

# Identification of prognostic factors for intrahepatic cholangiocarcinoma based on long non-coding RNAs-associated ceRNA network

**Zhichen Kang**

the Second Hospital of Jilin University

**Lixin Guo**

the Second Hospital of Jilin University

**Zhuo Zhu**

the Second Hospital of Jilin University

**Rongfeng Qu** (✉ [qurongfeng\\_qrf@yeah.net](mailto:qurongfeng_qrf@yeah.net))

The Second Hospital of Jilin University <https://orcid.org/0000-0002-1224-7452>

---

## Primary research

**Keywords:** Intrahepatic cholangiocarcinoma, Prognosis, Long non coding RNAs, competing endogenous RNA network, Sequence Read Archive, Cancer Genome Atlas

**Posted Date:** February 21st, 2020

**DOI:** <https://doi.org/10.21203/rs.2.24150/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on July 16th, 2020. See the published version at <https://doi.org/10.1186/s12935-020-01388-4>.

# Abstract

**Background** Accumulating evidence has highlighted the important roles of long non-coding RNAs (lncRNAs) acting as competing endogenous RNAs (ceRNAs) in tumor biology. However, the roles of cancer long non coding RNAs (lncRNAs) in lncRNA-related ceRNA network of intrahepatic cholangiocarcinoma (ICC) remain enigmatic. The current study aims to identify prognostic factors in the lncRNA-related ceRNA network of ICC.

**Methods** The transcriptome sequencing data of the lncRNAs, messenger RNA (mRNA) and microRNA (miR) were downloaded from SRA, and TCGA database respectively. Differentially expressed lncRNAs (DELncRNAs), DEmiRs and DEmRNAs were identified and adopted to construct lncRNA-miR-mRNA ceRNA network. ICC-associated DEmRNAs were adopted to construct the PPI network. The expression of the top 6 genes in the hub module was validated in mRNA transcriptome sequencing data and ICC-related gene expression dataset GSE45001, followed by GO and KEGG pathway enrichment analysis. The potential function of genes in hub module in the ceRNA network was finally subjected to GSEA.

**Results** Sixty coexpression DEmRNAs were identified in the ceRNA network. The 6 top genes in the hub module consists of FOS, HGF, IGF2, FOXO1, NTF3 (downregulated), and IGF1R (upregulated) in ICC tissues compared to normal adjacent tissues, followed by the successful construction of lncRNA-miR-hub network with 86 ceRNA modules. FOS, IGF2 or IGF1R might regulate the development of ICC by targeting “N-glycan biosynthesis”, “ $\alpha$ -linolenic acid metabolism” or “type II diabetes” pathways respectively.

**Conclusion** These results show FOS, IGF2 and IGF1R serve as prognostic factors of ICC patients.

## Background

Cholangiocarcinoma, an aggressive malignancy frequently occurs at the biliary tract with unfavorable prognosis [1]. Intrahepatic cholangiocarcinoma (ICC), is rarely seen as the least common subtype of cholangiocarcinoma and arises from the epithelial cells of the intrahepatic bile ducts [2]. Dominant risk factors for the pathogenesis of ICC consist of cirrhosis, chronic hepatitis B and C, alcohol use, diabetes, or even obesity [3]. The early combination of multiple means of treatment has been proposed to be beneficial to aggressive variant of ICC [4]. However, the later time of diagnosis is likely to cause locoregional extension of ICC, resulting in dismal prognosis [5]. Identifying prognostic factors for ICC is therefore critical to effective treatments for ICC.

The methylation of DLEC1 cilia and flagella associated protein is engaged in a favorable clinical outcome and prognosis in patients with small duct ICC [6]. However, CD90 expression contributed to lymph node metastasis and thereby dismal prognosis in human ICC [7]. Expression of VEGFR-3 in ICC also promotes the angiogenesis of lymph, thereby resulting in unfavorable prognosis [8]. Long non-coding RNAs (lncRNAs) are a multiple class of RNAs engaging in various biological processes [9]. LncRNAs also serve as both oncogene and suppression, making them target for tumorigenesis [10]. Competing endogenous RNA (ceRNA) is a newly-emerged regulatory network on a large scale across the transcriptome with

expansion to the genetically functional information in the human genome and exerts pivotal functions on cancer [11]. Interestingly, ceRNAs indicate a novel modulation on lncRNAs and genes that exert pivotal functions on cancer pathogenesis by binding microRNAs (miRs) in cholangiocarcinoma [12].

In the current study, we adopted the transcriptome sequencing data of the lncRNAs, messenger RNA (mRNA) and microRNA (miR) from Sequence Read Archive (SRA), and Cancer Genome Atlas (TCGA) database respectively. Differentially expressed lncRNAs (DELncRNAs), DEmiRs and DEmRNAs were identified and adopted to construct lncRNA-miR-mRNA ceRNA network. ICC-associated DEmRNAs were adopted to construct the PPI network. The expression of identified the top 6 genes in the hub module by PPI network, followed by GO and KEGG pathway enrichment analysis. The potential function of genes in hub module in the ceRNA network was finally subjected to Gene Set Enrichment Analysis (GSEA). Thus, the purpose of this study was to identify genetic alterations that underlie the prognostic factors of ICC.

## Methods

### Data collection and preprocessing

The original transcriptome sequencing data of the lncRNAs from human ICC and adjacent normal tissues were retrieved until September 2019 from SRA database (<https://www.ncbi.nlm.nih.gov/sra/>) and then subjected to SRP126672 dataset. SRP126672 was composed of 30 ICC tissues and 27 adjacent normal tissues. Fastqc and trimmomatic applications were adopted to control quality and filter the original sequencing data. lncRNAs were quantified based on the Genome Research Project of ENCYCLOPEDIA OF DNA ELEMENTS (GENCODE) (GRCh37) catalog (<http://www.genencodegenes.org/>). The RNA transcriptome sequencing data for ICC were downloaded from the Cancer Genome Atlas (TCGA) database. MiR sequencing and mRNA sequencing data were downloaded using a data transfer management tool (provided by GDC Apps) (<https://tcga-data.nci.nih.gov/>). Among them, miR sequencing data and mRNA sequencing data consisted of 33 ICC tissues and 8 normal tissues respectively. In addition, the gene expression GSE26566 dataset for ICC downloaded from the Gene Expression Omnibus (GEO) database included 10 samples of ICC and 10 non-tumor liver samples to test the expression level of the hub gene.

### Identification of DEGs

The original data of ICC tissues and adjacent normal tissues RNA-seq were corrected, normalized and their expression was calculated. Then, differentially expressed lncRNAs (DELncRNAs) were screened by the DESeq2 package. The standard is  $|\log_2(\text{fold change [FC]})| > 2$  and adjusted  $p < 0.05$ . In addition, the edgeR software package was used to screen differentially expressed miRs (DEmiRs) and mRNAs (DEmRNAs) with thresholds  $|\log_2(\text{fold change [FC]})| > 2$  and  $\text{FDR} < 0.01$ .

### Construction of lncRNA-miR-mRNA ceRNA network

The ceRNA network was constructed based on the DELncRNAs, DEmiRs and DEmRNAs. Briefly, following the identification of DELncRNAs, DEmiRs and DEmRNAs, the miRcode (<http://www.mircode.org/>) database was adopted to predict the markedly downregulated lncRNAs targeted by upregulated miRs and

markedly elevated lncRNAs targeted by repressed miRs in ICC. DEGs with correct trend and targeted relationships serve as candidate genes. Next, the TargetScan (<http://www.targetscan.org/>), miRDB (<http://www.mirdb.org/>) and miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw/php/index.php>) were adopted to predict markedly downregulated mRNAs targeted by upregulated miRs and markedly elevated mRNAs targeted by repressed miRs in ICC. The predicted mRNAs were selected by intersection among these databases. DEGs with correct trend and targeted relationships serve as candidate genes. The predicted lncRNA-miR and miR-mRNA was combined to construct the lncRNA-miR-mRNA ceRNA network. Finally, cytoscape v3.6.1 software was adopted to visualize and map the network.

## **GO and KEGG pathway enrichment analysis**

To clarify the potential biological processes of DEmRNAs related to the ceRNA network in the development of ICC, the DAVID database (<https://david.ncifcrf.gov/>) was used to perform GO enrichment analysis of DEmRNAs by setting the default parameters. The GO function enriched by  $p < 0.05$  was considered significant among transcriptome sequencing data. In order to understand the potential pathways of DEmRNAs involved in the ceRNA network, the KOBAS database (<http://kobas.cbi.pku.edu.cn/index.php>) was used to perform KEGG pathway enrichment analysis on DEmRNAs and the significance of KEGG pathway was evaluated at  $p < 0.001$ .

## **Construction of the protein-protein interaction (PPI) network and module analysis**

The interaction between DEGs identifies key genes in modules involved in the development of ICC with a combined score of  $> 0.4$  for a PPI network as the threshold. Then, the PPI information of DEmRNAs was obtained from STRING database (<http://www.string-db.org/>) and a PPI network was built using Cytoscape. Finally, the first six key genes in the hub module were obtained from the PPI network using the MCC network topology using the cytoHubba plug-in in Cytoscape.

## **Validation of key genes in the hub module**

The gene expression dataset GSE45001 of ICC from the Gene Expression Omnibus (GEO) database was used to verify the expression of key genes in the hub module. The gene expression dataset GSE26566 was also used hub gene function prediction.

## **Prediction of gene function in the hub module**

To better elucidate the potential function of the key gene in hub modules in the ceRNA network, a GSEA was performed. Briefly, according to the median expression of key genes in the RNA sequencing data, 33 ICC samples from the TCGA database were divided into the high expression group and low expression group. The reference gene set is the annotated c2.cp.kegg.v6.2.symbols.gmt gene set in the Molecular Signature Database (MSigDB). The critical criterion is  $p < 0.05$ .

## **Results**

# Identification of DEGs between ICC tissues and adjacent normal tissues

To identify DEmRNAs, DEmiRs and DElncRNAs in ICC, the SRP126672 was composed of 30 ICC tissues and 27 adjacent normal tissues downloaded from SRA database, followed by Fastqc and trimmomatic applications and then DElncRNAs were screened using the DESeq2 package. A total of 912 DElncRNAs were obtained, with 524 upregulated and 388 downregulated DElncRNAs (Fig. 1A,B). The different expression of DElncRNAs could help distinguish ICC tissues and 27 adjacent normal tissues determined by principal component analysis (Fig. 1C). DEmiR and DEmRNA sequencing data (counts) of ICC and adjacent normal tissues were obtained from the TCGA database, followed by the edgeR package to perform differential analysis. Finally, 66 DEmiRs (38 upregulated and 28 downregulated) (Fig. 1D,E) and 5,522 DEmRNAs (3158 upregulated and 2364 downregulated) (Fig. 1F,G) were obtained.

## The lncRNA-miR-mRNA ceRNA network in ICC

To better understand the role of lncRNAs and miRs in the ceRNA network of ICC tissues, a lncRNA-miR-mRNA ceRNA network was established. Initially, 66 DElncRNAs targeted by with 66 DEmiRs, and 31,452 lncRNA-miR pairs were predicted. DEGs with correct trend and targeted relationships serve as candidate genes. The selected candidate genes contain 66 lncRNA-miR pairs including 18 markedly up-regulated DElncRNAs, 19 down-regulated DElncRNAs, 10 significantly up-regulated miRs, and 2 significantly down-regulated miRs.

In addition, 16457, 7052, and 2081 target mRNAs were predicted from the databases TargetScan, miRDB, and miRTarBase respectively. Then, 10 downregulated candidate DEmiRs were intersected with 2364 downregulated DEmRNAs, yielding 43 shared genes (Fig. 2A), and 2 downregulated candidate DEmiRs were intersected with 3158 upregulated DEmRNAs, yielding 17 share genes (Fig. 2B). Next, the shared genes serve as candidate genes and generated 136 miR-mRNA pairs including 43 significantly down-regulated and 17 significantly up-regulated DEmRNAs. A total of 37 DElncRNAs, 12 DEmiRs, and 60 DEmRNAs were used to construct the ceRNA network. Based on these lncRNA-miR pairs and miR-mRNA pairs, a ceRNA network consisting of 37 lncRNA nodes, 12 miR nodes, and 60 mRNA nodes in ICC was constructed (Fig. 2C).

## GO and KEGG pathway enrichment analysis on DEmRNAs

To identify the biological functions and pathways of the 60 DEmRNAs in the ceRNA network, we used the DAVID database for GO function enrichment analysis and the KOBAS database for KEGG pathway enrichment analysis. GO enrichment analysis showed that DEmRNAs related to biological processes were mainly enriched in GO terms including the insulin receptor signaling pathway, positive regulation of cell proliferation, cell respiration and other items ( $p < 0.05$ ). DEmRNAs related to cellular component were most closely related to mitochondrial inner membrane ( $p < 0.05$ ). DEmRNAs related to molecular function were mainly enriched in GO terms of chemoattractant activity, growth factor activity, and transcriptional

coactivator binding ( $p < 0.05$ ) (Fig. 3A). KEGG analysis showed that DEmRNAs were significantly enriched in pathway such as cancer pathways, calcium signaling pathways, cytokin-cytokine receptor interaction, MAPK signaling pathway and proteoglycans in cancer ( $p < 0.001$ ) (Fig. 3B). Coherently, 60 significantly different DEmRNAs exert pivotal role in the occurrence and development of ICC.

## Identification of 6 genes in the hub module in PPI network

PPI network was constructed based on DEmRNAs, involving 60 nodes and 41 edges (Fig. 4A). To identify the first six genes in the hub module in the PPI network, the MCC network topology in the cytoHubba plugin was used to determine the intimacy of DEmRNAs. Fos proto-oncogene, AP-1 transcription factor subunit (FOS), insulin like growth factor 1 receptor (IGF1R), hepatocyte growth factor (HGF), insulin like growth factor 2 (IGF2), forkhead box O1 (FOXO1), and neurotrophin 3 (NTF3) in the hub module were obtained (Fig. 4B). Finally, the lncRNA-miR-hub gene subnetwork was constructed (Fig. 4C), including 86 ceRNA regulatory modules.

## Validation of FOS, HGF, IGF2, FOXO1, NTF3 and IGF1R expression in ICC

To better validate our analysis, the expression levels of 6 genes in the hub module were in 33 ICC tissues and 8 adjacent normal tissues. Concordant with our previous analysis, results (Fig. 5A) demonstrated that FOS, HGF, IGF2, FOXO1, and NTF3 remarkably downregulated while IGF1R potentially rose in ICC tissues ( $\log_{2}FC > 2$ ,  $FDR < 0.01$ ) compared to adjacent normal tissues. FOS, IGF2, and IGF1R were randomly selected and subjected to the gene expression dataset GSE45001 of ICC, and results exhibited the same trend (Fig. 5B-D). Taken together, these results validated our prediction.

## GSEA prediction of biological pathways related to FOS, IGF2, and IGF1R

To better identify the biological pathways related to FOS, IGF2, and IGF1R, we classified 33 IHC samples from the TCGA database into the high expression group and low expression group. The reference gene set is the annotated c2.cp.kegg.v6.2.symbols.gmt gene set in the MSigDB. Results revealed that the ICC tissues in the high expression group of FOS, IGF2 or IGF1R were markedly enriched in “N-glycan biosynthesis”, “ $\alpha$ -linolenic acid metabolism” or “type II diabetes” respectively (all  $p < 0.01$ ) (Fig. 6A-C), suggesting the regulatory roles of FOS, IGF2, and IGF1R in ICC may be exerted by these pathways.

## Discussion

The present study is the preliminary research of the prognostic factors for ICC patients. To identify and gain more insight into the molecules involved in the prognosis of ICC, we analyzed the transcriptome sequencing data of the lncRNAs, mRNA and microRNA miR to construct a lncRNA-miR-mRNA ceRNA network, where which identified 60 coexpression DEmRNAs associated with ICC. Main findings in the

current study displayed that the expression of FOS, HGF, IGF2, FOXO1, NTF3 were downregulated, but IGF1R expression was upregulated in ICC tissues compared with normal adjacent tissues. In addition, FOS, IGF2 or IGF1R might regulate the development of ICC by targeting “N-glycan biosynthesis”, “ $\alpha$ -linolenic acid metabolism” or “type II diabetes” pathways respectively.

FOS proteins were characterized by a leucine zipper and a basic region with a helix-turn-helix motif that binds DNA, and serve as an oncogene and transcription factors by binding DNA sequences [13]. C-terminally truncated FOS mutant has been previously reported to be expressed in epithelioid hemangioma [14]. The expression of FOS has been proposed to be positively related to c-Myb expression in colorectal cancer cells, while c-Myb expression is repressed in colorectal cancer tissues, suggesting that expression of FOS is also downregulated in colorectal cancer tissues [15], which is concordant with the current study. Meanwhile, the elevation of FOS-like antigen 1 expression is positively correlated with progression of perihilar cholangiocarcinoma [16]. N-glycan could determine the thermodynamic stability of proteins and exerts pivotal functions on various key biological processes [17]. Interestingly, the repressed genes in gemcitabine resistance-related of ICC belong to the N-glycan biosynthesis pathway [18], though the targeting relationship of FOS and N-glycan biosynthesis pathway has been scarcely documented.

IGFs (IGF1 and IGF2) expedite glucose metabolism with their availability modulated by IGF binding proteins, and function as be prognostic factors for type 1 diabetes [19]. The induction of IGF2 is also partially involved with the proliferation and survival of rhabdomyosarcoma cells [20]. Interestingly, IGF2 has been reported to be methylated in ICC compared to extrahepatic cholangiocarcinoma [21]. IGF-1 has been revealed to be associated with  $\alpha$ -linoleic acid in and exert function on porcine primary hepatocytes [22].  $\alpha$ -linolenic acid is a high unsaturated fatty acid content from tree peony species (woody oil crops) [23].  $\alpha$ -linolenic acid metabolism pathway is also engaged in liver cirrhosis [24].

IGF1 maintains the phenotype of tumor and survival of transformed murine pheochromocytoma cells [25]. The inhibition of IR/IGF1R reduced epithelial-mesenchymal transition and cancer stem cell-like traits in resistant cells of cholangiocarcinoma [26]. The elevated expression of IGF1R was observed in tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-resistant gastric cancer cells, enhancing TRAIL resistance in gastric cancer cells [27]. As genes encoding proteins related to insulin receptor, IGF1R could stimulate renal cancer cells [28]. Type II diabetes mellitus could reduce the risk of cholangiocarcinoma in patients with biliary tract diseases[29].

## Conclusion

In conclusion, our research identified several novel genetic alterations and pathway associated with prognosis of ICC, and better understood the potential role of miRs, lncRNAs and mRNAs in the ICC development via bioinformatic analysis. Based on the ceRNA network, we also found that FOS, IGF2 or IGF1R might target “N-glycan biosynthesis”, “ $\alpha$ -linolenic acid metabolism” or “type II diabetes” pathways respectively, thereby modulating the development of ICC by targeting, which provided information for understanding molecules involved in the prognosis of ICC, and even contributed to the treatment for ICC.

However, further studies with selective candidate genes and larger sample size are required to define their role in ICC. Additionally, more studies need to be performed to examine the underlying mechanisms of FOS and N-glycan biosynthesis pathway.

## Abbreviations

ceRNAs: competing endogenous RNAs; lncRNAs: long non-coding RNAs; ICC: intrahepatic cholangiocarcinoma; DElncRNAs: Differentially expressed lncRNAs; GENCODE: Genome Research Project of Encyclopedia of DNA Elements; TCGA: the Cancer Genome Atlas; GEO: Gene Expression Omnibus.

## Declarations

### Authors' contributions

RF Q and ZC K designed the study. LX G and Z Z collated the data, carried out data analyses and produced the initial draft of the manuscript. RF Q and ZC K contributed to drafting the manuscript. All authors have read and approved the final submitted manuscript.

### Author details

<sup>1</sup> Department of Rehabilitation, the Second Hospital of Jilin University, Changchun 130022, P.R. China. <sup>2</sup> Department of Anesthesiology, the Second Hospital of Jilin University, Changchun 130022, P.R. China. <sup>3</sup> Department of Hematology and Oncology, the Second Hospital of Jilin University, Changchun 130022, P.R. China

### Acknowledgments

The authors would like to acknowledge the helpful comments on this paper received from the reviewers.

### Competing interests

The authors declare that they have no competing interests.

### Availability of data and material

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

## Consent for publication

Not applicable.

## Funding

None.

## References

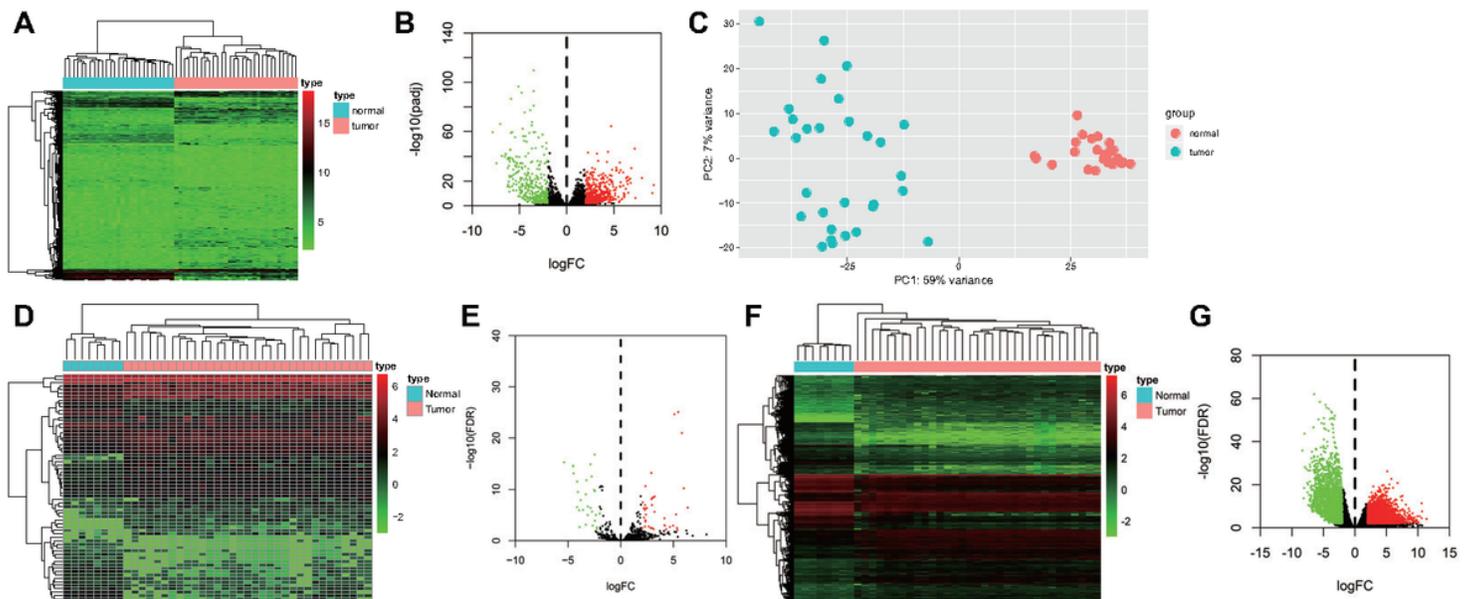
1. Mahipal A, Tella SH, Kommalapati A, Anaya D, Kim R. FGFR2 genomic aberrations: Achilles heel in the management of advanced cholangiocarcinoma. *Cancer Treat Rev.* 2019;78:1-7.
2. Zhang H, Yang T, Wu M, Shen F. Intrahepatic cholangiocarcinoma: Epidemiology, risk factors, diagnosis and surgical management. *Cancer Lett.* 2016;379:198-205.
3. Palmer WC, Patel T. Are common factors involved in the pathogenesis of primary liver cancers? A meta-analysis of risk factors for intrahepatic cholangiocarcinoma. *J Hepatol.* 2012;57:69-76.
4. Moro A, Paredes AZ, Farooq A, Sahara K, Tsilimigras DI, Mehta R, Endo I, Guglielmi A, Aldrighetti L, Alexandrescu S, Marques HP, Shen F, Koerkamp BG, Sasaki K, Pawlik TM and other members of the International Intrahepatic Cholangiocarcinoma Study G. Discordance in prediction of prognosis among patients with intrahepatic cholangiocarcinoma: A preoperative vs postoperative perspective. *J Surg Oncol.* 2019;120:946-55.
5. Le Roy B, Gelli M, Pittau G, Allard MA, Pereira B, Serji B, Vibert E, Castaing D, Adam R, Cherqui D, Sa Cunha A. Neoadjuvant chemotherapy for initially unresectable intrahepatic cholangiocarcinoma. *Br J Surg.* 2018;105:839-47.
6. Kim Y, Lee K, Jeong S, Wen X, Cho NY, Kang GH. DLEC1 methylation is associated with a better clinical outcome in patients with intrahepatic cholangiocarcinoma of the small duct subtype. *Virchows Arch.* 2019;475:49-58.
7. Yamaoka R, Ishii T, Kawai T, Yasuchika K, Miyauchi Y, Kojima H, Katayama H, Ogiso S, Fukumitsu K, Uemoto S. CD90 expression in human intrahepatic cholangiocarcinoma is associated with lymph node metastasis and poor prognosis. *J Surg Oncol.* 2018;118:664-74.
8. Sha M, Jeong S, Chen XS, Tong Y, Cao J, Sun HY, Xia L, Xu N, Wang X, Han LZ, Xi ZF, Zhang JJ, Kong XN, Xia Q. Expression of VEGFR-3 in intrahepatic cholangiocarcinoma correlates with unfavorable prognosis through lymphangiogenesis. *Int J Biol Sci.* 2018;14:1333-42.
9. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet.* 2016;17:47-62.

10. Slaby O, Laga R, Sedlacek O. Therapeutic targeting of non-coding RNAs in cancer. *Biochem J.* 2017;474:4219-51.
11. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* 2011;146:353-8.
12. Wan M, Zhang FM, Li ZL, Kang PC, Jiang PM, Wang YM, Wang ZD, Zhong XY, Li CL, Wang H, Zhao SY, Cui YF. Identifying survival-associated ceRNA clusters in cholangiocarcinoma. *Oncol Rep.* 2016;36:1542-50.
13. Alfonso-Gonzalez C, Riesgo-Escovar JR. Fos metamorphoses: Lessons from mutants in model organisms. *Mech Dev.* 2018;154:73-81.
14. van IDGP, Forghany Z, Liebelt F, Vertegaal AC, Jochemsen AG, Bovee J, Szuhai K, Baker DA. Functional analyses of a human vascular tumor FOS variant identify a novel degradation mechanism and a link to tumorigenesis. *J Biol Chem.* 2017;292:21282-90.
15. Qu X, Yan X, Kong C, Zhu Y, Li H, Pan D, Zhang X, Liu Y, Yin F, Qin H. c-Myb promotes growth and metastasis of colorectal cancer through c-fos-induced epithelial-mesenchymal transition. *Cancer Sci.* 2019;110:3183-96.
16. Liu Z, Sun R, Zhang X, Qiu B, Chen T, Li Z, Xu Y, Zhang Z. Transcription factor 7 promotes the progression of perihilar cholangiocarcinoma by inducing the transcription of c-Myc and FOS-like antigen 1. *EBioMedicine.* 2019;45:181-91.
17. Arigoni-Affolter I, Scibona E, Lin CW, Bruhlmann D, Souquet J, Broly H, Aebi M. Mechanistic reconstruction of glycoprotein secretion through monitoring of intracellular N-glycan processing. *Sci Adv.* 2019;5:eaax8930.
18. Sato J, Kimura T, Saito T, Anazawa T, Kenjo A, Sato Y, Tsuchiya T, Gotoh M. Gene expression analysis for predicting gemcitabine resistance in human cholangiocarcinoma. *J Hepatobiliary Pancreat Sci.* 2011;18:700-11.
19. Shapiro MR, Wasserfall CH, McGrail SM, Posgai AL, Bacher R, Muir A, Haller MJ, Schatz DA, Wesley JD, von Herrath M, Hagopian WA, Speake C, Atkinson MA, Brusko TM. Insulin-Like Growth Factor Dysregulation Both Preceding and Following Type 1 Diabetes Diagnosis. *Diabetes.* 2019.
20. Zheng Y, Xu L, Hassan M, Zhou X, Zhou Q, Rakheja D, Skapek SX. Bayesian modeling identifies PLAG1 as a key regulator of proliferation and survival in rhabdomyosarcoma cells. *Mol Cancer Res.* 2019.
21. Kim BH, Cho NY, Choi M, Lee S, Jang JJ, Kang GH. Methylation profiles of multiple CpG island loci in extrahepatic cholangiocarcinoma versus those of intrahepatic cholangiocarcinomas. *Arch Pathol Lab Med.* 2007;131:923-30.
22. Fang XL, Shu G, Zhang ZQ, Wang SB, Zhu XT, Gao P, Xi QY, Zhang YL, Jiang QY. Roles of alpha-linolenic acid on IGF-I secretion and GH/IGF system gene expression in porcine primary hepatocytes. *Mol Biol Rep.* 2012;39:10987-96.
23. Zhang Q, Yu R, Sun D, Rahman MM, Xie L, Hu J, He L, Kilaru A, Niu L, Zhang Y. Comparative Transcriptome Analysis Reveals an Efficient Mechanism of alpha-Linolenic Acid in Tree Peony

Seeds. *Int J Mol Sci.* 2018;20.

24. Yoo HJ, Jung KJ, Kim M, Kim M, Kang M, Jee SH, Choi Y, Lee JH. Liver Cirrhosis Patients Who Had Normal Liver Function Before Liver Cirrhosis Development Have the Altered Metabolic Profiles Before the Disease Occurrence Compared to Healthy Controls. *Front Physiol.* 2019;10:1421.
25. Martin A, Venara M, Matho C, Olea FD, Fernandez MC, Pennisi PA. Fibroblast deficiency of insulin-like growth factor 1 receptor type 1 (IGF1R) impairs initial steps of murine pheochromocytoma development. *Biochimie.* 2019;163:108-16.
26. Vaquero J, Lobe C, Tahraoui S, Claperon A, Mergey M, Merabtene F, Wendum D, Coulouarn C, Housset C, Desbois-Mouthon C, Praz F, Fouassier L. The IGF2/IR/IGF1R Pathway in Tumor Cells and Myofibroblasts Mediates Resistance to EGFR Inhibition in Cholangiocarcinoma. *Clin Cancer Res.* 2018;24:4282-96.
27. Guo T, Xu L, Che X, Zhang S, Li C, Wang J, Gong J, Ma R, Fan Y, Hou K, Zhou H, Hu X, Liu Y, Qu X. Formation of the IGF1R/CAV1/SRC tri-complex antagonizes TRAIL-induced apoptosis in gastric cancer cells. *Cell Biol Int.* 2017;41:749-60.
28. Solarek W, Koper M, Lewicki S, Szczylik C, Czarnecka AM. Insulin and insulin-like growth factors act as renal cell cancer intratumoral regulators. *J Cell Commun Signal.* 2019;13:381-94.
29. Tsai MS, Lee PH, Lin CL, Peng CL, Kao CH. Type II diabetes mellitus is associated with a reduced risk of cholangiocarcinoma in patients with biliary tract diseases. *Int J Cancer.* 2015;136:2409-17.

## Figures



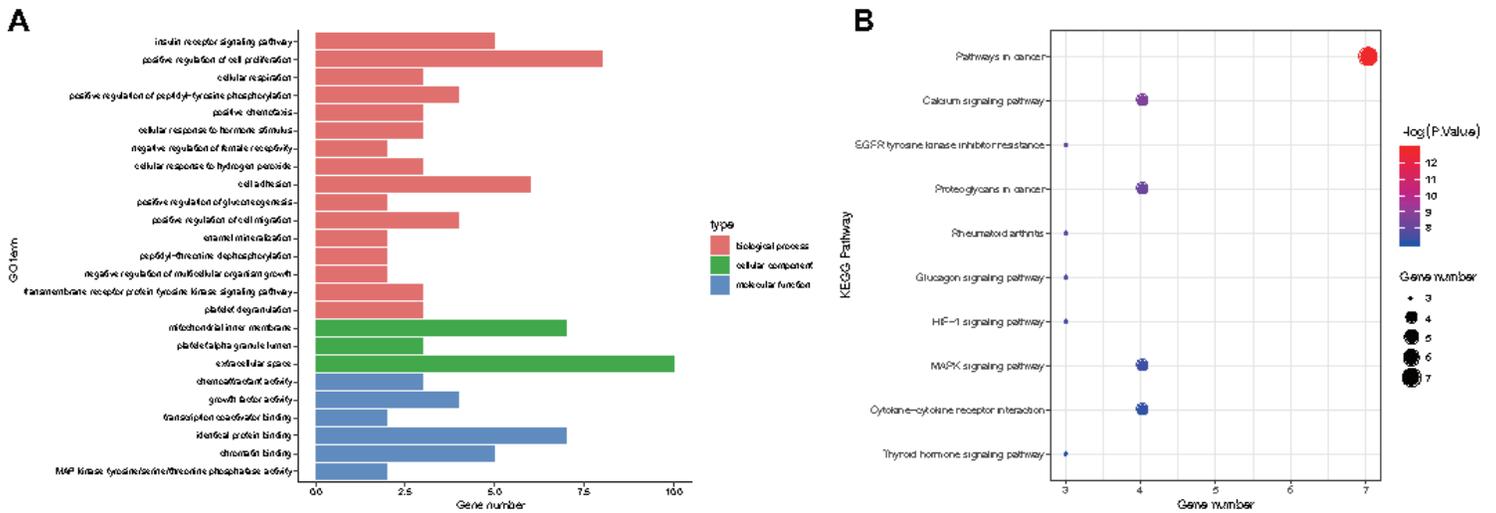
**Figure 1**

DEGs between ICC tissues and adjacent normal tissues were identified. A, The heat map of DElncRNAs. The x-axis represents the sample, and the y-axis represents a DElncRNAs. Red represents upregulated

DElncRNAs and green represents downregulated DElncRNAs. B, The volcano plot of DElncRNAs. The y-axis represents the negative base-10 logarithm of the corrected p-value, and the x-axis represents the base-2 logarithm of FC. Each point represents an DElncRNAs. Green dots indicate downregulated DElncRNAs, red dots indicate upregulated DElncRNAs, and black dots indicate lncRNAs that are not differentially expressed. C, DElncRNAs determined by principal component analysis. Adjacent normal tissues (red, n = 27) and ICC tissues (green, n = 30) were plotted along the axis around the first two main components (PC1 and PC2). D, The heat map of DEmiRs. The x-axis represents the sample, and the y-axis represents DEmiRs. Red represents upregulated DEmiRs and green represents downregulated DEmiRs. E, The volcano plot of DEmiRs. The y-axis represents the negative base-10 logarithm of FDR, and the x-axis represents the base-2 logarithm of FC, and each point represents a DEmiR. Green dots indicate downregulated DEmiRs, red dots indicate upregulated DEmiRs, and black dots indicate miRs that are not differentially expressed. F, The heat map of DEmiRs. The y-axis represents the sample, and the y-axis represents DEmiRs. Red represents upregulated DEmiRs and green represents downregulated DEmiRs. (G) The volcano plot of DEmRNAs. The y-axis represents the negative base-10 logarithm of FDR, the x-axis represents the base-2 logarithm of FC. Each dot represents a DEmRNA. Green dots indicate downregulated DEmRNAs, red dots indicate upregulated DEmRNAs, and black dots indicate mRNAs that are not differentially expressed.

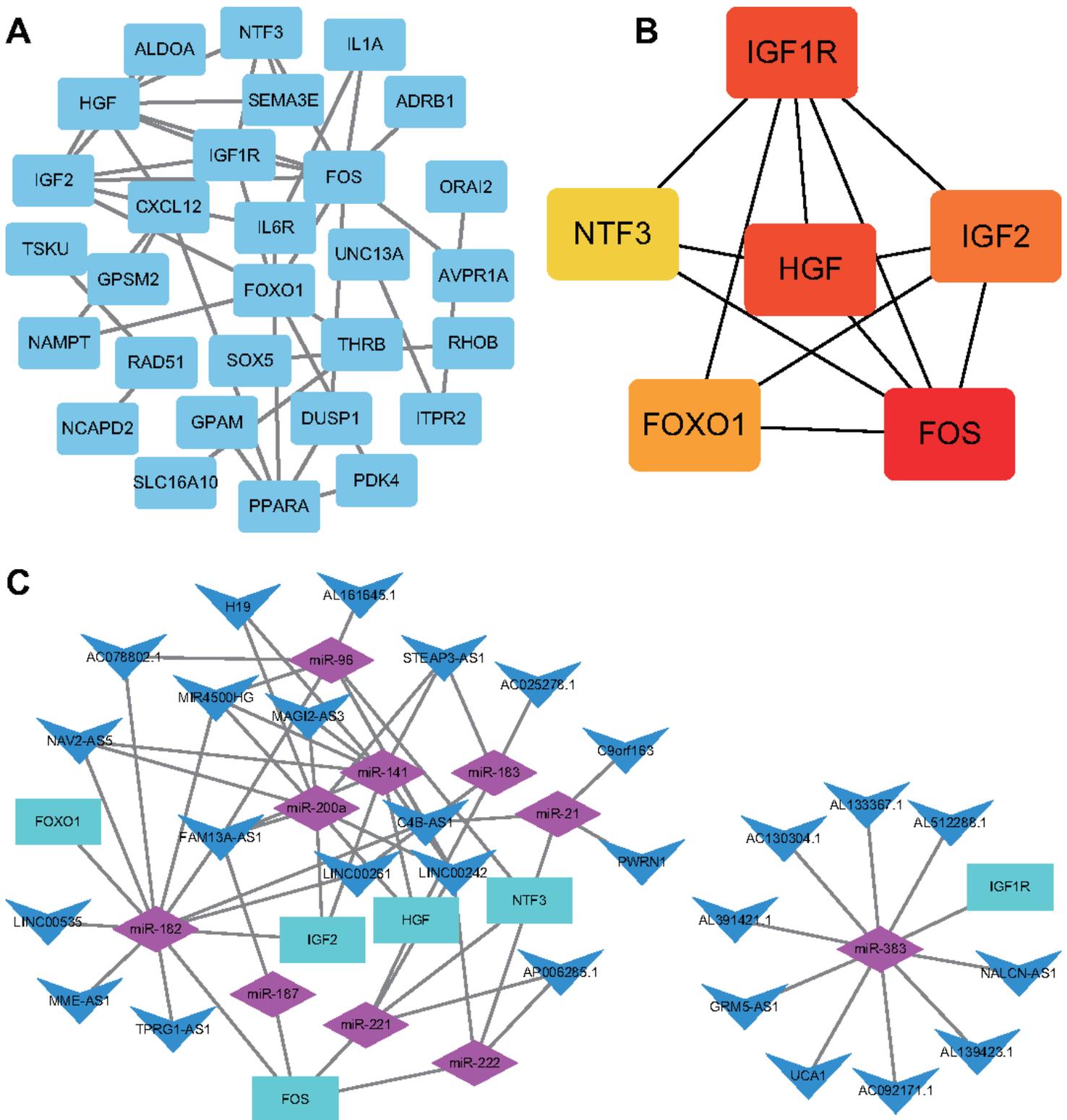


lncRNA nodes, 12 miR nodes, and 60 mRNAs. The triangle represents lncRNA, the circle represents miR, and the square represents mRNA. Nodes highlighted in red and blue indicate up and down, respectively.



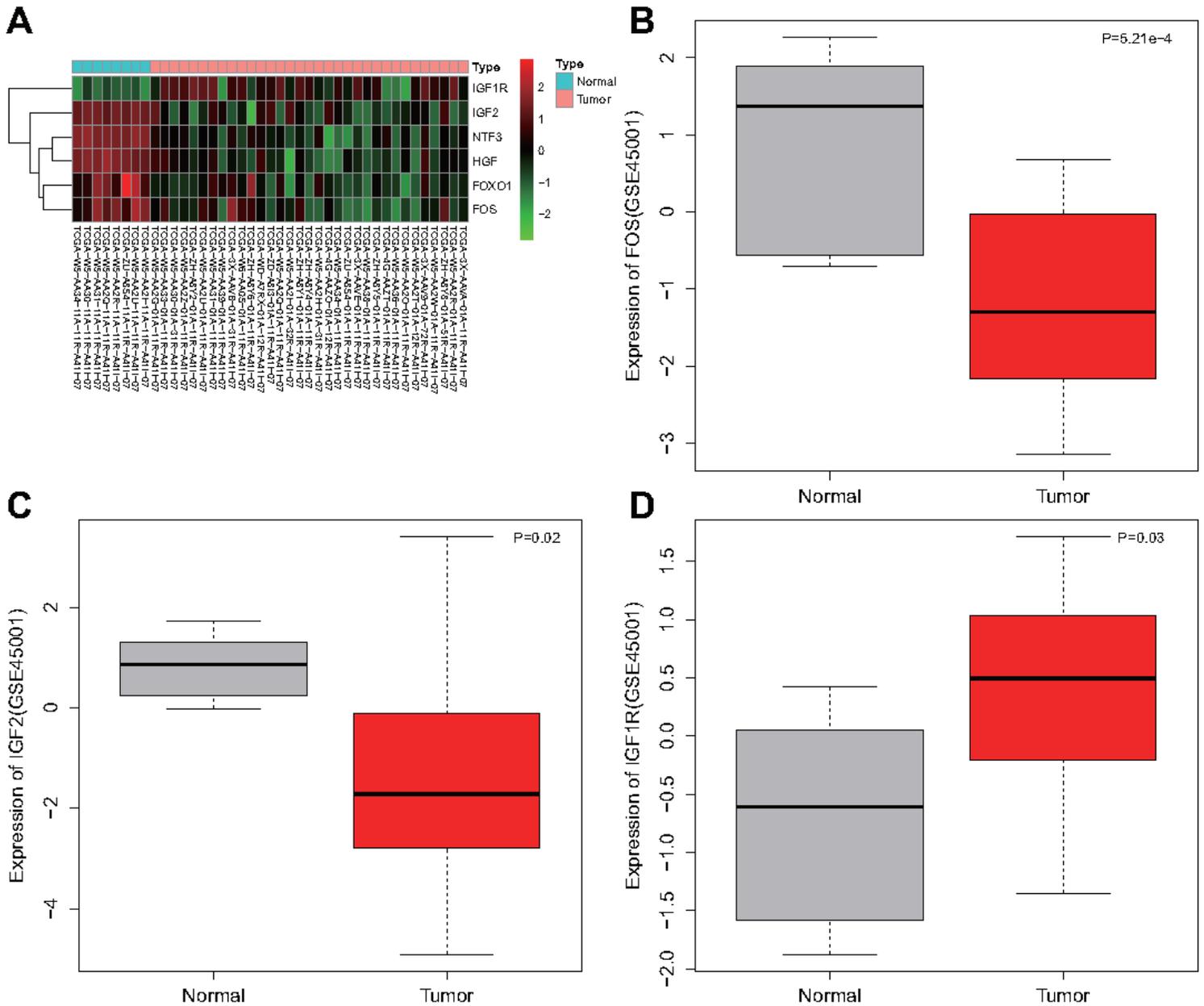
**Figure 3**

GO and KEGG pathway enrichment analysis were conducted on DEmRNAs A, The 60 DEmRNAs enriched GO terms were obtained using DAVID database ( $p < 0.05$ ). The y-axis represents GO terms, and the x-axis represents number of DEmRNAs enriched on the GO terms. B, The 60 DEmRNAs enriched KEGG pathways obtained from KOBAS database ( $p < 0.001$ ). The y-axis represents KEGG pathways, and the x-axis represents number of DEmRNAs enriched on the pathway. The larger bubbles indicate the increased number of enriched DEmRNAs, and the redder color represents more significantly enriched DEmRNAs.



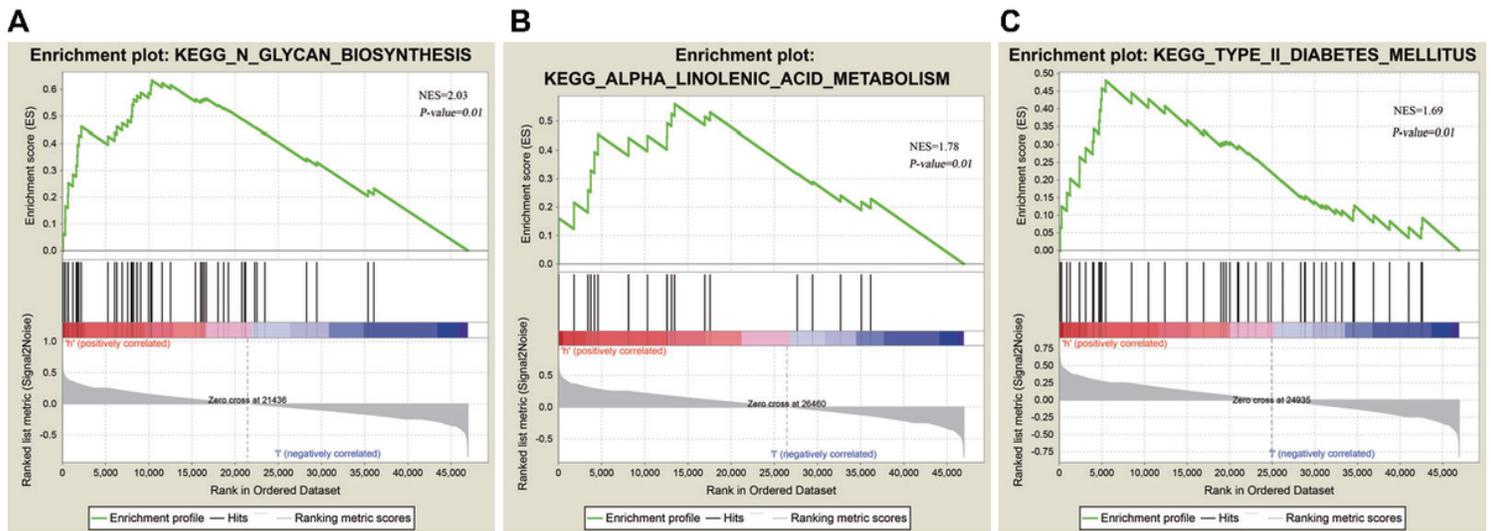
**Figure 4**

Six genes in the hub module were identified in the PPI network. A, PPI network of 60 DEmRNAs, consisting of 60 nodes and 41 edges. B, The first 6 genes from hub modules among the 60 DEmRNAs in PPI network. The node color gradually changes from yellow to red based on the log(2) fold change. C, lncRNA-miR-hub gene network. The network consists of 27 lncRNAs, 10 miRs, and 6 genes in the hub model. Vs indicate lncRNA, diamond indicates miR, and rectangle indicates mRNA.



**Figure 5**

The expression of FOS, HGF, IGF2, FOXO1, NTF3 and IGF1R was validated in ICC. A, The heat map of genes in the hub module expression in ICC tissues and adjacent normal tissues from the TCGA database. The x-axis represents the sample, and the y-axis represents the genes in the hub module. Red represents the up-regulated genes and green represents the down-regulated genes. B-D, The expression of FOS, IGF2, and IGF1R in gene expression dataset GSE45001 of ICC.



**Figure 6**

The biological pathways associated with FOS, IGF2, and IGF1R were predicted in ICC tissues using GSEA. A, The biological pathway enriched in the FOS high expression group. B, The biological pathway enriched in the IGF2 high expression group. C, The biological pathway enriched in the IGF1R high expression group.