


PRIMARY RESEARCH

Open Access



# Construction and comprehensive analysis of a ceRNA network to reveal potential prognostic biomarkers for hepatocellular carcinoma

Junyu Long<sup>1†</sup>, Yi Bai<sup>1†</sup>, Xiaobo Yang<sup>1†</sup>, Jianzhen Lin<sup>1</sup>, Xu Yang<sup>1</sup>, Dongxu Wang<sup>1</sup>, Li He<sup>2</sup>, Yongchang Zheng<sup>1\*</sup>  and Haitao Zhao<sup>1\*</sup>

## Abstract

**Background:** Long noncoding RNAs (lncRNAs) can act as microRNA (miRNA) sponges to regulate protein-coding gene expression; therefore, lncRNAs are considered a major part of the competitive endogenous RNA (ceRNA) network and have attracted growing attention. The present study explored the regulatory mechanisms and functional roles of lncRNAs as ceRNAs in hepatocellular carcinoma (HCC) and their potential impact on HCC patient prognosis.

**Methods:** In this study, we systematically studied the expression profiles and prognostic value of lncRNA, miRNA, and mRNA from a total of 838 HCC patients from five HCC cohorts (TCGA, GSE54236, GSE76427, GSE64041 and GSE14520). The TCGA, GSE54236 and GSE76427 HCC cohorts were utilized to establish a prognosis-related network of dysregulated ceRNAs by bioinformatics methods. The GSE64041 and GSE14520 HCC cohorts were utilized to verify the expression of candidate genes.

**Results:** In total, 721 lncRNAs, 73 miRNAs, and 1563 mRNAs were aberrantly expressed in HCC samples. A ceRNA network including 26 lncRNAs, four miRNAs, and six mRNAs specific to HCC was established. The survival analysis showed that four lncRNAs (MYCNOS, DLX6-AS1, LINC00221, and CRNDE) and two mRNAs (CCNB1 and SHCBP1) were prognostic biomarkers for patients with HCC in both the TCGA and GEO databases.

**Conclusion:** The proposed ceRNA network may help elucidate the regulatory mechanism by which lncRNAs function as ceRNAs and contribute to the pathogenesis of HCC. Importantly, the candidate lncRNAs, miRNAs, and mRNAs involved in the ceRNA network can be further evaluated as potential therapeutic targets and prognostic biomarkers for HCC.

**Keywords:** Competing endogenous RNA, Hepatocellular carcinoma, Long noncoding RNA, MicroRNA, Prognosis

## Background

Hepatocellular carcinoma (HCC) is a severe cancer with an increasing incidence and is the fifth most prevalent cancer worldwide [1–3]. Viral infections, such as hepatitis B and hepatitis C, are usually associated with

liver cirrhosis and HCC tumorigenesis [2, 4]. HCC remains an important global clinical challenge due to its high incidence, limited treatment strategies, and poor prognosis [5]. Therefore, strategies based on a personal need to treat HCC, such as discovering potential biomarkers and therapeutic targets, are urgently needed. The present study explored how HCC-related long noncoding RNAs (lncRNAs) serve as competitive endogenous RNAs (ceRNAs) to regulate target genes and how they affect the pathogenesis and prognosis of HCC.

\*Correspondence: zhengyongchang@pumch.cn; zhaoh@pumch.cn

†Junyu Long, Yi Bai and Xiaobo Yang contributed equally to this work

<sup>1</sup> Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

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



Although noncoding RNAs (ncRNAs) lack protein-encoding abilities, they are ubiquitous in organisms [6]. As a subtype of ncRNA greater than 200 nucleotides in length, lncRNAs were once considered transcriptional noise. Several studies have indicated that lncRNAs have many pivotal functions in tumor-related processes, including proliferation, invasion, and metastasis [7–9]. However, verification of the regulatory roles of lncRNAs in gene expression remains difficult. Currently, many researchers are working to reveal the different biological functions of lncRNAs in malignant tumors.

One ceRNA hypothesis was proposed by Salmena et al. [10]. A complicated posttranscriptional regulatory network was described that allowed lncRNAs, mRNAs, and other RNAs to compete with microRNAs (miRNAs) via acting as natural miRNA sponges by virtue of sharing no less than one miRNA response element (MRE). These ncRNAs act as ceRNAs to modulate mRNA expression and regulate protein levels, which contributes to the occurrence and development of tumors [11, 12]. Studies have shown that each miRNA can control the transcriptional expression levels of hundreds of proteins, and each mRNA contains different MREs and thus may be targeted by multiple miRNAs [12].

Recently, a growing number of studies have verified that the lncRNA–miRNA–mRNA regulatory network plays a critical role in the progression and pathogenesis of several tumors, including liver cancer, gallbladder cancer, and other malignant tumors [13–15]. lncRNAs with sequences similar to their target miRNAs can separate miRNAs from mRNAs. Wang et al. demonstrated that the lncRNA HULC influenced PRKACB gene expression by competitively combining with the miRNA miR-372 and thus participated in liver cancer pathogenesis [13]. Wang et al. confirmed that lncRNA H19 acts as a molecular sponge to absorb miR-342-3p in gallbladder cancer and upregulate FOXM1 gene expression [15].

Therefore, lncRNAs acting as ceRNAs have diverse biological functions that deserve further exploration. Here, we investigated differences in RNA expression patterns between 43 HCC tumor tissues and 43 paired nontumorous tissues and constructed a ceRNA network associated with HCC, including 26 lncRNAs, four miRNAs, and six mRNAs. The survival analysis showed that four lncRNAs (MYCNOS, DLX6-AS1, LINC00221, and CRNDE) and two mRNAs (CCNB1 and SHCBP1) were prognostic biomarkers for patients with HCC in both the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases. These candidate genes involved in the ceRNA network may become potential therapeutic targets or diagnostic biomarkers for HCC.

## Materials and methods

### Study population

A total of 371 patients with HCC participated in this study. Available miRNA-seq data from 371 HCC samples with survival data and 50 adjacent nontumorous samples (including 50 paired HCC samples) and mRNA-seq data for 367 HCC samples with survival data and 50 adjacent nontumorous samples (including 50 paired HCC samples) were obtained from the TCGA database (<https://tcga-data.nci.nih.gov/>) on May 1, 2018. Approval from the ethics committee was not necessary. This study fully meets the publication requirements of the TCGA. The RNA-seq data were obtained using the Illumina HiSeq\_miRNA-Seq and Illumina HiSeq\_RNA-Seq platforms. The genes identified via RNA expression profiling were annotated based on the Ensembl Gene ID. Log<sub>2</sub> transformation was performed on all gene expression data. We normalized the downloaded data by using the trimmed mean of M value (TMM) normalization method of the edgeR R package (Version: 3.24.3) in R software (Version: 3.5.2) [16]. When an RNA had duplicate data, the average RNA expression was used. The lncRNAs, miRNAs and mRNAs with an average expression value > 1 were retained, and low-abundance RNAs were eliminated.

### Microarray data

The gene expression profiles from GSE54236 (including 77 adjacent nontumorous samples and 78 HCC samples; platform: GPL6480), GSE76427 (including 52 adjacent nontumorous samples and 115 HCC samples; platform: GPL10558), GSE64041 (including 60 HCC samples and 60 paired adjacent nontumorous samples; platform: GPL6244), and GSE14520 (including 214 HCC samples and 214 paired adjacent nontumorous samples; platform: GPL571) were obtained from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). Log<sub>2</sub> transformation was performed on only the gene expression data for GSE76427. The average RNA expression value was used when duplicated data were found. The genes with an average expression value > 1 were retained, and low-abundance microarray data were removed. Two datasets (GSE54236 and GSE76427) with clinical data were integrated into the meta-GEO HCC cohort to identify the candidate genes associated with the overall survival of patients with HCC. We remove batch effects by using the sva package, and normalized the data by using the scale method in the limma R package (Version: 3.38.3) [17]. The GSE64041 and GSE14520 HCC cohorts with paired HCC samples were utilized to verify the expression patterns of the candidate genes. The paired t-test was used for normally distributed data; otherwise, the Wilcoxon rank test for paired data was used to evaluate whether a

gene was differentially expressed between normal tissue and tumor tissue.

#### Identification of differentially expressed genes (DEGs)

Ensembl genome browser 92 (<http://asia.ensembl.org/index.html>) was used to annotate gene symbols. Then, we compared 50 HCC samples with 50 paired nontumorous samples to identify differentially expressed lncRNAs (DELncRNAs), mRNAs (DEmRNAs), and miRNAs (DEmiRNAs) utilizing the edgeR R package with the threshold set at an adjusted P-value < 0.01 and  $|\log_2\text{-fold change (FC)}| > 2$  [16]. A hierarchical cluster heatmap based on Euclidean distance was generated using the pheatmap R package (Version: 1.0.12) and represents the expression intensity and direction of DEGs.

#### Constructing the ceRNA network

To ensure the functions of the lncRNAs, miRNAs, and mRNAs in the ceRNA network and to further improve the ceRNA network reliability, a coexpression network of DELncRNAs, DEmiRNAs, and DEmRNAs was constructed; the ggalluvial R package (Version: 0.9.1) was used to visualize the ceRNA network [18]. Using the miRcode database (Version 11; <http://www.mircode.org/>), we confirmed the interactions between DELncRNAs and DEmiRNAs. miRcode, which is an integrated, searchable map of putative target sites of miRNAs, includes the complete transcriptome annotated by the ENCYClopedia Of DNA Elements (ENCODE) [19]. The existing version contains 10,419 lncRNA genes. DEmRNAs targeted by the DEmiRNAs were retrieved from the miRDB (Version 5.0; <http://mirdb.org>), miRTarBase (Version 7.0; <http://mirtarbase.mbc.nctu.edu.tw/>), and TargetScan (Version 7.2; [http://www.targetscan.org/vert\\_72/](http://www.targetscan.org/vert_72/)) databases [20–22]. miRDB refers to an online database used to predict miRNA targets and make functional annotations [20]. MirTarget is a bioinformatics tool that was developed based on the analysis of thousands of interactions among miRNA targets from next-generation sequencing experiments [20] and was used to predict all targets in the miRDB database [20]. The DEmRNAs targeted by DEmiRNAs were verified by experimental studies using reporter assays, qRT-PCR, microarray analysis, Western blotting, and high-throughput sequencing experiments reported in miRTarBase [21]. TargetScan was used to search for conserved 6-mer, 7-mer, and 8-mer sites that corresponded to the seed regions of every miRNA to complete the prediction of biological targets of miRNAs [22].

#### Survival analysis

After combining the overall survival of patients with HCC in the TCGA, the survival R package (Version:

2.43-3) was used to perform a survival analysis of the samples with DELncRNAs, DEmiRNAs, and DEmRNAs to identify the prognostic genes. The patients were classified into two groups (high vs. low) using optimal cut-off values determined by the survminer R package (Version: 0.4.3). Log-rank  $P < 0.05$  was considered significant.

#### Gene set enrichment analysis (GSEA)

Based on the expression level of the candidate gene and the median expression value, 43 HCC samples from the RNA-seq data of the TCGA HCC cohort were classified into the low-expression and high-expression groups. To identify the underlying function of candidate genes, GSEA (Version: 3.0; <http://software.broadinstitute.org/gsea/index.jsp>) was conducted between the two groups [23]. The annotated gene sets file *c5.bp.v6.2.symbols.gm* was chosen for the reference gene sets. Gene size  $\geq 100$ ,  $|\text{enrichment score (ES)}| > 0.6$  and false discovery rate (FDR) < 0.01 were set as the cut-off criteria.

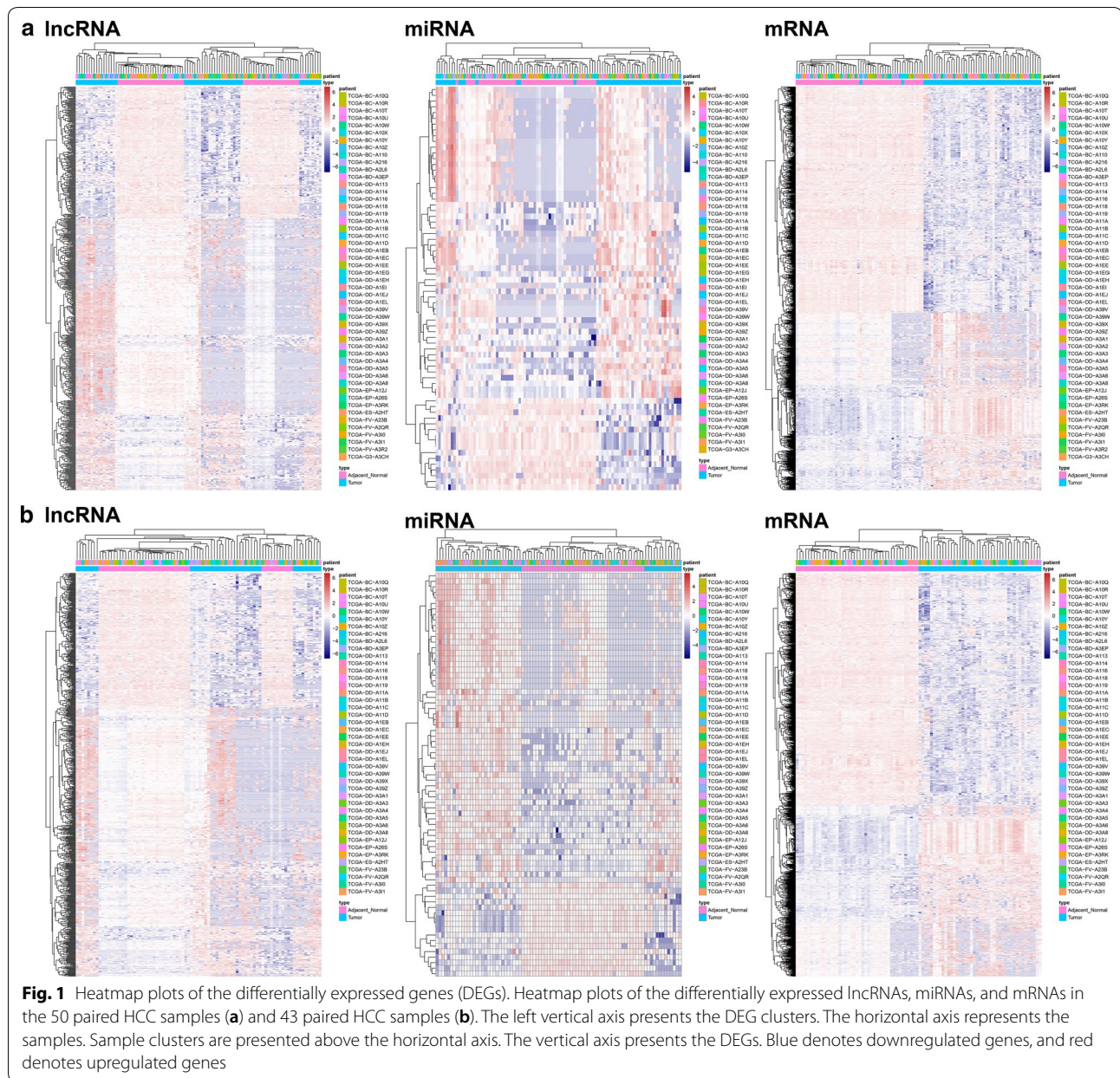
## Results

### DEGs in HCC

Using a cut-off threshold of  $|\log_2\text{ FC}| > 2$  and an adjusted P-value < 0.01 for the 50 HCC tissues compared with the paired 50 nontumorous samples, we identified 641 DELncRNAs, 70 DEmiRNAs, and 1392 DEmRNAs. Then, we used those DEGs for cluster analysis and removed seven abnormal samples (Fig. 1a). Next, we used the remaining 43 HCC tissues and the 43 paired nontumorous samples to perform differential expression analyses. As a result, 721 DELncRNAs, 73 DEmiRNAs, and 1563 DEmRNAs were identified (Additional file 1: Table S1). The heatmap of the lncRNAs, miRNAs and mRNAs showed that the tumors clustered separately from the paired nontumorous tissues (Fig. 1b).

### Construction of a ceRNA network for HCC

To better understand the effect of lncRNAs on mRNAs mediated by combination with miRNAs in HCC, we built a ceRNA network based on the abovementioned data and used the ggalluvial R package (Version: 0.9.1) to visualize the network. Additional file 2: Table S2 shows that 44 DELncRNAs interact with nine DEmiRNAs retrieved from the miRcode database. Eight of these nine DEmiRNAs were identified in the starBase database. Then, we searched for DEmRNAs based on eight DEmiRNAs in the miRDB, miRTarBase, and TargetScan databases. Six DEmRNAs that can interact with four of the eight DEmiRNAs according to all three of the databases were chosen (Additional file 3: Table S3). After removing the remaining four DEmiRNAs and the corresponding lncRNAs, 26 DELncRNAs, four DEmiRNAs, and six DEmRNAs were used to establish a ceRNA network (Fig. 2)

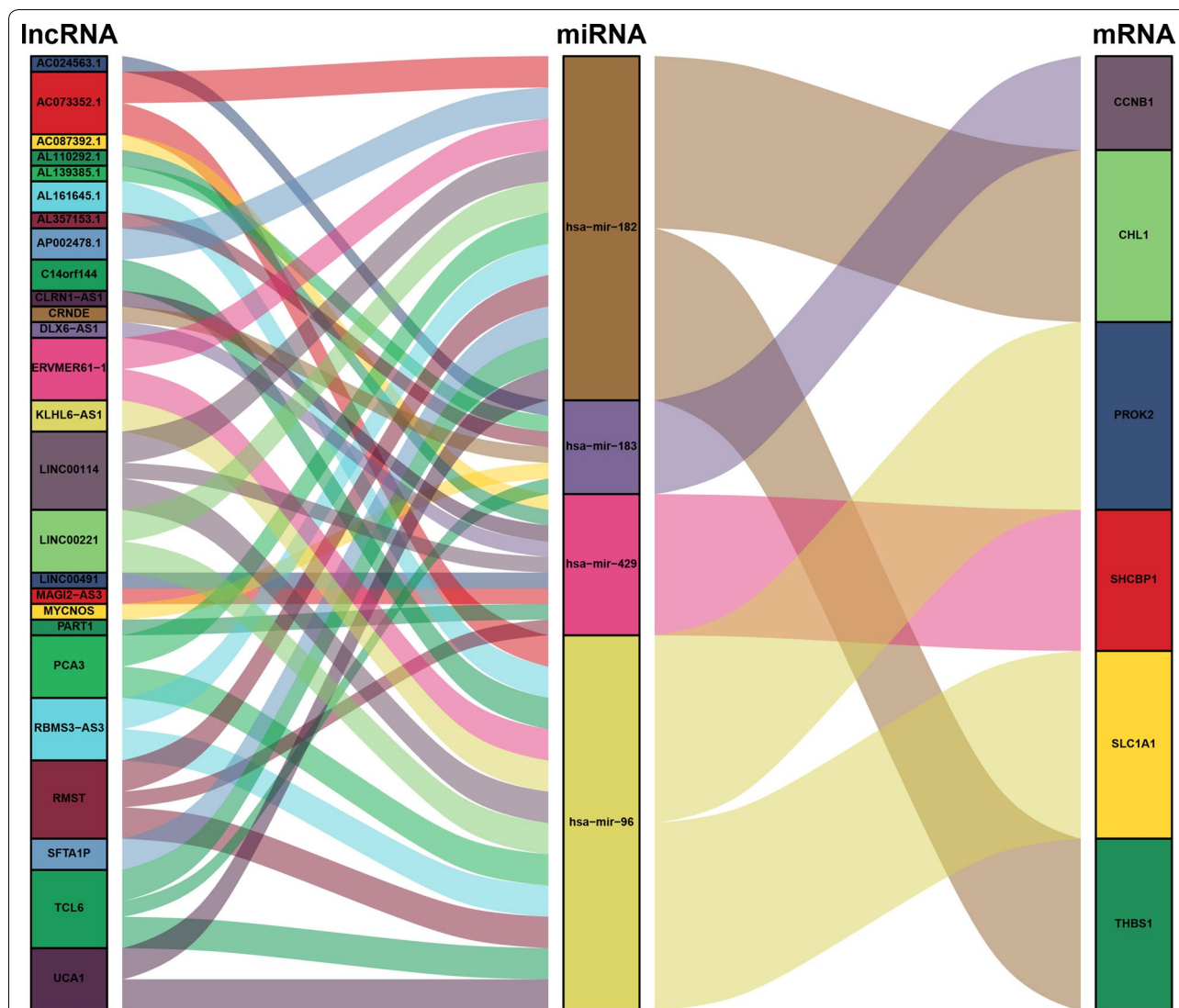


(Additional file 4: Table S4). Furthermore, the connection degree of each gene by topology was calculated to illustrate its importance in the ceRNA network (Fig. 2) (Additional file 5: Table S5).

**Prognosis assessment**

To determine which DELncRNAs, DEMiRNAs and DEMRNAs have an impact on the overall survival of patients with HCC, survival analyses were conducted to investigate the Kaplan–Meier curves for HCC patients with the 26 DELncRNAs, four DEMiRNAs, and six DEMRNAs in both the TCGA and meta-GEO HCC cohorts. In the

TCGA HCC cohort, 18 DELncRNAs and 73 DEMRNAs exhibited an obvious relationship to prognosis based on their respective optimal cutoffs ( $P < 0.05$ ) (Additional file 6: Table S6, Additional file 7: Table S7). Meanwhile, in the meta-GEO HCC cohort, seven DELncRNAs and five DEMRNAs were significantly relevant to prognosis based on their respective optimal cutoffs ( $P < 0.05$ ) (Additional file 8: Table S8, Additional file 9: Table S9). We consider that the genes with  $|\log_2 FC| > 2$  in the TCGA HCC cohort and  $HR > 1$  in both the TCGA and meta-GEO HCC cohorts were protective genes. In contrast, the genes with  $|\log_2 FC| < 2$  in the TCGA HCC cohort



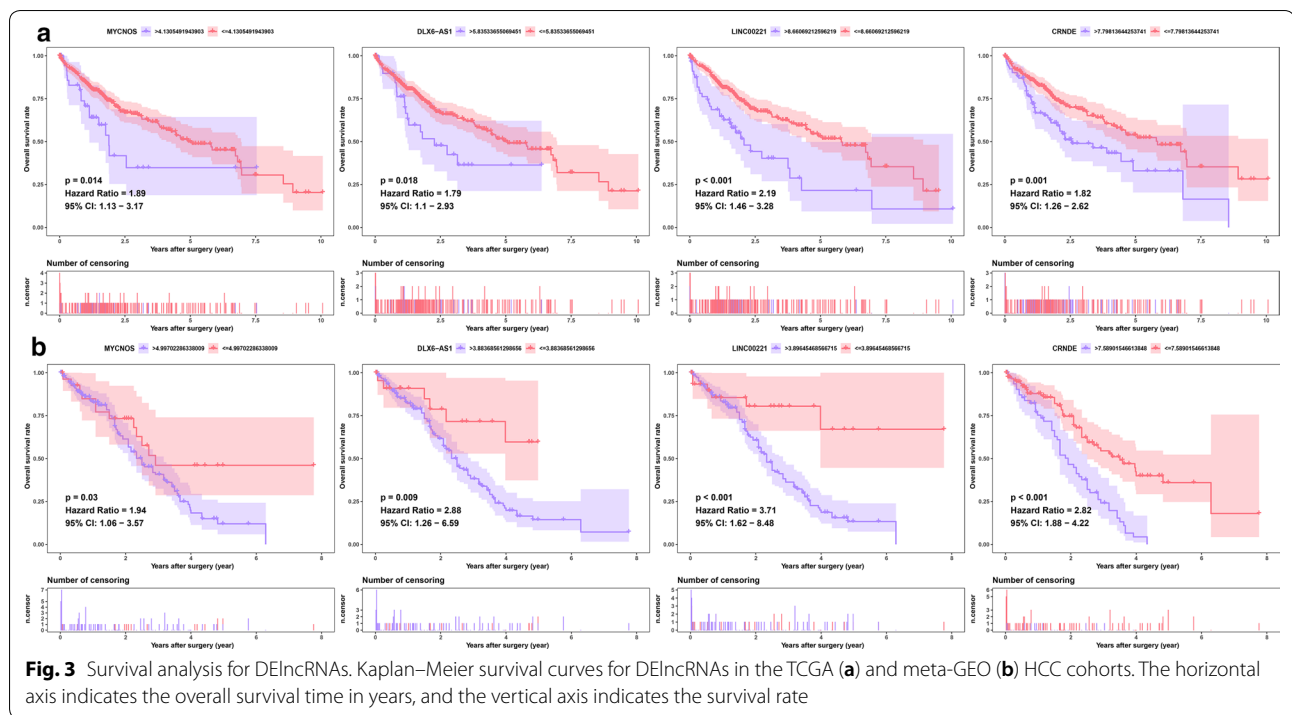
**Fig. 2** Sankey diagram for the ceRNA network in HCC. Each rectangle represents a gene, and the connection degree of each gene is visualized based on the size of the rectangle

and with  $HR < 1$  in both the TCGA and meta-GEO HCC cohorts were conferred a risk of poor prognosis (“risky genes”). As a result, four DElncRNAs (MYCNOS, DLX6-AS1, LINC00221, and CRNDE) and two DEmRNAs (CCNB1 and SHCBP1) were identified as risky genes in both the TCGA and GEO meta-GEO HCC cohorts (Figs. 3, 4). However, due to the lack of miRNAs in the GEO database, we could not validate the miRNAs in another HCC cohort. Based on the above screening criteria, two DEmiRNAs (hsa-miR-182 and hsa-miR-183) were identified as risky genes in the TCGA HCC cohort ( $P < 0.05$ ) (Fig. 5a) (Additional file 10: Table S10). In addition, we investigated whether the ratio of hsa-miR-182 with targeted RNAs (THBS1 and CHL1) and the ratio

of hsa-miR-183 with targeted RNAs (CCNB1) have an impact on the prognosis of patients with HCC. We found that the ratio of hsa-miR-182/CHL1 and the ratio of hsa-miR-183/CCNB1 were significantly associated with the overall survival of patients with HCC ( $P < 0.05$ ) (Fig. 5b).

**Validation of gene expression of DEmRNAs**

The two DEmRNAs (CCNB1 and SHCBP1) related to prognosis were chosen for validation. The expression of CCNB1 and SHCBP1 were significantly higher in the HCC tissues than in the paired nontumorous liver tissues in the GSE64041 and GSE14520 HCC cohorts, which is consistent with our results in the TCGA HCC



cohort (Fig. 4c–e). The results proved the reliability of our analysis.

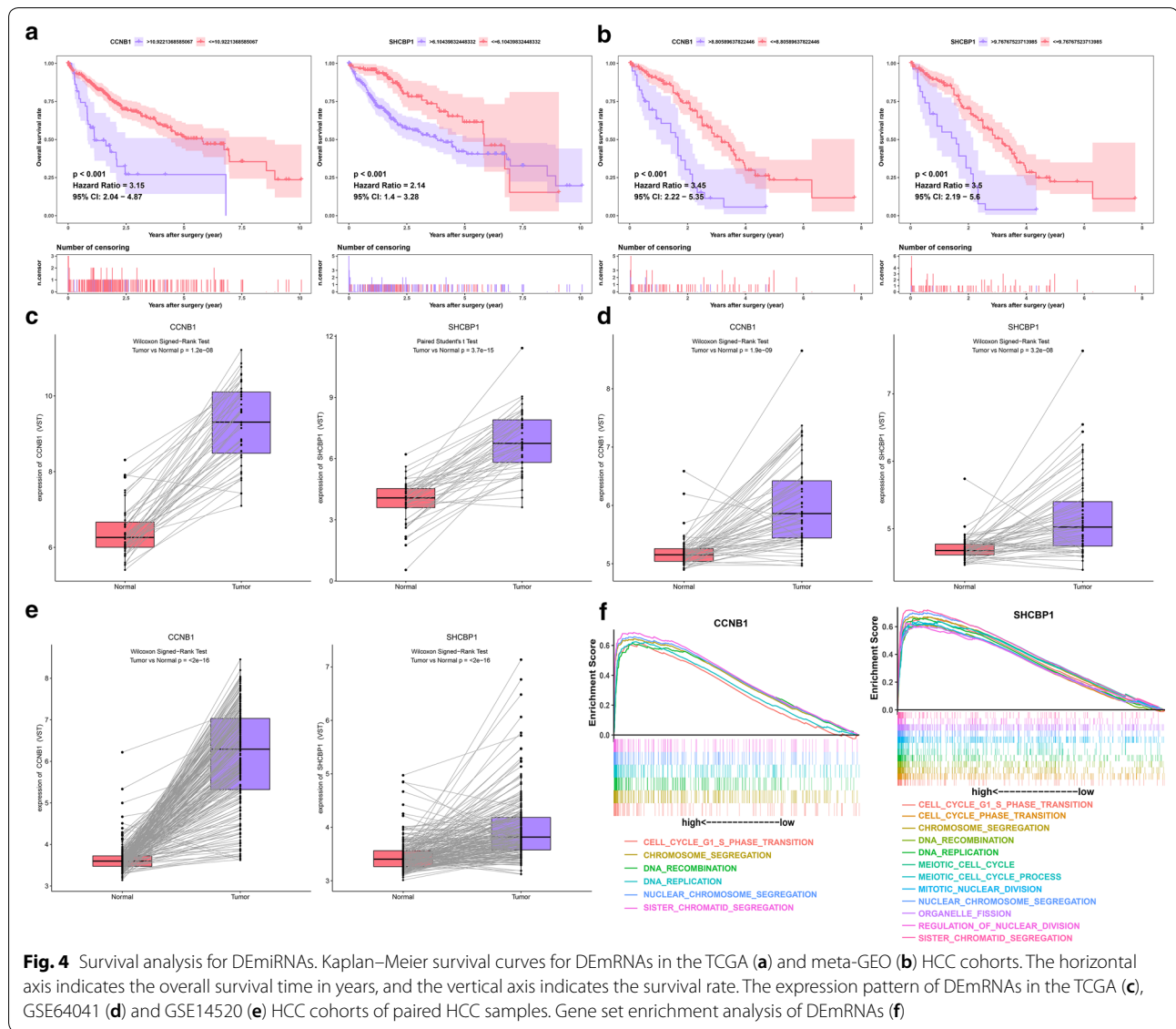
### Gene set enrichment analysis

To identify the biological pathways associated with CCNB1 and SHCBP1, which were highly expressed in HCC, we conducted GSEA of the HCC samples based on the TCGA HCC cohort. HCC samples in the CCNB1 high-expression group were most significantly enriched for “sister chromatid segregation” (Fig. 4e) (Additional file 11: Table S11); HCC samples in the SHCBP1 high-expression group were most significantly enriched for “regulation of nuclear division” (Fig. 4f) (Additional file 12: Table S12).

### Discussion

Approximately 782,500 new cases of HCC occur each year, making HCC the second greatest cause of cancer mortality worldwide [24]. Although many treatment strategies (i.e., radiofrequency ablation, surgical resection, and liver transplantation) have been adopted, the prognoses of HCC patients who undergo these treatments are still not satisfactory [25]. Most patients are diagnosed at advanced stages due to the asymptomatic nature of the disease. Elucidating the molecular mechanisms and processes underlying HCC is of great importance for identifying new therapeutic targets and improving the clinical outcomes of patients with this disease. More and more

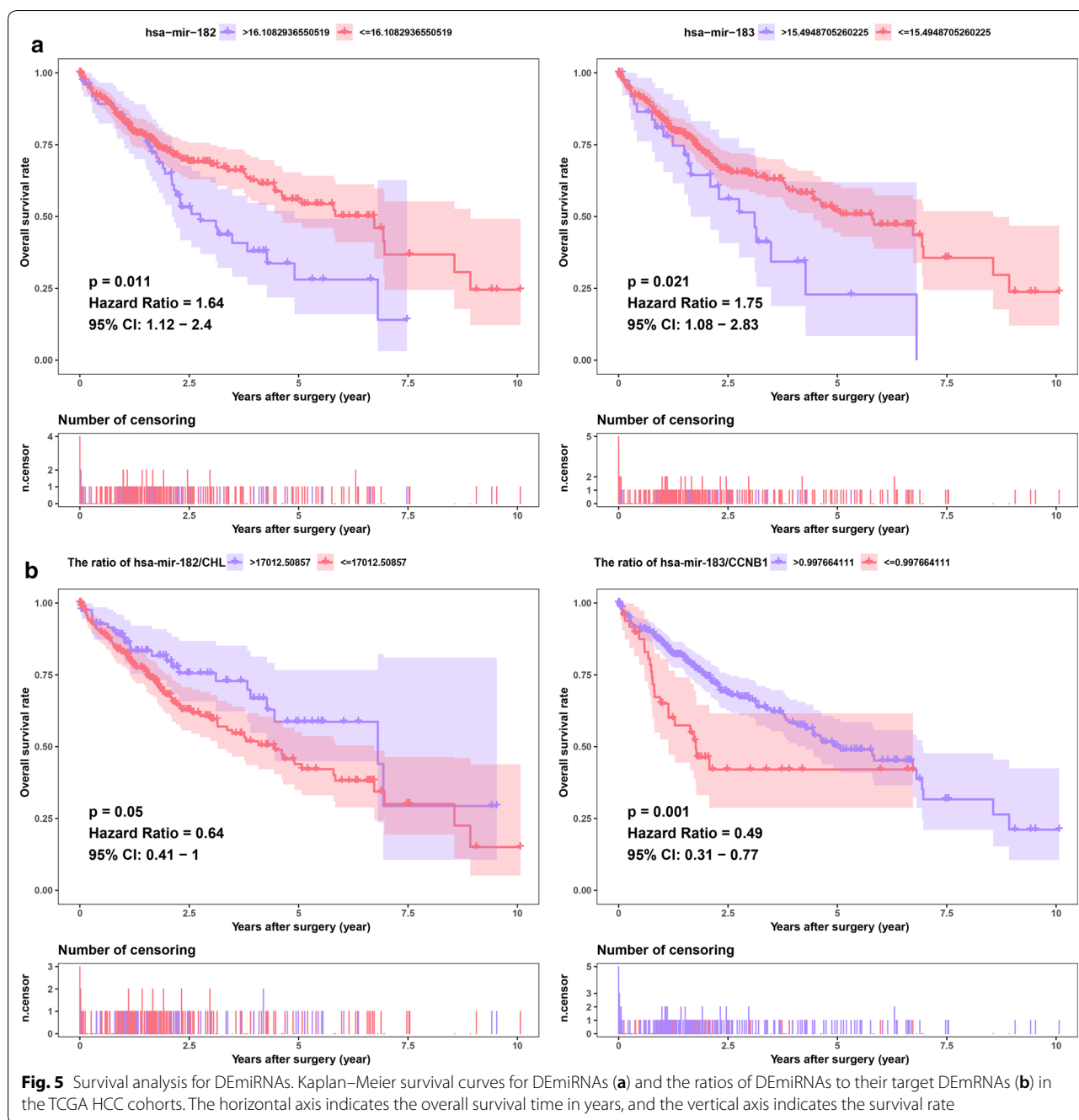
studies indicate that lncRNAs play a vital role in biological functions through multiple levels of regulation, which involve transcriptional, posttranscriptional, and epigenetic regulation [26, 27]. A great deal of research has shown that there is a complex and closely related regulatory network between miRNAs and lncRNAs. In addition, lncRNAs and miRNAs have crucial functions in the pathogenesis and progression of cancers [27, 28]. Thus, these lncRNAs may serve as tumor-related prognostic indicators. Recently, Robinson and Henderson constructed a network model to identify hypothetical mRNA–miRNA interaction networks associated with epithelial function [29]. A mRNA–miRNA interaction list was generated using functional-molecular databases and network modeling. The authors quantified and visualized inherent network structures by using R code and identified a subnetwork containing a large number of shared, targeting miRNAs, of genes related to cancer and cellular proliferation, including cyclin D and c-MYC. The complexity of the miRNA–mRNA interaction network represents an obstacle for predicting and verifying the function of the ceRNA network [29]. The ceRNA hypothesis was proposed to explain the mechanism of tumorigenesis; this hypothesis provides a novel guiding theory and suggests valuable strategies and research directions for the diagnosis and treatment of malignancies [10]. lncRNAs with sequences similar to those of their target miRNAs are able to regulate the expression of mRNAs by acting as sponges for miRNAs [10].



To the best of our knowledge, few studies on ceRNAs have focused on predicting HCC prognosis. Additionally, rare yet reliable lncRNAs, miRNAs, and mRNAs related to HCC can be treated as molecular biomarkers to detect HCC and stratify the HCC risk. Under this background and hypothesis, lncRNA, mRNA, and miRNA data from 371 HCC samples were collected from the TCGA. Based on comprehensive integration of the lncRNA, mRNA, and miRNA data from the 43 paired HCC samples, a ceRNA network associated with lncRNAs was constructed to study the regulatory mechanism of the ceRNAs. Although the biogenesis of HCC is extremely complex, the complexity of tumor growth and metastasis dissemination can be largely represented by DEGs; therefore, we used DEGs

to construct the ceRNA network. Here, we develop a ceRNA network including 26 lncRNAs, four miRNAs, and six mRNAs that are specific to HCC. Among them, four lncRNAs (MYCNOS, DLX6-AS1, LINC00221, and CRNDE) and two mRNAs (CCNB1 and SHCBP1) presented an obvious relevance to overall survival using patient data obtained from both the TCGA and GEO databases.

In the ceRNA network, the lncRNA LINC00221 (connection degree = 2) had the highest connection degree among the prognostic DE lncRNAs (MYCNOS, DLX6-AS1, LINC00221, and CRNDE) (Additional file 5: Table S5). Therefore, we concluded that it might exert a strong influence on HCC pathogenesis. Russell et al. showed that the expression of LINC00221 is 2.4-fold



higher in high-risk neuroblastomas than in low-risk neuroblastomas [30]. LINC00221 was involved in ceRNA networks in several tumor types. Fan et al. reported that LINC00221 is upregulated in breast cancer compared to adjacent-normal breast tissues and is involved in the lncRNA–miRNA–mRNA ceRNA network of breast cancer [31]. He et al. demonstrated that LINC00221 is dysregulated in gastric cancer and competes with six miRNAs (hsa-miR-96, hsa-miR-143, hsa-miR-204,

hsa-miR-372, hsa-miR-373 and hsa-miR-519d) to mediate target mRNA expression in gastric cancer [32]. Zhang et al. found that LINC00221 is associated with a poor prognosis in patients with gastric cancer, and an 11-lncRNA signature consisting of LINC00221 can predict the survival rate for gastric cancer [33]. Wang et al. noted that LINC00221 is involved in the lncRNA–miRNA–mRNA ceRNA network of muscle-invasive bladder cancer, and its high expression level



was significantly related to the progression of muscle-invasive bladder cancer [34]. Our study also showed that LINC00221 expression is 77-fold higher in HCC tissues than in paired nontumorous tissues. Increased expression of LINC00221 was associated with a poor prognosis in patients with HCC. Moreover, we found that LINC00221 might compete with two key DEmiRNAs (hsa-miR-182 and hsa-miR-96) to mediate target DEmRNA expression in HCC.

In addition to lncRNAs, miRNAs should also receive extensive attention. Undoubtedly, research related to tumorigenicity in terms of miRNA regulation is critically needed. miRNAs are RNA molecules of approximately 22 nucleotides in length that bind to the 3'-untranslated region (3'-UTR) of their respective target genes and exert their influence on gene expression by either inhibiting protein translation or degrading mRNA [35]. miRNAs are known to be an integral component of cancer development [35]. We noted that the DEmiRNA hsa-miR-182 (connection degree=13) had the highest connection degree among the prognostic DEmiRNAs (hsa-miR-182 and hsa-miR-183) in the ceRNA network, suggesting an obvious influence of hsa-miR-424 on HCC pathogenesis and prognosis (Additional file 5: Table S5). miR-182 is located on human chromosome 7q31-34 [36]. miR-182 promotes tumorigenesis in a variety of tumors and is one of the most frequently studied cancer-associated miRNAs [36]. In glioblastoma, miR-182-5p targets protein phosphatase 1 regulatory inhibitor subunit 1C [37]. By downregulating RAB27A expression, miR-182-5p improves the migration, mitosis, viability, and invasion capabilities of human gastric cancer cells [37]. Inhibiting miR-182-5p by regulating CASP9 expression confers pro-apoptotic and anti-proliferative effects in human breast cancer [38]. Activated STAT3 induces miR-182-5p expression, which enhances the growth of gliomas [39]. Tang et al. used real-time quantitative PCR to detect the expression of miR-182 in normal cervical epithelial cells and primary cervical cancer tissues and found that miR-182 was significantly upregulated in primary cervical cancer and that the expression level of miR-182 was significantly associated with the patient's International Federation of Gynecology and Obstetrics (FIGO) cancer stage level [40]. Aberrant expression of miR-182 promotes melanoma metastasis by inhibiting microphthalmia-associated transcription factor and FOXO3 [41, 42]. In HCC, Yu et al. reported that miR-182 was one of the most significantly overexpressed miRNAs in HCC [43]. Moreover, the upregulation of miR-182 expression was associated with unfavorable prognosis of HCC patients and intrahepatic metastasis. In cisplatin-treated HCC cells, upregulated miR-182-5p promotes drug resistance by targeting tumor protein 53-induced nuclear protein

1 (TP53INP1) [42]. By regulating metastasis suppressor 1 (MTSS1), miR-182-5p contributes to HCC metastasis [43]. Furthermore, upregulated expression of miR-182-5p may be a diagnostic and prognostic indicator for HCC patients [41]. Consistent with previous studies, our study showed that miR-182 is upregulated by sixfold in HCC, is associated with poor prognosis, and may act through downstream targets (Additional file 1: Table S1).

In addition, we investigated whether the ratios of hsa-miR-182 to its target RNAs (THBS1 and CHL1) and the ratio of hsa-miR-183 to its target RNA (CCNB1) have an impact on the prognosis of patients with HCC. We found that the hsa-miR-182/CHL1 ratio and the hsa-miR-183/CCNB1 ratio were significantly associated with the overall survival of patients with HCC. Zhu et al. reported that upregulation of miR-182 was significantly associated with CHL1 downregulation in papillary thyroid carcinoma (PTC) cell lines and human PTC tissues [44]. miR-182 inhibits the expression of CHL1 by directly targeting the 3'-UTR. Downregulation of miR-182 inhibited the invasion and growth of PTC cells. They concluded that miR-182 in PTC promotes cell invasion and proliferation by directly inhibiting CHL1. In glioblastoma, anti-miR-182 expression results in increased expression of biomarkers for proliferation and stem cell biomarkers (e.g., CCNB1, CD44, Sox2, and Nestin) [45]. Xu et al. reported that hsa-miR-183/CCNB1 may be related to the effect of ribavirin on HCC [46]. Therefore, we believe that the hsa-miR-182/CHL1 and hsa-miR-183/CCNB1 axes may have a strong influence on HCC pathogenesis and may provide novel therapeutic targets for HCC treatment.

Among the prognostic DEmRNAs, CCNB1 and SHCBP1 had the same connection degree (connection degree=1) in the ceRNA network (Additional file 5: Table S5). CCNB1, a rigorous quality control regulator and an important initiator of mitosis, is a key member of the cyclin family [47, 48]. By promoting the transition of the cell cycle from the G2 phase to mitosis, CCNB1 plays a key role in forming cyclin-dependent kinase 1 (CDK1) complexes [47, 48]. The dysregulated expression of CCNB1 is observed in many different cancers, including melanoma and esophageal squamous cell carcinoma [49, 50]. There is increasing evidence that CCNB1 is involved in checkpoint control and that its dysfunction is an early event in tumorigenesis [49, 50]. At the same time, there is evidence that inhibition of CCNB1 expression makes breast cancer cells more sensitive to the chemotherapy drug paclitaxel [51]. CCNB1 also has significant predictive power in monitoring hormone therapy efficacy and the prognosis of patients with ER+ breast cancer [52]. CCNB1 is also a biomarker for HBV-related HCC recurrence [53]. The overexpression of CCNB1 was an independent factor for unfavorable disease-free survival

and had an unfavorable prognosis in patients with lung adenocarcinoma [54]. Gu et al. found that CCNB1 overexpression is closely related to poor survival in patients with HCC [55]. Knockdown of CCNB1 by RNA interference significantly suppressed HCC cell invasion, migration, and proliferation. Furthermore, they also found that miR-144 inhibits CCNB1 expression by directly targeting CCNB1. Moreover, miR-144 negatively regulates CCNB1 to delay tumor formation. The study concluded that the miR-144/CCNB1 axis plays a vital role in HCC. In the present study, our results showed that CCNB1 was upregulated in HCC samples compared to the paired adjacent nontumorous samples from both the TCGA and GEO databases. Furthermore, the survival analysis indicated that patients with HCC exhibiting high CCNB1 expression levels presented a poor prognosis in both the TCGA and GEO databases.

SHC SH2-domain binding protein 1 (SHCBP1) is a cytoplasmic protein that couples signaling pathways to activated growth factor receptors [56]. The expression of SHCBP1 mRNA levels has significant individual differences in human normal tissues. SHCBP1 mRNA and protein are absent in normal, quiescent tissues but are selectively expressed in tissues containing proliferating cells or even cancer cells, demonstrating that SHCBP1 may be involved in tumor development [56]. Immunohistochemical analysis revealed that SHCBP1 was upregulated in breast cancer [57]. High expression level of SHCBP1 was related to poorer survival and advanced clinical stage [57]. In gliomas, by activating the NF- $\kappa$ B signaling pathway, increased expression of SHCBP1 contributed to invasion and migration [58]. In HCC, SHCBP1 was significantly overregulated [59]. Upregulation of SHCBP1 significantly promoted colony formation and survival and cell proliferation in HCC cell lines [59]. In parallel, knockdown of SHCBP1 suppressed cell proliferation and induced cell cycle delay [59]. Our findings indicate that SHCBP1 was upregulated in HCC and that its overexpression was related to poor prognosis in patients with HCC.

Although our ceRNA network identifies many HCC-related lncRNAs, miRNAs, and mRNAs, the correlation and the extent of ceRNA effects *in vivo* remain poorly understood. Recent experimental studies have shown that miRNA-mediated competition between ceRNAs plays a vital role in many biological contexts by constituting additional levels of posttranscriptional regulation [60]. Sensitivity analysis demonstrates that repression mechanisms and binding free energy are vital factors

for cross-talk between ceRNAs. Interactions that occur within a particular range of inhibitory values can be asymmetric (one ceRNA influences another but not the reverse) or symmetrical (one ceRNA influences another and vice versa) and can be limited by noise; at the same time, the interactions can be highly selective [60]. All in all, there are many criteria for validating ceRNA networks, such as cellular concentrations of RNA-binding proteins (RBPs) and miRNAs, timescales, steady-state parameters, kinetic parameters, the absolute concentration of the effective target pool, the miRNA:target ratio, miRNA concentration, the size and affinities of the competing target pool, quantitative measurements of miRNA, target abundance, and so on [61–65]. Therefore, our results still need to be verified through *in vivo* and *in vitro* experiments and clinical practice. Although the ceRNA network has many interference factors in experimental verification, this deficit did not hinder the reliability of the ceRNA network because our network was based on rigorous processes. First, we included only cancer-specific lncRNAs, miRNAs, and mRNAs that had an absolute fold change >4 and an adjusted P-value <0.01. Second, the interactions between DELncRNAs and DEMiRNAs and between DEMiRNAs and DEMRNAs were predicted by experiment-supported databases, such as miRTarBase. These two approaches guarantee that the interactions identified not only occur *in silico* but are also based on experimentally supported evidence. Therefore, we believe that the genes in the current ceRNA network are important for HCC. In the future, with the emergence of larger sample sizes, better databases, and better algorithms, a more comprehensive ceRNA network will be constructed. In addition, further research is warranted on the functions of key ceRNAs *in vivo* and *in vitro*.

## Conclusions

In conclusion, a ceRNA network including 26 DELncRNAs, four DEMiRNAs, and six DEMRNAs was successfully built. Importantly, four lncRNAs (MYCNOS, DLX6-AS1, LINC00221, and CRNDE) and two mRNAs (CCNB1 and SHCBP1) were remarkably related to the prognosis of patients with HCC in both the TCGA and GEO databases. Our research provides novel insights that will increase our understanding of the prognosis-related ceRNA network in HCC. Furthermore, the candidate lncRNAs, miRNAs, and mRNAs involved in the ceRNA network can be further evaluated as potential therapeutic targets and prognostic biomarkers for HCC.

## Additional files

**Additional file 1: Table S1.** Differentially expressed genes between HCC samples and paired nontumorous samples.

**Additional file 2: Table S2.** Forty-four DElncRNAs interacted with nine DEmiRNAs retrieved from the miRcode database.

**Additional file 3: Table S3.** Four DEmiRNAs interacted with six DEMRNAs retrieved from the miRDB, miRTarBase and TargetScan databases.

**Additional file 4: Table S4.** Interactions of the ceRNA network in HCC.

**Additional file 5: Table S5.** The connection degree of each gene in the ceRNA network.

**Additional file 6: Table S6.** Eighteen DElncRNAs were associated with the overall survival of patients with HCC in the TCGA HCC cohort.

**Additional file 7: Table S7.** Six DEMRNAs were associated with the overall survival of patients with HCC in the TCGA HCC cohort.

**Additional file 8: Table S8.** Seven DElncRNAs were associated with the overall survival of patients with HCC in the meta-GEO HCC cohort.

**Additional file 9: Table S9.** Five DEMRNAs were associated with the overall survival of patients with HCC in the meta-GEO HCC cohort.

**Additional file 10: Table S10.** Three DEmiRNAs were associated with the overall survival of patients with HCC in the TCGA HCC cohort.

**Additional file 11: Table S11.** Gene enrichment in the high CCNB1 expression group of patients with HCC in the TCGA HCC cohort.

**Additional file 12: Table S12.** Gene enrichment in the high SHCBP1 expression group of patients with HCC in the TCGA HCC cohort.

**Additional file 13.** R code.

## Abbreviations

HCC: hepatocellular carcinoma; ceRNA: competing endogenous RNA; lncRNA: long noncoding RNA; miRNA: microRNA; DEG: differentially expressed gene; DElncRNA: differentially expressed lncRNA; DEmiRNA: differentially expressed miRNA; DEMRNA: differentially expressed mRNA; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; FC: fold change; GSEA: gene set enrichment analysis; HR: hazard ratio; MRE: miRNA response element; TMM: trimmed mean of M value; FDR: false discovery rate.

## Authors' contributions

Writing, original draft preparation: JL, YB, and XY; software: JL; validation, XY; formal analysis: DW; investigation: LH; project administration: YZ and HZ. All authors read and approved the final manuscript.

## Author details

<sup>1</sup> Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China. <sup>2</sup> Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA.

## Acknowledgements

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Availability of data and materials

The authors declare that the data supporting the findings of this study are available within the article. The R script, which was used to generate figures and reproduce key findings in this study, was stored as Additional file 13.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Not applicable.

## Funding

The design of the study and the collection, analysis, and interpretation of the data in this work were supported by the International Science and Technology Cooperation Projects (2016YFE0107100), the Capital Special Research Project for Health Development (2014-2-4012), the Beijing Natural Science Foundation (L172055 and 7192158), the National Ten-thousand Talent Program, the Fundamental Research Funds for the Central Universities (3332018032), and the CAMS Innovation Fund for Medical Science (CIFMS) (2017-I2M-4-003 and 2018-I2M-3-001).

## Publisher's Note

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

Received: 25 September 2018 Accepted: 5 April 2019

Published online: 11 April 2019

## References

- Galun D, Bogdanovic A, Djokic Kovac J, Bulajic P, Loncar Z, Zuvella M. Pre-operative neutrophil-to-lymphocyte ratio as a prognostic predictor after curative-intent surgery for hepatocellular carcinoma: experience from a developing country. *Cancer Manag Res*. 2018;10:977–88.
- Llovet JM, Zucman-Rossi J, Pikarsky E, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers*. 2016;2:16018.
- Luo CL, Xu ZG, Chen H, et al. LncRNAs and EGFRVIII sequestered in TEPs enable blood-based NSCLC diagnosis. *Cancer Manag Res*. 2018;10:1449–59.
- Ye X, Wang X, Shang L, et al. Genetic variants of ALDH2-rs671 and CYP2E1-rs2031920 contributed to risk of hepatocellular carcinoma susceptibility in a Chinese population. *Cancer Manag Res*. 2018;10:1037–50.
- Mittal S, El-Serag HB. Epidemiology of hepatocellular carcinoma: consider the population. *J Clin Gastroenterol*. 2013;47(Suppl):S2–6.
- Yamamura S, Imai-Sumida M, Tanaka Y, Dahiya R. Interaction and cross-talk between non-coding RNAs. *Cell Mol Life Sci*. 2018;75(3):467–84.
- Sana J, Faltejskova P, Svoboda M, Slaby O. Novel classes of non-coding RNAs and cancer. *J Transl Med*. 2012;10:103.
- Takahashi K, Yan I, Haga H, Patel T. Long noncoding RNA in liver diseases. *Hepatology (Baltimore, MD)*. 2014;60(2):744–53.
- Ba Z, Gu L, Hao S, Wang X, Cheng Z, Nie G. Downregulation of lncRNA CASC2 facilitates osteosarcoma growth and invasion through miR-181a. *Cell Prolif*. 2018;51(1):e12409.
- Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 2011;146(3):353–8.
- Tan JY, Sirey T, Honti F, et al. Extensive microRNA-mediated crosstalk between lncRNAs and mRNAs in mouse embryonic stem cells. *Genome Res*. 2015;25(5):655–66.
- Karreth FA, Pandolfi PP. ceRNA cross-talk in cancer: when ce-bling rivalries go awry. *Cancer Dis*. 2013;3(10):1113–21.
- Wang J, Liu X, Wu H, et al. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res*. 2010;38(16):5366–83.
- Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*. 2010;465(7301):1033–8.
- Wang SH, Ma F, Tang ZH, et al. Long non-coding RNA H19 regulates FOXM1 expression by competitively binding endogenous miR-342-3p in gallbladder cancer. *J Exp Clin Cancer Res*. 2016;35(1):160.
- Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*. 2010;26(1):139–40.
- 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.
- Rosvall M, Bergstrom CT. Mapping change in large networks. *PLoS ONE*. 2010;5(1):e8694.
- Jeggari A, Marks DS, Larsson E. miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics (Oxford, England)*. 2012;28(15):2062–3.

20. Wong N, Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* 2015;43(Database issue):D146–52.
21. Chou CH, Shrestha S, Yang CD, et al. miRTarBase update 2018: a resource for experimentally validated microRNA–target interactions. *Nucleic Acids Res.* 2018;46(D1):D296–302.
22. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *eLife.* 2015;4:e05005.
23. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* 2005;102(43):15545–50.
24. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *Cancer J Clin.* 2017;67(1):7–30.
25. Chen D, Zhang C, Lin J, Song X, Wang H. Screening differential circular RNA expression profiles reveal that hsa\_circ\_0128298 is a biomarker in the diagnosis and prognosis of hepatocellular carcinoma. *Cancer Manag Res.* 2018;10:1275–83.
26. Muers M. RNA: genome-wide views of long non-coding RNAs. *Nat Rev Genet.* 2011;12(11):742.
27. Caley DP, Pink RC, Trujillano D, Carter DR. Long noncoding RNAs, chromatin, and development. *Sci World J.* 2010;10:90–102.
28. Augoff K, McCue B, Plow EF, Sossey-Alaoui K. miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer. *Mol Cancer.* 2012;11:5.
29. Robinson JM, Henderson WA. Modelling the structure of a ceRNA-theoretical, bipartite microRNA–mRNA interaction network regulating intestinal epithelial cellular pathways using R programming. *BMC Res Notes.* 2018;11(1):19.
30. Russell MR, Penikis A, Oldridge DA, et al. CASC15-5 is a tumor suppressor lncRNA at the 6p22 neuroblastoma susceptibility locus. *Cancer Res.* 2015;75(15):3155–66.
31. Fan CN, Ma L, Liu N. Systematic analysis of lncRNA–miRNA–mRNA competing endogenous RNA network identifies four-lncRNA signature as a prognostic biomarker for breast cancer. *J Transl Med.* 2018;16(1):264.
32. He W, Zhang D, Li X, et al. TCGA dataset-based construction and integrated analysis of aberrantly expressed long noncoding RNA mediated competing endogenous RNA network in gastric cancer. *Oncol Rep.* 2018;40(6):3511–22.
33. Zhang Y, Li H, Zhang W, Che Y, Bai W, Huang G. LASSO-based CoxPH model identifies an 11lncRNA signature for prognosis prediction in gastric cancer. *Mol Med Rep.* 2018;18(6):5579–93.
34. Wang H, Niu L, Jiang S, et al. Comprehensive analysis of aberrantly expressed profiles of lncRNAs and miRNAs with associated ceRNA network in muscle-invasive bladder cancer. *Oncotarget.* 2016;7(52):86174–85.
35. Patel N, Garikapati KR, Makani VKK, et al. Regulating BMI1 expression via miRNAs promote mesenchymal to epithelial transition (MET) and sensitizes breast cancer cell to chemotherapeutic drug. *PLoS ONE.* 2018;13(2):e0190245.
36. Arya D, Sachithanandan SP, Ross C, Palakodeti D, Li S, Krishna S. MiRNA182 regulates percentage of myeloid and erythroid cells in chronic myeloid leukemia. *Cell Death Dis.* 2017;8(1):e2547.
37. Liu L, Zhang X, Nan C, et al. MicroRNA-182 targets protein phosphatase 1 regulatory inhibitor subunit 1C in glioblastoma. *Oncotarget.* 2017;8(70):114677–84.
38. Sharifi M, Moridnia A. Apoptosis-inducing and antiproliferative effect by inhibition of miR-182-5p through the regulation of CASP9 expression in human breast cancer. *Cancer Gene Ther.* 2017;24(2):75–82.
39. Xue J, Zhou A, Wu Y, et al. miR-182-5p induced by STAT3 activation promotes glioma tumorigenesis. *Cancer Res.* 2016;76(14):4293–304.
40. Tang T, Wong HK, Gu W, et al. MicroRNA-182 plays an onco-miRNA role in cervical cancer. *Gynecol Oncol.* 2013;129(1):199–208.
41. Wang Y, Lee AT, Ma JZ, et al. Profiling microRNA expression in hepatocellular carcinoma reveals microRNA-224 up-regulation and apoptosis inhibitor-5 as a microRNA-224-specific target. *J Biol Chem.* 2008;283(19):13205–15.
42. Rossi S, Shimizu M, Barbarotto E, et al. microRNA fingerprinting of CLL patients with chromosome 17p deletion identify a miR-21 score that stratifies early survival. *Blood.* 2010;116(6):945–52.
43. Segura MF, Hanniford D, Menendez S, et al. Aberrant miR-182 expression promotes melanoma metastasis by repressing FOXO3 and microphthalmia-associated transcription factor. *Proc Natl Acad Sci USA.* 2009;106(6):1814–9.
44. Zhu H, Fang J, Zhang J, et al. miR-182 targets CHL1 and controls tumor growth and invasion in papillary thyroid carcinoma. *Biochem Biophys Res Commun.* 2014;450(1):857–62.
45. Kouri FM, Hurlley LA, Daniel WL, et al. miR-182 integrates apoptosis, growth, and differentiation programs in glioblastoma. *Genes Dev.* 2015;29(7):732–45.
46. Xu C, Luo L, Yu Y, et al. Screening therapeutic targets of ribavirin in hepatocellular carcinoma. *Oncol Lett.* 2018;15(6):9625–32.
47. Allemanni C, Weir HK, Carreira H, et al. Global surveillance of cancer survival 1995–2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). *Lancet (London, England).* 2015;385(9972):977–1010.
48. Morgan DO. Principles of CDK regulation. *Nature.* 1995;374(6518):131–4.
49. Yoshida T, Tanaka S, Mogi A, Shitara Y, Kuwano H. The clinical significance of Cyclin B1 and Wee1 expression in non-small-cell lung cancer. *Ann Oncol.* 2004;15(2):252–6.
50. Nozoe T, Korenaga D, Kabashima A, Ohga T, Saeki H, Sugimachi K. Significance of cyclin B1 expression as an independent prognostic indicator of patients with squamous cell carcinoma of the esophagus. *Clin Cancer Res.* 2002;8(3):817–22.
51. Androic I, Kramer A, Yan R, et al. Targeting cyclin B1 inhibits proliferation and sensitizes breast cancer cells to taxol. *BMC Cancer.* 2008;8:391.
52. Ding K, Li W, Zou Z, Zou X, Wang C. CCNB1 is a prognostic biomarker for ER+ breast cancer. *Med Hypotheses.* 2014;83(3):359–64.
53. Weng L, Du J, Zhou Q, et al. Identification of cyclin B1 and Sec62 as biomarkers for recurrence in patients with HBV-related hepatocellular carcinoma after surgical resection. *Mol Cancer.* 2012;11:39.
54. Liu W, Ouyang S, Zhou Z, et al. Identification of genes associated with cancer progression and prognosis in lung adenocarcinoma: analyses based on microarray from OncoPrint and The Cancer Genome Atlas databases. *Mol Genet Genomic Med.* 2018;7:e0052.
55. Gu J, Liu X, Li J, He Y. MicroRNA-144 inhibits cell proliferation, migration and invasion in human hepatocellular carcinoma by targeting CCNB1. *Cancer Cell Int.* 2019;19:15.
56. Schmandt R, Liu SK, McGlade CJ. Cloning and characterization of mPAL, a novel Shc SH2 domain-binding protein expressed in proliferating cells. *Oncogene.* 1999;18(10):1867–79.
57. Feng W, Li HC, Xu K, et al. SHCBP1 is over-expressed in breast cancer and is important in the proliferation and apoptosis of the human malignant breast cancer cell line. *Gene.* 2016;587(1):91–7.
58. Zhou Y, Tan Z, Chen K, et al. Overexpression of SHCBP1 promotes migration and invasion in gliomas by activating the NF-kappaB signaling pathway. *Mol Carcinog.* 2018;57(9):1181–90.
59. Tao HC, Wang HX, Dai M, et al. Targeting SHCBP1 inhibits cell proliferation in human hepatocellular carcinoma cells. *Asian Pac J Cancer Prev.* 2013;14(10):5645–50.
60. Figliuzzi M, Marinari E, De Martino A. MicroRNAs as a selective channel of communication between competing RNAs: a steady-state theory. *Biophys J.* 2013;104(5):1203–13.
61. Jens M, Rajewsky N. Competition between target sites of regulators shapes post-transcriptional gene regulation. *Nat Rev Genet.* 2015;16(2):113–26.
62. Martirosyan A, Figliuzzi M, Marinari E, De Martino A. Probing the limits to microRNA-mediated control of gene expression. *PLoS Comput Biol.* 2016;12(1):e1004715.
63. Bosson AD, Zamudio JR, Sharp PA. Endogenous miRNA and target concentrations determine susceptibility to potential ceRNA competition. *Mol Cell.* 2014;56(3):347–59.
64. Denzler R, Agarwal V, Stefano J, Bartel DP, Stoffel M. Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol Cell.* 2014;54(5):766–76.
65. Figliuzzi M, De Martino A, Marinari E. RNA-based regulation: dynamics and response to perturbations of competing RNAs. *Biophys J.* 2014;107(4):1011–22.