Pakistan J. Zool., vol. 55(4), pp 1553-1563, 2023 DOI: https://dx.doi.org/10.17582/journal.pjz/20211129131153

Identification of SET7/9-E2F1 as Novel Therapeutic Biomarkers in Hepatocellular Carcinoma

Lu Xie[1,2,3,4,5], Ye Gu[1,2,3,4], Qiang Liu[1,2,3,4], Hongzhang Shen[1,2,3,4], Yifeng Zhou[1,2,3,4], Jiangfeng Yang[1,2,3,4], Xiaofeng Zhang[1,2,3,4]* and Jinyu Huang[1,5]*

[1]The Affiliated Hangzhou Hospital of Nanjing Medical University
[2]Department of Gastroenterology, Key Laboratory of Clinical Cancer Pharmacology and Toxicology Research of Zhejiang Province, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310006, P.R. China.
[3]Hangzhou Hospital and Institute of Digestive Diseases, Hangzhou, Zhejiang 310006, P.R. China.
[4]Key Laboratory of Integrated Traditional Chinese and Western Medicine for Biliary and Pancreatic Diseases of Zhejiang Province, Hangzhou, Zhejiang 310006, P.R. China.
[5]Department of Cardiology, Key Laboratory of Clinical Cancer Pharmacology and Toxicology Research of Zhejiang Province, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310006, P.R. China.

Article Information
Received 29 November 2021
Revised 25 January 2022
Accepted 12 February 2022
Available online 06 June 2022
(early access)
Published 29 May 2023

Authors' Contribution
JH and XZ contributed to the initial design of the study. LX and YG prepared the manuscript. QL, HS, YZ and JY conducted bioinformatics analyses.

Key words
SET7/9, E2F1, Hepatocellular carcinoma, pathway

ABSTRACT

Our previous studies have shown that SET7/9 promotes hepatocellular carcinoma cells proliferation, invasion and migration via post-translational regulation of E2F1. In this study, we comprehensively analyzed the functions and mechanisms of the SET7/9-E2F1 axis using data mining. Data from the UALCAN database showed abnormal expression of both SET7/9 and E2F1 in multiple cancer types. Survival curves and correlation analysis by GEPIA supported the significant roles of SET7/9 and E2F1 in the progression of HCC. Functional enrichment analysis suggested that the SET7/9-E2F1 axis is involved in the regulation of cell cycle, DNA repair and replication, and gene transcription. Our results implicated the potential of SET7/9 in combination with E2F1 as novel therapeutic targets and prognostic biomarkers in hepatocellular carcinoma.

INTRODUCTION

Hepatocellular carcinoma (HCC) is a highly aggressive malignancy, which carries a 5-year survival rate of approximately 18% (Siegel *et al*., 2020). Surgical resection, transplantation, and radiofrequency ablation (RFA) are effective therapies for HCC at early stage (Yu, 2016). Uultrasound (US) and serum α-fetoprotein (AFP) are formally recommended screening and surveillant tools for HCC. However, the sensitivity and specificity of US and AFP can be influenced by several limitations, such as lesion size or different setting of cutoff values (Sauzay et al., 2016). Since most clinical cases are first diagnosed at an advanced stage, patient prognosis is extremely poor and symptomatic management is the only appropriate choice (Kulik and El-Sareg, 2019; Heimbach *et al*., 2018; Bruix *et al*., 2016; Colagrande *et al*., 2016; Chacko *et al*., 2016). Therefore, continued efforts are needed to improve the survival of HCC through development of new biomarkers.

Crosstalk between lysine methylation and other posttranslational modifications is crucial for HCC development. As a lysine methyltransferase, SET7/9 plays a prominent role in transcriptional gene regulation and epigenetic inheritance of histone and non-histone proteins (Pradhan *et al*., 2009). The potential functions

* Corresponding authors: zxf837@tom.com, huangjyls@163.com
0030-9923/2023/0004-1553 $ 9.00/0

Abbreviations
AFP, serum α-fetoprotein; HCC, hepatocellular carcinoma; Leading Edge Num, the number of leading edge genes; RFA, radiofrequency ablation; US, ultrasound.

of SET7/9 include gene expression regulation and chromatin architecture maintenance. Recent advances in understanding the molecular mechanisms of tumor genesis and progression have suggested that SET7/9 participates in multiple malignant processes in cancer (Si *et al*., 2020; Akiyama *et al*., 2016; Shen *et al*., 2015). Notably, Chen *et al*. (2016) showed that SET7/9 regulated tumor cell growth, which might be associated with HCC occurrence and progression (Chen *et al*., 2016). Our previous studies have also demonstrated that the expression levels of SET7/9 and E2F transcription factor 1 (E2F1) were up-regulated in HCC and were correlated with the pathological stage and lesion size in 68 clinical samples from HCC patients (Gu *et al*., 2018). Overexpression of SET7/9 promoted HCC cells proliferation, invasion and migration via post-translational regulation of E2F1 (Gu *et al*., 2018). However, there was still little research about the specific impacts of SET7/9 in the cellular regulatory system and relevant molecular mechanisms in HCC.

Hence, in this study, we re-evaluated the functions and mechanisms of the SET7/9-E2F1 axis through comprehensive bioinformatics analyses, which may provide potential significance for SET7/9-E2F1 as novel therapeutic target and prognostic biomarker in HCC.

MATERIALS AND METHODS

UALCAN

UALCAN (http://ualcan.path.uab.edu/index.html) is a comprehensive web resource for analyzing cancer genomics data, which provides easy access to The Cancer Genome Atlas (TCGA) and clinical data (Chandrashekar *et al*., 2017). To detect *SET7/9* and *E2F1* expression in cancer in more detail, we examined their expression pattern in pan-cancers according to TCGA database. Student *t* test was used to generate the adjusted *p* value after FDR (false discovery rate) correction. $p < 0.05$ was considered as statistically significant.

GEPIA

GEPIA (http://gepia.cancer-pku.cn/index.html) is an online analysis tool for easily exploring the TCGA and GTEx (Genotype-Tissue Expression) datasets (Tang *et al*., 2017). In this study, we analyzed the potential association between expression of SET7/9/E2F1 and patient survival and conducted gene correlation analysis of SET7/9 and E2F1 in HCC. Kaplan-Meier survival curves were used to assess the association between *SET7/9* and *E2F1* expression and overall survival rate in HCC. All the enrolled samples from HCC patients were categorized into high and low-expressed groups based on the median of *SET7/9* and *E2F1* expression levels. $p < 0.05$ was considered as statistically significant.

cBioPortal

cBioPortal (http://www.cbioportal.org/) is a web resource for exploring, visualizing, and analyzing multidimensional cancer genomics data, which integrates comprehensive research projects such as TCGA and ICGC and covers more than 28,000 clinical tumor specimens (Gao *et al*., 2013). A dataset involving 9,896 samples from 32 TCGA pan-cancer studies was used to explore the frequencies of genetic alteration of *SET7/9* and *E2F1* in various cancer types. Analysis of genetic alterations of *SET7/9* and *E2F1* in HCC was conducted based on a dataset of 366 TCGA HCC samples. The mRNA expression *z* scores (RNA Seq V2 RSEM) of both genes were obtained using a threshold of ±2.0.

STRING

The STRING database (https://string-db.org/) is a comprehensive and objective global network for collecting, scoring and integrating all publicly available sources of protein-protein interaction (PPI) information, and complementing these with computational predictions (Szklarczyk *et al*., 2019). In this study, we constructed a full STRING PPI network using SET7/9 and E2F1 as the query proteins. Protein interactors of SET7/9 and E2F1 with medium confidence interaction score (0.400) were presented in the network.

GeneMANIA

GeneMANIA (http://www.genemania.org) is an effective tool for in-depth analysis of a set of input genes, including protein and genetic interactions, pathways, co-expression, co-localization and protein domain similarity (Warde-Farley *et al*., 2010). In this study, we generated a gene network centered on SET7/9 and E2F1 for a better understanding of the functions of genes correlated with SET7/9 and E2F1.

Metascape

Metascape (https://metascape.org) is an effective and efficient tool for comprehensively integration of a broad set of biological databases (Zhou *et al*., 2019). In this study, we used Metascape for further enrichment analyses of SET7/9- and E2F1-correlated neighbor genes identified in the STRING database. GO and KEGG terms with a *p* value < 0.01, a minimum count of 3, and an enrichment factor > 1.5 were collected and grouped into clusters based on their membership similarities and visualized using Cytoscape. The Molecular Complex Detection (MCODE) plugin implemented in Cytospace was used for clustering analysis to identify highly interconnected nodes in PPI network. Three best-scoring terms were applied to each mCODE component independently.

LinkedOmics

The LinkedOmics database (http://www.linkedomics.org/) contains datasets of 32 different cancer types from TCGA (Vasaikar *et al*., 2018). In this study, a Pearson test was used to analyze the correlation between input genes (*SET7/9* and *E2F1*) and other differentially expressed genes in HCC. Genes showing an absolute value of log FC > 1 as the cutoff standard and p <0.05 as the statistical significance were considered as differentially expressed genes. The "LinkInterpreter" module was used to further explore the possible kinase, miRNA and transcription factor targets of SET7/9 and E2F1.

RESULTS

Expression levels and survival curves of SET7/9 *and* E2F1 *in HCC patients*

We first explored the abnormal expression of *SET7/9* and *E2F1* in tumors and normal tissues using UALCAN. *SET7/9* was found to significantly up-regulated in colon adenocarcinoma, kidney renal clear cell carcinoma, liver hepatocellular carcinoma, and lung adenocarcinoma, and down-regulated in bladder urothelial carcinoma, breast invasive carcinoma, cervical squamous cell carcinoma, rectum adenocarcinoma, and uterine corpus endometrial carcinoma. Meanwhile, *E2F1* was significantly up-regulated in almost all the cancer types except for glioblastoma multiforme, prostate adenocarcinoma, pheochromocytoma and paraganglioma, sarcoma, skin cutaneous melanoma, and thymoma, in which higher *E2F1* expression in tumor tissues was also detected (Fig. 1A). Consistent with our previous studies (Gu *et al*., 2018), the transcriptional levels of *SET7/9* (p = 1.16E-2) and *E2F1* (p = 1.62E-12) in HCC tissues were both significantly elevated (Fig. 1B). In addition, the expression level of E2F1 was positively correlated with HCC progression in terms of nodal metastasis and tumor grade with statistical significance (Fig. 1C).

We then assessed the effects of high- and low-expression of *SET7/9* and *E2F1* on disease-free survival and overall survival of HCC patients with GEPIA (Fig. 1B, C). As expected, HCC patients with low transcriptional levels of *SET7/9* (p = 0.013) and *E2F1* (p = 0.018) were associated with longer disease-free survival (Fig. 2A). Despite that the transcriptional level of *SET7/9* (p= 0.56) was not significantly associated with overall survival, the overall trend of SET7/9-related survival curve was similar with *E2F1*-related survival curve (p = 0.035) (Fig. 2B). These data suggest that aberrant expressions of *SET7/9* and *E2F1* may play critical roles in the progression of HCC.

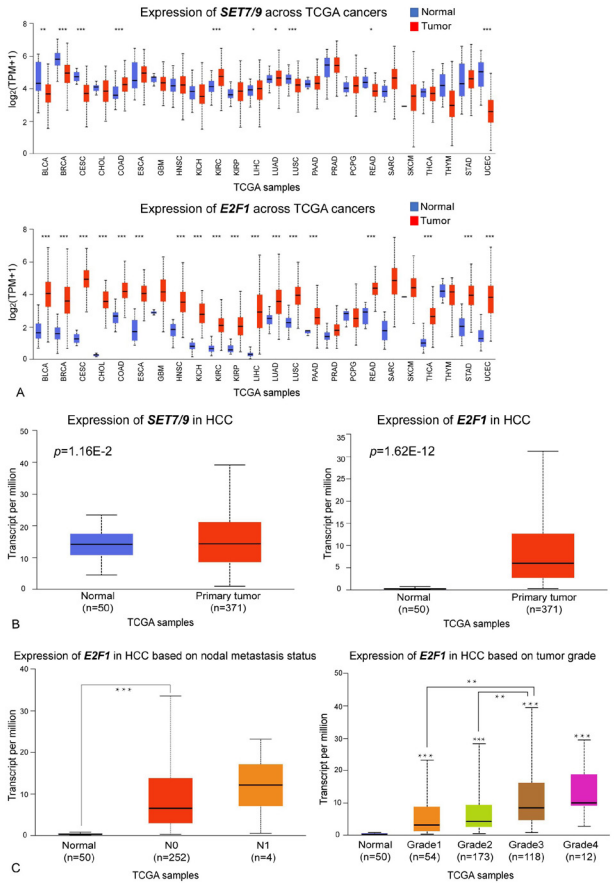

Fig. 1. Expression analyses of *SET7/9* and *E2F1* in tumor and normal tissues. (A) The transcriptional levels of *SET7/9* and *E2F1* in tumor and normal tissues of different cancer types. (B) The transcriptional levels of SET7/9 and E2F1 in HCC tumor and normal tissues. (C) Correlation between the expression level of E2F1 and tumor grade and metastasis status of clinical HCC samples. The asterisks on each bar in the right panel indicate the statistical significance between samples with different levels of tumor grade and normal sample. *p<0.05, **p<0.01. ***p<0.001.

Genetic alternations of SET7/9 *and* E2F1 *in HCC*

Given the significantly differential expression pattern of *SET7/9* and *E2F1* in HCC, we also analyzed the genetic alterations of *SET7/9* and *E2F1* using the cBioPortal database. The frequencies of genetic alteration were firstly explored based on a large patient cohort of 9,896 clinical samples with different cancer types. The alteration frequencies of *SET7/9* and *E2F1* ranged from 0.53% (acute myeloid leukemia) to 12.57% (uterine corpus endometrial carcinoma) in various cancer types. Generally, genetic mutation, gene amplification and deletion were most frequently occurred forms of genetic alteration, while

structural variation was detected in brain lower grade glioma, prostate adenocarcinoma, and sarcoma (Fig. 3A). Next, changes in genetic feature of *SET7/9* and *E2F1* were specifically examined based on a dataset of 366 TCGA HCC samples. Genetic alteration of *SET7/9* was detect in 18 cases (4.92%), including amplification in one case (0.27%), deep deletion in one case (0.27%), multiple alterations in one case (0.27%), mRNA up-regulation in nine cases (2.46%) and mRNA down-regulation in six cases (1.64%). By comparison, genetic alteration of *E2F1* was detect in 17 cases (4.64%), including genetic mutation in three cases (0.82%), amplification in three cases (0.82%), and mRNA up-regulation in 11 cases (3.01%) (Fig. 3B).

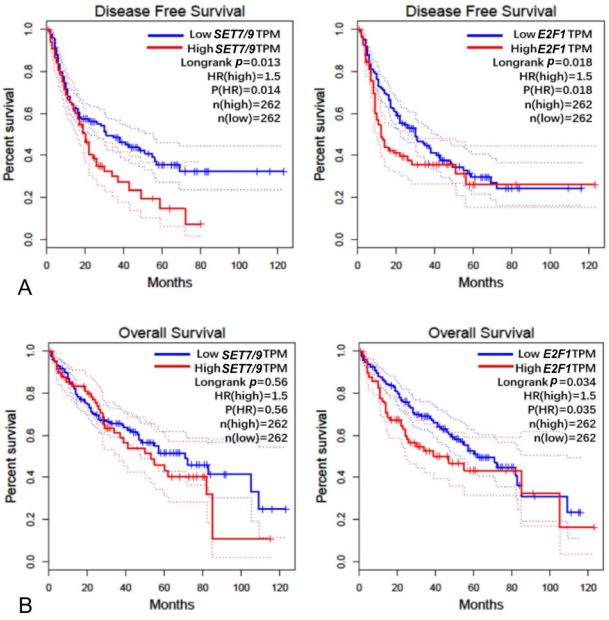

Fig. 2. The disease-free survival curve (A) and overall survival curve (B) of *SET7/9* and *E2F1* in HCC. The number of samples from high-expression group and low-expression group and the relevant *p* values were indicated on the right corner of each panel.

Co-expression pattern and interactive network of SET7/9 *and* E2F1 *in HCC*

Since our genetic alteration analysis indicated that change in mRNA expression was the most frequently-occurred genetic alteration form for both *SET7/9* and *E2F1* in HCC, we then analyzed the correlation of mRNA expression between SET7/9 and E2F1. A significant positive correlation between *SET7/9* and *E2F1* mRNA was detected in clinical samples of HCC ($p = 0.0003$, r = 0.16; Fig. 4A). To explore the potential protein partners and co-regulators of SET7/9 and E2F1, network analyses centered on SET7/9 and E2F1 were performed with the

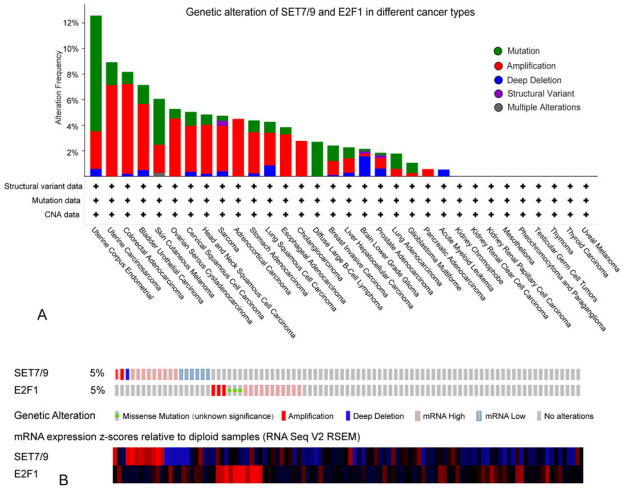

Fig. 3. Genetic alteration of *SET7/9* and *E2F1*. (A) Genetic alteration frequencies and alteration forms of *SET7/9* and *E2F1* detected in 9,896 clinical samples from 32 TCGA PanCancer Atlas Studies of different cancer types available from cBioPortal. (B) Genetic alteration and aberrant mRNA expression of *SET7/9* and *E2F1* in detected in a study involving 366 clinical samples of HCC. CNA, copy number aberration.

STRING and GeneMANIA tools (Fig. 4B, C). SET7/9 and E2F1 were proved to interact and co-express with each other in the PPI and gene networks from both STRING and GeneMANIA (Fig. 4B, C). The PPI network obtained from STRING database showed ten most important proteins directly interact with or co-expressed with SET7/9 and E2F1. Among the ten proteins, RB1, TBP, TP53, and HDAC1 are direct interactors and co-expressers of both SET7/9 and E2F1 (Fig. 4B). Using GeneMANIA database, 20 most important genes/proteins directly interact with, genetically interact with, co-expressed with, or involved in the same pathway with SET7/9 and E2F1 were detected (Fig. 4C). After combining the results of STRING and GeneMANIA, we identified 13 targets (HDAC1, CCNE1, RB1, DDB2, RBL1, TAF10, E2F2, E2F3, E2F4, TFDP1, TFDP2, DNMT1, SP1) co-expressed with both SET7/9 and E2F1 at either the protein or mRNA level and 14 proteins (HDAC1, TBP, RB1, BBC3, NDN, RBL1, E2F3, E2F4, TFDP1, TFDP2, TP53, DNMT1, FOXO3, SP1) that directly interact with both SET7/9 and E2F1 (Fig. 4D). Most of the identified gene/proteins targets related with SET7/9 and E2F1 were involved in transcriptional regulation, cell cycle transition, and cell apoptotic signaling pathways (Fig. 4C). In addition, the significant correlated expression patterns of *SET7/9* and *E2F1* with *CCNE1*, *FOXO3*, *RB1*, *TFDP1*, *TFDP2*, *SP1*, *RBL1*, *HDAC1*, *E2F3*, *E2F4*, *DNMT1*, and *DDB1* at the mRNA level were verified in a

TCGA RNAseq study involving 371 clinical samples of HCC. Among these genes, *RBL1*, *E2F3*, *SP1*, *RB1*, and *FOXO3* showed the most-significant correlation with *SET7/9*, while *RBL1*, *E2F1*, *TFDP1*, *CCNE1*, and *DNMT1* showed the most-significant correlation with *E2F1* (Fig. 4E, Supplementary Fig. S2). The co-expression patterns in other cancer types were also largely in consistent to those observed in HCC (Fig. 4E).

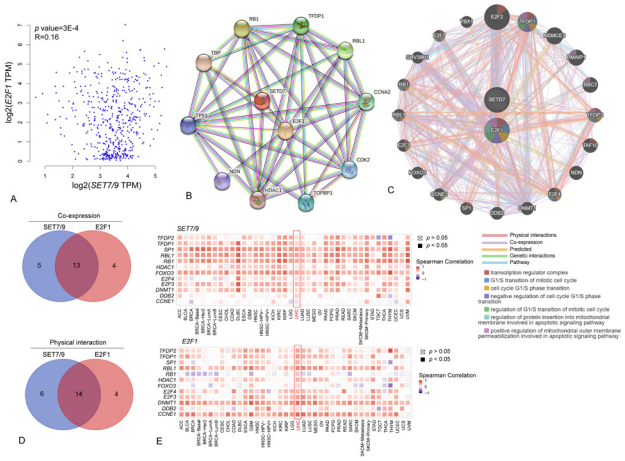

Fig. 4. Co-expression and PPI network of SET7/9 and E2F1. (A) Correlation between *SET7/9* and *E2F1* mRNA expression levels in HCC. (B) Full STRING PPI network of SET7/9 and E2F1 based on the STRING database. The known interactors, predicted interactors as well as co-expressed functional partners of SET7/9 and E2F1 were shown. The colored lines linking two protein pairs indicate the type of protein-protein association. Blue, known interaction from curated databases; Purple, known interactions determined by experiment; Green, predicted interaction by gene neighborhood analysis; Black, co-expression. (C) Gene network of SET7/9 and E2F1 obtained from the GeneMANIA portal. The node colors indicate function of correlated genes/proteins and line colors indicate the type of protein-protein association. (D) Venn diagrams showing the intersection of proteins co-expressed with SET7/9 and E2F1 (upper panel) or physically interacted with SET7/9 and E2F1 (bottom panel) based on the results of network analyses using STRING and GeneMANIA. (E) Correlations between *SET7/9* (upper panel) and *E2F1* (bottom panel) with co-expressed genes in various cancer types. The heatmaps are presented according to the purity-adjusted partial spearman's rho value as the degree of correlation.

Functional enrichment analyses

We next sought to further examine the functions of *SET7/9*- and *E2F1*-neighboring genes and explore the genetic pathways they participate in. All the differential expressing genes in 371 HCC tumor samples available from the TCGA database were screened to detect genes showing a significantly positive or negative relationship with *SET7/9* or *E2F1* in mRNA expression levels. A total of 12,041 differentially expressed genes correlated with *SET7/9* (5,619 positively correlated and 6,422 negatively correlated genes) and 9,721 differentially expressed genes correlated with *E2F1* (5,782 positively correlated and 3,939 negatively correlated genes) were identified (Fig. 5A). In consistent with network analysis, most of the co-expressed genes of SET7/9 and E2F1 identified in the STRING and GeneMANIA databases were among the positively correlated gene list of *SET7/9* (*RBL1*, *SP1*, *RB1*, *FOXO3*, *E2F4*, *HDAC1*, *CCNE1*) and *E2F1* (*TFDP1*, *TFDP2*, *SP1*, *RBL1*, *RB1*, *HDAC1*, *E2F4*, *E2F3*, *DNMT1*, *CCNE1*). Gene Set Enrichment Analysis (GSEA) showed that genes positively correlated with *SET7/9* were mainly enriched in the KEGG pathways of complement and coagulation cascades and chemical carcinogenesis, and in Gene Ontology (GO) terms of micro-body (cellular component), small molecule catabolic process (biological process), and lipid transporter activity and co-factor binding (molecular function) (Fig. 5B, Supplementary Fig. S1). Genes positively correlated with *E2F1* were mainly enriched in KEGG pathway of cell cycle regulation and GO terms of chromosomal region (cellular component), catalytic activity on DNA (molecular function), and chromosome segregation and mitotic cell cycle phase transition (biological process) (Fig. 5B, Supplementary Fig. S1).

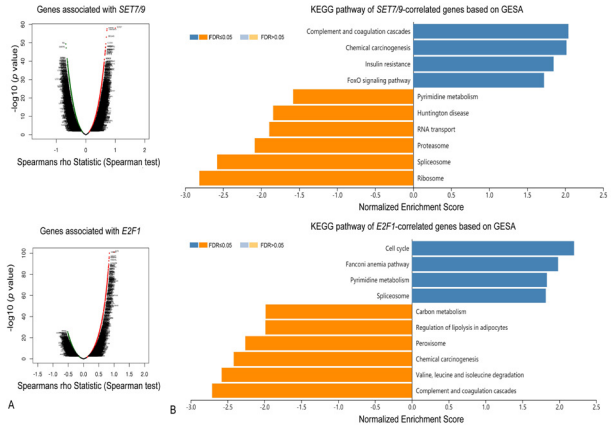

Fig. 5. Enrichment analysis of *SET7/9*- and *E2F1*-correlated genes in HCC. (A) Genes positively (red) and negatively (green) correlated with *SET7/9* and *E2F1* in HCC identified by Spearman's Correlation test of TCGA RNAseq data of 371 patients. (B) Top enriched KEGG pathways of *SET7/9*- and *E2F1*-correlated genes based on Gene Set Enrichment Analysis (GSEA). KEGG terms of positively and negatively correlated genes of *SET7/9* and *E2F1* are shown by blue and yellow bars, respectively.

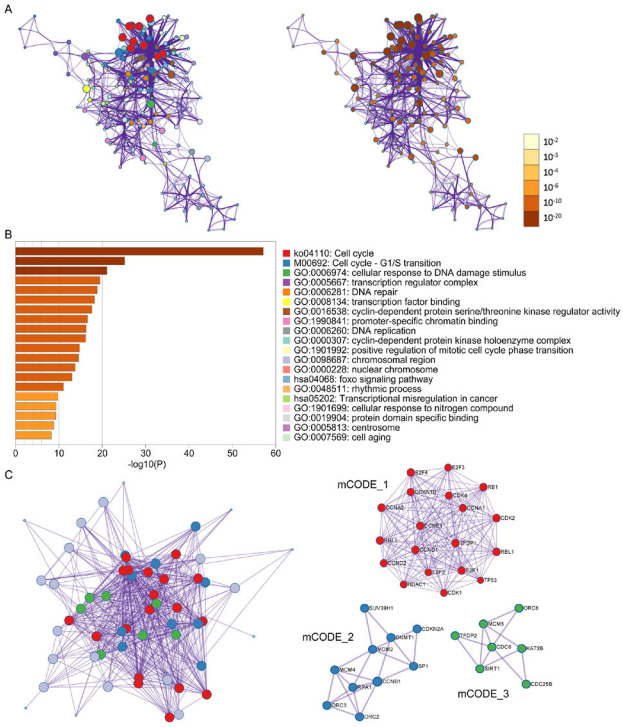

Fig. 6. Enrichment analysis of SET7/9- and E2F1-correlated proteins in PPI network. (A) The cnetplot of KEGG and GO analysis of SET7/9- and E2F1-correlated proteins. The functional terms and relative q values of each node were shown in the left and right panel. (B) Top 20 enriched terms of proteins in the PPI network. (C) Clustering analysis using Cytospace-mCODE. The overall PPI network colored by different cluster and three clustered PPI networks were shown in the left and right panel. Red, proteins clustered to mCODE_1. Blue, protein clustered to mCODE_2. Green, proteins clustered to mCODE_3.

Meanwhile, all the identified genes and proteins in the PPI and gene networks of SET7/9 and E2F1 from STRING and GeneMANIA databases were submitted to Metaspace for functional enrichment analyses (Supplementary Table SI). Our results showed that *SET7/9*- and *E2F1*-correlated gene and protein sets were significantly enriched in 85 KEGG pathways, 370 GO biological process terms, 39 GO cellular components terms, and 36 GO molecular function terms (Supplementary Table SII). Cell cycle, cellular response to DNA damage stimulus, and transcription regulator complex were the top listed KEGG pathways and GO terms (Fig. 6A, B). Clustering analysis was performed using Cytoscape-MCODE, which revealed three densely interconnected gene clusters based on the number of direct interactions and connectivity of proteins in the network. Cluster mCODE_1 has shown dense interactions with 19 proteins and 194 functional interactions. Whereas the mCODE_2 and mCODE_3 clusters include 19 and 10 proteins with 20 and 11 edges, respectively (Fig. 6C). Proteins clustered in mCODE_1 were mainly enriched in cell cycle and DNA repair, proteins clustered in mCODE_2 were mainly enriched in cell cycle and chromosome region, while proteins clustered in mCODE_3 were enriched in DNA replication, DNA replication initiation, and regulation of protein kinase activity (Table I; Fig. 6C).

Table I. Clustering analysis of *SET7/9*- and *E2F1*-correlated proteins using Cytospace-mCODE and the enriched GO and KEGG terms.

Network	Annotation	Category name	-Log10 (P)
mCODE_1	Ko04110	Cell cycle	46
	M00692	Cell cycle-G1/S transition	45.4
	GO: 0006281	DNA repair	30.8
mCODE_2	Ko04110	Cell cycle	12.1
	GO: 0098687	Chromosomal region	11.5
	GO: 0000228	Nuclear chromosome	8.2
mCODE_3	GO:0006260	DNA replication	10.7
	GO:0006270	DNA replication initiation	10.5
	GO:0045859	Regulation of protein kinase activity	7.3

Kinase, miRNA and transcription factor targets of SET7/9 *and* E2F1 *in HCC*

We finally explored the possible kinase, miRNA and transcriptional factor targets of SET7/9 and E2F1 in HCC by Linked Omics. Kinases PIM1 and STK4 were the top 2 kinase targets in the SET7/9 kinase-target network, while kinases CDK1 and PLK1 were predicted as the targets of E2F1 kinase-target network (Table II). (TTTGCAC) MIR-19A/MIR-19B and (TGAATGT) MIR-181A/MIR-181B/MIR-181C/MIR-181D, (CTCAAGA) MIR-526B and (ATATGCA) MIR-448 were predicted to be the top two targets in the SET7/9 and E2F1 miRNA-target networks, respectively (Table II). Components of the SET7/9 and E2F1 transcription factor targets were primarily related to FREAC2_01 and E4BP4_01, as well as E2F_Q6 and E2F_Q4 (Table II).

DISCUSSION

HCC is the sixth common cancer and the fourth leading cause of cancer related death worldwide (Siegel *et al.*, 2020). Same as most cancers, the occurrence of HCC is a multi-step process which might include formation of chronic inflammation, hyperplasia and malignant transformation

Table II. The Kinase, miRNA and transcription factor targets of *SET7/9* and *E2F1* in HCC.

Enriched category	Protein	Geneset	Leading Edge Num	FDR
Kinase target	SET7/9	Kinase_PIM1: Pim-1 proto-oncogene, serine/threonine kinase	10	0
		Kinase_STK4: serine/threonine kinase 4	3	0.004
	E2F1	Kinase_CDK1: cyclin dependent kinase 1	74	0
		Kinase_PLK1: polo like kinase 1	31	0
mRNA target	SET7/9	TTTGCAC, MIR-19A, MIR-19B	184	0
		TGAATGT, MIR-181A, MIR-181B, MIR-181C, MIR-181D	164	0
	E2F1	CTCAAGA, MIR-526B	20	0
		ATATGCA, MIR-448	67	0
Transcription factor target	SET7/9	V$FREAC2_01	77	0
		V$E4BP4_01	86	0
	E2F1	V$E2F_Q6	72	0
		V$E2F_Q4	71	0

*Leading Edge Num, the number of leading edge genes. V$, the annotation found in Molecular Signatures Database (MSigDB) for transcription factor (TF).

in the end. Abnormal activation of a variety of cell signal transduction pathways has contributed to the development of this long-term period. Recent advances in understanding the molecular mechanisms and signaling pathways underlying carcinogenesis have ushered in a new era of targeted therapies for treatment of HCC (Marquardt *et al*., 2012; Whittaker *et al*., 2010). Since the incidence of HCC is often not obvious and early symptoms are not typical, early diagnosis is one of the most important measures to prevent HCC occurrence and improve patient survival. Screening and identification of novel specific molecular markers for HCC using genomic or proteomic technologies based on a large patient cohort combed with bioinformatics analyses has become a priority for the establishment of a more comprehensive and effective molecular typing and stratification system, which may serve as the guidance for clinical diagnosis and targeted treatment of HCC patients.

In the past decades, growing evidences have indicated the involvement of SET7/9 in regulation of tumor metastasis, recurrence, as well as tumor cell proliferation and differentiation (Fu *et al*., 2016; Chen *et al*., 2016; Shen *et al*., 2015; Si *et al*., 2020). Of note, SET7/9 was shown to play different roles in different cancer types, which may be attributed to its multifarious substrates and the diverse biological pathways it participates in (Gu *et al*., 2018; Ea and Baltimore, 2009). Our previous studies of HCC have preliminarily investigated the expression of SET7/9 in HCC clinical samples and the effects of abnormal SET7/9 expression on the cellular behavior of HCC cells. The results showed that both SET7/9 and E2F1 are up-regulated in HCC and high-expression of SET7/9 in combination with E2F1 has a positive role in promoting the oncogenic processes of HCC (Gu *et al*., 2018). In consistent with our finding, the function of E2F1 in promoting HCC proliferation has been well recognized recently (Farra *et al*., 2017; Lin *et al*., 2019). In addition, E2F1 was found to participate in the oncogenic processes downstream of SET7/9 in both HCC cells, lung adenocarcinoma cells, and osteosarcoma cells (Gu *et al*., 2018; Lezina *et al*., 2014), which indicated an important role of the SET7/9-E2F1 axis in cancer development. However, the relevant signaling pathways and molecular partners of the SET7/9-E2F1 axis still remain to be further investigated in order to better understand the inner mechanisms of SET7/9-E2F1 in regulating HCC initiation and progression.

In this study, we explored the correlation between expression of SET7/9 and E2F1 and the risk and patient survival of HCC. Transcriptional sequencing data from 371 HCC patient cases from TCGA databases confirmed that expressions of *SET7/9* and *E2F1* are significantly higher in HCC compared with normal tissues (Fig. 1B), which have been observed in our previous study of 68 HCC tissues samples (Chen *et al*., 2016). Although significant correlation was only detected between the expression level of *E2F1* and tumor progression based on TCGA data (Fig. 1C), several clinical studies have proved the significant correlation between both SET7/9 and E2F1 expression and the pathological stage of HCC tumor at the protein level (Chen *et al*., 2016; Gu *et al*., 2018). Meanwhile, the expression level of *SET7/9* was significantly associated with disease-free survival (Fig. 2A), which supported a previous study showing a positive correlation between SET7/9 expression and tumor differentiation, tumor metastasis, and recurrence rate of HCC patients (Chen *et*

al., 2016). The expression level of *E2F1* was significantly associated with both disease-free survival and overall survival (Fig. 2A, B), which also accorded with results from a clinical study showing a positive correlation between E2F1 expression and HCC intrahepatic metastasis and distant metastasis and a negative correlation between E2F1 expression and overall survival rate of HCC patients (Lin *et al*., 2019). In addition, a positive correlation was detected between the mRNA expression levels of *SET7/9* and *E2F1* mRNA (p = 0.0003; Fig. 4A). Together, our results confirmed the synergistic role of SET7/9 and E2F1 in HCC.

As SET7/9 is a lysine methyltransferase and E2F1 is a transcription factor, they mainly participate in the carcinogenesis process by acting as a coordinator or transcriptional regulator that affects the activation of various downstream molecules involved in different signaling pathway. Subtle changes in SET7/9 or E2F1 expression at either mRNA or protein level may lead to dis-function of the related network that orchestrates tumor cell proliferation, growth, and differentiation. Indeed, genetic alteration analysis showed that aberrant mRNA expression of *SET7/9* and *E2F1* accounts for the majority of genetic variation forms (Fig. 3). Abnormal amplification, deep deletion, and mutation of both *SET7/9* and *E2F1* were also detected in HCC samples but with relatively low proportion (Fig. 3). Therefore, we further focused on the PPI and gene networks of SET7/9-E2F1 and characterized the enriched functions of SET7/9- and E2F1-correlated genes/proteins in the networks. Using the STRING database and GeneMANIA prediction server, we identified 15 proteins directly interact with SET7/9 and E2F1 and 13 genes/proteins co-expressed with SET7/9 and E2F1 (Fig. 4B, C). Noteworthy, the close relationship between SET7/9 and E2F1 and many proteins in the networks, such as TP53, CCNE1, RB1, DNMT1, HDAC1, SP1 and FOXO3 have been reported in several cancer types before (Lezina *et al*., 2014; Ivanov *et al*., 2007; Liu *et al*., 2018; López-Nieva *et al*., 2018; Zou *et al*., 2012; Tanaka *et al*., 2015; Shats *et al*., 2013; Carr *et al*., 2014; Calnan *et al*., 2012; Robertson *et al*., 2000; Montenegro *et al*., 2016). For example, TP53 is a methylation target of SET7/9 in colorectal cancer (CRC) and osteosarcoma tumor, and a transcriptional target of E2F1 in human T-cell lymphoblastic lymphomas (Ivanov *et al*., 2007; Liu *et al*., 2018; López-Nieva *et al*., 2018). CCNE1 is a downstream responder of the SET7/9-E2F1 axis, which is responsible for cell-cycle regulation of lung cancer cells upon DNA damage and cell proliferation of HCC cancer cells (Gu *et al*., 2018; Lezina *et al*., 2014). Some proteins identified in the PPI network tend to act as a reciprocal regulator with SET7/9 or E2F1 instead of a strict downstream regulator of the SET7/9-E2F1 axis. Transcription factors SP1, RB1, and FOXO3 can directly interact with E2F1 to regulate the expression of a series of downstream targets (Zou *et al*., 2012; Tanaka *et al*., 2015; Shats *et al*., 2013). RB1 can be methylated by SET7/9 (Carr *et al*., 2011, 2014; Munro *et al*., 2010), which is required for the formation of a chromatin-bound pRb/53BP1 complex on E2F target genes and participation of RB1 in E2F1-dependent cell cycle control and DNA-damage response (Carr *et al*., 2014). Similarly, methylation of FOXO3 by SET7/9 decreases the stability but increases the transcriptional activity of FOXO3, which may further lead to changes in the E2F1/FOXO transcriptional program by affecting the transcriptional specificity and apoptotic function of E2F1 (Shats *et al*., 2013; Calnan *et al*., 2012). Although these SET7/9-E2F1-related pathways have not been reported in HCC, co-expression analyses based on clinical HCC samples also showed that the mRNA expression levels of *RBL1*, *SP1*, *RB1*, *FOXO3*, *E2F4*, *HDAC1*, *CCNE1* were significantly correlated with *SET7/9* and those of *TFDP1*, *TFDP2*, *SP1*, *RBL1*, *RB1*, *HDAC1*, *E2F4*, *E2F3*, *DNMT1*, *CCNE1* were significantly correlated with *E2F1* (Fig. 4E; Supplementary Fig. S2). Meanwhile, a great portion of *SET7/9* co-expression genes in HCC were enrich in the KEGG pathway of FOXO signaling (Fig. 5B). The results suggested a similar regulatory network in HCC and in other cancer types. Future studies may provide further experimental evidence for the correlation of these proteins with SET7/9 and E2F1 and their involvement in controlling HCC development through the SET7/9-E2F1 pathway.

Functional enrichment analysis showed that most proteins related with SET7/9-E2F1 were involved in pathways controlling cell cycle, cellular response to DNA damage stimulus, and transcription regulator complex, which are closely correlated with malignant transformation of tumor cells. Clustering analysis further divided the target gene sets into three categories, each enriched in cell cycle and DNA repair, cell cycle and chromosome region, and DNA replication and protein kinase activity regulation, respectively (Fig. 6; Table I). Noteworthy, in lung cancer, colorectal cancer, and osteosarcoma tumor, SET7/9-catalyzed E2F1 methylation can lead to changes in the stability of E2F1 and the binding ability of E2F1 on its target genes, which serves as an important mechanism regulating the transcription of several E2F1 downstream targets controlling cell apoptosis and proliferation (Gu *et al*., 2018; Lezina *et al*., 2014; Carr *et al*., 2014). Our results are largely consistent with our current understanding on how the SET7/9-E2F1 pathway regulates cellular behavior of tumor cells and affects cancer progression.

In conclusion, our study confirmed a cancer-

promoting role of SET7/9 and E2F1 in HCC, predicted the potential co-regulators of the SET7/9-E2F1 axis and showed the involvement of SET7/9-E2F1-correlated pathway in the regulation of cell cycle, DNA repair and replication, and gene transcription. Our study supported the previous findings that SET7/9 and E2F1 may serve as valuable diagnostic and prognostic markers for HCC (Chen *et al*., 2016; Huang *et al*., 2019). Future *in vitro* and in vivo studies and molecular-level analyses in HCC cells are necessary to validate the bioinformatics predictions, especially the predicted protein co-regulators and kinase/miRNA/transcriptional factor targets of SET7/9-E2F1 based on transcriptome sequencing data and curated databases, which may provide novel insights into the molecular pathogenesis of HCC and the development of systemic therapy for HCC.

ACKNOWLEDGMENTS

The authors thank Dr. Martin Zulqarnain Muhammad for English editing and proof reading. This project was supported by Hangzhou Peak Discipline of Gastroenterology, the Key Laboratory of Integrated Traditional Chinese and Western Medicine for Biliary and Pancreatic Diseases of Zhejiang Province, the Key Laboratory of Clinical Cancer Pharmacology and Toxicology Research of Zhejiang Province (2020E10021), the Science and Technology Project of Hangzhou Health Commission (A20200119), the Zhejiang Medical and Health Science and Technology Plan (Grant No. WKJ-ZJ-2136 2019RC068 and 2021437779), and the Hangzhou Medical and Health Science and Technology Plan (Grant No. 2016ZD01, OO20190610 and A20200174). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Supplementary material

There is supplementary material associated with this article. Access the material online at: https://dx.doi.org/10.17582/journal.pjz/20211129131153

Statement of conflict of interest

The authors have declared no conflict of interest.

REFERENCES

Akiyama, Y., Koda, Y., Byeon, S.j., Shimada, S., Nishikawaji, T., Sakamoto. A, Chen, Y.X., Kojima, K., Kawano, T., Eishi, Y., Deng, D.J., Kim, W.H., Zhu, W.G., Yuasa, Y., and Tanaka, S., 2016. Reduced expression of SET7/9, a histone monomethyltransferase, is associated with gastric cancer progression. *Oncotarget*, **7**: 3966. https://doi.org/10.18632/oncotarget.6681

Bruix, J., Reig, M., and Sherman, M., 2016. Evidence-based diagnosis, staging, and treatment of patients with hepatocellular carcinoma. *Gastroenterology*, **150**: 835-853. https://doi.org/10.1053/j.gastro.2015.12.041

Calnan, D.R., Webb, A.E., White, J.L., Stowe, T.R., Goswami, T., Shi, X.B., Espejo, A., Bedford, M.T., Gozani, O., Gygi, S.P., and Brunet, A., 2012. Methylation by SET9 modulates foxo3 stability and transcriptional activity. *Aging*, **4**: 462-479. https://doi.org/10.18632/aging.100471

Carr, S.M., Munro, S., and Kessler, B., 2011. Interplay between lysine methylation and Cdk phosphorylation in growth control by the retinoblastoma protein. *EMBO J*., **30**: 317-327. https://doi.org/10.1038/emboj.2010.311

Carr, S.M., Munro, S., and Zalmas, L.P., 2014. Lysine methylation-dependent binding of 53BP1 to the pRb tumor suppressor. *Proc. natl. Acad. Sci.*, **111**: 11341-11346. https://doi.org/10.1073/pnas.1403737111

Chacko, S., and Samanta, S., 2016. Hepatocellular carcinoma: A life-threatening disease. *Biomed. Pharmacother*., **84**: 1679-1688. https://doi.org/10.1016/j.biopha.2016.10.078

Chandrashekar, D.S., Bashel, B., Balasubramanya, S.A.H., Creighton, C.J., Ponce-Rodriguez, I., Chakravarthi, B.V., and Varambally, S., 2017. Ualcan: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia*, **19**: 649-658. https://doi.org/10.1016/j.neo.2017.05.002

Chen, Y., Yang, S., Hu, J., Yu, C., He, M., and Cai, Z., 2016. Increased expression of SETD7 promotes cell proliferation by regulating cell cycle and indicates poor prognosis in hepatocellular carcinoma. *PLoS One*, **11**: e0154939. https://doi.org/10.1371/journal.pone.0154939

Colagrande, S., Inghilesi, A.L., Aburas, S., Taliani, G.G., Nardi, C., and Marra, F., 2016. Challenges of advanced hepatocellular carcinoma. *World J. Gastroenterol*., **22**: 7645. https://doi.org/10.3748/wjg.v22.i34.7645

Ea, C.K., and Baltimore, D., 2009. Regulation of NF-kappaB activity through lysine monomethylation of p65. *Proc. natl. Acad. Sci*., **106**: 18972-18977. https://doi.org/10.1073/pnas.0910439106

Farra, R., Grassi, G., Tonon, F., Abrami, M., Grassi, M., Pozzato, G., Fiotti, N., Forte, G., and Dapas, B.,

2017. The role of the transcription factor E2F1 in hepatocellular carcinoma. *Curr. Drug Deliv*., **14**: 272-281.

Fu, L., Wu, H.L., and Cheng, S.Y., 2016. SET7 mediated Gli3 methylation plays a positive role in the activation of sonic hedgehog pathway in mammals. *eLife*, **5**: e15690. https://doi.org/10.7554/eLife.15690

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y.C., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C., and Schultz, N., 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal*., **6**: 11. https://doi.org/10.1126/scisignal.2004088

Gu, Y., Wang, X., Liu, H., Li, G., Yu, W., and Ma, Q., 2018. SET7/9 promotes hepatocellular carcinoma progression through regulation of E2F1. *Oncol. Rep*., **40**: 1863-1874. https://doi.org/10.3892/or.2018.6621

Heimbach, J.K., Kulik, L.M., Finn, R.S., Sirlin, C.B., Abecassis, M.M., Roberts, L.R., Zhu, A.X., Murad, H., and Marrero, J.A., 2018. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology*, **67**: 358-380. https://doi.org/10.1002/hep.29086

Huang, Y.L., Ning, G., Chen, B.L., Lian, Y.F., Gu, Y.R., Wang, J.L., Chen, D.M., Wei, H., and Huang, Y.H., 2019. Promising diagnostic and prognostic value of e2fs in human hepatocellular carcinoma. *Cancer Manag. Res*., **11**: 1725-1740. https://doi.org/10.2147/CMAR.S182001

Ivanov, G.S., Ivanova, T., Kurash, J., Ivanov, A., and Barlev, N.A., 2007. Methylation acetylation interplay activates p53 in response to DNA damage. *Mol. Cell Biol*., **27**: 6756-6769. https://doi.org/10.1128/MCB.00460-07

Kulik, L., and El-Serag, H.B., 2019. Epidemiology and management of hepatocellular carcinoma. *Gastroenterology*, **156**: 477-491. https://doi.org/10.1053/j.gastro.2018.08.065

Lezina, L., Aksenova, V., Ivanova, T., Purmessur, N., Antonov, A.V., Tentler, D., Fedorova, O., Garabadgiu, A.V., Talianidis, I., and Melino, G., 2014. KMTase Set7/9 is a critical regulator of E2F1 activity upon genotoxic stress. *Cell Death Differ*., **21**: 1889-1899. https://doi.org/10.1038/cdd.2014.108

Lin, M., Liu, Y.M., Ding, X.K., Ke, Q.J., Shi, J.Y., Ma, Z.W., Gu, H.Q., Wang, H.X., Zhang, C.F., Yang, C.X., Fang, Z.J., Zhou, L.F., and Ye, M., 2019. E2F1 transactivates IQGAP3 and promotes proliferation of hepatocellular carcinoma cells through IQGAP3-mediated PKC-alpha activation. *Am. J. Cancer Res*., **9**: 285-299.

Liu, Z.L., Wu, X.H., Lv, J.J., Sun, H., and Zhou, F., 2018. Resveratrol induces p53 in colorectal cancer through SET7/9. *Oncol. Lett*., **17**: 3783-3789. https://doi.org/10.3892/ol.2019.10034

López-Nieva, P., Fernández-Navarro, P., Vaquero-Lorenzo, C., Villa-Morales, M., Graña-Castro, O., Cobos-Fernández, M.Á., López-Lorenzo, J.L., Llamas, P., González-Sanchez, L., Sastre, I., Pollan, M., Malumbres, M., Santos, J., and Fernández-Piqueras, J., 2018. Rna-seq reveals the existence of a cdkn1c-e2f1-tp53 axis that is altered in human t-cell lymphoblastic lymphomas. *BMC Cancer*, **18**: 430. https://doi.org/10.1186/s12885-018-4304-y

Marquardt, J.U., Galle, P.R., and Teufel, A., 2012. Molecular diagnosis and therapy of hepatocellular carcinoma (HCC): An emerging field for advanced technologies. *J. Hepatol*., **56**: 267-275. https://doi.org/10.1016/j.jhep.2011.07.007

Montenegro, M.F., Sánchz-d-Campo L., Ronzáz-urrro, R., Martínez-Barba, E., Piñero-Madrona, A., Cabezas-Herrera, J., and Rodríguez-López, J.N., 2016. Tumor suppressor SET9 guides the epigenetic plasticity of breast cancer cells and serves as an early-stage biomarker for predicting metastasis. *Oncogene*, **35**: 6143-6152. https://doi.org/10.1038/onc.2016.154

Munro, S., Khaire, N., and Inche, A., 2010. Lysine methylation regulates the pRb tumour suppressor protein. *Oncogene*, **29**: 2357-2367. https://doi.org/10.1038/onc.2009.511

Pradhan, S., Chin, H.G., Estève, P.O., and Jacobsen, S.E., 2009. SET7/9 mediated methylation of non-histone proteins in mammalian cells. *Epigenetics*, **4**: 383-387. https://doi.org/10.4161/epi.4.6.9450

Robertson, K.D., Aitsiali, S., Yokochi, T., Wade, P.A., Jones, P.L., and Wolffe, A.P., 2000. DNMT1 forms a complex with rb, e2f1 and hdac1 and represses transcription from e2f-responsive promoters. *Nat. Genet*., **25**: 338-342. https://doi.org/10.1038/77124

Sauzay, C., Petit, A., Bourgeois, A.M., Barbare, J.C., Chauffert, B., Galmiche, A., and Houessinon, A. 2016. Alpha-foetoprotein (AFP): A multi-purpose marker in hepatocellular carcinoma. *Clin. Chim. Acta*, **463**: 39-44. https://doi.org/10.1016/j.cca.2016.10.006

Shats, I., Gatza, M.L., Liu, B., Angus, S.P., You, L., and Nevins, J.R., 2013. Foxo transcription factors control e2f1 transcriptional specificity and apoptotic function. *Cancer Res*., **73**: 6056-6067.

https://doi.org/10.1158/0008-5472.CAN-13-0453

Shen, C., Wang, D., Liu, X., Gu, B., Du, Y., Wei, F.Z., Cao, L.L., Song, B.Y., Lu, X.P., Yang, Q.Y., Zhu, Q., Hou, T.Y., Li, M.T., Wang, L., Wang, H.Y., Zhao, Y., Yang, Y., and Zhu, W.G., 2015. SET7/9 regulates cancer cell proliferation by influencing β-catenin stability. *FASEB J.*, **29**: 4313-4323. https://doi.org/10.1096/fj.15-273540

Siegel, R.L., Miller, K.D., Goding, S.A., Fedewa, S.A., Butterly, L.F., Anderson, J.C., Cercek, A., Smith, R.A., and Dvm, A.J., 2020. Colorectal cancer statistics, CA. *Cancer J. Clin.*, **70**: 145-164. https://doi.org/10.3322/caac.21601

Si, W., Zhou, J., Zhao, Y., Zheng, J., and Cui, L., 2020. SET7/9 promotes multiple malignant processes in breast cancer development via RUNX2 activation and is negatively regulated by TRIM21. *Cell Death Dis.*, **11**: 1-15. https://doi.org/10.1038/s41419-020-2350-2

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., Jensen, L., and Mering, C.V., 2019. STRING v11: Protein protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucl. Acids Res.*, **47**: D607-D613. https://doi.org/10.1093/nar/gky1131

Tanaka, T., Kanatsu-Shinohara, M., and Shinohara, T., 2015. The cdkn1b-rb1-e2f1 pathway protects mouse spermatogonial stem cells from genomic damage. *J. Reprod. Develop.*, **61**: 305. https://doi.org/10.1262/jrd.2015-027

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z., 2017. GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucl. Acids Res.*, **45**: W98-W102. https://doi.org/10.1093/nar/gkx247

Vasaikar, S.V., Straub, P., Wang, J., and Zhang, B., 2018. LinkedOmics: Analyzing multi-omics data within and across 32 cancer types. *Nucl. Acids Res.*, **46**: D956-D963. https://doi.org/10.1093/nar/gkx1090

Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Max, F., Chris, G., Farzana, K., and Tannus, L.C., 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucl. Acids Res.*, **38**: W214-W220. https://doi.org/10.1093/nar/gkq537

Whittaker, S., Marais, R., and Zhu, A., 2010. The role of signaling pathways in the development and treatment of hepatocellular carcinoma. *Oncogene*, **29**: 4989-5005. https://doi.org/10.1038/onc.2010.236

Yu, S.J., 2016. A concise review of updated guidelines regarding the management of hepatocellular carcinoma around the world: 2010-2016. *Clin. Mol. Hepatol.*, **22**: 7. https://doi.org/10.3350/cmh.2016.22.1.7

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K., 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.*, **10**: 1-10. https://doi.org/10.1038/s41467-019-09234-6

Zou, L., Xu, H.G., Ren, W., Jin, R., Wang, Y., and Zhou, G.P., 2012. Transcriptional activation of the human CD2AP promoter by E2F1. *PLoS One*, **7**: e42774. https://doi.org/10.1371/journal.pone.0042774