver cancer progression. Their combined expression serves as an independent prognostic factor for poor disease-free survival, indicating their potential as a valuable biomarker for patient stratification.

4 Discussion

HCC remains a major global health challenge due to its increasing incidence, high mortality, and a persistent lack of effective long-term therapeutic strategies. Despite advances in locoregional and systemic treatments, including TACE, tyrosine kinase inhibitors (TKIs), and immune checkpoint inhibitors (ICIs), durable responses remain rare, in part because of the profound intratumoral heterogeneity and the persistence of CSCs [38]. These CSCs possess remarkable adaptive capabilities, enabling them to withstand therapeutic pressures through mechanisms such as enhanced DNA repair, ABC transporter activation, metabolic rewiring, and suppression of apoptosis. They further contribute to an immunosuppressive tumor microenvironment via intrinsic mechanisms (PD-L1 upregulation, TGF-β secretion) and extrinsic modulation (recruitment of regulatory T cells, M2 macrophages) [39], ultimately sustaining tumor survival, recurrence, and metastasis. Therefore, identifying upstream regulators of cancer stemness and effectively targeting CSCs are critical priorities for improving patient outcomes. Among the epigenetic regulators implicated in HCC, the histone methyltransferase SETDB1 emerges as one of the most upregulated and functionally relevant [40–43]. Beyond its well-established role as a transcriptional repressor implicated in tumorigenesis and immune evasion [40], the specific contribution of SETDB1 to liver CSC regulation had remained largely unexplored. Through integrative multiomic analyses with hallmark, KEGG, and curated stemness gene signatures, we highlight SETDB1 as a defining feature of a stemness-enriched, poorly differentiated, clinically aggressive HCC phenotype. Single-cell RNA-seq analyses revealed that SETDB1 expression is largely confined to malignant epithelial cells, with marked enrichment in progenitor-like subsets ($EPCAM^+$, $PROM1^+$, $ALDH1A1^+$, $KRT19^+$, $CD24^+$) previously linked to hepatic progenitor cells and CSC-like features [24]. Consistent with these findings, bulk TCGA-LIHC tumors stratified by SETDB1 expression revealed a clear dichotomy. SETDB1-low (SL) tumors retained hepatocytic differentiation, exhibited low proliferative indices, and were enriched in favorable-prognosis molecular subclasses, whereas SETDB1-high (SH) tumors aligned with aggressive subclasses characterized by strong MYC activation, cell-cycle progression, TGF-β signaling, and a marked loss of hepatocytic identity.

In addition, SETDB1-high tumors displayed a significantly higher burden of TP53 mutations, most notably the R249S hotspot. Mutant TP53 has been shown to enhance SETDB1 transcription and protein stability [44], fostering cell cycle progression and apoptosis evasion [45], a feedback loop that may fuel aggressive, therapy-resistant clones. The tumor microenvironment also contributes to SETDB1 accumulation, particularly through hypoxia, a well-established driver of CSC maintenance. Acting via hypoxia-inducible factors (HIFs), metabolic rewiring, and epigenetic regulation, hypoxia sustains CSC properties and promotes tumor progression, therapy resistance, and recurrence. Our transcriptomic analyses and in vitro data demonstrate that hypoxic stress markedly increases SETDB1 expression. Mechanistic experiments have shown that,

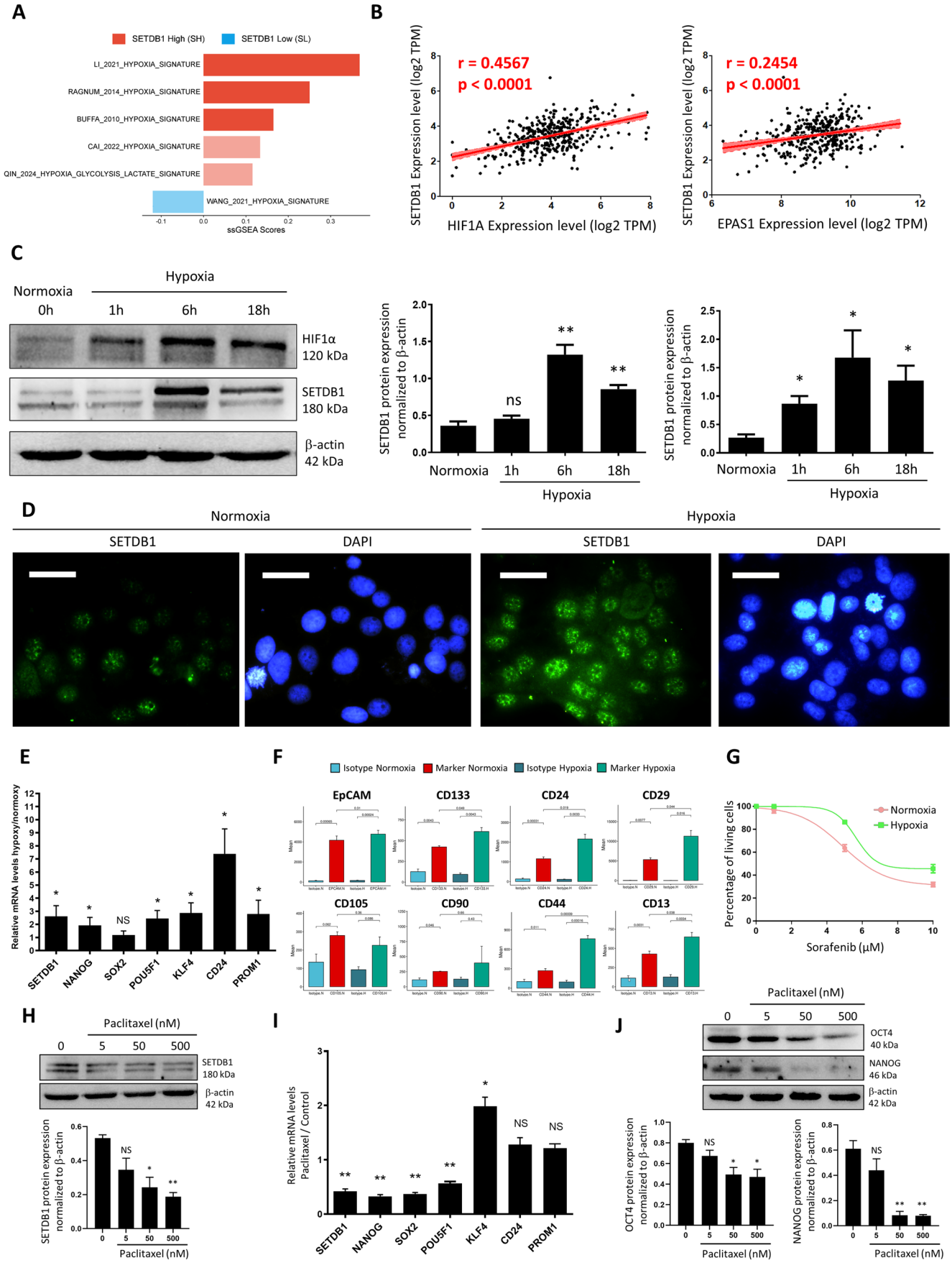

A
SETDB1 High (SH)
SETDB1 Low (SL)
LI_2021_HYPOXIA_SIGNATURE
RAGNUM_2014_HYPOXIA_SIGNATURE
BUFFA_2010_HYPOXIA_SIGNATURE
CAI_2022_HYPOXIA_SIGNATURE
QIN_2024_HYPOXIA_GLYCOLYSIS_LACTATE_SIGNATURE
WANG_2021_HYPOXIA_SIGNATURE
ssGSEA Scores
B
r = 0.4567
p < 0.0001
SETDB1 Expression level (log2 TPM)
HIF1A Expression level (log2 TPM)
r = 0.2454
p < 0.0001
EPAS1 Expression level (log2 TPM)
C
Normoxia
Hypoxia
0h
1h
6h
18h
HIF1α
120 kDa
SETDB1
180 kDa
β-actin
42 kDa
SETDB1 protein expression normalized to β-actin
ns
Normoxia
Hypoxia
D
Normoxia
Hypoxia
SETDB1
DAPI
E
Relative mRNA levels hypoxy/normoxy
NS
SETDB1
NANOG
SOX2
POU5F1
KLF4
CD24
PROM1
F
Isotype Normoxia
Marker Normoxia
Isotype Hypoxia
Marker Hypoxia
EpCAM
CD133
CD24
CD29
CD105
CD90
CD44
CD13
G
Normoxia
Hypoxia
Percentage of living cells
Sorafenib (μM)
H
Paclitaxel (nM)
0
5
50
500
SETDB1
180 kDa
β-actin
42 kDa
SETDB1 protein expression normalized to β-actin
NS
Paclitaxel (nM)
I
Relative mRNA levels
Paclitaxel / Control
NS
SETDB1
NANOG
SOX2
POU5F1
KLF4
CD24
PROM1
J
Paclitaxel (nM)
OCT4
40 kDa
NANOG
46 kDa
β-actin
42 kDa
OCT4 protein expression normalized to β-actin
NANOG protein expression normalized to β-actin
NS
Paclitaxel (nM)

Fig. 8 SETDB1 expression mirrors stemness dynamics in HuH7WT cells under hypoxia and paclitaxel treatment. **(A)** ssGSEA scores of hypoxia-related gene signatures in SH and SL groups. The bar graph shows the enrichment of hypoxia-related gene sets. Red bars indicate enrichment in the SH group, and blue bars indicate enrichment in the SL group. Saturated colors indicate significant gene sets ($P<0.05$, Wilcoxon test); attenuated colors are non-significant ($P>0.05$). **(B)** Scatter plots showing Spearman's correlation between SETDB1 expression and hypoxia-inducible factor genes (left, HIF1A and right, EPAS1, HIF-2α) in the TCGA-LIHC cohort. **(C)** Western blot analysis of SETDB1 and HIF-1α protein expression under hypoxic conditions. The blot and corresponding bar graph show the expression of SETDB1 and HIF-1α in Huh7 cells after 6 h of hypoxia compared to normoxia. β-actin served as the internal loading control. Data are represented as means ± SE, $n=3$. Statistical significance is indicated as follows: $*P<0.05$, $**P<0.01$, $***P<0.001$, $****P<0.0001$. **(D)** Immunofluorescence analysis of SETDB1 localization under hypoxic conditions. The images show nuclear (DAPI, blue) and cytoplasmic/nuclear (SETDB1, green) staining of SETDB1 in Huh7 cells under normoxic and 6-hour hypoxic conditions. Data are represented as means ± SE, $n=3$. Scale bar = 200 µm. **(E)** qPCR analysis of stemness markers in Huh7 cells after 24-hour hypoxia, showing relative mRNA levels of *NANOG*, *SOX2*, *POU5F1*, *KLF4*, *CD24*, and *PROM1* normalized to 18 S and expressed versus normoxia. **(F)** Flow cytometry of stemness-associated surface markers (CD133, CD24, CD44) in Huh7 cells after 3 days of hypoxia versus normoxia. **(G)** Sorafenib sensitivity assay under hypoxic conditions. The graph shows the dose-response curve of sorafenib in Huh7 cells cultured under normoxic and 3-day hypoxic conditions, highlighting the change in IC50. **(H)** Western blot of SETDB1 protein expression SETDB1 in Huh7 cells treated with increasing concentrations of paclitaxel (5-500 nM, 48 h). The blot and bar graph depict SETDB1 protein levels, with β-actin as loading control. **(I)** qPCR analysis of stemness markers in Huh7 cells treated with paclitaxel (500 nM, 24 h), showing relative mRNA levels normalized to 18 S and expressed versus control. **(J)** Western blot of OCT4, and NANOG protein expression SETDB1 in Huh7 cells treated with increasing concentrations of paclitaxel (5-500 nM, 48 h). The blot and bar graph depict SETDB1 protein levels, with β-actin as loading control

under low-oxygen conditions, SETDB1 evades $CRL2^{VHL}$-mediated proteasomal degradation, leading to its stabilization, as the protein escapes the oxygen-dependent activity of this E3 ubiquitin ligase complex [46]. Adding to these layers of regulation, SETDB1 is also integrated into an oncogenic feed-forward loop with MYC, wherein MYC promotes SETDB1 expression and SETDB1 reinforces MYC-driven transcriptional programs [47]. This feedback architecture may help stabilize a dedifferentiated, stemness-dominated cellular state that confers high proliferative potential and therapeutic resistance. Collectively, the convergence of TP53 mutations, hypoxia-driven stabilization, and MYC-mediated transcriptional reinforcement provides a compelling explanation for why SETDB1 is selectively overexpressed in the most aggressive HCC subtypes.

Beyond the upstream factors driving SETDB1 overexpression, its functional impact on CSC regulation is mediated through several interconnected epigenetic and transcriptional mechanisms. The strong correlation between SETDB1 expression and multiple independent stemness signatures suggests that SETDB1 is embedded within a broader pluripotency-associated regulatory network. This is consistent with evidence from embryonic stem cells demonstrating that SETDB1 is essential for maintaining pluripotency and early embryonic viability. Indeed, SETDB1-deficient embryos exhibit peri-implantation lethality, with developmental arrest around E4.5 [11], and SETDB1 depletion in vitro prevents the formation of stable ESC colonies [11]. Mechanistically, SETDB1 sustains pluripotency through two complementary modes of action. First, it physically interacts with OCT4, and loss of SETDB1 reduces OCT4 expression [48, 49], alters colony morphology, and triggers differentiation programs. Second, SETDB1 deposits the repressive H3K9me3 mark on differentiation-associated genes, thereby silencing their expression and indirectly stabilizing the pluripotent state [50]. For example, Wu et al. demonstrated that SETDB1 represses the totipotency factor Dux, preventing activation of 2 C-like programs and ensuring the transition from totipotency to pluripotency [51]. Together, these mechanisms highlight a conserved role for SETDB1 in maintaining stemness, a function that CSCs appear to exploit in oncogenic contexts. SETDB1 also supports CSC traits through its interaction with PML nuclear bodies (PML-NBs), non-membranous nuclear structures that coordinate chromatin modifiers, transcriptional regulators, and DNA repair factors essential for chromatin organization and CSC maintenance. By stabilizing PML-NBs, SETDB1 preserves the nuclear scaffold required for assembling repressive chromatin domains, whereas its inhibition disrupts these compartments [52]. Given that PML degradation reduces OCT4 levels and depletes CSCs [53, 54], SETDB1 may promote stemness not only through H3K9me3-mediated repression but also by stabilizing PML-NB–dependent nuclear architectures that preserve stem cell identity.

Among the targets silenced through these combined mechanisms, SETDB1 represses endogenous retroelements (ERVs) under hypoxic conditions, preventing the accumulation of transposon-derived dsRNAs that would otherwise activate innate immune pathways and induce lethal DNA damage [55]. By maintaining ERV silencing, SETDB1 acts as a genome-protective factor that enables CSCs to withstand oxygen deprivation, metabolic stress, and immune pressure. This transposon-silencing function, previously shown to be essential for preserving stem cell identity and genomic stability [56], likely contributes to CSC maintenance in hostile microenvironmental niches.

SETDB1 further contributes to HCC heterogeneity by modulating miRNA programs that govern stemness and differentiation [57]. SH tumors show upregulation of oncogenic miRNAs (e.g., miR-196b, miR-181 family, miR-454-3p) that promote CSC features and resistance [25, 33,

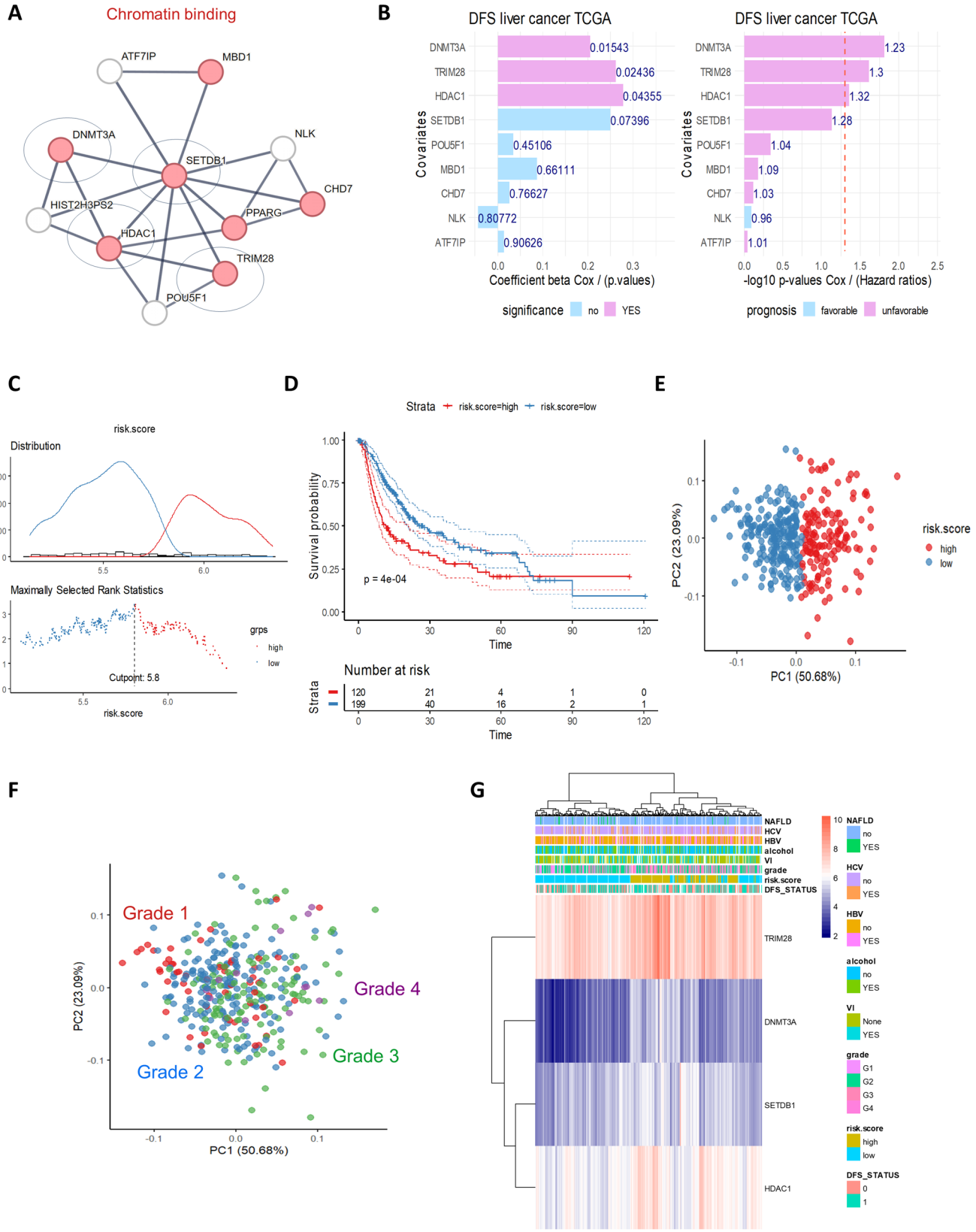

A
Chromatin binding
ATF7IP
MBD1
DNMT3A
NLK
SETDB1
CHD7
HIST2H3PS2
PPARG
HDAC1
TRIM28
POU5F1
B
DFS liver cancer TCGA
DNMT3A 0.01543
TRIM28 0.02436
HDAC1 0.04355
SETDB1 0.07396
POU5F1 0.45106
MBD1 0.66111
CHD7 0.76627
NLK 0.80772
ATF7IP 0.90626
Covariates
Coefficient beta Cox / (p.values)
significance no YES
DFS liver cancer TCGA
DNMT3A 1.23
TRIM28 1.3
HDAC1 1.32
SETDB1 1.28
POU5F1 1.04
MBD1 1.09
CHD7 1.03
NLK 0.96
ATF7IP 1.01
-log10 p-values Cox / (Hazard ratios)
prognosis favorable unfavorable
C
risk.score
Distribution
Density
Maximally Selected Rank Statistics
Standardized Log-Rank Statistic
Cutpoint: 5.8
grps high low
risk.score
D
Strata risk.score=high risk.score=low
Survival probability
p = 4e-04
Time
Number at risk
Strata
120 21 4 1 0
199 40 16 2 1
E
PC2 (23.09%)
PC1 (50.68%)
risk.score high low
F
Grade 1
Grade 2
Grade 3
Grade 4
PC2 (23.09%)
PC1 (50.68%)
G
NAFLD
HCV
HBV
alcohol
VI
grade
risk.score
DFS_STATUS
TRIM28
DNMT3A
SETDB1
HDAC1
NAFLD no YES
HCV no YES
HBV no YES
alcohol no YES
VI None YES
grade G1 G2 G3 G4
risk.score high low
DFS_STATUS 0 1

Fig. 9 A SETDB1-Related Expression Risk Score Stratifies HCC Patients in the TCGA Liver Cancer Cohort. **(A)** Protein-protein interaction (PPI) network of SETDB1 and its partners. The network, analyzed using the STRING server, highlights SETDB1's chromatin-binding partners (red nodes) used to calculate the expression risk score. **(B)** Univariate Cox analyses for disease-free survival (DFS). The forest plot shows the coefficient beta and -log10 p-values for each marker, highlighting partners with a significant correlation to prognosis. The dashed red line indicates a hazard ratio of 1.0. **(C)** Risk score distribution and optimal cutpoint determination. The upper panel shows the distribution of the risk score, while the lower panel shows the determination of the optimal cutoff threshold using standardized log-rank statistics, which stratifies patients into low- and high-risk groups. **(D)** Kaplan-Meier survival curves for DFS. The plot shows DFS stratified by the low- and high-risk score categories. The table below the graph indicates the number of patients at risk over time. **(E)** Principal component analysis (PCA) plot stratified by risk score. The plot is based on the expression of SETDB1 and its partners (*TRIM28*, *HDAC1*, and *DNMT3A*). Patients are colored according to their risk score (low risk in blue, high risk in red), showing a clear separation of the groups. **(F)** PCA plot stratified by pathological grade. The same patient cohort is stratified by pathological grade (Grade 1 to Grade 4), indicating that the expression of these markers correlates with tumor differentiation. **(G)** Unsupervised hierarchical clustering and heatmap analysis. The heatmap displays the expression levels of the SETDB1-related risk score markers. The top annotation bar includes key clinical variables, such as viral status (HCV, HBV), vascular invasion (VI), and pathological grade, showing that the clustering of high-risk samples corresponds to high-grade tumors

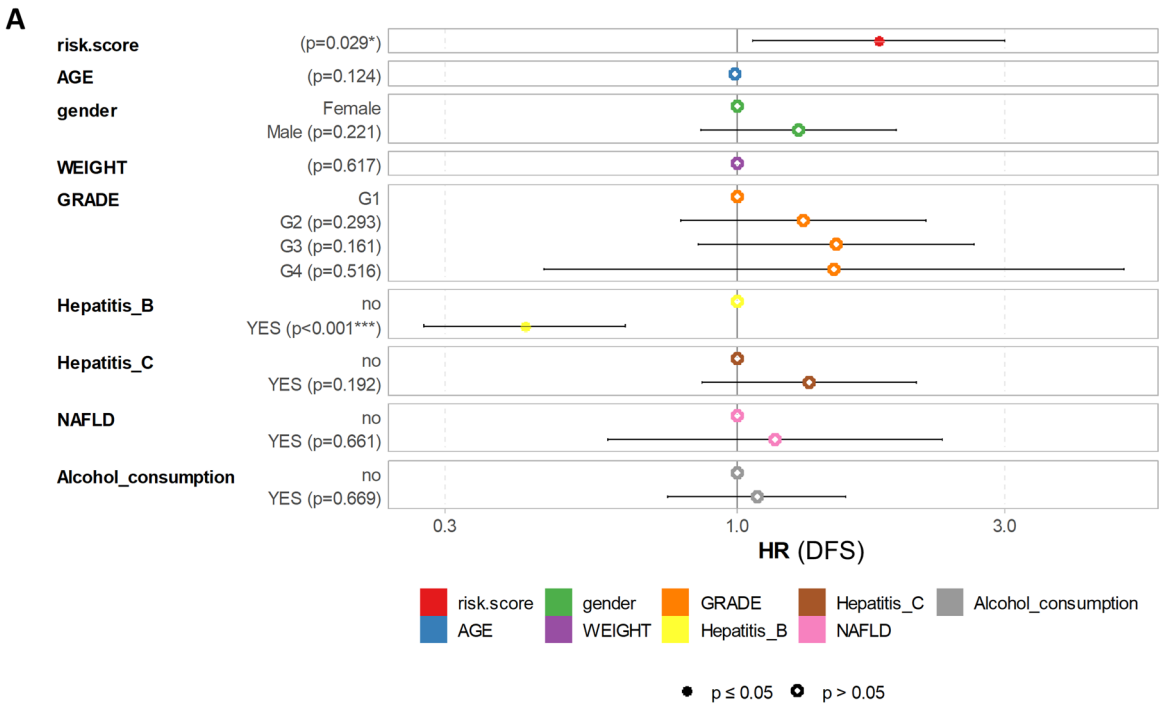
34], while tumor-suppressive miRNAs (e.g., miR-29a/b, miR-122) are downregulated [26, 30, 31]. This dual regulation suggests that SETDB1 reinforces tumor plasticity not only through chromatin remodeling but also by orchestrating a pro-oncogenic miRNA environment. This is further supported by the downregulation of miRNAs that directly target the SETDB1 3′-UTR in SH tumors [43, 58], contributing to SETDB1 overexpression, CSC enrichment, and therapeutic resistance.

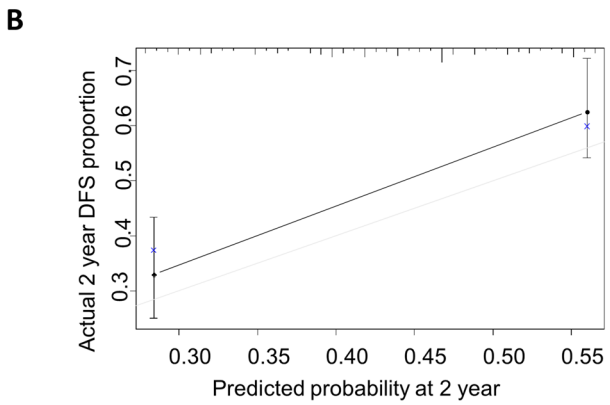
SETDB1 may also contribute to the regulation of CSCs through EMT. Several studies have shown that SETDB1 promotes EMT [59, 60], tumor progression, and the acquisition of CSC traits [61, 62] in solid cancers, including colorectal and breast cancer. Conversely, SETDB1 can oppose TGF-β–induced EMT by repressing SNAI1, as reported in breast epithelial cells [62] and pulmonary fibrosis models [63]. Together, these observations highlight the context-dependent dual role of SETDB1, which can either promote or inhibit EMT. This duality underscores the complexity of the relationships between EMT and CSCs. While early studies proposed a direct and linear association between EMT activation and CSC acquisition [64], more recent data have refined this paradigm. CSCs are now understood as dynamic states shaped by epithelial–mesenchymal plasticity (EMP), with stemness properties confined to specific "stemness windows" often associated with hybrid E/M states [65, 66]. These hybrid states exhibit maximal plasticity, enhanced metastatic potential, and superior tumor-initiating capacity compared to cells locked in purely epithelial or fully mesenchymal states [67, 68]. The positioning of these windows varies across cellular models, ranging from configurations close to the epithelial pole to intermediate hybrid E/M states or more advanced mesenchymal states [65]. In this context, the absence of association between SETDB1 expression and EMT signatures observed in the TCGA HCC cohort suggests that the relevant window for CSC regulation by SETDB1 likely lies closer to the epithelial side.

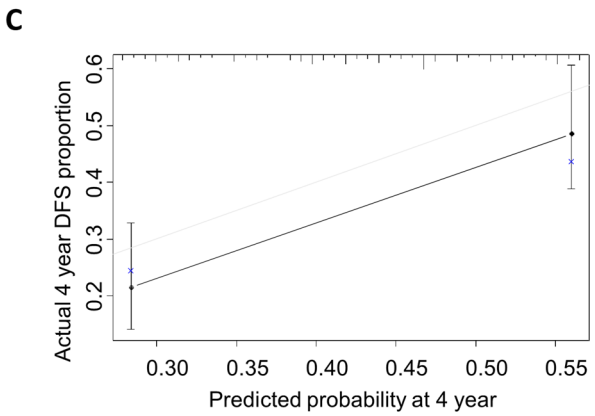
A major clinical implication of our findings is the strong association between SETDB1 overexpression and resistance to a broad range of HCC therapies, including immunotherapy, TACE, and TKIs. SL tumors display a "hot tumor" immune profile, characterized by higher infiltration of cytotoxic T and NK cells and increased immunogenic signaling, suggesting a greater likelihood of response to immune checkpoint inhibitors (ICIs). In contrast, SH tumors exhibit an "immune-excluded" phenotype with low infiltration of cytotoxic T and NK cells, and enrichment in immunosuppressive pathways. Mechanistically, SETDB1 contributes to this resistance by epigenetically repressing ERVs and interferon-stimulated genes, thereby dampening type I interferon signaling and reducing tumor immunogenicity [16, 55]. Moreover, SETDB1-high tumors are enriched for gene signatures associated with TACE non-response and sorafenib resistance. These findings align with previous studies [69, 70] and are corroborated by our analysis of the GSE104580 cohort, showing elevated SETDB1 levels in TACE non-responders, thus reinforcing its potential as a resistance biomarker. Importantly, the CSC compartment is a recognized reservoir for therapeutic escape. Our data, together with specific gene signatures, suggest that SETDB1 contributes to this phenotype by modulating drug transporters and anti-apoptotic pathways [27, 28], thereby sustaining survival under therapeutic pressure. Furthermore, SH tumors are specifically enriched in ferroptosis-resistant gene programs, suggesting that SETDB1 may suppress ferroptosis-related pathways, a key mechanism of therapeutic failure in HCC.

Taken together, our data showing the strong correlation between SETDB1 expression, stemness features, and tumor aggressiveness highlight SETDB1 as a compelling therapeutic target in HCC. Our chromatin-modifier–based risk score (SETDB1, DNMT3A, TRIM28, HDAC1) shows robust prognostic performance and may help stratify patients for future epigenetic-based interventions. However, selective SETDB1 inhibitors are still in early development. In the absence of direct SETDB1 inhibitors, currently available therapies that modulate SETDB1 indirectly become particularly relevant.

In this context, our experiments with paclitaxel provide important, albeit preliminary, insights. We observed


A
risk.score (p=0.029*)
AGE (p=0.124)
gender
Female
Male (p=0.221)
WEIGHT (p=0.617)
GRADE
G1
G2 (p=0.293)
G3 (p=0.161)
G4 (p=0.516)
Hepatitis_B
no
YES (p<0.001***)
Hepatitis_C
no
YES (p=0.192)
NAFLD
no
YES (p=0.661)
Alcohol_consumption
no
YES (p=0.669)
0.3
1.0
3.0
HR (DFS)
risk.score
AGE
gender
WEIGHT
GRADE
Hepatitis_B
Hepatitis_C
NAFLD
Alcohol_consumption
p ≤ 0.05
p > 0.05


B
Actual 2 year DFS proportion
0.3
0.4
0.5
0.6
0.7
0.30
0.35
0.40
0.45
0.50
0.55
Predicted probability at 2 year


C
Actual 4 year DFS proportion
0.2
0.3
0.4
0.5
0.6
0.30
0.35
0.40
0.45
0.50
0.55
Predicted probability at 4 year

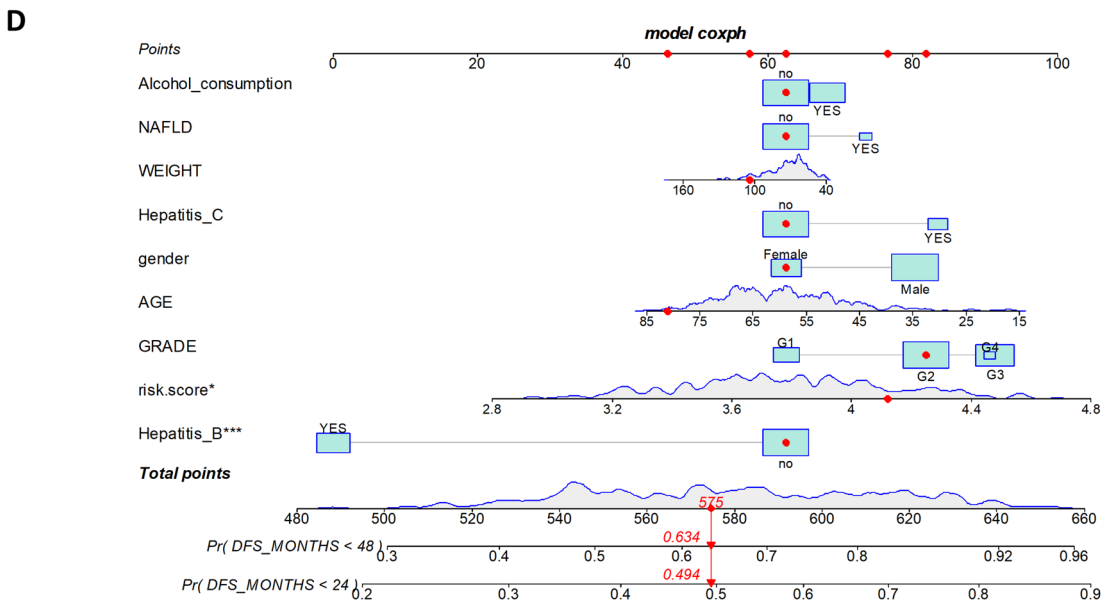

D
model coxph
Points
0
20
40
60
80
100
Alcohol_consumption
no
YES
NAFLD
no
YES
WEIGHT
160
100
40
Hepatitis_C
no
YES
gender
Female
Male
AGE
85
75
65
55
45
35
25
15
GRADE
G1
G2
G3
G4
risk.score*
2.8
3.2
3.6
4
4.4
4.8
Hepatitis_B***
YES
no
Total points
480
500
520
540
560
580
600
620
640
660
575
Pr(DFS_MONTHS < 48)
0.3
0.4
0.5
0.6
0.7
0.8
0.92
0.96
0.634
Pr(DFS_MONTHS < 24)
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0.494

Fig. 10 The SETDB1-Related Risk Score is an Independent Unfavorable Prognostic Factor for Disease-Free Survival in Liver Cancer. **(A)** Multivariable Cox model for disease-free survival (DFS) in the TCGA cohort. The forest plot shows the hazard ratio (HR) and p-values for the SETDB1-related risk score and various clinical and demographic factors, including age, gender, weight, pathological grade, and liver cancer etiology (hepatitis B, hepatitis C, NAFLD, and alcohol consumption). **(B)** Calibration plot of the multivariable DFS model at two years. This plot, performed by bootstrapping, shows the predicted probability of disease-free survival at two years compared to the observed outcome, indicating the model's predictive accuracy. **(C)** Calibration plot of the multivariable DFS model at four years. Similar to panel B, this plot shows the model's predictive performance for the probability of disease-free survival at four years. **(D)** Nomogram for predicting DFS. The nomogram illustrates the probability of disease-free survival at two and four years of follow-up. The red markers highlight the values for a representative individual case from the TCGA cohort

that paclitaxel treatment reduced SETDB1 expression and decreased stemness markers. However, paclitaxel does not directly inhibit SETDB1. Instead, it represses SETDB1 through a p53-dependent mechanism, whereby activated p53 binds the SETDB1 promoter and recruits SUV39H1 to increase H3K9me3 deposition, leading to transcriptional repression [37]. Furthermore, SETDB1 can directly methylate p53 at lysine 370, stabilizing oncogenic mutant p53 proteins by preventing their MDM2-mediated degradation [44]. This p53-mediated repression of SETDB1 is particularly meaningful, as SETDB1 itself transcriptionally silences tumor-suppressor genes, including p53, through H3K9 methylation at specific promoter sites [71], establishing a self-reinforcing oncogenic feedback loop. By reducing SETDB1 levels, paclitaxel disrupts this oncogenic circuitry, allowing reactivation of p53-mediated apoptotic pathways while simultaneously destabilizing gain-of-function p53 mutants that depend on SETDB1-driven methylation for their stability. This dual effect provides a plausible mechanistic basis for the reduced CSC-associated traits observed following paclitaxel exposure.

The reduction in CSC features observed following paclitaxel exposure may therefore reflect broader p53-driven transcriptional reprogramming rather than direct SETDB1 inhibition. As we did not perform SETDB1 knockdown or rescue assays, these findings should be interpreted as correlational. Future mechanistic studies employing genetic SETDB1 inhibition will be essential to establish causality and confirm SETDB1 as a therapeutic vulnerability in HCC.

Future therapeutic approaches may benefit from combination therapies targeting SETDB1 in conjunction with ICIs, ferroptosis inducers, or synthetic lethality partners. Given the functional heterogeneity within SH tumors, precision targeting will require selective inhibitors, refined molecular stratification, and context-specific therapeutic combinations.

In conclusion, our integrative analyses identify SETDB1 as a core epigenetic regulator, supporting CSC identity, immunosuppression, and broad therapy resistance in HCC. SETDB1 defines a clinically aggressive molecular subtype marked by poor differentiation, heightened plasticity, and unfavorable outcomes. These findings highlight SETDB1 as an attractive biomarker and therapeutic target, and they lay the groundwork for future mechanistic studies and translational efforts aimed at incorporating SETDB1-directed strategies into precision medicine for HCC.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13402-025-01157-3.

Author contributions MP and AD performed, analyzed, and interpreted in vitro data. CD and JGM performed and analyzed computational analyses. MP, CD, GU, AL and JGM wrote the manuscript. GU and AL supervised the study. All authors read and approved the final manuscript.

Funding Open access funding provided by Université Paris-Saclay. This research received no specific grant.

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethics approvals and consent to participate Not applicable.

Consent for publication All authors have given their approval for the manuscript to be submitted in its current form.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. H. Rumgay, M. Arnold, J. Ferlay, O. Lesi, C.J. Cabasag, J. Vignat et al., Global burden of primary liver cancer in 2020 and predictions to 2040. J. Hepatol. **77**, 1598–1606 (2022). https://doi.org/10.1016/j.jhep.2022.08.021
2. A.G. Singal, E. Zhang, M. Narasimman, N.E. Rich, A.K. Waljee, Y. Hoshida et al., HCC surveillance improves early Detection, curative treatment Receipt, and survival in patients with cirrhosis: A systematic review and Meta-Analysis. J. Hepatol. **77**, 128–139 (2022). https://doi.org/10.1016/j.jhep.2022.01.023

3. Z.-H. Liu, B.-F. Lian, Q.-Z. Dong, H. Sun, J.-W. Wei, Y.-Y. Sheng et al., Whole-exome mutational and transcriptional landscapes of combined hepatocellular cholangiocarcinoma and intrahepatic cholangiocarcinoma reveal molecular diversity. Biochim. Biophys. Acta Mol. Basis Dis. **1864**, 2360–2368 (2018). https://doi.org/10.1016/j.bbadis.2018.01.027
4. Y. Shiozawa, B. Nie, K.J. Pienta, T.M. Morgan, R.S. Taichman, Cancer stem cells and their role in metastasis. Pharmacol. Ther. **138**, 285–293 (2013). https://doi.org/10.1016/j.pharmthera.2013.01.014
5. L.T.H. Phi, I.N. Sari, Y.-G. Yang, S.-H. Lee, N. Jun, K.S. Kim et al., Cancer stem cells (CSCs) in drug resistance and their therapeutic implications in cancer treatment. Stem Cells Int. **2018**, 5416923 (2018). https://doi.org/10.1155/2018/5416923
6. A. Kreso, J.E. Dick, Evolution of the cancer stem cell model. Cell. Stem Cell. **14**, 275–291 (2014). https://doi.org/10.1016/j.stem.2014.02.006
7. J. Liu, P. Li, L. Wang, M. Li, Z. Ge, L. Noordam et al., Cancer-Associated fibroblasts provide a stromal niche for liver cancer organoids that confers trophic effects and therapy resistance. Cell. Mol. Gastroenterol. Hepatol. **11**, 407–431 (2020). https://doi.org/10.1016/j.jcmgh.2020.09.003
8. N. Wang, S. Wang, M.-Y. Li, B.-G. Hu, L.-P. Liu, S.-L. Yang et al., Cancer stem cells in hepatocellular carcinoma: an overview and promising therapeutic strategies. Ther. Adv. Med. Oncol. **10**, 1758835918816287 (2018). https://doi.org/10.1177/1758835918816287
9. Y. Chen, B. Ren, J. Yang, H. Wang, G. Yang, R. Xu et al., The role of histone methylation in the development of digestive cancers: a potential direction for cancer management. Signal. Transduct. Target. Ther. **5**, 143 (2020). https://doi.org/10.1038/s41392-020-00252-1
10. A.A. Emran, D.M. Marzese, D.R. Menon, H. Hammerlindl, F. Ahmed, E. Richtig et al., Commonly integrated epigenetic modifications of differentially expressed genes lead to adaptive resistance in cancer. Epigenomics. **11**, 732–737 (2019). https://doi.org/10.2217/epi-2018-0173
11. J.E. Dodge, Y.-K. Kang, H. Beppu, H. Lei, E. Li, Histone H3-K9 methyltransferase ESET is essential for early development. Mol. Cell. Biol. **24**, 2478–2486 (2004). https://doi.org/10.1128/MCB.24.6.2478-2486.2004
12. S. Koide, M. Oshima, K. Takubo, S. Yamazaki, E. Nitta, A. Saraya et al., Setdb1 maintains hematopoietic stem and progenitor cells by restricting the ectopic activation of nonhematopoietic genes. Blood. **128**, 638–649 (2016). https://doi.org/10.1182/blood-2016-01-694810
13. J. Batham, P.S. Lim, S. Rao, SETDB-1: A potential epigenetic regulator in breast cancer metastasis. Cancers (Basel). **11**, 1143 (2019). https://doi.org/10.3390/cancers11081143
14. Y.-H. Chang, S. Goyama, The role of histone methyltransferase SETDB1 in normal and malignant hematopoiesis. Cancer Sci. (2025). https://doi.org/10.1111/cas.70116
15. E. Verdier, N. Gaspar, M.E. Marques Da Costa, A. Marchais, SETDB1 amplification in osteosarcomas: insights from its role in healthy tissues and other cancer types. Oncotarget. **16**, 51–62 (2025). https://doi.org/10.18632/oncotarget.28688
16. C.-Q. Yin, C.-Q. Song, Tumor intrinsic immunogenicity suppressor SETDB1 worsens the prognosis of patients with hepatocellular carcinoma. Cells. **13**, 2102 (2024). https://doi.org/10.3390/cells13242102
17. J. Zhou, J. Che, L. Xu, W. Yang, Y. Li, W. Zhou et al., Enhancer of Zeste homolog 2 promotes hepatocellular cancer progression and chemoresistance by enhancing protein kinase B activation through microRNA-381-mediated SET domain bifurcated 1. Bioengineered. **13**, 5737–5755 (2022). https://doi.org/10.1080/21655979.2021.2023792
18. J. Liu, T. Lichtenberg, K.A. Hoadley, L.M. Poisson, A.J. Lazar, A.D. Cherniack et al., An integrated TCGA Pan-Cancer clinical data resource to drive High-Quality survival outcome analytics. Cell. **173**, 400–416e11 (2018). https://doi.org/10.1016/j.cell.2018.02.052
19. J. Gao, B.A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S.O. Sumer et al., Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. **6**, pl1 (2013). https://doi.org/10.1126/scisignal.2004088
20. S.V. Vasaikar, P. Straub, J. Wang, B. Zhang, LinkedOmics: analyzing multi-omics data within and across 32 cancer types. Nucleic Acids Res. **46**, D956–D963 (2018). https://doi.org/10.1093/nar/gkx1090
21. D.A. Barbie, P. Tamayo, J.S. Boehm, S.Y. Kim, S.E. Moody, I.F. Dunn et al., Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. **462**, 108–112 (2009). https://doi.org/10.1038/nature08460
22. D. Chicco, G. Jurman, The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genom. **21**, 6 (2020). https://doi.org/10.1186/s12864-019-6413-7
23. K. Schulze, S. Imbeaud, E. Letouzé, L.B. Alexandrov, J. Calderaro, S. Rebouissou et al., Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat. Genet. **47**, 505–511 (2015). https://doi.org/10.1038/ng.3252
24. L. Ma, M.O. Hernandez, Y. Zhao, M. Mehta, B. Tran, M. Kelly et al., Tumor cell biodiversity drives microenvironmental reprogramming in liver cancer. Cancer Cell. **36**, 418–430e6 (2019). https://doi.org/10.1016/j.ccell.2019.08.007
25. Q. Chen, S.-B. Yang, Y.-W. Zhang, S.-Y. Han, L. Jia, B. Li et al., miR-3682-3p directly targets FOXO3 and stimulates tumor stemness in hepatocellular carcinoma via a positive feedback loop involving FOXO3/PI3K/AKT/c-Myc. World J. Stem Cells. **14**, 539–555 (2022). https://doi.org/10.4252/wjsc.v14.i7.539
26. Y.-L. Yang, Y.-H. Chang, C.-J. Li, Y.-H. Huang, M.-C. Tsai, P.-Y. Chu et al., New insights into the role of miR-29a in hepatocellular carcinoma: implications in mechanisms and theragnostics. J. Pers. Med. **11**, 219 (2021). https://doi.org/10.3390/jpm11030219
27. J.-P. Gillet, J.B. Andersen, J.P. Madigan, S. Varma, R.K. Bagni, K. Powell et al., A gene expression signature associated with overall survival in patients with hepatocellular carcinoma suggests a new treatment strategy. Mol. Pharmacol. **89**, 263–272 (2016). https://doi.org/10.1124/mol.115.101360
28. C. Li, X. Cui, Y. Li, D. Guo, S. He, Identification of ferroptosis and drug resistance related hub genes to predict the prognosis in hepatocellular carcinoma. Sci. Rep. **13**, 8681 (2023). https://doi.org/10.1038/s41598-023-35796-z
29. H.G. Woo, J.-H. Lee, J.-H. Yoon, C.Y. Kim, H.-S. Lee, J.J. Jang et al., Identification of a Cholangiocarcinoma-Like gene expression trait in hepatocellular carcinoma. Cancer Res. **70**, 3034–3041 (2010). https://doi.org/10.1158/0008-5472.CAN-09-2823
30. D. Pan, Y. Du, R. Li, A. Shen, X. Liu, C. Li et al., miR-29b-3p increases radiosensitivity in stemness cancer cells via modulating oncogenes axis. Front. Cell. Dev. Biol. **9**, (2021). https://doi.org/10.3389/fcell.2021.741074
31. L. Boix, J.M. López-Oliva, A.C. Rhodes, J. Bruix, Restoring miR122 in human stem-like hepatocarcinoma cells, prompts tumor dormancy through Smad-independent TGF-β pathway. Oncotarget. **7**, 71309–71329 (2016). https://doi.org/10.18632/oncotarget.11885
32. X.-F. Fu, H.-C. Zhao, C.-L. Yang, C.-Z. Chen, K. Wang, F. Gao et al., MicroRNA-203-3p inhibits the proliferation, invasion and migration of pancreatic cancer cells by downregulating fibroblast growth factor 2. Oncol. Lett. **22**, 626 (2021). https://doi.org/10.3892/ol.2021.12887

33. J. Azumi, T. Tsubota, T. Sakabe, G. Shiota, miR-181a induces Sorafenib resistance of hepatocellular carcinoma cells through downregulation of RASSF1 expression. Cancer Sci. **107**, 1256–1262 (2016). https://doi.org/10.1111/cas.13006
34. F. Zhong, Y. Wang, YY1-regulated LncRNA SOCS2-AS1 suppresses HCC cell stemness and progression via miR-454-3p/CPEB1. Biochem. Biophys. Res. Commun. **679**, 98–109 (2023). https://doi.org/10.1016/j.bbrc.2023.08.056
35. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. **36**, 411–420 (2018). https://doi.org/10.1038/nbt.4096
36. Di M. Giovannantonio, F. Hartley, B. Elshenawy, A. Barberis, D. Hudson, H.S. Shafique et al., Defining hypoxia in cancer: A landmark evaluation of hypoxia gene expression signatures. Cell. Genom. **5**, 100764 (2025). https://doi.org/10.1016/j.xgen.2025.100764
37. H.-J. Noh, K.-A. Kim, K.-C. Kim, p53 Down-regulates SETDB1 gene expression during Paclitaxel induced-cell death. Biochem. Biophys. Res. Commun. **446**, 43–48 (2014). https://doi.org/10.1016/j.bbrc.2014.02.053
38. J. Giron-Michel, M. Padelli, E. Oberlin, H. Guenou, J.-C. Duclos-Vallée, State-of-the-Art liver cancer organoids: modeling cancer stem cell heterogeneity for personalized treatment. BioDrugs. **39**, 237–260 (2025). https://doi.org/10.1007/s40259-024-00702-0
39. X. Dai, Y. Guo, Y. Hu, X. Bao, X. Zhu, Q. Fu et al., Immunotherapy for targeting cancer stem cells in hepatocellular carcinoma. Theranostics. **11**, 3489–3501 (2021). https://doi.org/10.7150/thno.54648
40. E. Zhang, P. He, The function of histone methyltransferase SETDB1 and its roles in liver cancer. Front. Cell. Dev. Biol. **12**, 1500263 (2024). https://doi.org/10.3389/fcell.2024.1500263
41. E. Orouji, A. Federico, L. Larribère, D. Novak, D.B. Lipka, Y. Assenov et al., Histone methyltransferase SETDB1 contributes to melanoma tumorigenesis and serves as a new potential therapeutic target. Int. J. Cancer. **145**, 3462–3477 (2019). https://doi.org/10.1002/ijc.32432
42. L. Yu, F. Ye, Y.-Y. Li, Y.-Z. Zhan, Y. Liu, H.-M. Yan et al., Histone methyltransferase SETDB1 promotes colorectal cancer proliferation through the STAT1-CCND1/CDK6 axis. Carcinogenesis. **41**, 678–688 (2020). https://doi.org/10.1093/carcin/bgz131
43. C.-M. Wong, L. Wei, C.-T. Law, D.W.-H. Ho, F.H.-C. Tsang, S.L.-K. Au et al., Up-regulation of histone methyltransferase SETDB1 by multiple mechanisms in hepatocellular carcinoma promotes cancer metastasis. Hepatology. **63**, 474–487 (2016). https://doi.org/10.1002/hep.28304
44. Q. Fei, K. Shang, J. Zhang, S. Chuai, D. Kong, T. Zhou et al., Histone methyltransferase SETDB1 regulates liver cancer cell growth through methylation of p53. Nat. Commun. **6**, 8651 (2015). https://doi.org/10.1038/ncomms9651
45. D. Strepkos, M. Markouli, A. Klonou, A.G. Papavassiliou, C. Piperi, Histone methyltransferase SETDB1: A common denominator of tumorigenesis with therapeutic potential. Cancer Res. **81**, 525–534 (2021). https://doi.org/10.1158/0008-5472.CAN-20-2906
46. S. Park, J.H. Cho, J.-H. Kim, M. Park, S. Park, S.-Y. Kim et al., Hypoxia stabilizes SETDB1 to maintain genome stability. Nucleic Acids Res. **51**, 11178–11196 (2023). https://doi.org/10.1093/nar/gkad796
47. J.-F. Xiao, Q.-Y. Sun, L.-W. Ding, W. Chien, X.-Y. Liu, A. Mayakonda et al., The c-MYC-BMI1 axis is essential for SETDB1-mediated breast tumourigenesis. J. Pathol. **246**, 89–102 (2018). https://doi.org/10.1002/path.5126
48. S. Cho, J.S. Park, S. Kwon, Y.-K. Kang, Dynamics of Setdb1 expression in early mouse development. Gene Expr Patterns. **12**, 213–218 (2012). https://doi.org/10.1016/j.gep.2012.03.005
49. L.-S. Yeap, K. Hayashi, M.A. Surani, ERG-associated protein with SET domain (ESET)-Oct4 interaction regulates pluripotency and represses the trophectoderm lineage. Epigenetics Chromatin. **2**, 12 (2009). https://doi.org/10.1186/1756-8935-2-12
50. Y.-H. Loh, Q. Wu, J.-L. Chew, V.B. Vega, W. Zhang, X. Chen et al., The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. **38**, 431–440 (2006). https://doi.org/10.1038/ng1760
51. K. Wu, H. Liu, Y. Wang, J. He, S. Xu, Y. Chen et al., SETDB1-Mediated cell fate transition between 2 C-Like and pluripotent States. Cell. Rep. **30**, 25–36e6 (2020). https://doi.org/10.1016/j.celrep.2019.12.010
52. S. Cho, J.S. Park, Y.-K. Kang, Dual functions of histone-lysine N-methyltransferase Setdb1 protein at promyelocytic leukemia-nuclear body (PML-NB): maintaining PML-NB structure and regulating the expression of its associated genes. J. Biol. Chem. **286**, 41115–41124 (2011). https://doi.org/10.1074/jbc.M111.248534
53. X. Zhang, B. Hu, Y.-F. Sun, X.-W. Huang, J.-W. Cheng, A. Huang et al., Arsenic trioxide induces differentiation of cancer stem cells in hepatocellular carcinoma through Inhibition of LIF/JAK1/STAT3 and NF-kB signaling pathways synergistically. Clin. Transl Med. **11**, e335 (2021). https://doi.org/10.1002/ctm2.335
54. C. Fracassi, M. Ugge', M. Abdelhalim, E. Zapparoli, M. Simoni, D. Magliulo et al., PML modulates epigenetic composition of chromatin to regulate expression of pro-metastatic genes in triple-negative breast cancer. Nucleic Acids Res. **51**, 11024–11039 (2023). https://doi.org/10.1093/nar/gkad819
55. G.K. Griffin, J. Wu, A. Iracheta-Vellve, J.C. Patti, J. Hsu, T. Davis et al., Epigenetic Silencing by SETDB1 suppresses tumour intrinsic immunogenicity. Nature. **595**, 309–314 (2021). https://doi.org/10.1038/s41586-021-03520-4
56. Y. Lyu, S.J. Kim, E.S. Humphrey, R. Nayak, Y. Guan, Q. Liang et al., Stem cell activity-coupled suppression of endogenous retrovirus governs adult tissue regeneration. Cell. **187**, 7414–7432e26 (2024). https://doi.org/10.1016/j.cell.2024.10.007
57. X. Jiao, X. Qian, L. Wu, B. Li, Y. Wang, X. Kong et al., MicroRNA: the impact on cancer stemness and therapeutic resistance. Cells. **9**, 8 (2019). https://doi.org/10.3390/cells9010008
58. Y. Shao, X. Song, W. Jiang, Y. Chen, Z. Ning, W. Gu et al., MicroRNA-621 acts as a tumor radiosensitizer by directly targeting SETDB1 in hepatocellular carcinoma. Mol. Ther. **27**, 355–364 (2019). https://doi.org/10.1016/j.ymthe.2018.11.005
59. N. Cao, Y. Yu, H. Zhu, M. Chen, P. Chen, M. Zhuo et al., SETDB1 promotes the progression of colorectal cancer via epigenetically Silencing p21 expression. Cell. Death Dis. **11**, 351 (2020). https://doi.org/10.1038/s41419-020-2561-6
60. W. Yang, Y. Su, C. Hou, L. Chen, D. Zhou, K. Ren et al., SETDB1 induces epithelial–mesenchymal transition in breast carcinoma by directly binding with snail promoter. Oncol. Rep. **41**, 1284–1292 (2019). https://doi.org/10.3892/or.2018.6871
61. H. Zhang, K. Cai, J. Wang, X. Wang, K. Cheng, F. Shi et al., MiR-7, inhibited indirectly by LincRNA HOTAIR, directly inhibits SETDB1 and reverses the EMT of breast cancer stem cells by downregulating the STAT3 pathway. Stem Cells. **32**, 2858–2868 (2014). https://doi.org/10.1002/stem.1795
62. D. Du, Y. Katsuno, D. Meyer, E.H. Budi, S. Chen, H. Koeppen et al., Smad3-mediated recruitment of the methyltransferase SETDB1/ESET controls Snail1 expression and epithelial–mesenchymal transition. EMBO Rep. **19**, 135–155 (2018). https://doi.org/10.15252/embr.201744250
63. T. Liu, P. Xu, S. Ke, H. Dong, M. Zhan, Q. Hu et al., Histone methyltransferase SETDB1 inhibits TGF-β-induced epithelial-mesenchymal transition in pulmonary fibrosis by regulating SNAI1 expression and the ferroptosis signaling pathway. Arch.

Biochem. Biophys. **715**, 109087 (2022). https://doi.org/10.1016/j.abb.2021.109087
64. S.A. Mani, W. Guo, M.-J. Liao, E. ENg, A. Ayyanan, A.Y. Zhou et al., The Epithelial-Mesenchymal transition generates cells with properties of stem cells. Cell. **133**, 704–715 (2008). https://doi.org/10.1016/j.cell.2008.03.027
65. M.K. Jolly, S.C. Tripathi, J.A. Somarelli, S.M. Hanash, H. Levine, Epithelial/mesenchymal plasticity: how have quantitative mathematical models helped improve our understanding? Mol. Oncol. **11**, 739–754 (2017). https://doi.org/10.1002/1878-0261.12084
66. M.K. Jolly, R.J. Murphy, S. Bhatia, H.J. Whitfield, A. Redfern, M.J. Davis et al., Measuring and modelling the Epithelial- mesenchymal hybrid state in cancer: clinical implications. Cells Tissues Organs. **211**, 110–133 (2022). https://doi.org/10.1159/000515289
67. C. Kröger, A. Afeyan, J. Mraz, E.N. Eaton, F. Reinhardt, Y.L. Khodor et al., Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. Proc. Natl. Acad. Sci. U.S.A. **116**, 7353–7362 (2019). https://doi.org/10.1073/pnas.1812876116
68. V.M. Perelmuter, E.S. Grigoryeva, V.V. Alifanov, A.Y. Kalinchuk, E.S. Andryuhova, O.E. Savelieva et al., Characterization of EpCAM-Positive and EpCAM-Negative tumor cells in Early-Stage breast cancer. Int. J. Mol. Sci. **25**, 11109 (2024). https://doi.org/10.3390/ijms252011109
69. G. Zhang, W. Fan, H. Wang, J. Wen, J. Tan, M. Xue et al., Non-Apoptotic programmed cell Death-Related gene signature correlates with stemness and immune status and predicts the responsiveness of transarterial chemoembolization in hepatocellular carcinoma. Front. Cell. Dev. Biol. **10**, 844013 (2022). https://doi.org/10.3389/fcell.2022.844013
70. Q. Zhu, S. Liao, T. Wei, S. Liu, C. Yang, J. Tang, Development of a novel prognostic signature based on cytotoxic T lymphocyte-evasion genes for hepatocellular carcinoma patient management. Discov Onc. **16**, 144 (2025). https://doi.org/10.1007/s12672-025-01909-5
71. S. Ogawa, A. Fukuda, Y. Matsumoto, Y. Hanyu, M. Sono, Y. Fukunaga et al., SETDB1 inhibits p53-Mediated apoptosis and is required for formation of pancreatic ductal adenocarcinomas in mice. Gastroenterology. **159**, 682–696e13 (2020). https://doi.org/10.1053/j.gastro.2020.04.047

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.