

RESEARCH ARTICLE

Open Access



Expression and gene regulatory network of SNHG1 in hepatocellular carcinoma

Chaoran Zheng^{1*} and Shicheng Yu^{1,2,3*}

Abstract

Background: Small nucleolar RNA host gene 1 (SNHG1), a long noncoding RNA (lncRNA), is a transcript that negatively regulates tumour suppressor genes, such as p53. Abnormal SNHG1 expression is associated with cell proliferation and cancer. We used sequencing data downloaded from Genomic Data Commons to analyse the expression and interaction networks of SNHG1 in hepatocellular carcinoma (HCC).

Methods: Expression was examined using the limma package of R and verified by Gene Expression Profiling Interactive Analysis. We also obtained miRNA expression data from StarBase to determine the lncRNA-miRNA-mRNA-related RNA regulatory network in HCC. Kaplan–Meier (KM) analysis was performed using the survival package of R. Gene Ontology annotation of genes was carried out using Metascape.

Results: We found that SNHG1 was overexpressed and often amplified in HCC patients. In addition, SNHG1 upregulation was associated with the promotion of several primary biological functions, including cell proliferation, transcription and protein binding. Moreover, we found similar trends of small nucleolar RNA host gene 1 (SNHG1), *E2F8* (E2F transcription factor 8), *FANCE* (FA complementation group E) and *LMNB2* (*encodes lamin B2*) expression. In the SNHG1-associated network, high expression levels of SNHG1 (log-rank P value = 0.0643), *E2F8* (log-rank P value = 0.000048), *FANCE* (log-rank P value = 0.00125) and *LMNB2* (log-rank P value = 0.0392) were significantly associated with poor survival. Single-cell analysis showed that *E2F8* may play an important role in tumorigenesis or cancer development.

Conclusions: Our results highlight the benefit of utilizing multiple datasets to understand the functional potential regulatory networks of SNHG1 and the role of SNHG1 in tumours.

Keywords: LIHC, HCC, TCGA, ceRNA, SNHG1, Survival

Background

Liver cancer is the fourth leading cause of death from cancer in the world [1] and is one of the most fatal and algetic tumours. Hepatocellular carcinoma (HCC) is the most common histological type of primary liver cancer, occurring in 75–85% of patients with primary liver cancer, and Africa and Asia have the highest incidences [1,

2]. Recent research shows that HCC patients tend to have a family history of HCC and that the pathological mechanisms are complex [3, 4]. Overall, medical practice has changed rapidly in the last few years with the aid of rapidly expanding genome sequencing techniques [5], and precision medicine is gradually becoming an emerging field in the treatment and prevention of diseases. Although medical researchers have overcome the challenges of assessing large amounts of data, the application of genetic testing in cancer treatment is progressing slowly. A primary reason for this slow progress is that large amounts of accumulated patient data are independent and discrete. Thus, researchers need to determine

*Correspondence: zhengchaoran@whu.edu.cn; yu_shicheng@gibh.ac.cn; shyu@whu.edu.cn

¹ Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China

Full list of author information is available at the end of the article



relationships between diverse data and their correlations with disease.

Noncoding RNAs (ncRNAs) are a class of RNA molecules that do not participate in protein coding but do regulate gene expression, contributing to many physiological and pathological processes, including tumours [6]. Long noncoding RNAs (lncRNAs) are an important subcategory of ncRNAs with lengths longer than 200 nucleotides [7, 8]. Recent studies show that lncRNAs participate in various biological processes, such as cell differentiation [9], signal transduction [10] and posttranscriptional regulation [11], that are conducive to the progression of tumours and many other diseases [12]. Increasing evidence suggests that lncRNAs can function alone or in cooperation with microRNAs (miRNAs). Typically, lncRNAs act as ceRNAs to regulate messenger RNAs (mRNAs). This mode is known as the “lncRNA-miRNA-mRNA network”. This regulation has been reported in some tumour-related studies, but traditional research methods are insufficient for analysing multi-omics data [13, 14]. Nevertheless, ceRNA network analysis can shed light on ncRNA expression alterations in cancers and uncover relationships between altered RNAs.

To reveal the roles of the lncRNA-miRNA-mRNA network in human HCC, we sought to identify ceRNA networks by identifying similar regulatory networks from an extensive collection of data from Genomic Data Commons (GDC) using a stringent bioinformatics pipeline. The identified ceRNA networks may help in identifying molecular subtypes or new biological markers to classify HCC patients. Moreover, by examining differences in patient survival based on the expression of each RNA subtype, our ceRNA networks offer potential clinical utility for personalized treatments and therapeutic targets.

Methods

Data downloading and organization

The `gdc-client` method was used to download RNA-seq, miRNA and clinical data from the GDC database via `TCGA-Assembler` (<http://www.compgenome.org/TCGA-Assembler/>) of R (3.6.2) [15]. A total of 424 RNA-seq, 425 miRNA and 377 clinical data points were included in the study. Duplicates were filtered out before analysis of the metadata download from the GDC database. We also removed samples that were not from primary tumours or normal solid tissues. The TMM method implemented in `edgeR` of R (3.6.2) [16] was used to normalize RNA-seq and miRNA raw count data, after which we removed genes with a `logcpm` of expression below 1 in more than half of the samples. We used the `limma` [17] package in R (3.6.2) (<https://www.r-project.org/>) to screen genes and miRNAs showing significant differential expression between tumour and normal tissues.

Significant differentially expressed genes (DEGs) (based on $P < 0.01$ and fold change (FC) > 2) were identified in the RNA-seq and miRNA expression profiles, and volcano plots and bar plots were constructed in R to visualize these DEGs. Expression analysis of different sample types (tumour vs. normal tissues) was performed with GEPIA (<http://gepia.cancer-pku.cn/>) [18]. Next, we assessed the levels of protein expression for these genes based on immunohistochemistry data obtained from the Human Protein Atlas database (<https://www.proteinatlas.org/>) [19–21].

ceRNA network analysis

Competing lncRNA-mRNA pairs were identified using the `gdcCEAnalysis` function of `GDCRNATools` [15]. The precise identification of lncRNA-miRNA-mRNA interactions was based on three lines of evidence. First, the hypergeometry of shared miRNAs was evaluated to reduce the number of lncRNA-mRNA pairs; hypergeometric analysis was performed to determine whether a lncRNA-mRNA pair shares many miRNAs. Second, we excluded all lncRNA and mRNA pairs that exhibited variable correlations in expression. Finally, the lncRNA-mRNA pairs retained were required to have regulation similar to that of all shared miRNAs. After these steps, several lncRNA-miRNA-mRNA interactions were identified, and functional analysis was then carried out. To identify the number of miRNAs related to lncRNAs and mRNAs, miRNA-mRNA and miRNA-lncRNA interactions were obtained from StarBase v2.0 [22]. The network was visualized using `Cytoscape` (3.4.0) (<https://cytoscape.org>) [23, 24].

Univariate survival analysis

`GdcSurvival Analysis` in `GDCRNATools` was employed to analyse the association with overall survival for each gene [15]. First, we divided the patients into high- and low-expression groups based on the median expression level. Then, Kaplan–Meier (KM) analysis was performed using the survival package, and the results were examined in univariate survival analysis. The final results are presented as the hazard ratio (HR) and significance (log-rank P value).

Functional analysis

Gene Ontology (GO) annotation for the genes was carried out using the `Metascape` database (<http://metascape.org/gp/index.html#/main/step1>) [25]. Then, Online Mendelian Inheritance in Man (OMIM) (<https://www.omim.org/>) was applied to analyse associations between diseases and the genes in the lncRNA-miRNA-mRNA network.

Single-cell analysis

To assess gene expression in different cell types, we obtained single-cell sequencing data from a previous report [26]. Gene expression was normalized by $\log_2(\text{filteredexpr} + 1)$ as mentioned in [26]. Cell clusters were generated with *seurat* software [27], and the cell clustering information was retrieved from the same report [26]. The expression of these genes in endothelial cells and macrophages was evaluated, and the results of gene expression in different cell clusters are illustrated in a violin plot.

Results

RNA expression in HCC

Using the *limma* package, we identified 807 upregulated DEGs, 1590 downregulated DEGs, 59 upregulated differentially expressed miRNAs (DEMs), and 71 downregulated DEMs (Fig. 1). In total, we identified 2397 DEGs and 130 DEMs.

Biological interaction networks in HCC

We next sought to determine the biological interaction networks in HCC by using internally incorporated miRNA-mRNA and miRNA-lncRNA interactions, and we identified seven mRNA-miRNA-lncRNA networks (Fig. 2). All the ceRNA networks share similar regulation of all miRNAs interacting with the lncRNA-mRNA pairs. For example, the neighbouring genes of *MAGI2-AS3*, namely, *ADAMTSL3*, *AUTS2*, *BHLHE40*, *DUSP6*, *ERRFI1*, *GADD45A*, *GNPNAT1*, *HIVEP1*, *MYLK*,

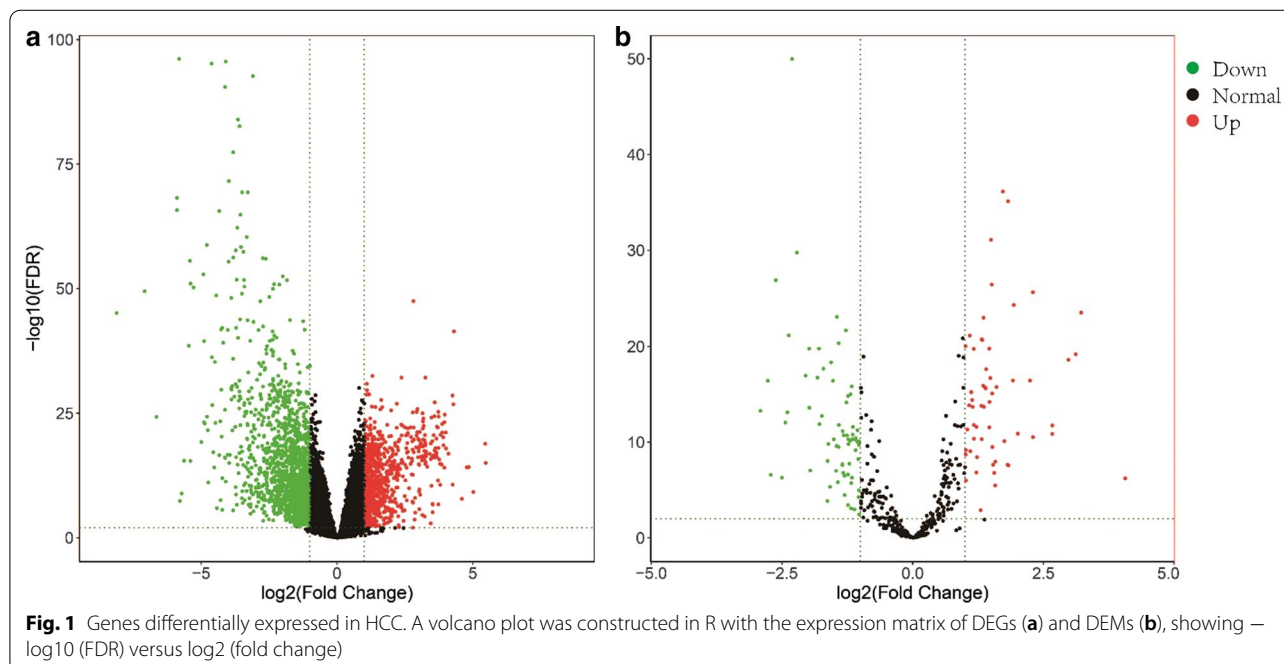
MYO10, *PNRC1*, *RIPOR2*, *SMAD6*, *SPRY2*, and *SYBU*, share the same miRNAs. Additionally, two miRNAs (*hsa-miR-374a-5p* and *hsa-miR-374b-5p*) mediated the *MAGI2-AS3*-mRNA network.

Univariate survival analysis of coexpressed genes in HCC

Abnormal expression of many genes affects the duration of human survival. By constructing a ceRNA network, we identified hub genes with which miRNAs might interact. In addition, we established survival-related networks. Then, we focused on *SNHG1* and related genes that predict the same direction of correlation with prognosis.

We evaluated *SNHG1* transcription levels in HCC studies from GDC. Data in the GDC database revealed significantly higher mRNA expression of *SNHG1* in HCC samples than in normal solid tissue samples ($\log\text{-rank P value} = 6.43\text{E-}18$) (Table 1), with a fold difference above 2. The results of the expression analysis for *SNHG1*, *LMNB2*, *FANCE*, and *E2F8* were verified by GEPIA (Fig. 3). We also verified the low protein levels of these four genes in normal samples using the Human Protein Atlas database (Fig. 4). Four miRNAs (*hsa-miR-421*, *hsa-miR-377-3p*, *hsa-miR-330-5p*, and *hsa-miR-326*) mediated the *SNHG1* network. Thus, *SNHG1* expression may serve as a prognostic indicator in HCC.

To better understand the functions of the coexpressed genes (*E2F8*, *FANCE* and *LMNB2*) correlated with *SNHG1*, we compared survival time among patients with different levels of gene expression and found that high expression of *SNHG1* was associated with decreased



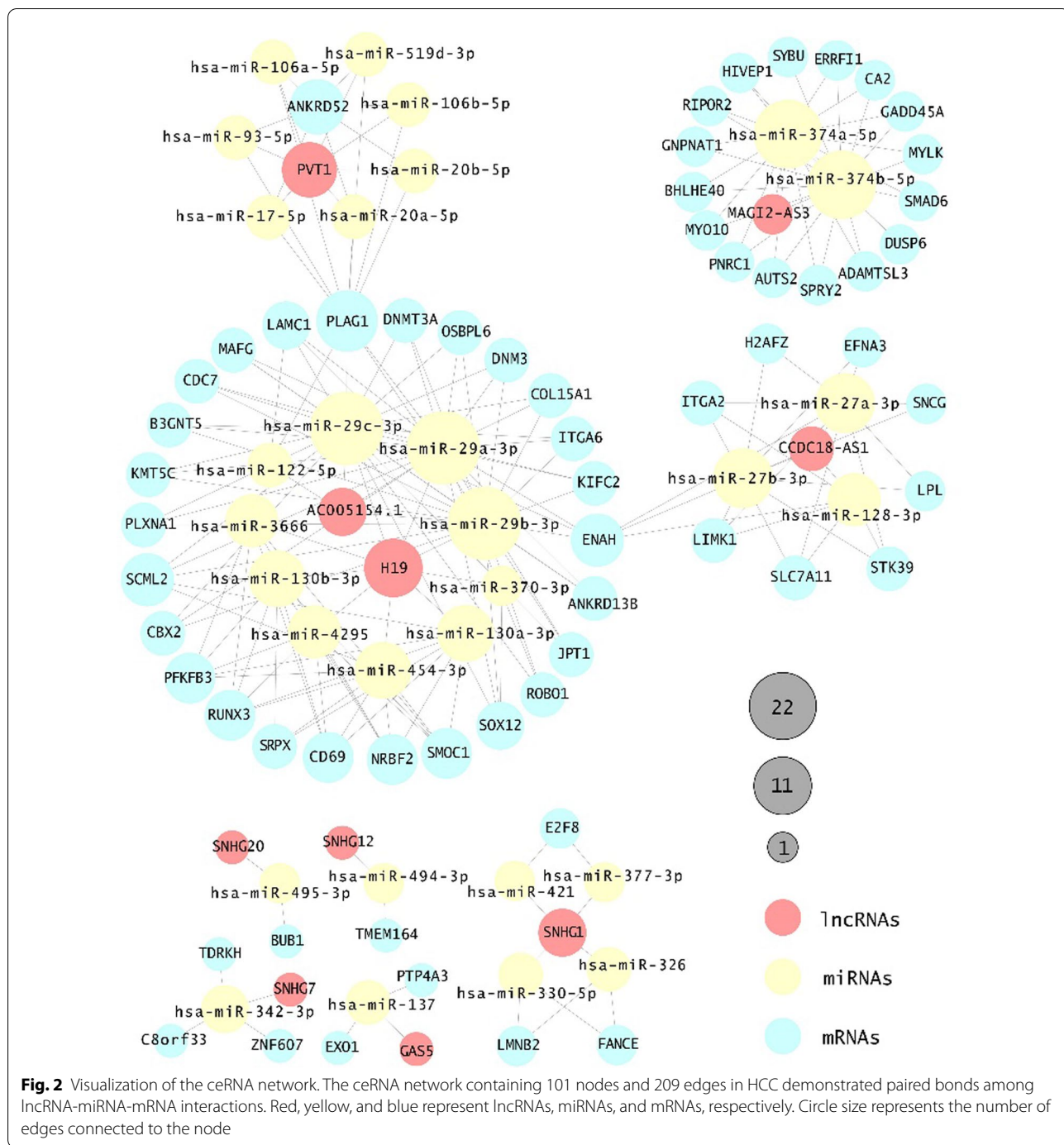
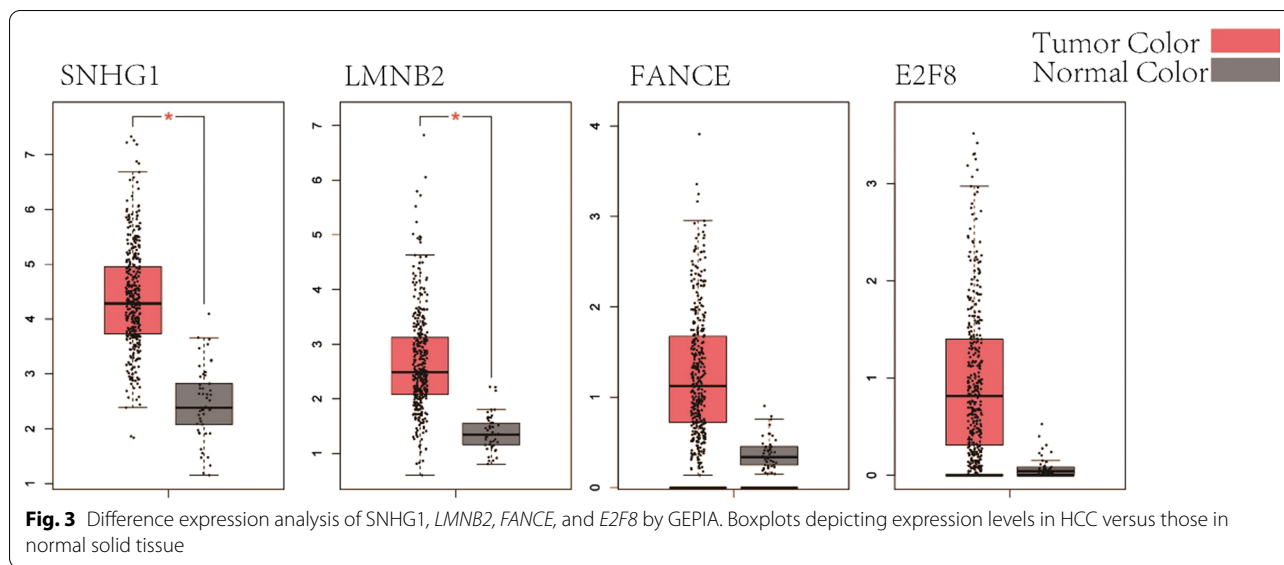


Table 1 Biological functions of the genes in the SNHG1-related network

Gene symbol	Description	OMIM	Type	LogFC	P value
SNHG1	Small nucleolar RNA host gene 1	603222	Long non coding	1.495397836	6.43E-18
LMNB2	E2F transcription factor 8	612047	Protein coding	1.115366226	5.95E-13
FANCE	FA complementation group E	613976	Protein coding	1.459672439	9.12E-14
E2F8	Lamin B2	150341	Protein coding	3.729047657	8.89E-27

The table lists 5 genes in the SNHG1-related network and the following information: Gene symbol, Description, OMIM, Type, LogFC (trimmed binary logarithm of average expression level fold change), and P value



survival (log-rank P value = 0.0643) (Fig. 5). Additionally, in the *SNHG1*-associated network, high expression levels of *E2F8* (log-rank P value = 0.000048), *FANCE* (log-rank P value = 0.00125) and *LMNB2* (log-rank P value = 0.0392) correlated significantly with poor survival (Fig. 5); these genes affect transcription activity, DNA repair, and protein binding, respectively. Overall, we detected interaction effects of *SNHG1* and the other three genes on survival.

Functional analysis of the coexpressed genes correlated with *SNHG1* in HCC

To examine the roles that the 3 coexpressed genes we identified play in the cell, we annotated all three genes and *SNHG1* using the Custom Analysis function of Metascape (<http://metascape.org/>) (Table 1). All of these genes are involved in primary life processes, such as transcription activity, DNA repair, and protein binding. Two of these genes encode plasma proteins, and *E2F8* encodes a transcription factor with a winged helix-turn-helix DNA-binding domain.

SNHG1-related mRNA expression in endothelial and macrophage populations

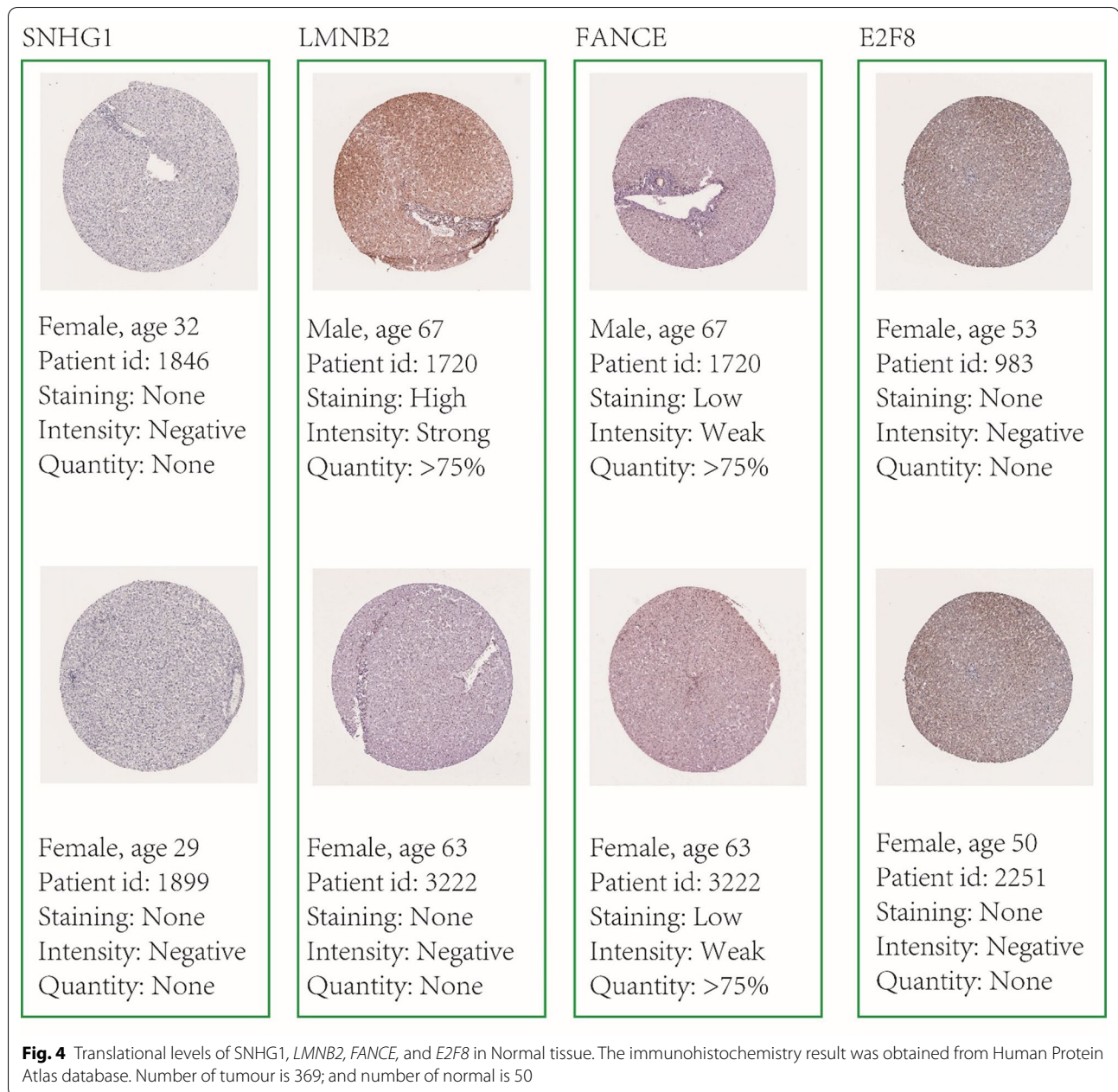
We analysed the expression of the three coexpressed genes described above in endothelial cells (ECs) and macrophages, which were significantly different based on single-cell data [26] (Figs. 6 and 7). *LMNB2* was found to be uniformly expressed in the 11 EC clusters. Specifically, *E2F8* was expressed at low levels in clusters 0, 1, 4, and 8, corresponding to normal cells [26]. At the same time, *FANCE* was expressed at low levels in clusters 8 (adjacent normal), 9 (core tumour), and 10 (peripheral

tumour), indicating no difference between tumour and normal cells. Mononuclear phagocyte-based subclustering analysis also suggested low levels of *E2F8* expression in peripheral tumours and adjacent normal tissues. However, *LMNB2* and *FANCE* were uniformly expressed in the 9 macrophage clusters. These results suggest that several clusters of *E2F8*-deficient endothelial cells and macrophages may be tuned to regulate cancer processes in HCC.

Discussion

Using RNA-seq, miRNA and clinical data downloaded from the GDC database, we identified 807 and 1590 upregulated and downregulated DEGs and 59 and 71 upregulated and downregulated DEMs, respectively. A total of 2397 DEGs and 130 DEMs were found, and by constructing a ceRNA network, we detected hub genes with which miRNAs may interact. We observed 7 lncRNA-miRNA-mRNA interactions and performed survival analysis on all genes in the networks.

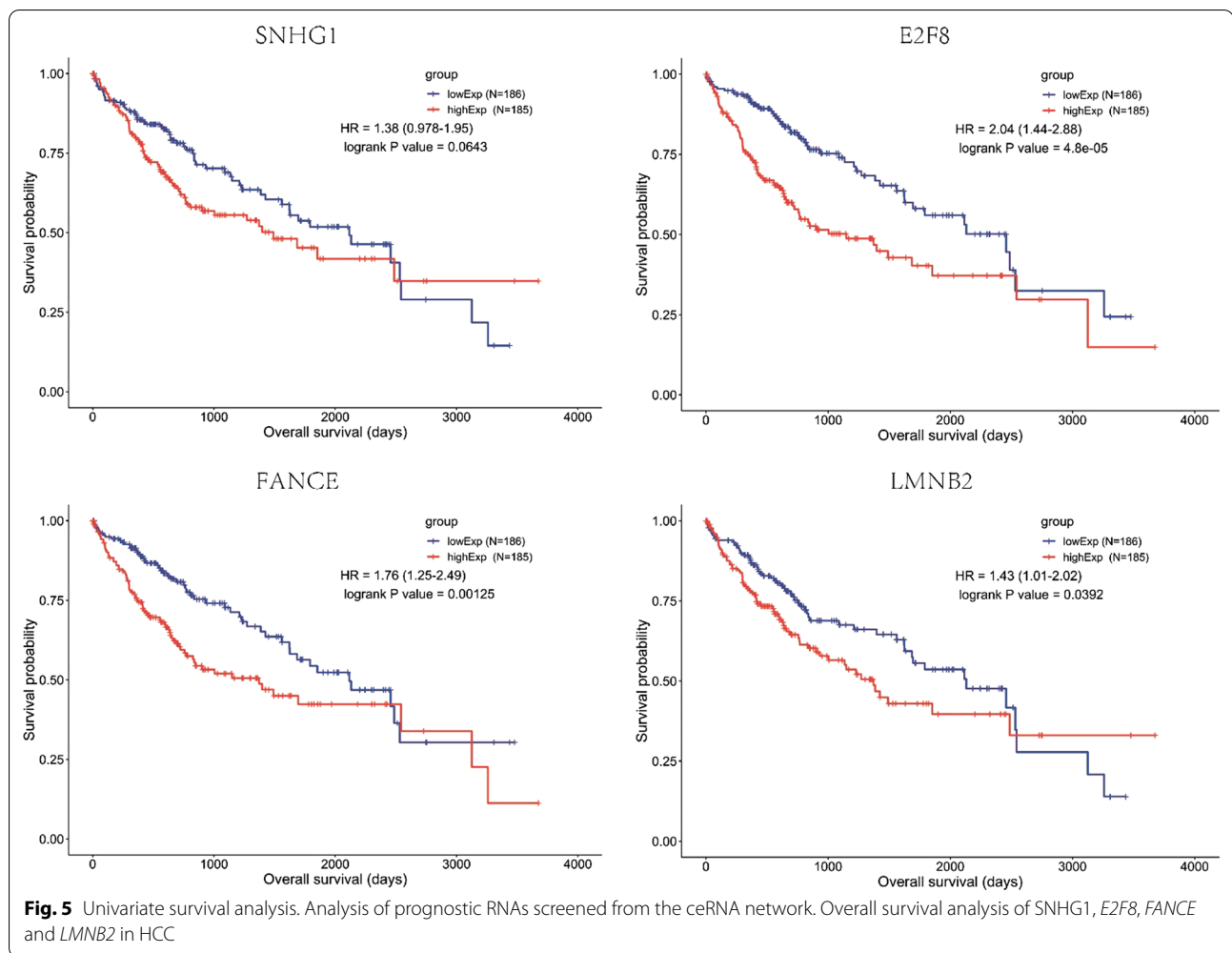
We analysed data for 377 patients from the GDC database and found that high *SNHG1* expression indicated poor survival; high *E2F8*, *FANCE* and *LMNB2* levels were also associated with poor patient survival. A previous study revealed a strong negative correlation between *SNHG1* expression and survival time in patients with high-risk neuroblastoma (NB) [28]. In addition, downregulation of *E2f8* resulted in a series of embryonic defects affecting vasculature and cell survival [29]. *FANCE* is reported to be a vital factor influencing the poor survival of *FANCE*-deficient cells among breast cancer cell lines [30], and upregulation of *LMNB2* is associated with significantly poorer prognosis in lung



adenocarcinoma patients [31]. Therefore, SNHG1 may affect survival through *E2F8*, *FANCE* and *LMNB2*. Further experiments are needed to examine interactions in this SNHG1-related network.

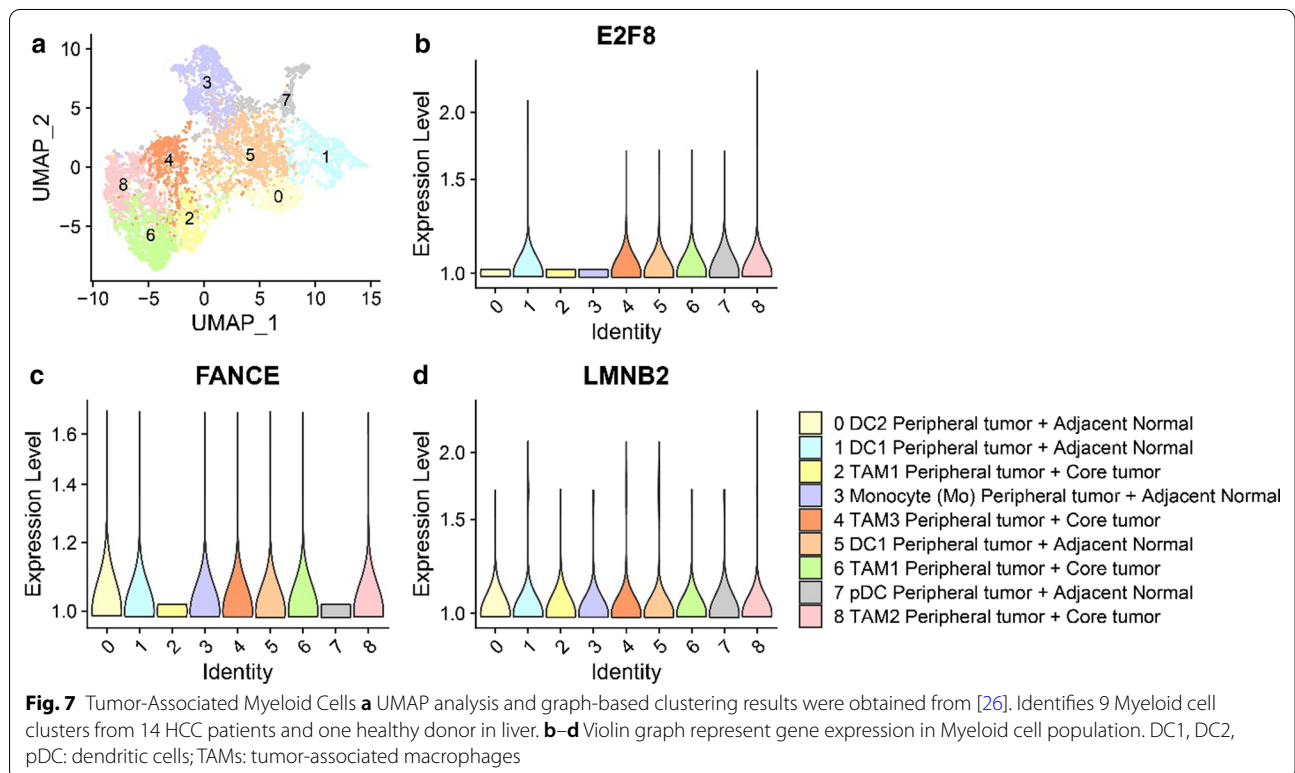
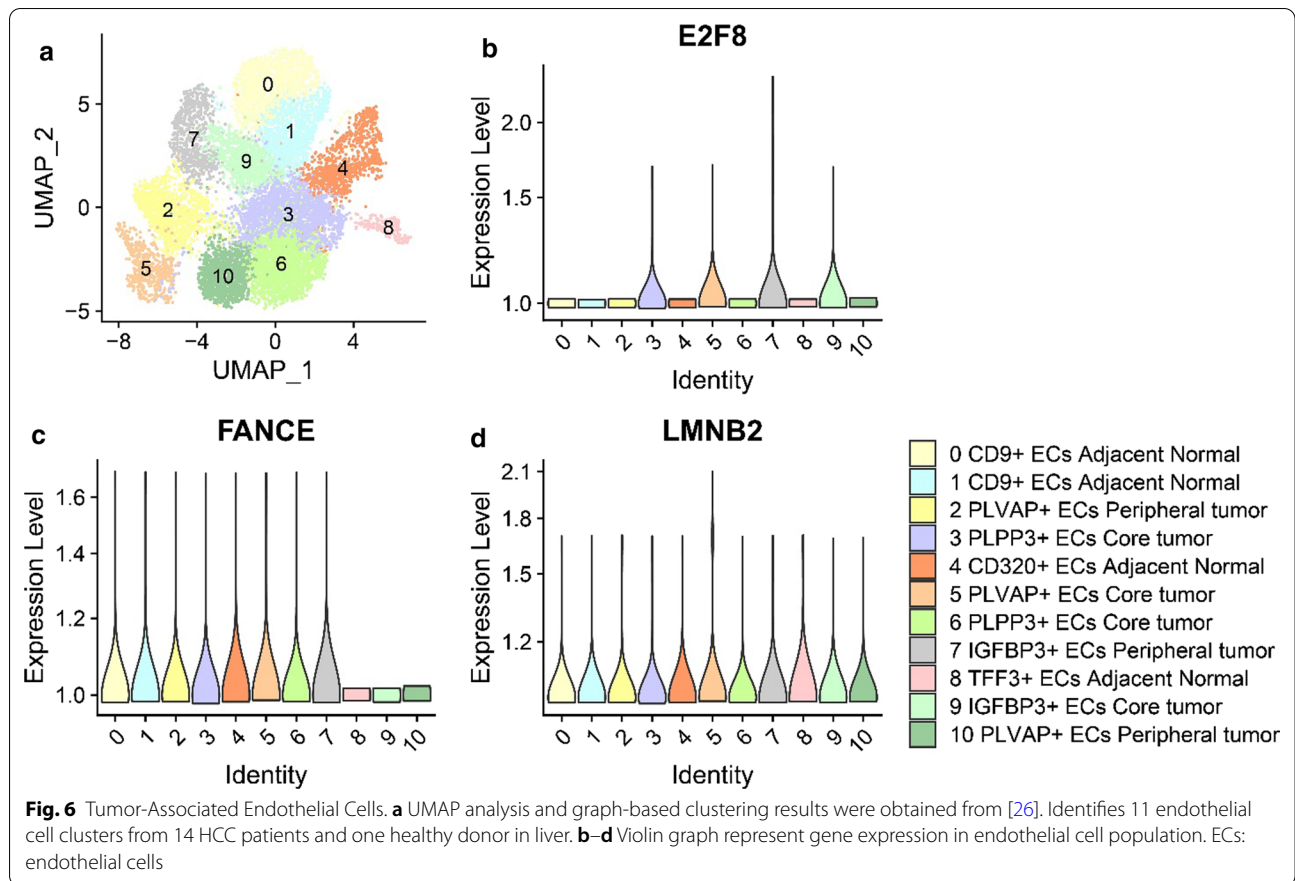
CeRNAs are RNA molecules that affect mRNA expression by competing for miRNAs in the same interaction network. Disruption of interactive networks can cause disease, including various cancers [32]. It is widely recognized that SNHG1 plays roles in cell proliferation and tumours. The patients showed poor prognosis when SNHG1 was upregulated, which indicated the

prognostic value of SNHG1 [33, 34]. They also identified several miRNA-SNHG1 subnetworks (such as hsa-miR-195-5p-SNHG1, hsa-miR-497-5p-SNHG1, and hsa-miR-101-3p-SNHG1) and several mRNA-SNHG1 pairs (such as *MYLK-AS1*-SNHG1, *CHEK1*-SNHG1, and *KIF2C*-SNHG1) in the ceRNA regulatory network of HCC [33, 34]. Through ceRNA network studies, SNHG1 was found to be important in HCC (SNHG1-hsa-miR-101-3p-*KIF2C*) [33], colorectal cancer (SNHG1-hsa-miR-484-*ORC6*, SNHG1-hsa-miR-423-5p-*EZH2*, and SNHG1-let-7b-5p-*ATP6V1F*) [35], lung adenocarcinoma



(*SNHG1*-*hsa-miR-21-5p-RALGPS2*, *SNHG1*-*hsa-miR-326-RALGPS2*, *SNHG1*-*hsa-miR-377-3p-RALGPS2*) [36], nasopharyngeal carcinoma (*SNHG1*-*miR-145-5p-NUAK1*) [37] and prostate cancer (*SNHG1*-*miR-199a-3p-CDK7*) [38]. *SNHG1* was also found to be highly expressed in posterior longitudinal ligament patients [39], cholangiocarcinoma [40], and osteosarcoma [41, 42]. In the ceRNA network established in this study, we found *SNHG1*-*hsa-miR-421-E2F8*, *SNHG1*-*hsa-miR-377-3p-E2F8*, *SNHG1*-*hsa-miR-330-5p-LMNB2*, and *SNHG1*-*hsa-miR-326-FANCE* axes in HCC. *E2F8* was found to be required for cell cycle progression and proliferation [43], *FANCE* to be essential for protection against chromosome breakage [44], and *LMNB2* to be required for early replication and transcription [45]. These genes mediate links between miRNAs, which affect the development of HCC as a network. However, studies with large samples are needed to confirm how these miRNAs are connected.

SNHG1 is a lncRNA that is generally thought to be produced by a small nuclear RNA host gene. Cell proliferation and changes in the tumour protein p53 occur when *SNHG1* is abnormally expressed [46]. Considering the very important functions of *SNHG1*, its 3 coexpressed genes can be expected to be related to disease, with altered expression in many human diseases. Of these genes, *E2F8* is usually accompanied by overexpression of genes during progression of the cell cycle and cell proliferation [47]. Expression of *E2F* target genes is required in the S phase of the cell cycle. As a part of this nuclear complex, *FANCE* interacts with *BRCA1* at sites of DNA damage by activating the downstream protein *FANCD2* [44]. *LMNB2*, replicated in the first minute of S phase in humans, is essential for early replication and transcription [48]. In addition, *LMNB2*, which is located on the inner side of the nuclear envelope, is required for nuclear stability, chromatin structure, and gene expression [48].



To explore the cell type specifically of these genes in tumour and normal samples, we analysed their expression in endothelial cells and macrophages and found low expression of *E2F8* in normal epithelial cells, which may indicate that its presence promotes cancer. This is also consistent with the conclusion drawn from TCGA data, whereby no difference between *LMNB2* and *FANCE* was detected in these two cell types from normal and tumour samples.

Conclusions

In this work, we identified 7 lncRNA-miRNA-mRNA interactions. Furthermore, we demonstrated that *ADAMTSL3*, *AUTS2*, *BHLHE40*, *DUSP6*, *ERRF1*, *GADD45A*, *GNPNAT1*, *HIVEP1*, *MYLK*, *MYO10*, *PNRC1*, *RIPOR2*, *SMAD6*, *SPRY2*, and *SYBU* show the same expression trends as *SNHG1* in HCC. By generating lncRNA-miRNA-mRNA interaction networks and performing univariate survival analysis of the genes coexpressed with lncRNAs in HCC, we demonstrated that *SNHG1* has the same effect on patient survival as *E2F8*, *FANCE* and *LMNB2*. Single-cell analysis also demonstrated that *E2F8* may play an important role in tumorigenesis or cancer development. Therefore, *SNHG1*, *E2F8*, *FANCE* and *LMNB2* may represent therapeutic targets for HCC with complex regulation. Taken together, our research provides genetic resources for capturing the majority of essential genes involved, which will yield a global view of the minimum set of genes and pathways. The biological interaction network of *SNHG1*-related lncRNA-miRNA-mRNA interactions provides insight into the function of *SNHG1* and its molecular regulation in HCC.

Abbreviations

HCC: Hepatocellular carcinoma; GEPIA: Gene Expression Profiling Interactive Analysis; LIHC: Liver hepatocellular carcinoma; *SNHG1*: Small nucleolar RNA host gene 1; miRNA: MicroRNA; lncRNA: Long noncoding RNA; ceRNAs: Competing endogenous RNAs; DEGs: Differentially expressed genes; DEMs: Differentially expressed miRNAs; GO: Gene Ontology; PPI: Protein-protein interaction; TCGA: The Cancer Genome Atlas; *E2F8*: E2F transcription factor 8; *FANCE*: FA complementation group E; *LMNB2*: Lamin B2.

Acknowledgements

We thank the editor and reviewers for taking the time and effort to provide insightful and critical discussion.

Authors' contributions

CZ and SY conceived the study. SY collected raw data and analysed the GDC data. CZ performed the Cytoscape analysis. SY performed the gene function and analysis. CZ and SY wrote the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by an unrestricted grant from the Guangdong Postdoctoral Research Foundation (CN) (O0390302). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and material

The RNA-seq, miRNA and clinical data of LIHC (Liver hepatocellular carcinoma) were downloaded from the DATA PORTAL of GDC (Genomic Data Commons) <https://portal.gdc.cancer.gov/projects/TCGA-LIHC>. Immunohistochemistry data were downloaded from the Human Protein Atlas database (<https://www.proteinatlas.org/>) under gene symbols described in Table 1. Single-cell sequencing data is available from NCBI Gene Expression Omnibus (GEO: GSE156337).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

We declare that we have no conflict of interest.

Author details

¹ Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China.

² Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, 190 Kaiyuan Avenue, Guangzhou Science Park, Luogang District, Guangzhou 510530, China. ³ Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Bioland Laboratory, Guangzhou 510005, China.

Received: 10 October 2020 Accepted: 17 January 2021

Published online: 26 January 2021

References

- Bray F, Ferlay J, Soerjomataram I, et al. 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.
- Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359–86.
- Chayanupatkul M, Omino R, Mittal S, et al. Hepatocellular carcinoma in the absence of cirrhosis in patients with chronic hepatitis B virus infection. *J Hepatol*. 2017;66(2):355–62.
- Turati F, Edefonti V, Talamini R, et al. Family history of liver cancer and hepatocellular carcinoma. *Hepatology*. 2012;55(5):1416–25.
- Fu J, Wang H. Precision diagnosis and treatment of liver cancer in China. *Cancer Lett*. 2018;412:283–8.
- Anastasiadou E, Jacob LS, Slack FJ. Non-coding RNA networks in cancer. *Nat Rev Cancer*. 2018;18(1):5–18.
- Bonasio R, Shiekhattar R. Regulation of transcription by long noncoding RNAs. *Annu Rev Genet*. 2014;48:433–55.
- St Laurent G, Wahlestedt C, Kapranov P. The landscape of long noncoding RNA classification. *Trends Genet*. 2015;31(5):239–51.
- Luo J, Wang K, Yeh S, et al. lncRNA-p21 alters the antiandrogen enzalutamide-induced prostate cancer neuroendocrine differentiation via modulating the EZH2/STAT3 signaling. *Nat Commun*. 2019;10(1):2571.
- Qu L, Wu Z, Li Y, et al. A feed-forward loop between lncARSR and YAP activity promotes expansion of renal tumour-initiating cells. *Nat Commun*. 2016;7:12692.
- Chowdhury TA, Koceja C, Eisa-Beygi S, et al. Temporal and spatial post-transcriptional regulation of zebrafish tie1 mRNA by long noncoding RNA during brain vascular assembly. *Arterioscler Thromb Vasc Biol*. 2018;38(7):1562–75.
- Augoff K, McCue B, Plow EF, et al. miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer. *Mol Cancer*. 2012;11:5.
- Tan X, Banerjee P, Liu X, et al. The epithelial-to-mesenchymal transition activator ZEB1 initiates a prometastatic competing endogenous RNA network. *J Clin Invest*. 2018;128(7):3198.
- Braconi C, Kogure T, Valeri N, et al. microRNA-29 can regulate expression of the long non-coding RNA gene MEG3 in hepatocellular cancer. *Oncogene*. 2011;30(47):4750–6.

15. Li R, Qu H, Wang S, et al. GDCRNATools: an R/Bioconductor package for integrative analysis of lncRNA, miRNA and mRNA data in GDC. *Bioinformatics*. 2018;34(14):2515–7.
16. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–40.
17. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
18. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45(W1):W98–102.
19. Uhlen M, Bjorling E, Agaton C, et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics*. 2005;4(12):1920–32.
20. Thul PJ, Lindskog C. The human protein atlas: a spatial map of the human proteome. *Protein Sci*. 2018;27(1):233–44.
21. Zhang X, Feng H, Li D, et al. Identification of differentially expressed genes induced by aberrant methylation in oral squamous cell carcinomas using integrated bioinformatic analysis. *Int J Mol Sci*. 2018;19(6):1698.
22. Li JH, Liu S, Zhou H, et al. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2014;42(Database issue):92–7.
23. Otasek D, Morris JH, Boucas J, et al. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol*. 2019;20(1):185.
24. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.
25. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
26. Sharma A, Seow JJW, Dutertre CA, et al. Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell*. 2020;183:377–94.
27. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell*. 2019;177(7):1888–1902 e21.
28. Sahu D, Hsu CL, Lin CC, et al. Co-expression analysis identifies long noncoding RNA SNHG1 as a novel predictor for event-free survival in neuroblastoma. *Oncotarget*. 2016;7(36):58022–37.
29. Li J, Ran C, Li E, et al. Synergistic function of E2F7 and E2F8 is essential for cell survival and embryonic development. *Dev Cell*. 2008;14(1):62–75.
30. Bouffard F, Plourde K, Belanger S, et al. Analysis of a FANCE splice isoform in regard to DNA repair. *J Mol Biol*. 2015;427(19):3056–73.
31. Misono S, Seki N, Mizuno K, et al. Dual strands of the miR-145 duplex (miR-145-5p and miR-145-3p) regulate oncogenes in lung adenocarcinoma pathogenesis. *J Hum Genet*. 2018;63(10):1015–28.
32. Liu XH, Sun M, Nie FQ, et al. Lnc RNA HOTAIR functions as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer. *Mol Cancer*. 2014;13:92.
33. Zhang J, Lou W. A key mRNA-miRNA-lncRNA competing endogenous RNA triple sub-network linked to diagnosis and prognosis of hepatocellular carcinoma. *Front Oncol*. 2020;10:340.
34. Lin P, Wen DY, Li Q, et al. Genome-wide analysis of prognostic lncRNAs, miRNAs, and mRNAs forming a competing endogenous RNA network in hepatocellular carcinoma. *Cell Physiol Biochem*. 2018;48(5):1953–67.
35. Qi X, Lin Y, Liu X, et al. Biomarker discovery for the carcinogenic heterogeneity between colon and rectal cancers based on lncRNA-associated ceRNA network analysis. *Front Oncol*. 2020;10:535985.
36. Tan J, Wang W, Song B, et al. Integrative analysis of three novel competing endogenous RNA biomarkers with a prognostic value in lung adenocarcinoma. *Biomed Res Int*. 2020;2020:2837906.
37. Lan X, Liu X. LncRNA SNHG1 functions as a ceRNA to antagonize the effect of miR-145a-5p on the down-regulation of NUAK1 in nasopharyngeal carcinoma cell. *J Cell Mol Med*. 2019;23(4):2351–61.
38. Li J, Zhang Z, Xiong L, et al. SNHG1 lncRNA negatively regulates miR-199a-3p to enhance CDK7 expression and promote cell proliferation in prostate cancer. *Biochem Biophys Res Commun*. 2017;487(1):146–52.
39. Wang Y, Niu H, Liu Y, et al. Promoting effect of long non-coding RNA SNHG1 on osteogenic differentiation of fibroblastic cells from the posterior longitudinal ligament by the microRNA-320b/IFNGR1 network. *Cell Cycle*. 2020;19(21):2836–50.
40. Li Z, Li X, Du X, et al. The interaction between lncRNA SNHG1 and miR-140 in regulating growth and tumorigenesis via the TLR4/NF-kappaB pathway in cholangiocarcinoma. *Oncol Res*. 2019;27(6):663–72.
41. Wang J, Cao L, Wu J, et al. 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.
42. Jiang Z, Jiang C, Fang J. Up-regulated lnc-SNHG1 contributes to osteosarcoma progression through sequestration of miR-577 and activation of WNT2B/Wnt/beta-catenin pathway. *Biochem Biophys Res Commun*. 2018;495(1):238–45.
43. Maiti B, Li J, de Bruin A, et al. Cloning and characterization of mouse E2F8, a novel mammalian E2F family member capable of blocking cellular proliferation. *J Biol Chem*. 2005;280(18):18211–20.
44. Pace P, Johnson M, Tan WM, et al. FANCE: the link between Fanconi anaemia complex assembly and activity. *EMBO J*. 2002;21(13):3414–23.
45. Giacca M, Zentilin L, Norio P, et al. Fine mapping of a replication origin of human DNA. *Proc Natl Acad Sci USA*. 1994;91(15):7119–23.
46. Shen Y, Liu S, Fan J, et al. Nuclear retention of the lncRNA SNHG1 by doxorubicin attenuates hnRNP-p53 protein interactions. *EMBO Rep*. 2017;18(4):536–48.
47. Christensen J, Cloos P, Toftgaard U, et al. Characterization of E2F8, a novel E2F-like cell-cycle regulated repressor of E2F-activated transcription. *Nucleic Acids Res*. 2005;33(17):5458–70.
48. Biamonti G, Giacca M, Perini G, et al. The gene for a novel human lamin maps at a highly transcribed locus of chromosome 19 which replicates at the onset of S-phase. *Mol Cell Biol*. 1992;12(8):3499–506.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

