



An IncRNA Model for Predicting the Prognosis of Hepatocellular Carcinoma Patients and ceRNA Mechanism

Hao Zhang¹, Renzheng Liu¹, Lin Sun² and Xiao Hu¹*

¹Department of Hepatobiliary Pancreatic Surgery, The Affiliated Hospital of Qingdao University, Qingdao, China, ²Department of ICU, The Affiliated Hospital of Qingdao University, Qingdao, China

Liver cancer is a highly malignant tumor. Notably, recent studies have found that long noncoding RNAs (IncRNAs) play a prominent role in the prognosis of patients with liver cancer. Herein, we attempted to construct an IncRNA model to accurately predict the survival rate in liver cancer. Based on The Cancer Genome Atlas (TCGA) database, we first identified 1066 IncRNAs with differential expression. The patient data obtained from TCGA were divided into the experimental group and the verification group. According to the difference in IncRNAs, we used single-factor and multi-factor Cox regression to select the genes needed to build the model in the experimental group, which were verified in the verification group. The results showed that the model could accurately predict the survival rate of patients in the high and low risk groups. The reliability of the model was also confirmed by the area under the receiver operating characteristic curve. Our model is significantly correlated with different clinicopathological features. Finally, we built a ceRNA network based on IncRNAs, which was used to display miRNAs and mRNAs related to IncRNAs. In summary, we constructed an IncRNA model to predict the survival rate of patients with hepatocellular carcinoma.

OPEN ACCESS

Edited by:

Yunpeng Hua, The First Affiliated Hospital of Sun Yat-sen University, China

Reviewed by:

Gao Liu, Meizhou People's Hospital, China Shunli Shen, The First Affiliated Hospital of Sun Yat-sen University, China

*Correspondence:

Xiao Hu 8371270@qq.com

Specialty section:

This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

Received: 29 July 2021 Accepted: 08 October 2021 Published: 12 November 2021

Citation:

Zhang H, Liu R, Sun L and Hu X (2021) An IncRNA Model for Predicting the Prognosis of Hepatocellular Carcinoma Patients and ceRNA Mechanism. Front. Mol. Biosci. 8:749313. doi: 10.3389/fmolb.2021.749313 Keywords: IncRNA, model, prognosis, hepatocellular carcinoma, ceRNA

INTRODUCTION

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer, with high mortality (Farazi and DePinho, 2006). Despite several available treatment methods for HCC, patients in different stages require different treatment measures and the long-term survival remains poor. At present, the staging and grading of HCC is not comprehensive (Zhu et al., 2011), which makes assessment of prognosis difficult. Therefore, it is important to identify predictive indicators and develop a prognostic model for patients with HCC.

As the biological functions of long non-coding RNAs (lncRNAs) have been gradually uncovered, they have become a research hotspot. The function of lncRNAs can be divided into two types: regulating the local chromatin structure and/or gene expression of cis vs. trans, and those leaving the transcription site and performing cell function with trans transcripts (Gil and Ulitsky, 2020). After

Abbreviations: HCC, hepatocellular carcinoma; lncRNAs, long non-coding RNAs; qPCR, quantitative real-time PCR; ROC, receiver operating characteristic; TCGA, The Cancer Genome Atlas.

TABLE 1	Primer	sequences	used	in	the	study
---------	--------	-----------	------	----	-----	-------

AC016717.2	Forward:5'-GATGCTGATGCTGCCTGTCCATAG-3
	Reverse:5'-ATGCTGTGCTGTTGGTCTCTGAAG-3'
DDX11-AS1	Forward:5'-CTGGCTACTCTTCCTCCTGG-3'
	Reverse:5'-CAGAGGACATGTGGGAGGTT-3'
LINC00462	Forward:5'-ACTAGGTCCTTCTGGTGTT-3'
	Reverse:5'-GTAAAACTTGCTGCTGATG-3'
ZFPM2-AS1	Forward:5'-GGTGGCACCTGAAATCACAGA-3'
	Reverse:5'-TGCAAGATGACGCTCAGTCG-3'
GAPDH	Forward:5'-TCGGAGTCAACGGATTTGGT-3'
	Reverse:5'-TTCCCGTTCTCAGCCTTGAC-3'

the role of lncRNAs in regulating tumorigenesis and oncoprotein expression was established, lncRNAs were also found to alter the prognosis of numerous human cancers. Various studies have shown the effect of lncRNAs on the survival of patients with HCC. However, there is no effective prognostic model to evaluate the prognosis and guide the treatment of HCC. The Cancer Genome Atlas (TCGA) database has abundant transcriptome expression profiles of different cancers. Bioinformatics facilitates screening of biomarkers and constructing prediction models using public databases. Through the analysis of gene transcription level using TCGA and bioinformatics, this study found new molecular markers related to HCC survival and established a prognosis model of lncRNAs in HCC. Cox regression analysis, and least absolute shrinkage and selection operator were implemented to validate the lncRNAs associated with prognosis. The power of the prognostic model including four lncRNAs was tested by the receiver operating characteristic (ROC) curve. Statistics and graphic visualization were completed in R software.

Next, we used miRcode database to validate LINC00462 and find the downstream miRNAs. We used miRDB, miRTarBase, and TargetScan databases to find miRNAs targeting specific genes. Finally, we used the Cytoscape to structure the ceRNA network.

MATERIALS AND METHODS

Clinical Samples and Ethics Statement

The ethics committee of the Affiliated Hospital of Qingdao University approved this research. This study included 10 HCC patients (5 males and 5 females), who were diagnosed by pathology, between June 2018 and June 2019, at the Affiliated Hospital of Qingdao University (Shandong, China). The age range was 50–80 years. Tissues were sliced into small sections and stored at -80° C. Corresponding adjacent non-cancerous tissues (ANT) and cancerous tissues were simultaneously collected. The hospital's Institutional Review Board strictly adhered to the Declaration of Helsinki protocol for approving the study. Written informed consent was signed by each patient. The institutional approval number for human studies is QYFY ZWLL 26019.

Gene Expression Profiles of HCC Patients

The expression data of lncRNAs, miRNAs, and mRNAs, as well as clinical features were retrieved from TCGA (https://portal.gdc.cancer.





gov/). We downloaded the annotation file of transcriptome from Ensembl online database (https://asia.ensembl.org/index.html) and used it to distinguish lncRNAs and miRNAs. There are not copyright issues. Thereafter, we selected differentially expressed genes for this



study. The screening of genes was performed in the R software. We used edgeR package to identify differentially expressed genes.

Definition Of Prognostic Risk Model in the Training Dataset

We randomly divided 342 patients into the training dataset (172) and validation dataset (170). Cox regression analysis, and least

absolute shrinkage and selection operator were used to screen prognosis related lncRNAs. Four lncRNAs were screened out in the training dataset. Finally, the following model was established to assess the prognostic risk. Risk score = coefficient₁ * Expression₁ + coefficient₂ * Expression₂ + coefficient_N * Expression_N. N is the above-selected lncRNAs, expression is the lncRNA expression in every HCC patient. Cox regression analysis was used to calculate the corresponding lncRNA coefficients in the training dataset. The

"survival", "caret", "glmnet", "survminer", and "survivalROC" packages were used in the analysis process.

RNA Extraction, Reverse Transcription, and Quantitative Real-Time PCR

We used RNAiso Plus (TaKaRa, Tokyo, Japan) to extract the total RNA from the tissues. The quality of RNA was detected by NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Inc.). Next, reverse transcription was performed with T100TM thermal cycler (BIO-RAD, US) using PrimeScriptTMRT reagent kit and gDNA eraser (TaKaRa, Tokyo, Japan). Quantitative real-time PCR (qPCR) was performed in LightCycler[®] 96 (Roche, Switzerland) with TB Green[®]Premix Ex TaqTM II (Tli RNase H Plus) (TaKaRa, Tokyo, Japan). Primer sequences are listed in Table 1.

Statistical Analysis and the Formation of ceRNA Network

According to the median risk score, the patients were divided into the high-risk group and the low-risk group. All the analyses were performed using the R program and software package. The receiver operating characteristic (ROC) curve was used to test the predictive ability of the prognosis model composed of four lncRNAs (Mandrekar, 2010).

We used miRcode (http://www.mircode.org/) (Jeggari et al., 2012), miRDB (http://mirdb.org/) (Chen and Wang, 2020), miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/index.html) (Chou et al., 2018), and TargetScan (http://www.targetscan. org) (Lewis et al., 2005) databases to predict the downstream genes. We used Cytoscape to visualize the ceRNA network.

RESULTS

Identification of Differentially Expressed IncRNAs

We downloaded the related data of 422 patients, 50 normal cases and 372 tumor cases from TCGA database. There were 1066 out of 6739 lncRNAs with logFC >2 and FDR <0.01 in the TCGA dataset, after removing lncRNAs with extremely similar gene expression between HCC and normal patients (**Figure 1**).

Identification of Prognostic IncRNAs From the Training Dataset

We randomly split 342 cancer patients. The training dataset included 172 patients and the test group had 170 patients. We further screened the predicting lncRNAs from the 1066 lncRNAs by univariate Cox analysis. Thereafter, 104 lncRNAs were selected that had an impact on the survival of patients (**Figure 2**).

Construction of the Prognostic Model and Checking the Expression of IncRNAs

The number of genes was reduced to 5 lncRNAs by least absolute shrinkage and selection operator (**Figure 3A**). Finally, four lncRNAs

(DDX11-AS1, ZFPM2-AS1, AC016717.2, and LINC00462), which were closely related to patient survival, were selected by multivariate Cox analysis (**Figure 3B**). A prognostic model composed of the four lncRNAs was established as follows: Risk score = 0.0257 * DDX11-AS1 + 0.0008 * ZFPM2-AS1 + 0.0026 * AC016717.2 + 0.0306 * LINC00462. We verified the transcription levels of these four genes in 10 pairs of human tissue samples (**Figure 3C**).

The Model for Survival Prediction

Based on the median risk score, we divided the patients from the training dataset into the high-risk group (n = 86) and the low-risk group (n = 86) and compared the survival of the two risk groups by Kaplan–Meier survival analysis. The survival rate of the high-risk group was significantly lower than that of the low-risk group (**Figure 4A**). To confirm the predictive power of the prognostic model, we divided the patients in the test dataset into two risk groups according to the median risk score. The Kaplan–Meier curve showed that in the trial cohort, the lifespan of patients in the low-risk group was prominently longer than that in the high-risk group, which was similar to the results obtained from the training dataset (**Figure 4B**).

The survival status of HCC patients obtained from the training dataset and test dataset is depicted in the plot. The heat map shows the different expression patterns of the four prognosis-related lncRNAs in the low-risk and high-risk populations (**Figures 4C,D**). Similar to the training dataset, patients with high-risk scores tended to have higher levels of the four lncRNAs than patients with low-risk scores. According to the ROC (AUC) curve, the size of the AUC was related to the prognosis. The ROC curves at 1 year of the whole, training, and validation datasets are shown in **Figures 4E–G**.

Prognostic Model and Clinicopathological Features

The Chi square test was used to analyze the significance of clinicopathological features between high- and low-risk groups (**Figure 5A**). In addition, by analyzing the data of patients with different characteristics, we found that the risk score can significantly distinguish the prognostic differences in different age, gender, G stage, M0, N0, stage, and T group (**Figure 5B**), which significantly improves the prognostic value of our model.

Construction of the ceRNA Network with the Prognostic IncRNAs

Using the TCGA database, 1976 differentially expressed mRNAs were selected from 424 HCC patients (logFC > 2 or logFC < -2,FDR<0.01). From 50 normal samples and 376 tumor samples, 247 differentially expressed miRNAs were selected (logFC > 1 or logFC < -1, FDR < 0.01). We compared DDX11-AS1, ZFPM2-AS1, AC016717.2, and LINC00462 in the miRcode database to identify the downstream miRNAs, and only LINC00462 could be found in the miRcode database. The miRNAs found in the miRcode intersected with the differentially expressed miRNAs in TCGA. The miRDB, miRTarBase, and TargetScan data were combined to predict the downstream mRNAs with the miRNAs obtained in the previous step. The downstream target genes were obtained using the intersection of the differentially expressed



mRNAs in the previous TCGA database and the predicted mRNAs in the three databases (Figure 6).

DISCUSSION

In recent years, with the gradual advancements in the study of lncRNAs, their function in the occurrence and development of HCC has become a research hotspot. Therefore, it is critical to understand the mechanism of action of lncRNAs and comprehensively analyze their clinical significance in patients with liver cancer. The purpose of this study was to establish lncRNA markers related to prognosis of HCC and explore their molecular mechanisms. This study was undertaken to provide a better understanding of the prognosis of patients with liver cancer and facilitate the discovery of biomarkers.

In this study, we downloaded related data of 422 patients from TCGA database and screened out 1066 lncRNAs. We randomly

split 342 cancer patients. We further selected the 104 predictive lncRNAs from 1066 lncRNAs by univariate Cox analysis in the training dataset. Finally, four lncRNAs (DDX11-AS1, ZFPM2-AS1, AC016717.2, and LINC00462), which were closely related to patient survival, were selected by least absolute shrinkage and selection operator, and multivariate Cox analysis. A prognostic model or a risk score was constituted with the above lncRNAs. We divided patients from the training dataset into the high-risk group and the low-risk group. The survival of patients in the high-risk group was significantly less than those in the low-risk group. The survival of patients in the test dataset was similar to the training dataset. Furthermore, ROC curve analysis showed that the lncRNA risk score model had good prediction efficiency in HCC. The AUC value of our model is significantly higher than that of Wu et al. (2020) and Zhou et al. (2021). Finally, we used the miRcode, miRDB, miRTarBase, and TargetScan databases to find the downstream miRNAs and mRNAs, and create the ceRNA network.





DDX11-AS1 was identified as an oncogene in HCC. Silencing the expression of this gene reduced the proliferation, migration, and invasion ability of HCC cells, and the gene could promote the growth of tumors in vivo (Wan et al., 2020). DDX11-AS1 was also reported to be related to overall survival (Li et al., 2019). DDX11-AS1 displayed a cancer-promoting role by regulating the expression of related genes directly or indirectly in various malignant tumors, such as osteosarcoma, colorectal cancer, gastric cancer, non-small cell lung cancer, and bladder cancer (Feng et al., 2020). ZFPM2-AS1 regulates the expression of GDF10 through competitive binding to miRNA to promote cell proliferation, migration, invasion, and inhibit apoptosis in HCC (He et al., 2020). As an oncogenic gene, ZFPM2-AS1 plays a role in retinoblastoma (Lyv et al., 2020), breast cancer (Zhao et al., 2020), small cell lung cancer (Yan et al., 2020), cervical cancer (Dai et al., 2020), esophageal squamous cell carcinoma (Sun and Wu, 2020), and other tumors. AC016717.2 is located on chromosome 2 and the gene maps to 226, 563, 230-226, 568, 370 in GRCh37 coordinates. There are few studies on this lncRNA. LINC00462 can significantly increase the invasive ability of HCC cells (Gong et al., 2017). It can also promote pancreatic cancer (Zhou et al., 2018).

REFERENCES

- Chen, Y., and Wang, X. (2020). miRDB: an Online Database for Prediction of Functional microRNA Targets. *Nucleic Acids Res.* 48 (D1), D127–D131. doi:10.1093/nar/gkz757
- Chou, C.-H., Shrestha, S., Yang, C.-D., Chang, N.-W., Lin, Y.-L., Liao, K.-W., et al. (2018). miRTarBase Update 2018: a Resource for Experimentally Validated

In summary, this was the first study to propose a prediction model based on four lncRNAs, in which the effects of DDX11-AS1, ZFPM2-AS1, and lncRNAs on the prognosis of patients with liver cancer have been confirmed. AC016717.2 and LINC00462 are new biomarkers in liver cancer. Although the lncRNAs selected by the predictive model showed good performance, this study only included internal validation results but no external data validation due to resource constraints. Hence, clear evidence with multi-center clinical data is required to validate the results of this study.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The ethics committee of the Affiliated Hospital of Qingdao University approved this research. This study included 10 HCC patients (5 males and 5 females), who were diagnosed by pathology, between June 2018 and June 2019, at the Affiliated Hospital of Qingdao University (Shandong, China). The age range was 50–80 years. Tissues were sliced into small sections and stored at –80°C. Corresponding adjacent non-cancerous tissues (ANT) and cancerous tissues were simultaneously collected. The hospital's Institutional Review Board strictly adhered to the Declaration of Helsinki protocol for approving the study. Written informed consent was signed by each patient. The institutional approval number for human studies is QYFY ZWLL 26019.

AUTHOR CONTRIBUTIONS

H.Z. contributed to the study design, acquired and analyzed data, and wrote the manuscript. R.L. and L.S. contributed to statistical analysis and data acquisition. X.H. provided intellectual input. All authors read and approved the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.749313/full#supplementary-material

microRNA-Target Interactions. Nucleic Acids Res. 46 (D1), D296–D302. doi:10.1093/nar/gkx1067

- Dai, J., Wei, R., Zhang, P., and Liu, P. (2020). Long Noncoding RNA ZFPM2-AS1 Enhances the Malignancy of Cervical Cancer by Functioning as a Molecular Sponge of microRNA-511-3p and Consequently Increasing FGFR2 Expression. *Cmar* 12, 567–580. doi:10.2147/cmar.s238373
- Farazi, P. A., and DePinho, R. A. (2006). Hepatocellular Carcinoma Pathogenesis: from Genes to Environment. Nat. Rev. Cancer 6 (9), 674–687. doi:10.1038/nrc1934

- Feng, Y., Wu, M., Hu, S., Peng, X., and Chen, F. (2020). LncRNA DDX11-AS1: a Novel Oncogene in Human Cancer. *Hum. Cel* 33 (4), 946–953. doi:10.1007/ s13577-020-00409-8
- Gil, N., and Ulitsky, I. (2020). Regulation of Gene Expression by Cis-Acting Long Non-coding RNAs. Nat. Rev. Genet. 21 (2), 102–117. doi:10.1038/s41576-019-0184-5
- Gong, J., Qi, X., Zhang, Y., Yu, Y., Lin, X., Li, H., et al. (2017). Long Noncoding RNA Linc00462 Promotes Hepatocellular Carcinoma Progression. *Biomed. Pharmacother.* 93, 40–47. doi:10.1016/j.biopha.2017.06.004
- He, H., Wang, Y., Ye, P., Yi, D., Cheng, Y., Tang, H., et al. (2020). Long Noncoding RNA ZFPM2-AS1 Acts as a miRNA Sponge and Promotes Cell Invasion through Regulation of miR-139/GDF10 in Hepatocellular Carcinoma. J. Exp. Clin. Cancer Res. 39 (1), 159. doi:10.1186/s13046-020-01664-1
- Jeggari, A., Marks, D. S., and Larsson, E. (2012). miRcode: a Map of Putative microRNA Target Sites in the Long Non-coding Transcriptome. *Bioinformatics* 28 (15), 2062–2063. doi:10.1093/bioinformatics/bts344
- Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005). Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes Are microRNA Targets. *Cell* 120 (1), 15–20. doi:10.1016/ j.cell.2004.12.035
- Li, Y., Zhuang, W., Huang, M., and Li, X. (2019). Long Noncoding RNA DDX11-AS1 Epigenetically Represses LATS2 by Interacting with EZH2 and DNMT1 in Hepatocellular Carcinoma. *Biochem. Biophysical Res. Commun.* 514 (4), 1051–1057. doi:10.1016/j.bbrc.2019.05.042
- Lyv, X., Wu, F., Zhang, H., Lu, J., Wang, L., and Ma, Y. (2020). Long Noncoding RNA ZFPM2-AS1 Knockdown Restrains the Development of Retinoblastoma by Modulating the MicroRNA-515/HOXA1/Wnt/β-Catenin Axis. *Invest. Ophthalmol. Vis. Sci.* 61 (6), 41. doi:10.1167/iovs.61.6.41
- Mandrekar, J. N. (2010). Receiver Operating Characteristic Curve in Diagnostic Test Assessment. J. Thorac. Oncol. 5 (9), 1315–1316. doi:10.1097/ jto.0b013e3181ec173d
- Sun, G., and Wu, C. (2020). ZFPM2-AS1 Facilitates Cell Growth in Esophageal Squamous Cell Carcinoma via Up-Regulating TRAF4. *Biosci. Rep.* 40 (4). doi:10.1042/BSR20194352
- Wan, T., Zheng, J., Yao, R., Yang, S., Zheng, W., and Zhou, P. (2020). LncRNA DDX11-AS1 Accelerates Hepatocellular Carcinoma Progression via the miR-195-5p/MACC1 Pathway. Ann. Hepatol. 20, 100258. doi:10.1016/ j.aohep.2020.09.003

- Wu, H., Liu, T., Qi, J., Qin, C., and Zhu, Q. (2020). Four Autophagy-Related IncRNAs Predict the Prognosis of HCC through Coexpression and ceRNA Mechanism. *Biomed. Res. Int.* 2020, 3801748. doi:10.1155/2020/3801748
- Yan, Z., Yang, Q., Xue, M., Wang, S., Hong, W., and Gao, X. (2020). YY1-induced lncRNA ZFPM2-AS1 Facilitates Cell Proliferation and Invasion in Small Cell Lung Cancer via Upregulating of TRAF4. *Cancer Cel Int* 20, 108. doi:10.1186/ s12935-020-1157-7
- Zhao, Y. F., Li, L., Li, H. J., Yang, F. R., Liu, Z. K., Hu, X. W., et al. (2020). LncRNA ZFPM2-AS1 Aggravates the Malignant Development of Breast Cancer via Upregulating JMJD6. *Eur. Rev. Med. Pharmacol. Sci.* 24 (21), 11139–11147. doi:10.26355/eurrev_202011_23601
- Zhou, B., Guo, W., Sun, C., Zhang, B., and Zheng, F. (2018). Linc00462 Promotes Pancreatic Cancer Invasiveness through the miR-665/tgfbr1-Tgfbr2/smad2/3 Pathway. *Cell Death Dis* 9 (6), 706. doi:10.1038/s41419-018-0724-5
- Zhou, X., Dou, M., Liu, Z., Jiao, D., Li, Z., Chen, J., et al. (2021). Screening Prognosis-Related lncRNAs Based on WGCNA to Establish a New Risk Score for Predicting Prognosis in Patients with Hepatocellular Carcinoma. *J. Immunol. Res.* 2021, 5518908. doi:10.1155/2021/5518908
- Zhu, A. X., Duda, D. G., Sahani, D. V., and Jain, R. K. (2011). HCC and Angiogenesis: Possible Targets and Future Directions. *Nat. Rev. Clin. Oncol.* 8 (5), 292–301. doi:10.1038/nrclinonc.2011.30

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Zhang, Liu, Sun and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.