# 1 Integrating bulk and single-cell RNA sequencing reveals SH3D21

2 promotes hepatocellular carcinoma progression by activating the

# **3 PI3K/AKT/mTOR pathway**

# 4 Wangxia Tong<sup>1</sup>, Tao Lu<sup>2</sup>, Li Liu<sup>1</sup>, Rong Liu<sup>1</sup>, Jibing Chen<sup>3</sup>, Ning Luo<sup>4</sup>

- 5 <sup>1</sup> Department of Hepatology, Ruikang Hospital Affiliated to Guangxi University of Chinese Medicine,
- 6 Nanning, Guangxi, China.
- 7 <sup>2</sup> Department of hepatobiliary surgery, Ruikang Hospital Affiliated to Guangxi University of Chinese
- 8 Medicine, Nanning, Guangxi, China.
- 9 <sup>3</sup>Center for Translational Medicine of Integrated Traditional Chinese and Western Medicine, Ruikang
- 10 Hospital Affiliated to Guangxi University of Chinese Medicine, Nanning, Guangxi, China.
- <sup>4</sup> Department of Neurology, RuiKang Hospital Affiliated to Guangxi University of Chinese Medicine,
- 12 Nanning, Guangxi, China.
- 13 \*Corresponding author
- 14 Ning Luo
- 15 E-mail: <u>ln760320@163.com</u>
- 16 Wangxia Tong and Tao Lu contributed equally to this work.
- 17 &Jibing Chen and Ning Luo contributed equally to this work.

## 18 Abstract

- 19 As a novel genetic biomarker, the potential role of *SH3D21* in hepatocellular carcinoma remains unclear.
- 20 Here, we decipher the expression and function of SH3D21 in human hepatocellular carcinoma. The
- 21 expression level and clinical significance of SH3D21 in hepatocellular carcinoma patients, the
- 22 relationship between SH3D21 and the features of tumor microenvironment (TME) and role of SH3D21
- 23 in promoting hepatocellular carcinoma progression were analyzed based on the bulk samples obtained
- from The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC)

#### NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

25	databases. Single-cell sequencing samples from Gene Expression Omnibus (GEO) database were
26	employed to verify the prediction mechanism. Additionally, different biological effects of SH3D21 on
27	hepatocellular carcinoma cells were investigated by qRT-PCR, CCK-8 assay, colony forming assay and
28	Western blot analysis. Bioinformatics analysis and <i>in vitro</i> experiments revealed that the expression level
29	of SH3D21 was up-regulated in hepatocellular carcinoma and correlated with the poor prognosis in
30	hepatocellular carcinoma patients. SH3D21 effectively promoted the proliferation, invasion, and
31	migration as well as the formation of immunosuppressive microenvironment of hepatocellular carcinoma.
32	In addition, SH3D21 can activate the PI3K/AKT/mTOR signaling pathway. SH3D21 stimulates the
33	progression of hepatocellular carcinoma by activating the PI3K/AKT/mTOR signaling pathway, and
34	SH3D21 can serve as a prognostic biomarker and therapeutic target for hepatocellular carcinoma.

35 Keywords: Hepatocellular carcinoma, SH3 domain containing 21 (SH3D21), Single-cell RNA
36 sequencing, PI3K/AKT/mTOR pathway

## 37 1. Introduction

38 Primary hepatoma is one of the most common malignant tumors, primarily hepatocellular carcinoma 39 (HCC). It is often characterized by a high degree of malignancy, with a propensity for metastasis and 40 recurrence following the treatment [1, 2]. Despite the availability of various treatment modalities, 41 including surgical resection, radiotherapy, and chemotherapy [2, 3], the prognosis for HCC patients 42 remains significantly poor [4]. Recent advancements in targeted therapy for HCC, such as application of 43 vascular endothelial growth factor receptor (VEGFR) inhibitors and programmed cell death 1 ligand 44 1(PD-L1) inhibitors, have sparked newfound hope [5, 6]. However, the efficacy of targeted therapy is 45 limited, thereby benefiting fewer than one-third of patients and only marginally prolonging the median

46 survival by a mere three months [7]. Hence, identification of effective biomarker is urgently needed for
47 targeted therapy of HCC.

48 The PI3K/Akt/mTOR pathway represents a crucial intracellular signaling cascade that can regulate 49 diverse cellular processes including proliferation, apoptosis, metabolism, and angiogenesis, by 50 functioning through intricate crosstalk with upstream or downstream molecules [8, 9]. Aberrant 51 activation of the PI3K/Akt/mTOR pathway has been reported in approximately 50% of hepatoma cases 52 [10], thereby exerting significant influence on the various facets of hepatoma initiation as well as 53 progression, including tumor cell proliferation, differentiation, metabolism, autophagy, and immune 54 regulation [11, 12]. Although multi-targeted tyrosine kinase inhibitors are being used as first-line 55 therapeutic agents targeting the PI3K/Akt/mTOR pathway, their efficacy in improving patient survival 56 remains modest. Notably, patients often develop acquired drug resistance during the treatment, thus 57 limiting the therapeutic options available [13]. This could be possibly attributed to the limited 58 understanding of the intricate interactions and metabolic roles of the PI3K/Akt/mTOR signaling axis in 59 hepatoma [14].

60 SH3 domain-containing 21 (*SH3D21*) is a SH3 domain containing protein that is localized within 61 the nucleus and plasma membrane. It functions by forming protein complexes and thereby modulating 62 intracellular signaling to regulate various key biological processes, including cell division, differentiation, 63 and growth and development [15]. Despite its biological significance, the current understanding of the 64 role of *SH3D21* in human tumors remains limited. Given the dearth of experimental studies related to 65 the role of *SH3D21* in tumorigenesis, this study was designed to elucidate its involvement in tumors and 66 evaluate its potential utility both as a novel biomarker and therapeutic target.

67 The association between *SH3D21* and prognosis as well as clinical features in HCC patients were

68	analyzed in this study. The mechanisms through which SH3D21 can promote HCC and its effect on
69	PI3K/Akt/mTOR pathway activation were also investigated. The clinical samples and cell experiments
70	were employed to verify the experimental results.
71	2. Materials and methods
72	Data source
73	The transcriptome and clinical characteristics analysis data of SH3D21 in HCC were obtained from
74	$TCGA\ database\ (\ https://portal.gdc.cancer.gov\ )\ and\ ICGC\ database\ (\ https://dcc.icgc.org/\ )$
75	(download time: 20221224). We have also incorporated the genetic variation information obtained from
76	the biological portal website (https://www.cbioportal.org/) [16]. Specifically, we have utilized a dataset
77	comprising 442 liver samples obtained from TCGA and Firhose Legacy. The data of human pan-cancer
78	was retrieved from the UCSC Xena online database ( <u>https://xena.ucsc.edu/</u> ) [17]. The downloaded data
70	
79	included mRNA expressions and the clinical features of 33 different tumors.
79 80	included mRNA expressions and the clinical features of 33 different tumors. Differentially expressed genes (DEGs) and analysis of characterization of tumor
79 80 81	included mRNA expressions and the clinical features of 33 different tumors. Differentially expressed genes (DEGs) and analysis of characterization of tumor microenvironment
79 80 81 82	<ul> <li>included mRNA expressions and the clinical features of 33 different tumors.</li> <li>Differentially expressed genes (DEGs) and analysis of characterization of tumor microenvironment</li> <li>The "Limma" software package in the R language was used to identify the various differentially</li> </ul>
79 80 81 82 83	<ul> <li>included mRNA expressions and the clinical features of 33 different tumors.</li> <li>Differentially expressed genes (DEGs) and analysis of characterization of tumor microenvironment</li> <li>The "Limma" software package in the R language was used to identify the various differentially expressed genes (DEGs) and the expression levels of <i>SH3D21</i> in both the tumor tissues and control</li> </ul>
79 80 81 82 83 84	included mRNA expressions and the clinical features of 33 different tumors. <b>Differentially expressed genes (DEGs) and analysis of characterization of tumor</b> <b>microenvironment</b> The "Limma" software package in the R language was used to identify the various differentially expressed genes (DEGs) and the expression levels of <i>SH3D21</i> in both the tumor tissues and control tissues in the dataset. The differential gene screening was identified using the criteria of log <sub>2</sub> FC>1.0 and
79 80 81 82 83 83 84 85	included mRNA expressions and the clinical features of 33 different tumors. <b>Differentially expressed genes (DEGs) and analysis of characterization of tumor microenvironment</b> The "Limma" software package in the R language was used to identify the various differentially expressed genes (DEGs) and the expression levels of <i>SH3D21</i> in both the tumor tissues and control tissues in the dataset. The differential gene screening was identified using the criteria of log <sub>2</sub> FC>1.0 and FDR<0.05, thereby defining the genes exhibiting significant differential expression. We have stratified
79 80 81 82 83 83 84 85 86	included mRNA expressions and the clinical features of 33 different tumors. <b>Differentially expressed genes (DEGs) and analysis of characterization of tumor</b> <b>microenvironment</b> The "Limma" software package in the R language was used to identify the various differentially expressed genes (DEGs) and the expression levels of <i>SH3D21</i> in both the tumor tissues and control tissues in the dataset. The differential gene screening was identified using the criteria of log <sub>2</sub> FC>1.0 and FDR<0.05, thereby defining the genes exhibiting significant differential expression. We have stratified the tumor-derived samples into up- and down-regulation groups based on the median <i>SH3D21</i> expression
<ol> <li>79</li> <li>80</li> <li>81</li> <li>82</li> <li>83</li> <li>84</li> <li>85</li> <li>86</li> <li>87</li> </ol>	included mRNA expressions and the clinical features of 33 different tumors. <b>Differentially expressed genes (DEGs) and analysis of characterization of tumor</b> <b>microenvironment</b> The "Limma" software package in the R language was used to identify the various differentially expressed genes (DEGs) and the expression levels of <i>SH3D21</i> in both the tumor tissues and control tissues in the dataset. The differential gene screening was identified using the criteria of log <sub>2</sub> FC>1.0 and FDR<0.05, thereby defining the genes exhibiting significant differential expression. We have stratified the tumor-derived samples into up- and down-regulation groups based on the median <i>SH3D21</i> expression level. The IOBR package in R language was employed to analyze the potential relationship between
79 80 81 82 83 83 84 85 86 87 88	included mRNA expressions and the clinical features of 33 different tumors. <b>Differentially expressed genes (DEGs) and analysis of characterization of tumor</b> <b>microenvironment</b> The "Limma" software package in the R language was used to identify the various differentially expressed genes (DEGs) and the expression levels of <i>SH3D21</i> in both the tumor tissues and control tissues in the dataset. The differential gene screening was identified using the criteria of log <sub>2</sub> FC>1.0 and FDR<0.05, thereby defining the genes exhibiting significant differential expression. We have stratified the tumor-derived samples into up- and down-regulation groups based on the median <i>SH3D21</i> expression level. The IOBR package in R language was employed to analyze the potential relationship between <i>SH3D21</i> expression level and the characteristics of tumor microenvironment by ssGSEA algorithm.

## 90 Genes and Genomes) analyses of bulk samples

91	The tumor samples in TCGA data set were divided into up- and down-regulation groups based on the
92	median expression level of SH3D21 for the differential gene analysis. In addition, GO, KEGG and GSEA
93	enrichment analysis of the differential genes were performed. The tumor samples in ICGC data set were
94	also divided into up- and down-regulation groups according to the median expression level of SH3D21
95	for differential gene analysis. Thereafter, the correlation between the expression level of SH3D21 and
96	the differential genes was analyzed (CoR=0.4, $P \le 0.05$ ). Subsequently, the high and low expression
97	differential genes were fed into WebGestalt (www.WebGestalt.org), an online bioinformatics website
98	that utilizes the KEGG database [18]. P<0.05 was considered to be a statistically different enrichment
99	pathway.
99 100	pathway. Single-cell sequencing sample quality control and cell type identification
99 100 101	pathway. Single-cell sequencing sample quality control and cell type identification The quality control process of the single-cell sequencing samples was conducted using Seurat (version
99 100 101 102	pathway. Single-cell sequencing sample quality control and cell type identification The quality control process of the single-cell sequencing samples was conducted using Seurat (version 4.15) [19]. To ensure high-quality cells for the downstream analysis, the following criteria were applied:
99 100 101 102 103	pathway. <b>Single-cell sequencing sample quality control and cell type identification</b> The quality control process of the single-cell sequencing samples was conducted using Seurat (version 4.15) [19]. To ensure high-quality cells for the downstream analysis, the following criteria were applied: The cells expressing fewer than 500 or more than 6000 total genes were filtered out, cells with over 15%
99 100 101 102 103 104	<ul> <li>pathway.</li> <li>Single-cell sequencing sample quality control and cell type identification</li> <li>The quality control process of the single-cell sequencing samples was conducted using Seurat (version 4.15) [19]. To ensure high-quality cells for the downstream analysis, the following criteria were applied:</li> <li>The cells expressing fewer than 500 or more than 6000 total genes were filtered out, cells with over 15%</li> <li>mitochondrial genes were excluded and the genes expressed in fewer than three cells were removed.</li> </ul>
99 100 101 102 103 104 105	<ul> <li>pathway.</li> <li>Single-cell sequencing sample quality control and cell type identification</li> <li>The quality control process of the single-cell sequencing samples was conducted using Seurat (version 4.15) [19]. To ensure high-quality cells for the downstream analysis, the following criteria were applied:</li> <li>The cells expressing fewer than 500 or more than 6000 total genes were filtered out, cells with over 15%</li> <li>mitochondrial genes were excluded and the genes expressed in fewer than three cells were removed.</li> <li>Overall, a total of 60,437 cells were detected and the analysis incorporated 19,125 different genes. The</li> </ul>

107 to eliminate effects of the cell cycle. To identify the principal clusters, the Seurat FindClusters function

108 (resolution = 0.5) was utilized and the top 30 principal component analysis (PCA) components were

selected for further analysis using the gravel plot. The resulting clusters were thereafter visualized using

110 "tSNE" and "UMAP" plots. The FindAllMarkers function of Seurat was employed to identify cell

111 markers for each cluster. To identify major cell types and provide annotations, the CellMarker database

112 (http://xteam.xbio.top) was combined with the data source of the cell surface markers [20].

### 113 InferCNV

- 114 The R software copyKAT package was used for CNV analysis of the hepatocytes present in the samples,
- and the T cells were used as normal reference cells. The normal hepatocytes and HCC cells were
- 116 distinguished according to the changes of CNV.

## 117 Differential gene identification and GSVA (Gene-set Variation Analysis) enrichment analysis in

- 118 the single-cell samples
- 119 A nonparametric and unsupervised algorithm for gene set variation analysis (GSVA) was utilized to
- 120 evaluate the pathway activity in normal hepatocytes, *SH3D21*-positive HCC cells, and *SH3D21*-negative
- 121 HCC cells. The geneset used for GSVA enrichment analysis was Hallmark geneset.

## 122 Pseudotime trajectory analysis

123 The Monocle2 package (v2.8.0) of R software was employed to perform the pseudo-time series analysis

124 of HCC cells [21]. The input for the analysis consisted of the original count matrix obtained from the

125 data processed by Seurat. The new Cell Data Set function was utilized to create an expression family

- 126 object, facilitating the specification of the lower detection limit as 0.5. An unsupervised analysis method
- 127 was employed to explore the cell development trajectory using the highly variable genes selected by
- 128 Monocle. The dispersion empirical parameter was set to 1, whereas the remaining data analysis
- 129 parameters were selected as the default settings.

## 130 Immunotherapy prediction and chemotherapy sensitivity analysis

- 131 Three distinct immunotherapy cohort studies GSE78220, GSE67501 and IMvigor210 were collected
- 132 from GEO database to examine the potential correlation between SH3D21 and immunotherapy. The
- tumor immune atlas database ( The Cancer Immunome Atlas, https://tcia.at/ ) was also used to observe

- the treatment response of *SH3D21* to PD1 and CTLA4. In addition, based on Cancer Drug Sensitivity
- 135 Genomics (GDSC), a public pharmacology portal, the chemotherapeutic agents sensitive to SH3D21
- 136 were estimated and the top 10 drugs were shortlisted for further analysis.
- 137 Test and verify
- 138 *Cell lines and cell culture*
- 139 LX-2, Huh7, HepG2, SK-HEP-1 cell line was provided by CTCC, Ltd (Wuxi, China). The frozen HepG2
- 140 (in Dulbecco's modified Eagle medium (DMEM, high glucose)), Huh-7 (in DMEM, high glucose, and
- 141 SK-HEP-1 (MEM-EBSS) cells were thawed at 37 °C and then resuspended into the corresponding
- 142 complete culture media. When the cells grew to approximately 80%, cell passage was executed.
- 143 Additionally, the cells were grouped and cultured in an incubator (Thermo, USA) at 37 °C. The cells
- 144 were thereafter collected and then stored for subsequent experiments.
- 145 Tissue source
- 146 The HCC tissues and normal hepatic tissue were procured from the different individuals undergoing liver
- 147 cancer resection or liver biopsy at the Department of Hepatobiliary Surgery, Ruikang Hospital Affiliated
- to Guangxi University of Traditional Chinese Medicine between December 2022 and October 2023. The
- 149 studies were approved by the Ethics Committees of Ruikang Hospital Affiliated to Guangxi University
- 150 of Traditional Chinese Medicine. The assigned ethical review approval number: KY2022-045.
- 151 *Real-time, quantitative reverse transcription-polymerase chain reaction (qRT-PCR)*
- 152 Total RNA was extracted by using Trizol reagent (Invitrogen, Grand Island, New York, USA). Both the
- 153 concentration and purity of RNA were measured using a Nanodrop spectrophotometer (ThermoFisher
- 154 Science, Waltham, Massachusetts, USA). RNA was reverse transcribed into cDNA using a reverse
- transcription kit (Takara, Dalian, China) in accordance with the manufacturer's instructions. SYBR

156	Select Master Mix (ThermoFisher Science, Waltham, Massachusetts, USA) was used for qRT-PCR
157	analysis. A QuantStudio <sup>TM</sup> 6 Flex real-time quantitative PCR system was employed for the data
158	collection. The samples obtained amplified in accordance with the following protocol: 95°C 5 mins,
159	95°C 40 cycles 30s, 60°C 40s, 72°C 1 min. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was
160	used as an endogenous control. Each sample was prepared in triplicate. The $2^{-\Delta\Delta CT}$ method was used to
161	calculate the relative expression. The sequences of the qRT-PCR primers used for SH3D21 were: forward:
162	5'-AGCAAGGAGGGCAATGACTCT-3', reverse: 5'-ACGCAGTACTTGCCACTCT-3'; GAPDH:
163	forward: 5'-GCAGCGTGATCCCTGCAAAAT-3', reverse: 5'-GCCTGTAGTCAGCAACTCATCG-3'.
164	Cell transfection
165	Huh7 and HepG2 cells were seeded in six-well culture plate for 24 h ( $1 \times 10^5$ cells/well). Thereafter, the
166	cells were divided into the HepG2 and Huh7 group, the HepG2/Huh7+siRNA NC group and the
167	HepG2/Huh7+si-SH3D21-1, 2, 3 group. HepG2/Huh7+si-SH3D21-1, which was validated by RT-PCR
168	to be the optimized knockdown target sequence, was selected for the subsequent experiments. The DNA
169	sequence of SH3D21 gene was amplified by PCR, digested, and then cloned into the expression vector
170	(PCI-neo). The expression vector was then transfected into HepG2 cells to enable up-expression of the
171	SH3D21 gene; the group was confirmed to be the OE-SH3D21 if the up-expression was validated by RT-
172	PCR. Additionally, Lipofectamine 2000 transfection reagent (1.5 $\mu$ L/mL, ThermoFisher, USA) was used
173	to transfect 50 nM gene fragments or 1 $\mu g$ of plasmid to HepG2 and Huh7 cells. After 48-hour
174	transfection, the samples were collected and stored.
175	Colony formation

176 The samples were divided into the control group, the si-control group and the si-SH3D21 group. After

177 transfection, the cell suspension was prepared once the cell growth was in the logarithmic phase. The

178	cells were then seeded into the six-well plates (Corning, USA) and incubated at 37 $^\circ$ C, 5% CO <sub>2</sub> and
179	saturated humidity for two weeks. The culture was terminated if the clones visible to the naked eyes
180	appeared in the dish. Afterward, the cells were rinsed with PBS twice and fixed with 5 mL of pure
181	methanol (Beijing Dingguo Changsheng Biotechnology Co., Ltd., China) for 15 min. The methanol was
182	then removed, and Giemsa staining solution (Beyotime, China) was added and incubated for 30 min. The
183	staining solution was removed, and the sample was dried in air. The cells were counted manually from
184	obtained images. Alternatively, clones with more than 10 cells were counted using a light microscope
185	(Olympus, Japan). The colony formation rate was calculated by using the formula (number of
186	clones)/(number of inoculated cells) $\times$ 100%.
187	Determination of cell viability
188	Cell Counting Kit-8 (CCK-8) cell proliferation and cytotoxicity assay kit (Sigma-Aldrich, Merck KGaA,
189	Germany) was used to measure the viability of transfected Huh7 and HepG2 cells. The cells were seeded
190	in the wells of a 96-well plate and thereafter incubated for 24h, 48h, 72h, respectively. CCK-8 solution

- 191 (15  $\mu$ L) was added to each well at the designated time intervals and incubated at 37°C for 4h. The
- 192 absorbance at 492 nm was measured using a microplate reader (Thermo Scientific, Waltham,

193 Massachusetts, USA). Each experiment was repeated independently in triplicate.

- 194 Western blot analysis
- 195 HCC cells and tissues were lysed in radioimmunoprecipitation assay (RIPA) buffer (KeyGEN, Nanjing,
- 196 China) containing protease suppressors for 30 min. The protein concentration was quantified using a
- 197 bicinchoninic acid (BCA) kit (KeyGEN, Nanjing, China) in accordance with the manufacturer's protocol.
- 198 The total protein was subjected to heating at 95°C for 5 min. Thereafter, identical quantities of protein
- 199 were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The

200	separated proteins were transferred to polyvinylidene fluoride (PVDF) membranes after which
201	electrometastasis was applied and then blocked with 5% skimmed milk at the room temperature for 60
202	min. The membranes were then stored overnight at 4°C with an appropriate primary antibody. The panel
203	of antibodies used was as follows: SH3D21(1:1000; Proteintech No. 25767-1-AP, Chicago, USA),
204	mTOR (1:1000; Proteintech No. 66888-1-lg, Chicago, USA), p-AKT (1:1000; Proteintech No. 66444-1-
205	lg, Chicago, USA), p-PI3K (1:1000; Proteintech No. 4228, Chicago, USA), PI3K (1:1000; Proteintech
206	No. 20584-1-AP, Chicago, USA), AKT (1:4000; Proteintech No. 60203-2-lg, Chicago, USA), and
207	GAPDH (1:5000; Proteintech No. 60004-1-Ig, Chicago, USA). After washing 3 times in TBST buffer,
208	the membranes were incubated at the room temperature for 1h with rabbit horseradish peroxidase (HRP)-
209	conjugated secondary antibody (1:1000, Beyotime. No. A0208, Shanghai, China) then washed with the
210	blocking solution and visualized by enhanced chemiluminescence (ECL, Thermo Fisher Scientific,
211	Waltham, Massachusetts, USA). Finally, quantity One gel analysis software was used to detect the signal
212	intensity of each membrane. The intensity was measured relative to that of GAPDH.

## 213 Immunohistochemical staining

214 Following the dewaxing and antigen retrieval steps, the tissue sections were subjected to additional 215 treatments. First, incubation with 3% H2O2 (Beyotime, China) for 10 min and 0.1% Triton X-100 216 (Beyotime, China) at the room temperature for 10 min was conducted. Subsequently, 50 µL of 5% BSA 217 (Wuxi Puhe, China) was applied to the tissues and incubated at 37 °C for 1 hour to block non-specific 218 binding sites. The primary antibody, specifically SH3D21 (dilution: 1:1000; Proteintech No. 25767-1-219 AP, Chicago, USA), was proportionally diluted and then applied to the tissues, followed by incubation 220 at 37 °C for 2 h. After washing with PBS, a secondary antibody was added: Rabbit secondary antibody 221 (dilution: 1:200, Zhongshan Jinqiao, China) or Murine dimab (dilution: 1:200, Zhongshanjinqiao, China),

222	depending on the species sources of the primary antibodies used. Thereafter, incubation with the
223	secondary antibody was performed at 37°C for 1h. For visualization, diaminobenzidine (DAB) (Nakasugi
224	Kinbashi, China) color development was employed. Hematoxylin restaining was then conducted,
225	followed by sealing with the neutral gum. Finally, the prepared tissue sections were observed and
226	photographed under a microscope (Olympus, Japan) to evaluate the target staining and overall tissue
227	morphology.

- 228 Invasion and Migration Experiments
- 229 Matrigel matrix glue (BD Biosciences, USA) was diluted in a serum-free medium at 1:6, and 50ul was
- evenly spread into the upper chamber of Transwell chamber (Corning, USA). The chamber was placed
- in a 24-wellplateandincubated at 37°C for 4h to allow it to gelatinize. The cell density was adjusted to
- $2\times 105$  cells/mL, and  $100\mu$ L per well was inoculated into the upper chamber of transwell chamber. The
- cells in the upper chamber were removed after 24h, fixed with methanol and 0.1% crystal violet (Beijing
- 234 Dingguo Changsheng Biotechnology, CHN) for 20min, stained for 10min, and washed twice with PBS.
- 235 The cells were then counted under an inverted optical microscope (Olympus, JPN).
- 236 Scratch Wound-Healing Assay
- Huh7 and HepG2 cells were seeded in 6-well culture plates (Corning, USA) at 95–100% confluence.
- 238 The cells were harvested 48h after transfection. The scratch was created using a 200 µL pipette tip. PBS
- was used to wash the plates thrice to remove the cellular debris. The images for the wound closure were
- captured at the different times (12, 24, 36 and 48 h) by microscope (Olympus, JPN).

## 241 Statistical analysis

- 242 The biological information analysis in this study was conducted by using R programming (version 4.15,
- 243 https://www.r-project.org/). Statistical analysis was performed by SPSS 27 and GraphPad Prism V7.0

244	software. The data has been presented as mean ± standard deviation (SD). For comparison of means
245	between the two groups, T-test was applied. Analysis of variance (ANOVA) was employed to compare
246	multiple sets of values. The relationship between the expression of SH3D21 and clinical features was
247	assessed using the Chi-square test. Survival analysis was performed using the Kaplan-Meier method, and
248	the log-rank test was used for the group comparisons. $P < 0.05$ was considered statistically significant.

#### 249 **3. Results**

## 250 SH3D21 was highly expressed in HCC and related to poor outcomes of HCC patients

251 The differential analysis of HCC transcriptome data from TCGA database (https://portal.gdc.cancer.gov) 252 revealed that the mRNA expression of SH3D21 in HCC tissued was significantly higher than that in 253 normal hepatic tissue (logFC=1.937) (Fig 1A). Moreover, determination of the relationship between 254 SH3D21 expression and clinical characteristics of HCC patients revealed that SH3D21 was highly 255 expressed in females (P=0.0099) and increased with the elevation of tumor T-stage (Fig 1B), histological grading (Fig 1C) and clinical stage (Fig 1D). High expression of SH3D21 resulted in shorter survival of 256 257 HCC patients (Fig 1E). However, no correlation was observed between Progression Free Survival (PFS), 258 Disease Free Interval (DFI), and Disease-Specific Survival (DSS) (P>0.05). Univariate and multivariate 259 analyses were conducted to investigate the various clinical features and the possible relationship between 260 SH3D21 expression and the survival status of patients. High expression of SH3D21 remained a 261 significant independent predictor for patients with HCC (Fig 11, J). In addition, HCC dataset (LIRI-JP) 262 was downloaded for validation from the ICGC (https://dcc.icgc.org/) database. The results revealed that 263 SH3D21 was highly expressed in HCC (logFC= 1.442) (Fig 1F) and increased with both the elevated 264 histological grading (Fig 1G) and clinical stage of tumor tissues (Fig 1H). The above results illustrated 265 that SH3D21 was highly expressed in HCC and was associated with poor clinical outcomes in HCC

- 266 patients. To experimentally validate the predicted results, three HCC-derived samples and normal liver
- tissue-derived samples were collected. Immunohistochemical staining and Western blot analysis to
- determine the expression of *SH3D21* protein in HCC and normal liver tissues was performed. The results
- demonstrated that the level of *SH3D21* protein was significantly higher in the tumor samples compared
- to the normal liver tissues (Fig 1K-M) (Full length gel Supplementary Fig 1). These results confirmed
- the prediction that the expression of *SH3D21* was increased in HCC.



272

Fig: 1 *SH3D21* was highly expressed and associated with poor clinical outcomes in HCC patients. A Differential
analysis of mRNA expression of *SH3D21* in normal control and HCC tissues obtained from TCGA database source
samples. B-D Relationship between the expression level of *SH3D21* and T Stage (B), Grade (C), and stage (D) of

276	HCC patients in TCGA database. E Impact of high expression of SH3D21 on prognosis of HCC patients. F
277	Differential analysis of mRNA expression of SH3D21 in the normal control and HCC tissues from ICGC database
278	source samples. G, H Correlation between SH3D21 expression level and Grade (G) and Stage (H) in ICGC database
279	source samples. I, J Univariate and multifactorial analysis of the clinical characteristics of HCC patients and
280	correlation between SH3D21 gene expression and survival status of patients from TCGA database. K, M The
281	expression of SH3D21 protein in HCC and normal liver tissue samples was detected by immunohistochemical
282	staining and Western blotting.
283	Determination of Mutation as well as methylation status of SH3D21 in HCC and relationship
284	between <i>SH3D21</i> expression and TME
285	To investigate the impact of SH3D21 genetic variation on mRNA expression, the mutation of SH3D21
286	in HCC samples was next analyzed using the online database ccBioPortal ( <u>https://www.cbioportal.org/</u> ).
287	As demonstrated, our analysis revealed that SH3D21 was genetically altered in 6.88% (24/349) of the
288	samples. Among these alterations, the highest frequency was observed in cases of high mRNA expression
289	(20/349, 5.73%), followed by mutations (2/349, 0.57%), amplification, and deep deletion (1/349, 0.29%)
290	(Fig 2A). The most common mutation in the SH3D21 gene was the I428T missense mutation (0.07%),
291	which was diploid in nature. The second most frequent mutation was the M355Cfs*38 frameshift
292	mutation (0.06%), resulting in a mild copy number loss (Fig 2B). Further correlation analysis between
293	the mutation type and mRNA expression of SH3D21 revealed that the increase in mRNA expression
294	could be primarily attributed to gene expansion and significant amplification events (Fig 2C). Therefore,
295	mRNA expression of SH3D21 was negatively related to the methylation level (Fig 2D). Overall, these
296	findings indicated that SH3D21 mRNA expression levels were associated with gene amplification.
297	To explore the possible function of SH3D21 in promoting the clinical progression of HCC, TCGA and

298	ICGC datasets were employed to analyze the correlation between the expression of SH3D21 and the
299	microenvironment characteristics of HCC. The results revealed that increase of SH3D21 expression was
300	associated with EMT2, EMT3, cell cycle, DNA replication, TH1/TH2 cell regulation, MDSC and TIP
301	tumor immune cell infiltration. Therefore, we speculated that there was a significant correlation between
302	the above characteristics and the expression level of SH3D21. In addition, the above characteristics
303	displayed a consistent trend in TCGA (Fig 2E, F) and ICGC data sets (Fig 2G, H), thereby indicating
304	that SH3D21 can play a vital role in regulating the proliferation, invasion, migration, and formation of
305	immunosuppressive microenvironment of HCC.

306



307

308 Fig. 2 Mutation and methylation analysis of SH3D21 in HCC and the relationship between SH3D21 expression
309 and TME features. A The type and frequency of the genetic changes in SH3D21. B SH3D21 mutation type, site, and
310 proportion. C Potential relationship between SH3D21 mRNA expression and mutation. D The relationship between
311 SH3D21 expression and methylation. E Relationship between SH3D21 gene expression and TME
312 characteristics in the TCGA dataset. F Correlation analysis between SH3D21 gene expression

- and TME characteristics in TCGA dataset. G Relationship between SH3D21 gene expression and
- 314 TME characteristics in the ICGC dataset. H Correlation analysis between SH3D21 gene
  315 expression and TME characteristics in ICGC dataset.
- 316 SH3D21 can promote the proliferation, invasion, and migration of HCC cells as confirmed by in
- 317 *vitro* experiments
- 318 Three distinct kinds of HCC cells, SK-HEP-1, Huh-7 and HepG2 were selected to observe the mRNA
- and protein expression of *SH3D21*. The results revealed that the mRNA and protein expression levels of
- 320 *SH3D21* were markedly higher in the three HCC cells in comparison to the control LX-2 cells (Fig 3A,
- B) (Full length gel Supplementary Fig 2). Interestingly, silencing of SH3D21 in HepG2 and Huh-7
- 322 cells resulted in reduced cell viability at 48h and 72h (P<0.01) (Fig 3C, F) and a significant decrease in
- 323 the cell colony forming ability (P<0.01) (Fig 3D, E). Moreover, silencing the expression of SH3D21
- 324 significantly inhibited the wound healing ability (Fig 3G, I), as well as reduced the invasive ability (Fig
- 325 3H, K) and the number of migrating HCC cells (Fig 3H, J). The above results confirmed that
- 326 high expression of SH3D21 can promote the proliferation, migration, and invasion of
- HCC cells.





Fig. 3 *SH3D21* was highly expressed and promoted proliferation, invasion, and migration of HCC cells.
A, B *SH3D21* exhibited higher mRNA and protein expression levels in HCC cell lines than in human
hepatic stellate cells. C, F Silencing of *SH3D21* gene significantly reduced the survival of HepG2 and

Huh-7 cells at 48h and 72h, respectively. **D**, **E** Silencing SH3D21 markedly inhibited colony-forming

- ability of HCC cells. G, I Silencing SH3D21 significantly suppressed wound healing ability of HCC cells.
- **334 H**, **K** Silencing *SH3D2* significantly reduced the invasive ability of HCC cells. **H**, **J** Silencing *SH3D2*
- 335 significantly attenuated the number of migrating HCC cells.

## 336 Bulk sample analysis of TCGA and ICGC databases revealed that SH3D21 can cause

337 activation of PI3K/AKT signaling pathway in HCC

338 HCC samples in the TCGA dataset were divided into high- and low- expression groups 339 according to the median value of SH3D21 expression for the differential analysis, as 340 well as GO (Fig 4A) and KEGG enrichment analysis (Fig 4B) were performed for the 341 differential genes. Biological process (BP) results of GO enrichment analysis revealed 342 that SH3D21 was mainly involved in organic acid metabolic process, carboxylic acid 343 metabolic process and ion transport. Cellular component (CC) results demonstrated that 344 SH3D21 was located on the cell surface, intrinsic component of plasma membrane and 345 integral component of plasma membrane. Molecular function (MF) of SH3D21 included 346 transmembrane transporter activity, oxidoreductase activity and transporter activity. 347 KEGG enrichment analysis revealed that the differential genes were mainly enriched in 348 regulation of retinol metabolism, metabolism of xenobiotics by cytochrome P450 and 349 metabolic pathways. In addition, HCC samples in the TCGA dataset were divided into 350 high- and low- expression groups according to the median value of SH3D21 expression 351 for GSEA pathway enrichment analysis. GSEA enrichment analysis further indicated that 352 the differential genes associated with SH3D21 high-expression group were enriched in 353 cell cycle, DNA replication, mismatch repair, tight linking and PI3K-AKT-mTOR

354 pathway. All these pathways are primarily related to HCC cell proliferation and invasion.

The differential genes associated with *SH3D21* low-expression group were found to be enriched in drug metabolism cytochrome P450, fatty acid metabolism, complement and coagulation cascades and PPAR signaling pathway (Fig 4C).

358 KEGG enrichment analysis of SH3D21-associated differential genes was performed 359 using the ICGC dataset to validate GSEA results from the TCGA dataset. The results 360 suggested that the differential genes associated with SH3D21 high-expression group 361 were mainly enriched in pathways related to the tumor proliferation and invasion, such 362 as cell cycle, PI3K-AKT signaling pathway, DNA replication, mismatch repair, gap 363 linking, and VEGF angiogenesis pathway. The differential genes associated with 364 SH3D21 low-expression group were enriched in complement and coagulation cascades, 365 drug metabolism, arachidonic acid metabolism and PPAR signaling pathway (Fig 4D). 366 The above results were similar to the GSEA enrichment analysis results of TCGA dataset. 367 Among the pathways enriched by the differential genes in SH3D21 high- expression 368 group, PI3K-AKT signaling pathway was identified. The PI3K-AKT signaling pathway 369 can promote multiple signaling processes and regulate a variety of cellular functions, 370 such as metabolism, proliferation, cell survival, growth, and angiogenesis. Activation of 371 PI3K pathway can not only promote cell cycle progression, migration, but also inhibit 372 cellular inflammation and apoptosis. Therefore, the above results indicated that the high 373 expression of SH3D21 has a potential role in activating PI3K signaling cascade.

21



374

Fig. 4 *SH3D21* promoted substantial activation of PI3K/AKT signaling pathway in HCC. A, B GO and
KEGG enrichment analysis of differential genes associated with *SH3D21* high and low
expression groups in TCGA dataset. C GSEA enrichment analysis of differential genes associated with *SH3D21* high and low expression groups in TCGA dataset. D KEGG enrichment analysis of differential

379 genes associated with SH3D21 high and low expression groups in ICGC dataset.

## 380 Single-cell sequencing analysis and *in vitro* experiments revealed that SH3D21 can activate

- 381 PI3K/AKT signaling pathway in HCC cells
- 382 Following data acquisition from Lu *et al.* [22], which involved single-cell sequencing of hepatoma from
- the GEO database, we focused on analyzing the single-cell sequencing results obtained from 10 different
- 384 HCC tissues and 8 para-cancerous control tissues in the dataset. After the batch correction was performed
- using the harmony algorithm to eliminate its potential effect on the cell clustering (Fig 5B). This resulted
- in the formation of distinct 18 cell groups (Fig 5A). We utilized marker genes provided in the source
- 387 literature to perform the cell type annotation (Fig 5C). Subsequently, HCC cells, macrophages,
- 388 monocytes, B cells, T cells and NK cells were selected for further analysis (Fig 5D). Notably, our cell
- 389 type annotation revealed that *SH3D21* was expressed in all the above 6 cell type (Fig 5E).
- 390 The hepatocyte subsets were further analyzed, and cellular CNV analysis was performed on hepatocytes 391 using Copykat (Fig 5F). After the diploid and heteroploidy of hepatocytes were identified, normal 392 hepatocytes and HCC cells were distinguished (Figs 5F-H). The differential analysis between the 393 hepatocytes and HCC cells revealed that SH3D21 was highly expressed in HCC cells (Fig 5I). The 394 hepatocyte subsets were further divided into SH3D21 positive HCC cells, SH3D21 negative HCC cells 395 and normal hepatocytes, and GSVA enrichment analysis was conducted on the above three groups. The 396 results of GSVA enrichment analysis revealed that SH3D21 positive HCC cells exhibited enrichment in 397 the different pathways associated with tumor proliferation, invasion, and immunosuppression, such as 398 PI3K/AKT/mTOR, angiogenesis, EMT, IL6-JAK-STAT3. Interestingly, SH3D21 negative HCC cells 399 exhibited enrichment in various pathways associated with hedgehog signaling, apical junction,
- 400 coagulation, and hypoxia. GSVA enrichment analysis revealed that normal hepatocytes exhibited

401	enrichment in various pathways associated with inflammation and immune response, such as TNFA
402	signaling via NFKB, inflammatory response, complement, P53, apoptosis (Fig 5J). These results
403	indicated that SH3D21 was highly expressed in HCC cells and can potentially activate the
404	PI3K/AKT/mTOR pathway, thus promoting the growth of HCC.
405	Furthermore, <i>in vitro</i> experiments to examine above results. we observed that overexpression of SH3D21
406	led to increased protein levels of mTOR, p-PI3K/PI3K, and p-AKT/AKT in HepG2 and Huh-7 cells
407	(Figs 6A, B) (Full length gel Supplementary Fig 3). Conversely, when SH3D21 was knocked down,
408	reverse effects on the aforementioned proteins in HepG2 and Huh-7 cells were observed (Figs 6C, D)

- 409 (Full length gel Supplementary Fig 4). These findings indicated that SH3D21 could activate
- 410 PI3K/AKT/ mTOR signaling pathway in HCC cells.





412 Fig. 5 Single-cell sequencing revealed that *SH3D21* caused activation of PI3K/AKT signaling pathway
413 in HCC. A For clustering of cells after the batch correction. B For clustering of cells after the batch
414 correction was performed using the harmony algorithm. C Expression of the various cell marker genes
415 in clustering cells. D, E Cell types that were included in subsequent analyses and expression of *SH3D21*416 in various cell types. F-H For clustering of subsets of hepatocytes(F), Copykat was used to perform the

- 417 cellular CNV analysis on hepatocytes(G). Distribution of the normal hepatocytes and HCC cells (H). I
- 418 Difference analysis of *SH3D21* expression between hepatocytes and HCC cells. J GSVA enrichment
- 419 analysis of *SH3D21*-positive, *SH3D21*-negative HCC cells and the control hepatocytes.



420

421 Fig. 6 *In vitro* experiments revealed that *SH3D21* can activate the PI3K/AKT signaling pathway. A, B
422 Protein expression electrophoretograms of key proteins in PI3K/AKT signaling pathway after *SH3D21*423 silencing and overexpression in HepG2 and Huh-7 cells, respectively. C, D Ratio of the gray value of
424 key proteins in PI3K/AKT signaling pathway after *SH3D21* silencing and overexpression in HepG2 and
425 Huh-7 cells.

### 426 Immune landscape of *SH3D21* in HCC and its response to immunotherapy

427 The correlation analysis between *SH3D21* expression and TME based on TCGA database revealed that

- 428 high expression group of *SH3D21* possessed higher immune score than low expression group of *SH3D21*
- 429 (*P* < 0.001) (Fig 7A). This result is also confirmed in ICGC database analysis (Fig 7B). However, the
- 430 analysis results of TCGA and ICGC databases revealed that there was no statistically significant
- 431 correlation observed between the expression level of *SH3D21* and immune scores (*P*>0.05). In addition,

432	TCGA and ICGC datasets also revealed that there was no statistical difference observed in tumor matrix
433	score and tumor purity score between high- and low- expression groups of SH3D21. The CIBERSORT
434	method was employed to analyze the possible relationship between expression of SH3D21 and immune
435	cell infiltration in TCGA data set. The results revealed that high expression of SH3D21 accompanied by
436	M0 immune cell infiltration ( $P \le 0.001$ ) (Fig 7C). Moreover, the high expression of SH3D21 was
437	positively correlated with M0 cell infiltration ( $R=0.5$ , $P=7.1E-08$ ) (Fig 7D). This result was also
438	confirmed in ICGC database analysis (Fig 7E). Based on analysis results of the immune
439	microenvironment characteristics before, it was suggested that SH3D21 could regulate immunity through
440	affecting macrophages, MDSC and Th2 cells.
441	The correlation analysis between SH3D21 and immune checkpoints in TCGA (Fig 7G) and ICGC (Fig
442	7F) datasets demonstrated that high expression of SH3D21 was positively correlated with the expression
443	of immune checkpoint molecule. This observation indicated that high expression of SH3D21 can affect
444	immunosuppression microenvironment of HCC. In addition, the correlation between SH3D21 and the
445	expression level of tumor-related factors MKI67 , TP53 , KRAS, CTNNB1 and IDH1 were also
446	investigated. The analysis based on TCGA (Fig 7H) and ICGC (Fig 7I) datasets revealed that the
447	expression of SH3D21 was positively correlated with MIK67, TP53, KRAS as well as CTNNB1, and
448	the expression of SH3D21 was negatively correlated with IDH1. In addition, the correlation analysis
449	between SH3D21 and the stemness of HCC cells revealed that SH3D21 was negatively correlated with
450	DNAsi of HCC cells (Fig 7H), but there was no significant correlation between SH3D21 and RNAsi.
451	Immunotherapy data in GEO database was further analyzed to observe the intervention effect of
452	immunotherapy on SH3D21. These data included Melanoma Treatment study GSE78220, Renal Cancer
453	Treatment Study GSE67501, Urothelial Cancer Treatment Study IMvigor, and immunotherapy data from

- 454 TCIA database ( https://tcia.at/patients). It was found that none of the above immunotherapies had a
- 455 statistically significant effect on SH3D21 (P > 0.05), thus indicating that SH3D21 was not sensitive to
- 456 immunotherapy. Thereafter, cancer-related drug sensitivity genomics (GDSC, Genomics of Drug
- 457 Sensitivity in Cancer ) were used to screen SH3D21 sensitive drugs. The top 10 sensitive drugs identified
- 458 were AMG-706, ATRA, CEP-701, GSK1904529A, JNK-9L, JQ12, LFM-A13, PFI-1, SB505124 and
- 459 YK4-279.



460

461 Fi. 7 Immune landscape of SH3D21 in HCC. A, B The correlation between expression of SH3D21 and



468 between SH3D21 and MIK67, TP53, KRAS, CTNNB1 and IDH1 based on ICGC dataset.



469

470 Fig. 8 Top 10 sensitive drugs to *SH3D21* in drug sensitivity experiments.

### 471 Alterations in SH3D21 during HCC cell development, construction of predictive models and

472 SH3D21 expression in pan-cancer

473 A mimetic time-series analysis was performed to observe the potential changes in SH3D21 during the

474 development of HCC cells. The results revealed that *SH3D21* was at a relatively stable level during the

475 development of HCC cells and did not show any significant decrease or increase with the progression of

476 HCC (Fig 9A, B).

477	To aid clinicians assess patients' prognoses based on the aforementioned parameters, we established a
478	nomograph of different prognostic factors including SH3D21 expression, patient age, gender,
479	hepatocellular carcinoma stage, grade, and several other variables to predict the OS of the patients (Fig
480	9C). The multivariate ROC curve analysis demonstrated that <i>SH3D21</i> expression possessed predictive
481	value for one-year (AUC=0.764), three-year (AUC=0.750) and five-year (AUC=0.752) survival rates in
482	HCC patients (Fig 9D), the fitting curve revealed that the model had good reliability.
483	Finally, the mRNA expression and clinical data of SH3D21 in 33 tumors were downloaded from the
484	UCSC Xena online database (https://xena.ucsc.edu/) to observe the differential expression of SH3D21 in
485	different tumors and the effect on patient prognosis. We set the number of normal sample cases in the
486	tumors to be greater than or equal to 1. A total of 24 tumors were included in the analysis. Interestingly,
487	the differential expression of SH3D21 was observed in 19 tumors in comparison to the normal tissues.
488	The expression of <i>SH3D21</i> was elevated in 17 tumors ( $P < 0.05$ ), decreased in 2 tumors ( $P < 0.05$ ), and did
489	not differ in the remaining 5 tumors (P>0.05) (Fig 9F). We divided SH3D21 into two groups, high and
490	low, according to the median value of SH3D21 expression in 19 tumors, and observed the effect of
491	different expression on patients' prognosis. High SH3D21 expression could lead to poor prognosis in
492	patients with HCC in addition to kidney renal clear cell carcinoma (KIRC) (Fig 9G), whereas high
493	SH3D21 expression could prolong patients' survival in head and neck squamous cell carcinoma (HNSC)
494	(Fig 9I) and breast invasive carcinoma (BRCA) (Fig 9H). The results of a further univariate analysis
495	indicated that SH3D21 could be used as an influential factor affecting patient prognosis in LIHC, KIRC,
496	BRCA and HNSC tumor types (Fig 9E). This result suggested that the function of SH3D21 in various
497	tumors varies depending on the tumor type.



498

499 Fig. 9 Alterations of SH3D21 during HCC development, construction of predictive models and SH3D21

expression in pan-cancer. A, B Proposed temporal analysis of *SH3D21* in development of HCC. C
Nomograph of prognostic factors predicting OS of patients based on multiple variables (including: *SH3D21* expression, patient age, gender, hepatocellular carcinoma stage, grade). D ROC curve
construction of expression level of *SH3D21* and its predictive value for survival of HCC. E Univariate
analysis of *SH3D21* on the prognosis of 33 tumor patients. F Differential analysis of *SH3D21* expression

505 in 24 tumors. G-I Effect of high *SH3D21* expression on the prognosis of patients with KIRC (G), BRCA
506 (H) and HNSC (I).

### 507 4. Discussion

- 508 In this study, the role of *SH3D21* in HCC was investigated in detail. The results revealed that expression
- 509 level of SH3D21 was increased in HCC and associated with poor prognosis of HCC patients. SH3D21
- 510 was found to promote the proliferation, invasion, and migration of HCC cells, and can activate
- 511 PI3K/AKT/ mTOR signaling pathway.

512 SH3D21 gene is located at 1p34.3 locus and encodes a protein with a molecular weight of 71 kDa. In

513 terms of subcellular localization, *SH3D21* is found in both the nucleus and the cytoplasmic membrane.

514 Functionally, *SH3D21* can function as a signaling molecule within the cells by forming complexes with

515 other proteins. For instance, in a previous study by Jacklyn N. Hellwege et al., it was discovered that two

516 single nucleotide polymorphisms within the SH3D21 gene were associated with Resting Metabolic Rates

517 (RMR) based on exome array analysis [15]. Moreover, in another study, Mohammad Masoudi et al.

518 utilized the Pickles database and identified the requirement of SH3D21 gene for maintaining the survival

519 of NCIH526 lung cancer cell line. Furthermore, in the context of pancreatic cancer research, SH3D21

520 has been identified as a novel sensitizer to gemcitabine treatment [23]. Despite these findings, our

521 understanding of the exact role played by SH3D21 in tumorigenesis remains relatively limited. Hence,

522 further research is warranted to comprehensively elucidate the potential involvement of *SH3D21* in HCC.

523 In the analysis of HCC samples from the TCGA database, SH3D21 was found to be highly expressed in

- 524 HCC, and the increased expression of SH3D21 was positively correlated with the histological grade,
- 525 clinical stage, and T stage of HCC, thereby suggesting that SH3D21 has a key role in promoting the
- 526 clinical progression of HCC. This conclusion was also confirmed by the analysis of the HCC dataset

527 from the ICGC database. The IOBR package was further employed to analyze the correlation between 528 SH3D21 and TME characteristics. TCGA and ICGC database analysis consistently revealed that high-529 expression of SH3D21 was associated with EMT, cell cycle, DNA replication, TH1, TH2 cell regulation, 530 MDSC and TIP immune cell infiltration. These results indicated that SH3D21 can effectively promote 531 the clinical progression of HCC by stimulating the proliferation and invasion of HCC cells and generating 532 tumor immunosuppressive microenvironment. To further investigate the cell types that SH3D21 533 specifically acts on, HCC single-cell sequencing dataset from the GEO database was employed and by 534 analyzing the expression of SH3D21 in different types of cells, the primary cellular role of SH3D21 was 535 determined. The results revealed that SH3D21 was expressed on HCC cells, macrophages, monocytes, 536 B cells, T cells and NK cells. This indicated that SH3D21 had important functional effects on all the 537 above cells. In this study, only HCC cells were selected for follow-up research. A single-cell dataset from the GEO database revealed that SH3D21 was expressed at higher levels in 538 539 HCC samples in comparison to the normal hepatocytes. The role of SH3D21 in promoting the 540 proliferation and invasion of HCC cells was further verified by CCK8, cloning, invasion, and migration 541 experiments. In order to further explore the specific mechanism of SH3D21 promoting the progression 542 of HCC cells, GSEA analysis results of TCGA database, KEGG analysis results of ICGC database and 543 GSVA analysis results of single cell sequencing dataset were used. The results of enrichment analysis 544 revealed that cell cycle, PI3K-Akt, DNA replication, mismatch repair, gap link, and VEGF angiogenesis 545 pathways were significantly enriched in all the above three datasets. However, the potential effect of 546 SH3D21 on activation of PI3K-Akt signaling pathway was the primary focus of this study. The PI3K-547 Akt signaling pathway has been implicated in the regulation of cell proliferation, invasion, apoptosis 548 signaling processes. Activation of PI3K pathway can not only promote the cell cycle progression and

549 tumor cell migration, but also inhibit immune cells, thereby constructing a tumor immunosuppressive550 microenvironment.

551	Activation of the PI3K/AKT/mTOR signaling pathway can occur through diverse mechanisms, including
552	AKT-induced phosphorylation of apoptosis and cell cycle blocking factors such as BAD (BCL2-related
553	cell death agonist), MDM2 (ubiquitin ligase of p53), FOXO (apoptosis-inducing transcription factor),
554	IKK (inhibitor of kappa B kinase), and IkB (IkB kinase, which acts as a inhibitor of the transcription
555	factor NF- $\kappa$ B). Interestingly, targeted inhibition of these factors can lead to enhanced activity of cell
556	cycle, DNA replication, and other associated signaling pathways [24, 25]. In addition, due to metabolic
557	reprogramming in tumors, abnormal activation of PI3K/AKT/mTOR pathway can promote the
558	immunosuppressive response [26], which in turn can lead to the formation of the immunosuppressive
559	network of tumors [27]. PI3K/AKT/ mTOR signaling pathway also plays a role in regulating M1/M2
560	polarization of macrophages [28], similar to the tumor-associated macrophage (TAM) switch between
561	immune stimulation and immune suppression [29]. In addition, PI3K/AKT/mTOR signaling pathway
562	also plays a significant role in promoting glycolytic metabolism in HCC cells, which can cause the
563	abnormal accumulation of lactic acid. As a product of glycolysis, lactic acid produced by the tumor cells
564	can promote M2-like polarization of tumor-associated macrophages [30-32]. Moreover, M2-like
565	polarization of tumor-associated macrophages can inhibit anti-tumor immunity, stimulate tumor
566	angiogenesis, enhance tumor cell invasion as well as infiltration, and further promote the tumor growth
567	[33-35], thus forming a vicious cycle. In this study, the activation of PI3K/AKT/ mTOR signaling
568	pathway by high expression of SH3D21 was confirmed by further experiments.

569 Considering that there are few studies reported on potential role of SH3D21 in tumorigenesis, the

570 expression of *SH3D21* in pan-cancer and its influence on pan-cancer were also analyzed. Among the 33

571	different	tumor	types	investigated,	SH3D21	exhibited	altered	expression	in	19	tumor	tissues	in
-----	-----------	-------	-------	---------------	--------	-----------	---------	------------	----	----	-------	---------	----

- 572 comparison to the normal tissues. Specifically, *SH3D21* expression was increased in 17 tumors (*P*<0.05)
- 573 and decreased in 2 tumors (*P*<0.05). Interestingly, although *SH3D21* expression was elevated in certain
- 574 tumors such as BLCA, ESCA, and READ, it did not significantly impact the overall patient survival
- 575 (P>0.05). These observations indicate that the influence of SH3D21 might significantly vary across
- 576 different tumor types.

## 577 **5.** Conclusion

- 578 In summary, *SH3D21* can display elevated expression levels in various tumor types, including primary
- 579 HCC. Notably, SH3D21 can promote the proliferation and invasion of HCC by activating
- 580 PI3K/AKT/mTOR signaling pathway. Consequently, SH3D21 could serve as a promising candidate for
- 581 targeted intervention in hepatoma treatment.

## 582 Supporting information

- 583 S1 Raw images (PDF)
- 584 S2 Raw images (PDF)
- 585 Author contributions
- 586 Conceptualization: Ning Luo
- 587 Data curation: Wangxia Tong, Tao Lu
- 588 Formal analysis: Li Liu, Jibing Chen
- 589 Methodology: Rong Liu, Li Liu, Wangxia Tong, Tao Lu
- 590 Resources: Wangxia Tong, Ning Luo and Jibing Chen
- 591 Supervision: Jibing Chen
- 592 Validation: Jibing Chen

- 593 Writing-original draft: Wangxia Tong
- 594 Writing-review & editing: Wangxia Tong, Ning Luo
- 595 Acknowledgments
- 596 Not applicable
- 597 Consent for publication
- 598 This article has been authorized by all authors and agreed to be published.

## 599 Ethical approval

- 600 The studies involving human participants were reviewed and approved by the Ethics Committees of
- 601 Ruikang Hospital Affiliated to Guangxi University of Chinese Medicine. The assigned ethical review
- 602 approval number: KY2022-045. The atients/participants provided their written informed consent to
- 603 participate in this study.
- 604

## 605 References

- 606 1. Villanueva A. Hepatocellular Carcinoma. New Engl J Med. 2019;380(15):1450-62.
   607 http://doi.org/10.1056/NEJMra1713263 PMID:30970190
- Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet. 2018;391(10127):1301-14.
   http://doi.org/10.1016/S0140-6736(18)30010-2 PMID:29307467
- 610 3. Wong KM, King GG, Harris WP. The Treatment Landscape of Advanced Hepatocellular
  611 Carcinoma. Curr Oncol Rep. 2022;24(7):917-27. <u>http://doi.org/10.1007/s11912-022-01247-7</u>
  612 PMID:35347594 PMID:35347594
- 613 4. Hartke J, Johnson M, Ghabril M. The diagnosis and treatment of hepatocellular carcinoma. Semin
- 614 Diagn Pathol. 2017;34(2):153-9. <u>http://doi.org/10.1053/j.semdp.2016.12.011</u> PMID:28108047
- 615 5. Kelley RK, Rimassa L, Cheng AL, Kaseb A, Qin S, Zhu AX, et al. Cabozantinib plus atezolizumab
- 616 versus sorafenib for advanced hepatocellular carcinoma (COSMIC-312): a multicentre, open-label,
- 617 randomised, phase 3 trial. Lancet Oncol. 2022;23(8):995-1008. <u>http://doi.org/10.1016/S1470-</u>
- 618 <u>2045(22)00326-6</u> PMID:35798016

- 6. Yau T, Park JW, Finn RS, Cheng AL, Mathurin P, Edeline J, et al. Nivolumab versus sorafenib in
- 620 advanced hepatocellular carcinoma (CheckMate 459): a randomised, multicentre, open-label,
- 621 phase 3 trial. Lancet Oncol. 2022;23(1):77-90. <u>http://doi.org/10.1016/S1470-2045(21)00604-5</u>
- 622 PMID:34914889
- 623 7. McGlynn KA, Petrick JL, El-Serag HB. Epidemiology of Hepatocellular Carcinoma. Hepatology.
- 624 2021;73 Suppl 1(Suppl 1):4-13. <u>http://doi.org/10.1002/hep.31288</u> PMID:32319693
- Manning BD, Toker A. AKT/PKB Signaling: Navigating the Network. Cell. 2017;169(3):381-405.
   http://doi.org/10.1016/j.cell.2017.04.001 PMID:28431241
- 627 9. Sever R, Brugge JS. Signal transduction in cancer. Csh Perspect Med. 2015;5(4).
   628 <u>http://doi.org/10.1101/cshperspect.a006098</u> PMID:25833940
- 629 10. Sun EJ, Wankell M, Palamuthusingam P, McFarlane C, Hebbard L. Targeting the PI3K/Akt/mTOR
  630 Pathway in Hepatocellular Carcinoma. Biomedicines. 2021;9(11).
- 631 <u>http://doi.org/10.3390/biomedicines9111639</u> PMID:34829868
- 632 11. Porta C, Paglino C, Mosca A. Targeting PI3K/Akt/mTOR Signaling in Cancer. Front Oncol.
  633 2014;4:64. http://doi.org/10.3389/fonc.2014.00064 PMID:24782981
- 12. Yu XN, Chen H, Liu TT, Wu J, Zhu JM, Shen XZ. Targeting the mTOR regulatory network in
  hepatocellular carcinoma: Are we making headway? Bba-Rev Cancer. 2019;1871(2):379-91.
  http://doi.org/10.1016/j.bbcan.2019.03.001 PMID:30951815
- 13. Yang C, Zhang H, Zhang L, Zhu AX, Bernards R, Qin W, et al. Evolving therapeutic landscape of
  advanced hepatocellular carcinoma. Nat Rev Gastro Hepat. 2023;20(4):203-22.
  http://doi.org/10.1038/s41575-022-00704-9 PMID:36369487
- 640 14. Tian LY, Smit DJ, Jucker M. The Role of PI3K/AKT/mTOR Signaling in Hepatocellular
  641 Carcinoma Metabolism. Int J Mol Sci. 2023;24(3). <u>http://doi.org/10.3390/ijms24032652</u> PMID:
  642 36768977
- 643 15. Hellwege JN, Velez ED, Acra S, Chen K, Buchowski MS, Edwards TL. Association of gene coding
  644 variation and resting metabolic rate in a multi-ethnic sample of children and adults. Bmc Obes.
  645 2017;4:12. http://doi.org/10.1186/s40608-017-0145-5 PMID:28417008
- 646 16. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics647 portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov.

648 2012;2(5):401-4. <u>http://doi.org/10.1158/2159-8290.CD-12-0095</u> PMID:22588877

- 649 17. Goldman MJ, Craft B, Hastie M, Repecka K, McDade F, Kamath A, et al. Visualizing and
  650 interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38(6):675-8.
  651 http://doi.org/10.1038/s41587-020-0546-8 PMID:32444850
- 18. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with
  revamped UIs and APIs. Nucleic Acids Res. 2019;47(W1): W199-205.
  http://doi.org/10.1093/nar/gkz401 PMID:31114916
- 19. Slovin S, Carissimo A, Panariello F, Grimaldi A, Bouche V, Gambardella G, et al. Single-Cell RNA
  Sequencing Analysis: A Step-by-Step Overview. Methods Mol Biol. 2021; 2284:343-65.
  http://doi.org/10.1007/978-1-0716-1307-8 19 PMID:33835452
- 658 20. Hu C, Li T, Xu Y, Zhang X, Li F, Bai J, et al. CellMarker 2.0: an updated database of manually
  659 curated cell markers in human/mouse and web tools based on scRNA-seq data. Nucleic Acids Res.
  660 2023;51(D1):D870-6. <u>http://doi.org/10.1093/nar/gkac947</u> PMID:36300619
- 21. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves
  complex single-cell trajectories. Nat Methods. 2017;14(10):979-82.
  http://doi.org/10.1038/nmeth.4402 PMID:28825705
- Lu Y, Yang A, Quan C, Pan Y, Zhang H, Li Y, et al. A single-cell atlas of the multicellular
  ecosystem of primary and metastatic hepatocellular carcinoma. Nat Commun. 2022;13(1):4594.
  http://doi.org/10.1038/s41467-022-32283-3 PMID:35933472
- 667 23. Masoudi M, Seki M, Yazdanparast R, Yachie N, Aburatani H. A genome-scale CRISPR/Cas9
  668 knockout screening reveals SH3D21 as a sensitizer for gemcitabine. Sci Rep-Uk. 2019;9(1):19188.
  669 http://doi.org/10.1038/s41598-019-55893-2 PMID:31844142
- 670 24. Mossmann D, Park S, Hall MN. mTOR signalling and cellular metabolism are mutual determinants
  671 in cancer. Nat Rev Cancer. 2018;18(12):744-57. <u>http://doi.org/10.1038/s41568-018-0074-8</u>
  672 PMID:30425336
- 673 25. Giannone G, Ghisoni E, Genta S, Scotto G, Tuninetti V, Turinetto M, et al. Immuno-Metabolism
  674 and Microenvironment in Cancer: Key Players for Immunotherapy. Int J Mol Sci. 2020;21(12).
- 675 <u>http://doi.org/10.3390/ijms21124414</u> PMID:32575899
- 676 26. Okkenhaug K. Signaling by the phosphoinositide 3-kinase family in immune cells. Annu Rev

- 677
   Immunol.
   2013;31:675-704.
   <a href="http://doi.org/10.1146/annurev-immunol-032712-095946">http://doi.org/10.1146/annurev-immunol-032712-095946</a>
- 678 PMID:23330955
- 679 27. Speiser DE, Ho PC, Verdeil G. Regulatory circuits of T cell function in cancer. Nat Rev Immunol.
  680 2016;16(10):599-611. http://doi.org/10.1038/nri.2016.80 PMID:27526640
- 681 28. Tsai CF, Chen GW, Chen YC, Shen CK, Lu DY, Yang LY, et al. Regulatory Effects of Quercetin
- on M1/M2 Macrophage Polarization and Oxidative/Antioxidative Balance. Nutrients. 2021;14(1).
   http://doi.org/10.3390/nu14010067 PMID:35010945
- 29. Zhang Q, Lou Y, Bai XL, Liang TB. Immunometabolism: A novel perspective of liver cancer
  microenvironment and its influence on tumor progression. World J Gastroentero.
  2018;24(31):3500-12. http://doi.org/10.3748/wjg.v24.i31.3500 PMID:30131656
- 687 30. Kaneda MM, Messer KS, Ralainirina N, Li H, Leem CJ, Gorjestani S, et al. PI3Kgamma is a
  688 molecular switch that controls immune suppression. Nature. 2016;539(7629):437-42.
  689 <u>http://doi.org/10.1038/nature19834</u> PMID:27642729
- 690 31. Wu X, Chen H, Wang Y, Gu Y. Akt2 Affects Periodontal Inflammation via Altering the M1/M2
  691 Ratio. J Dent Res. 2020;99(5):577-87. http://doi.org/10.1177/0022034520910127 PMID:32228353
- 692 32. Vergadi E, Ieronymaki E, Lyroni K, Vaporidi K, Tsatsanis C. Akt Signaling Pathway in
  693 Macrophage Activation and M1/M2 Polarization. J Immunol. 2017;198(3):1006-14.
  694 http://doi.org/10.4049/jimmunol.1601515 PMID:28115590
- 695 33. Colegio OR, Chu NQ, Szabo AL, Chu T, Rhebergen AM, Jairam V, et al. Functional polarization
- of tumour-associated macrophages by tumour-derived lactic acid. Nature. 2014;513(7519):559-63.
   <a href="http://doi.org/10.1038/nature13490">http://doi.org/10.1038/nature13490</a> PMID:25043024
- Sunakawa Y, Stintzing S, Cao S, Heinemann V, Cremolini C, Falcone A, et al. Variations in genes
   regulating tumor-associated macrophages (TAMs) to predict outcomes of bevacizumab-based
   treatment in patients with metastatic colorectal cancer: results from TRIBE and FIRE3 trials. Ann
- 701 Oncol. 2015;26(12):2450-6. <u>http://doi.org/10.1093/annonc/mdv474</u> PMID:26416897
- Xiang X, Wang J, Lu D, Xu X. Targeting tumor-associated macrophages to synergize tumor
  immunotherapy. Signal Transduct Tar. 2021;6(1):75. <u>http://doi.org/10.1038/s41392-021-00484-9</u>
  PMID: 33619259
- 705