



Original Article



Expression and Regulatory Network Analysis of Function of Small Nucleolar RNA Host Gene 4 in Hepatocellular Carcinoma

Jing Cao^{1#}, Cuicui Xiao^{2#}, Christ-Jonathan Tsia Hin Fong³, Jiao Gong⁴, Danyang Li¹, Xiangyong Li¹, Yusheng Jie^{1*} and Yutian Chong^{1*}

¹Department of Infectious Diseases, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China;

²Department of Anesthesiology, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China;

³Department of Gastroenterology, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China;

⁴Department of Laboratory Medicine, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China

Received: 22 December 2020 | Revised: 30 March 2021 | Accepted: 2 June 2021 | Published: 2 July 2021

Abstract

Background and Aims: Long non-coding RNA small nucleolar RNA host genes (SNHG) play a critical role in the occurrence and development of tumors. In this study, we aimed to investigate the role of *SNHG4* in hepatocellular carcinoma (HCC) and its underlying mechanism. **Methods:** Datasets were acquired from The Cancer Genome Atlas (TCGA) database. InLoco 2.0 was used to identify the distribution of *SNHG4* in HCC cells. Gene expression, Kaplan-Meier survival, microRNA and transcription factor target analyses were performed with the University of Alabama Cancer (UALCAN) Database, Kaplan-Meier Plotter, LinkedOmics, WebGestalt and gene set enrichment analysis, respectively. Gene Ontology and pathway enrichment analyses and assessment of RNA binding proteins were performed by R software, circIncrNA and Encyclopedia of RNA Interactomes (ENCORI). In addition, CircIncrNA and ENCORI were used to find the correlation between *SNHG4* and important proteins, while the prognostic value was assessed with the Human Protein Atlas database and Kaplan-Meier Plotter. **Results:**

Expression of *SNHG4* in HCC is higher in HCC tissue than in normal healthy liver tissues and is mainly distributed in the nucleus. *SNHG4* positively correlated with poor prognosis ($p < 0.01$ for overall survival and recurrence-free survival). Functional enrichment analysis revealed *SNHG4* involvement with regulation of ribosomal RNA synthesis and the RNA processing and surveillance pathway. *SNHG4* is closely associated with miR-154 and miR-206, transcription factor target E2F family and the signaling pathway for MAPK/ERK and mTOR. U2 auxiliary factor 2 (U2AF2) showed strong correlation with *SNHG4*, while low-expression of U2AF2 showed good prognosis. **Conclusions:** Based on our findings, we infer *SNHG4* may play a role in the formation of HCC via regulation of tumor-related pathways.

Citation of this article: Cao J, Xiao C, Fong C-JTH, Gong J, Li D, Li X, et al. Expression and Regulatory Network Analysis of Function of Small Nucleolar RNA Host Gene 4 in Hepatocellular Carcinoma. J Clin Transl Hepatol 2022;10(2):297–307. doi: 10.14218/JCTH.2020.00175.

Keywords: Small nucleolar RNA host gene 4; Hepatocellular carcinoma; Prognosis; Bioinformatic analysis.

Abbreviations: BP, biological process; CC, cellular component; CI, confidence interval; CNV, copy number variation; DEGs, differentially expressed genes; ENCORI, Encyclopedia of RNA Interactomes; ERK, extracellular-signal-regulated kinase; ES, enrichment score; FDR, false discovery rate; GO, Gene Ontology; GSEA, gene set enrichment analysis; HCC, hepatocellular carcinoma; HEATR1, HEAT repeat-containing protein 1; HGF, hepatocyte growth factor; HPA, Human Protein Atlas; HR, hazard ratio; KEGG, Kyoto Encyclopedia of Genes and Genomes; LIHC, liver hepatocellular carcinoma; lncRNA, long non-coding RNA; MAPK, mitogen-activated protein kinase; MF, molecular function; miR/miRNA, microRNA; mTOR, mammalian target of rapamycin; ncRNA, non-coding RNA; OS, overall survival; NES, normalized enrichment score; RBP, RNA-binding protein; RFS, recurrence-free survival; RNA-seq, RNA sequencing; SNHG, small nucleolar RNA host genes; snoRNA, small nucleolar RNAs; snRNP, small nuclear ribonucleoprotein particles; RNF44, RING finger protein 44; TCGA, The Cancer Genome Atlas; TRAP, Transcription factor Affinity Prediction; U2AF2, U2 auxiliary factor 2; UALCAN, University of Alabama Cancer.

*Contributed equally to this work.

***Correspondence to:** Yusheng Jie and Yutian Chong, Department of Infectious Diseases, Third Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong 510630, China. ORCID: <https://orcid.org/0000-0003-3756-0653> (YJ), <https://orcid.org/0000-0001-8215-4393> (YC). Tel: +86-20-8525-2372, Fax: +86-20-8525-2250, E-mail: jieyush@mail.sysu.edu.cn (YJ), chongyt@mail.sysu.edu.cn (YC)

Introduction

Hepatocellular carcinoma (HCC), which represents about 85–90% of cases of primary liver cancer, is the fourth most common malignancy and the third leading cause of tumor-related deaths in China.¹ Worldwide, HCC is the second and sixth leading cause of cancer-related deaths in males and females, respectively.² Risk factors, including viral infection, alcoholic cirrhosis and fatty liver, along with chronic inflammatory disorders of the biliary tract, genetic diseases and carcinogens, contribute to the risk of HCC. These are responsible for abnormal gene expression, resulting in increased cancer cell proliferation and escape from immune surveillance, which can speed up the process of development into HCC.³ Patients with HCC are usually hospitalized at an advanced stage and are prone to recurrence and metastasis. Therefore, it is of primary concern to explore the potential mechanisms of HCC and discover sensitive biomarkers to screen out high-risk patients.

With the advent of molecular biology, bioinformatics and

sequencing technology have become increasingly used in research. The Cancer Genome Atlas (TCGA) database has accumulated abundant genomic and gene expression profiles for different diseases in the past decade. Through analysis of these data, various key genes and signaling pathways related to the disease may be identified, which may aid in better understanding of the disease mechanism and provide a theoretical basis for various clinical applications, including prognostic markers and potential therapeutic targets. Recent studies have shown that non-coding RNA (ncRNA) characterized by microRNA (miR/miRNA) and long non-coding RNA (lncRNA) are important regulatory molecules involved in various physiological and pathological cellular processes.⁴ Small nucleolar RNAs (snoRNA) are a class of nuclear-enriched intron-derived lncRNAs, between 60–300 nucleotides in length, predominantly found in the nucleolus and playing an important role in the process of chemical RNA modification, pre-RNA processing and alternative splicing control. Recently, snoRNAs and their host genes (small nucleolar RNA host genes, SNHGs) have been reported in a wide spectrum of cancers.^{5–7} *SNHG4*, which is localized at the 5q31.2 region and has five exons, can promote tumor growth in patients with osteosarcoma⁸ and has been associated with a shorter overall survival (OS) in patients with HCC.⁹ miRNAs are small endogenous RNAs involved in regulating gene-expression post-transcriptionally. *SNHG4* was found to promote tumor growth by sponging miR-224-3p in osteosarcoma and miRNA-204-5p in gastric cancer.^{8,10}

Understanding this novel RNA crosstalk can lead to significant insight into gene regulatory networks and have implications in disease pathogenesis.¹¹ However, the functional network and relationship between *SNHG4* and HCC has not yet been elucidated. Therefore, the aim of the present study was to investigate the relationship between the expression of *SNHG4* and HCC using online databases and to predict *SNHG4*-related pathways as well as related binding proteins and miRNAs (Fig. 1).

Methods

Expression of *SNHG4* in HCC

The TCGA portal is the largest and most commonly used public resource providing somatic mutation, gene expression, gene methylation and copy number variation (CNV) datasets for several thousands of tumor samples.¹² Various computational tools have been developed to aid researchers in performing specific TCGA data analyses. UALCAN (<http://ualcan.path.uab.edu>) is an interactive website used to perform in-depth analyses of TCGA gene expression data and analyze relative expression of a query gene(s) across tumor and normal samples, as well as in various tumor subgroups, based on individual cancer stages, tumor grade or other clinicopathological features.^{13,14} Among 377 samples of liver hepatocellular carcinoma (LIHC) present in the TCGA database, *SNHG4* expression was found in only 371 samples. The IncLocator 2.0 is a cell-line-specific predictor, which trains an end-to-end deep model per cell line, for predicting lncRNA subcellular localization from sequences. IncLocator 2.0 was used to identify the distribution of *SNHG4* in the HepG2 cell line.¹⁵ Experimental validation was then conducted (see Supplementary Materials).

Survival analysis of *SNHG4* in HCC

The Kaplan-Meier Plotter (www.kmplot.com) is an online database which can assess the effects of 54,675 genes on survival using 10,461 cancer samples.¹⁶ In this study, it was

used to analyze OS and the recurrence-free survival (RFS) of HCC patients; patient samples were split into two groups, according to whether the expression value was above or below the median. The Mantel-Cox test, also known as the log-rank test, was used to determine significance of difference between survival curves. The number-at-risk is indicated below each panel, while the hazard ratio (HR) with 95% confidence interval (CI) and log-rank *p*-value are shown in each plot.

Functional and signaling pathway analysis

The functional annotation of *SNHG4*-related genes was performed using Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. GO enrichment analysis is commonly used to annotate the degree of gene function terms in differentially expressed genes (DEGs), which include molecular function (MF), biological process (BP), and cellular component (CC). KEGG enrichment analysis was used to demonstrate enriched signaling pathways in DEGs. Both GO and KEGG pathway analyses were performed using the R package “clusterProfiler”. Enrichment maps which allow visualization of pathway enrichment results were drawn using R software (<http://www.r-project.org/>). Gene set enrichment analysis (GSEA; <http://www.broad.mit.edu/gsea>) was conducted to perform analysis of *SNHG4*-related canonical pathways in HCC.

Analysis of genes, miRNA and transcription factor targets related to *SNHG4* in HCC

Analysis of genes which interact with *SNHG4*, miRNA and transcription factor targets of *SNHG4* in HCC was carried out using LinkedOmics. The LinkedOmics database (<http://www.linkedomics.org/login.php>) harbors multiomics data and clinical data for 32 cancer types and a total of 11,158 patients from TCGA.¹⁷ LinkedOmics consists of three analysis modules: LinkFinder, LinkCompare, and LinkInterpreter. The LinkFinder module in LinkedOmics was used to study genes differentially expressed in correlation with *SNHG4*. Results were graphically presented as volcano plots, heat maps or scatter plots, and were statistically analyzed using Pearson’s correlation coefficient. A *p*-value of <0.05 was considered significant. Subsequently, GSEA was performed using the LinkInterpreter module in LinkedOmics. The results were further analyzed by KEGG and Wikipathway, conducted in WebGestalt for analysis of genes interacting with *SNHG4*, miRNA targets and transcription factor-target enrichment. The rank criterion was a false discovery rate (FDR) <0.05, and 500 simulations were performed. In order to validate the transcription factors obtained by LinkedOmics, we used Transcription Factor Affinity Prediction (TRAP) Web Tools (<http://trap.molgen.mpg.de/cgi-bin/home.cgi>). TRAP Web Tools was developed at the Max Planck Institute for Molecular Genetics to predict transcription factor (TF) binding affinities to DNA.¹⁸

OncoPrint is an online platform that provides data from cancer microarray datasets, which can also be used for data mining in multiple cancers, including HCC.¹⁹ Wurmbach Liver dataset was selected from the OncoPrint to evaluate the expression characteristics of RNF44 and HEATR1 in HCC, compared to that in normal liver tissues. A *p*-value <0.05 was considered as statistically significant.

Prediction and analysis of RNA-binding proteins (RBPs) related to *SNHG4*

CircIncrNet (<http://120.126.1.61/circlnc/index.php>) is

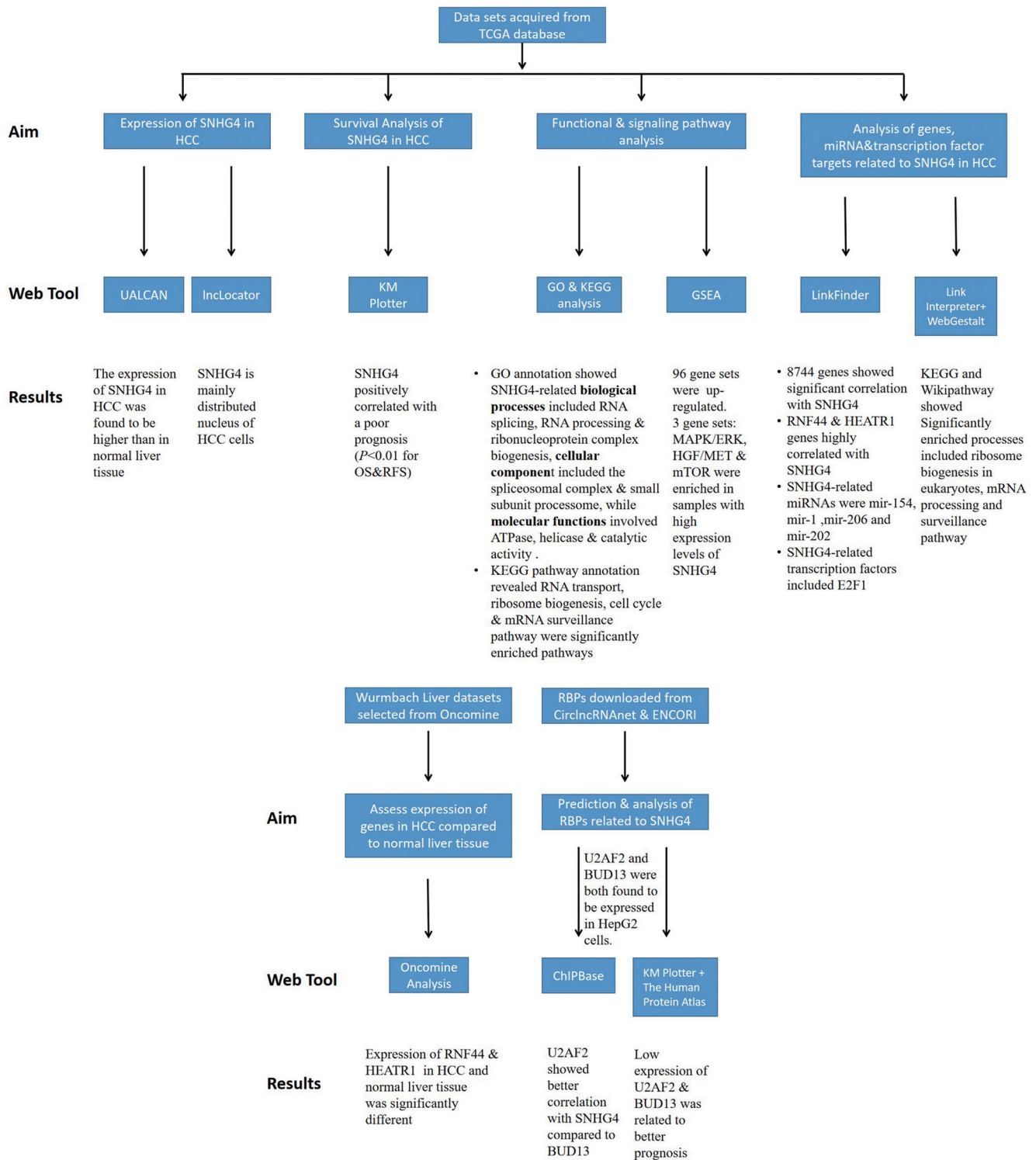


Fig. 1. Workflow chart for this research study.

an integrated web-based resource for mapping of functional networks of ncRNAs.²⁰ The Encyclopedia of RNA Interactomes (ENCORI) is an open-source web tool used for studying the ncRNA interactions of CLIP-seq, degradome-seq, and RNA-RNA interactome and RBPs data.²¹ The RBPs rel-

evant to *SNHG4* were downloaded from the CircIncrNA.net and ENCORI. Subsequently, RBPs coexistent in both databases and which were expressed in HepG2 cells (Human hepatoma cell lines), were screened.

The Human Protein Atlas (HPA) (<https://www.proteinatlas>

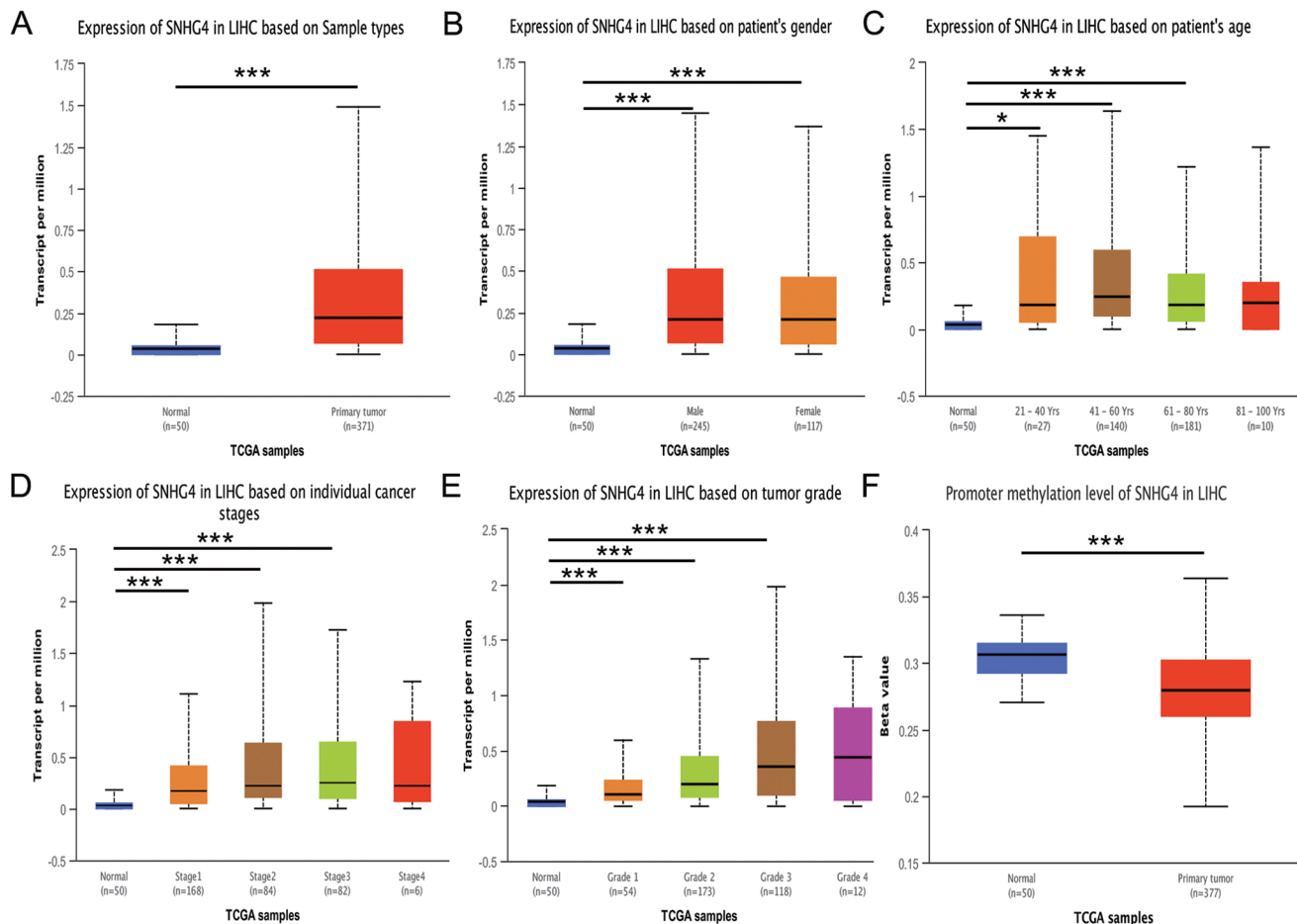


Fig. 2. *SNHG4* expression in subgroups of patients with HCC, stratified based on sex, age and other criteria (University of Alabama Cancer Database). Relative expression of *SNHG4* is shown in boxplots. (A) In normal and LIHC samples. (B) In normal individuals of both sexes compared to male and female patients with LIHC. (C) In normal individuals of any age or in patients with LIHC aged 21–40, 41–60, 61–80 or 81–100 years. (D) In normal individuals or in patients with LIHC in stages I, II, III or IV. (E) In normal individuals or patients with LIHC with grade 1, 2, 3 or 4 tumors. (F) Promoter methylation in normal individuals or patients with LIHC. Data are expressed as mean±standard error. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *SNHG4*, small nucleolar RNA host gene 4; HCC, hepatocellular carcinoma; LIHC, liver hepatocellular carcinoma.

las.org/about) is a Swedish-based program initiated in 2003, aiming to map all the human proteins in cells, tissues and organs, by integrating various omics technologies. The HPA consists of six separate parts: the Tissue Atlas, the Single Cell Type Atlas, the Pathology Atlas, the Brain Atlas, the Blood Atlas and the Cell Atlas, each focusing on a particular aspect of the genome-wide analysis of the human proteins. The Cell Type Atlas shows single cell RNA sequencing (scRNAseq) data from 13 different human tissues, together with immunohistochemically-stained tissue sections, allowing for visualization of the corresponding spatial protein expression patterns.

Results

SNHG4 expression in HCC

RNA sequencing (RNA-seq) datasets were obtained from TCGA. *SNHG4* was found to be upregulated in patients with HCC compared to healthy subjects (Fig. 2A). The HCC group was further stratified based on sex (Fig. 2B), age (Fig. 2C), individual cancer stages (Fig. 2D), tumor grade (Fig. 2E) and

promoter methylation levels (Fig. 2F) by UALCAN. Moreover, the OS and RFS of patients with HCC expressing high *SNHG4* levels and low *SNHG4* levels were significantly different (Fig. 3A and B). Based on results of InLocor 2.0, *SNHG4* was found to be mainly distributed in the nucleus of HCC cells. We then performed *in vitro* analysis of lncRNA *SNHG4* expression in cytoplasm and nucleus. RNA was extracted from the cytoplasm and nucleus of two different kinds of hepatoma cell lines (HepG2 and HuH-7) using the PARIS™ Kit, and subjected to reverse transcription for analysis of lncRNA *SNHG4* expression using real-time fluorescent quantitative PCR. These results showed that lncRNA *SNHG4* was mainly located in the nucleus (For details, see Supplementary Fig. 1).

Functional and signaling pathway analysis

GO annotation showed that *SNHG4*-related biological processes included RNA splicing, RNA processing and ribonucleoprotein complex biogenesis; cellular component included the spliceosomal complex and small subunit processome, while molecular functions involved ATPase, helicase and catalytic activity (Fig. 4A, B and C). In addition, KEGG

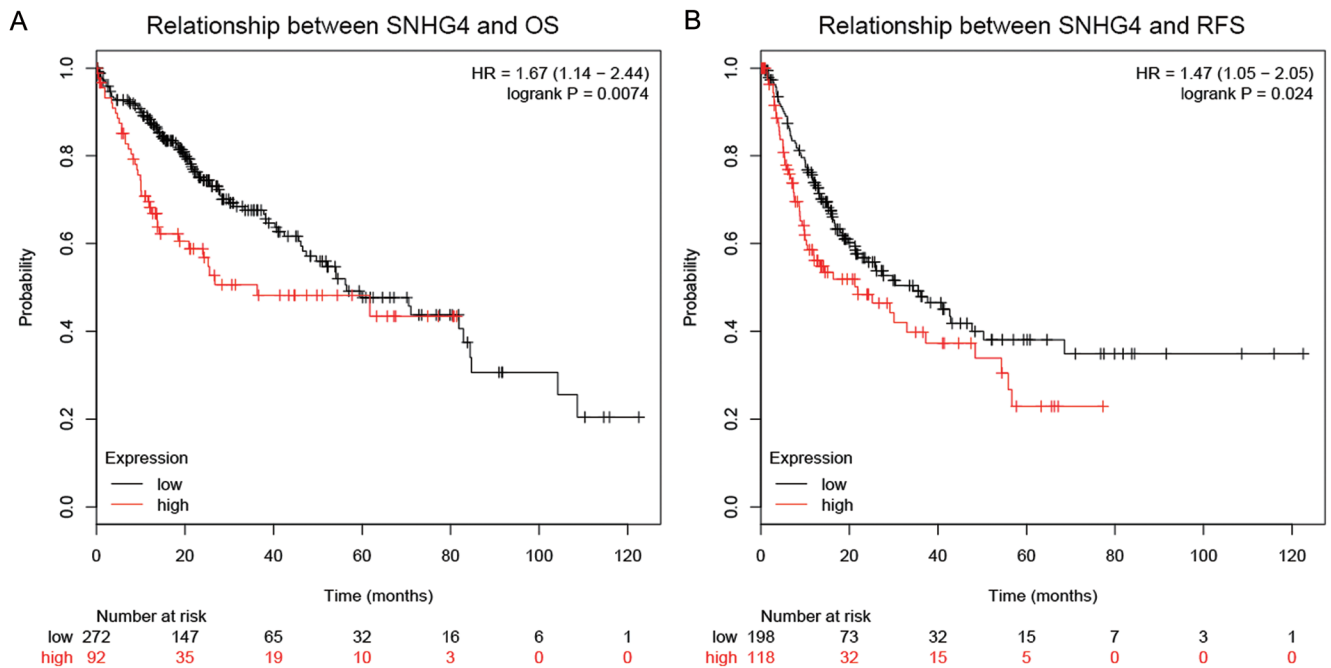


Fig. 3. Mantel-Cox survival curves. (A) Disease-specific survival (HCC-related death). (B) HCC recurrence-free survival. SNHG4, small nucleolar RNA host gene 4; HCC, hepatocellular carcinoma; HR, hazard ratio; OS, overall survival; RFS, recurrence-free survival.

pathway annotation revealed that RNA transport, ribosome biogenesis, cell cycle and RNA surveillance pathway were significantly enriched pathways (Fig. 4D). To identify the potential function of *SNHG4*, GSEA was conducted to search for biological processes enriched in samples with high levels of *SNHG4*. A total of 96 gene sets were found to be upregulated. Three gene sets, namely mitogen-activated protein kinase (MAPK)/ERK, hepatocyte growth factor (HGF)/MET and mTOR, were enriched in samples with high expression levels of *SNHG4* (Fig. 5A, B and C).

Analysis of genes, miRNA and transcription factor targets related to SNHG4 in HCC

LinkedOmics was used to analyze RNA sequencing data of 371 HCC patients from TCGA LIHC. A total of 8,744 genes showed significant correlation with *SNHG4* (Fig. 6A), with a false discovery rate of <0.01 and *p*-value < 0.05. We found 49 positive and 50 negative gene correlations with *SNHG4*, as shown in the heatmap in Figure 6B and C, respectively. Among these positively correlated genes, RING finger protein 44 (*RNF44*; Pearson’s correlation coefficient: 0.66, *p*=3.76×10⁻⁴⁷) and HEAT repeat-containing protein 1 (*HEATR1*; Pearson’s correlation coefficient: 0.64, *p*=2.3×10⁻⁴⁴) significantly correlated with *SNHG4*. KEGG and Wikipathway cancer enrichment analyses of genes interacting with *SNHG4*, which had been identified with LinkedOmics LinkFinder, were conducted in WebGestalt. Ribosome biogenesis in eukaryotes and the RNA processing and surveillance pathway were among the top terms (Fig. 6D and E). Oncomine analysis of the data from the Wurmbach Liver dataset validated the significant difference in expression of *RNF44* and *HEATR1* in HCC (Fig. 6F and G), and normal healthy liver tissue was significantly different; the levels of expression were based on the gene rank percentile, fold change and *p*-value. The target transcription factors and miRNAs are shown in Figure 7A and B, respectively. *SNHG4*-related miRNAs were

(ATAACCT) mir-154 [enrichment score (ES): 0.684; normalized enrichment score (NES): 1.834], (ACATTCC) mir-1, mir-206 (ES: 0.595; NES: 1.832) and (ATAGGAA) mir-202 (ES: 0.659; NES: 1.831), while *SNHG4*-related transcription factors included the E2F transcription factor (E2F) family, consisting of E2F1DP1_01 (ES: 0.622; NES: 1.896), E2F1DP2_01 (ES: 0.622; NES: 1.896) and E2F4DP2_01 (ES: 0.622; NES: 1.896). Moreover, by using TRAP web tools to validate the above findings, we obtained transcription factors related to *SNHG4*, which were found to be related to the E2F family. (See Supplementary Table 1; in the supplementary table, we ranked them from highest to lowest, according to corrected *p*-value.)

RBP’s related to SNHG4

A total of 65 *SNHG4* binding proteins were downloaded from CircLncRNAnet and 111 *SNHG4* binding proteins were downloaded from ENCORI. U2AF2 and BUD13 were coexistent in both databases and found to be expressed in HepG2 cells. Moreover, both U2AF2 and BUD13 had prognostic significance as shown in Kaplan–Meier plotter (Fig. 8A and C) and the HPA (Fig. 8B and D). Low expression of U2AF2 and BUD13 was found to be related to a better prognosis, while U2AF2 showed better correlation with *SNHG4* compared to BUD13 (Fig. 8E and F).

Discussion

HCC is considered to be one of the main causes of cancer mortality worldwide with high morbidity and mortality rates. HCC accounts for about 80–90% of liver cancer cases and globally results in an annual death toll of 600,000.²² Therefore, identifying novel prognostic biomarkers is essential to improving current HCC treatment. LncRNAs, which have recently attracted much attention, are dysregulated

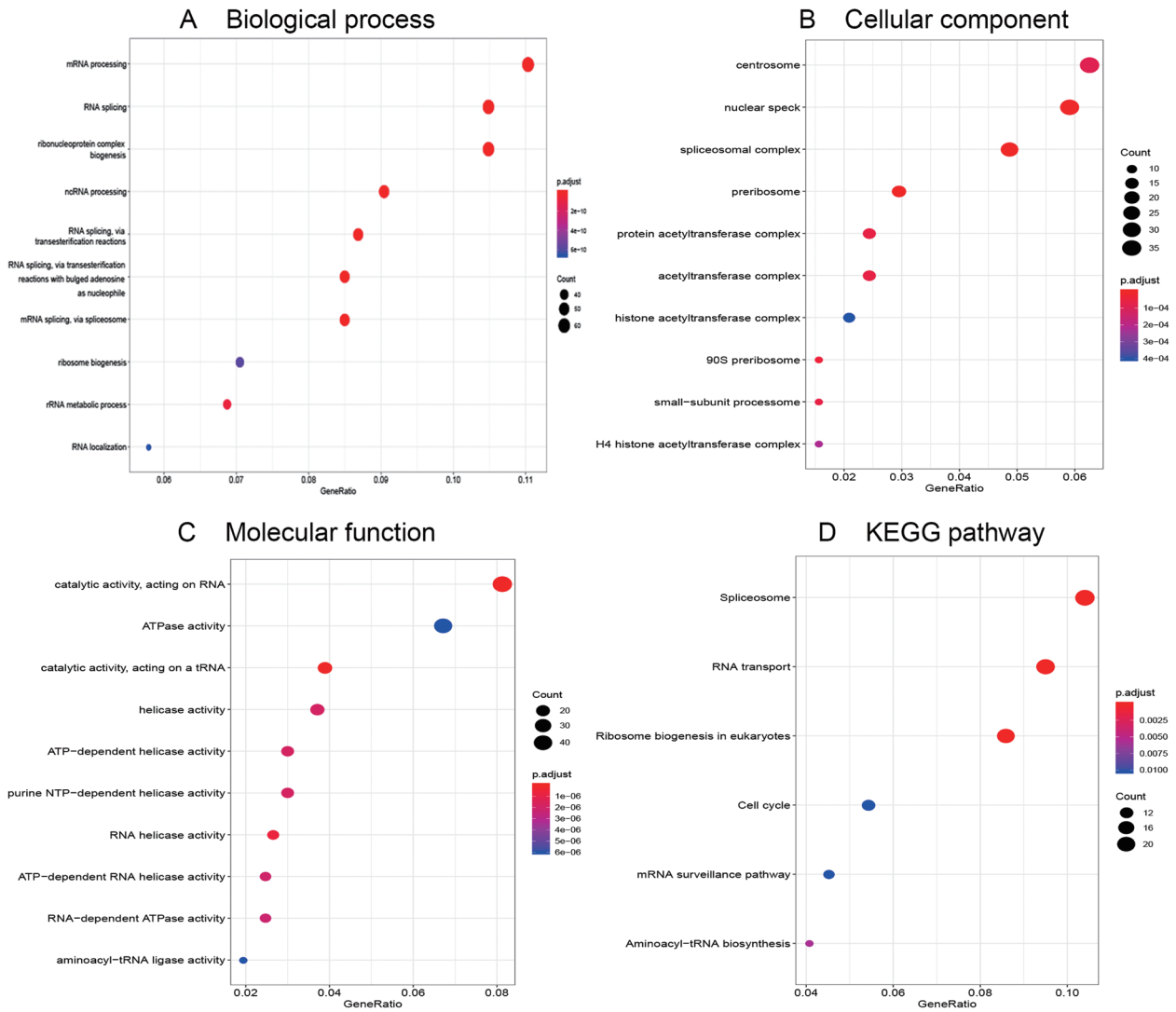


Fig. 4. GO and KEGG pathway of genes relevant to SNHG4 in HCC. (A) Biological process. (B) Cellular component. (C) Molecular function. (D) KEGG pathway enrichment of SNHG4. SNHG4, small nucleolar RNA host gene 4; HCC, hepatocellular carcinoma; KEGG, Kyoto Encyclopedia of Genes and Genomes.

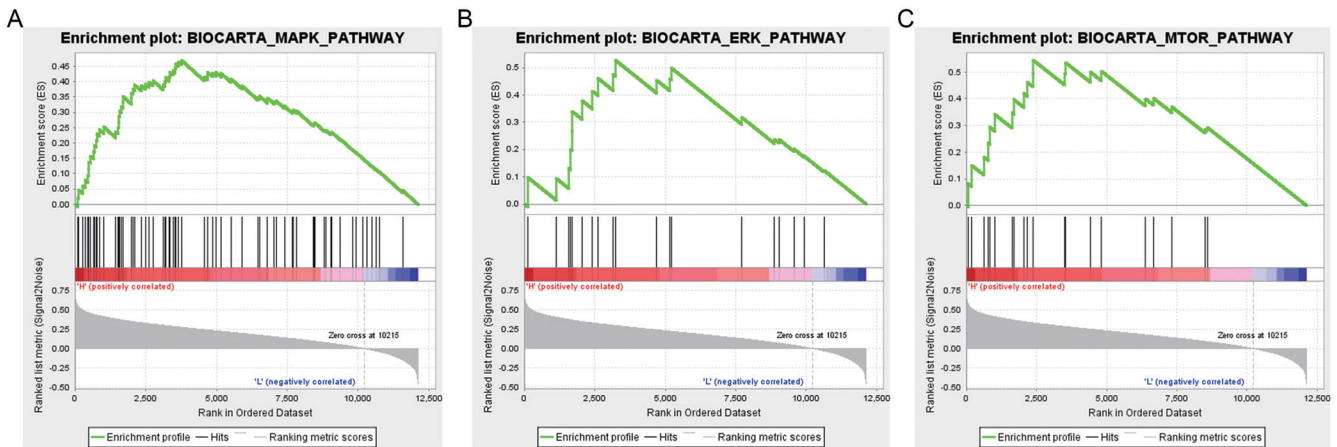


Fig. 5. GSEA for gene sets related with SNHG4 expression. SNHG4, small nucleolar RNA host gene 4; GSEA, Gene set enrichment analysis; MAPK, mitogen-activated protein kinase; ERK, extracellular-signal-regulated kinase; MTOR, mammalian target of rapamycin.

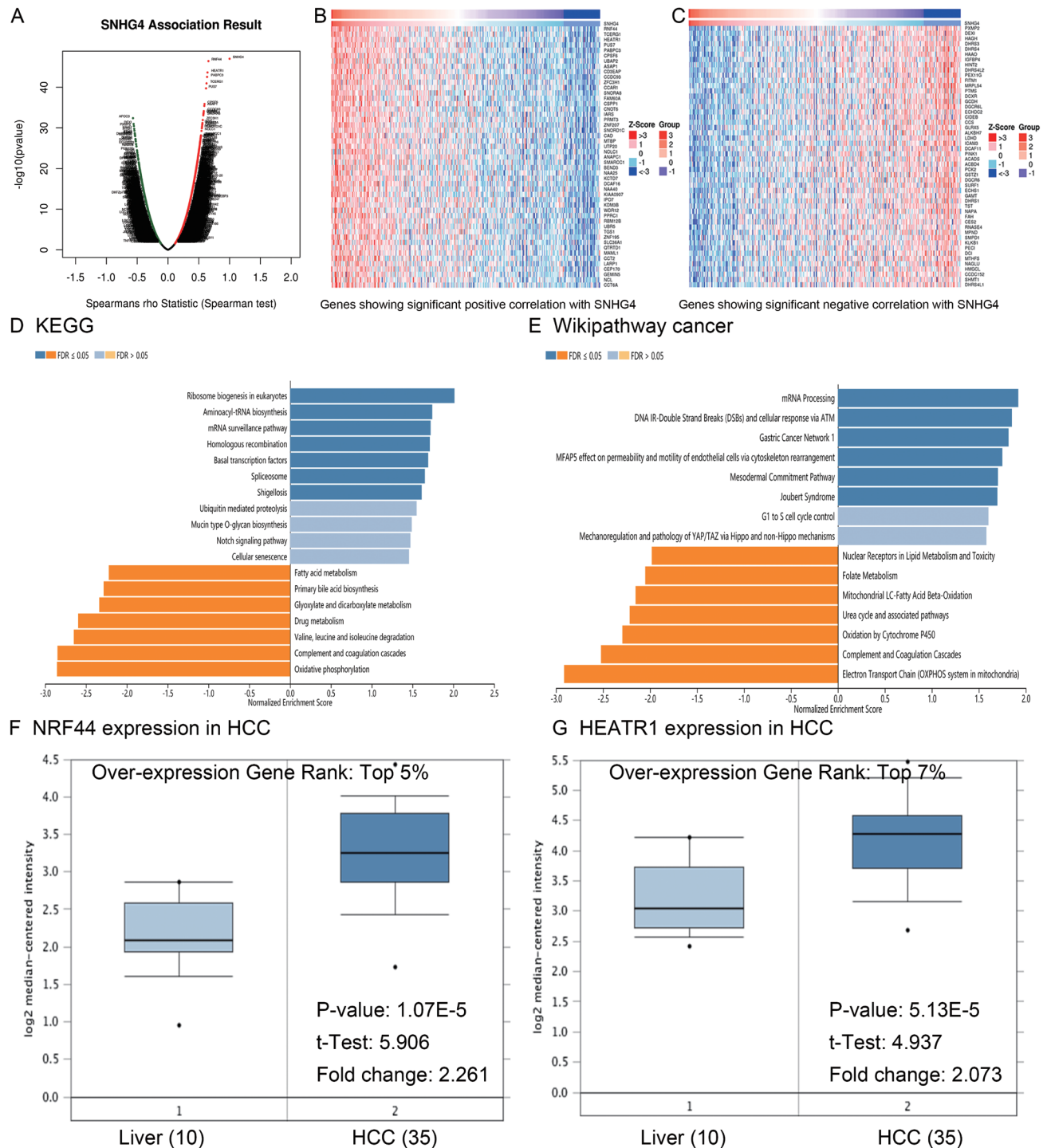


Fig. 6. Genes related to SNHG4 in HCC (LinkedOmics). (A) Volcano map showing the relationship between SNHG4 and the genes expressed in LIHC. Red indicates positive correlation; green indicates negative correlation. (B) Heat map showing genes positively correlated with SNHG4 in LIHC. (C) Heat map showing genes negatively correlated with SNHG4 in LIHC. (D, E) KEGG (D) and WikiPathway (E) analyses results of SNHG4-related genes. (F, G) OncoPrint analysis showed high expression of RNF44 (F) and HEATR1 (G) in HCC. SNHG4, small nucleolar RNA host gene 4; HCC, hepatocellular carcinoma; LIHC, liver hepatocellular carcinoma; KEGG, Kyoto Encyclopedia of Genes and Genomes; RNF44, RING finger protein 44; HEATR1, HEAT repeat-containing protein 1.

in a wide spectrum of cancers and have potential roles in tumorigenesis and tumor progression. Increasing evidence has demonstrated that snoRNA, a type of lncRNA, plays an important role in tumorigenesis.²³ Moreover, SNHG4 which

is an important type of snoRNA, has been reported to promote tumor growth and is a poor prognostic factor in patients with osteosarcoma.⁸

The significance of the role of SNHGs has been docu-

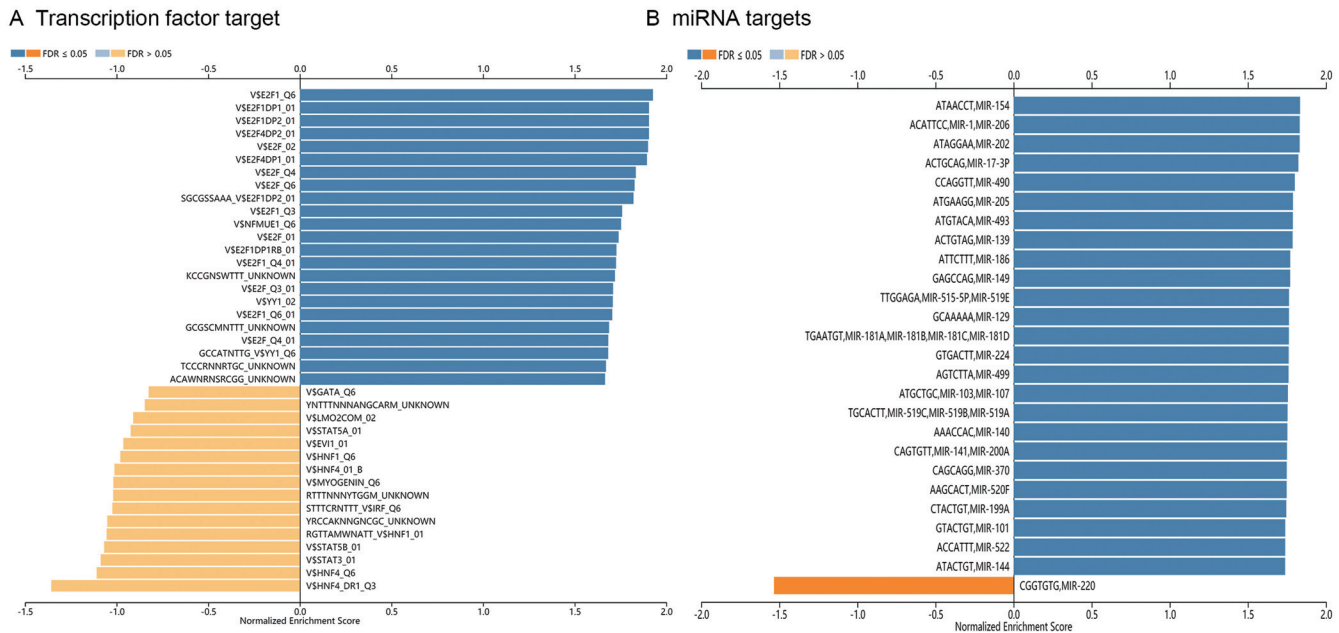


Fig. 7. miRNA and transcription factor targets related to SNHG4 in HCC (LinkedOmics). (A, B) SNHG4-related transcription factor targets and miRNAs, respectively. SNHG4, small nucleolar RNA host gene 4; HCC, hepatocellular carcinoma.

mented in the development of various cancers. For instance, SNHG6 has been reported to promote both cell proliferation in gastric cancer and tumor progression in HCC.²⁴ Additionally, knockdown of SNHG1 inhibited cell growth and metastasis of osteosarcoma *in vitro* and *in vivo*,⁶ while several genes including SNHG1, SNHG3, SNHG4, SNHG5, SNHG6, SNHG7, SNHG10, SNHG11 and SNHG12, were found to be upregulated in HCC. Increased SNHG4 expression was also found to be associated with shorter OS (HR: 1.319, 95% CI: 1.131–1.537, $p < 0.001$).⁹

The significant correlation between SNHG4 expression in liver cancer and histological type, histologic grade, stage, T classification and survival status, was reported by Jiao *et al.*;²⁵ however, only GSEA analysis was performed during that study. The present study performed a comprehensive bioinformatics analysis of SNHG4 using TCGA datasets and found that SNHG4 was significantly upregulated in patients with HCC and showed poor prognosis in terms of OS and RFS. Based on functional and pathway enrichment analyses, SNHG4 was identified to be mainly involved in cancer-related signaling pathways, potentially facilitating tumorigenesis and progression in HCC. KEGG pathway analysis results showed SNHG4 may be involved in RNA transport, ribosome biogenesis in cell cycle, mRNA surveillance pathway and aminoacyl-transfer RNA biosynthesis. Among the genes interacting with SNHG4, RNF44 and HEATR1 showed the most significant correlation. Interestingly, both have been previously reported to be involved in DNA repair, regulation of rRNA synthesis, RNA stability, and initiation of translation.^{26,27} Using WebGestalt, KEGG and Wikipathway cancer enrichment analyses of SNHG4-related genes were performed; significantly enriched processes included ribosome biogenesis in eukaryotes and the RNA processing and surveillance pathway, results similar to the KEGG pathway analysis of SNHG4 with R software.

MiRNAs are short ncRNA sequences containing ~22 nucleotides that act as important regulators of gene expression by specifically binding and cleaving mRNAs or inhibiting their translation.²⁸ SNHG4 was reported to promote tumor growth in patients with osteosarcoma, where miRNA-154

is downregulated in liver tissue and can inhibit tumorigenesis and the G1/S transition in cancer cells when overexpressed.²⁹ miRNA-206, which is downregulated in human HCC tissues, and abnormally increased miR-206 expression (in HCC cell lines) have been reported to attenuate cell viability, migration, invasion and increased apoptosis.³⁰ E2F1 plays a crucial role in the control of cell cycle and action. E2F1 acts as a transcription activator for E2F target genes, and can regulate transcriptional activities by binding to the promoters of a target gene via their response elements during the cell cycle progression of the late G1/S phase.³¹ Studies have shown that SNHG3 could be transcriptionally activated by E2F1. Choi *et al.*³² reported that E2F1 acts as an inhibitor on the hepatitis B virus life cycle, which mediated HCC by overcoming the control of virus-encoded HBx function on the host p53 promoter, while E2F1 activated the p53 promoter via the E2F1 binding site. Moreover, Palaiologou *et al.*³³ reported E2F1 overexpression and its pro-apoptotic nature in human HCC.

It has long been established that alternative splicing contributes to tumorigenesis by producing splice isoforms which can stimulate cell proliferation and cell migration or induce resistance to apoptosis and anticancer agents.³⁴ Among RBPs interacting with SNHG4, low expression of BUD13 and U2AF2 showed good prognosis in HCC patients, while U2AF2 showed good correlation with SNHG4 ($r = 0.6745$, $p < 0.001$). U2AF, comprised of a large and a small subunit, is a non small nuclear ribonucleoprotein particles (snRNP) protein required for the binding of U2 snRNP to the pre-mRNA branch site. Pre-mRNA splicing is a crucial step in eukaryotic gene expression. The recognition of the splice sites is initiated by U2AF2, which binds to the poly-pyrimidine tract (referred to also as Py-tract) RNA upstream of exons to assemble the spliceosome. Li *et al.* found that U2AF2 was related to activation of the AKT/mTOR pathway in non-small cell lung cancer.³⁵ Similarly, from our GSEA analysis results, we can infer that the SNHG4 related-biological pathways were MAPK/ERK and mTOR, the former presenting with a significantly higher number of gene hits before the peak (see Fig. 5A) and hence contributing more to the

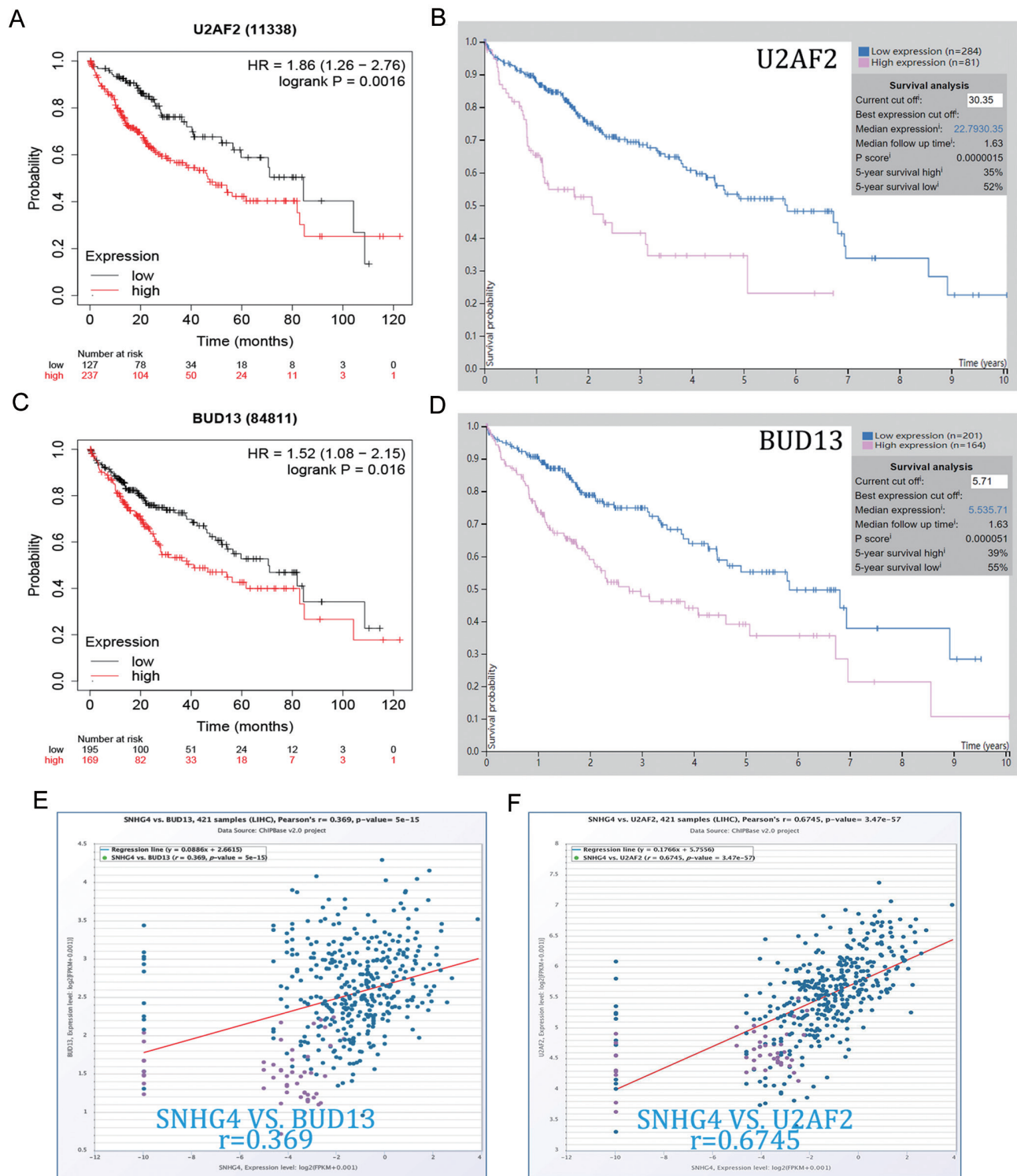


Fig. 8. Prognostic value of U2AF2 and BUD13 in LIHC patients. (A-D) Low expression of U2AF2 and BUD13 was associated with better prognosis. (E, F) Correlation of BUD13 (E) and U2AF2 (F) with SNHG4. SNHG4, small nucleolar RNA host gene 4; LIHC, liver hepatocellular carcinoma; U2AF2, U2 auxiliary factor 2.

enrichment score of the gene set. This suggests that SNHG4 might promote tumor progression by regulating the MAPK/ERK pathway. Both MAPK/ERK and mTOR have been reported as majorly important cellular signaling pathways during

the genesis and development of HCC.³⁶ Taking into account that numerous studies have found that SNHG4 plays a crucial role in carcinogenesis,^{8,9} it can be inferred that SNHG4 contributes to the progression of HCC and influences prog-

nosis by regulating the MAPK/ERK or mTOR pathway.

SNHG4 has the potential to be the basis of future drug treatment or gene therapy in HCC. The present findings suggest that *SNHG4* may play a critical role in hepatocarcinogenesis by influencing different steps of gene expression, from RNA splicing and ribosome biogenesis to the mRNA processing and surveillance pathway. A limitation of our study is the lack of validation at experimental level, which we shall perform in the future. Secondly, HCC has many different causes, which may lead to different pathogenesis. We did not explore the role of *SNHG4* in HCC of different etiology. However, based on patient medical history and risk factors of HCC, patients in our study could not be simply subclassified into distinct groups, such as virus-induced HCC and alcohol-related HCC, since for most patients HCC was found to be of mixed etiology, as is often the case at clinical level. To our knowledge, this is a limitation of any study on HCC involving data from the TCGA database. In addition, the findings of the present study are preliminary and purely descriptive.

Funding

This work was supported by grants from the National key research and development program (2018YFC1315400) and the National Natural Science Foundation of China (Nos. 81773176, 81870449).

Conflict of interest

The authors have no conflict of interests related to this publication.

Author contributions

Designed the study and take responsibility for the integrity of the data and the accuracy of the data analysis (JG, YJ, YC), contributed to collection and analysis of the data (JC, CX, CJTHF, DL, XL), and wrote the manuscript (JC, CX, CJTHF). All authors contributed to data interpretation, and reviewed and approved the final version.

Data sharing statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

References

- Zhou J, Sun HC, Wang Z, Cong WM, Wang JH, Zeng MS, *et al*. Guidelines for diagnosis and treatment of primary liver cancer in China (2017 Edition). *Liver Cancer* 2018;7(3):235–260. doi:10.1159/000488035.
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68(6):394–424. doi:10.3322/caac.21492.
- Roderburg C, Luedde T. The role of the gut microbiome in the development and progression of liver cirrhosis and hepatocellular carcinoma. *Gut Microbes* 2014;5(4):441–445. doi:10.4161/gmic.29599.
- Siprashvili Z, Webster DE, Johnston D, Shenoy RM, Ungewickell AJ, Bhaduri A, *et al*. The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. *Nat Genet* 2016;48(1):53–58. doi:10.1038/ng.3452.
- Li Z, Qiu R, Qiu X, Tian T. SNHG6 promotes tumor growth via repression of P21 in colorectal cancer. *Cell Physiol Biochem* 2018;49(2):463–478. doi:10.1159/000492986.
- Wang J, Cao L, Wu J, Wang Q. Long non-coding RNA SNHG1 regulates NOB1 expression by sponging miR-326 and promotes tumorigenesis in osteosarcoma. *Int J Oncol* 2018;52(1):77–88. doi:10.3892/ijo.2017.4187.
- You J, Fang N, Gu J, Zhang Y, Li X, Zu L, *et al*. Noncoding RNA small nucleolar RNA host gene 1 promote cell proliferation in nonsmall cell lung cancer. *Indian J Cancer* 2014;51(Suppl 3):e99–e102. doi:10.4103/0019-509X.154092.
- Xu R, Feng F, Yu X, Liu Z, Lao L. LncRNA SNHG4 promotes tumour growth by sponging miR-224-3p and predicts poor survival and recurrence in human osteosarcoma. *Cell Prolif* 2018;51(6):e12515. doi:10.1111/cpr.12515.
- Zhu Q, Yang H, Cheng P, Han Q. Bioinformatic analysis of the prognostic value of the lncRNAs encoding snoRNAs in hepatocellular carcinoma. *Biofactors* 2019;45(2):244–252. doi:10.1002/biof.1478.
- Cheng XB, Zhang T, Zhu HJ, Ma N, Sun XD, Wang SH, *et al*. Knockdown of lncRNA SNHG4 suppresses gastric cancer cell proliferation and metastasis by targeting miR-204-5p. *Neoplasia* 2021;68(3):546–556. doi:10.4149/neo_2021_200914N981.
- Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA cross-talk and competition. *Nature* 2014;505(7483):344–352. doi:10.1038/nature12986.
- Deng M, Bragelmann J, Schultze JL, Perner S. Web-TCGA: an online platform for integrated analysis of molecular cancer data sets. *BMC Bioinformatics* 2016;17:72. doi:10.1186/s12859-016-0917-9.
- Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, *et al*. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 2017;19(8):649–658. doi:10.1016/j.neo.2017.05.002.
- Lin Y, Liang R, Qiu Y, Lv Y, Zhang J, Qin G, *et al*. Expression and gene regulation network of RBM8A in hepatocellular carcinoma based on data mining. *Aging (Albany NY)* 2019;11(2):423–447. doi:10.18632/aging.101749.
- Lin Y, Pan X, Shen HB. InLocatator 2.0: a cell-line-specific subcellular localization predictor for long non-coding RNAs with interpretable deep learning. *Bioinformatics* 2021;btab127. doi:10.1093/bioinformatics/btab127.
- Mongre RK, Jung S, Mishra CB, Lee BS, Kumari S, Lee MS. Prognostic and clinicopathological significance of SERTAD1 in various types of cancer risk: a systematic review and retrospective analysis. *Cancers (Basel)* 2019;11(3):337. doi:10.3390/cancers11030337.
- Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res* 2018;46(D1):D956–D963. doi:10.1093/nar/gkx1090.
- Thomas-Chollier M, Hufton A, Heinig M, O’Keeffe S, Masri NE, Roeder HG, *et al*. Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nat Protoc* 2011;6(12):1860–1869. doi:10.1038/nprot.2011.409.
- Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, *et al*. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 2004;6(1):1–6. doi:10.1016/s1476-5586(04)80047-2.
- Wu SM, Liu H, Huang PJ, Chang IY, Lee CC, Yang CY, *et al*. circLncRNA.net: an integrated web-based resource for mapping functional networks of long or circular forms of noncoding RNAs. *Gigascience* 2018;7(1):1–10. doi:10.1093/gigascience/gix118.
- Fan K, Zebisch A, Horny K, Schrama D, Becker JC. Highly expressed miR-375 is not an intracellular oncogene in Merkel cell polyomavirus-associated Merkel cell carcinoma. *Cancers (Basel)* 2020;12(3):529. doi:10.3390/cancers12030529.
- Roberts LR. Sorafenib in liver cancer—just the beginning. *N Engl J Med* 2008;359(4):420–422. doi:10.1056/NEJMe0802241.
- Gong J, Li Y, Liu CJ, Xiang Y, Li C, Ye Y, *et al*. A pan-cancer analysis of the expression and clinical relevance of small nucleolar RNAs in human cancer. *Cell Rep* 2017;21(7):1968–1981. doi:10.1016/j.celrep.2017.10.070.
- Chang L, Yuan Y, Li C, Guo T, Qi H, Xiao Y, *et al*. Upregulation of SNHG6 regulates ZEB1 expression by competitively binding miR-101-3p and interacting with UPF1 in hepatocellular carcinoma. *Cancer Lett* 2016;383(2):183–194. doi:10.1016/j.canlet.2016.09.034.
- Jiao Y, Li Y, Jia B, Chen Q, Pan G, Hua F, *et al*. The prognostic value of lncRNA SNHG4 and its potential mechanism in liver cancer. *Biosci Rep* 2020;40(1):BSR20190729. doi:10.1042/BSR20190729.
- He S, Ma X, Ye Y, Zhang M, Zhuang J, Song Y, *et al*. HEATR1 modulates cell survival in non-small cell lung cancer via activation of the p53/PUMA signaling pathway. *Onco Targets Ther* 2019;12:4001–4011. doi:10.2147/OTT.S195826.
- Li YY, Wu C, Shah SS, Chen SM, Wangpaichitr M, Kuo MT, *et al*. Degradation of AMPK- α 1 sensitizes BRAF inhibitor-resistant melanoma cells to arginine deprivation. *Mol Oncol* 2017;11(12):1806–1825. doi:10.1002/1878-0261.12151.
- Wang W, Peng B, Wang D, Ma X, Jiang D, Zhao J, *et al*. Human tumor microRNA signatures derived from large-scale oligonucleotide microarray datasets. *Int J Cancer* 2011;129(7):1624–1634. doi:10.1002/ijc.25818.
- Zhou H, Zhang M, Yuan H, Zheng W, Meng C, Zhao D. MicroRNA-154 functions as a tumor suppressor in osteosarcoma by targeting Wnt5a. *Oncol Rep* 2016;35(3):1851–1858. doi:10.3892/or.2015.4495.
- Wang Y, Tai Q, Zhang J, Kang J, Gao F, Zhong F, *et al*. MiRNA-206 inhibits hepatocellular carcinoma cell proliferation and migration but promotes apoptosis by modulating cMET expression. *Acta Biochim Biophys Sin (Shanghai)* 2019;51(3):243–253. doi:10.1093/abbs/gmy119.
- Cam H, Dynlacht BD. Emerging roles for E2F: beyond the G1/S transition and DNA replication. *Cancer Cell* 2003;3(4):311–316. doi:10.1016/s1535-6108(03)00080-1.
- Choi M, Lee H, Rho HM. E2F1 activates the human p53 promoter and overcomes the repressive effect of hepatitis B viral X protein (Hbx) on the p53 pro-

- moter. *IUBMB Life* 2002;53(6):309–317. doi:10.1080/15216540213466.
- [33] Palaiologou M, Koskinas J, Karanikolas M, Fatourou E, Tiniakos DG. E2F-1 is overexpressed and pro-apoptotic in human hepatocellular carcinoma. *Virchows Arch* 2012;460(5):439–446. doi:10.1007/s00428-012-1220-4.
- [34] Shkreta L, Bell B, Revil T, Venables JP, Prinos P, Elela SA, *et al*. Cancer-associated perturbations in alternative pre-messenger RNA splicing. *Cancer Treat Res* 2013;158:41–94. doi:10.1007/978-3-642-31659-3_3.
- [35] Li J, Cheng D, Zhu M, Yu H, Pan Z, Liu L, *et al*. OTUB2 stabilizes U2AF2 to promote the Warburg effect and tumorigenesis via the AKT/mTOR signaling pathway in non-small cell lung cancer. *Theranostics* 2019;9(1):179–195. doi:10.7150/thno.29545.
- [36] Whittaker S, Marais R, Zhu AX. The role of signaling pathways in the development and treatment of hepatocellular carcinoma. *Oncogene* 2010;29(36):4989–5005. doi:10.1038/onc.2010.236.