

# Identification of Function LncRNAs Associated With Hepatocellular Carcinoma

**Xiaolei Wang**

Qingdao Third People's Hospital

**Jun Fang**

Qingdao Third People's Hospital

**An Wang**

Qingdao Third People's Hospital

**Wuhui Zhu** (✉ [zhuwuhui@126.com](mailto:zhuwuhui@126.com))

Qingdao Third People's Hospital

---

## Research Article

**Keywords:** hepatocellular carcinoma, survival analysis, prognosis

**Posted Date:** March 8th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1303494/v2>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

## Background

Hepatocellular carcinoma (HCC) is one of the most common malignant tumors of the digestive tract. New diagnostic and prognostic biomarkers are urgently needed for HCC. To investigate potential pathogenic and prognostic long noncoding RNAs (lncRNAs) associated with HCC, we conduct this study.

## Material and Methods

Gene expression matrix were obtained from TCGA and GEO databases. The overlapping differentially expressed genes (DEGs) between TCGA and GEO databases were analyzed using the limma package. Then we constructed mRNA-lncRNA-miRNA network. In the end, Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and survival analysis were performed to find hub lncRNAs related to diagnosis and prognosis of HCC.

## Results

In total, 299 DEGs including 10 lncRNAs and 289 mRNAs were acquired. The mRNA-lncRNA-miRNA network comprised of four lncRNAs (GSEC, LINC00844, LINC01146 and MIR99AHG), 49 miRNAs and 25 mRNAs. GO and KEGG analysis shown that four lncRNAs were mainly enriched in cell cycle. In addition, survival analysis demonstrated that there is a significant correlation between GSEC and overall survival rate. Finally, the expression of LINC00844, LINC01146 and MIR99AHG were down-regulated in HCC tissues, and GSEC was up-regulated.

## Conclusion

In summary, four lncRNAs including GSEC, LINC00844, LINC01146 and MIR99AHG may play an important part in the development of HCC. The function of these lncRNAs may have a great effect to the overall survival rate of HCC.

## Background

Hepatocellular carcinoma (HCC) is the most common histological type of primary liver cancer (90%) and the fourth leading cause of cancer-related death in the world [1]. TNM staging of tumors has an important value in treatment and diagnosis of tumors, but its predictive value for clinical prognosis is limited. Despite there has a great progress in pathology, epidemiology, and molecular mechanisms at present, there is still room for improvement in predicting the clinical prognosis of HCC. In terms of therapeutic methods, many new methods such as targeted therapy and immunotherapy have emerged. However, therapeutic treatments only benefit part of patients. Thus, novel molecular markers associated with diagnosis and prognosis are still urgently needed to ameliorate bad outcome of HCC. It's important to discriminate the high-risk patients in the early stage. Developing effective treatment methods is also critical to reduce the high reoccurrence of HCC.

An increasing body of research demonstrated that lncRNAs play a more and more important part in various cancers, including breast cancer [2], gallbladder cancer [3], colorectal carcinoma [4], and HCC [5]. The abnormal expression of lncRNAs in HCC is closely associated with oncogenesis, metastasis and prognosis. Some carcinogenic and anticancer lncRNAs associated with HCC have been found [5–7]. Up-regulated lncRNAs, such as ATB, MCM3AP-AS1, PCNAP1 and H19 have been proved to be closely related to various clinical features and cell phenotypes [5, 8–10]. For instance, m5C-modified H19, which is a lncRNA, may promote the development of carcinoma by recruiting G3BP1 [10]. Of the down-regulated lncRNAs, uc.134 suppresses HCC progression by inhibiting the CUL4A-mediated ubiquitination of LATS1. It also can restrain HCC development by increasing YAP<sup>S127</sup> phosphorylation [11]. While cytosporone-B, which is an agonist for Nur77, could promote WFDC21P, then it can inhibit HCC through WFDC21P [12].

Many studies had demonstrated that lncRNAs had various effects in HCC. Competing endogenous RNA (ceRNA) network is one of the most popular mechanism, that is to say, mRNA-lncRNA-miRNA network. The ceRNA network hypothesis considered that lncRNAs are important endogenous molecular sponges, which can competitively integrate miRNAs. This process can regulate mRNA expression levels. Salmena et al. first proposed this hypothesis [13]. Lots of researches have demonstrated that mRNA-lncRNA-miRNA network plays an important part in tumorigenesis, development and metastasis. For instance, lncRNA SNHG8 acted as a sponge of miR-149 which can balance the tumor inhibiting effects of miR-149 in HCC [14]. In addition, LINC01093/miR-96-5p/ZFAND5/NF-κB signaling axis may have an important role in the pathogenesis of HCC [15]. Thus, the ceRNA network is beneficial to gain insight into intricated gene interactions in HCC.

Although there are many researches had studied the diagnostic capability of lncRNAs, lots of lncRNAs are worth of further exploration [16]. Take the important function of lncRNAs of HCC into consideration, the purpose of this research was to explore the role of lncRNAs in HCC patients.

Gene expression matrix of mRNA and lncRNA were took into integrated analysis. We constructed ceRNA network on the basis of ceRNA theory. In the first place, DEGs were acquired from HCC and normal tissue samples which were obtained from GEO and TCGA database. Then the overlapping DEGs between two databases were obtained by using the limma package. In the end, in order to analyze the underlying roles of lncRNAs, we constructed lots of analysis, such as mRNA-lncRNA-miRNA network, GO analysis, KEGG pathway analysis, protein-protein interaction (PPI) network and survival analysis. This research may provide novel views to the molecular mechanism of HCC.

## 2 Material And Methods

### 2.1 Data source

The microarray data of HCC was obtained from GEO ([www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo)) and TCGA database. First, we got the eligible data from GEO database. It was considered eligible when studies meet the following criteria: (1) Studies including HCC and adjacent normal tissues. (2) Studies containing matched platform information which can be used for further analyze. On the basis of these criteria, two independent datasets of HCC were acquired from GEO database. The detailed information of each dataset is shown in Table 1. Then, clinical data of HCC patients and the corresponding gene expression profile data were downloaded from TCGA. Clinical data of these HCC patients is displayed in Table 2. Our present study was performed as shown in the flow chart (Fig. 1).

### 2.2 Reannotation of raw data

In the begining, the raw data and expression matrix were obtained from GEO datasets. The probes which had unique gene were retained for analysis. Then, the retained probes should match the corresponding lncRNAs or mRNAs retrieved from the GENCODE. We used average expression values when lots of probes corresponded to one gene.

### 2.3 Identification of differential expression analysis

The raw data were transformed into expression matrix using Affy R package. Then, the limma package was used to identify DEGs from TCGA and GEO datasets. A  $|log FCI| > 1$  and  $adjust P value < 0.05$  was regarded as significantly different. The intersected DEGs among TCGA, GSE101728 and GSE93789 datasets were obtained. And then these DEGs were used to do further analyze.

### 2.4 Construction of mRNA-lncRNA-miRNA network

To analyze whether these lncRNAs and miRNAs can construct ceRNA network, we used DIANA Tools to collect lncRNA-targeted miRNAs. MiRDB was used to predict miRNA-mRNA interactions. The DEGs were used to make mRNA-lncRNA-miRNA network. We used the Cytoscape display mRNA-lncRNA-miRNA network.

### 2.5 Functional enrichment analysis

GO and KEGG pathway enrichment analysis were performed using cluster Profiler package. Protein-protein interaction (PPI) network were conducted using STRING v 9.1 (<http://string-db.org/>).

### 2.6 Survival analysis

In order to analyze if the hub lncRNAs and corresponding mRNAs were associated to the prognosis of HCC patients, Kaplan-Meier (KM) analysis was constructed using the survival package. All cases were split into two groups using the median value of a gene's expression. It was considered as statistically significant when  $P$  value was less than 0.05.

### 2.7 Data availability

TCGA (The Cancer Genome Atlas) dataset is available using the following link: <https://portal.gdc.cancer.gov/>. GEO (Gene Expression Omnibus) dataset is available using the following link: <https://www.ncbi.nlm.nih.gov/geo/>.

## 3 Results

### 3.1 Characteristics of gene expression from GEO

The gene expression data (GSE93789, GSE101728) were obtained from the GEO data repository (<http://www.ncbi.nlm.nih.gov/geo/>). After reannotation, 33591 genes in GSE93789 (GPL16956) and 46102 genes in GSE101728 (GPL21047) were obtained. On the basis of the criteria of  $|log FC| > 1$  and  $p\text{-adjusted value} < 0.05$ , gene expression data were analyzed for differentially expressed genes using limma package. The GSE93789 had 4553 DEGs, containing 2735 up-regulated and 1818 down-regulated genes. The GSE101728 had 1599 DEGs, containing 657 up-regulated and 942 down-regulated genes. The particular information of GSE93789 and GSE101728 data set and the number of DEGs acquired

from the two data sets are displayed in Table 1. We constructed volcano plots to show the significantly DEGs (Fig. 2). We used the heatmap to display the gene expression changes (Fig. 3).

### 3.2 Clinical data of HCC and the intersected DEGs

The clinical data of HCC patients were obtained from TCGA database. The information of HCC patients mainly containing sex, age, TNM stage, tumor grade, Child–Pugh, residual tumor classification and Ishak fibrosis score were collected, as shown in Table 2. Limma package was used to analyze DEGs on the basis of the criteria of  $|Log\ FCI| > 1$  and adjust  $P$  value  $< 0.05$ . In summary, 148 lncRNAs with 68 up-regulated and 80 down-regulated, 2181 mRNAs with 728 up-regulated and 1453 down-regulated were screened. Venn diagram demonstrates intersected genes both coming from GEO and TCGA. Totally, 289 differentially expressed mRNAs and 10 differentially expressed lncRNAs were acquired (Fig. 4 and Table 3).

### 3.3 ceRNA network

In order to analyze if these hub lncRNAs and mRNAs existing mRNA-lncRNA-miRNA network, we used the intersected lncRNAs and mRNAs to perform mRNA-lncRNA-miRNA network. Ultimately, four lncRNAs, forty-nine miRNAs and twenty-five mRNAs were involved in mRNA-lncRNA-miRNA network (Fig. 5).

### 3.4 GO, KEGG pathway analysis, and PPI network

In order to analyze the underlying biological roles of the hub genes contained in the mRNA-lncRNA-miRNA network, we performed GO and KEGG analysis (Fig. 6). The outcome shown that the biological functions mainly related to nuclear division, sister chromatid segregation, organelle fission and chromosome segregation, while the molecular functions gathered in monooxygenase activity, oxidoreductase activity, iron ion binding and heme binding. The KEGG pathway analysis demonstrated that most of the genes were related to cell cycle, p53 signaling pathway, oocyte meiosis, complement and coagulation cascades and progesterone-mediated oocyte maturation. Furthermore, differentially expressed mRNAs were used to construct PPI network by STRING. The outcome demonstrated that most genes are associated with cell cycle.

### 3.5 Prognostic value of hub lncRNAs and mRNAs in patients with HCC

In order to find out whether the hub lncRNAs and mRNAs involved in the mRNA-lncRNA-miRNA network were related to the prognosis of HCC patients, we analyzed the hub gene expression and the survival data in TCGA. Because some patients didn't have intact information to do survival analysis, 415 cases were obtained in the end. In the four hub lncRNAs, the result demonstrated that lncRNA GSEC was obviously significant associated with the overall survival (OS) ( $P < 0.05$ ). The expression of GSEC was negatively correlated with survival time of HCC. In the 25 hub mRNAs, the result shown that 12 mRNAs were related to the OS of HCC: B3GNT5( $P < 0.001$ ), SORBS2( $P = 0.037$ ), ENAH( $P = 0.003$ ), SUCO( $P = 0.004$ ), ESR1( $P = 0.006$ ), FAT4( $P = 0.047$ ), SLC38A2( $P = 0.012$ ), MASP1( $P = 0.031$ ), GHR( $P = 0.007$ ), PARPBP( $P = 0.002$ ), RRM2( $P < 0.001$ ) and MMS22L( $P < 0.001$ ). The high levels of SORBS2, ESR1, FAT4, MASP1 and GHR predicted good prognosis, but seven mRNAs had a negative correlation with the OS of HCC. The results of KM analysis are displayed in Figure 7.

## 4 Discussion

LncRNAs are becoming more and more popular with the rapid development of high-throughput sequencing and bioinformatic technologies. At present, we obtained the gene expression of HCC from GEO and TCGA. Then we explored the hub lncRNAs of HCC. We constructed mRNA-lncRNA-miRNA network, GO and KEGG pathway analysis, PPI network and survival analysis. In the end, the result shown that four lncRNAs including GSEC, LINC00844, LINC01146 and MIR99AHG may have a relation with HCC development. The underlying roles of these hub lncRNAs might be related to the OS of HCC.

There are many integrated bioinformatics which can be used to find DEGs in HCC. With the rapid development of bioinformation technology, lots of novel genes have been found. In order to find novel lncRNAs and mRNAs associated with the prognosis of HCC, the RNA expression data from GEO was analyzed. In addition, different individual had different gene expression level of the same gene. In order to minimize the error of different individual, we only analyzed paired tissue samples in GEO. After we obtained the DEGs from TCGA, intersected DEGs from GEO and TCGA were used to further analyze. In summary, 10 co-differentially expressed lncRNAs and 289 mRNAs were used to continue next investigation.

More and more studies have shown that ceRNAs have great effects in HCC [17, 18]. To explore the molecular mechanisms of these hub lncRNAs, we constructed mRNA-lncRNA-miRNA network. In the end, the mRNA-lncRNA-miRNA network included 4 lncRNAs (GSEC, LINC00844, LINC01146 and MIR99AHG), 49 miRNAs and 25 mRNAs. We found that the hub lncRNAs can affect mRNAs expression through competitively uniting miRNA.

MIR99AHG may suppress lung adenocarcinoma's proliferation and metastasis and it can promote autophagy [19]. Studies showing that not only MIR99AHG but also miR-99a expression level was positively related to the survival of lung adenocarcinoma patients. Furthermore, when the expression of MIR99AHG of H1299 cells was overexpressed, the nuclear expression of ANXA2 was obviously decreased. The expression of

MIR99AHG was obviously up-regulated in gastric cancer [20]. In addition, MIR99AHG was obviously overexpressed in pancreatic cancer. It can promote the proliferation and metastasis of pancreatic cancer. Studies demonstrated that up-regulated MIR99AHG was induced by transcription factor FOXA1 [21].

LINC00844 was found to be significantly decreased in HCC tissue and was correlated with poor tumor characteristics, such as portal vein invasion, high α-fetoprotein (AFP), and a high rate of tumor recurrence [22]. Studies demonstrated that LINC00844 had an important role in suppressing tumor development and metastasis [23, 24]. Besides, LINC00844 is important in facilitating AR binding to the chromatin. In the end, LINC00844 works partly by activating the expression level of NDRG1 which is an important tumor metastasis suppressor. It has been suggested that LINC01146 was a hub lncRNA in both subtypes of chronic rhinosinusitis with nasal polyps [25]. However, there was no researches have demonstrated the role of LINC01146 in HCC up to now.

Unlike LINC00844 and MIR99AHG, the expression level of GSEC can be up-regulated in many tumors, including colon cancer [26], breast cancer [27], lung cancer [28] and osteosarcoma [29]. Studies have revealed that GSEC has an important influence in cell migration of colon cancer by suppressing the role of DHX36 through its G-quadruplex structure [26]. Moreover, GSEC was significantly up-regulated in triple negative breast cancer [27]. GSEC functioned as a underlying oncogene and it can promote AXL-mediated TNBC progression by sponging miR-202-5p. In addition, High expression level of GSEC promoted the proliferating and migratory capacity, and suppressed the apoptosis of osteosarcoma [29].

There are some limitations in the present study. First of all, this study only contains two GEO data sets, which are both noncoding RNA data. Then, although we acquire the expression matrix of genes by using probe reannotation method, some genes which can't match the probes will not be found because of the platform focused on noncoding RNA data. But the aim of this research was to find hub lncRNA of HCC patients, so the roles of 4 lncRNAs could not be further explored because the mRNAs which were used to analyze were acquired from noncoding RNA data. That is to say, the hub mRNAs were associated with lncRNAs. In total, the results of our analysis are precise and reliable, and it can provide more valuable information to the clinical work. Although we find 4 lncRNAs were statistically related to OS of HCC, more experiments are needed to validate the function of these hub lncRNAs.

## 5 Conclusion

In total, four lncRNAs including GSEC, LINC00844, LINC01146 and MIR99AHG may play an important part in the development of HCC. The function of these lncRNAs may have a great effect to the overall survival rate of HCC.

## Declarations

**Author contributions** WZ: conception and design. XW: development of methodology, analysis and writing of the manuscript. JF and AW: review of the manuscript. No funding was received.

**Funding** No funding was received.

**Conflict of interest** The authors declare that there is no conflict of interests.

**Ethical approval** Not applicable.

**Consent to publish** All authors have given their consent to participate in this report and submit it to Molecular Biology Reports.

## References

1. Hepatocellular carcinoma. *Nature reviews Disease primers*. 2021; 7(1):7.
2. Liang Y, Song X, Li Y, et al. LncRNA BCRT1 promotes breast cancer progression by targeting miR-1303/PTBP3 axis. *Mol Cancer*. 2020; 19(1):85.
3. Hu YP, Jin YP, Wu XS, et al. LncRNA-HGBC stabilized by HuR promotes gallbladder cancer progression by regulating miR-502-3p/SET/AKT axis. *Mol Cancer*. 2019; 18(1):167.
4. Wang FW, Cao CH, Han K, et al. APC-activated long noncoding RNA inhibits colorectal carcinoma pathogenesis through reduction of exosome production. *J Clin Invest*. 2019; 129(2):727-43.
5. Yuan JH, Yang F, Wang F, et al. A long noncoding RNA activated by TGF-β promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell*. 2014; 25(5):666-81.
6. Song X, Du R, Gui H, et al. Identification of potential hub genes related to the progression and prognosis of hepatocellular carcinoma through integrated bioinformatics analysis. *Oncol Rep*. 2020; 43(1):133-46.

7. Yu B, Ding Y, Liao X, Wang C, Wang B, Chen X. Overexpression of PARPBP Correlates with Tumor Progression and Poor Prognosis in Hepatocellular Carcinoma. *Dig Dis Sci.* 2019; 64(10):2878-92.
8. Wang Y, Yang L, Chen T, et al. A novel lncRNA MCM3AP-AS1 promotes the growth of hepatocellular carcinoma by targeting miR-194-5p/FOXA1 axis. *Mol Cancer.* 2019; 18(1):28.
9. Feng J, Yang G, Liu Y, et al. LncRNA PCNAP1 modulates hepatitis B virus replication and enhances tumor growth of liver cancer. *Theranostics.* 2019; 9(18):5227-45.
10. Sun Z, Xue S, Zhang M, et al. Aberrant NSUN2-mediated mC modification of H19 lncRNA is associated with poor differentiation of hepatocellular carcinoma. *Oncogene.* 2020; 39(45):6906-19.
11. Ni W, Zhang Y, Zhan Z, et al. A novel lncRNA uc.134 represses hepatocellular carcinoma progression by inhibiting CUL4A-mediated ubiquitination of LATS1. *J Hematol Oncol.* 2017; 10(1):91.
12. Guan Y-F, Huang Q-L, Ai Y-L, et al. Nur77-activated lncRNA WFDC21P attenuates hepatocarcinogenesis via modulating glycolysis. *Oncogene.* 2020; 39(11):2408-23.
13. 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):353-58.
14. Dong J, Teng F, Guo W, Yang J, Ding G, Fu Z. lncRNA SNHG8 Promotes the Tumorigenesis and Metastasis by Sponging miR-149-5p and Predicts Tumor Recurrence in Hepatocellular Carcinoma. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology.* 2018; 51(5):2262-74.
15. Zheng Y, Yu K, Huang C, et al. Integrated bioinformatics analysis reveals role of the LINC01093/miR-96-5p/ZFAND5/NF-κB signaling axis in hepatocellular carcinoma. *Exp Ther Med.* 2019; 18(5):3853-60.
16. Lim LJ, Wong SYS, Huang F, et al. Roles and Regulation of Long Noncoding RNAs in Hepatocellular Carcinoma. *Cancer Res.* 2019; 79(20):5131-39.
17. Wang H, Huo X, Yang X-R, et al. STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4. *Mol Cancer.* 2017; 16(1):136.
18. Li X, Ding J, Wang X, Cheng Z, Zhu Q. NUDT21 regulates circRNA cyclization and ceRNA crosstalk in hepatocellular carcinoma. *Oncogene.* 2020; 39(4):891-904.
19. Han C, Li H, Ma Z, et al. MIR99AHG is a noncoding tumor suppressor gene in lung adenocarcinoma. *Cell Death Dis.* 2021; 12(5):424.
20. Meng Q, Wang X, Xue T, Zhao Q, Wang W, Zhao K. Long noncoding RNA MIR99AHG promotes gastric cancer progression by inducing EMT and inhibiting apoptosis via miR577/FOXP1 axis. *Cancer Cell Int.* 2020; 20:414.
21. Xu J, Xu W, Yang X, Liu Z, Zhao Y, Sun Q. LncRNA MIR99AHG mediated by FOXA1 modulates NOTCH2/Notch signaling pathway to accelerate pancreatic cancer through sponging miR-3129-5p and recruiting ELAVL1. *Cancer Cell Int.* 2021; 21(1):674.
22. Lin B, He H, Zhang Q, et al. Long non-coding RNA00844 inhibits MAPK signaling to suppress the progression of hepatocellular carcinoma by targeting AZGP1. *Annals of translational medicine.* 2020; 8(21):1365.
23. Lingadahalli S, Jadhao S, Sung YY, et al. Novel lncRNA Regulates Prostate Cancer Cell Migration and Invasion through AR Signaling. *Molecular cancer research : MCR.* 2018; 16(12):1865-78.
24. Qiu K, Zheng Z, Huang Y. Long intergenic noncoding RNA 00844 promotes apoptosis and represses proliferation of prostate cancer cells through upregulating GSTP1 by recruiting EBF1. *J Cell Physiol.* 2020; 235(11):8472-85.
25. Wang M, Bu X, Luan G, et al. Distinct type 2-high inflammation associated molecular signatures of chronic rhinosinusitis with nasal polyps with comorbid asthma. *Clinical and translational allergy.* 2020; 10:26.
26. Matsumura K, Kawasaki Y, Miyamoto M, et al. The novel G-quadruplex-containing long non-coding RNA GSEC antagonizes DHX36 and modulates colon cancer cell migration. *Oncogene.* 2017; 36(9):1191-99.
27. Zhang J, Du C, Zhang L, Wang Y, Zhang Y, Li J. lncRNA GSEC Promotes the Progression of Triple Negative Breast Cancer (TNBC) by Targeting the miR-202-5p/AXL Axis. *Onco Targets Ther.* 2021; 14:2747-59.
28. Tian Y, Yu M, Sun L, et al. Distinct Patterns of mRNA and lncRNA Expression Differences Between Lung Squamous Cell Carcinoma and Adenocarcinoma. *J Comput Biol.* 2020; 27(7):1067-78.
29. Liu R, Ju C, Zhang F, et al. LncRNA GSEC promotes the proliferation, migration and invasion by sponging miR-588/ EIF5A2 axis in osteosarcoma. *Biochem Biophys Res Commun.* 2020; 532(2):300-07.

## Tables

**Table 1.** Details information of HCC researches and related microarray data sets from GEO

GSE	Total differentially expressed genes	Up-regulated	Down-regulated	Platform	Sample size	Age	Sex
GSE93789	4553	2735	1818	GPL16956	total:10; tumor:5;  adjeacent:5	44,45,49,48,40,44,45,49,48,40	Male:8; Female:2
GSE101728	1599	657	942	GPL21047	total:14; tumor:7;  adjeacent:7	57,57,67,67,58,58,52,52,62,62,54,54,27,27	Male:10; Female:4

**Table 2.** The clinical data of HCC in TCGA

Characteristics	Number of cases	Percentages
Sex		
Female	143	33.97
Male	278	66.03
Age		
>60	226	53.68
≤60	195	46.32
Stage		
Stage I	189	44.89
Stage II	97	23.04
Stage III	97	23.04
Stage IV	6	1.43
Histologic grade		
G1	60	14.25
G2	204	48.46
G3	137	32.54
G4	12	2.85
Topography		
T1	201	47.74
T2	108	25.65
T3	93	22.09
T4	16	3.80
Child-Pugh stage		
A	240	57.01
B	30	7.13
C	1	0.24
Residual tumor		
R0	364	86.46
R1	21	4.99
R2	2	0.48
Rx	25	5.94
Ishak fibrosis score		
0	96	22.80
1,2	35	8.31
3,4	31	7.36
5	10	2.38
6	75	17.81

**Table 3.** The intersected hub lncRNAs and mRNAs

lncRNA	AC018467.1 AC115619.1 AL445524.1 ELF3-AS1 GSEC LINC00844 LINC01146 LINC02037 LINC02362 MIR99AHG
mRNA	AADAT AASS ABCA8 ACADSB ADRA1A AKAP12 AKR1C3 AKR1D1 ALDH2 ALDOB ANGPTL4 ANLN ANXA2 APOF APOL6 ARHGEF39 ARID3C ASPA ASPG ASPM AASS1 AURKA AVPI1 B3GNT5 BARD1 BCAS4 BCKDHB BCO2 BDH2 BIRC5 BUB1 BUB1B C8A CAP2 CCDC34 CCL14 CCL4 CCL5 CCNA2 CCNB1 CCNB2 CCT3 CD302 CD34 CD4 CD5L CD69 CDC25C CDC45 CDC6 CDC8 CDHR2 CDK1 CDK5R1 CDKN2C CDKN3 CENPA CENPE CENPF CENPI CENPL CENPM CENPW CETP CFD CFP CHAF1B CHEK1 CHML CKAP2L CLRN3 CMYA5 CNDP1 CPS1 CST7 CTH CTHRC1 CXCL12 CYP2A6 CYP2A7 CYP2B6 CYP2C18 CYP2C8 CYP2C9 CYP39A1 CYP4A11 CYP4A22 DBH DEPDC1 DEPDC7 DNASE1L3 DSCC1 DSN1 DTYMK DUSP1 ECM1 EGR1 EHD3 ENAH EPHX2 ERRFI1 ESCO2 ESM1 ESR1 EXOC3L4 F9 FABP5 FAM111B FAM83D FAT4 FBP1 FCGR2B FCN3 FETUB FLVCR1 FOS FYB2 GABARAPL1 GADD45B GCGR GCH1 GHR GINS4 GJC1 GLS2 GLYAT GNAZ GNMT GPT1 GREM2 HAAO HAMP HAO2 HGF HSD17B13 IGF2BP1 IL33 KBTBD11 KCNK5 KIF14 KIF18A KIF20A KIF23 KIF2C KIFC1 KLKB1 KLRD1 KPNA2 LAG3 LAPTM4B LCAT LCN2 LIFR MAF2L1 MASP1 MASP2 MAT1A MCM10 MEALK MFPAP4 MFSD2A MKI67 MMS22L MND1 MT1E MT1F MT1G MT1M MT1X MT2A MYCT1 MYO10 N4BP2L1 NAT1 NAT2 NCAPG NDC80 NDRG2 NEK2 NKG7 NMB NNMT NOCT NPY1R NUF2 NUSAP1 OAT OIT3 OTUB2 OXER1 PAIP2B PAMR1 PARPBP PBK PCLAF PCOLCE PDLM5 PEX11G PGLYRP2 PHPT1 PLG PLGLB1 PLGLB2 PLVAP PON3 PRC1 PRODH PSMC3 P PTGIR PTP4A3 PTTG1 RAB3B RAD51 RAD51AP1 RDH16 RETREG1 RFC4 RIDA RNASE4 RNASEH2A RND3 RNF152 RRM2 SERPINA10 SERPINA11 SERPINA4 SERPINA5 SERPINI1 SFN SGO1 SHCBP1 SKA1 SKA3 SLC10A1 SLC22A1 SLC22A10 SLC25A15 SLC38A2 SLC7A2 SMAD6 SMG5 SMIM24 SNX22 SOC2 SORBS2 SORL1 SPINK1 SPP2 SPRYD4 SPSB2 SQLE SRD5A1 SRD5A2 SRPX SSR2 STEAP3 STEAP4 STIL SUCO SULF2 SULT1E1 TACSTD2 TAT TDRKH TESMIN THBS1 TIAM1 TKT TMEM100 TOP2A TPM2 TPX2 TRIM45 TRIP13 TTC36 TTC9 TTK UBE2C UBE2T UROC1 WDR62 WDR76 XDHD ZFP1 ZG16 ZKSCAN3 ZNF607 ZWINT

## Figures

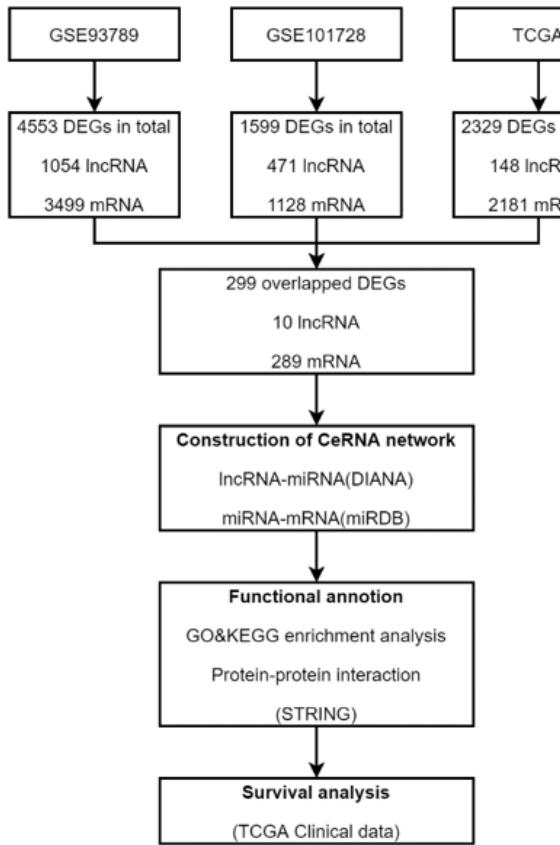
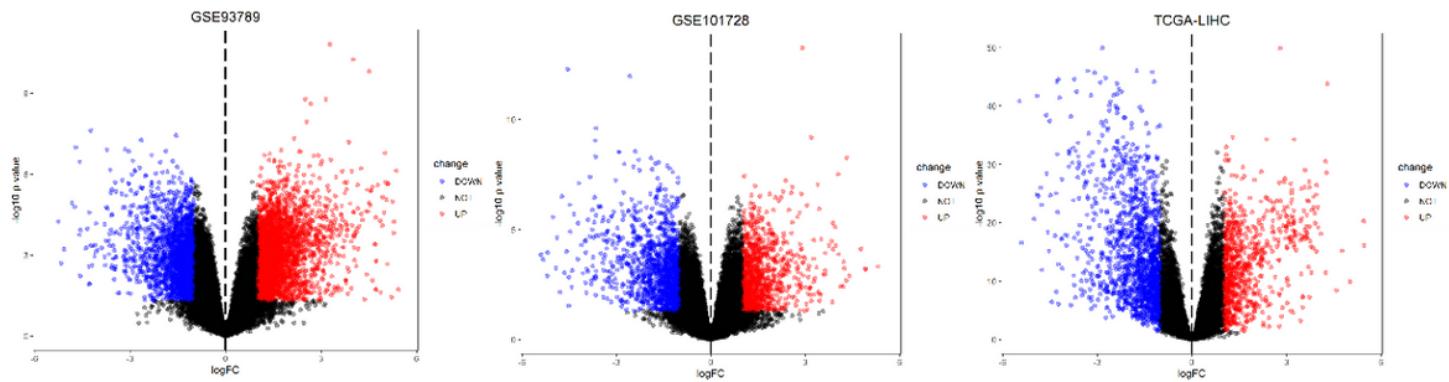


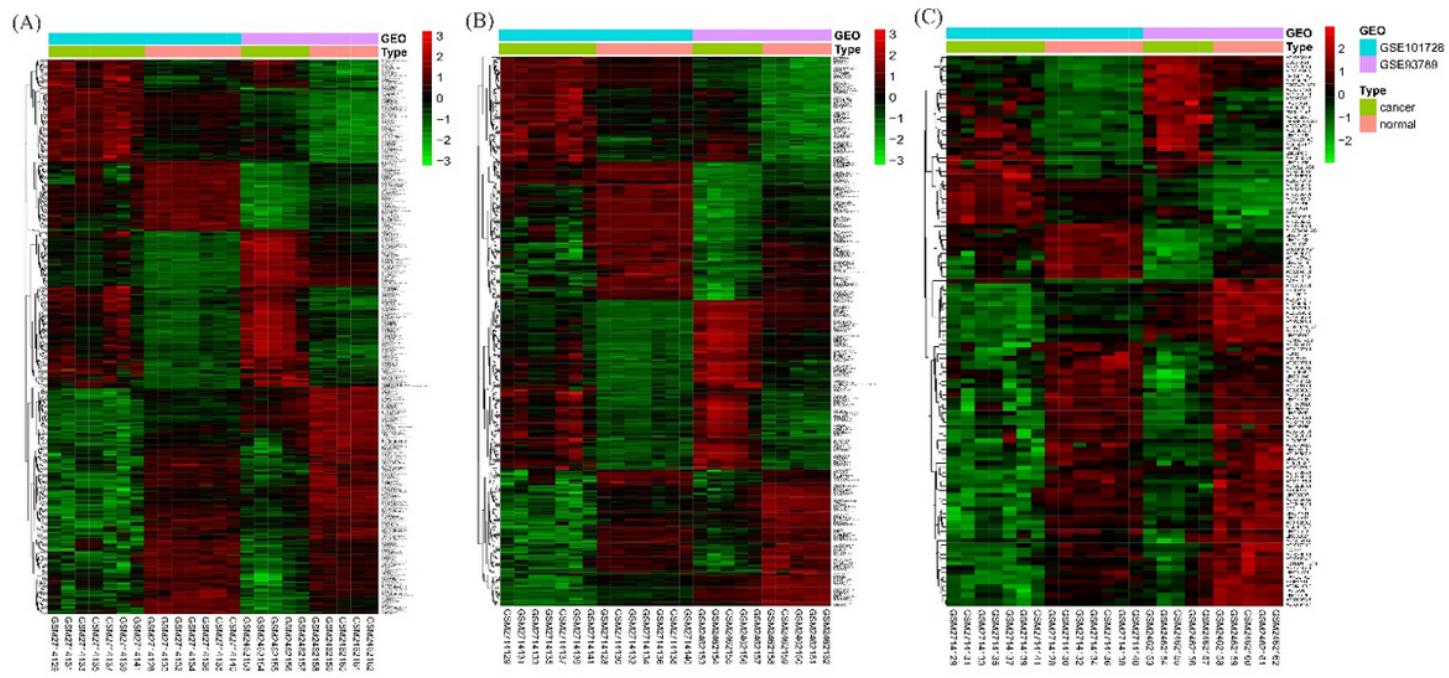
Figure 1

The flow chart shows the process of how to analyze the gene expression data from GEO and TCGA databases. lncRNA, long noncoding RNA; miRDB, microRNA database; miRNA, microRNA; mRNA, messenger RNA; TCGA, The Cancer Genome Atlas



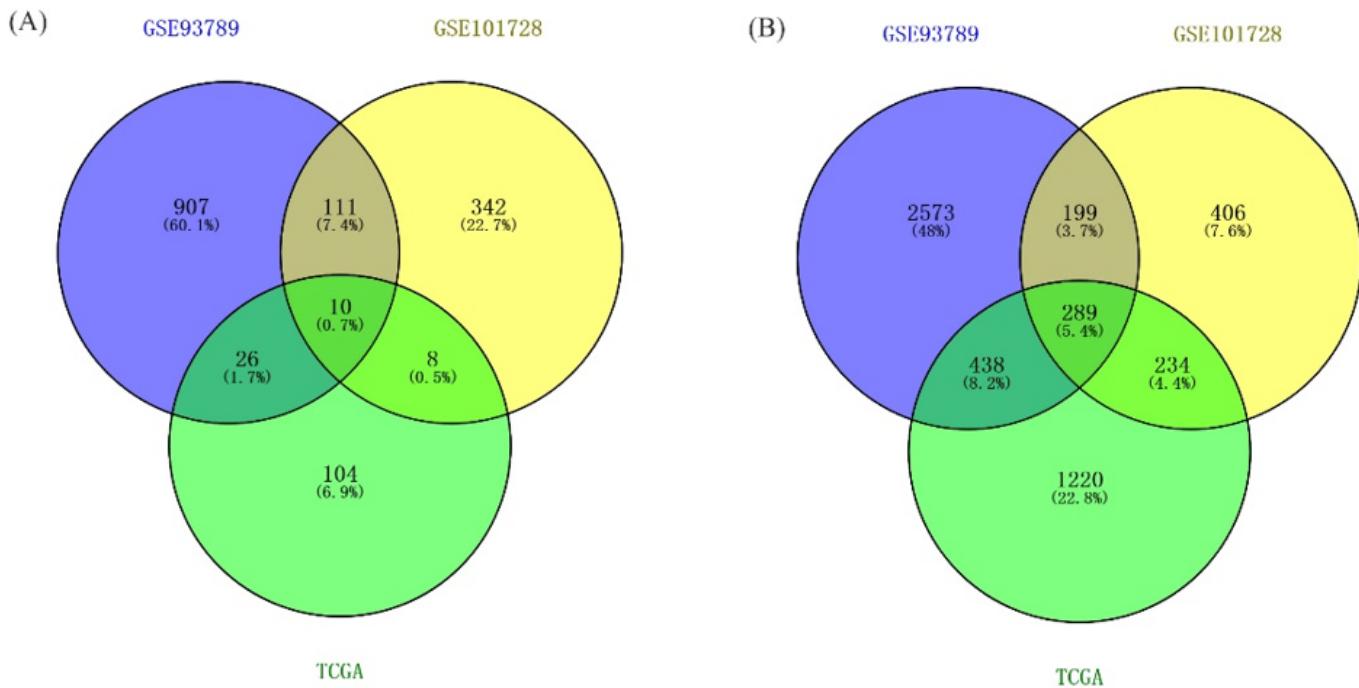
**Figure 2**

Volcano plots of DEGs in GSE93789, GSE101728 and TCGA database. The red dots and the blue dots represent the up-regulated and down-regulated genes respectively



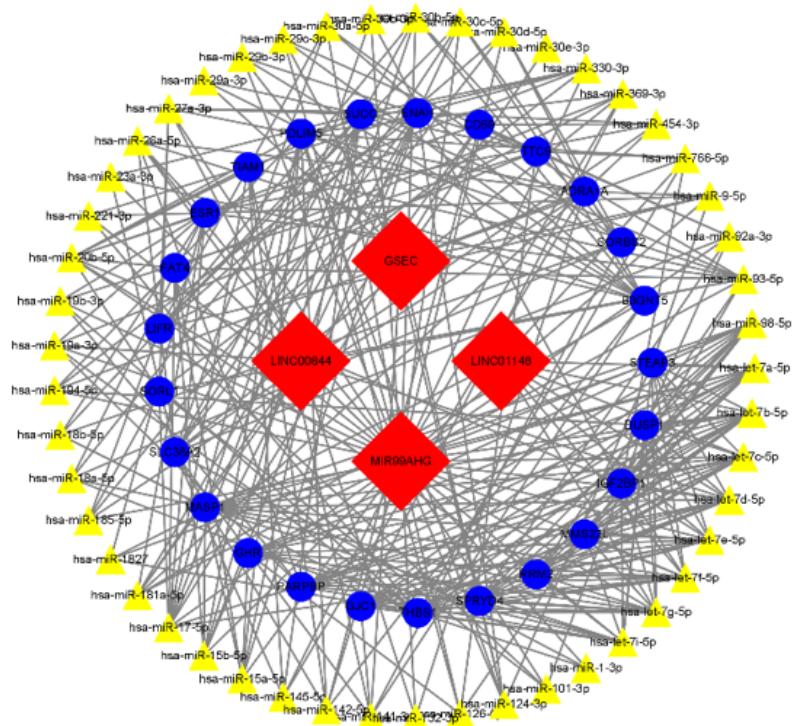
**Figure 3**

Heatmap of DEGs in GSE93789 and GSE101728 data set. A, total DEGs. B, lncRNAs. C, mRNAs. The red color and the green color represent the up-regulated and down-regulated genes respectively



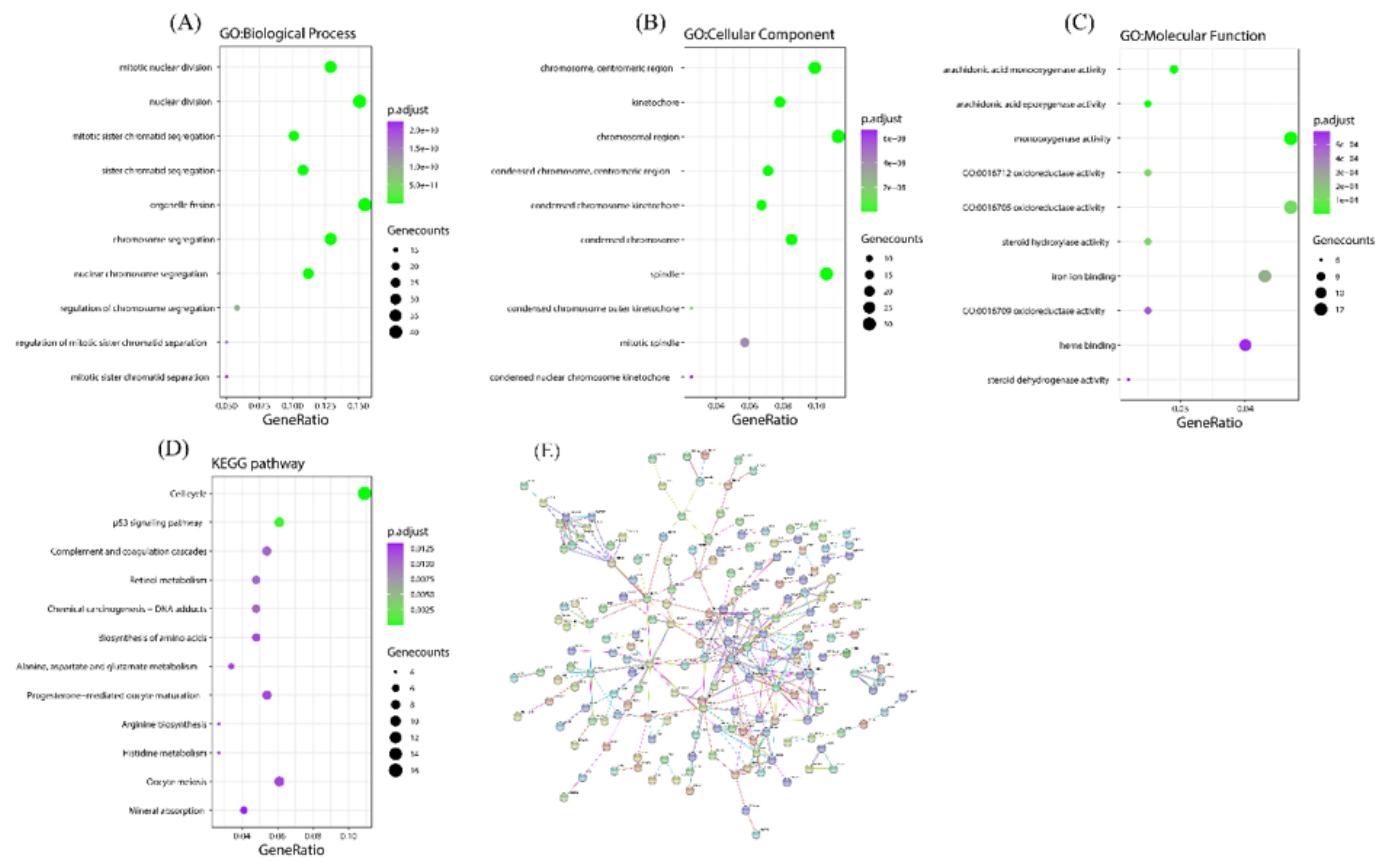
**Figure 4**

Venny diagram shows the intersected genes both in GEO and TCGA. A, intersected lncRNAs; B, intersected mRNAs



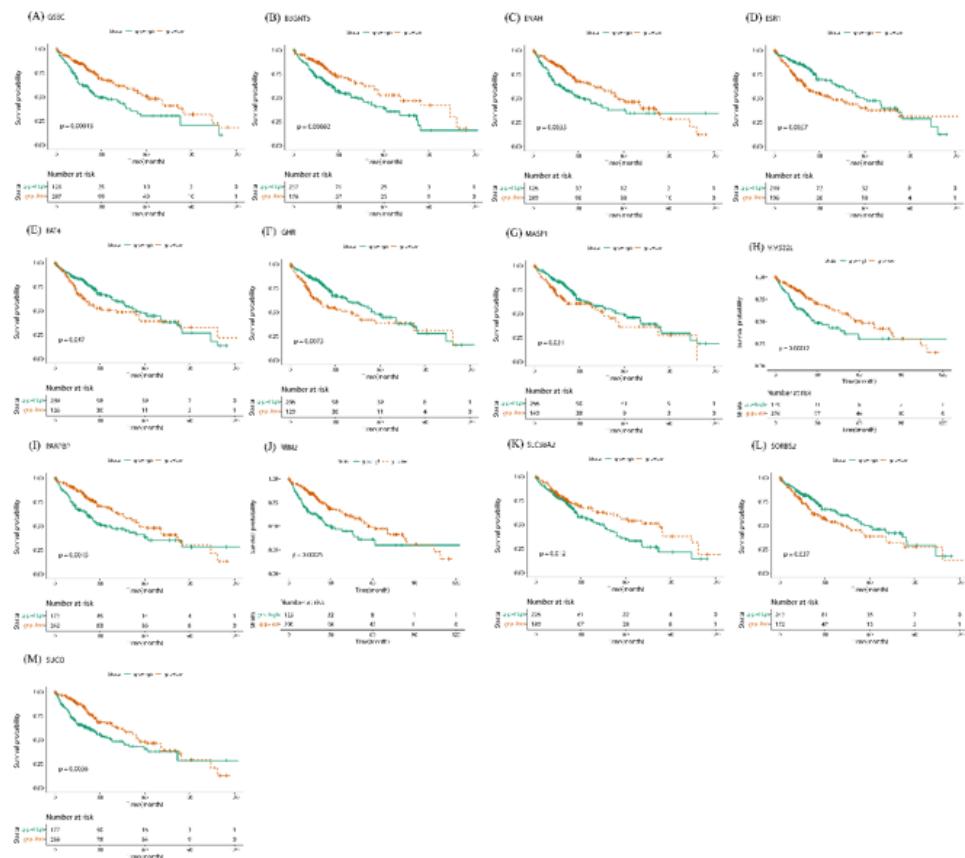
**Figure 5**

mRNA-IncRNA-miRNA network. Red color means IncRNAs; yellow color means miRNAs; blue color means mRNAs



**Figure 6**

Functional analysis of DEGs A: GO-biological processes, B: GO-cellular component, C: GO-molecular functions. D: KEGG pathway analysis. E: PPI network



**Figure 7**

The results of KM analysis of hub genes. We found 13 hub genes were related to the OS of HCC. A, GSEC B, B3GNT5 C, ENAH D, ESR1 E, FAT4 F, GHR G, MASP1 H, MMS22L I, PARPBP J, RRM2 K, SLC38A2 L, SORBS2 M, SUCO