DECODING NON-CODING RNA IMPLICATED IN CANCER CELL SURVIVAL & GROWTH MODULATION

EDITED BY: Ramkrishna Mitra and Wei Jiang
PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-320-7
DOI 10.3389/978-2-88976-320-7

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

DECODING NON-CODING RNA IMPLICATED IN CANCER CELL SURVIVAL & GROWTH MODULATION

Topic Editors:
Ramkrishna Mitra, Thomas Jefferson University, United States
Wei Jiang, Nanjing University of Aeronautics and Astronautics, China

Citation: Mitra, R., Jiang, W., eds. (2022). Decoding Non-Coding RNA Implicated in Cancer Cell Survival & Growth Modulation. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-320-7

Table of Contents

frontiers in Genetics

ORIGINAL RESEARCH
published: 28 January 2021
doi: 10.3389/fgene.2020.625977

A Seven-Long Non-coding RNA Signature Improves Prognosis Prediction of Lung Adenocarcinoma: An Integrated Competing Endogenous RNA Network Analysis

Rang Li[1†], Kedong Han[2†], Dehua Xu[1], Xiaolin Chen[1], Shujin Lan[1], Yuanjun Liao[1], Shengnan Sun[1] and Shaoqi Rao[1]*

[1] Institute of Medical Systems Biology, School of Public Health, Guangdong Medical University, Dongguan, China,
[2] Department of Cardiology, Maoming People's Hospital, Maoming, China

OPEN ACCESS

Edited by:
Wei Jiang,
Nanjing University of Aeronautics and Astronautics, China

Reviewed by:
Antonio Mora,
Guangzhou Medical University, China
Jinming Li,
Southern Medical University, China

***Correspondence:**
Shaoqi Rao
raoshaoq@gdmu.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *13 November 2020*
Accepted: *21 December 2020*
Published: *28 January 2021*

Citation:
Li R, Han K, Xu D, Chen X, Lan S, Liao Y, Sun S and Rao S (2021) A Seven-Long Non-coding RNA Signature Improves Prognosis Prediction of Lung Adenocarcinoma: An Integrated Competing Endogenous RNA Network Analysis. Front. Genet. 11:625977. doi: 10.3389/fgene.2020.625977

Early and precise prediction is an important way to reduce the poor prognosis of lung adenocarcinoma (LUAD) patients. Nevertheless, the widely used tumor, node, and metastasis (TNM) staging system based on anatomical information only often could not achieve adequate performance on foreseeing the prognosis of LUAD patients. This study thus aimed to examine whether the long non-coding RNAs (lncRNAs), known highly involved in the tumorigenesis of LUAD through the competing endogenous RNAs (ceRNAs) mechanism, could provide additional information to improve prognosis prediction of LUAD patients. To prove the hypothesis, a dataset consisting of both RNA sequencing data and clinical pathological data, obtained from The Cancer Genome Atlas (TCGA) database, was analyzed. Then, differentially expressed RNAs (DElncRNAs, DEmiRNAs, and DEmRNAs) were identified and a lncRNA–miRNA–mRNA ceRNA network was constructed based on those differentially expressed RNAs. Functional enrichment analysis revealed that this ceRNA network was highly enriched in some cancer-associated signaling pathways. Next, lasso-Cox model was run 1,000 times to recognize the potential survival-related combinations of the candidate lncRNAs in the ceRNA network, followed by the "best subset selection" to further optimize these lncRNA-based combinations, and a seven-lncRNA prognostic signature with the best performance was determined. Based on the median risk score, LUAD patients could be well distinguished into high-/low-risk subgroups. The Kaplan–Meier survival curve showed that LUAD patients in the high-risk group had significantly shorter overall survival than those in the low-risk group (log-rank test $P = 4.52 \times 10^{-9}$). The ROC curve indicated that the clinical genomic model including both the TNM staging system and the signature had a superior performance in predicting the patients' overall survival compared to the clinical model with the TNM staging system only. Further stratification analysis suggested that the signature could work well in the different strata of the stage, gender, or age, rendering it to be a wide application. Finally, a ceRNA subnetwork related to the signature was extracted, demonstrating its high involvement in the

tumorigenesis mechanism of LUAD. In conclusion, the present study established a lncRNA-based molecular signature, which can significantly improve prognosis prediction for LUAD patients.

Keywords: lung adenocarcinoma, prognosis, lncRNA, molecular signature for survival, ceRNA network

INTRODUCTION

Lung adenocarcinoma (LUAD), a major type of non-small cell lung cancer (NSCLC), has a low survival rate and an increasing incidence (Matsuda and Machii, 2015; Denisenko et al., 2018). The etiology of LUAD is multifactorial, involving a large number of environmental factors and internal factors (Rajer et al., 2014). So far, because of the lack of specific symptoms and signs, LUAD coupling with complex and diverse clinical manifestations is easily missed and misdiagnosed (Del et al., 2017). Hence, when most LUAD patients are diagnosed, they are already in an advanced stage. Although many kinds of treatments, including surgical resection, chemotherapy, radiotherapy, and chemo-radiotherapy, were applied to improve patient's survival rate, the overall 5-year survival rate is still extremely bleak (Wakeam et al., 2017). Therefore, early detection and diagnosis is vital to improve LUAD patient's poor prognosis.

The TNM staging system is currently the most common tumor prognosis predictor and a powerful tool for guiding adjuvant therapy at present (Mittendorf et al., 2015; Pontius et al., 2017). According to the invasion extent of the primary tumor stage (T stage), regional lymph node metastasis stage (N stage), and distant metastasis stage (M stage), the total pathologic stage of the malignant tumor (the TNM stage) could be determined (Hutter, 1991). In general, the higher the TNM stage, the higher the degree of the malignant tumor is. However, the TNM staging system, which is limited to the anatomical extent rather than the biological behavior of the disease, has obvious limitations compared with the multifactorial prognostic index (Fouad et al., 2017; Ball, 2019). Given the shortcomings of the TNM staging system in LUAD patient's prognosis prediction, it is highly demanding to develop a molecular diagnostic and predictive biomarker.

Long non-coding RNAs (lncRNAs) are defined as any ncRNA that is 200 nucleotides to 100 kb in length (Yin et al., 2018). Many studies reported that lncRNA plays an important role in the pathogenesis of cancer and has significant clinical value in prognosis and diagnosis (Knoll et al., 2015; Evans et al., 2016; Huang et al., 2017; Tripathi et al., 2018). It was also demonstrated that lncRNA can act as a "sponge" to regulate the targeted gene expression by competitively binding with miRNA (Zhang et al., 2016). This novel model of gene regulation is a part of the competing endogenous RNA (ceRNA) hypothesis, which was first proposed in 2011 (Salmena et al., 2011). ceRNAs (including lncRNA, circRNA, and mRNA) competitively bind with microRNA via sharing microRNA response elements (MREs) to weaken the inhibition effect for the target gene. The regulatory relations among lncRNAs, miRNA, and mRNA form a complex ceRNA network, and the abnormal expression of lncRNA would destroy the balance of the ceRNA network to lead to the initiation and progression of cancer (Karreth and Pandolfi, 2013).

Owing to the heterogeneity and polygenic mutation in lung cancer, a single genomic mutation is difficult to explain the various phenotypes and the variable risks of complex disease (Andreassen et al., 2014). Compared with a single gene and single factor, a biomolecular network(s) including multiple disease-related factors, which perform their dysfunctions through physical and biochemical interactions in a network (Zhao and Liu, 2019), represents various molecular relationships underlying complex diseases and depicts a clear global picture of interactions among disease-related factors (Jinawath et al., 2016). As a biomolecular network, the ceRNA network, describing post-transcriptional interactions between ceRNAs and miRNAs, had great value in prognosis, diagnosis, and therapy of cancers (Lin et al., 2018; Zhang Y. et al., 2018; Eissa et al., 2019). In recent years, there are several successful attempts that use the ceRNA networks to identify prognostic signatures for different cancers (Hu et al., 2019; Wang et al., 2019b).

The aim of the present study was to establish a multi-lncRNA prognosis predictor. To this end, a LUAD-related lncRNA–miRNA–mRNA ceRNA network was constructed based on integrated transcriptome data from The Cancer Genome Atlas (TCGA) database. Then, by using lasso-Cox regression model, a seven-lncRNA prognostic signature was identified from the LUAD-related ceRNA network. Survival analysis and the receiver operating characteristic (ROC) curve suggested the seven-lncRNA prognostic signature is a robust and independent prognostic factor. Most importantly, our study demonstrated that the seven-lncRNA prognostic signature effectively enhanced the prognosis prediction performance over the conventional TNM staging system.

MATERIALS AND METHODS

Data Retrieval and Processing

RNA and miRNA sequencing raw count data and corresponding clinical data of LUAD patients were obtained from TCGA database by using the GDC Data Transfer Tool. Then, individual sample expression files were merged into an expression matrix using the Perl language for further processing. To eliminate the adverse effect of low abundance, RNAs with an average value of less than 1 were excluded in further analysis. Finally, the trimmed mean of the *M*-values (TMM) method was used to normalize RNA sequencing data (Smid et al., 2018).

Identifying Differentially Expressed RNAs

Differentially expressed RNAs (DE-lncRNAs, DE-miRNAs, and DE-mRNAs) were identified by comparing the expression

values between LUAD samples and adjacent normal tissue samples based on the edgeR package of R platform (Robinson et al., 2010). The cutoff criterion was set at $|\log FC| > 1$ and FDR < 0.05 for the screening of DE-lncRNAs, DE-miRNAs, and DE-mRNAs. Volcano plot was used to display the differentially expressed RNAs.

Constructing the ceRNA Network

Regulatory relationships among DE-RNAs were identified by mining knowledge of several public databases. miRcode database (Jeggari et al., 2012) was used to predict the regulatory relationships between lncRNAs and miRNAs, while miRDB, miRTarBase, and TargetScan databases (Zheng et al., 2019) were used to define the regulatory relationship between miRNAs and mRNAs. According to the ceRNA theory, there should be a negative regulatory relationship between ceRNAs and miRNAs (Taulli et al., 2013). Therefore, the Pearson correlation coefficient between lncRNAs/mRNAs and miRNAs was calculated to identify negatively correlated RNA–RNA regulatory pairs ($P < 0.05$). In short, the regulatory relationships between lncRNAs/mRNAs and miRNAs were determined by three facts: (1) having a biological basis, supported by knowledge bases; (2) being a negative relationship, which agrees with the competing endogenous RNA theory; and (3) achieving the significance level in the Pearson correlation analysis based on their expression data. Finally, based on the shared miRNAs among these regulatory pairs, the lncRNA–miRNA–mRNA ceRNA network was built by connecting negative lncRNA–miRNA and miRNA–mRNA regulatory pairs. Cytoscape v3.7.1 was used for network visualization (Shannon et al., 2003).

To reveal the biological function(s) that ceRNA regulatory network involved, the Kyoto Encyclopedia of Genes and Genomes (KEGG)-based enrichment analysis was conducted to assess the ceRNA regulatory network using clusterProfiler package in R (Yu et al., 2012). The enriched KEGG pathway(s) with a FDR less than 0.05 was considered as statistically significant.

Defining the Prognostic Signature

In order to identify the lncRNAs and optimal subset(s) related to the overall survival of LUAD patients, lasso-Cox model (R package glmnet) was run 1,000 times to recognize the potential survival-related combinations of the candidate lncRNAs in the ceRNA network, followed by the "best subset selection" (the area under the ROC curve, AUC > 0.70 with the minimal set size) to further optimize these lncRNA-based combinations (i.e., the survival-related signatures). Finally, in order to evaluate the joint effect of the best signature, a risk score was calculated based on a linear combination of the expression levels of the included lncRNAs weighted by their regression coefficients derived from the multivariate Cox regression analysis.

The risk score formula was defined as following:

$$Riskscore = \sum_{i=1}^{n} \exp_i \beta_i$$

Here, n represents the number of lncRNAs in the model, $\exp_i$ represents the expression level of lncRNA i, and β_i is the regression coefficient of lncRNA i in multivariate Cox regression model.

Assessing the Prognosis Value of the Newly Identified Signature

First, in order to evaluate its potential to classify the LUAD patients, according to the median risk score of the signature, LUAD patients were divided into low- and high-risk groups. The Kaplan–Meier method was used to display the difference in survival time between low-risk and high-risk LUAD groups. The statistical significance of the difference between the survival profiles of the two groups was determined by using the log-rank test. ROC curve was used to estimate its sensitivity and specificity. Second, to assess whether combining the lncRNA-based signature with the TNM stages could improve the prognosis prediction for LUAD, a clinical genomic model with the TNM stages and the lncRNA-based signature combined was constructed, and AUC was compared to the model with the TNM stages only. Third, in order to explore its applicability, a stratification analysis by the TNM stages, gender, or age was performed. Finally, to explore its biological role(s), a core ceRNA network was constructed and its functional involvements were identified by a KEGG-based enrichment analysis.

The detailed workflows of the proposed strategies for identifying and assessing the survival-related lncRNA-based signature are illustrated in **Figure 1**.

RESULTS

Aberrantly Expressed lncRNAs, miRNAs, and mRNAs

In total, RNA-seq data for 535 LUAD samples and 59 adjacent normal samples, miRNA-seq data for 519 LUAD samples and 48 adjacent normal samples, and the corresponding clinical data for 504 LUAD patients were obtained. According to the cutoff criteria ($|\log FC| > 1$ and FDR < 0.05), 5,537 mRNAs (3,721 upregulated and 1,816 downregulated), 352 miRNAs (273 upregulated and 79 downregulated), and 3,939 lncRNAs (3,202 upregulated and 737 downregulated) were found differentially expressed (named DE-lncRNAs, DE-miRNAs, and DE-mRNAs, respectively). Their volcano plots are shown in **Figure 2**.

The ceRNA Network for LUAD

Among these DE-RNAs, 475 lncRNA–miRNA pairs between 197 lncRNAs and 39 miRNAs and 198 miRNA–mRNA pairs between 39 miRNAs and 140 mRNAs were found showing a significant negative correlation ($P < 0.05$), after excluding positively correlated pairs. Based on 39 shared miRNAs among these regulatory pairs, the lncRNA–miRNA–mRNA (ceRNA) network was established (**Figure 3A**). The top 15 KEGG pathways (P-value < 0.05) that the network was involved are shown in **Figure 3B**, indicating that the ceRNA network for LUAD was closely related to some cancer-associated pathways, such as

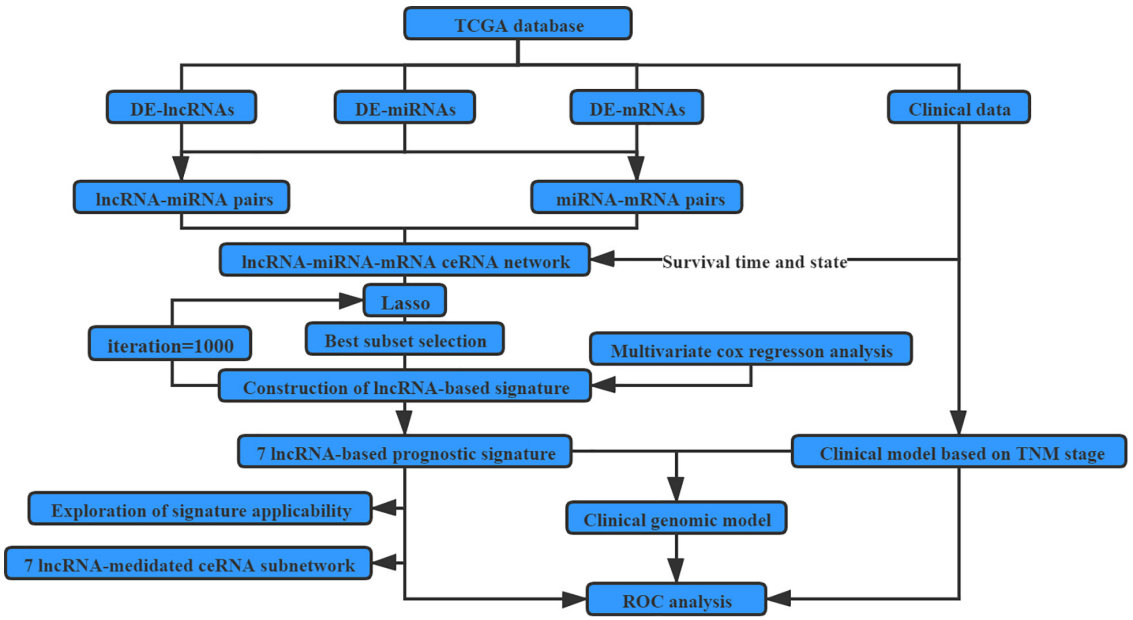


FIGURE 1 | Flow chart for identifying and assessing the survival-related LncRNA-based signature.

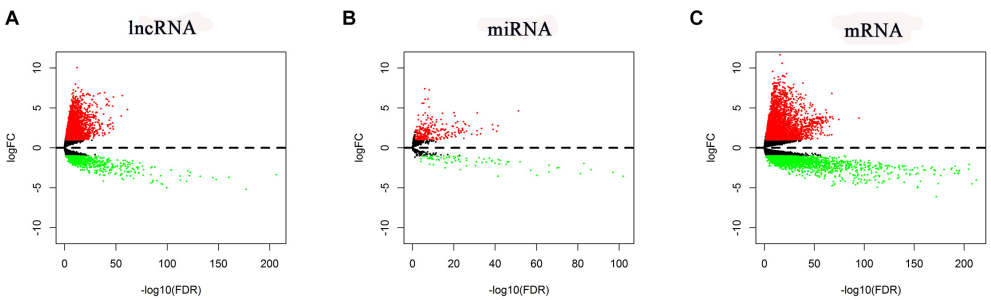


FIGURE 2 | Volcano plots of the DE-lncRNAs, DE-miRNAs, and DE-mRNAs. **(A)** DE-lncRNAs. **(B)** DE-miRNAs. **(C)** DE-mRNAs.

microRNAs in cancer, transcriptional misregulation in cancer, cellular senescence, cell cycle, p53 signaling pathway, small cell lung cancer, and so on.

The lncRNA-Based Prognostic Signature for LUAD

Among 1,000 lncRNA sets, constructed by using lasso-Cox regression analysis of 197 lncRNAs included in the ceRNA network for LUAD, five had AUC $>$ 0.7. The optimal set with minimal size was selected as the lncRNA-based prognostic signature for LUAD, which was consisted of seven lncRNAs (*SNHG*12, *DLEU7_AS1*, *FAM*41*C*, *FAM*181*A_AS*1, *AC*022148.1, *CCDC*13_*AS*1, and *LINC*00319). **Figures 4A,B** shows the convergence of the lasso-based variable selection (or called feature shrinkage) with the log of the penalty parameter lambda (λ) as well as the changes of model fitting statistics (partial likelihood deviance).

Then, a predictive model for the lncRNA set was constructed according to their lncRNA expression values and their corresponding coefficients derived from the multivariate Cox regression analysis. The risk score based on the lncRNA set was defined as following:

$$\begin{aligned}\text{Riskscore} = &(-0.2196 \times \textit{SNHG}12)\\ &+ (-0.14458 \times \textit{DLEU7\_AS}1)\\ &+ (-0.12641 \times \textit{FAM}41C)\\ &+ (-0.08658 \times \textit{FAM}181A\_\textit{AS}1)\\ &+ (-0.11915 \times \textit{AC}022148.1)\\ &+ (-0.15204 \times \textit{CCDC}13\_\textit{AS}1)\\ &+ (0.12557 \times \textit{LINC}00319)\end{aligned}$$

Based on the median value of the risk score, 504 patients were subdivided into high-risk and low-risk groups. The expression patterns (heat maps) for patients of two groups (low-risk group and high-risk group) are shown in **Figure 4C**,

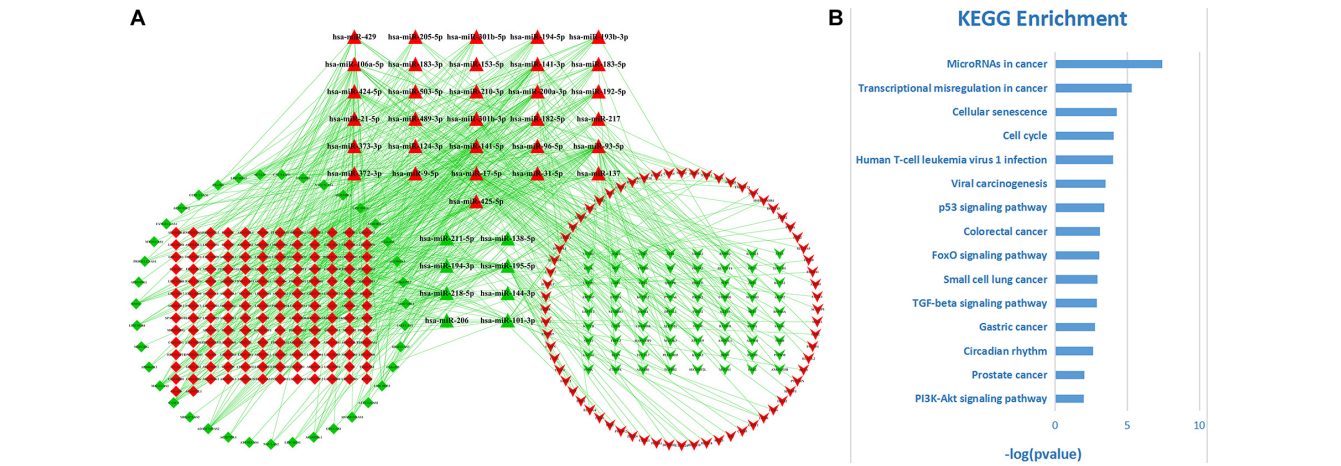


FIGURE 3 | The ceRNA network for LUAD and its functional involvements. **(A)** The topological structure of the ceRNA network for LUAD. Red nodes represent upregulation DE-RNAs and green ones downregulation DE-RNAs. Diamond, triangle, and inverse-triangle represent lncRNA, miRNA, and mRNA, respectively. **(B)** The top 15 KEGG pathways (*P*-value < 0.05) that the network was involved.

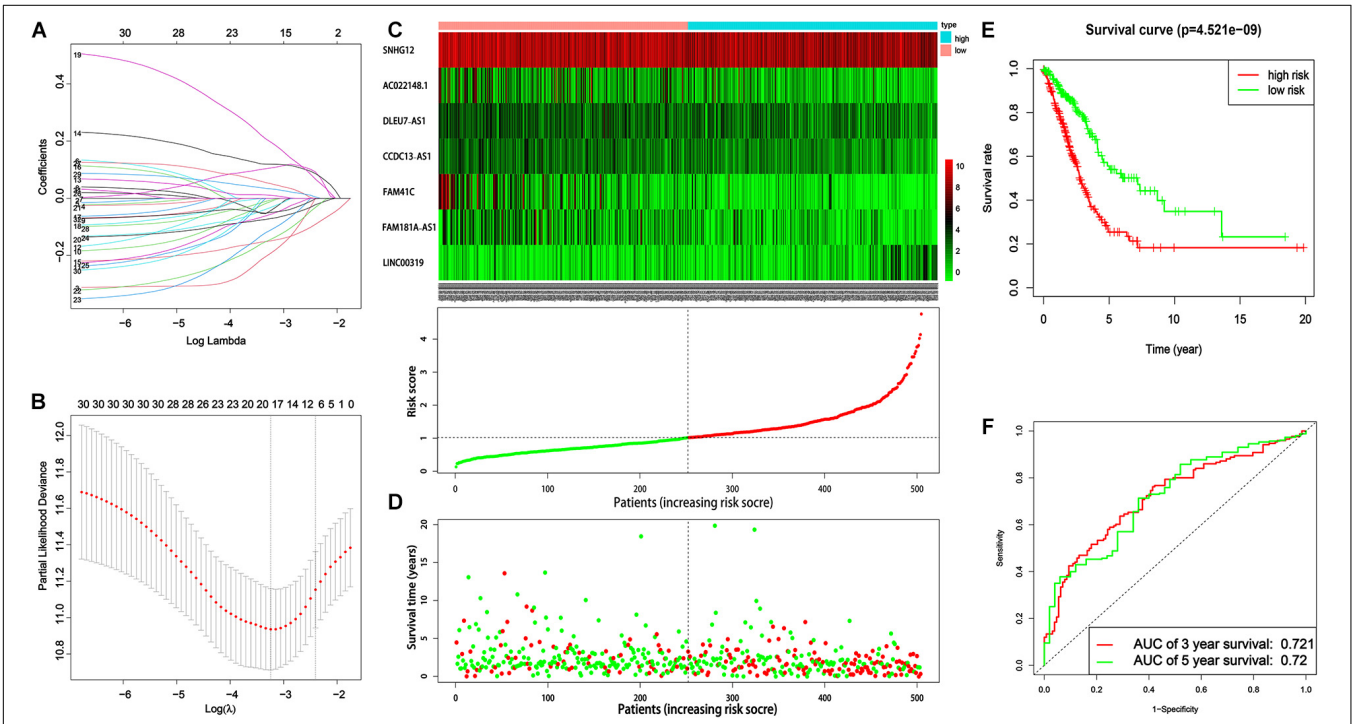


FIGURE 4 | The lncRNA-based prognostic signature for LUAD. **(A)** The convergence of the lasso-based variable selection with the log of the penalty parameter lambda (λ). **(B)** The changes of model fitting statistic, partial likelihood deviance, and its range, obtained from 10-fold cross-validation. **(C)** The expression patterns (heat maps) of the seven lncRNAs for patients of two groups (low-risk group, type low and high-risk group, type high). **(D)** The dot plot of survival time for patients sorted by the lncRNA-based risk score. Red dots indicate death events while green ones survivals. **(E)** Kaplan–Meier curves for patients of two categories (low-risk groups versus high-risk groups) defined by the lncRNA-based prognostic signature. **(F)** ROC analysis of the lncRNA-based prognostic signature to estimate AUC values of survival over two different periods.

while their risk score distribution, as well as their survival, data are shown in **Figure 4D**, which clearly indicated that with the increase of the risk score, LUAD patients tended to have a shorter survival time. Kaplan–Meier survival analysis showed that the patients in the low-risk group had a better prognosis than those with high-risk scores ($P = 4.52e^{-9}$) (**Figure 4E**). ROC analysis estimated that the AUC values of the lncRNA-based prognostic signature for survival were 0.721 for the third year and 0.72 for the fifth year, respectively (**Figure 4F**).

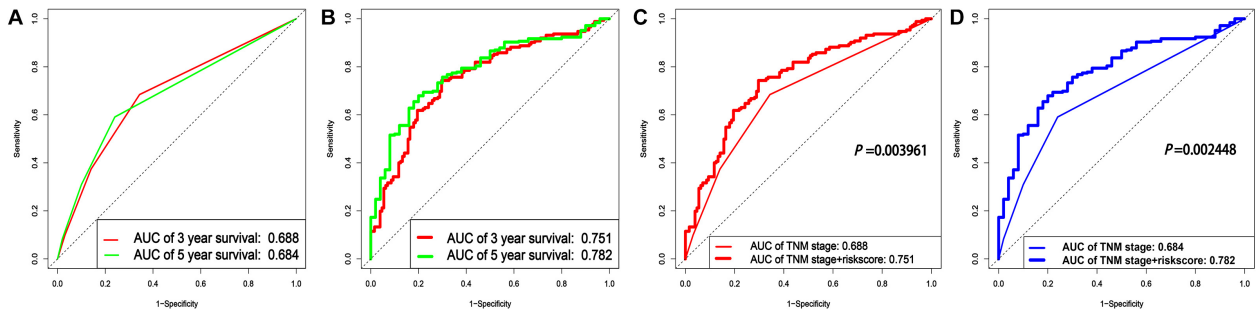


FIGURE 5 | A comparison between the clinical genomic model with both the TNM stages and the lncRNA signature and with the clinical model with the TNM stages only. **(A)** The ROCs of two periods of survival for the conventional model with the TNM stages only. **(B)** The ROCs of two periods of survival for the clinical genomic model with both the TNM stages and the lncRNA-based signature included. **(C,D)** Comparison of the two models for prognosis prediction over the two periods (3 and 5 years), respectively.

Comparison Between the Clinical Genomic Model With Both the TNM Stages and the lncRNA Signature Included and the Clinical Model With the TNM Stages Only

In order to verify whether combining the lncRNA-based signature with the TNM stages could improve the prognosis prediction for LUAD, a clinical genomic model with the TNM stages and the lncRNA-based signature combined was constructed, and its area under the ROC curve (AUC) was compared to the model with the TNM stages only. As shown in **Figure 5A**, the AUC values for the model with the TNM stages only were 0.688 and 0.684 for 3- and 5-year survival, respectively, which were markedly lower than the estimates for the clinical genomic model with both the TNM stages and the lncRNA-based signature (0.751 and 0.782 for the two periods, respectively) (**Figure 5B**). A comparison of the two models for prognosis prediction over the two periods separately all demonstrated that the clinical genomic model performed significantly better than the conventional model with the TNM stage (all $P < 0.01$) (**Figures 5C,D**). These results suggest that lncRNAs could provide additional information on the prognosis prediction for LUAD, and more importantly, this fact may render an earlier prognosis prediction for LUAD to be practical.

Applicability of the lncRNA-Based Prognostic Signature

In order to explore its applicability, a stratification analysis by the TNM stages (overall rating, divided into four ranks), gender, or age was performed. The results indicated that the lncRNA-based prognostic signature could work well in most strata of the TNM stages, gender, or age (**Figures 6A–C,E–H**), i.e., having a good capacity in separating the LUAD patients into high-risk and low-risk groups. For the TNM stages, stage IV was the only exception where no statistical significance was found between the low- and high-risk groups defined by the lncRNA-based signature (**Figure 6D**), which might be due to the small sample size ($n = 26$). These results suggested that the lncRNA-based prognostic signature was largely independent of tumor stage, gender, and age, which was agreed well with the multivariate Cox regression analysis where the lncRNA-based prognostic signature remained to be the most significant factor ($P < 0.001$) even after adjusting for several demographic and clinical factors (**Figure 6I**).

The ceRNA Subnetwork Related to the lncRNA-Based Signature

On basis of the seven lncRNAs contained in the prognostic signature for LUAD, a core network (**Figure 7A**) was extracted from the primary ceRNA network. KEGG-based functional enrichment analysis demonstrated that this lncRNA-mediated ceRNA subnetwork was highly involved in several cancer-associated signaling pathways (**Figure 7B**), which implicated that the seven lncRNAs play vital roles in the tumorigenesis mechanism of LUAD via regulating related gene expressions by competitively sponging several miRNAs.

DISCUSSION

As the most common malignancy, LUAD had an unfavorable 5-year survival rate at an advanced stage. Early detection and diagnosis was an important way to improve the LUAD patient's prognosis. Although it had been confirmed as an effective prognosis predictor for LUAD patients (Carter et al., 2018), the TNM staging system, which was founded on the anatomical information only, could not perfectly perform prognosis prediction. An increasing amount of evidence demonstrated that genetic disorders and alterations were of significance in tumorigenesis and the progression of LUAD, suggesting that molecular markers had great value in the prediction of overall survival of LUAD patients (Trimarchi et al., 2014; Timmer et al., 2016).

With the development of high-throughput sequencing technology, the roles of long non-coding RNAs (lncRNAs) in human cancers had received more and more attention. Previous studies demonstrated that lncRNAs play an important role in tumor proliferation, migration, and invasion (Zhang G. et al., 2018) and have potential value in applications to early

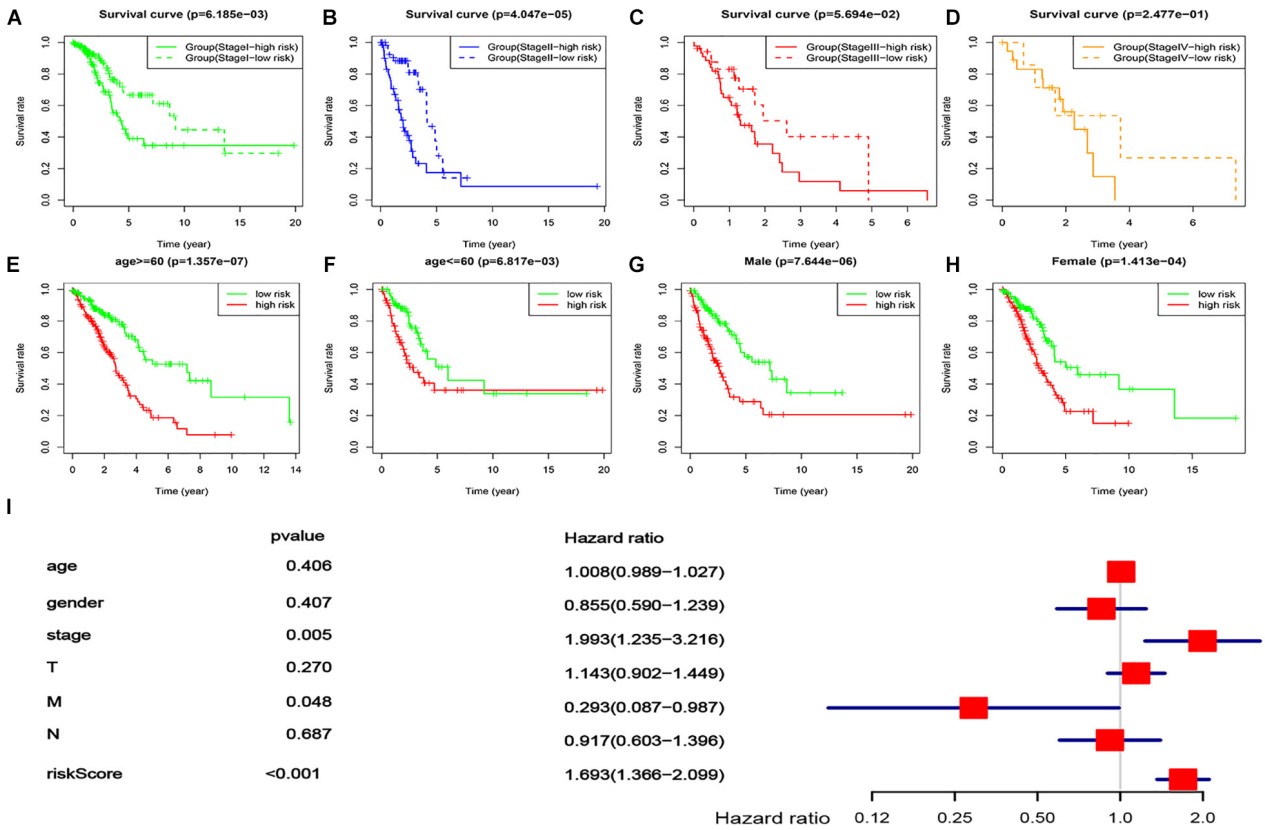


FIGURE 6 | Applicability of the lncRNA-based prognostic signature. **(A–D)** survival curves of the LUAD patients in different risk groups, stratified by four tumor stages [stage I **(A)**, stage II **(B)**, stage III **(C)**, and stage IV **(D)**]. **(E,F)** Survival curves of the LUAD patients in different risk groups, stratified by two ages [age ≤60 **(E)** and age <60 **(F)**]. **(G,H)** Survival curves of the LUAD patients in different risk groups, stratified by gender [male **(G)** and female **(H)**]. **(I)** Multivariate Cox regression with the lncRNA-based signature (risk score) and all available demographic and clinical factors included.

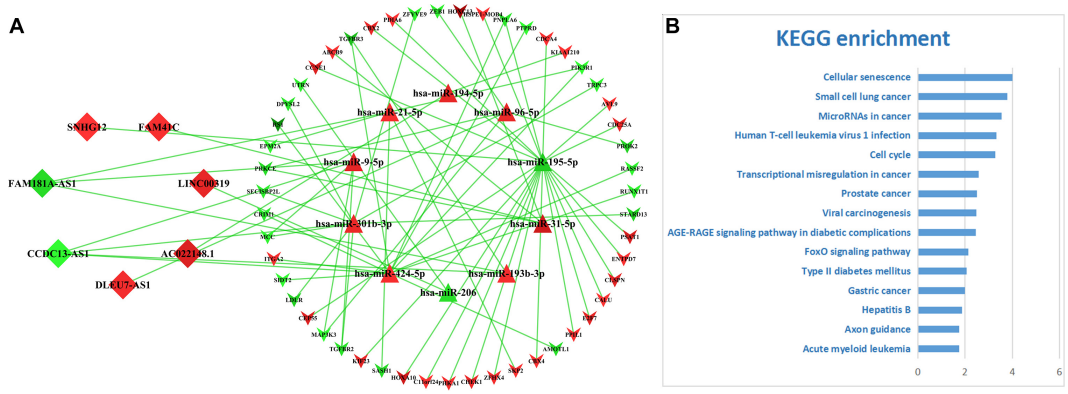


FIGURE 7 | The ceRNA subnetwork mediated by the lncRNA-based signature. **(A)** The ceRNA subnetwork depicting the relationships between seven lncRNAs and their counterparts. Diamonds represent lncRNAs, triangles miRNAs, and inverse-triangles mRNAs. The red color depicts upregulation and green downregulation. **(B)** The top 15 significantly enriched KEGG pathways.

prognosis and diagnosis for cancers (Takahashi et al., 2014). Zhai et al. (2017) demonstrated that the lower expression of lncRNA-*SARCC* influenced downstream genes such as *K-RAS*, *MMP-13*, *AKT*, and *P-ERK* expression by suppressing *miR-143-3p* expression, which could enhance RCC cell invasion, migration, and proliferation. Wu et al. (2017) first reported that lncRNA *CASC9*, as an oncogene, promoted ESCC cell growth by negatively regulating *PDCD4* expression via recruiting *EZH2*, which could be a potential diagnosis and prognosis biomarker for ESCC. All these facts highlight the importance of the large

number of lncRNAs to be the molecular biomarkers for early prognosis prediction or early diagnosis of cancers.

Compared with the coding RNAs, it was more complex to study the functional meanings of lncRNAs. The ceRNA hypothesis provided a new solution for achieving better functional studies of lncRNAs. In the ceRNA theory, lncRNAs regulate the expression of the targeted genes by competitively absorbing miRNAs at the post-transcriptional level, forming a huge ceRNA regulatory network (Calloni and Bonatto, 2019). In most scenarios, "communications" between ceRNAs and miRNAs were in dynamic balance (Cai and Wan, 2018). However, an abnormal expression of lncRNA destroyed the balance of the ceRNA network, which was closely related to tumorigenesis (Shao et al., 2015). In this study, a total of 3,939 DE-lncRNAs, 352 DE-miRNA, and 5,537 DE-mRNAs were identified. According to the ceRNA theory, the negatively correlated RNA–RNA regulatory pairs were built among DE-lncRNAs, DE-miRNA, and DE-mRNA. Subsequently, a lncRNA–miRNA–mRNA ceRNA network including 197 DE-lncRNAs, 39 DE-miRNAs, and 140 DE-mRNAs was constructed by connecting the negatively correlated RNA–RNA regulatory pairs. Further functional enrichment analysis showed that the ceRNA network was mainly involved in some cancer-related pathways including "microRNAs in cancer," "transcriptional misregulation in cancer," "cell cycle," "p53 signaling pathway," "colorectal cancer," "small cell lung cancer," etc., which was not surprising at all to us because more and more evidence indicated that at the molecular levels various cancers were interconnected. In brief, the ceRNA regulatory network including complex molecular regulatory relationships not only had potential value to mine prognosis-related biomarkers (Bai et al., 2019) but also provided a new avenue to broaden our knowledge on massive lncRNAs and their functional involvements in the pathogenic mechanisms for cancers like LUAD.

Most of the preceding studies focused on single lncRNAs related to lung cancer. Nie et al. (2016) found that *LncRNA-UCA1* upregulated *ERBB4* by sponging *miR-193a-3p* to exert oncogenic functions. Guan et al. (2019) demonstrated that *LINC00673-v4* enhanced cancer cell invasion, migration, and metastasis by overactivating WNT/β-catenin signaling and could be a candidate for the therapeutic target of LUAD patients. Nevertheless, as a complex disease, LUAD was thought to be a series of biological cascades resulting from the perturbations of intracellular and intercellular elements (Fiscon et al., 2018), and it was impossible to have a global picture about the sophisticated pathogenic mechanism of LUAD by studying only a single biomarker. Therefore, compared with a single biomarker, a prognostic signature integrating multiple biomarkers could achieve more power in prognosis prediction for LUAD (Kratz et al., 2019). This study demonstrated that the newly identified lncRNA-based signature with seven lncRNAs could provide >9% improvement in prognosis prediction over various periods for LUAD and was deemed to a robust complement for the conventional TNM staging system.

Among the seven lncRNAs containing the prognostic signature, four (*AC022148.1*, *DLEU7_AS1*, *LINC00319*, and *SNHG12*) were found to be involved in tumorigenesis, migration, and metastasis in cancers. Qi et al. (2019) demonstrated that a seven-lncRNA prognostic model including *AC022148.1* was a robust indicator to assess the prognosis risk of lung squamous cell carcinoma patients. Liu et al. (2018) found that *DLEU7_AS1*, as an adverse prognosis factor, was closely associated with colorectal cancer (CRC) staging, lymph node metastasis, and distant metastasis and may regulate the Wnt/β-catenin pathway to promote the occurrence and development of CRC. Zhou et al. (2017) demonstrated that *LINC00319* strengthened proliferation and invasion of lung cancer cells by downregulating the expression of miR-32 and upregulating the expression levels of miR-32 target genes. Wang et al. (2019a) revealed that knockdown of *SNHG12* inhibited migration and invasion of NSCLC cells via the Slug/zinc finger E-box-binding homeobox 2 EMT signaling pathway by upregulating the expression of miR-218 and could be a potential prognostic marker and therapeutic target for NSCLC. Up to date, there is dearth of information about the roles of the remaining three lncRNAs in tumors, waiting for further studies to clarify.

Compared with several previous studies, our approach has the following improvements. First, for constructing the ceRNA network, we used only negatively correlated lncRNAs/mRNAs–miRNAs regulatory pairs, which fits well the definition of the ceRNA network compared to the previous method without this distinction (Li et al., 2018; Fan et al., 2020; Wu et al., 2020). Second, we applied 1,000 rounds of lasso-Cox regression model fittings to identify the optimal lncRNA-based set, which is deemed to be more robust than the conventional Cox model utilized in a previous study (Zheng et al., 2017). Third, consequently, our study achieved better performance on predicting the survival of the LUAD patients by using a seven-lncRNA-based signature (AUC = 0.72 and 0.721 for 3- and 5-year survival, respectively) than a previous study (Yao et al., 2020), who identified an eight-lncRNA signature for LUAD (AUC = 0.702 and 0.671 for 3- and 5-year survival). Up to date, only the clinical genomic model built by the present study has achieved adequate improvement over the conventional TNM staging system (P = 0.003961 and 9.16% for the third year; P = 0.002448 and 14.33% for the fifth year), compared to a previous similar study (Zheng et al., 2017), who reported a non-significant improvement ($P > 0.05$ and 4.24% for the fifth year).

In conclusion, we applied an integrated ceRNA network analysis to identify a lncRNA-based signature for predicting the prognosis of LUAD patients. The established molecular signature with seven lncRNAs, derived from the ceRNA network, was demonstrated to be a robust and independent factor for the survival prediction of LUAD patients and, hence, could be an important complement for the conventional TNM staging system.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

AUTHOR CONTRIBUTIONS

SR, RL, and KH conceived and designed the study. RL, KH, SR, YL, SS, and SL performed data analysis. RL, XC, and DX contributed in the software and programming. RL, SR, KH, XC, and DX wrote, reviewed, and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 31071166 and 81373085 to SR) and the High-level Hospital Construction Research Project of Maoming People's Hospital (ZX2020013 to KH).

REFERENCES

Andreassen, O. A., Zuber, V., Thompson, W. K., Schork, A. J., Bettella, F., Djurovic, S., et al. (2014). Shared common variants in prostate cancer and blood lipids. *Int. J. Epidemiol.* 43, 1205–1214.doi: 10.1093/ije/dyu090

Bai, Y., Long, J., Liu, Z., Lin, J., Huang, H., Wang, D., et al. (2019). Comprehensive analysis of a ceRNA network reveals potential prognostic cytoplasmic lncRNAs involved in HCC progression. *J. Cell. Physiol.* 234, 18837–18848.doi: 10.1002/jcp.28522

Ball, D. (2019). TNM in non-small cell lung cancer: A staging system for all oncologists or just for surgeons? *Ann. Transl. Med.* 7:S103. doi: 10.21037/atm.2019.04.84

Cai, Y., and Wan, J. (2018). Competing endogenous RNA regulations in neurodegenerative disorders: current challenges and emerging insights. *Front. Mol. Neurosci.* 11:370. doi: 10.3389/fnmol.2018.00370

Calloni, R., and Bonatto, D. (2019). Characteristics of the competition among RNAs for the binding of shared miRNAs. *Eur. J. Cell Biol.* 98, 94–102.doi: 10.1016/j.ejcb.2019.04.001

Carter, B. W., Lichtenberger, J. R., Benveniste, M. K., de Groot, P. M., Wu, C. C., Erasmus, J. J., et al. (2018). Revisions to the TNM staging of lung cancer: rationale, significance, and clinical application. *Radiographics* 38, 374–391.doi: 10.1148/rg.2018170081

Del, C. A., Franchi, P., Contegiacomo, A., Cicchetti, G., Bonomo, L., and Larici, A. R. (2017). Missed lung cancer: When, where, and why? *Diagn. Interv. Radiol.* 23, 118–126.doi: 10.5152/dir.2016.16187

Denisenko, T. V., Budkevich, I. N., and Zhivotovsky, B. (2018). Cell death-based treatment of lung adenocarcinoma. *Cell Death Dis.* 9:117. doi: 10.1038/s41419-017-0063-y

Eissa, S., Safwat, M., Matboli, M., Zaghloul, A., El-Sawalhi, M., and Shaheen, A. (2019). Measurement of Urinary Level of a Specific Competing endogenous RNA network (FOS and RCAN mRNA/ miR-324-5p, miR-4738-3p, /lncRNA miR-497-HG) enables diagnosis of bladder cancer. *Urol. Oncol.* 37, 292.e19–292.e27. doi: 10.1016/j.urolonc.2018.12.024

Evans, J. R., Feng, F. Y., and Chinnaiyan, A. M. (2016). The bright side of dark matter: lncRNAs in cancer. *J. Clin. Invest.* 126, 2775–2782.doi: 10.1172/JCI84421

Fan, F., Ping, Y., Yang, L., Duan, X., Resegofetse, M. N., Li, B., et al. (2020). Characterization of a non-coding RNA-associated ceRNA network in metastatic lung adenocarcinoma. *J. Cell. Mol. Med.* 24, 11680–11690.doi: 10.1111/jcmm.15778

Fiscon, G., Conte, F., Farina, L., and Paci, P. (2018). Network-based approaches to explore complex biological systems towards network medicine. *Genes* 9:437.doi: 10.3390/genes9090437

Fouad, T. M., Barrera, A., Reuben, J. M., Lucci, A., Woodward, W. A., Stauder, M. C., et al. (2017). Inflammatory breast cancer: a proposed conceptual shift in the UICC-AJCC TNM staging system. *Lancet Oncol.* 18, e228–e232. doi: 10.1016/S1470-2045(17)30192-4

Guan, H., Zhu, T., Wu, S., Liu, S., Liu, B., Wu, J., et al. (2019). Long noncoding RNA LINC00673-v4 promotes aggressiveness of lung adenocarcinoma via activating WNT/beta-catenin signaling. *Proc. Natl. Acad. Sci. U.S.A.* 116, 14019–14028.doi: 10.1073/pnas.1900997116

Hu, J., Xu, L., Shou, T., and Chen, Q. (2019). Systematic analysis identifies three-lncRNA signature as a potentially prognostic biomarker for lung squamous cell carcinoma using bioinformatics strategy. *Transl. Lung Cancer Res.* 8, 614–635.doi: 10.21037/tlcr.2019.09.13

Huang, M., Hou, J., Wang, Y., Xie, M., Wei, C., Nie, F., et al. (2017). Long noncoding RNA LINC00673 is activated by SP1 and exerts oncogenic properties by interacting with LSD1 and EZH2 in Gastric Cancer. *Mol. Ther.* 25, 1014–1026.doi: 10.1016/j.ymthe.2017.01.017

Hutter, R. V. (1991). The role of the pathologist in the management of breast cancer. *CA Cancer J. Clin.* 41, 283–299.doi: 10.3322/canjclin.41.5.283

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, 2062–2063.doi: 10.1093/bioinformatics/bts344

Jinawath, N., Bunbanjerdsuk, S., Chayanupatkul, M., Ngamphaiboon, N., Asavapanumas, N., Svasti, J., et al. (2016). Bridging the gap between clinicians and systems biologists: from network biology to translational biomedical research. *J. Transl. Med.* 14:324. doi: 10.1186/s12967-016-1078-3

Karreth, F. A., and Pandolfi, P. P. (2013). ceRNA cross-talk in cancer: when ce-bling rivalries go awry. *Cancer Discov.* 3, 1113–1121.doi: 10.1158/2159-8290.CD-13-0202

Knoll, M., Lodish, H. F., and Sun, L. (2015). Long non-coding RNAs as regulators of the endocrine system. *Nat. Rev. Endocrinol.* 11, 151–160.doi: 10.1038/nrendo.2014.229

Kratz, J. R., Haro, G. J., Cook, N. R., He, J., Van Den Eeden, S. K., Woodard, G. A., et al. (2019). Incorporation of a molecular prognostic classifier improves conventional non-small cell lung cancer staging. *J. Thorac. Oncol.* 14, 1223–1232.doi: 10.1016/j.jtho.2019.03.015

Li, L., Peng, M., Xue, W., Fan, Z., Wang, T., Lian, J., et al. (2018). Integrated analysis of dysregulated long non-coding RNAs/microRNAs/mRNAs in metastasis of lung adenocarcinoma. *J. Transl. Med.* 16:372. doi: 10.1186/s12967-018-1732-z

Lin, P., Wen, D. Y., Li, Q., He, Y., Yang, H., and Chen, G. (2018). Genome-wide analysis of prognostic lncRNAs, miRNAs, and mRNAs forming a competing endogenous rna network in hepatocellular carcinoma. *Cell. Physiol. Biochem.* 48, 1953–1967.doi: 10.1159/000492519

Liu, X. B., Han, C., and Sun, C. Z. (2018). Long non-coding RNA DLEU7-AS1 promotes the occurrence and development of colorectal cancer via Wnt/beta-catenin pathway. *Eur. Rev. Med. Pharmacol. Sci.* 22, 110–117. doi: 10.26355/eurrev_201801_14107

Matsuda, T., and Machii, R. (2015). Morphological distribution of lung cancer from Cancer Incidence in Five Continents Vol. X. *Jpn. J. Clin. Oncol.* 45:404.doi: 10.1093/jjco/hyv041

Mittendorf, E. A., Ballman, K. V., McCall, L. M., Yi, M., Sahin, A. A., Bedrosian, I., et al. (2015). Evaluation of the stage IB designation of the American Joint Committee on Cancer staging system in breast cancer. *J. Clin. Oncol.* 33, 1119–1127.doi: 10.1200/JCO.2014.57.2958

Nie, W., Ge, H. J., Yang, X. Q., Sun, X., Huang, H., Tao, X., et al. (2016). LncRNA-UCA1 exerts oncogenic functions in non-small cell lung cancer by targeting miR-193a-3p. *Cancer Lett.* 371, 99–106.doi: 10.1016/j.canlet.2015.11.024

Pontius, L. N., Oyekunle, T. O., Thomas, S. M., Stang, M. T., Scheri, R. P., Roman, S. A., et al. (2017). Projecting survival in papillary thyroid cancer: a comparison of the seventh and eighth editions of the American joint commission on cancer/union for international cancer control staging systems in two contemporary national patient cohorts. *Thyroid* 27, 1408–1416.doi: 10.1089/thy.2017.0306

Qi, L., Zhang, T., Yao, Y., Zhuang, J., Liu, C., Liu, R., et al. (2019). Identification of lncRNAs associated with lung squamous cell carcinoma prognosis in the competitive endogenous RNA network. *PeerJ* 7:e7727.doi: 10.7717/peerj.7727

Rajer, M., Zwitter, M., and Rajer, B. (2014). Pollution in the working place and social status: co-factors in lung cancer carcinogenesis. *Lung Cancer* 85, 346–350.doi: 10.1016/j.lungcan.2014.06.012

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.doi: 10.1093/bioinformatics/btp616

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? *Cell* 146, 353–358.doi: 10.1016/j.cell.2011.07.014

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504.doi: 10.1101/gr.1239303

Shao, T., Wu, A., Chen, J., Chen, H., Lu, J., Bai, J., et al. (2015). Identification of module biomarkers from the dysregulated ceRNA-ceRNA interaction network in lung adenocarcinoma. *Mol. Biosyst.* 11, 3048–3058.doi: 10.1039/c5mb00364d

Smid, M., Coebergh, V. D. B. R., van de Werken, H., van Riet, J., van Galen, A., de Weerd, V., et al. (2018). Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. *BMC Bioinformatics* 19:236. doi: 10.1186/s12859-018-2246-7

Takahashi, K., Yan, I., Haga, H., and Patel, T. (2014). Long noncoding RNA in liver diseases. *Hepatology* 60, 744–753.doi: 10.1002/hep.27043

Taulli, R., Loretelli, C., and Pandolfi, P. P. (2013). From pseudo-ceRNAs to circ-ceRNAs: a tale of cross-talk and competition. *Nat. Struct. Mol. Biol.* 20, 541–543.doi: 10.1038/nsmb.2580

Timmer, M. R., Martinez, P., Lau, C. T., Westra, W. M., Calpe, S., Rygiel, A. M., et al. (2016). Derivation of genetic biomarkers for cancer risk stratification in Barrett's oesophagus: a prospective cohort study. *Gut* 65, 1602–1610.doi: 10.1136/gutjnl-2015-309642

Trimarchi, T., Bilal, E., Ntziachristos, P., Fabbri, G., Dalla-Favera, R., Tsirigos, A., et al. (2014). Genome-wide mapping and characterization of Notch-regulated long noncoding RNAs in acute leukemia. *Cell* 158, 593–606.doi: 10.1016/j.cell.2014.05.049

Tripathi, M. K., Doxtater, K., Keramatnia, F., Zacheaus, C., Yallapu, M. M., Jaggi, M., et al. (2018). Role of lncRNAs in ovarian cancer: defining new biomarkers for therapeutic purposes. *Drug Discov. Today* 23, 1635–1643.doi: 10.1016/j.drudis.2018.04.010

Wakeam, E., Acuna, S. A., Leighl, N. B., Giuliani, M. E., Finlayson, S., Varghese, T. K., et al. (2017). Surgery versus chemotherapy and radiotherapy for early and locally advanced small cell lung cancer: a propensity-matched analysis of survival. *Lung Cancer* 109, 78–88. doi: 10.1016/j.lungcan.2017.04.021

Wang, Y., Liang, S., Yu, Y., Shi, Y., and Zheng, H. (2019a). Knockdown of SNHG12 suppresses tumor metastasis and epithelial-mesenchymal transition via the Slug/ZEB2 signaling pathway by targeting miR-218 in NSCLC. *Oncol. Lett.* 17, 2356–2364. doi: 10.3892/ol.2018.9880

Wang, Y., Liu, X., Guan, G., Xiao, Z., Zhao, W., and Zhuang, M. (2019b). Identification of a five-pseudogene signature for predicting survival and its ceRNA network in Glioma. *Front. Oncol.* 9:1059. doi: 10.3389/fonc.2019.01059

Wu, X., Sui, Z., Zhang, H., Wang, Y., and Yu, Z. (2020). Integrated analysis of lncRNA-mediated ceRNA network in lung Adenocarcinoma. *Front. Oncol.* 10:554759. doi: 10.3389/fonc.2020.554759

Wu, Y., Hu, L., Liang, Y., Li, J., Wang, K., Chen, X., et al. (2017). Up-regulation of lncRNA CASC9 promotes esophageal squamous cell carcinoma growth by negatively regulating PDCD4 expression through EZH2. *Mol. Cancer* 16:150.

Yao, Y., Zhang, T., Qi, L., Liu, R., Liu, G., Wang, J., et al. (2020). Comprehensive analysis of prognostic biomarkers in lung adenocarcinoma based on aberrant lncRNA-miRNA-mRNA networks and Cox regression models. *Biosci. Rep.* 40:BSR20191554. doi: 10.1042/BSR20191554

Yin, H., Wang, X., Zhang, X., Wang, Y., Zeng, Y., Xiong, Y., et al. (2018). Integrated analysis of long noncoding RNA associated-competing endogenous RNA as prognostic biomarkers in clear cell renal carcinoma. *Cancer Sci.* 109, 3336–3349.doi: 10.1111/cas.13778

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287.doi: 10.1089/omi.2011.0118

Zhai, W., Sun, Y., Guo, C., Hu, G., Wang, M., Zheng, J., et al. (2017). LncRNA-SARCC suppresses renal cell carcinoma (RCC) progression via altering the androgen receptor(AR)/miRNA-143-3p signals. *Cell Death Differ.* 24, 1502–1517.doi: 10.1038/cdd.2017.74

Zhang, G., Li, S., Lu, J., Ge, Y., Wang, Q., Ma, G., et al. (2018). LncRNA MT1JP functions as a ceRNA in regulating FBXW7 through competitively binding to miR-92a-3p in gastric cancer. *Mol. Cancer* 17:87. doi: 10.1186/s12943-018-0829-6

Zhang, P., Cao, L., Fan, P., Mei, Y., and Wu, M. (2016). LncRNA-MIF, a c-Myc-activated long non-coding RNA, suppresses glycolysis by promoting Fbxw7-mediated c-Myc degradation. *EMBO Rep.* 17, 1204–1220.doi: 10.15252/embr.201642067

Zhang, Y., Li, X., Zhou, D., Zhi, H., Wang, P., Gao, Y., et al. (2018). Inferences of individual drug responses across diverse cancer types using a novel competing endogenous RNA network. *Mol. Oncol.* 12, 1429–1446. doi: 10.1002/1878-0261.12181

Zhao, X., and Liu, Z. P. (2019). Analysis of topological parameters of complex disease genes reveals the importance of location in a biomolecular network. *Genes* 10:143.doi: 10.3390/genes10020143

Zheng, H., Liu, J., Tycksen, E., Nunley, R., and McAlinden, A. (2019). MicroRNA-181a/b-1 over-expression enhances osteogenesis by modulating PTEN/PI3K/AKT signaling and mitochondrial metabolism. *Bone* 123, 92–102.doi: 10.1016/j.bone.2019.03.020

Zheng, S., Zheng, D., Dong, C., Jiang, J., Xie, J., Sun, Y., et al. (2017). Development of a novel prognostic signature of long non-coding RNAs in lung adenocarcinoma. *J. Cancer Res. Clin. Oncol.* 143, 1649–1657.doi: 10.1007/s00432-017-2411-9

Zhou, B., Yuan, W., and Li, X. (2017). Long intergenic noncoding RNA 319 (linc00319) promotes cell proliferation and invasion in lung cancer cells by directly downregulating the tumor suppressor MiR-32. *Oncol. Res.* doi: 10.3727/096504017X15016337254650

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 26 March 2021
doi: 10.3389/fgene.2021.659788

RNA-Associated Co-expression Network Identifies Novel Biomarkers for Digestive System Cancer

Zheng Chen[1,2†], Zijie Shen[2†], Zilong Zhang[2], Da Zhao[1,2], Lei Xu[3] and Lijun Zhang[1]**

[1] School of Applied Chemistry and Biological Technology, Shenzhen Polytechnic, Shenzhen, China, [2] Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China, [3] School of Electronic and Communication Engineering, Shenzhen Polytechnic, Shenzhen, China

Cancers of the digestive system are malignant diseases. Our study focused on colon cancer, esophageal cancer (ESCC), rectal cancer, gastric cancer (GC), and rectosigmoid junction cancer to identify possible biomarkers for these diseases. The transcriptome data were downloaded from the TCGA database (The Cancer Genome Atlas Program), and a network was constructed using the WGCNA algorithm. Two significant modules were found, and coexpression networks were constructed. CytoHubba was used to identify hub genes of the two networks. GO analysis suggested that the network genes were involved in metabolic processes, biological regulation, and membrane and protein binding. KEGG analysis indicated that the significant pathways were the calcium signaling pathway, fatty acid biosynthesis, and pathways in cancer and insulin resistance. Some of the most significant hub genes were *hsa-let-7b-3p*, *hsa-miR-378a-5p*, *hsa-miR-26a-5p*, *hsa-miR-382-5p*, and *hsa-miR-29b-2-5p* and *SECISBP2 L*, *NCOA1*, *HERC1*, *HIPK3*, and *MBNL1*, respectively. These genes were predicted to be associated with the tumor prognostic reference for this patient population.

Keywords: cancer, long non-coding RNA, WGCNA, MicroRNA, hub genes

OPEN ACCESS

Edited by:
Wei Jiang,
Nanjing University of Aeronautics and Astronautics, China

Reviewed by:
Yi Xiong,
Shanghai Jiao Tong University, China
Lei Deng,
Central South University, China

****Correspondence:***
Lei Xu
csleixu@szpt.edu.cn
Lijun Zhang
c7zlj@szpt.edu.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *28 January 2021*
Accepted: *25 February 2021*
Published: *26 March 2021*

Citation:
Chen Z, Shen Z, Zhang Z, Zhao D, Xu L and Zhang L (2021) RNA-Associated Co-expression Network Identifies Novel Biomarkers for Digestive System Cancer. Front. Genet. 12:659788. doi: 10.3389/fgene.2021.659788

INTRODUCTION

Cancer is one of the most common and lethal disease types in humans, and the molecular mechanisms governing cancer progression have not been elucidated to date (Liang et al., 2019; Yang et al., 2020; Zhang Z.M. et al., 2020). Tumors of the digestive system are a serious problem for human health, and esophageal tumors, colorectal tumors, gastric tumors and rectosigmoid-junction tumors are among the ten most common cancers worldwide. These cancers cause the most deaths, and the incidence rate is increasing year by year (Fitzmaurice et al., 2017, 2018, 2019; Bray et al., 2018; Doja et al., 2020; Hong et al., 2020; Li S. et al., 2020; Zeng et al., 2020a; Zhang L. et al., 2020). Therefore, the study of digestive system tumors has important practical significance.

Digestive system tumors are a common malignancy in the clinic. Colorectal cancer is the combination of colon and rectal cancer cases (Slattery et al., 2007; Ghosh and Yan, 2020). Colorectal cancer ranks third in men and second in women in the rankings of cancer incidence and is commonly diagnosed despite the increased death risk worldwide (Torre et al., 2015; Ozkan et al., 2019; Li J. et al., 2020). Colorectal cancer (CRC) is derived from the accumulation of epigenetic and genetic changes that lead to cancer-related deaths worldwide. With significant advancements in the early diagnosis of colorectal cancer, the cancer mortality rate has increased in the United States

(Rahman et al., 2015; Callahan et al., 2019; Ma et al., 2020). In European countries, colorectal cancer also has a high incidence, and the annual rate is increasing worldwide (Ferlay et al., 2015; Amer et al., 2020). Although the cancer-related death rate increases every year, treatments can be administered at the early stages of cancer (Pawa et al., 2011; Zhang Y. et al., 2019). Other digestive system tumors, such as esophageal cancer (ESCC), have been reported as the eighth most common cancer worldwide with a high cancer-related death rate (Ferlay et al., 2010; Abbas and Krasna, 2017). According to previous studies, almost 400,000 patients die from this disease among 450,000 diagnosed patients, the cancer-related death rate is approximately equal to 87.8% worldwide, and the 5-year survival rate is as low as approximately 25% (Rustgi and El-Serag, 2014, 2015). The most common cancer in Asia is gastric cancer (GC), which is the fifth ranking cancer leading to the third highest cancer-related death rate in the world. At stages I and IIA, the GC 5-year survival rate is 81.8–93.6%, but at stage IIIC, it is very low at only 17.9% (Amin et al., 2017; Lin et al., 2019).

Recently, network analysis has become a popular method for large-scale data (Cheng et al., 2019b; Zeng X. et al., 2019; Jin et al., 2020; Liu et al., 2020; Zeng et al., 2020b), and the WGCNA is a popular method for these coexpression network constructs (Jiang et al., 2010; Zeng W. et al., 2019; Iliopoulos et al., 2020). Five steps were included in this method: 1. a gene coexpression network was constructed by an adjacency matrix; 2. network modules were identified by the hierarchical cluster method; 3. modules with related phenotypes were analyzed; 4. module relationships were analyzed; and 5. key modules were found. The WGCNA method is widely used in biological networks and genomic data mining analysis, and also gives rise to network-based meta-analysis techniques (Langfelder et al., 2013).

The theoretical basis of the ceRNA (competitive endogenous RNA) hypothesis is that lncRNAs interact directly to regulate the expression of targeted genes and indirectly combine with miRNA sequences via the common miRNA response elements (MREs) of lncRNAs (Sardina et al., 2017). Cytoplasmic lncRNAs mainly affect mRNA reliability and translation mechanisms by binding miRNAs in the ceRNA regulatory network (Cao et al., 2018). Recently, analysis of lncRNA, miRNA, and mRNA networks has been identified to be involved in the progression of cancer (Zhang X. et al., 2017; Zeng et al., 2018), including endogenous cancer, hepatocellular carcinoma cancer and other malignant tumors, using the WGCNA method (Wang et al., 2010; Tay et al., 2014; Wei et al., 2014, 2018, 2019a,b; Thomson and Dinger, 2016; Jiang et al., 2018, 2019; Bai et al., 2019; Cheng et al., 2019a; Jin et al., 2019; Long et al., 2019; Shen et al., 2019). To date, there are few studies on digestive system malignant tumors. Therefore, integrated analysis of the regulatory functions of digestive system malignant tumors for lncRNA-miRNA-mRNA interaction networks requires large sample data and methods (Jiang et al., 2014; Zeng et al., 2016; Zou et al., 2016; Liu et al., 2017; Cheng et al., 2018; Cheng, 2019).

Competitive endogenous RNAs (ceRNAs) represent a novel gene expression regulation model and have attracted much attention from the academic community in recent years (Zhao et al., 2020). Compared with the miRNA regulation network, the ceRNA regulation network is more sophisticated and complex because the ceRNA regulation network includes more RNA molecules, such as mRNAs, pseudogenes, long non-coding RNAs (lncRNAs) and miRNAs. The ceRNA network provides a new perspective for transcriptomic and biological research.

In our studies, we used RNA-seq data from TCGA for free-scale gene coexpression network construction of digestive system malignant tumors. This study was able to provide novel biomarkers for these malignant tumors. Therefore, potential biomarkers were identified by using bioinformatics methods for integrative analysis based on the large amount of RNA-seq. According to the results of the analysis of malignant tumors, *NCOA1*, *HERC1*, *HIPK3*, *MBNL1*, *hsa-let-7b-3p*, *hsa-miR-378a-5p*, and *hsa-miR-26a-5p* were predictive or prognostic factors for malignant tumors.

DATA AND METHODS

Date Collection

For data analysis in this study, gdc-client 1.5.0 ww[1] and TCGAbiolinks 2.16.3 tools[2] (Colaprico et al., 2016) were used from the TCGA database. A total of 437 samples of mRNA data, 388 samples of miRNA data and 385 samples of clinical data for colon cancer were obtained from TCGA, including colon cancer samples and normal samples. For colon cancer, there were over 500 individuals, including 480 tumor and 41 normal samples. Eighty-seven samples of mRNA data, 97 samples of miRNA data, and 87 samples of clinical data for ESCC were obtained from TCGA, including esophageal cancer (ESCC) samples and normal samples. There were 90 samples of mRNA data, 79 samples of miRNA data and 87 samples of clinical data for rectal cancer obtained from TCGA, including ESCC samples and normal samples. A total of 373 samples of mRNA data, 452 samples of miRNA data and 406 samples of clinical data for gastric cancer (GC) were obtained from TCGA, including ESCC samples and normal samples. Sixty-six samples of mRNA data, 63 samples of miRNA data and 63 samples of clinical data for rectosigmoid junction cancer were obtained from TCGA, including ESCC samples and normal samples. The guidelines of this study are from the TCGA website[3].

Computational Analysis of RNA-Seq Data

The level three data of these cancers from TCGA by Illumina HiSeq 2000 platform (Illumina Inc., San Diego, California, United States) including miRNA, lncRNA, and mRNA data were analyzed by Digital Gene Expression Data package (edgeR 3.30.3)[4] and limma (limma 3.44.3)[5] package in R. The genes were annotated gene symbols through the Ensembl database[6]

[1]https://github.com/NCI-GDC/gdc-client/

[2]http://www.bioconductor.org/packages/release/bioc/html/TCGAbiolinks.html

[3]http://cancergenome.nih.gov/publications/publicationguidelines

[4]http://www.bioconductor.org/packages/release/bioc/html/edgeR.html

[5]http://www.bioconductor.org/packages/release/bioc/html/limma.html

[6]https://uswest.ensembl.org/index.html

(Zerbino et al., 2018). According to these analyses, absolute $\log_2$(fold change) $\geq$ 2.0 and FDR $\leq$ 0.01 were set to screen the significant differentially expressed genes of mRNAs (DEGs), lncRNAs (DElncRNAs) and miRNAs (DEmiRNAs) between tumor tissues and normal tissues. Non-cancer-specific expression genes were filtered, and upregulated or downregulated genes were saved. The package ggplot2 3.3.2[7] was used in RStudio to generate volcano plots of upregulated or downregulated genes. Meanwhile, the VennDiagram R package[8] was used for the Venn diagram of the datasets.

Construction of the Coexpression Network

The WGCNA package was used to analyze the data as described previously (Langfelder and Horvath, 2008). For the different conditions of tumor tissues, some samples may deviate from the actual situation. These tissue samples will affect the accuracy of the results. Therefore, these samples were removed before sample analysis. The function of goodSamplesGenes was used to verify if there were many missing values for each sample and to remove these samples from the total data. Then, the function hculst was used for cluster analysis of tissue samples. Appropriate soft thresholds need to be selected to build the network when the WGCNA method is performed. The soft threshold value should not only reach a scale-free fitting index of more than 0.9 but also requires greater mean connectivity of the network. Thus, 4 was selected for the soft threshold value. A good soft threshold is selected based on this rule. At the same time, the size of the gene module will be divided, and at least 30 genes will be identified for one module. If the correlation coefficient is greater than 0.75 for two modules, then these two modules will be merged into one module. To enhance the productivity of the modules, a cutoff (< 0.25) was selected to combine similar modules. After these modules were divided, the sample traits were associated with each module such that the most important associated modules for the traits were searched. Cytoscape 3.8.1 was used for the visualization of gene coexpression.

Gene Ontology and Pathway Enrichment Analysis

The DAVID online tool was applied for the GO (Gene Ontology) and KEGG analysis and functional annotation of DEmRNAs. The DAVID[9] database website was used for annotation and GO enrichment. KEGG pathways and GO terms enriched adopt adjusted P-value $\leq$ 0.05. WebGestalt[10] is used for enrichment analysis that supports three well-established and complementary methods for enrichment analysis. The genes of the network were analyzed by the WebGestalt website. The R package of GOplot (version 1.0.2) was used for the GO analysis of the DEmRNAs of the five cancers.

[7]https://cloud.r-project.org/package=ggplot2

[8]https://cran.r-project.org/web/packages/VennDiagram/

[9]http://david.abcc.ncifcrf.gov/

[10]http://www.webgestalt.org/

Hub Gene Identification and Validation

To identify the gene connectivity, Pearson's correlation was used for the test. In general, hub genes existed in modules closely linked to traits (cor.geneTraitSignificance $>$ 0.2), the modules with hub genes had high connectivity (cor.geneModuleMembership $>$ 0.8), and the packages of cytoHubba and MCODE in Cytoscape 3.8.1 were used to search the hub genes.

RESULTS

Identification and Validation of Differentially Expressed mRNAs, miRNAs, and lncRNAs

To identify the differentially expressed mRNAs, lncRNAs and miRNAs, the standard of $\log_2$(fold change) $\geq$ 1 and q-value $<$ 0.05 was used, and the EdgeR package of R (Law et al., 2016; Maza, 2016; Chen et al., 2017) was adopted to calculate separately for the normal samples and cancer samples of five cancers. For colon cancer, 4982 DEmRNAs, 149 DEmiRNAs and 1636 DElncRNAs were identified. There were 1409 differentially expressed genes in lncRNAs, 78 DEmiRNAs and 3659 differentially expressed mRNAs genes found between healthy and cancer-treated samples for esophageal cancer. A total of 3337 DEmRNAs, 168 DEmiRNAs and 871 DElncRNAs were identified between the rectosigmoid junction cancer tissues and matched normal control tissues; in addition, 2438 DEmRNAs, 156 DEmiRNAs and 607 DElncRNAs were identified for rectal cancer, and 4325 DEmRNAs, 62 DEmiRNAs and 2120 DElncRNAs were identified for gastric cancer compared with normal control samples. The distributions of differentially expressed DEmRNAs, DEmiRNAs and DElncRNAs were identified through volcano plots (**Figure 1**). A number of differences were found, and 87 DElncRNAs, 6 DEmiRNAs and 602 DEmRNAs were shared among the five cancers as shown in **Figure 2**.

GO and KEGG Pathway Analysis of DEmRNAs

GO and KEGG methods were adopted to investigate the annotation of DEmRNAs for the five cancers (**Figures 3**, **4**). Standards of $\log_2$(fold change) $\geq$ 3 and q-value $<$ 0.01 were used for the analysis.

A total of 766 DEmRNAs were primarily enriched in biological processes (BP) such as organic anion transport for colon cancer (GO:0015711) and cell components (CC) such as the apical part of the cell (GO:0045177) and the apical plasma membrane (GO:0016324) as well as ligand activity (GO:0048018) and signaling receptor activator activity (GO:0030546) of molecular function (MF) receptor for colon cancer (**Figure 3**). In addition, KEGG enrichment analysis revealed that the significant pathways were "bile secretion," "drug metabolism - cytochrome P450," "chemical carcinogenesis," "metabolism of xenobiotics by cytochrome P450" and "retinol metabolism," which are associated with the progression of colon cancer (**Figure 4**).

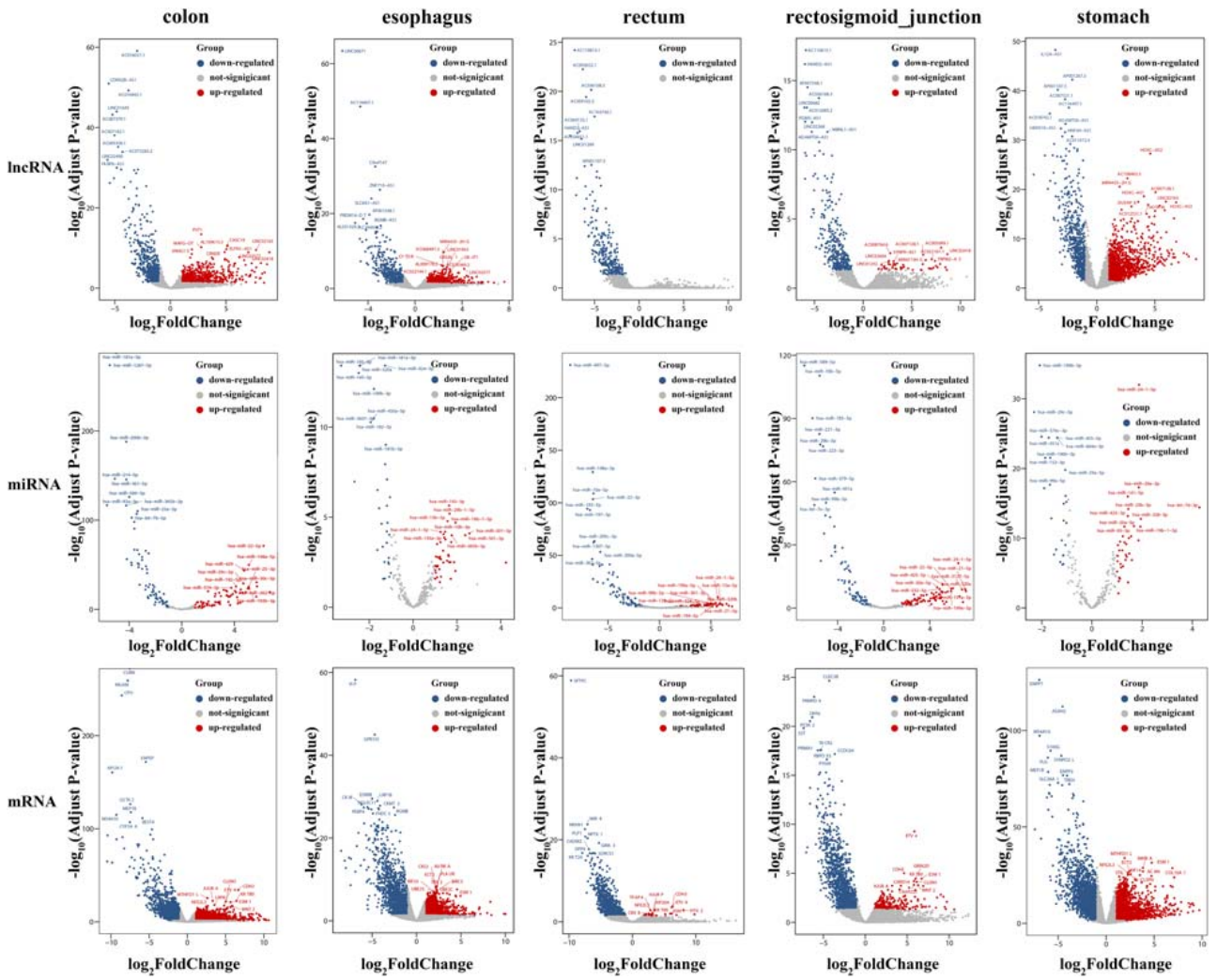


FIGURE 1 | Volcano plots of DEmRNAs, DEmiRNAs and DElncRNAs in five cancers. The volcano map of differentially expressed lncRNAs (top), miRNAs (middle), and mRNAs (bottom). Red and blue spots represent significant up- and down-regulated RNAs, respectively. The remark for the gene symbols represent the significant up- and down-regulated RNAs.

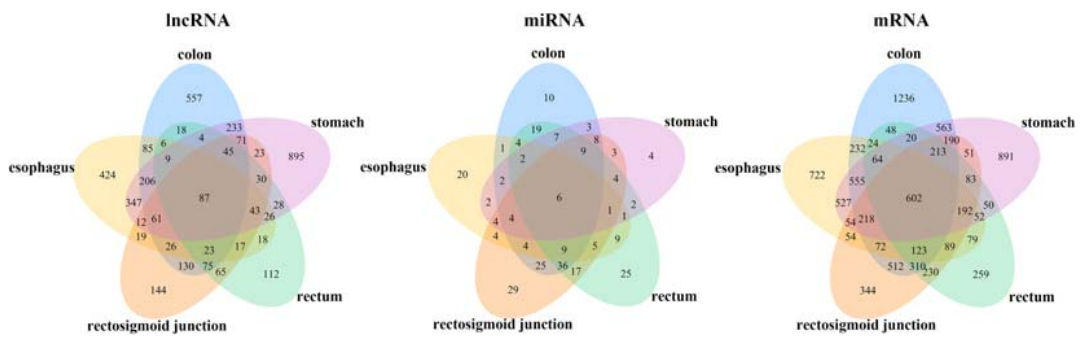


FIGURE 2 | Venn diagram of the DEmRNAs, DEmiRNAs, and DElncRNAs in five cancers. Different colors indicate the different cancers. The numbers represent the common RNAs among different cancers. The color of blue, orange, pink, green and yellow represent for colon, rectosigmoid junction, stomach, rectum, and esophagus cancer, respectively.

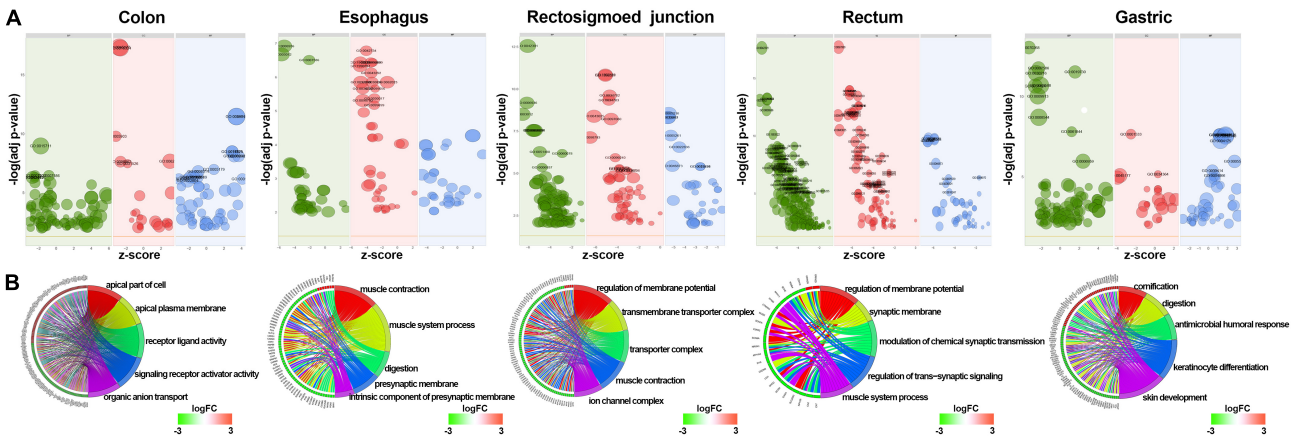


FIGURE 3 | Identification of differentially expressed RNAs and GO enrichment of DEmRNAs. **(A)** Bubble plot of GO enrichment for cancers. Green represents biological processes; red represents cellular components; and blue represents molecular functions. **(B)** The top 5 significantly enriched GO pathways and relevant genes.

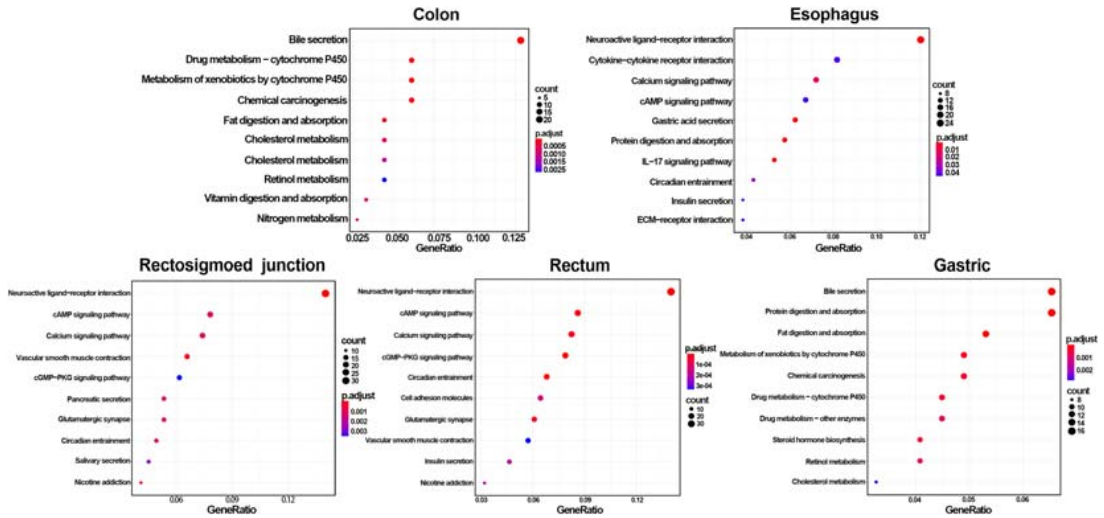


FIGURE 4 | Identification of differentially expressed RNAs and KEGG enrichment of DEmRNAs. The color of the bubble represent the significant of p.adjust.

To investigate the enrichment in esophageal cancer, 448 DEmRNAs were used and enriched in muscle contraction (GO:0006936), the muscle system process (GO:0003012), digestion (GO:0007586), the presynaptic membrane (GO:0042734), the intrinsic component of the presynaptic membrane (GO:0098889) of biological processes (BP) and cell components (CC) (**Figure 3**). In addition, "gastric acid secretion," "neuroactive ligand-receptor interaction," "protein digestion and absorption," "IL-17 signaling pathway," and "calcium signaling pathway" were the top 5 pathways of the KEGG analysis (**Figure 4**).

Rectosigmoid junction cancer was also analyzed by GO and KEGG enrichment; 623 DEmRNAs were screened out, and the top 5 GO enrichments were the regulation of membrane potential (GO:0042391), the transmembrane transporter complex (GO:1902495), the transporter complex (GO:1990351), muscle contraction (GO:0006936), and the ion channel complex (GO:0034702) of biological processes (BP) and cell components (CC) (**Figure 3**). The top 5 KEGG pathways were "neuroactive ligand-receptor interaction," "nicotine addiction," "vascular smooth muscle contraction," "pancreatic secretion," and "cAMP signaling pathway" (**Figure 4**).

Rectal cancer was also analyzed by GO and KEGG enrichment; 691 DEmRNAs were screened out, and the top 5 GO enrichments were the regulation of membrane potential (GO:0042391), the regulation of transsynaptic signaling (GO:0099177), the synaptic membrane (GO:0097060), the modulation of chemical synaptic transmission (GO:0050804), and the muscle system process (GO:0003012) of biological processes (BP) and cell components (CC) (**Figure 3**). The top 5 KEGG pathways were "neuroactive ligand-receptor

interaction," "circadian entrainment," "cGMP-PKG signaling pathway," "glutamatergic synapse" and "cAMP signaling pathway" (**Figure 4**).

Gastric cancer was also subjected to GO and KEGG enrichment analyses. A total of 637 DEmRNAs were screened out, and the top 5 GO enrichments were cornification (GO:0070268), digestion (GO:0007586), antimicrobial humoral response (GO:0019730) (Ji et al., 2019; Nadia and Ramana, 2020), keratinocyte differentiation (GO:0030216), and skin development (GO:0043588) of biological processes (BP) (**Figure 3**). The top 5 KEGG pathways were "Protein digestion and absorption," "Bile secretion," "Fat digestion and absorption," "Chemical carcinogenesis," and "Metabolism of xenobiotics by cytochrome P450" (**Figure 4**).

Construction of the WGCNA Network

To explore further the key biological genes related to cancers, the WGCNA method was applied to select the DEmRNAs, DEmiRNAs and DElncRNAs. For further analysis, a soft threshold (β = 4) was adopted to guarantee a scale-free network with high scale independence and low mean connectivity (near 0) (**Figure 5A**). DEGs from the five cancers were divided into several modules by cluster analysis (**Figure 5B**). Approximately 43 modules were generated for the five cancers. The module trait relationship is shown in **Figure 5C**. The dark green module related to the colon, rectosigmoid junction and rectal tumors were the deepest (cor = 0.21, 0.16, and 0.52 P = 4E-11, 6e-07, and 4e-67). For the other two cancers, the esophageal and gastric cancers, the brown module was chosen according to the correlation (cor = 0.46 and 0.46 P = 4E-52, 3e-52). Dark green and brown modules were selected for further analysis (**Figure 5C**). The high correlation and high P-value indicated that these modules are suitable for hub gene identification in these cancers.

Identification of the Hub Genes in the Five Cancers

The dark green module was selected for the edges signifying the correlations in colon cancer, rectosigmoid junction cancer and rectal cancer with the filter criterion of a weight value greater than 0.02 by the WGCNA algorithm. A total of 629 edges and 61 nodes were obtained and input into Cytoscape 3.8.1 (**Figure 6A**). A network was generated by Cytoscape software, and the hub genes were obtained from the network by applying CytoHubba. The most significant hub gene network was discovered by CytoHubba, as shown in **Figure 6B**. The identified hub miRNAs included *hsa-let-7b-3p*, *hsa-miR-378a-5p*, *hsa-miR-26a-5p*, *hsa-miR-382-5p*, and *hsa-miR-29b-2-5p*. In addition, the brown module was selected for the edges signifying the correlations in esophageal cancer and gastric cancer with the filter criterion of a weight value greater than 0.12 by the WGCNA algorithm. A total of 1027 edges and 279 nodes were obtained and input into Cytoscape 3.8.1 (**Figure 6C**). The most significant hub genes were discovered by CytoHubba as shown in **Figure 6D**. The identified hub genes were *SECISBP2 L*, *NCOA1*, *HERC1*, *HIPK3*, and *MBNL1*.

We discover that most of the genes in dark green module that were significantly correlated with colon cancer, rectosigmoid junction cancer and rectal cancer were miRNAs. This dominance of miRNAs in dark green module suggests the possibility that lncRNAs play a significant role in these cancers through regulation of coding genes in key pathways. miRNA can act by guiding histone modifiers and chromatin modifiers to regulate transcription and play crucial roles in cell differentiation that ultimately determine cell fate.

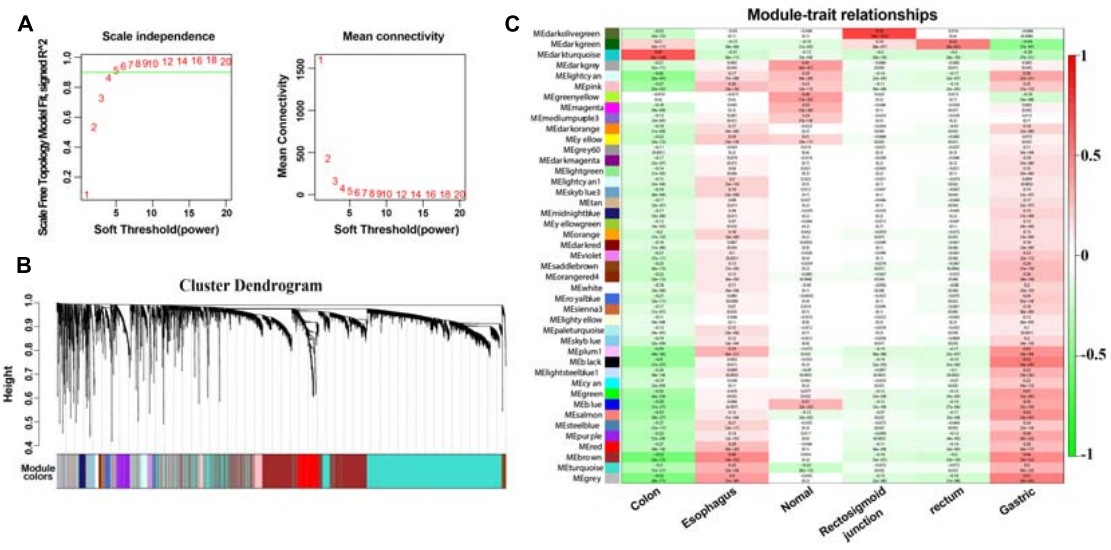


FIGURE 5 | WGCNA method and module identification of the cancers. **(A)** Scale independence and mean connectivity of soft-thresholding values (β). **(B)** Cluster dendrogram of filtered genes. **(C)** Correlation between the cancers and MEs.

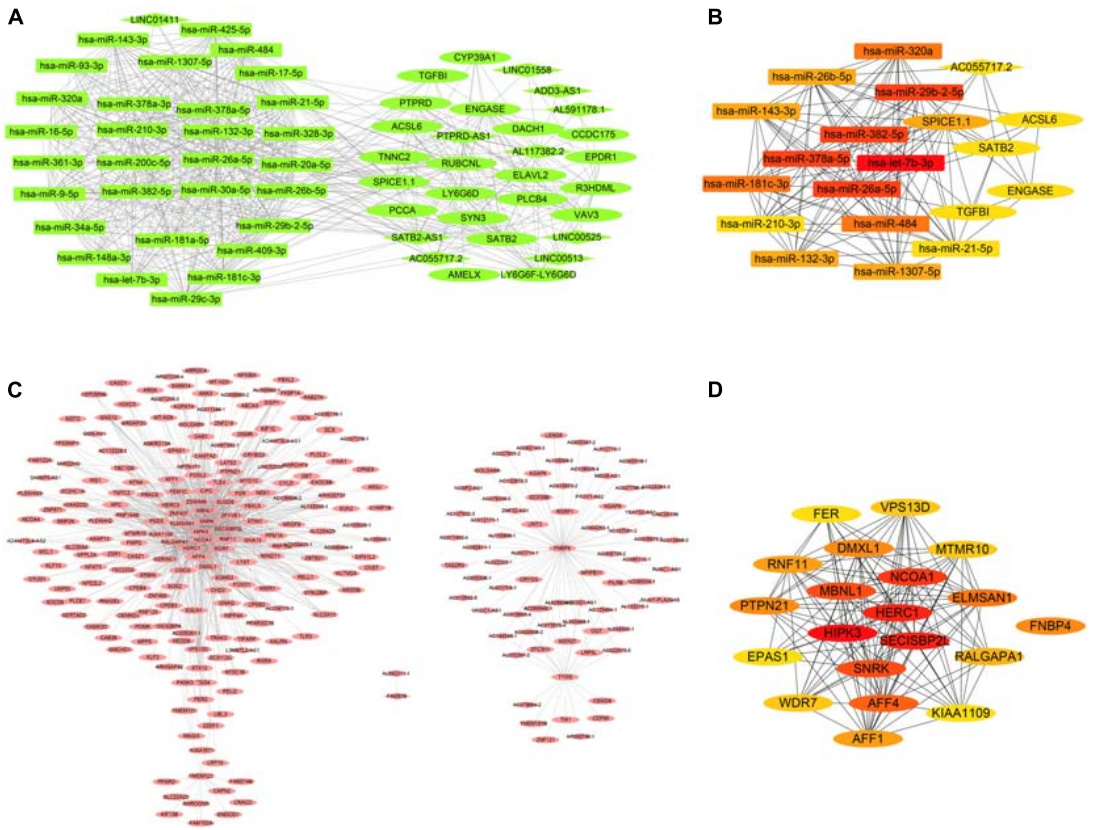


FIGURE 6 | Coexpression network and hub gene network. **(A)** The significant gene coexpression network of colon cancer, rectosigmoid junction cancer, and rectal cancer in the dark green module. **(B)** The most significant hub gene network generated by CytoHubba from the dark green module. **(C)** The significant gene coexpression network of esophageal cancer and gastric cancer in the brown module, which has 1027 edges and 279 nodes. **(D)** The most significant hub gene network generated by CytoHubba from the brown module.

Functional Annotation of the Module of Interest and the Network Genes

GO and KEGG analyses of the two modules of dark green and brown with 61 and 279 genes, respectively, in WebGestalt are shown in **Figure 7**. These were involved in molecular and functional biological processes and cellular components. For the dark green module, the genes were associated with metabolic process, membrane and protein binding, membrane among others (**Figure 7A**), and the genes in the brown module were enriched in biological regulation, membrane, and protein binding among others (**Figure 7B**). Analysis of KEGG pathways for the dark green module, the calcium signaling pathway, fatty acid biosynthesis and the chemokine signaling pathway were significant among all the pathways (**Figure 7C**), and pathways in cancer, insulin resistance and renal cell carcinoma were significant pathways in the brown module (**Figure 7D**).

DISCUSSION

Cancer of the digestive system is a common cancer worldwide. Recently, RNAs have been reported to be related to cancers. The WGCNA method was applied to analyze the related modules with different tumors and to identify the important genes of these tumors.

ESCC remains a serious burden on the health system worldwide, and some research has suggested that RNAs are associated with ESCC. For example, 2,046 circRNAs were frequently altered in ESCC tissues (Song et al., 2019). Several induced circRNAs were identified in radioresistant ESCC cells compared to normal ESCC cells (Su et al., 2016). Eight lncRNAs could be practical and reliable prognostic tools for esophageal cancers (Li W. et al., 2020). Additionally, the special functions of RNAs have been clearly demonstrated. For the analysis in gastric cancer (GC), 9 miRNAs, 41 lncRNAs, and 10 mRNAs were suggested and significant for GC (Pan et al., 2019). In addition, 15 lncRNAs were identified to be significantly related to the clinical features of colon cancer (Qian et al., 2019). In rectal cancer, COL1A1 and MZB1, as the key genes of rectal cancer, can interact with other genes correlated with shorter survival for patients (Wu et al., 2019). However, these studies presented research on only one cancer for the digestion system. Therefore, there is still an urgent need to identify the hub genes in digestive system cancers.

In this study, data from the TCGA website were adopted, and DEmRNAs, 168 DEmiRNAs, and 871 DElncRNAs were identified for the different tumors with the RNA-sequencing

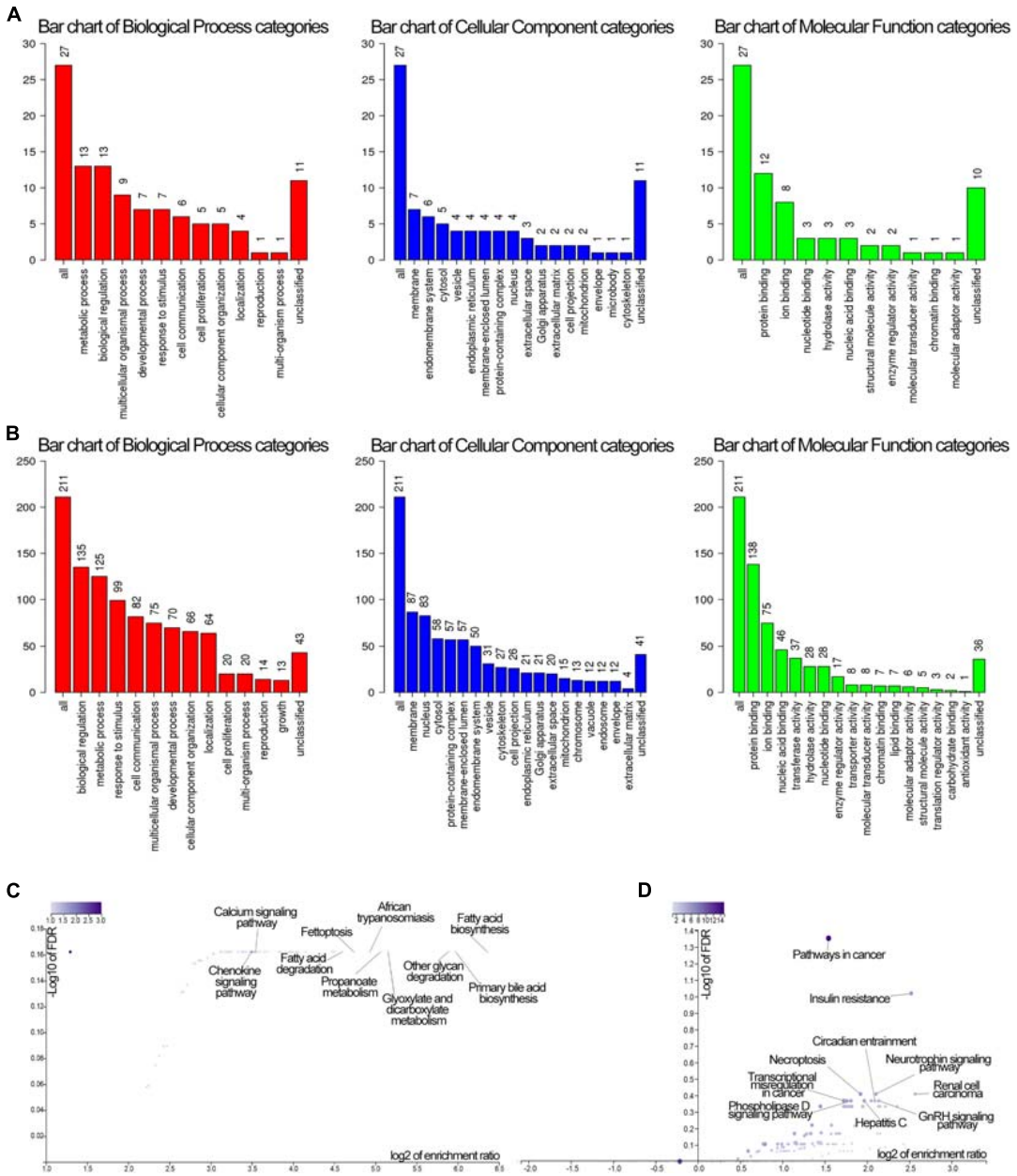


FIGURE 7 | GO and KEGG enrichment of genes in the dark green and brown modules. **(A)** GO enrichment of the darkgreen module. Red color stand for biological process, blue stand for cellular component, while green stand for molecular function. **(B)** GO enrichment of the brown module. Red color stand for biological process, blue stand for cellular component, while green stand for molecular function. **(C)** KEGG pathway of the dark green module. **(D)** KEGG pathway of the brown module. The GO enrichment vertical axis suggests the number of enriched genes.

method and integrated bioinformatics analysis. Based on the WGCNA algorithm, the modules associated with different tumors were generated under the appropriate thresholds. Therefore, a network related to the different cancers was constructed in this study. Based on the network analysis, two modules, dark green and brown, were selected for further research. CytoHubba from Cytoscape 3.8.1 was used for this analysis, and several hub genes were chosen for the prediction.

Twenty hub genes were associated with the staging of esophageal cancer and gastric cancer, such as *SECISBP2 L*, *NCOA1*, *HERC1*, *HIPK3*, and *MBNL1*. *SECISBP2 L*, which encodes the SECIS binding protein 2 (SBP2), is a selenoprotein that exists in yeast, fungi and higher plants (Donovan and Copeland, 2009). However, the molecular mechanism of this protein has not been elucidated. *NCOA1* is nuclear receptor coactivator 1 (also known as SRC-1) belonging to the SRC gene family (Oñate et al., 1995; Xu et al., 2009). The abnormal

expression level of SRCs has triggered cancers, such as breast, endometrial and ovarian cancers. The *SRC-3* and *SRC-1* genes have high expression levels in breast cancer. *SRC-1* regulates polyoma enhancer activator 3 (PEA3) and promotes breast-to-lung metastasis via transcription factors (Redmond et al., 2009; Walsh et al., 2012; Qin et al., 2015). HERC1 is a HECT and RLD domain containing E3 ubiquitin protein ligase family member 1 in humans. E3 ubiquitin ligases play a critical role in catalyzing ubiquitin transfer from E2 enzymes to the substrate in the ubiquitylation system. Similar to most E3 ubiquitin ligases, HERC1 has been shown to be a potential target for cancer therapy (Schneider et al., 2018; Dong et al., 2020). Homeodomain-interacting protein kinase 3 (*HIPK3*) is encoded by the *HIPK3* gene in humans. In NSCLC tissues, low expression of *HIPK3* was associated with poor survival rates. Therefore, *HIPK3* is considered a valuable biomarker for the survival of NSCLC patients (Liu et al., 2018). Muscleblind-like splicing regulator 1 (MBNL1), encoded by the *MBNL1* gene, is an RNA splicing protein and essential for MLL-rearranged leukemia cell growth (Itskovich et al., 2020). The isoforms of the MBNL1 protein affect cancer development and are targets for drug development (Li Z. et al., 2020). The results of the network indicated that these genes are also related to esophageal cancer and gastric cancer, and further research is necessary to determine the molecular mechanisms of these genes.

We eventually obtained 20 miRNAs associated with the staging of esophageal cancer and gastric cancer (Jiang et al., 2009), including *hsa-let-7b-3p*, *hsa-miR-378a-5p*, *hsa-miR-26a-5p*, *hsa-miR-382-5p*, and *hsa-miR-29b-2-5p*. Previous studies indicate that gastric cancer stem-like cells are derived from MKN-45- by the expression change of *hsa-let-7b-3p* (Salehi et al., 2018). Moreover, *microRNA-378a-5p* increased the expression level in melanoma cells and is a novel positive regulator of melanoma progression (Tang et al., 2018; Tupone et al., 2020). The expression of *hsa-miR-26a-5p* was downregulated in breast cancer tissues and appeared to have a poor prognosis for patients (Huang et al., 2019). However, overexpression *of hsa-miR-382-5p* increased oral squamous cell carcinoma cell invasion and migration (Sun et al., 2019). These studies suggested that miRNAs are also biomarkers for various cancers; however, the prediction of the interaction network indicated that these miRNAs are also related to colon cancer, rectosigmoid junction cancer and rectal cancer (Jiang et al., 2013).

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga.

AUTHOR CONTRIBUTIONS

LX and LZ designed the research. ZC and ZS performed the research. ZC, ZS, ZZ, and DZ analyzed the data. ZC wrote the manuscript. All authors read and approved the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 62002244, 62002242, and 62001311).

REFERENCES

Abbas, G., and Krasna, M. (2017). Overview of esophageal cancer. *Annals of Cardiothoracic Surgery* 6, 131–136. doi: 10.21037/acs.2017.03.03

Amer, A., Nagah, A., Tian, T., and Zhang, X. (2020). Mutation Mechanisms of Breast Cancer among the Female Population in China. *Current Bioinformatics* 15, 253–259. doi: 10.2174/1574893615666191220141548

Amin, M. B., Greene, F. L., Edge, S. B., Compton, C. C., Gershenwald, J. E., Brookland, R. K., et al. (2017). The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA: A Cancer Journal for Clinicians* 67, 93–99. doi: 10.3322/caac.21388

Bai, Y., Long, J., Liu, Z., Lin, J., Huang, H., Wang, D., et al. (2019). Comprehensive analysis of a ceRNA network reveals potential prognostic cytoplasmic lncRNAs involved in HCC progression. *Journal of Cellular Physiology* 234, 18837–18848. doi: 10.1002/jcp.28522

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68, 394–424. doi: 10.3322/caac.21492

Callahan, K., Ponce, C., Cross, C., Sy, F., and Pinheiro, P. (2019). Low colorectal cancer survival in the Mountain West state of Nevada: A population-based analysis. *PLoS One* 14:e0221337. doi: 10.1371/journal.pone.0221337

Cao, Z., Pan, X., Yang, Y., Huang, Y., and Shen, H. B. (2018). The lncLocator: a subcellular localization predictor for long non-coding RNAs based on a stacked ensemble classifier. *Bioinformatics* 34, 2185–2194. doi: 10.1093/bioinformatics/bty085

Chen, Y., Pal, B., Visvader, J. E., and Smyth, G. K. (2017). Differential methylation analysis of reduced representation bisulfite sequencing experiments using edgeR. *F1000Res* 6, 2055. doi: 10.12688/f1000research.13196.2

Cheng, L. (2019). Computational and Biological Methods for Gene Therapy. *Current Gene Therapy* 19, 210–210.

Cheng, L., Hu, Y., Sun, J., Zhou, M., and Jiang, Q. (2018). DincRNA: a comprehensive web-based bioinformatics toolkit for exploring disease associations and ncRNA function. *Bioinformatics* 34, 1953–1956. doi: 10.1093/bioinformatics/bty002

Cheng, L., Wang, P., Tian, R., Wang, S., Guo, Q., Luo, M., et al. (2019a). LncRNA2Target v2.0: a comprehensive database for target genes of lncRNAs in human and mouse. *Nucleic Acids Res* 47, D140–D144. doi: 10.1093/nar/gky1051

Cheng, L., Zhao, H., Wang, P., Zhou, W., Luo, M., Li, T., et al. (2019b). Computational Methods for Identifying Similar Diseases. *Molecular therapy. Nucleic acids* 18, 590–604.

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 44, e71. doi: 10.1093/nar/gkv1507

Doja, M. N., Kaur, I., and Ahmad, T. (2020). Current State of the Art for Survival Prediction in Cancer Using Data Mining Techniques. *Current Bioinformatics* 15, 174–186. doi: 10.2174/1574893614666190902152142

Dong, Y.-M., Bi, J.-H., He, Q.-E., and Song, K. (2020). ESDA: An Improved Approach to Accurately Identify Human snoRNAs for

Precision Cancer Therapy. *Current Bioinformatics* 15, 34–40. doi: 10.2174/1574893614666190424162230

Donovan, J., and Copeland, P. R. (2009). Evolutionary history of selenocysteine incorporation from the perspective of SECIS binding proteins. *BMC Evolutionary Biology* 9:229. doi: 10.1186/1471-2148-9-229

Ferlay, J., Shin, H.-R., Bray, F., Forman, D., and Parkin, D. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN. *International journal of cancer. Journal international du cancer* 127, 2893–2917. doi: 10.1002/ijc.25516

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 136, E359–E386. doi: 10.1002/ijc.29210

Fitzmaurice, C., Abate, D., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdel-Rahman, O., et al. (2019). Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 5, 1749–1768. doi: 10.1001/jamaoncol.2019.2996

Fitzmaurice, C., Akinyemiju, T. F., Al Lami, F. H., Alam, T., Alizadeh-Navaei, R., Allen, C., et al. (2018). Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2016: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 4, 1553–1568. doi: 10.1001/jamaoncol.2018.2706

Fitzmaurice, C., Allen, C., Barber, R. M., Barregard, L., Bhutta, Z. A., Brenner, H., et al. (2017). Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 3, 524–548. doi: 10.1001/jamaoncol.2016.5688

Ghosh, A., and Yan, H. (2020). Stability Analysis at Key Positions of EGFR Related to Non-small Cell Lung Cancer. *Current Bioinformatics* 15, 260–267. doi: 10.2174/1574893614666191212112026

Hong, Z., Zeng, X., Wei, L., and Liu, X. J. B. (2020). Identifying enhancer–promoter interactions with neural network based on pre-trained DNA vectors and attention mechanism. *Bioinformatics* 36, 1037–1043.

Huang, Z.-M., Ge, H.-F., Yang, C.-C., Cai, Y., Chen, Z., Tian, W.-Z., et al. (2019). MicroRNA-26a-5p inhibits breast cancer cell growth by suppressing RNF6 expression. *The Kaohsiung Journal of Medical Sciences* 35, 467–473. doi: 10.1002/kjm2.12085

Iliopoulos, A. C., Beis, G., Apostolou, P., and Papasotiriou, I. (2020). Complex Networks, Gene Expression and Cancer Complexity: A Brief Review of Methodology and Applications. *Current Bioinformatics* 15, 629–655. doi: 10.2174/1574893614666191017093504

Itskovich, S. S., Gurunathan, A., Clark, J., Burwinkel, M., Wunderlich, M., Berger, M. R., et al. (2020). MBNL1 regulates essential alternative RNA splicing patterns in MLL-rearranged leukemia. *Nature Communications* 11, 2369. doi: 10.1038/s41467-020-15733-8

Ji, J., Tang, J., Xia, K.-J., and Jiang, R. (2019). LncRNA in Tumorigenesis Microenvironment. *Current Bioinformatics* 14, 640–641. doi: 10.2174/1574893614071909171616 54

Jiang, L., Ding, Y., Tang, J., and Guo, F. (2018). MDA-SKF: Similarity Kernel Fusion for Accurately Discovering miRNA-Disease Association. *Frontiers in Genetics* 9:1–13. doi: 10.3389/fgene.2018.00618

Jiang, L., Xiao, Y., Ding, Y., Tang, J., and Guo, F. (2019). FKL-Spa-LapRLS: an accurate method for identifying human microRNA-disease association. *BMC Genomics* 19:11–25.

Jiang, Q., Hao, Y., Wang, G., Juan, L., Zhang, T., Teng, M., et al. (2010). Prioritization of disease microRNAs through a human phenome-microRNAome network. *BMC Syst Biol* 4(Suppl. 1):S2. doi: 10.1186/1752-0509-4-S1-S2

Jiang, Q., Wang, G., Jin, S., Li, Y., and Wang, Y. (2013). Predicting human microRNA-disease associations based on support vector machine. *Int J Data Min Bioinform* 8, 282–293.

Jiang, Q., Wang, J., Wang, Y., Ma, R., Wu, X., and Li, Y. (2014). TF2LncRNA: identifying common transcription factors for a list of lncRNA genes from ChIP-Seq data. *Biomed Res Int* 2014, 317642. doi: 10.1155/2014/317642

Jiang, Q., Wang, Y., Hao, Y., Juan, L., Teng, M., Zhang, X., et al. (2009). miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res* 37, D98–D104. doi: 10.1093/nar/gkn714

Jin, S., Zeng, X., Fang, J., Lin, J., Chan, S. Y., Erzurum, S. C., et al. (2019). A network-based approach to uncover microRNA-mediated disease comorbidities and potential pathobiological implications. *NPJ systems biology and applications* 5, 1–11.

Jin, S., Zeng, X., Xia, F., Huang, W., and Liu, X. (2020). Application of deep learning methods in biological networks. *Briefings in Bioinformatics* 2, bbaa043. doi: 10.1093/bib/bbaa043

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Langfelder, P., Mischel, P. S., and Horvath, S. (2013). When Is Hub Gene Selection Better than Standard Meta-Analysis? *PLoS One* 8:e61505. doi: 10.1371/journal.pone.0061505

Law, C., Alhamdoosh, M., Su, S., Smyth, G., and Ritchie, M. (2016).). RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research* 5, 1408. doi: 10.12688/f1000research.9005.1

Li, J., Chang, M., Gao, Q., Song, X., and Gao, Z. (2020). Lung Cancer Classification and Gene Selection by Combining Affinity Propagation Clustering and Sparse Group Lasso. *Current Bioinformatics* 15, 703–712. doi: 10.2174/1574893614666191017103557

Li, W., Liu, J., and Zhao, H. (2020). Identification of a nomogram based on long non-coding RNA to improve prognosis prediction of esophageal squamous cell carcinoma. *Aging* 12, 1512–1526. doi: 10.18632/aging.102697

Li, Z., Zhang, T., Lei, H., Wei, L., Liu, Y., Shi, Y., et al. (2020). Research on Gastric Cancer's Drug-resistant Gene Regulatory Network Model. *Current Bioinformatics* 15, 225–234. doi: 10.2174/1574893614666190722102557

Li, S., Jiang, L., Tang, J., Gao, N., and Guo, F. (2020). Kernel Fusion Method for Detecting Cancer Subtypes via Selecting Relevant Expression Data. *Front. Genet.* 11:979. doi: 10.3389/fgene.2020.00979

Liang, C., Changlu, Q., He, Z., Tongze, F., and Xue, Z. (2019). gutMDisorder: a comprehensive database for dysbiosis of the gut microbiota in disorders and interventions. *Nucleic Acids Research* 48, 7603.

Lin, M., Li, X., Guo, H., Ji, F., Ye, L., Ma, X., et al. (2019). Identification of Bone Metastasis-associated Genes of Gastric Cancer by Genome-wide Transcriptional Profiling. *Current Bioinformatics* 14, 62–69. doi: 10.2174/1574893612666171121154017

Liu, X., Hong, Z., Liu, J., Lin, Y., Rodríguez-Patón, A., Zou, Q., et al. (2020). Computational methods for identifying the critical nodes in biological networks. *Briefings in Bioinformatics* 21, 486–497.

Liu, Y., Qian, L., Yang, J., Huang, H., Feng, J., Li, X., et al. (2018). The expression level and prognostic value of HIPK3 among non-small-cell lung cancer patients in China. *OncoTargets and therapy* 11, 7459–7469. doi: 10.2147/OTT.S166878

Liu, Y., Zeng, X., He, Z., and Zou, Q. (2017). Inferring MicroRNA-Disease Associations by Random Walk on a Heterogeneous Network with Multiple Data Sources. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 14, 905–915. doi: 10.1109/tcbb.2016.2550432

Long, J., Bai, Y., Yang, X., Lin, J., Yang, X., Wang, D., et al. (2019). Construction and comprehensive analysis of a ceRNA network to reveal potential prognostic biomarkers for hepatocellular carcinoma. *Cancer Cell International* 19, 90. doi: 10.1186/s12935-019-0817-y

Ma, X., Xi, B., Zhang, Y., Zhu, L., Sui, X., Tian, G., et al. (2020). A Machine Learning-based Diagnosis of Thyroid Cancer Using Thyroid Nodules Ultrasound Images. *Current Bioinformatics* 15, 349–358. doi: 10.2174/1574893614666191017091959

Maza, E. (2016). In Papyro Comparison of TMM (edgeR), RLE (DESeq2), and MRN Normalization Methods for a Simple Two-Conditions-Without-Replicates RNA-Seq Experimental Design. *Front Genet* 7:164. doi: 10.3389/fgene.2016.00164

Nadia, and Ramana, J. (2020). The Human OncoBiome Database: A Database of Cancer Microbiome Datasets. *Current Bioinformatics* 15, 472–477. doi: 10.2174/1574893614666190902152727

Oñate, S. A., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (1995). Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* 270, 1354–1357. doi: 10.1126/science.270.5240.1354

Ozkan, A., Isgor, S. B., Sengul, G., and Isgor, Y. G. (2019). Benchmarking Classification Models for Cell Viability on Novel Cancer Image Datasets. *Current Bioinformatics* 14, 108–114. doi: 10.2174/1574893614666181120093740

Pan, H., Guo, C., Pan, J., Guo, D., Song, S., Zhou, Y., et al. (2019). Construction of a Competitive Endogenous RNA Network and Identification of Potential Regulatory Axis in Gastric Cancer. *Frontiers in Oncology* 9:912. doi: 10.3389/fonc.2019.00912

Pawa, N., Arulampalam, T., and Norton, J. D. (2011). Screening for colorectal cancer: established and emerging modalities. *Nat Rev Gastroenterol Hepatol* 8, 711–722. doi: 10.1038/nrgastro.2011.205

Qian, W., Feng, Y., Li, J., Peng, W., Gu, Q., Zhang, Z., et al. (2019). Construction of ceRNA networks reveals differences between distal and proximal colon cancers. *Oncol Rep* 41, 3027–3040. doi: 10.3892/or.2019.7083

Qin, L., Xu, Y., Xu, Y., Ma, G., Liao, L., Wu, Y., et al. (2015). NCOA1 promotes angiogenesis in breast tumors by simultaneously enhancing both HIF1α- and AP-1-mediated VEGFa transcription. *Oncotarget* 6, 23890–23904. doi: 10.18632/oncotarget.4341

Rahman, R., Schmaltz, C., Jackson, C. S., Simoes, E. J., Jackson-Thompson, J., and Ibdah, J. A. (2015). Increased risk for colorectal cancer under age 50 in racial and ethnic minorities living in the United States. *Cancer Med* 4, 1863–1870. doi: 10.1002/cam4.560

Redmond, A. M., Bane, F. T., Stafford, A. T., McIlroy, M., Dillon, M. F., Crotty, T. B., et al. (2009). Coassociation of estrogen receptor and p160 proteins predicts resistance to endocrine treatment; SRC-1 is an independent predictor of breast cancer recurrence. *Clin Cancer Res* 15, 2098–2106. doi: 10.1158/1078-0432.Ccr-08-1649

Rustgi, A., and El-Serag, H. (2014). Esophageal Carcinoma. *The New England journal of medicine* 371, 2499–2509. doi: 10.1056/NEJMra1314530

Rustgi, A., and El-Serag, H. B. (2015). Esophageal carcinoma. *N Engl J Med* 372, 1472–1473. doi: 10.1056/NEJMc1500692

Salehi, Z., Akrami, H., and Sisakhtnezhad, S. (2018). The effect of placenta growth factor knockdown on hsa-miR-22-3p, hsa-let-7b-3p, hsa-miR-451b, and hsa-mir-4290 expressions in MKN-45-derived gastric cancer stem-like cells. *Middle East Journal of Cancer* 9, 113–122.

Sardina, D. S., Alaimo, S., Ferro, A., Pulvirenti, A., and Giugno, R. (2017). A novel computational method for inferring competing endogenous interactions. *Brief Bioinform* 18, 1071–1081. doi: 10.1093/bib/bbw084

Schneider, T., Martinez-Martinez, A., Cubillos-Rojas, M., Bartrons, R., Ventura, F., and Rosa, J. L. (2018). The E3 ubiquitin ligase HERC1 controls the ERK signaling pathway targeting C-RAF for degradation. *Oncotarget* 9, 31531–31548. doi: 10.18632/oncotarget.25847

Shen, C., Jiang, L., Ding, Y., Tang, J., and Guo, F. (2019). LPI-KTASLP: Prediction of lncRNA-Protein Interaction by Semi-Supervised Link Learning with Multivariate Information. *IEEE Access* 7, 13486–13496.

Slattery, M. L., Wolff, R. K., Herrick, J. S., Caan, B. J., and Potter, J. D. (2007). IL6 genotypes and colon and rectal cancer. *Cancer Causes & Control* 18, 1095–1105. doi: 10.1007/s10552-007-9049-x

Song, J., Lu, Y., Sun, W., Han, M., Zhang, Y., and Zhang, J. (2019). Changing expression profiles of lncRNAs, circRNAs and mRNAs in esophageal squamous carcinoma. *Oncology Letters* 18, 5363–5373. doi: 10.3892/ol.2019.10880

Su, H., Lin, F., Deng, X., Shen, L., Fang, Y., Fei, Z., et al. (2016). Profiling and bioinformatics analyses reveal differential circular RNA expression in radioresistant esophageal cancer cells. *Journal of translational medicine* 14, 225–225. doi: 10.1186/s12967-016-0977-7

Sun, L.-P., Xu, K., Cui, J., Yuan, D.-Y., Zou, B., Li, J., et al. (2019). Cancer-associated fibroblast-derived exosomal miR-382-5p promotes the migration and invasion of oral squamous cell carcinoma. *Oncology reports* 42, 1319–1328. doi: 10.3892/or.2019.7255

Tang, W., Wan, S., Yang, Z., Teschendorff, A. E., and Zou, Q. (2018). Tumor origin detection with tissue-specific miRNA and DNA methylation markers. *Bioinformatics* 34, 398–406. doi: 10.1093/bioinformatics/btx622

Tay, Y., Rinn, J., and Pandolfi, P. P. (2014). The multilayered complexity of ceRNA crosstalk and competition. *Nature* 505, 344–352. doi: 10.1038/nature12986

Thomson, D. W., and Dinger, M. E. (2016). Endogenous microRNA sponges: evidence and controversy. *Nature Reviews Genetics* 17, 272–283. doi: 10.1038/nrg.2016.20

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. *CA Cancer J Clin* 65, 87–108. doi: 10.3322/caac.21262

Tupone, M. G., D'Aguanno, S., Di Martile, M., Valentini, E., Desideri, M., Trisciuoglio, D., et al. (2020). microRNA-378a-5p iS a novel positive regulator of melanoma progression. *Oncogenesis* 9, 22. doi: 10.1038/s41389-020-0203-6

Walsh, C. A., Qin, L., Tien, J. C.-Y., Young, L. S., and Xu, J. (2012). The function of steroid receptor coactivator-1 in normal tissues and cancer. *International journal of biological sciences* 8, 470–485. doi: 10.7150/ijbs.4125

Wang, J., Liu, X., Wu, H., Ni, P., Gu, Z., Qiao, Y., et al. (2010). CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res* 38, 5366–5383. doi: 10.1093/nar/gkq285

Wei, L., Liao, M., Gao, Y., Ji, R., He, Z., and Zou, Q. (2014). Improved and Promising Identification of Human MicroRNAs by Incorporating a High-Quality Negative Set. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 11, 192–201. doi: 10.1109/tcbb.2013.146

Wei, L., Su, R., Wang, B., Li, X., Zou, Q., and Gao, X. (2019a). Integration of deep feature representations and handcrafted features to improve the prediction of N-6-methyladenosine sites. *Neurocomputing* 324, 3–9. doi: 10.1016/j.neucom.2018.04.082

Wei, L., Xing, P., Shi, G., Ji, Z., and Zou, Q. (2019b). Fast Prediction of Protein Methylation Sites Using a Sequence-Based Feature Selection Technique. *Ieee-Acm Transactions on Computational Biology and Bioinformatics* 16, 1264–1273. doi: 10.1109/tcbb.2017.2670558

Wei, L., Zhou, C., Chen, H., Song, J., and Su, R. (2018). ACPred-FL: a sequence-based predictor using effective feature representation to improve the prediction of anti-cancer peptides. *Bioinformatics* 34, 4007–4016. doi: 10.1093/bioinformatics/bty451

Wu, W., Yang, Z., Long, F., Luo, L., Deng, Q., Wu, J., et al. (2019). COL1A1 and MZB1 as the hub genes influenced the proliferation, invasion, migration and apoptosis of rectum adenocarcinoma cells by weighted correlation network analysis. *Bioorganic Chemistry* 95, 103457. doi: 10.1016/j.bioorg.2019.103457

Xu, J., Wu, R. C., and O'Malley, B. W. (2009). Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. *Nat Rev Cancer* 9, 615–630. doi: 10.1038/nrc2695

Yang, L., Gao, H., Wu, K., Zhang, H., Li, C., and Tang, L. (2020). Identification of Cancerlectins By Using Cascade Linear Discriminant Analysis and Optimal g-gap Tripeptide Composition. *Current Bioinformatics* 15, 528–537. doi: 10.2174/1574893614666190730103156

Zeng, W., Wang, F., Ma, Y., Liang, X., and Chen, P. (2019). Dysfunctional Mechanism of Liver Cancer Mediated by Transcription Factor and Non-coding RNA. *Current Bioinformatics* 14, 100–107. doi: 10.2174/1574893614666181119121916

Zeng, X., Zhu, S., Liu, X., Zhou, Y., Nussinov, R., and Cheng, F. J. B. (2019). deepDR: a network-based deep learning approach to in silico drug repositioning. *Bioinformatics* 35, 5191–5198.

Zeng, X., Lin, Y., He, Y., Lv, L., and Min, X. (2020a). Deep collaborative filtering for prediction of disease genes. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 17, 1639–1647.

Zeng, X., Zhu, S., Lu, W., Liu, Z., Huang, J., Zhou, Y., et al. (2020b). Target identification among known drugs by deep learning from heterogeneous networks. *Chemical Science* 11, 1775–1797.

Zeng, X., Zhang, X., and Zou, Q. (2016). Integrative approaches for predicting microRNA function and prioritizing disease-related microRNA using biological interaction networks. *Briefings in Bioinformatics* 17, 193–203. doi: 10.1093/bib/bbv033

Zeng, X. X., Liu, L., Lu, L. Y., and Zou, Q. (2018). Prediction of potential disease-associated microRNAs using structural perturbation method. *Bioinformatics* 34, 2425–2432. doi: 10.1093/bioinformatics/bty112

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Research* 46, D754–D761. doi: 10.1093/nar/gkx1098

Zhang, L., He, Y., Song, H., Wang, X., Lu, N., Sun, L., et al. (2020). Elastic Net Regularized Softmax Regression Methods for Multi-subtype Classification in Cancer. *Current Bioinformatics* 15, 212–224. doi: 10.2174/1574893613666181112141724

Zhang, Z. M., Tan, J. X., Wang, F., Dao, F. Y., Zhang, Z. Y., and Lin, H. (2020). Early Diagnosis of Hepatocellular Carcinoma Using Machine Learning Method. *Front Bioeng Biotechnol* 8:254. doi: 10.3389/fbioe.2020.00254

Zhang, X., Zou, Q., Rodriguez-Paton, A., and Zeng, X. (2017). Meta-path methods for prioritizing candidate disease miRNAs. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 16, 283–291.

Zhang, Y., Kou, C., Wang, S., and Zhang, Y. (2019). Genome-wide Differential-based Analysis of the Relationship between DNA Methylation and Gene Expression in Cancer. *Current Bioinformatics* 14, 783–792. doi: 10.2174/1574893614666190424160046

Zhao, T., Hu, Y., Peng, J., and Cheng, L. (2020). DeepLGP: a novel deep learning method for prioritizing lncRNA target genes. *Bioinformatics* 36, 4466–4472. doi: 10.1093/bioinformatics/btaa428

Zou, Q., Li, J., Song, L., Zeng, X., and Wang, G. (2016). Similarity computation strategies in the microRNA-disease network: a survey. *Briefings in Functional Genomics* 15, 55–64. doi: 10.1093/bfgp/elv024

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 01 April 2021
doi: 10.3389/fgene.2021.638980

A Glycolysis-Based Long Non-coding RNA Signature Accurately Predicts Prognosis in Renal Carcinoma Patients

*Honghao Cao[1,2], Hang Tong[1,3], Junlong Zhu[1,3], Chenchen Xie[3], Zijia Qin[1,3], Tinghao Li[1,3], Xudong Liu[1] and Weiyang He[1]**

[1] Department of Urology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China, [2] Department of Urology, Rongchang Traditional Chinese Medicine Hospital, Chongqing, China, [3] Central Laboratory, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

Background: The prognosis of renal cell carcinoma (RCC) varies greatly among different risk groups, and the traditional indicators have limited effect in the identification of risk grade in patients with RCC. The purpose of our study is to explore a glycolysis-based long non-coding RNAs (lncRNAs) signature and verify its potential clinical significance in prognostic prediction of RCC patients.

Methods: In this study, RNA data and clinical information were downloaded from The Cancer Genome Atlas (TCGA) database. Univariate and multivariate cox regression displayed six significantly related lncRNAs (AC124854.1, AC078778.1, EMX2OS, DLGAP1-AS2, AC084876.1, and AC026401.3) which were utilized in construction of risk score by a formula. The accuracy of risk score was verified by a series of statistical methods such as receiver operating characteristic (ROC) curves, nomogram and Kaplan-Meier curves. Its potential clinical significance was excavated by gene enrichment analysis.

Results: Kaplan-Meier curves and ROC curves showed reliability of the risk score to predict the prognosis of RCC patients. Stratification analysis indicated that the risk score was independent predictor compare to other traditional clinical parameters. The clinical nomogram showed highly rigorous with index of 0.73 and precisely predicted 1-, 3-, and 5-year survival time of RCC patients. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene set enrichment analysis (GSEA) depicted the top ten correlated pathways in both high-risk group and low-risk group. There are 6 lncRNAs and 25 related mRNAs including 36 lncRNA-mRNA links in lncRNA-mRNA co-expression network.

Conclusion: This research demonstrated that glycolysis-based lncRNAs possessed an important value in survival prediction of RCC patients, which would be a potential target for future treatment.

Keywords: glycolysis, lncRNA, prognosis, carcinoma, renal cell

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Aimin Li,
Xi'an University of Technology, China
Saurav Mallik,
Harvard University, United States

***Correspondence:**
Weiyang He
weiyang361@aliyun.com

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *08 December 2020*
Accepted: *16 March 2021*
Published: *01 April 2021*

Citation:
Cao H, Tong H, Zhu J, Xie C, Qin Z, Li T, Liu X and He W (2021) A Glycolysis-Based Long Non-coding RNA Signature Accurately Predicts Prognosis in Renal Carcinoma Patients. Front. Genet. 12:638980. doi: 10.3389/fgene.2021.638980

INTRODUCTION

Renal cell carcinoma (RCC) is one of the most aggressive urinary system tumor, which accounts for nearly 4% of adult malignancies (Zhai et al., 2019). The latest statistical researches show that there are 403,262 new cases and 175,098 deaths in 2018 (Khadirnaikar et al., 2019). Histologically, there are many subtypes of RCC, which can be used to evaluate the prognosis. Despite significant advances in its diagnosis and treatment, the prognosis of RCC patients remains poor (Cheng et al., 2019). Due to the low sensitivity to radiotherapy and chemotherapy, the treatment of RCC is a systematic treatment based on radical resection (Heo et al., 2020). Recently, targeted therapy for RCC has been used in clinical cases, but the overall survival (OS) of patients is not as good as expected, especially for some terminal stage patients. Therefore, it is urgent to find the molecular atypia of RCC in order to distinguish between high- and low-risk patients, so as to facilitate early diagnosis and improve prognosis.

Long non-coding RNAs (lncRNAs) are the member of the non-coding RNA family, which have over 200 nucleotides in length (Michal et al., 2020). Previous genome sequencing studies have explicated that lncRNAs do not have the function of transcribing proteins, but play a vitally important role in variety of physiologic activities (Zhang Y. et al., 2020). Several studies have explicitly pointed out that lncRNAs participate in important cellular biological functions for their contributions in regulation of transcription and organization of nuclear domains (Yang et al., 2020). Meanwhile, lncRNAs play crucial roles in many cellular processes, which include glycolysis, cell differentiation and DNA repair (Wang Y. et al., 2019). According to these findings, abnormal expression of lncRNAs in tumor cell reveals that lncRNAs may be as potential bio-markers in survival prediction of RCC patients. Additionally, latest researches have revealed relation between non-mutational regulation of gene expression and drug resistance, during which lncRNAs could affect drug sensitivity to tumor cells as the major modulators (Xu et al., 2020). Based on the highly specific subtype of tumor cell, lncRNAs have been considered to have the potential ability to predict prognosis and provide novel therapeutic options.

The occurrence and migration of RCC have been verified to be related to glycolysis (Zhang C. et al., 2020). Glycolysis is a normal physiological metabolic process of human cells, which provides limited energy for the body. However, Glycolysis, known as the "Warburg effect," is extremely active in tumor cell metabolism which decompose glucose or glycogen into pyruvate or lactic acid, even under aerobic environment (Chen et al., 2019; Kim et al., 2020). The glycolysis can product sufficient energy to tumor cell and inhibits apoptosis which promote proliferation and migration of cancer cell under severe environment (Wang S. et al., 2019). Interestingly, different solid tumors show different heterogeneity of glycolysis, such as breast cancer and non-small cell lung cancer (Guo et al., 2020; Smolle et al., 2020). Several studies have demonstrated that glycolysis can be used as a biological metabolic marker to predict the prognosis of tumors and further study of the mechanism of glycolysis can explore new targets to guide treatment.

Several studies have confirmed that the glycolysis-based lncRNAs affect solid tumor cells (such as colon cancer, breast cancer, and ovarian cancer) through epigenetic regulation (Chen et al., 2019; Lin et al., 2020; Zhang Z. et al., 2020). The role of the glycolysis-based lncRNAs in RCC is still unclear. In our study, we hypothesis that there are several lncRNAs related to glycolysis genes and may contribute to predict prognosis of RCC patients. By combining the clinicopathological types and tumor molecular characteristics, we can more effectively reveal the heterogeneity of RCC, and provide theoretical basis for the clinical diagnosis and prognosis of RCC. Finally, we developed a new scoring system based on six lncRNAs related to glycolysis aiming to accurately assess the prognosis of patients.

Abbreviations: AJCC, The American Joint Committee on Cancer; AUC, area under the curve; RCC, renal cell carcinoma; GO, Gene Ontology; GSEA, Gene Set Enrichment Analysis; HR, hazard ratio; KEGG, Kyoto Encyclopedia of Genes and Genomes; LncRNA, long non-coding RNA; OS, overall survival; ROC, receiver operating characteristic; TCGA, The Cancer Genome Atlas.

MATERIALS AND METHODS

Collection of Transcriptional Data and Clinical Sample

RNA expression data and clinical information were downloaded from the [The Cancer Genome Atlas (TCGA)-GDC][1]. The extracted clinical data included OS, age, sex, grade, stage, tumor size, distant metastasis, and lymph node metastasis. The glycolysis-related gene expression profiles downloaded from Gene Set Enrichment Analysis (GSEA)[2]. Patients with incomplete data or vague living status were excluded from our study. The clinical data were available in the database publicly, so this research required neither written approval from patients nor agreements from Ethics committee. All the data of this article was downloaded on October 15, 2020.

Identification of Glycolysis-Related LncRNAs and Construction of Prognosis Model

Pearson correlation analysis was conducted to identify glycolysis-related lncRNAs. The correlation was calculated according to the expression value between lncRNAs and glycolysis genes. Our selection criteria were $|R| > 0.4$ and $P < 0.001$. Survival package of R was utilized for univariate and multivariate Cox regression analysis which provided hazard ratio (HR), β (cox) and *P*-values. The HRs were used to identify risk-related lncRNAs (HR > 1) and protective lncRNAs (HR < 1). Subsequently, we identified six target glycolysis-related lncRNAs for the prognostic signature model. Finally, the predictive model was constructed with six glycolysis-related lncRNAs through a previous formula (risk score $= \sum_{(i=1)}^{n} \text{Coef}\ (i) \times x\ (i)$) (Sun et al., 2020), where Coef (*i*) and *x* (*i*) represent the estimated regression coefficient and the expression value of each glycolysis-related lncRNAs, respectively.

[1] https://portal.gdc.cancer.gov/

[2] http://www.gsea-msigdb.org/gsea/index.jsp

Evaluation and Verification of the Accuracy of Prognostic Signature

The Kaplan-Meier survival curve was utilized to show the survival correlation between high-risk group and low-risk group. Then, heatmap and scatter dot plot were used to observe the gene expression and prognosis of different groups. Finally, the receiver operating characteristic (ROC) curve was used to evaluate the accuracy of prediction. The correlation analysis was utilized to verify the connections between the risk score and clinicopathological characteristics of patients. We proved the independent prediction ability of glycolysis-related lncRNAs related prediction model through univariate and multivariate Cox regression analyses. In addition, stratified analysis was utilized to examine the precision of prognostic prediction in patient survival outcomes according to other clinicalpathological features.

Establishment and Validation of Nomogram

A nomogram based on the risk scores and other clinicalpathological parameters was constructed to provide a reliable clinical prediction tool for RCC patients in 3- and 5-years survival time. Next, the calibration curves were applied to assess the concordance between predicted and actual observed patients. Finally, we also examined the area under the curve (AUC) values of ROC to determine the accuracy of our nomogram in predicting prognosis of patients.

Gene Set Enrichment Analysis (GSEA)

In order to discover significant functional phenotype between high-risk group and low-risk group, we conducted GSEA. The GSEA software were downloaded from (see text footnote 2) and were conducted to distinguish differential expressed genes in high- and low-risk groups. The enriched gene sets were obtained based on a P-value < 0.05 and a false discovery rate (FDR) value < 0.25 after performing 1,000 permutations (Bandyopadhyay et al., 2014). Then the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was used to explore the exact signaling pathways related to high- and low-risk group.

Construction of the LncRNA-mRNA Co-expression Network

The lncRNA-mRNA co-expression network was constructed using Cytoscape to research the correlation between the glycolysis-related lncRNAs and their target mRNAs and the potential functions of the six glycolysis-related lncRNAs in renal carcinoma.

Statistical Analysis and Software Support

The data was processed via PERL programming langue[3], Version (strawberry-perl-5.32.0.1-64bit.msi). Statistical analysis and figure constructing were conducted with R software[4], version R x64 4.0.2.

[3]https://www.perl.org/

RESULTS

Acquisition of Prognostically Significant Glycolysis-Related LncRNAs in RCC Patient Tissue Samples

A total of 539 renal cancer patients were included for the subsequent study and the clinical information of 537 patients with renal clear cell carcinoma was downloaded at the same time. The complete clinical characteristics of the patients were included in **Table 1**. We also extracted glycolysis-related genes from the GSEA (see text footnote 2) and next acquired 1199 glycolysis-related lncRNAs by performing Pearson correlation analysis between the lncRNAs samples and the glycolysis-related genes setting $|R| > 0.4$ and $P < 0.001$ as the selection criteria. Univariate Cox regression analysis combining the clinical surviving statistics exhibited that expression of 18 lncRNAs significantly correlated with the survival time of RCC patients ($P < 0.05$; **Figure 1A**). Multivariate Cox regression analysis filtered 6 well candidate lncRNAs for constructing the prognostic signature (AC124854.1, AC078778.1, EMX2OS, DLGAP1-AS2, AC084876.1, and AC026401.3). Among the 6 lncRNAs,

[4]https://www.r-project.org/

TABLE 1 | Summary of baseline clinical characteristics of patients with renal cancer.

Characteristic	TCGA
Age (years)	
≤65	340
>65	174
Gender	
Male	335
Female	179
Grade	
Low	230
High	276
T stage	
T1	263
T2	66
T3	75
T4	11
N stage	
N0	226
N1-3	17
AJCC stage	
I–II	310
III–IV	201
Survival status	
Alive	348
Deceased	166

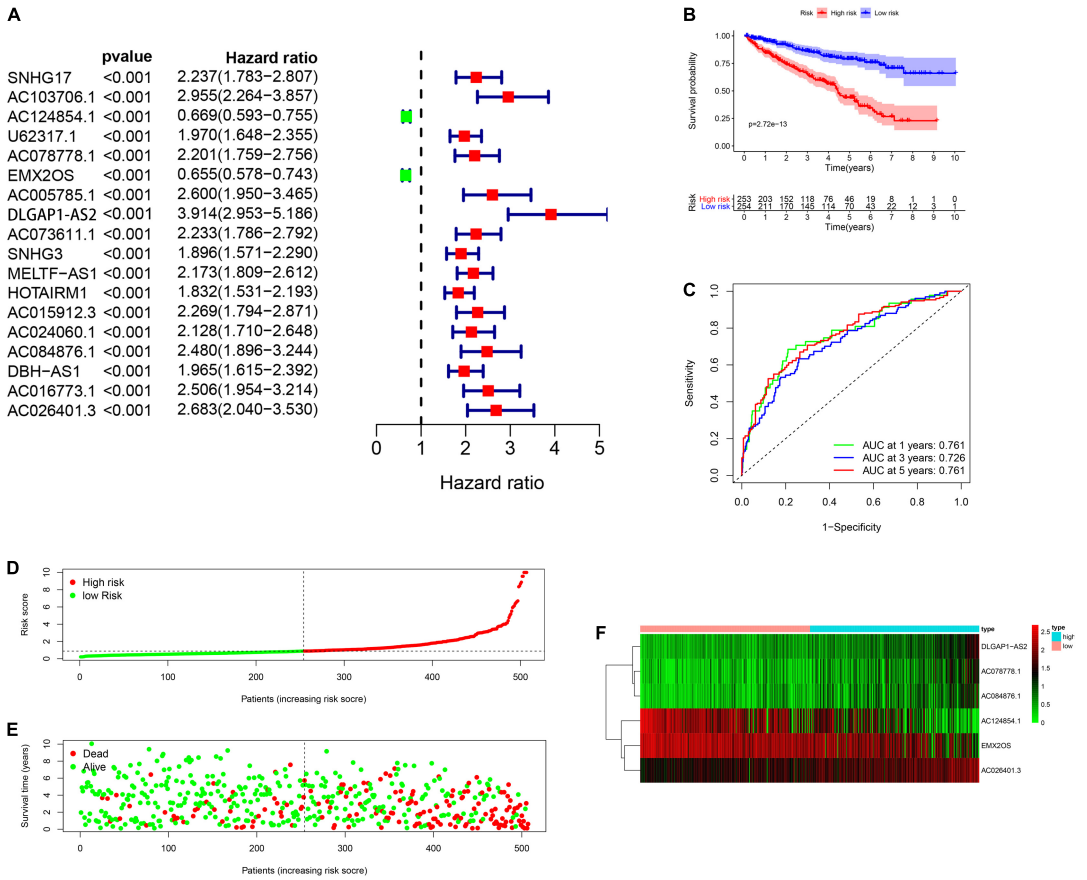


FIGURE 1 | Construction and verification of glycolysis-related lncRNAs prognostic signatures in RCC patients. **(A)** The univariate Cox regression analysis combining the clinical surviving statistics exhibit that expression of 18 lncRNAs significantly correlated with OS of RCC patients. **(B)** Kaplan-Meier survival curve analysis shows that the significantly different survival time between the low-risk and the high-risk group. **(C)** ROC curve indicate the reliability of applying the risk score to predict the prognosis of RCC patients. **(D)** Distribution of high-risk and low-risk patients with renal cell carcinoma based on the glycolysis-related lncRNAs prognosis signature. **(E)** The scatter dot plot shows that the relationship between survival rates of the RCC patients and the risk score. **(F)** The heatmap shows the different expression of prognostic signature-related lncRNAs in different risk groups.

DLGAP1-AS2, AC084876.1, and AC026401.3 possessed the HR larger than 1 while the others smaller than 1 (**Table 2**).

Verifies the Accuracy of Six Glycolysis-Related LncRNAs Prognostic Signature

Then, the clinical patients were distinguished into two groups which included high-risk group (n = 253) and low-risk group (n = 254) according to their corresponding median cut off value. Kaplan-Meier survival curve analysis revealed that patients in the low-risk group presented a significant higher survival time and a better prognosis than those in the high-risk group (**Figure 1B**). ROC curve indicated the reliability of applying the risk scores to predict the prognosis of RCC patients in 1-, 3- and 5-years and all the AUC (area under the ROC) value greater than 0.7 (**Figure 1C**). RCC patients were then ranked according to the risk scores based on the glycolysis-related lncRNAs prognosis

TABLE 2 | The information of target lncRNAs from the multivariate cox regression analysis.

Gene	Ensemble ID	Location	β(cox)	HR	*P*
AC124854.1	ENSG00000249776	chr5:92,410,256-92,671,581	−0.216164701	0.805603	0.006005
AC078778.1	ENSG00000258344	chr12:54,276,631-54,345,083	−0.388143004	0.678315	0.1036
EMX2OS	ENSG00000229847	chr10:117,470,801-117,545,068	−0.19260207	0.82481	0.020958
DLGAP1-AS2	ENSG00000262001	chr18:3,603,000-3,610,089	0.862480852	2.369031	0.000234
AC084876.1	ENSG00000277299	chr12:109,948,389-109,949,029	0.550076597	1.733386	0.010455
AC026401.3	ENSG00000280206	chr16:15,701,237-15,702,118	0.303059387	1.353995	0.074331

signature (**Figure 1D**). The scatter dot plot showed that the relationship between survival rates of the RCC patients and the risk score; patients with a higher risk score demonstrated lower survival time (**Figure 1E**). The heatmap showed the different expression of prognostic signature-related lncRNAs in the low-risk group and high-risk group. The risk factors (DLGAP1-AS2, AC084876.1, and AC078778.1) presented higher level in high-risk patients while low-risk patients presented higher levels of protective factors (AC124854.1, DLGAP1-AS2, and AC026401.3) (**Figure 1F**).

Correlation Analysis of the Glycolysis-Related LncRNAs Prognosis Signature With Clinical Features

To discover the connections between the risk scores and clinicopathological characteristics of patients, we then analyzed the correlation and the results displayed that except for gender (**Figure 2B**, $P = 0.5908$), other parameters including age (**Figure 2A**, $P = 0.017$), grade (**Figure 2C**, $P < 0.001$), M stage (**Figure 2D**, $P < 0.001$), American Joint Committee on Cancer (AJCC), stage (**Figure 2E**, $P < 0.0001$), and T stage (**Figure 2F**, $P < 0.001$) were significant to risk scores. These results demonstrated that glycolysis-related lncRNAs risk signature was significantly related to renal carcinoma progression.

Assess the Possibility of Independent Prediction of Glycolysis-Related LncRNAs Risk Signature

Then, we verified whether the glycolysis-related lncRNAs prognostic signature was an independent prognostic factor in RCC patients by univariate and multivariate Cox regression analyses. Univariate analysis revealed that age ($P < 0.001$), AJCC stage ($P < 0.001$), grade ($P < 0.001$), T stage ($P < 0.001$), M stage ($P < 0.001$), and glycolysis-related lncRNAs prognostic risk score ($P < 0.001$) were significantly correlated with survival time except gender ($P = 0.740$) (**Figure 3A**). Multivariate analysis showed that age ($P < 0.001$), AJCC stage ($P < 0.05$), grade ($P < 0.05$), and glycolysis-related lncRNAs prognostic risk score ($P < 0.001$) were significant associated with survival time (**Figure 3B**). Meanwhile, Multi-parameter ROC curve analyses showed that the AUC value of glycolysis-related lncRNAs prognostic risk score was 0.750 (**Figure 3C**). All these statistical data demonstrated that the glycolysis-related lncRNAs risk score predict the prognosis independently in patients of RCC.

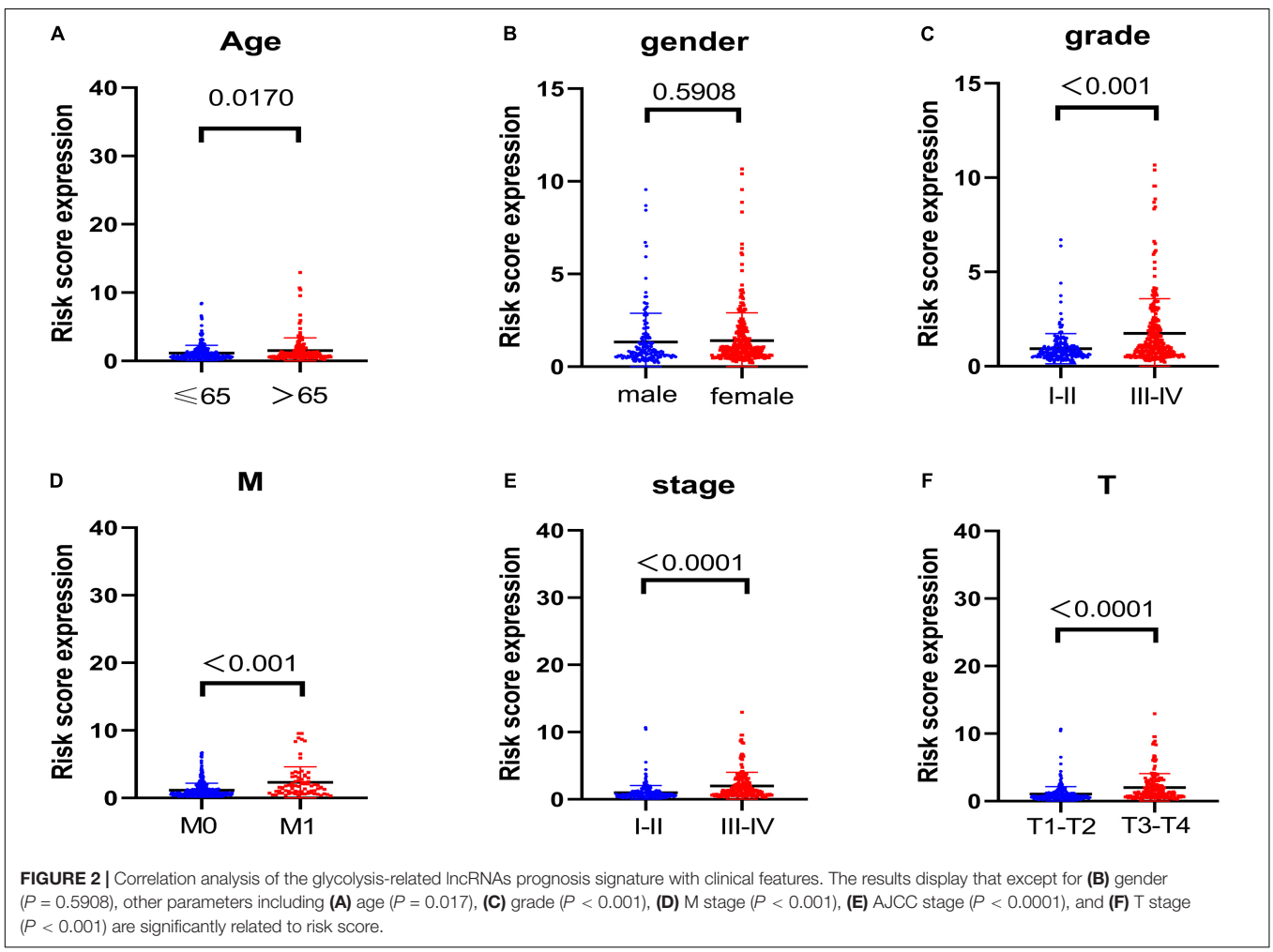


FIGURE 2 | Correlation analysis of the glycolysis-related lncRNAs prognosis signature with clinical features. The results display that except for **(B)** gender ($P = 0.5908$), other parameters including **(A)** age ($P = 0.017$), **(C)** grade ($P < 0.001$), **(D)** M stage ($P < 0.001$), **(E)** AJCC stage ($P < 0.0001$), and **(F)** T stage ($P < 0.001$) are significantly related to risk score.

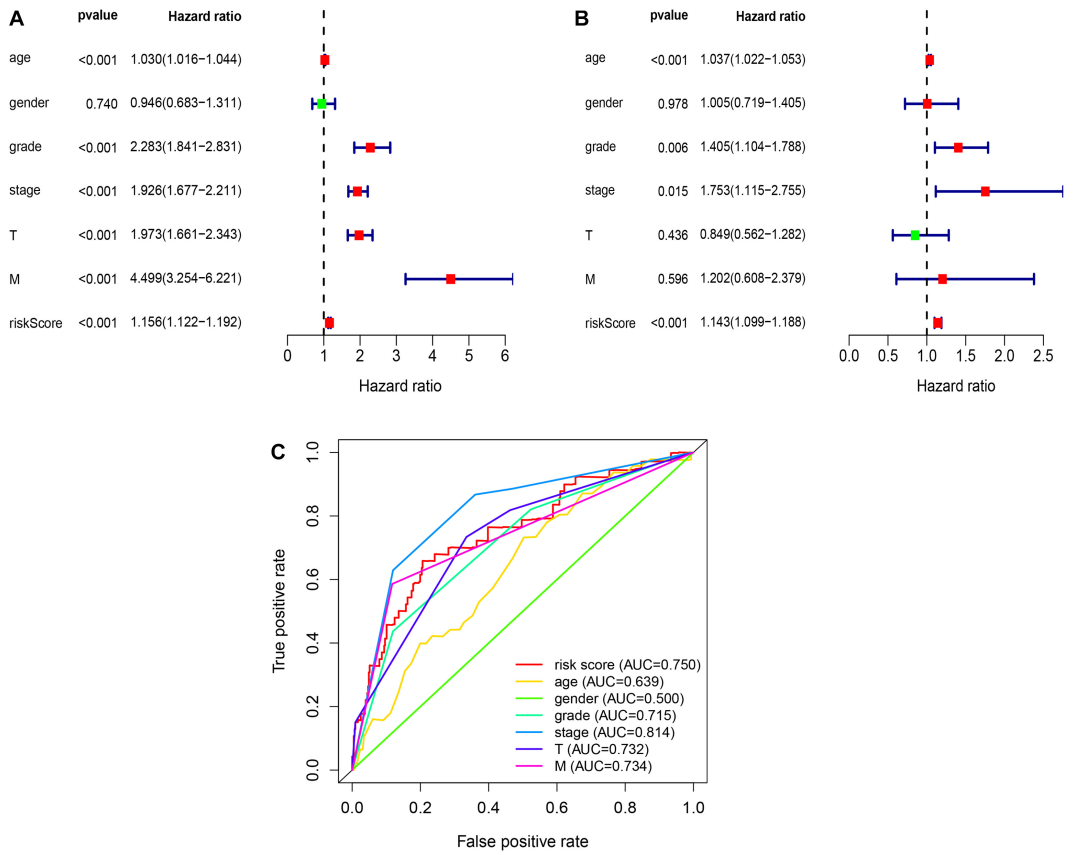


FIGURE 3 | The independent prediction of glycolysis-related lncRNAs risk signature. **(A)** Univariate analysis reveal that glycolysis-related lncRNAs prognostic risk scores ($P < 0.001$) and other clinical features were significantly correlated with survival time except gender ($P = 0.740$). **(B)** Multivariate analysis shows that age ($P < 0.001$), AJCC stage ($P < 0.05$), grade ($P < 0.05$), and glycolysis-related lncRNAs prognostic risk score ($P < 0.001$) were independent prognostic indicators. **(C)** Multi-parameter ROC curve analysis shows the different AUC values among the clinicopathological characteristics.

Stratification Analysis According to Other Clinicalpathological Features

In the whole cohort, we performed a stratification analysis based on the clinicopathological characteristics of RCC patients (e.g., tumor stage, age, gender, and stage). As shown in **Figure 4**, the Kaplan-Meier survival curve analysis showed that glycolysis-related lncRNAs risk score was still significantly correlated with OS rate among the patients with age > 65 or age < 65, female or male, high grade or low grade, metastasis or non-metastasis, advanced-stage (Stage III–IV) or early stage (Stage I–II) and T1-2 or T3-4. All these results clarified that the risk score was still an accurate model for predicting prognosis in different patient groups.

Quantification of Clinical Parameters and Evaluation of Prognostic Accuracy of Risk Score by Nomogram

In this section, the nomogram based on clinicalpathological parameters and glycolysis-related lncRNAs prognostic signature was applied to calculate the score for evaluating the precision of the prediction model. We constructed the nomogram (**Figure 5A**) combined multiple clinicalpathological characteristics including gender, age, T stage, AJCC stage, and the six glycolysis-related lncRNAs risk score to accurately estimate the 3-, and 5-year survival time by in RCC patients. The calibration curve analysis elucidated the concordance between predicted and actually observed 3- and 5-year OS of RCC patients (**Figures 5B,C**). Finally the time-dependent ROC (**Figure 5D**) showed that the AUC values of 3- and 5-year survival were 0.788 and 0.766, respectively. The above results verified that the risk score was reliable and accurate.

Gene Set Enrichment Analysis (GSEA)

To depict the potential pathway and their corresponding functions of the glycolysis-related lncRNAs signature implicated in RCC progression, we performed GSEA between the high-risk group and low-risk group. The findings revealed that in the high-risk group, glycolysis-related lncRNAs signature exhibited a significant enrichment in energy metabolism related signal pathway and cancer signal pathway including base excision repair, glycerophospholipid metabolism, homologous recombination, linoleic acid metabolism, and P53 signaling pathway (**Figure 6A**). In addition, GSEA in low-risk group

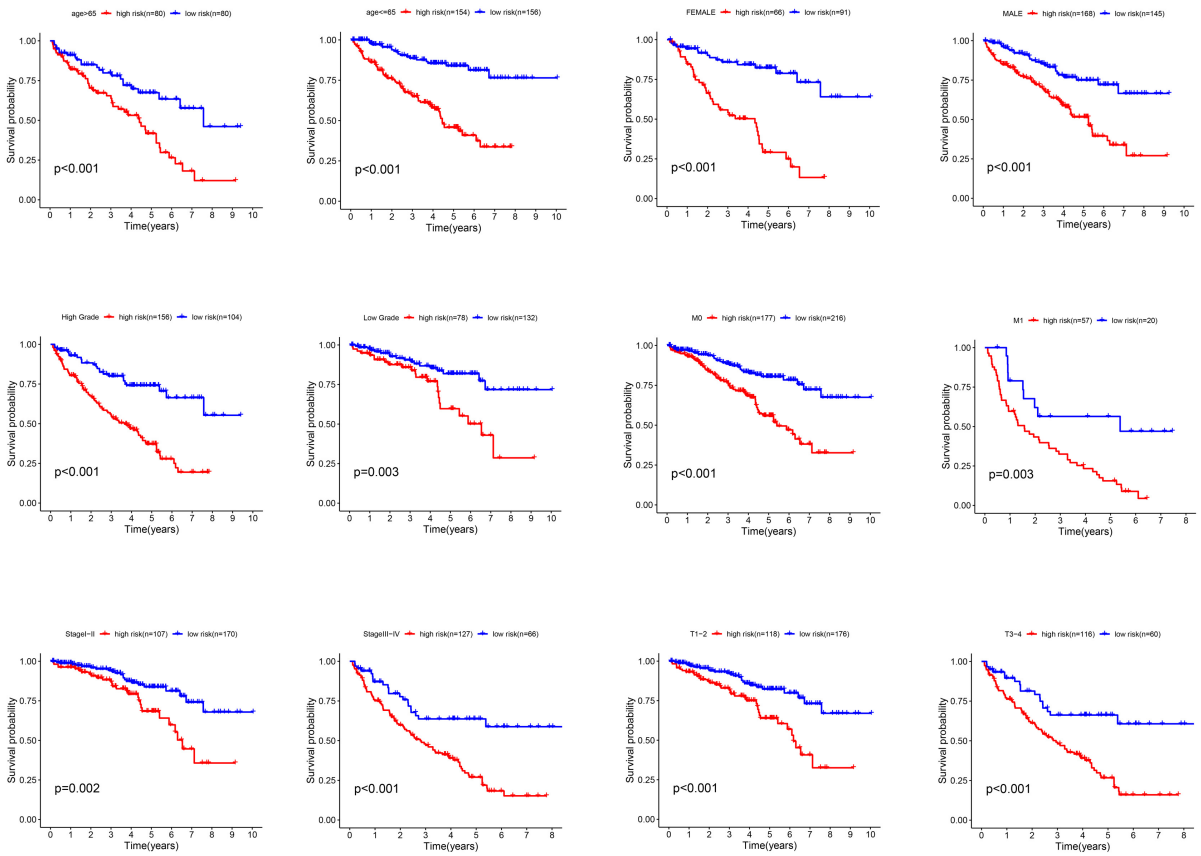


FIGURE 4 | The different survival time of high- and low-risk groups classified by clinicopathological parameters. The Kaplan-Meier survival curve analysis shows that the correlation of risk score and OS among the patients with different characteristics such as age > 65 or age < 65, female or male, high grade or low grade, metastasis or non-metastasis, advanced-stage (Stage III–IV) or early stage (Stage I–II) and T1-2 or T3-4.

(**Figure 6B**) identified that immune related signal pathway and Amino acid metabolism signal Pathway such as endocytosis, neurotrophin signaling pathway, propanoate metabolism, tryptophan metabolism, valine leucine, and isoleucine degradation which suggested that the glycolysis-related lncRNAs signature participated in immune-related regulation. Our study provided some valuable insights for future investigations to discover new individualized treatments and achieve full-process management of RCC patients in different risk groups.

Construction and Analysis of the LncRNA-mRNA Co-expression Network

In our subject, the lncRNA-mRNA co-expression network was constructed using Cytoscape to research the potential functions of the six glycolysis-related lncRNAs in renal carcinoma. We found 36 lncRNA-mRNA links among 6 lncRNAs and 25 related mRNAs according to Pearson correlation coefficient $|R| > 0.3$ and $P < 0.05$ (**Figure 7A**). The Sankey diagram exhibited the correlation between the 25 mRNAs and 6 lncRNAs (risk/protective) (**Figure 7B**). According to these, 6 lncRNAs significantly correlated with the 25 mRNAs in the prognostic signature. Meanwhile, the top GO term for molecular biological metabolic processes were isomerase activity and racemase and epimerase activity, acting on carbohydrates and derivatives (**Figure 7C**). Finally, the KEGG pathway analysis showed that amino sugar and nucleotide sugar metabolism was the most significantly enriched pathway (**Figure 7D**).

DISCUSSION

Renal cell carcinoma is a typical heterogeneous malignant tumor with different subtypes and clinical manifestations in the genitourinary tumor family (Zhai et al., 2019; Patel et al., 2020). The treatment of RCC is mainly characterized by radical resection (Greef and Eisen, 2016) for insensitive to radiotherapy and chemotherapy and it is difficult for clinician to choose adjuvant therapy for post-operative recurrence. The biggest trouble faced by clinicians is how to identify high-risk patients early and make corresponding treatment strategies (Pierorazio et al., 2016). Because the prognosis of RCC is not positively related to the size of the tumor, the traditional TNM staging and clinical imaging data cannot meet this requirement (Lallas et al., 2015; Zabell et al., 2018). It is urgent to find new biomarkers to make molecular diagnosis and guide clinical strategies.

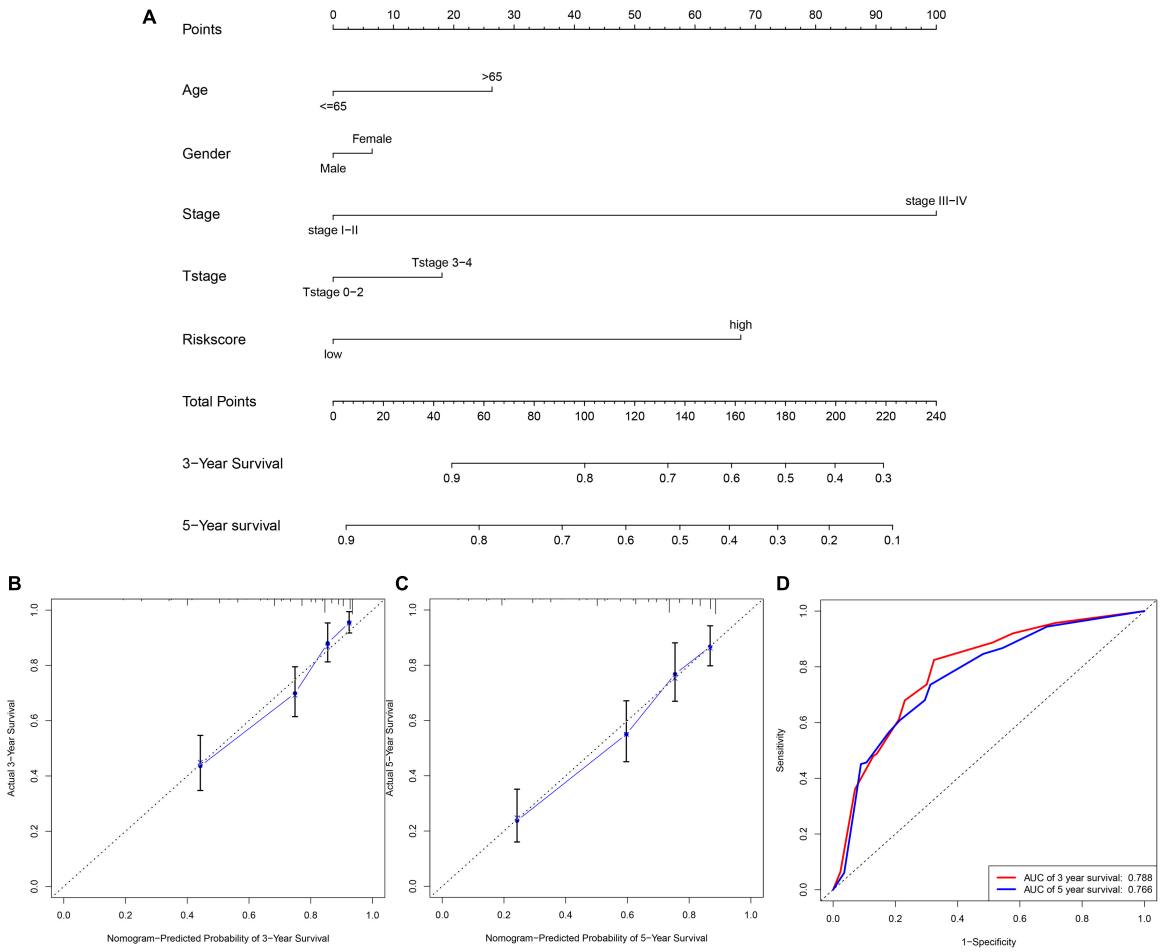


FIGURE 5 | The nomogram based on clinicalpathological parameters and risk score. **(A)** The nomogram shows that the risk score accurately estimate the 3-, and 5-year survival time. **(B,C)** Calibration curves elucidate the concordance between predicted and observed 3- and 5-year OS of RCC patients. **(D)** ROC curve shows the AUC values (0.788 and 0.766) of 3- and 5-year survival.

Previous studies have confirmed that lncRNAs as an important member of the non-coding RNA family, involved in the invasion and progression of RCC (Li D. et al., 2020). Shi et al. (2020) found that knocking down lncRNA DILA1 decreased Cyclin D1 protein expression and inhibited breast cancer cell growth both *in vitro* and *in vivo*. The lncRNA PVT1 is a novel oncogenic enhancer and its activity is controlled through aberrant methylation in colorectal cancer (Shigeyasu et al., 2020). MYC-mediated glycolysis of rectal tumor cells could be weakened with knocking out lncRNA LINRIS, which inhibited the progression of rectal cancer (Wang Y. et al., 2019). The glycolysis process was verified to be the energy source of tumor cell growth and invasion, which was regulated by downstream transcriptional activities mediated by lncRNA (Zhou et al., 2020). Glycolysis, as the main metabolic process in the cell metabolic microenvironment, participates in the proliferation of tumor cells. The glycolysis intermediates such as glucose 6-phosphate/pyruvate can participate in the regulation of fat and nucleic metabolism and other biological metabolic processes (Costanza et al., 2019; Orang et al., 2019; Smolle et al., 2020). In addition, the acidification of the microenvironment during glycolysis process is not conducive to the growth of normal cells, but is conducive to the infiltration and metastasis of tumor cells (Avella et al., 2020; Birts et al., 2020).

There are a lot of survival and prognosis analysis methods in the previous article. Some studies showed that "Gene mutations and copy number variants analysis" can be used to assess the difference between high- and low-risk groups (Xue et al., 2019). Some studies utilized logistic regression analysis and Cox proportional hazard analysis to verify the effect of prediction model (Kandimalla et al., 2020). In present study, we applied six glycolysis related lncRNAs (AC124854.1, AC078778.1, EMX2OS, DLGAP1-AS2, AC084876.1, and AC026401.3) which have not been previously reported to calculate the risk score, then confirmed the independent predictive ability by multivariate regression analysis with other clinical related parameters and verified the accuracy by ROC curve. Finally, the advantage of the predictive score of glycolysis related lncRNAs was shown in nomogram based on other clinical parameters (age, stage, T stage, etc.). GSEA enrichment analysis showed the enrichment of metabolic pathways in high- and low-risk patients, which further

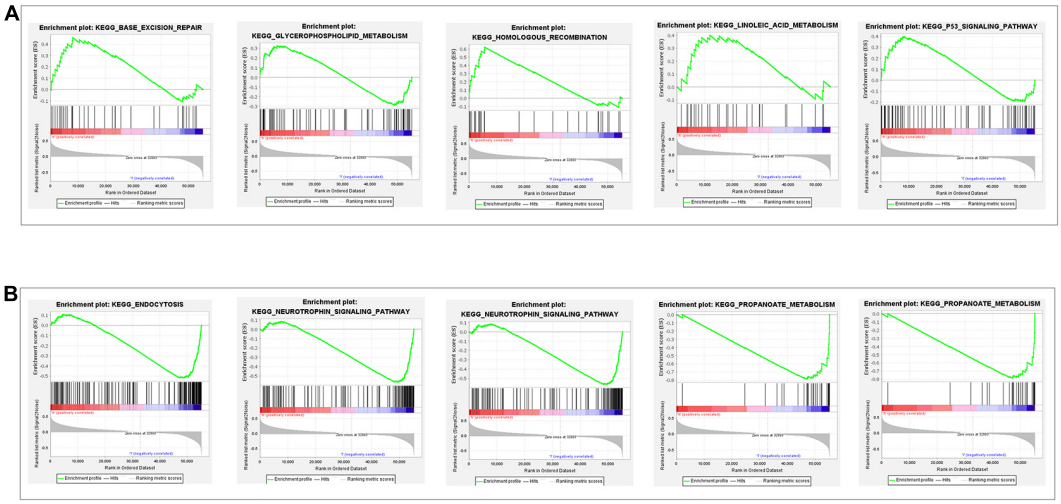


FIGURE 6 | Gene set enrichment analysis of glycolysis-related gene sets between the high- and low-risk groups. **(A)** The results show significant enrichment of energy metabolism related signal pathway and cancer signal pathway in the high-risk group. **(B)** The results show immune related signal pathway and amino acid metabolism signal pathway in the low-risk group.

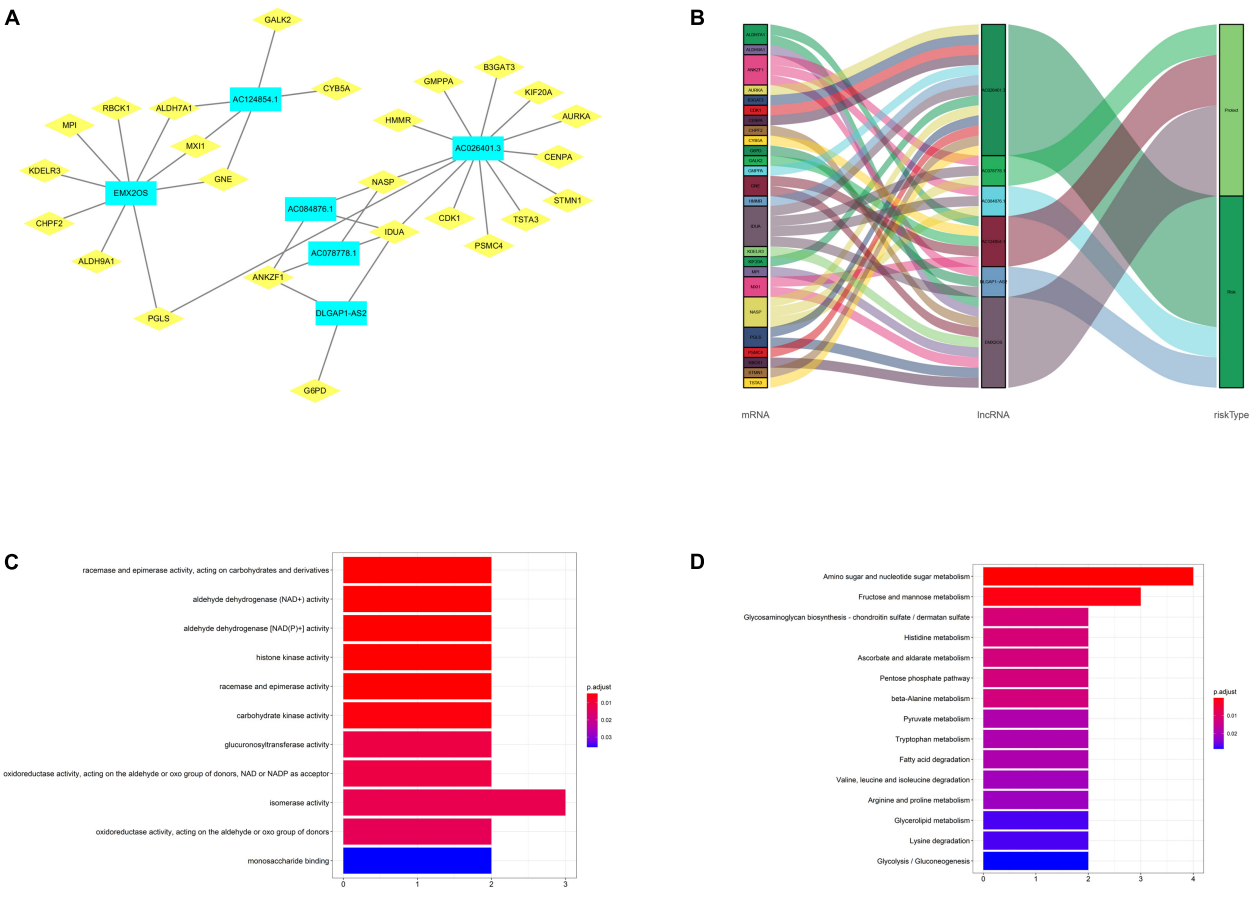


FIGURE 7 | Construction and analysis of the lncRNA-mRNA co-expression network. **(A)** The diagram shows 36 lncRNA-mRNA links among 6 lncRNAs and 25 related mRNAs. **(B)** The Sankey diagram exhibited the correlation between the 25 mRNAs and 6 lncRNAs (risk/protective). **(C)** The GO term shows the enriched molecular biological metabolic processes. **(D)** The KEGG pathway analysis shows the significantly enriched pathways.

confirmed that glycolysis-related lncRNAs played a regulatory role in energy metabolism. High-risk group analysis indicated fat metabolism and tumor signal pathway enrichment; low-risk group analysis showed amino acid metabolism pathway enrichment, which also suggested that amino acid metabolism was involved in the occurrence of RCC and provided a potential target for the treatment of RCC in the future. It is generally believed that lncRNAs does not directly encode proteins, but through a variety of pathways to regulate gene expression, so as to promote tumorigenesis and tumor metastasis (Athie et al., 2020). Our study explored the lncRNA-mRNA regulatory network and described the biological function and signal pathway. There were many signal pathways about lncRNA-mRNA regulation, among which the classical tri-molecular regulatory network (lncRNA-miRNA-TF/gene) had been widely described in other malignant tumors (Ye et al., 2018; Li A. et al., 2020). By sharing the common miRNA-binding sites, lncRNA affected the function of downstream target genes and regulated the progress of the disease (Abdollahzadeh et al., 2019). In present study, GO term showed that the differential mRNA in the lncRNA-mRNA co-expression network mainly affected the biological functions such as isomerase activity. KEGG pathway enrichment analysis showed that the differential mRNA in the co-expression network was highly enriched in amino sugar and fructose and nucleotide sugar metabolism and mannose metabolism. However, these pathways were rarely studied in the pathogenesis of renal cancer cells, which provided a new target for the research of tumor progression.

Note that, some problems of distribution deviation have also attracted our attention in this research. Multivariate regression analysis showed that high-risk lncRNA (such as AC026401.3) was highly expressed in low-risk patients, while low-risk lncRNA (such as AC078778.1) showed high expression in high-risk patients. Previous reports have shown that tumor cells show a high degree of heterogeneity (Baliu-Piqué et al., 2020; Louault et al., 2020), and there may be individualized internal mechanisms or even opposite biological processes in different tumor stages. Therefore, we suspect that this inconsistent result may be related to the stage of tumor progression, and there would be different metabolic conditions between local and progressive tumor.

There are several limitations in our subject. First, the transcriptome sequencing data and clinical information downloaded from the public TCGA database which might be incomplete or biased in this research. Moreover, considering the lack of external sequence verification of other databases, it is need to prove the robustness of lncRNAs signature. Finally, our study is a theoretical research based on bio-informatics and statistical analysis, biochemical researches and animal experiments are required to confirm the results in further investigations.

CONCLUSION

In summary, this study reveals that a glycolysis-based lncRNA signature (six candidate lncRNAs) exhibit diagnostic and prognostic values in RCC patients. Several related signaling pathway and metabolic relating activities may provide a potential target in clinical management and treatment of renal carcinoma. Since no relevant clinical data were found in other databases, we did not conduct external verification, which is indeed one of the shortcomings of this study. We will verify it in the clinical samples we collect later. At present, clinical patients are being followed up.

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/s.

AUTHOR CONTRIBUTIONS

HC contributed to design the study and draft the manuscript. HT and JZ performed data acquisition and analysis. TL and ZQ performed the bioinformatics and the statistical data analysis. CX and XL constructed the figures and tables. WH was responsible for the integrity of the entire study and manuscript review. All authors read and approved the final manuscript.

FUNDING

This study was supported by funds from the National Natural Science Foundation of China (Grant No. 81874092).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.638980/full#supplementary-material

REFERENCES

Abdollahzadeh, R., Daraei, A., Mansoori, Y., Sepahvand, M., Amoli, M., and Tavakkoly-Bazzaz, J. (2019). Competing endogenous RNA (ceRNA) cross talk and language in ceRNA regulatory networks: a new look at hallmarks of breast cancer. *J. Cell. Physiol.* 234, 10080–10100. doi: 10.1002/jcp.27941

Athie, A., Marchese, F., González, J., Lozano, T., Raimondi, I., Juvvuna, P., et al. (2020). Analysis of copy number alterations reveals the lncRNA ALAL-1 as a regulator of lung cancer immune evasion. *J. Cell. Biol.* 219:e201908078. doi: 10.1083/jcb.201908078

Avella, D., Manjunath, Y., Singh, A., Deroche, C., Kimchi, E., Staveley-O'Carroll, K., et al. (2020). F-FDG PET/CT total lesion glycolysis is associated with circulating tumor cell counts in patients with stage I to IIIA non-small cell lung cancer. *Transl. Lung Cancer Res.* 9, 515–521. doi: 10.21037/tlcr.2020.04.10

Baliu-Piqué, M., Pandiella, A., and Ocana, A. J. C. (2020). Breast cancer heterogeneity and response to novel therapeutics. *Cancers* 12:3271. doi: 10.3390/cancers12113271

Bandyopadhyay, S., Mallik, S., and Mukhopadhyay, A. (2014). A survey and comparative study of statistical tests for identifying differential expression from

microarray data. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 11, 95–115. doi: 10.1109/tcbb.2013.147

Birts, C., Banerjee, A., Darley, M., Dunlop, C., Nelson, S., Nijjar, S., et al. (2020). p53 is regulated by aerobic glycolysis in cancer cells by the CtBP family of NADH-dependent transcriptional regulators. *Sci. Signal.* 13:eaau9529. doi: 10.1126/scisignal.aau9529

Chen, F., Chen, J., Yang, L., Liu, J., Zhang, X., Zhang, Y., et al. (2019). Extracellular vesicle-packaged HIF-1alpha-stabilizing lncRNA from tumour-associated macrophages regulates aerobic glycolysis of breast cancer cells. *Nat. Cell Biol.* 21, 498–510. doi: 10.1038/s41556-019-0299-0

Cheng, G., Liu, D., Liang, H., Yang, H., Chen, K., and Zhang, X. J. A. (2019). A cluster of long non-coding RNAs exhibit diagnostic and prognostic values in renal cell carcinoma. *Aging (Albany N. Y.)* 11, 9597–9615. doi: 10.18632/aging.102407

Costanza, B., Rademaker, G., Tiamiou, A., De Tullio, P., Leenders, J., Blomme, A., et al. (2019). Transforming growth factor beta-induced, an extracellular matrix interacting protein, enhances glycolysis and promotes pancreatic cancer cell migration. *Int. J. Cancer* 145, 1570–1584. doi: 10.1002/ijc.32247

Greef, B., and Eisen, T. (2016). Medical treatment of renal cancer: new horizons. *Br. J. Cancer* 115, 505–516. doi: 10.1038/bjc.2016.230

Guo, F., Li, S., Guo, C., Xu, X., Zhou, X., Ma, D., et al. (2020). Circular RNA circMAGI3 accelerates the glycolysis of non-small cell lung cancer through miR-515-5p/HDGF. *Am. J. Transl. Res.* 12, 3953–3963.

Heo, J., Park, C., Ghosh, S., Park, S., Zivkovic, M., and Rascati, K. (2020). A network meta-analysis of efficacy and safety of first-line and second-line therapies for the management of metastatic renal cell carcinoma. *J. Clin. Pharm. Ther.* 46, 35–49. doi: 10.1111/jcpt.13282

Kandimalla, R., Shimura, T., Mallik, S., Sonohara, F., Tsai, S., Evans, D., et al. (2020). Identification of serum miRNA signature and establishment of a nomogram for risk stratification in patients with pancreatic ductal adenocarcinoma. *Ann. Surg.* doi: 10.1097/SLA.0000000000003945 [Epub ahead of print].

Khadirnaikar, S., Kumar, P., Pandi, S. N., Malik, R., Dhanasekaran, S. M., and Shukla, S. K. (2019). Immune associated LncRNAs identify novel prognostic subtypes of renal clear cell carcinoma. *Mol. Carcinog.* 58, 544–553. doi: 10.1002/mc.22949

Kim, N., Sung, N., Youn, H., and Park, S. (2020). Gremlin-1 activates Akt/STAT3 signaling, which increases the glycolysis rate in breast cancer cells. *Biochem. Biophys. Res. Commun.* 533, 1378–1384. doi: 10.1016/j.bbrc.2020.10.025

Lallas, C., Trabulsi, E., Kaffenberger, S., and Touijer, K. (2015). Treatment of exophytic renal cancer smaller than 3 cm: surgery versus active surveillance. *J. Urol.* 193, 16–18. doi: 10.1016/j.juro.2014.10.052

Li, A., Mallik, S., Luo, H., Jia, P., Lee, D., and Zhao, Z. (2020). H19, a long non-coding RNA, mediates transcription factors and target genes through interference of microRNAs in pan-cancer. *Mol. Ther. Nucleic Acids* 21, 180–191. doi: 10.1016/j.omtn.2020.05.028

Li, D., Li, C., Chen, Y., Teng, L., Cao, Y., Wang, W., et al. (2020). LncRNA HOTAIR induces sunitinib resistance in renal cancer by acting as a competing endogenous RNA to regulate autophagy of renal cells. *Cancer Cell Int.* 20:338. doi: 10.1186/s12935-020-01419-0

Lin, X., Feng, D., Li, P., and Lv, Y. (2020). LncRNA LINC00857 regulates the progression and glycolysis in ovarian cancer by modulating the Hippo signaling pathway. *Cancer Med.* 9, 8122–8132. doi: 10.1002/cam4.3322

Louault, K., Li, R., and DeClerck, Y. (2020). Cancer-associated fibroblasts: understanding their heterogeneity. *Cancers (Basel)* 12:3108. doi: 10.3390/cancers12113108

Michal, F., Febu, Júlia, B., Alexandr, P., and Dalibor, P. (2020). Long non-coding RNAs and renal cell carcinoma. *Klin. Onkol.* 33, 340–349. doi: 10.14735/amko2020340

Orang, A., Petersen, J., McKinnon, R., and Michael, M. (2019). Micromanaging aerobic respiration and glycolysis in cancer cells. *Mol. Metab.* 23, 98–126. doi: 10.1016/j.molmet.2019.01.014

Patel, V., Elias, R., Formella, J., Schwartzman, W., Christie, A., Cai, Q., et al. (2020). Acute interstitial nephritis, a potential predictor of response to immune checkpoint inhibitors in renal cell carcinoma. *J. Immunother. Cancer* 8:e001198. doi: 10.1136/jitc-2020-001198

Pierorazio, P., Johnson, M., Patel, H., Sozio, S., Sharma, R., Iyoha, E., et al. (2016). Management of renal masses and localized renal cancer: systematic review and meta-analysis. *J. Urol.* 196, 989–999. doi: 10.1016/j.juro.2016.04.081

Shi, Q., Li, Y., Li, S., Jin, L., Lai, H., Wu, Y., et al. (2020). LncRNA DILA1 inhibits Cyclin D1 degradation and contributes to tamoxifen resistance in breast cancer. *Nat. Commun.* 11:5513. doi: 10.1038/s41467-020-19349-w

Shigeyasu, K., Toden, S., Ozawa, T., Matsuyama, T., Nagasaka, T., Ishikawa, T., et al. (2020). The PVT1 lncRNA is a novel epigenetic enhancer of MYC, and a promising risk-stratification biomarker in colorectal cancer. *Mol. Cancer* 19:155. doi: 10.1186/s12943-020-01277-4

Smolle, E., Leko, P., Stacher-Priehse, E., Brcic, L., El-Heliebi, A., Hofmann, L., et al. (2020). Distribution and prognostic significance of gluconeogenesis and glycolysis in lung cancer. *Mol. Oncol.* 14, 2853–2867. doi: 10.1002/1878-0261.12780

Sun, Z., Jing, C., Xiao, C., and Li, T. (2020). An autophagy-related long non-coding RNA prognostic signature accurately predicts survival outcomes in bladder urothelial carcinoma patients. *Aging (Albany N. Y.)* 12, 15624–15637. doi: 10.18632/aging.103718

Wang, S., Zheng, W., Ji, A., Zhang, D., and Zhou, M. (2019). Overexpressed miR-122-5p promotes cell viability, proliferation, migration and glycolysis of renal cancer by negatively regulating PKM2. *Cancer Manag. Res.* 11, 9701–9713. doi: 10.2147/cmar.S225742

Wang, Y., Lu, J., Wu, Q., Jin, Y., Wang, D., Chen, Y., et al. (2019). LncRNA LINRIS stabilizes IGF2BP2 and promotes the aerobic glycolysis in colorectal cancer. *Mol. Cancer* 18:174. doi: 10.1186/s12943-019-1105-0

Xu, Q., Jia, X., Wu, Q., Shi, L., Ma, Z., Ba, N., et al. (2020). Esomeprazole affects the proliferation, metastasis, apoptosis and chemosensitivity of gastric cancer cells by regulating lncRNA/circRNA-miRNA-mRNA ceRNA networks. *Oncol. Lett.* 20:329. doi: 10.3892/ol.2020.12193

Xue, M., Shang, J., Chen, B., Yang, Z., Song, Q., Sun, X., et al. (2019). Identification of prognostic signatures for predicting the overall survival of uveal melanoma patients. *J. Cancer.* 10, 4921–4931. doi: 10.7150/jca.30618

Yang, F., Liu, C., Zhao, G., Ge, L., Song, Y., Chen, Z., et al. (2020). Long non-coding RNA LINC01234 regulates proliferation, migration and invasion via HIF-2α pathways in clear cell renal cell carcinoma cells. *PeerJ.* 8:e10149. doi: 10.7717/peerj.10149

Ye, Y., Li, S., and Wang, S. (2018). Construction and analysis of mRNA, miRNA, lncRNA, and TF regulatory networks reveal the key genes associated with prostate cancer. *PLoS One* 13:e0198055. doi: 10.1371/journal.pone.0198055

Zabell, J., Demirjian, S., Lane, B., Derweesh, I., Isharwal, S., Suk-Ouichai, C., et al. (2018). Predictors of long-term survival after renal cancer surgery. *J. Urol.* 199, 384–392. doi: 10.1016/j.juro.2017.08.096

Zhai, W., Zhu, R., Ma, J., Gong, D., Zhang, H., Zhang, J., et al. (2019). A positive feed-forward loop between LncRNA-URRCC and EGFL7/P-AKT/FOXO3 signaling promotes proliferation and metastasis of clear cell renal cell carcinoma. *Mol. Cancer* 18:81. doi: 10.1186/s12943-019-0998-y

Zhang, C., Gou, X., He, W., Yang, H., and Yin, H. (2020). A glycolysis-based 4-mRNA signature correlates with the prognosis and cell cycle process in patients with bladder cancer. *Cancer Cell Int.* 20:177. doi: 10.1186/s12935-020-01255-2

Zhang, Y., Zhang, L., Xu, Y., Wu, X., Zhou, Y., and Mo, J. (2020). Immune-related long noncoding RNA signature for predicting survival and immune checkpoint blockade in hepatocellular carcinoma. *J. Cell. Physiol.* 235, 9304–9316. doi: 10.1002/jcp.29730

Zhang, Z., Yang, W., Li, N., Chen, X., Ma, F., Yang, J., et al. (2020). LncRNA MCF2L-AS1 aggravates proliferation, invasion and glycolysis of colorectal cancer cells via the crosstalk with miR-874-3p/FOXM1 signaling axis. *Carcinogenesis* 42, 263–271. doi: 10.1093/carcin/bgaa093

Zhou, X., Gao, W., Hua, H., and Ji, Z. (2020). LncRNA-BLACAT1 facilitates proliferation, migration and aerobic glycolysis of pancreatic cancer cells by repressing CDKN1C via EZH2-induced H3K27me3. *Front. Oncol.* 10:539805. doi: 10.3389/fonc.2020.539805

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 28 May 2021
doi: 10.3389/fgene.2021.666833

MiR-3614-5p Is a Potential Novel Biomarker for Colorectal Cancer

*Lin Han[1], Yanjun Sun[2], Cansheng Lu[3], Chungeng Ma[4], Jian Shi[3] and Dengqun Sun[2]**

[1] Graduate School, Anhui University of Traditional Chinese Medicine, Hefei, China, [2] Department of General Surgery, The Armed Police Corps Hospital of Anhui, Hefei, China, [3] Department of Anus and Colon Surgery, The First Affiliated Hospital of Anhui University of Traditional Chinese Medicine, Hefei, China, [4] Department of Anus and Colon Surgery, The Second Affiliated Hospital of Anhui University of Traditional Chinese Medicine, Hefei, China

MiR-3614-5p has been found in a variety of cancers including colorectal cancer. However, the association of miR-3614-5p with colorectal cancer is still unclear. Based on the Cancer Genome Atlas (TCGA) database, the relationship between miR-3614-5p and colorectal cancer can be proved. Wilcoxon rank-sum test was used to compare the miR-3614-5p expression in colorectal cancer tissues and under normal conditions, respectively. The logistic regression method was further employed to analyze the relationship between miR-3614-5p and clinicopathological characteristics. Also, the correlation between miR-3614-5p and survival rate was evaluated by Kaplan-Meier and Cox regression analysis. Besides, gene set enrichment analysis (GSEA) was used to investigate the biological functions of miR-3614-5p. The decrease of miR-3614-5p expression of colorectal cancer was significantly correlated with N stage (OR) = 0.7 for N1&N2 vs. N0), M stage (OR = 0.5 for M1 vs. M0), pathologic stage (OR = 0.7 for Stage III & Stage IV vs. Stage I & Stage II), neoplasm type (OR = 0.5 for rectum adenocarcinoma vs. colon adenocarcinoma), and lymphatic invasion (OR = 0.6 for YES vs. NO) (all p-values < 0.05). Kaplan-Meier survival analysis showed that colorectal cancer with low miR-3614-5p has a poorer prognosis than that of high miR-3614-5p ($p = 0.005$). According to univariate analysis, low miR-3614-5p was associated with poor overall survival (OS) [hazard ratio (HR) = 0.599; 95% confidence interval (CI): 0.418-0.857; $p = 0.005$]. In multivariate analysis, miR-3614-5p was closely related to OS (HR = 0.630; 95% CI: 0.405-0.978, $p = 0.021$). GSEA showed that the high expression phenotype of miR-3614-5p differentially enriches the P53 pathway. Meanwhile, the high expression phenotype of miR-3614-5p enhanced NK T cell activation, negative T cell selection, response to interleukin 2, and response to tumor cells. MiR-3614-5p is a possible prognostic marker of low survival rate for patients with colorectal cancer. Moreover, the P53 pathway and P38MAPK pathway may be the key pathways regulated by miR-3614-5p in colorectal cancer.

Keywords: colorectal cancer, MiR-3614-5p, prognosis, TCGA, GSEA

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Christos K. Kontos,
National and Kapodistrian University of Athens, Greece
Saurav Mallik,
Harvard University, United States

****Correspondence:***
Dengqun Sun
sundengqunsyl@126.com

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *11 February 2021*
Accepted: *04 May 2021*
Published: *28 May 2021*

Citation:
Han L, Sun Y, Lu C, Ma C, Shi J and Sun D (2021) MiR-3614-5p Is a Potential Novel Biomarker for Colorectal Cancer. Front. Genet. 12:666833. doi: 10.3389/fgene.2021.666833

INTRODUCTION

Colorectal cancer (CRC), a common malignant tumor in the gastrointestinal tract/intestine or large intestine, is a major threat to health around the globe. Nowadays, CRC is the fourth fatal cancer and causes nearly 900,000 deaths annually worldwide. In addition to the aging population and unhealthy eating habits in high-income countries, adverse risk factors such as the lack of physical exercise, obesity, and smoking also increase the risk of CRC (Dekker et al., 2019). CRC usually occurs when healthy colonic epithelial cells become benign adenomas and eventually end up in malignancies (Binefa et al., 2014). Through quantitative analysis of CRC estimates that mutation from stem cells to malignant cells happens at first. Then it takes some time for these tumor cells to acquire metastatic ability. The window period is about 10 years (Jones et al., 2008). Therefore, early diagnosis of CRC is a challenging task for clinicians. According to statistics of CRC patients, nearly 43% of patients have liver metastases and 25% of patients have liver and lung metastases. Besides, the 5-year survival rate of stage IV is less than 10% (Chuang et al., 2017; Bora et al., 2021). Surgery remains to be the primary clinical treatment, supplemented by chemotherapy and immunotherapy. However, the efficacy of these treatments remains poor for the advanced stage of the disease. In summary, there is great urgency to develop new diagnostic and therapeutic targets.

MicroRNAs (miRNAs) are small non-coding RNAs with a size of 19-25 nucleotides (Bartel, 2004; Chen et al., 2019). Due to their short structure, they may be crucial biomarkers. They are easy to be used in the early detection and treatment of various cancers and have been proved workable in real practice (Kapp et al., 2015). Moreover, miRNA is of great use in regulating biological and pathological processes throughout the development of cancer (Bartel, 2004; Lin and Gregory, 2015).

Exosomes derived from tumor cells have an active role in carcinogenesis, metastasis, and response to treatment through the transfer of oncogenes and onco-miRNAs between CRS and tumor stroma cells (Nedaeinia et al., 2017). MiRNA, which is often dysregulated in cancer, has shown great potential as a tissue-based marker for cancer classification and prognosis (Hayes et al., 2014; Okugawa et al., 2015). Certain miRNAs could be secreted into the blood as cell-free miRNAs. They can be detected in serum as a highly stable form. Therefore, circulating miRNA has become a promising invasive biomarker for the diagnosis and monitoring of human cancers (Zen and Zhang, 2012; Huang et al., 2017).

MicroRNAs can act as biomarkers in colorectal cancer (Schee et al., 2010).

MiR-3614-5p with a length of 24 nt was located on chromosome 17q22. It has been reported that miR-3614-5p can antagonize the dengue virus by regulating adenosine deaminase which acts on RNA 1 (ADAR1) in human macrophages (Diosa-Toro et al., 2017). It is reported that miR-3614-5p plays a crucial role in the progress of non-small cell lung cancer. These findings may provide potential targets for the development of treatment strategies for patients with non-small cell lung cancer (Li et al., 2020). Besides, overexpression of miR-3614-5p significantly inhibited the proliferation of breast cancer cells (Wang et al., 2019). Bioinformatics analysis showed that miR-3614-5p may inhibit the WNT signaling pathway by targeting NFATC2 in NSCLC cells (Shang et al., 2019). Another report shows that there is a close relationship between the expression of miR-3614-5p and the risk of autoimmune diseases (Wohlers et al., 2018). Previous studies also used a miRNA risk-stratification signature that can be used as a non-invasive assay for the identification of high-risk patients and potential disease monitoring in patients with PDAC (Kandimalla et al., 2020; Lv et al., 2020). MicroRNAs are involved in many biological and pathological processes such as cell growth, differentiation, apoptosis, etc. Dysregulation of miRNAs expression patterns has been reported in many tumors including Colorectal Cancer. Various studies indicate that miRNAs can be utilized as diagnostic and prognostic biomarkers for evaluation of tumor initiation, development, invasion, metastasis, and response to chemotherapeutic drugs. Numerous investigations have also shown dysregulation of miRNAs in tissue samples and body fluids such as serum, plasma, and fecal samples from CRC patients (Shirafkan et al., 2018).

In this study, CRC-related miRNA-Seq data were acquired from the Cancer Genome Atlas (TCGA) for further accessing the prognostic value of the miR-3614-5p expression. It was found that CRC patients with miR-3614-5p have significantly lower expression. Thus, the miR-3614-5p can be a sensitive biomarker and prognosis standard.

MATERIALS AND METHODS

Sequencing Data and Clinic Information of miRNA From TCGA Data Repository

The level 3 BCGSC miRNA profiling miRNA-Seq data and clinical information were downloaded from the TCGA[1] colorectal cancer COAD and READ projects. Among the statistics, miRNA-Seq data without clinical information were discarded. A total of 602 miRNA-Seq data with clinical information were used for further association analysis (**Table 1**). The detailed clinicopathologic characteristics including age, pathologic stage, neoplasm type (rectum adenocarcinoma vs. colon adenocarcinoma), height, weight, gender, race, history of colon polyps, colon polyps present, and lymphatic invasion were recorded for analysis. Besides, other information for records included TP53 status, KRAS status, PIK3CA status, and TNM stage (TNM stage is the tumor stage, where T refers to the tumor, representing the range of primary tumor; N is a lymph node, showing whether there are lymph node metastasis and the range of the representative area; and M represents the existence of a distant transfer). This study does not include direct research of human participants or animals exerted by any authors.

Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) is a computational method and determines whether predetermined gene sets have statistically significant and consistent differences between two biological states (Subramanian et al., 2005). Based on the co-expression gene analysis of miR-3614-5p, GSEA was performed

[1]https://portal.gdc.cancer.gov/

TABLE 1 | TCGA colorectal cancer patient characteristics.

Clinical characteristics	Level	Overall (602)
T stage (%)	T1	20 (3.3%)
	T2	101 (16.8%)
	T3	411 (68.5%)
	T4	68 (11.3%)
N stage (%)	N0	337 (56.3%)
	N1	148 (24.7%)
	N2	114 (19.0%)
M stage (%)	M0	441 (83.4%)
	M1	88 (16.6%)
Pathologic stage (%)	Stage I	102 (17.5%)
	Stage II	216 (37.1%)
	Stage III	175 (30.1%)
	Stage IV	89 (15.3%)
Gender (%)	Female	285 (47.3%)
	Male	317 (52.7%)
Lymphatic invasion (%)	NO	322 (59.5%)
	YES	219 (40.5%)
History of colon polyps (%)	NO	354 (68.9%)
	YES	160 (31.1%)
Colon polyps present (%)	NO	208 (69.3%)
	YES	92 (30.7%)
Neoplasm type (%)	Colon adenocarcinoma	442 (73.4%)
	Rectum adenocarcinoma	160 (26.6%)
TP53 status (%)	Mut	306 (59.3%)
	WT	210 (40.7%)
KRAS status (%)	Mut	207 (40.1%)
	WT	309 (59.9%)
PIK3CA status (%)	Mut	130 (25.2%)
	WT	386 (74.8%)
Age (median [IQR])		68.00 (58.00, 76.00)
Height (median [IQR])		170.00 (162.00, 176.00)
Weight (median [IQR])		78.90 (65.00, 92.00)

IQR, InterQuartile Range; WT, wild type; MUT, mutant.

on the miR-3614-5p low and high expression groups through the TCGA colorectal cancer COAD and READ expression matrix, and clusterPorfiler package (Yu et al., 2012). In this study, an ordered list of all genes was made based on their correlation with miR-3614-5p expression via GSEA. Then GSEA was performed to clarify the significant observed survival differences between the high and low miR-3614-5p groups. We used a preliminary version of GSEA to analyze data (Mootha et al., 2003). Each analysis performs 10,000 permutations of the genome. The level of miR-3614-5p expression = functions as a phenotypic marker. By adjusting the *p*-value, the enrichment pathway for each phenotype was classified via the standardized enrichment score (NES).

Analysis of Immune Cell Characteristics by ssGSEA

The CRC immune infiltrate was analyzed by ssGSEA (single sample GSEA GSEA) and GSVA package in R (3.6.3) (Barbie et al., 2009). Thereby, GSEA was conducted on 24 types of immune cells in tumor samples, which include B cells, CD8 T cells, Eosinophils, Mast cells, Treg, NK CD56 bright cells, Th1 cells, immature DCs[iDCs], DCs, Th17 cells, T helper cells, NK cells, T cells, NK CD56 dim cells, Tgd, pDCs, Tfh, T effector memory[Tem], Cytotoxic cells, Tcm, Neutrophils, activated DCs[aDCs], Macrophages and Th2 cells. Then, according to the characteristic genes of 24 immune cells in the literature (Bindea et al., 2013), the relative enrichment fraction of each immune cell was quantified from the gene expression profile of each tumor sample. Spearman correlation and Wilcoxon rank-sum test analysis were used for analyzing the correlation between miR-3614-5p and immune cell infiltration level and the correlation between immune cell infiltration and different expression groups of miR-3614-5p.

Statistical Analysis

All statistical analyses were conducted using R (3.6.3). Wilcoxon rank sum test and logistic regression were used for analyzing the relationship between clinicopathological characteristics and miR-3614-5p. Cox regression and Kaplan-Meier method were employed in the analysis of clinical and pathological characteristics related to overall survival (OS) in TCGA patients.

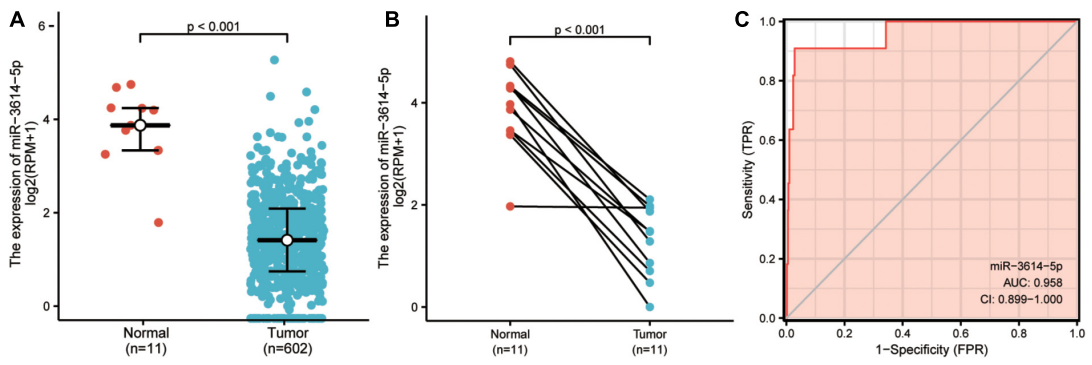


FIGURE 1 | The expression and diagnostic value of miR-3614-5p in colorectal tissues. **(A)** miR-3614-5p showed significantly lower expression in cancer tissues than in normal tissues. **(B)** miR-3614-5p was prominently lowexpressed in colorectal cancer ($P < 0.001$) compared with 11 pairs non-cancerous adjacent tissues. **(C)** Receiver Operating Characteristic Curve, FPR, False Positive Rate; TPR, True Positive Rate; CI, Confidence interval.

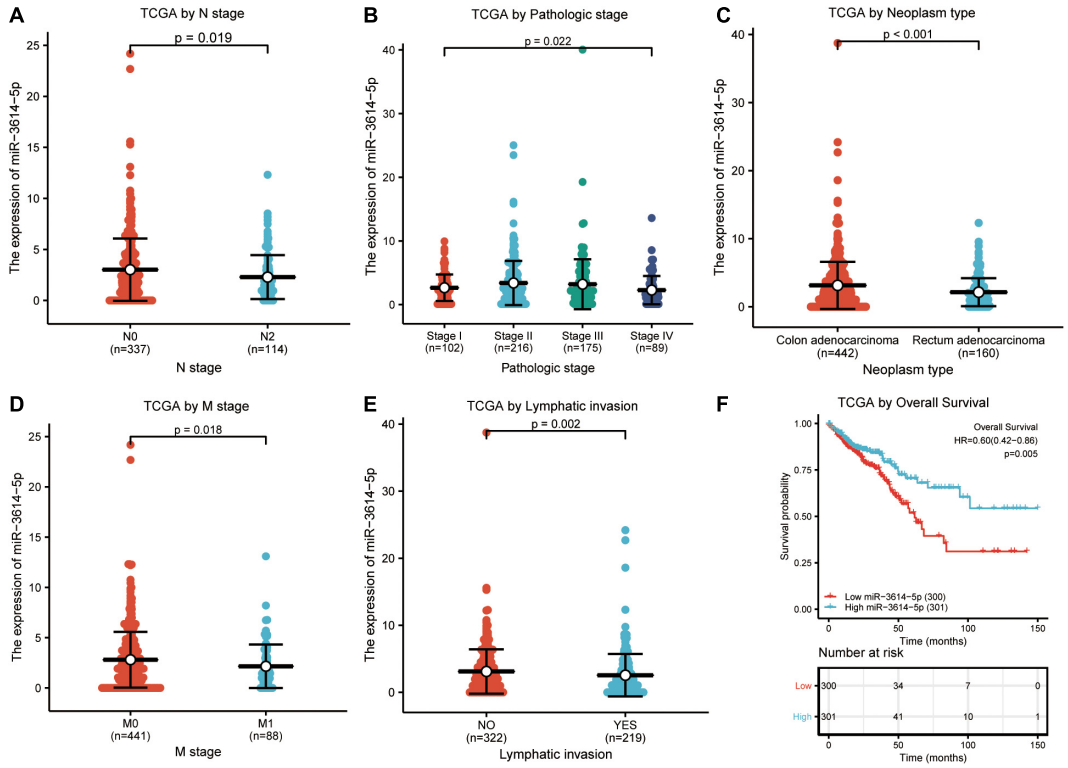


FIGURE 2 | Correlation analysis between miR-3614-5p expression and clinicopathological factors. **(A)** N stage. **(B)** Pathologic stage. **(C)** Neoplasm type. **(D)** M stage. **(E)** Lymphatic invasion. **(F)** Effect of miR-3614-5p expression on OS of CRC patients in TCGA cohort. TCGA, The Cancer Genome Atlas; OS, overall survival; CRC, colorectal cancer.

Multivariate Cox analysis was used to compare the effect of the miR-3614-5p expression on survival and other clinical characteristics (pathological stage, tumor type, gender, history of colon polyps, presence of colon polyps, lymphatic invasion, TP53 status, KRAS status, and PIK3CA status). The critical value of the miR-3614-5p expression was determined by its median value (in all tests, *p*-values less than 0.5 were considered significant).

TABLE 2 | miR-3614-5p expression* associated with clinical pathological characteristics (logistic regression).

Characteristics	Total (N)	Odds Ratio in miR-3614-5p expression	*P*-value
T stage (T3&T4 vs. T1&T2)	600	0.87 (0.58–1.30)	0.502
N stage (N1&N2 vs. N0)	599	0.69 (0.50–0.95)	0.023
M stage (M1 vs. M0)	529	0.52 (0.32–0.83)	0.007
Pathologic stage (Stage III&Stage IV vs. Stage I&Stage II)	582	0.68 (0.49–0.95)	0.023
Neoplasm type (Rectum adenocarcinoma vs. Colon adenocarcinoma)	602	0.50 (0.34–0.72)	<0.001
Lymphatic invasion (YES vs. NO)	541	0.62 (0.44–0.88)	0.007
TP53 status (Mut vs. WT)	516	1.08 (0.76–1.54)	0.652
KRAS status (Mut vs. WT)	516	0.94 (0.66–1.34)	0.730
PIK3CA status (Mut vs. WT)	516	1.04 (0.70–1.55)	0.844

**Categorical dependent variable, greater or less than the median expression level.*

RESULTS

Patients and Samples

As shown in **Table 1**, the characteristics of patients may affect survival rate. A total of 285 female patients and 317 male patients were analyzed in this study. Among them, patients with lymphatic invasion account for 219 (40.5%). 31.1% (n = 160) had a history of colon polyps before treatment, and 92 cases (30.7%) had colonic polyps. Following the statistical analysis of the pathologic stage, patients with stage I accounted for 17.5% (n = 102), while 37.1% (n = 216), 30.1% (n = 175), and 15.3% (n = 89) patients were in stage II, III, and IV, respectively. In the light of the anatomy, 442 tumors cases (73.4%) were colon adenocarcinoma, and 160 cases (26.6%) were rectal adenocarcinoma. Meanwhile, the topography distribution contained T stage: 3.293% T1 (n = 20), 16.8% T2 (n = 101), 68.5% T3 (n = 411), and 11.3% T4 (n = 68); N stage: 56.3% N0 (n = 337), 24.7% N1 (n = 148), 19.0% N2 (n = 114); and M stage: 83.4% M0 (n = 441), 16.6% M1 (n = 88). Besides, the expression of TP53, KRAS and PIK3CA status were, respectively, studied according to wild type (WT) and mutant (MUT).

TABLE 3 | Univariate and multivariate Cox proportional hazards regression analysis of miR-3614-5p expression.

Characteristics	Total (N)	HR (95% CI) univariate analysis	*P*-value univariate analysis	HR (95% CI) multivariate analysis	*P*-value multivariate analysis
T stage (T3&T4 vs. T1&T2)	599	2.383 (1.281–4.434)	0.006		
N stage (N1&N2 vs. N0)	598	2.630 (1.826–3.788)	<0.001		
Lymphatic invasion (YES vs. NO)	540	2.076 (1.422–3.032)	<0.001		
M stage (M1 vs. M0)	528	4.087 (2.741–6.092)	<0.001	2.126 (1.256–3.599)	0.005
Pathologic stage (Stage III&Stage IV vs. Stage I&Stage II)	581	2.999 (2.040–4.409)	<0.001	6.678 (1.985–22.461)	0.002
Age (> 65 vs. == 65)	601	2.034 (1.377–3.005)	<0.001	3.088 (1.908–4.997)	<0.001
miR-3614-5p (High vs. Low)	601	0.599 (0.418–0.857)	0.005	0.630 (0.405–0.978)	0.039

HR, hazard ratio.

The Expression and Diagnostic Value of miR-3614-5p in Colorectal Tissues

Next, Wilcoxon rank-sum test was used to examine the expression of miR-3614-5p in 602 CRC tissues and 11 normal tissues. The expression of miR-3614-5p in cancer tissues was much lower than that in normal ones ($P < 0.001$) (**Figure 1A**). Besides, the Wilcoxon single-rank test was employed to further analyze the expression of miR-3614-5p in 11 pairs of CRC tissues and healthy adjacent tissues. The result shows that miR-3614-5p was prominently low-expressed in CRC ($P < 0.001$) (**Figure 1B**), indicating that miR-3614-5p may facilitate the occurrence of CRC. From the TCGA database, ROC (receiver operating characteristics) was used for predicting the outcome of CRC and adjacent tissues. The diagnostic efficacy of miR-3614-5p for CRC was also analyzed. The area under the curve (AUC) of miR-3614-5p in **Figure 1C** is 0.958, which indicates that the expression of miR-3614-5p has a good discrimination ability in tumor and healthy tissues (Under the ROC curve area value from 0.5 to 1, the closer AUC is to 1, the better the diagnostic effect will be. AUC 0.5 ∼ 0.7 leads to lower accuracy, while AUC 0.7 ∼ 0.9 results in moderate accuracy, and higher accuracy can be guaranteed as AUC is higher than 0.9).

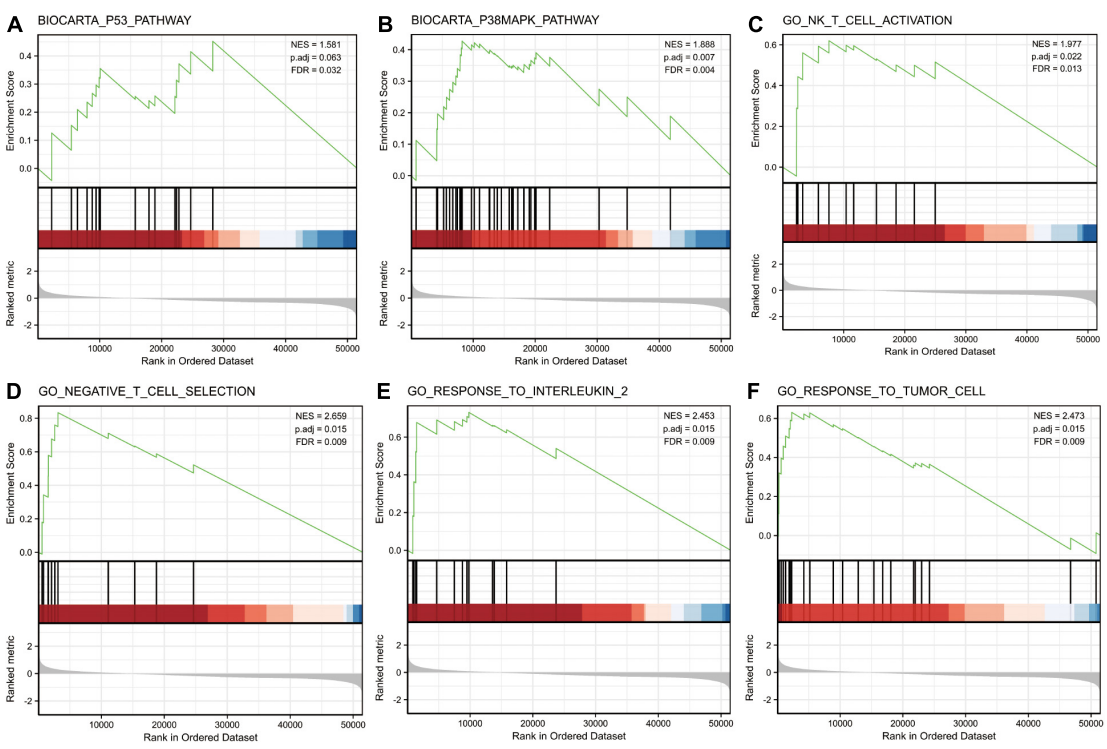


FIGURE 3 | Enrichment plots from gene set enrichment analysis (GSEA). GSEA results showing P53 pathway **(A)**, P38MAPK pathway **(B)**, NK T cell activation **(C)**, negative T cell selection **(D)**, response to interleukin 2 **(E)** and response to tumor cell **(F)** are differentially enriched in miR-3614-5p-related CRC. NES, normalized ES; p.adj, adjust *p*-value; FDR, false discovery rate.

TABLE 4 | Gene sets enriched in phenotype high.

MSigDB collection	Gene set name	NES	p.adj	FDR
c2.cp.v7.0.symbols.gmt	BIOCARTA_P53_PATHWAY	1.581	0.063	0.032
	BIOCARTA_P38MAPK_PATHWAY	1.888	0.007	0.004
c5.all.v7.0.symbols.gmt	GO_NK_T_CELL_ACTIVATION	1.977	0.022	0.013
	GO_NEGATIVE_T_CELL_SELECTION	2.659	0.015	0.009
	GO_RESPONSE_TO_INTERLEUKIN_2	2.453	0.015	0.009
	GO_RESPONSE_TO_TUMOR_CELL	2.473	0.015	0.009

NES, normalized enrichment score; p.adj, adjust p-value; FDR, false discovery rate.

TABLE 4-1 | The corresponding transcript names for each pathway in this table.

Gene set name	Corresponding transcript name
BIOCARTA_P53_PATHWAY	BCL2/MDM2/ATM/CDK2/APAF1/CCND1/RB1/CCNE1/CDKN1A/E2F1/PCNA/BAX/CDK4/GADD45A/TP53/TIMP3
BIOCARTA_P38MAPK_PATHWAY	STAT1/RPS6KA5/MAP3K5/PLA2G4A/CDC42/GRB2/MAX/HMGN1/MAP3K1/CREB1/MAP2K6/DDIT3/MAP2K4/ATF2
GO_NK_T_CELL_ACTIVATION	IL6R/IL12B/IL18/IL15/RASAL3/CD300A
GO_NEGATIVE_T_CELL_SELECTION	PTPRC/THEMIS/DOCK2/CD74/CCR7/CD3E/CD28
GO_RESPONSE_TO_INTERLEUKIN_2	CITED1/IL2/IL2RB/IL2RA/JAK3/PTPN2/JAK1/STAT5A/CDC5L
GO_RESPONSE_TO_TUMOR_CELL	CD274/CRTAM/KLRC4-KLRK1/CD226/KLRK1/PRF1/HAVCR2/CD160/IL12B

Association Between miR-3614-5p Expression Level and Clinicopathologic Variables

In total, 602 CRC samples with miR-3614-5p expression data were analyzed from TCGA, which covered all characteristics of the patients. As shown in **Figures 2A–E**, the low expression of miR-3614-5p was significantly correlated with tumor N stage (N2 vs. N0, $p = 0.019$), pathological stage (Stage I vs. Stage II vs. Stage III vs. Stage IV, $p = 0.022$), tumor type (Rectum adenocarcinoma vs. colon adenocarcinoma, $p < 0.001$), M stage (M1 vs. M0, $p = 0.018$), and lymphatic invasion (YES vs. NO, $p = 0.002$). The results of the univariate analysis which employs logistic regression show that miR-3614-5p expression as a categorical dependent variable (based on median expression value of 2.5) is related to poor prognostic clinicopathological characteristics (**Table 2**). Decreased expression of miR-3614-5p in CRC was related to the following factors: N stage (OR of N1 and N2 and N0 = 0.7), M stage (OR of M1 and M0 = 0.5), pathological stage (OR of stage III and IV = 0.7 and I Stage and II), tumor type (OR of rectal adenocarcinoma and colon adenocarcinoma = 0.5), and lymphatic infiltration (OR of YES vs. NO = 0.6) (all $p < 0.05$). These results indicate that compared with people with high miR-3614-5p expression, people with low miR-3614-5p expression are more likely to enter the late stage and may have a lymphatic invasion.

Univariate and Multivariate Survival Analyses

According to **Figure 2F**, Kaplan-Meier survival analysis shows that CRC with low miR-3614-5p was associated with a worse prognosis than CRC with high miR-3614-5p ($p = 0.005$). Univariate analysis shows that low miR-3614-5p was closely related to worse OS [hazard ratio (HR): 0.599; 95% confidence interval (CI): 0.418–0.857; $p = 0.005$]. Other clinicopathological variables related to low survival rate include TNM stage, pathological stage, age, and lymphatic invasion. In multivariate analysis, miR-3614-5p was independently associated with OS, HR of 0.630 (95% CI: 0.405–0.978, $p = 0.021$), as well as M stage, pathological stage, and age (**Table 3**).

GSEA Identifies a miR-3614-5p-Related Signaling Pathway

GSEA was performed on the low and high miR-3614-5p expression datasets to identify signaling pathways that are differentially activated in CRC. It revealed significant differences in the enrichment of MSigDB collections (c2.cp.v7.0. and c5.all.v7.0. symbols) (FDR < 0.05, NOM p-val < 0.05). The most enriched signaling pathways were selected based on the normalized enrichment score (NES) (**Figure 3** and **Table 4**). **Figures 3A–F** illustrates that the high expression phenotype of miR-3614-5p differentially enriches the P53 pathway, P38MAPK pathway, NK T cell activation, negative T cell selection, response to interleukin 2, and response to tumor cells.

The Correlation Between miR-3614-5p Expression and Immune Infiltration

Then, the correlation between the expression level (TPM) of miR-3614-5p and immune cell enrichment level (generated by ssGSEA) was analyzed by Spearman correlation. As a result, the miR-3614-5p expression was found negatively correlated with the abundances of immunocytes (Eosinophils, iDCs, NK CD56bright cells, etc.), while positively correlated with the abundances of immunocytes (Th2 cells, aDCs, CD8 T cells, Th1 cells, Cytotoxic cells, T cells, etc.) (**Figure 4A**). Wilcoxon rank sum test also showed that the enrichment score of Th2 cells was significantly higher in miR-3614-5p high expression samples (**Figure 4B**). Furthermore, the difference in Th2

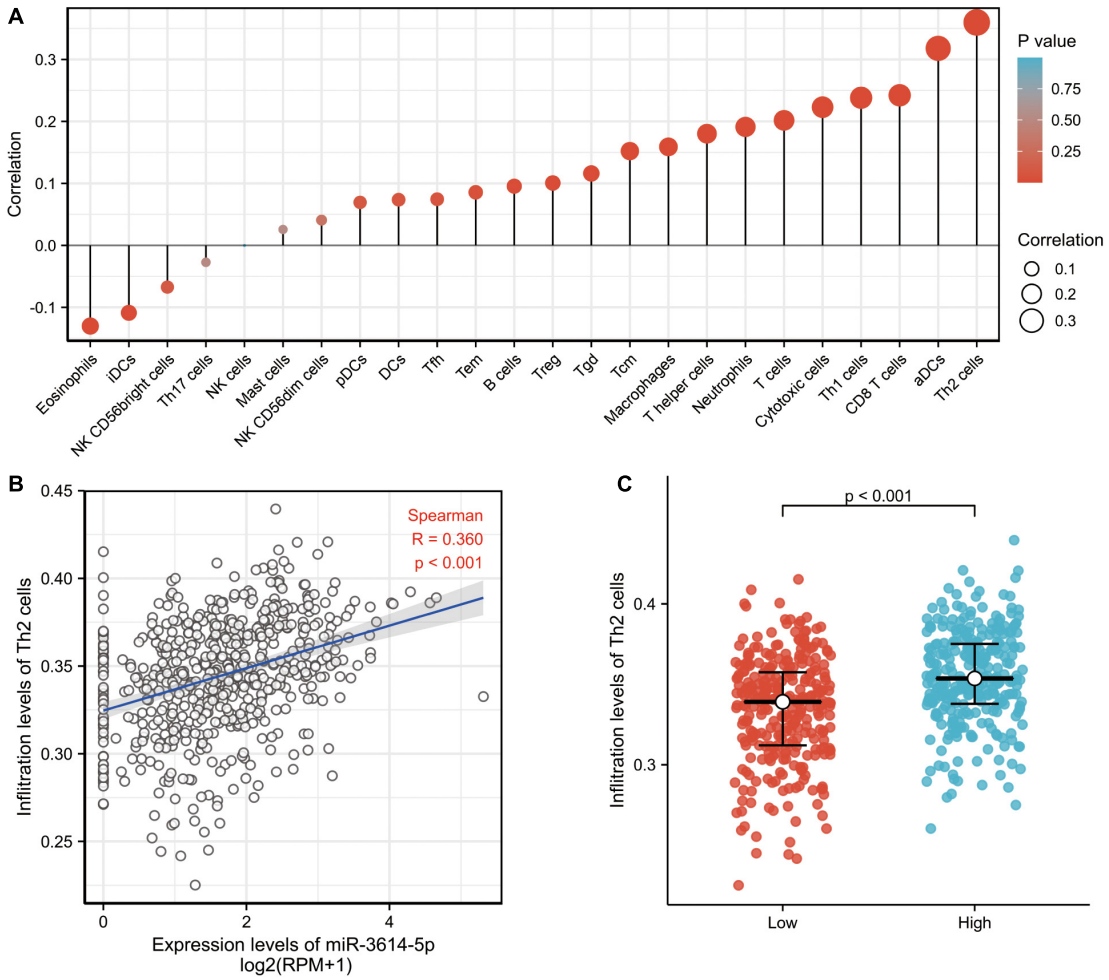


FIGURE 4 | The correlation between miR-3614-5p expression and immune invitation. **(A)** The forest plot shows the correlation between miR-3614-5p expression and immune cell subsets. **(B)** Correlation between the relative enrichment score of Th2 cells and the expression level (TPM) of miR-3614-5p. **(C)** Difference in Th2 cells infiltration level between miR-3614-5p high and low expression groups.

cells infiltration level was analyzed in miR-3614-5p high and low expression groups. The results are statistically significant ($p < 0.001$) (**Figure 4C**).

Compute the Matthews Correlation Coefficient (MCC)

Simultaneously, we recalculate the matthews correlation coefficient (MCC) value by querying the data on colorectal cancer in TCGA (Bandyopadhyay et al., 2014). Use the calculation formula of MCC to substitute the value into the formula. After the calculation, the MCC value is 0, it can be understood as an average random prediction, which can better reflect a correct prediction in this study (the MCC is in essence a correlation coefficient between the observed and predicted binary classifications; it returns a value between -1 and + 1. A coefficient of + 1 represents a perfect prediction, 0 an average random prediction and -1 an inverse prediction).

DISCUSSION

According to recent studies, miR-3614-5p plays an essential role in various cancers. For example, it can be concluded that overexpression of CHAIP2 may hinder LUAD cell proliferation and invasion due to the modulation of the WNT signal pathway targeted by miR-3614-5p (Shang et al., 2019). The genome-wide association study (GWAS) signal for rheumatoid arthritis and Crohn's disease were enriched in genes predicted to be targeted by miR-3614-5p. What's more, it points out the potential pathophysiological role of miR-3614-5p in autoimmunity (Wohlers et al., 2018). However, few studies have been carried out on the correlation between miR-3614-5p and CRC. However, little correlation has been proved between miR-3614-5p and CRC. Therefore, this paper is aimed at clarifying the expression of miR-3614-5p in CRC tissue, and its potential therapeutic and prognostic value. In this research, the miR-3614-5p was significantly downregulated in CRC tissue compared to normal or adjacent normal tissue. Thus, the potential

role of miR-3614-5p low expression in CRC is the focus of the present study.

Herein, CRC data based on high throughput RNA sequencing were collected from the TCGA database. Meanwhile, it was demonstrated that miR-3614-5p was significantly down-regulated in CRC tissues compared with normal or adjacent normal tissues. It is proved that reduced expression of miR-3614-5p in CRC is associated with clinical pathologic characteristics in advanced periods (age, clinical stage, pathologic stage, histological type, lymphatic invasion), short survival time, and poor prognosis. Furthermore, GSEA results show that miR-3614-5p phenotypes with high expression are significantly related to the P53 pathway, P38MAPK pathway, NK T cell activation, negative T cell selection, response to interleukin 2, and response to the tumor cell. These pathways are said to be responsible for the proliferation of cancer cells, invasion, and metastasis (Fan et al., 2014; Wang et al., 2017; Kimura, 2018; Yu-Lee et al., 2018; Najafi et al., 2019). The results suggest that the miR-3614-5p may be a new therapeutic and prognostic target for CRC. Studies have shown that Brahma-related gene 1 (BRG1) plays an important role in cell aging and tumor growth. Thus, it promotes a new mechanism of cell senescence in CRC by affecting the p53 signal axis, which is a new potential target for cancer treatment. Current studies have shown that there are functionally important single nucleotide polymorphisms (SNPs) in certain genes of the pathway p53. They can change the signal transduction amplitude of the protein. Besides, those variants may influence cancer risk, progression, and efficacy of radiation and chemotherapy. Besides, the p53 pathway is of great significance in other biological processes, including metabolism and reproductive adaptability. Thereby, these variants also have the potential to change other diseases (Basu and Murphy, 2016). Studies have also shown that the migration and invasion of human gastric cancer cells SGC7901 can be inhibited by the P38MAPK signaling pathway through the expression of MMP-2 and MMP-9 (Lu et al., 2017). This study further tests the phosphorylation of P38MAPK to completely stratify the prognosis, showing the combined survival advantage of P38MAPK in cMMR BRAF mutant stage III CRC patients (Roseweir et al., 2018). Studies have shown that Barbaloin increases the apoptotic rate of A549 cells by reducing cell growth and Ki-67 expression levels while proliferates cell nuclear antigen (PCNA).

Besides, Barbaloin induces G2/M phase accumulation, thereby inactivating the P38MAPK signaling pathway. In the end, the proliferation and metastasis of small cell lung cancer can be inhibited (Zhang et al., 2017). Phosphorylated p38 (p-p38) is important in the regulation of disease progression. It is related to tumor prognosis. Poor OS of CRC can be predicted by overexpression of p-p38, which is also an important independent factor leading to death, recurrence, and distant metastasis (Fan et al., 2014). Studies have shown that the overexpression of methyltransferase 3 (METTL3) exerts a tumor suppressor effect on the proliferation, migration, and invasion of CRC cells through the p38 pathway (Deng et al., 2019). This indicates that miR-3614-5p can affect the growth, metastasis, and low survival rate of colorectal cancer cells through the P53 and P38MAPK pathways.

Another important aspect of this study is that the expression of miR-3614-5p is correlated with multiple levels of immune infiltration in CRC. What's more, the results also indicate that in CRC expression, there is a moderate to a strong positive correlation between the expression level of miR-3614-5p and the infiltration level of Th2 cells and aDC. Besides, the infiltration level of CD8 T cells, Th1 cells, cytotoxic cells, and T cells have significant positive correlations with the expression of miR-3614-5p. Besides, the correlation between miR-3614-5p expression and immune cell marker genes suggests the role of miR-3614-5p in regulating CRC tumor immunology. First, iDC has a weak correlation with miR-3614-5p expression, while aDC shows a strong correlation. These results reveal the potential regulatory role of miR-3614-5p in tumor-related DC polarization. In addition, a significant correlation can be found between the expression of miR-3614-5p and the regulation of several markers of T helper cells in CRC (Th2, Th1, Tfh). These correlations may indicate the potential mechanism by which miR-3614-5p regulates T cell function in CRC. These findings collectively indicate that miR-3614-5p may largely affect the recruiting and regulating of immune infiltrating cells in CRC.

In this paper, these methods deepened our understanding of the relationship between miR-3614-5p and CRC. However, further improvements need to be made. First, to fully reveal the specific role of miR-3614-5p in the occurrence of CRC, various clinical factors, such as detailed information about the patients being treated should be considered. However, considering that the experiments are conducted in multiple laboratories, the processing methods in the public database may be inconsistent. Second, the number of healthy subjects used for controls in this study is significantly variant from the number of cancer patients. As a consequence, follow-up studies are needed to maintain a balanced sample size. Finally, multi-center studies in public databases are designed to make up for the shortcomings of single-center studies. However, there are also limitations in retrospective studies, especially differences in interventions and the lack of specific information. As a result, prospective studies are supposed to be conducted in the future to avoid analysis bias due to the retrospective nature of the current research. Moreover, it is impossible to clearly assess the direct mechanism of miR-3614-5p involved in CRC development. Therefore, future research should carry out wet experiments on the direct mechanism of CRC.

CONCLUSION

According to our study, low miR-3614-5p expression is closely associated with CRC cancer progression, low survival rate, and immune infiltration. Thus, it can promote tumor formation through abnormal inflammation and immune responses. This study provides promising insights to exhume the clinical-pathological significance and molecular etiology of CRC. However, further randomized clinical trials and supplementary studies are needed to validate the basic molecular mechanism and clinical application of CRC patients.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

AUTHOR CONTRIBUTIONS

DS: conceptualization and design of the study and funding. LH, YS, and CL: details of the experimental design. LH: carrying out the experiment, preparation of the data. CM, JS, YS, and LH: implementation of the scenario. DS and LH: analysis. LH, DS, and YS: writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors would like to acknowledge the financial supports from the Natural Science Foundation of Anhui Province (No. 1808085MH238).

REFERENCES

Bandyopadhyay, S., Mallik, S., and Mukhopadhyay, A. (2014). A survey and comparative study of statistical tests for identifying differential expression from microarray data. *IEEE ACM Trans. Comput. Biol. Bioinf.* 11, 95–115. doi: 10.1109/TCBB.2013.147

Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462, 108–112. doi: 10.1038/nature08460

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297. doi: 10.1016/s0092-8674(04)00045-5

Basu, S., and Murphy, M. E. (2016). Genetic modifiers of the p53 pathway. *Cold Spring Harb. Perspect. Med.* 6:a026302. doi: 10.1101/cshperspect.a026302

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39, 782–795. doi: 10.1016/j.immuni.2013.10.003

Binefa, G., Rodríguez-Moranta, F., Teule, A., and Medina-Hayas, M. (2014). Colorectal cancer: from prevention to personalized medicine. *World J. Gastroenterol.* 20, 6786–6808. doi: 10.3748/wjg.v20.i22.6786

Bora, K., Bhuyan, M. K., Kasugai, K., Mallik, S., and Zhao, Z. (2021). Computational learning of features for automated colonic polyp classification. *Sci. Rep.* 11:4347.

Chen, B., Xia, Z., Deng, Y.-N., Yang, Y., Zhang, P., Zhu, H., et al. (2019). Emerging microRNA biomarkers for colorectal cancer diagnosis and prognosis. *Open Biol.* 9, 180212–180212. doi: 10.1098/rsob.180212

Chuang, H. Y., Jiang, J. K., Yang, M. H., Wang, H. W., Li, M. C., Tsai, C. Y., et al. (2017). Aminopeptidase A initiates tumorigenesis and enhances tumor cell stemness via TWIST1 upregulation in colorectal cancer. *Oncotarget* 8, 21266–21280. doi: 10.18632/oncotarget.15072

Dekker, E., Tanis, P. J., Vleugels, J., Kasi, P. M., and Wallace, M. B. (2019). Colorectal cancer. *Lancet* 394, 1467–1480. doi: 10.1016/S0140-6736(19)32319-0

Deng, R., Cheng, Y., Ye, S., Zhang, J., Huang, R., Li, P., et al. (2019). m6A methyltransferase METTL3 suppresses colorectal cancer proliferation and migration through p38/ERK pathways. *OncoTargets Ther.* 12, 4391–4402. doi: 10.2147/OTT.S201052

Diosa-Toro, M., Echavarría-Consuegra, L., Flipse, J., Fernández, G. J., Kluiver, J., van den Berg, A., et al. (2017). MicroRNA profiling of human primary macrophages exposed to dengue virus identifies miRNA-3614-5p as antiviral and regulator of ADAR1 expression. *PLoS Neglect. Trop. Dis.* 11:e0005981. doi: 10.1371/journal.pntd.0005981

Fan, X. J., Wan, X. B., Fu, X. H., Wu, P. H., Chen, D. K., Wang, P. N., et al. (2014). Phosphorylated p38, a negative prognostic biomarker, complements TNM staging prognostication in colorectal cancer. *Tumour Biol.* 35, 10487–10495. doi: 10.1007/s13277-014-2320-3

Hayes, J., Peruzzi, P. P., and Lawler, S. (2014). MicroRNAs in cancer: biomarkers, functions and therapy. *Trends Mol. Med.* 20, 460–469. doi: 10.1016/j.molmed.2014.06.005

Huang, G., Li, S., Yang, N., Zou, Y., Zheng, D., and Xiao, T. (2017). Recent progress in circular RNAs in human cancers. *Cancer Lett.* 404, 8–18. doi: 10.1016/j.canlet.2017.07.002

Jones, S., Chen, W. D., Parmigiani, G., Diehl, F., Beerenwinkel, N., Antal, T., et al. (2008). Comparative lesion sequencing provides insights into tumor evolution. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4283–4288. doi: 10.1073/pnas.0712345105

Kandimalla, R., Shimura, T., Mallik, S., Sonohara, F., Tsai, S., Evans, D. B., et al. (2020). identification of serum mirna signature and establishment of a nomogram for risk stratification in patients with pancreatic ductal adenocarcinoma. *Ann. Surg.* [Epub Ahead of Print].

Kapp, J. R., Diss, T., Spicer, J., Gandy, M., Schrijver, I., Jennings, L. J., et al. (2015). Variation in pre-PCR processing of FFPE samples leads to discrepancies in BRAF and EGFR mutation detection: a diagnostic RING trial. *J. Clin. Pathol.* 68, 111–118. doi: 10.1136/jclinpath-2014-202644

Kimura, H. (2018). EBV in T-/NK-Cell Tumorigenesis. *Adv. Exp. Med. Biol.* 1045, 459–475. doi: 10.1007/978-981-10-7230-7_21

Li, F., Yang, H., Kong, T., Chen, S., Li, P., Chen, L., et al. (2020). PGAM1, regulated by miR-3614-5p, functions as an oncogene by activating transforming growth factor-β (TGF-β) signaling in the progression of non-small cell lung carcinoma. *Cell Death Dis.* 11:710. doi: 10.1038/s41419-020-02900-4

Lin, S., and Gregory, R. I. (2015). MicroRNA biogenesis pathways in cancer. *Nat. Rev. Cancer* 15, 321–333. doi: 10.1038/nrc3932

Lu, S., Zhang, Z., Chen, M., Li, C., Liu, L., and Li, Y. (2017). Silibinin inhibits the migration and invasion of human gastric cancer SGC7901 cells by downregulating MMP-2 and MMP-9 expression via the p38MAPK signaling pathway. *Oncol. Lett.* 14, 7577–7582. doi: 10.3892/ol.2017.7080

Lv, Y., Duanmu, J., Fu, X., Li, T., and Jiang, Q. (2020). Identifying a new microRNA signature as a prognostic biomarker in colon cancer. *PLoS One* 15:e0228575. doi: 10.1371/journal.pone.0228575

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34, 267–273. doi: 10.1038/ng1180

Najafi, M., Farhood, B., and Mortezaee, K. (2019). Contribution of regulatory T cells to cancer: a review. *J. Cell. Physiol.* 234, 7983–7993. doi: 10.1002/jcp.27553

Nedaeinia, R., Manian, M., Jazayeri, M. H., Ranjbar, M., Salehi, R., Sharifi, M., et al. (2017). Circulating exosomes and exosomal microRNAs as biomarkers in gastrointestinal cancer. *Cancer Gene Ther.* 24, 48–56. doi: 10.1038/cgt.2016.77

Okugawa, Y., Grady, W. M., and Goel, A. (2015). Epigenetic alterations in colorectal cancer: emerging biomarkers. *Gastroenterology* 149, 1204.e12–1225.e12. doi: 10.1053/j.gastro.2015.07.011

Roseweir, A. K., Halcrow, E. S., Chichilo, S., Powell, A. G., McMillan, D. C., Horgan, P. G., et al. (2018). ERK and p38MAPK combine to improve survival in patients with BRAF mutant colorectal cancer. *Br. J. Cancer* 119, 323–329. doi: 10.1038/s41416-018-0174-y

Schee, K., Fodstad, Ø, and Flatmark, K. (2010). MicroRNAs as biomarkers in colorectal cancer. *Am. J. Pathol.* 177, 1592–1599.

Shang, J., Wang, Z., Chen, W., Yang, Z., Zheng, L., Wang, S., et al. (2019). Pseudogene CHIAP2 inhibits proliferation and invasion of lung

adenocarcinoma cells by means of the WNT pathway. *J. Cell. Physiol.* 234, 13735–13746. doi: 10.1002/jcp.28053

Shirafkan, N., Mansoori, B., Mohammadi, A., Shomali, N., Ghasbi, M., and Baradaran, B. (2018). MicroRNAs as novel biomarkers for colorectal cancer: new outlooks. *Biomed. Pharmacother.* 97, 1319–1330. doi: 10.1016/j.biopha.2017.11.046

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102

Wang, G., Fu, Y., Hu, F., Lan, J., Xu, F., Yang, X., et al. (2017). Loss of BRG1 induces CRC cell senescence by regulating p53/p21 pathway. *Cell Death Dis.* 8:e2607. doi: 10.1038/cddis.2017.1

Wang, Z., Tong, D., Han, C., Zhao, Z., Wang, X., Jiang, T., et al. (2019). Blockade of miR-3614 maturation by IGF2BP3 increases TRIM25 expression and promotes breast cancer cell proliferation. *EBioMedicine* 41, 357–369. doi: 10.1016/j.ebiom.2018.12.061

Wohlers, I., Bertram, L., and Lill, C. M. (2018). Evidence for a potential role of miR-1908-5p and miR-3614-5p in autoimmune disease risk using integrative bioinformatics. *J. Autoimmun.* 94, 83–89. doi: 10.1016/j.jaut.2018.07.010

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Yu-Lee, L. Y., Yu, G., Lee, Y. C., Lin, S. C., and Pan, J. (2018). Osteoblast-secreted factors mediate dormancy of metastatic prostate cancer in the bone via activation of the TGFβRIII-p38MAPK-pS249/T252RB pathway. *Cancer Res.* 78, 2911–2924. doi: 10.1158/0008-5472.CAN-17-1051

Zen, K., and Zhang, C. Y. (2012). Circulating microRNAs: a novel class of biomarkers to diagnose and monitor human cancers. *Med. Res. Rev.* 32, 326–348. doi: 10.1002/med.20215

Zhang, Z., Rui, W., Wang, Z. C., Liu, D. X., and Du, L. (2017). Anti-proliferation and anti-metastasis effect of barbaloin in non-small cell lung cancer via inactivating p38MAPK/Cdc25B/Hsp27 pathway. *Oncol. Rep.* 38, 1172–1180. doi: 10.3892/or.2017.5760

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 05 July 2021
doi: 10.3389/fgene.2021.681277

Identification and Validation of Immune-Related LncRNA Prognostic Signature for Lung Adenocarcinoma

*Guomin Wu[1], Qihao Wang[1], Ting Zhu[2], Linhai Fu[2], Zhupeng Li[2], Yuanlin Wu[2] and Chu Zhang[3]**

[1] School of Medicine, Shaoxing University, Shaoxing, China, [2] Department of Thoracic Surgery, Shaoxing People's Hospital, Shaoxing Hospital, Zhejiang University School of Medicine, Shaoxing, China, [3] Department of Thoracic Surgery, The First Affiliated Hospital of Shaoxing University (Shaoxing People's Hospital), Shaoxing, China

This study aimed to establish a prognostic risk model for lung adenocarcinoma (LUAD). We firstly divided 535 LUAD samples in TCGA-LUAD into high-, medium-, and low-immune infiltration groups by consensus clustering analysis according to immunological competence assessment by single-sample gene set enrichment analysis (ssGSEA). Profile of long non-coding RNAs (lncRNAs) in normal samples and LUAD samples in TCGA was used for a differential expression analysis in the high- and low-immune infiltration groups. A total of 1,570 immune-related differential lncRNAs in LUAD were obtained by intersecting the above results. Afterward, univariate COX regression analysis and multivariate stepwise COX regression analysis were conducted to screen prognosis-related lncRNAs, and an eight-immune-related-lncRNA prognostic signature was finally acquired (AL365181.2, AC012213.4, DRAIC, MRGPRG-AS1, AP002478.1, AC092168.2, FAM30A, and LINC02412). Kaplan–Meier analysis and ROC analysis indicated that the eight-lncRNA-based model was accurate to predict the prognosis of LUAD patients. Simultaneously, univariate COX regression analysis and multivariate COX regression analysis were undertaken on clinical features and risk scores. It was illustrated that the risk score was a prognostic factor independent from clinical features. Moreover, immune data of LUAD in the TIMER database were analyzed. The eight-immune-related-lncRNA prognostic signature was related to the infiltration of B cells, CD4+ T cells, and dendritic cells. GSEA enrichment analysis revealed significant differences in high- and low-risk groups in pathways like pentose phosphate pathway, ubiquitin mediated proteolysis, and P53 signaling pathway. This study helps to treat LUAD patients and explore molecules related to LUAD immune infiltration to deeply understand the specific mechanism.

Keywords: lung adenocarcinoma, prognosis, lncRNA, immune infiltration, risk score

OPEN ACCESS

Edited by:
Wei Jiang,
Nanjing University of Aeronautics and Astronautics, China

Reviewed by:
Xiaowen Chen,
Harbin Medical University, China
Hongyi Zhang,
University of Texas Southwestern Medical Center, United States

***Correspondence:**
Chu Zhang
zhangchu@usx.edu.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *16 March 2021*
Accepted: *20 May 2021*
Published: *05 July 2021*

Citation:
Wu G, Wang Q, Zhu T, Fu L, Li Z, Wu Y and Zhang C (2021) Identification and Validation of Immune-Related LncRNA Prognostic Signature for Lung Adenocarcinoma. Front. Genet. 12:681277. doi: 10.3389/fgene.2021.681277

INTRODUCTION

Lung cancer has the highest morbidity and mortality among all cancers in the world (The Lancet, 2018). More than 50% of lung cancer patients have tumor metastasis when diagnosed, with a 5-year survival rate of only 5% (Siegel et al., 2020). According to histology, lung cancer is divided into small cell lung cancer and non-small cell lung cancer (NSCLC), and the latter accounts for about 85% of total lung cancer cases as a predominant subtype (Siegel et al., 2020). Lung adenocarcinoma (LUAD) accounts for 40% of all lung cancers, with a 5-year survival rate of only 15%, since patients

are diagnosed when the cancer has been locally advanced and developed metastasis (Imielinski et al., 2012). To assist the diagnosis of early-stage LUAD and treat different patients with reasonable treatment regimens without wasting medical resources and delaying patient's conditions, a suitable method is needed for predicting patient's survival status.

Malignant phenotype of cancers is both influenced by cell characteristics itself and tumor microenvironment (TME) (Catalano et al., 2013). TME encompasses blood vessels around tumor cells, fibroblasts, stromal cells, immune cells, and different signaling molecules. Accumulating studies have suggested the correlation between the infiltration level of immune cells in TME and clinical outcomes. A study on colon carcinoma elaborated the impact of T cell infiltration on cancer recurrence rate (Mlecnik et al., 2011). Above all, tumor immune infiltration is crucial in cancer progression.

Long non-coding RNAs (lncRNAs) are a class of non-protein-coding RNAs more than 200 nt in length, and they directly or indirectly participate in regulating various life activities (Geisler and Coller, 2013). Moreover, they have been proved to be pertinent to immune infiltration of tumor tissue and to regulate cancer progression through TME. For example, downregulated lncRNA SATB2-AS1 hampers tumor metastasis *via* regulating SATB2 in colorectal cancer, and lncRNA SATB2-AS1 expression negatively correlates immune infiltration (Xu et al., 2019). LncRNA NEAT1 interacts with DNMT1 to affect lung cancer progression and immune cell infiltration *via* cGAS/STING and p53 pathways (Ma et al., 2020). Knocking down lncRNA NEAT1 constrains the viability, proliferation, and invasion of lung cancer cells, while facilitating the infiltration of cytotoxic T cells in tumor tissue (Ma et al., 2020). Therefore, lncRNAs related to tumor immune infiltration can work as potential prognostic biomarkers.

Lung adenocarcinoma patients in the TCGA-LUAD dataset were grouped by immune viability in TME. Then, immune-related differentially expressed lncRNAs (DElncRNAs) were obtained by differential expression analysis. The screened DElncRNAs were applied to establish a risk score model. The model could predict patient's prognosis using the expression level of lncRNA marker to assist with choosing suitable therapeutic plans. The obtained lncRNAs can serve as researching objects in the subsequent exploration of immune infiltration regulation on LUAD to deeply understand immune escape mechanism and to dig novel therapeutic targets.

MATERIALS AND METHODS

Patient's Samples Grouped by Immunological Competence

Long non-coding RNA count data, mRNA FPKM data (normal: 59, tumor: 535), and corresponding clinical information were accessed from The Cancer Genome Atlas (TCGA)-LUAD in the TCGA database[1]. Single-sample gene set enrichment analysis (ssGSEA) was conducted on mRNA expression profile in cancer tissue of LUAD patients with R package "GSVA." Cancer tissue enrichment was scored by a gene set based on prognosis-related immune microenvironment markers obtained by Bindea et al. (2013). Next, to classify cancer tissue by immunological competences, the *k*-means algorithm in R package "ConsensusClusterPlus" and the two-phase sampling method were applied, with sampling 80% each time and 1000 times of repetition. A consensus clustering analysis was performed on cancer tissue samples according to scores of ssGSEA to obtain the most stable group.

Validation of Grouping Efficiency

Profiles of mRNA expression of cancer tissue samples were analyzed using R package "ESTIMATE" to score each sample. The immune score, stromal score, ESTIMATE score, and tumor purity of each sample were acquired. Human lymphocyte antigen (HLA) family genes and CD274 gene expression of cancer tissue in each group were analyzed by using R package "ggpubr." The efficiency of cancer tissue grouping was detected with the above results.

Screening of Immune-Related lncRNA in LUAD

A differential expression analysis ($|\log FC|>1$, $FDR<0.01$) was carried out on lncRNA expression profiles in normal tissue and cancer tissue in TCGA-LUAD using R package "edgeR." DElncRNAs related to LUAD occurrence were obtained. Differential expression analysis ($|\log FC|>1$, $FDR<0.01$) was undertaken on lncRNAs of cancer tissue in line with groups classified by immunological competence, and immune-related differentially expressed lncRNAs were obtained. DElncRNAs in the above two groups were intersected to acquire final immune-related DElncRNAs in LUAD.

Construction of Immune-Related lncRNA Prognostic Signature

Samples in TCGA-LUAD with a follow-up time less than 30 days were removed. Univariate COX regression analysis ($p < 0.001$) was conducted on immune-related DElncRNAs in LUAD using R package "survival" to screen lncRNAs that were correlated with LUAD prognosis. Then, multivariate COX regression analysis was undertaken through stepwise regression using dual-choice assay. The model with the smallest Akaike Information Criterion (AIC) was chosen as the immune-related lncRNA prognostic model. Patient's risk score was calculated by the prognostic model, with the expression level of lncRNA as the risk factor. The higher the risk score, the poorer the predicted prognosis of LUAD patients. The formula is $Risk\ score = \sum_{i=1}^{n}(Coefi^{*}\chi i)$ ($Coef_i$: risk coefficient, χ_i: gene expression level standardized by Z-score). Clinical information of samples used in the analysis is shown in **Supplementary Table 1**.

Validation of the Accuracy of the Prognostic Model

To validate the predictive accuracy of the immune-related lncRNA prognostic signature, risk score of patients in TCGA-LUAD was obtained from the prognostic model. Patients were

[1]https://portal.gdc.cancer.gov/

divided into high- and low-risk groups based on the median score. Kaplan–Meier survival analysis was undertaken on two groups using R package "survival" to compare the survival rate. Afterward, R package "survivalROC" was employed to draw the receiver operating characteristic (ROC) curves of 3-year overall survival (OS) and 5-year OS. Thereafter, the area under curve (AUC) was calculated. The risk score distribution and survival status of patients were plotted to determine the accuracy of the model. At the same time, a heatmap of the expression level of lncRNAs in the prognostic model in cancer tissue of patients in high- and low-risk groups was drawn.

Validation of the Independence of the Prognostic Model

To detect whether the risk score of patients evaluated by the immune-related lncRNA prognostic signature is independent of clinical features as a prognostic factor of LUAD, univariate and multivariate COX regression analyses were performed combining with risk scores and clinical traits (age, gender, T stage, N stage, and clinical stage). Next, R package "survivalROC" was applied to draw time-dependent ROC curves of the above prognostic factors, and AUCs were calculated, respectively.

Correlation Analysis Between the Risk Score of the Prognostic Model and Infiltration of Each Immune Cell Subgroup in LUAD Samples

Immune infiltration data of B cells, CD4+ T cells, CD8+ T cells, dendritic cells, macrophages, and neutrophils were downloaded from the TIMER database[2]. The correlation between the risk score and infiltration of each immune cell subgroup was detected, respectively, by using Pearson correlation analysis.

GSEA Enrichment Analysis in the High- and Low-Risk Groups

Pathway differences between the high- and low-risk groups were investigated by GSEA enrichment analysis (Mootha et al., 2003; Subramanian et al., 2005). mRNA expression microarray in the high- and low-risk groups in TCGA-LUAD was taken as the expression dataset. c2.cp.kegg.v7.4.symbols.gmt was selected as enrichment analysis gene set to run GSEA software. Displacement test times were set as 1000 and gene set with FDR < 0.25 was considered significant enrichment.

RESULTS

LUAD Patients Are Grouped According to Immunological Competence

To explore the immune infiltration and prognosis of LUAD, this study established an immune-related LUAD prognostic model by bioinformatics method (**Figure 1**). mRNA expression profile of cancer tissue samples in TCGA-LUAD was analyzed *via* ssGSEA, and the immunological competence of each cancer tissue sample was scored. Afterward, all cancer tissue samples were divided into three groups by consensus clustering analysis according to the score: high-immune infiltration group (Immunity_H, 217 samples), medium-immune infiltration group (Immunity_M, 200 samples), and low-immune infiltration group (Immunity_L, 118 samples) (**Figures 2A–C**). Chi-square test unveiled remarkable differences in T stage distribution of patients in different immune infiltration groups ($p < 0.05$). To verify the effect of grouping, ESTIMATE algorithm was applied to score each sample in three groups, and it was found that the immune cell infiltration in cancer tissue was positively correlated with the immune score, stromal score, and ESTIMATE score, while it was negatively correlated with tumor purity (**Figure 3A**). It was also displayed that differences in infiltration of immune cells in varying immune groups were highly statistically notable ($p < 0.01$) (**Figure 3B**). Besides, the expression of HLA family genes and CD274 gene of cancer tissue sample in three groups was analyzed. The expression of immune-related genes elevated as immunological competence increased. It was also illustrated that differences in the expression of these genes in three groups were highly statistically significant ($p < 0.01$) (**Figures 3C,D**). All in all, three groups of LUAD subtypes divided by immunological competence in this study had different degrees of immune infiltration.

Screening of Immune-Related Differential lncRNAs in LUAD

Differential expression analysis was undertaken on LUAD tissue and normal tissue. A total of 3,202 differentially upregulated lncRNAs and 690 downregulated lncRNAs in LUAD tumor tissue were obtained ($|\log FC| > 1$, FDR < 0.01) (**Figure 4A**). Thereafter, differential expression analysis was performed on lncRNAs in the Immunity_H group and Immunity_L group. A total of 397 upregulated lncRNAs and 2,047 downregulated lncRNAs in the Immunity_H group were obtained ($|\log FC| > 1$, FDR < 0.01) (**Figure 4B**). The duplicate lncRNAs in two groups were chosen to acquire 1,570 immune-related differential lncRNAs in LUAD (**Figure 4C**), and these lncRNAs could be used to screen LUAD prognostic signature in the following experiments.

Construction of the Immune-Related lncRNA Prognostic Signature and Validation of Its Accuracy

Patient's clinical data in TCGA-LUAD were obtained, and samples with less than 30 days of follow-up were removed. Then, univariate COX regression analysis was conducted on 1,570 immune-related differential lncRNAs in LUAD, and 15 lncRNAs highly significantly correlating patient's prognosis were obtained ($p < 0.001$) (**Supplementary Table 2**). Next, multivariate COX regression analysis was applied on the 15 lncRNAs to acquire lncRNAs (AL365181.2, AC012213.4, DRAIC, MRGPRG-AS1, AP002478.1, AC092168.2, FAM30A, and LINC02412), and the expression of the eight lncRNAs was taken as prognostic factors of the risk score model (**Figure 5A** and **Supplementary Table 3**).

[2]https://cistrome.shinyapps.io/timer/

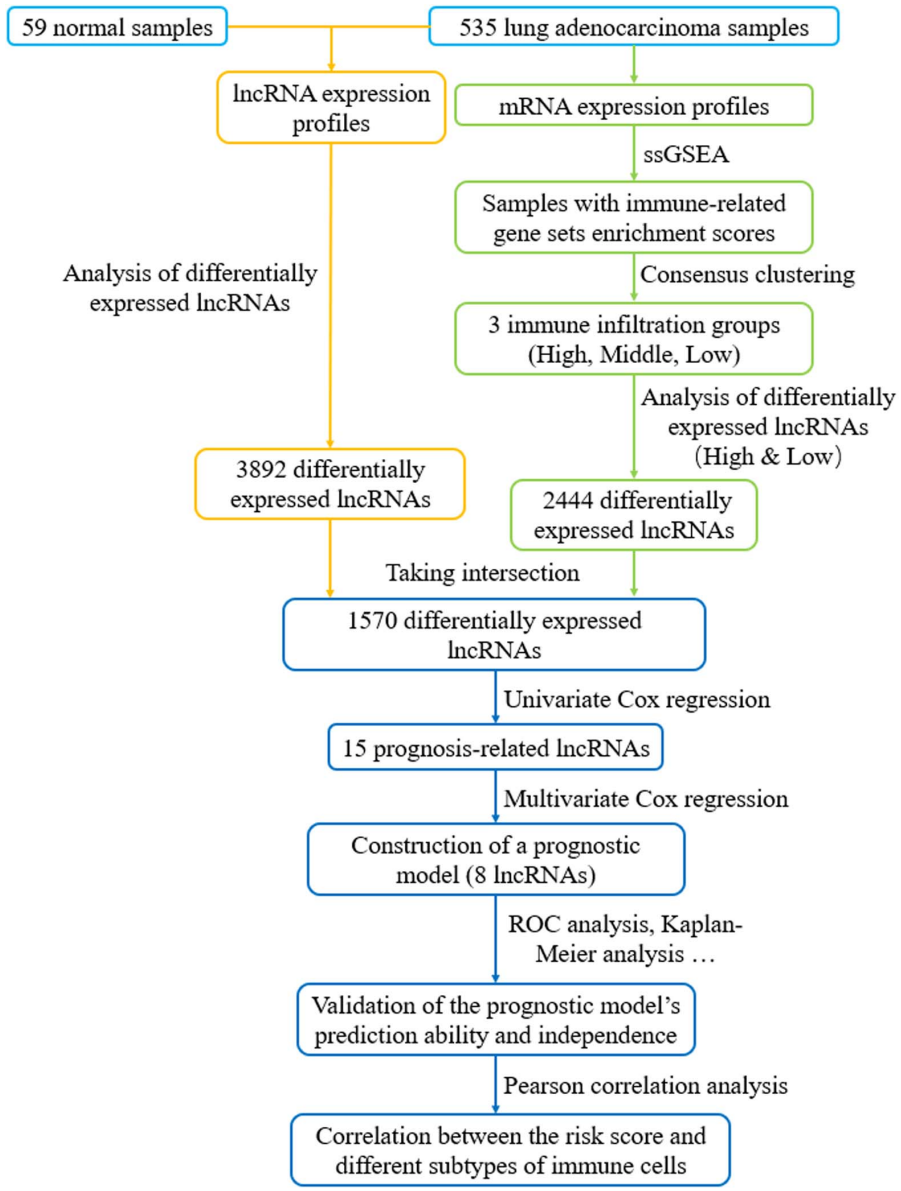


FIGURE 1 | Flow chart of constructing a prognostic signature based on eight immune-related lncRNAs. LncRNAs related to immune infiltration and cancer occurrence were screened to build an eight-lncRNA prognostic signature. The predictive accuracy and independence of the signature in prognosis of LUAD were testified, and correlation between the signature-based risk score and abundance of several immune infiltrates was detected.

The model predicted patient's prognosis by calculating risk scores. The higher the risk score, the poorer the prognosis. LUAD patients in TCGA-LUAD were divided into high-risk group and low-risk group according to the median risk score of all tumor samples (**Figure 5D**). Kaplan–Meier survival analysis was carried out on patients in two groups, and it was found that the survival rate of patients in the high-risk group was significantly lower than that in the low-risk group ($p < 0.001$) (**Figure 5B**). ROC curves of 3-year and 5-year survival predicted by the prognostic model were drawn, and AUCs were calculated (3-year AUC: 0.734, 5-year AUC: 0.696). These indicated that the prognostic model could predict 3-year survival rate and 5-year survival rate of LUAD patients to a certain extent, and the accuracy and sensitivity in predicting 3-year survival rate were relatively high (**Figure 5C**). Meanwhile, it was found through statistical analysis that the mortality of patients in the high-risk group was higher than that in the low-risk group (**Figure 5E**). Moreover, compared with the expression in the low-risk group, the expression of lncRNAs (AL365181.2, AC012213.4, MRGPRG-AS1, AP002478.1, AC092168.2, and LINC02412) that were not conductive to positive prognosis in the high-risk group was higher, while the expression of the other two lncRNAs (DRAIC and FAM30A) that were conducive to positive prognosis in the high-risk group was lower (**Figure 5F**). The above

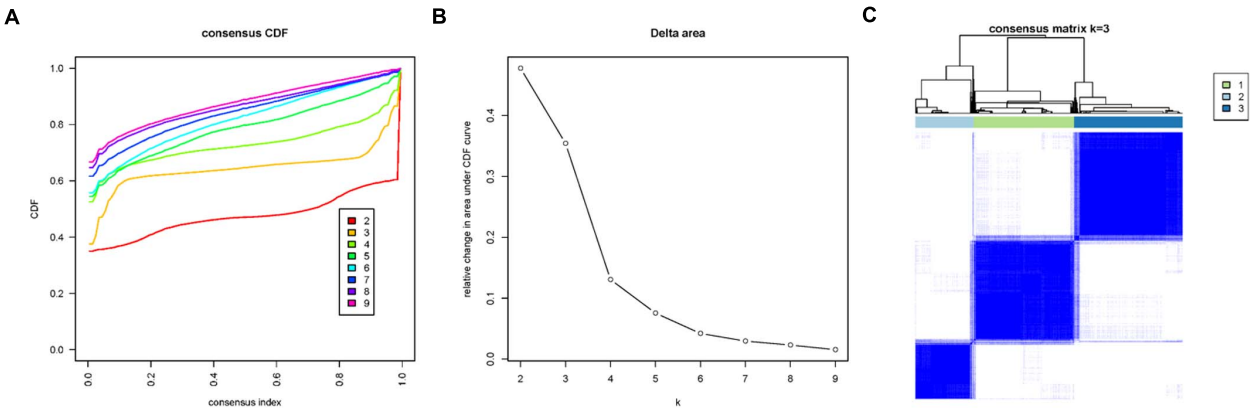


FIGURE 2 | Consensus clustering analysis of LUAD tissue samples in TCGA-LUAD according to immune infiltration degree. **(A)** Consensus clustering cumulative distribution function (CDF) diagram with a tissue clustering number from $k = 2$ to $k = 9$. **(B)** Delta area plot of consensus clustering analysis of LUAD tissue, and the relative changes in CDF AUC in different k-values compared with that in k-1. k-value was finally chosen as 3, combining a series of analysis results. **(C)** Heatmap of sample consensus after grouping.

results indicated that the constructed eight-immune-related-lncRNA prognostic signature could accurately predict LUAD patient's prognosis.

Examination of the Independence of the Eight-Immune-Related-lncRNA Prognostic Signature

Since many clinical traits are related to prognosis of cancer, we focused on detecting the independence of the eight-immune-related-lncRNA prognostic signature. A univariate COX regression analysis was conducted on clinical traits (age, gender, T stage, N stage, and clinical stage) and the risk score calculated by the prognostic model. It was exhibited that T stage, N stage, clinical stage, and the risk score were significantly correlated with LUAD patient's prognosis ($p < 0.001$) (**Figure 6A**). Afterward, a multivariate COX regression analysis was performed on clinical traits and the risk score. It was presented that only clinical stage and the risk score were remarkably related to LUAD patient's prognosis ($p < 0.05$), and only the risk score displayed a highly significant correlation ($p < 0.001$) (**Figure 6B**). In addition, the AUC value for 3-year survival of the risk score was the highest (0.731) among the AUC values of other factors (**Figure 6C**). In conclusion, the eight-immune-related-lncRNA prognostic signature was an independent prognostic factor of LUAD prognosis.

Analysis of the Correlation Between the Eight-Immune-Related-lncRNA Prognostic Signature and the Infiltration of Each Immune Cell Subgroup in LUAD Samples

Since all factors in the prognostic model were related to immune infiltration, to explore the correlation between the eight-immune-related-lncRNA prognostic signature and immune infiltration, Pearson correlation analysis was performed on the infiltration of varying immune cell subgroups in LUAD samples and the eight-immune-related-lncRNA prognostic signature using data in the TIMER database. It was revealed that the risk score was negatively correlated with the infiltration of B cells, CD4+ T cells, and dendritic cells ($p < 0.01$) (**Figures 7A–F**). Thus, it was suggested that the risk score predicted for patients by the eight-immune-related-lncRNA prognostic signature could reflect the infiltration of B cells, CD4+ T cells, and dendritic cells.

GSEA Enrichment Analysis in the High- and Low-Risk Groups

The above analyses discovered significant correlation between the prognostic characteristics of eight immune-related lncRNAs and immune cell infiltration. In a bid to investigate the specific signaling pathways affected by these eight lncRNAs, GSEA enrichment analysis was undertaken in the high- and low-risk groups in this study. The results showed significant differences in signaling pathways (pentose phosphate pathway, ubiquitin mediated proteolysis, and P53 signaling pathway) between the high- and low-risk groups (**Figures 8A–C**). These pathways were found to be vital in tumor progression and directly or indirectly modulate immune characteristics of cancers (Liu et al., 2018; Giacomini et al., 2020; Li and Song, 2020). Above all, eight immune-related lncRNA prognostic characteristics were closely associated with some important cancer-related signaling pathways.

DISCUSSION

In this study, consensus clustering analysis was performed on patients in TCGA-LUAD for analyzing immunological competence. The immune infiltration of the patients was determined by using ESTIMATE algorithm plus the expression level of HLA family genes and the CD274 gene. The CD274 gene, which plays an important role in the immune escape of cancer cells, encodes programmed death ligand-1 (PD-L1)

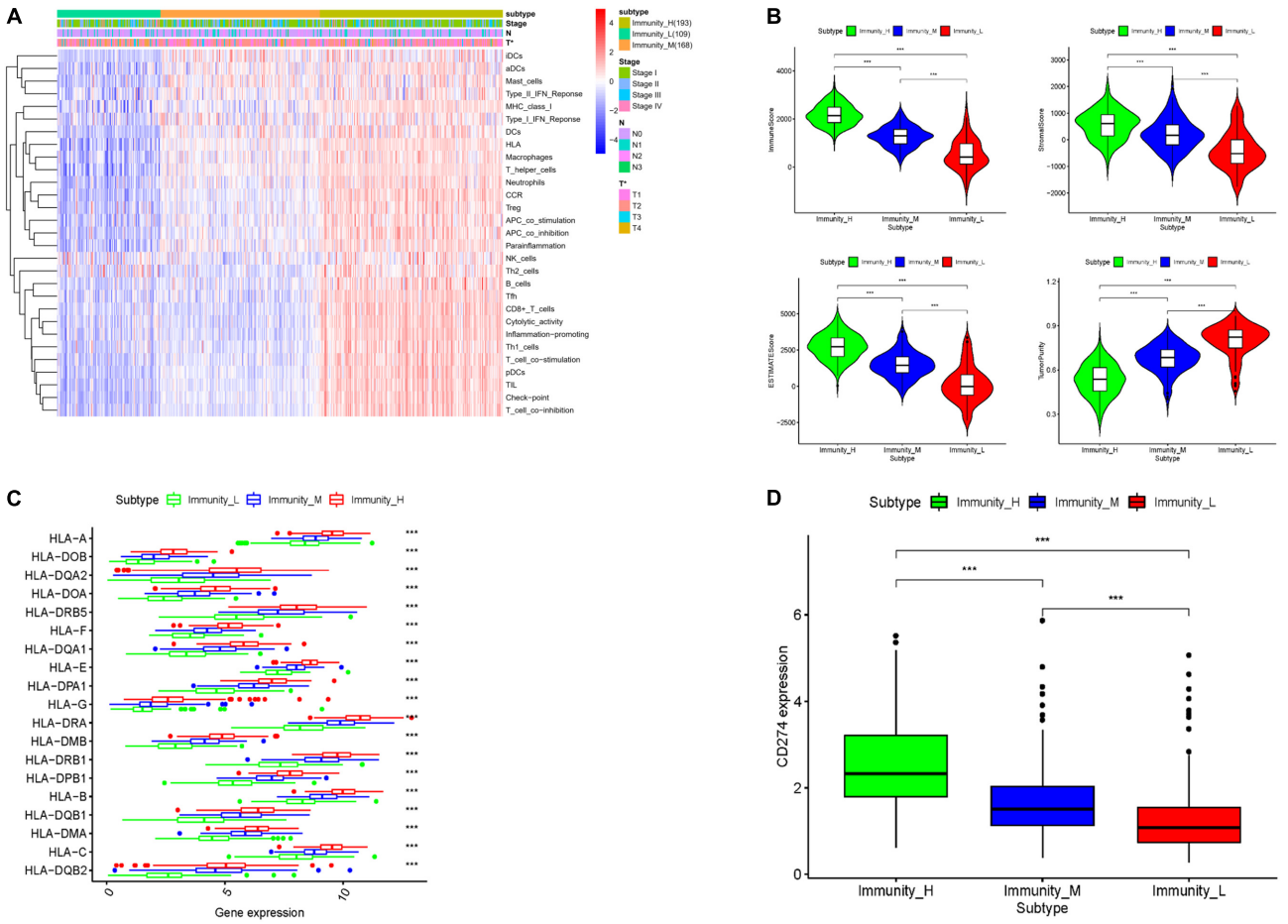


FIGURE 3 | Validation of grouping effect of LUAD tissue samples. **(A)** Samples were divided into Immunity_H, Immunity_M, and Immunity_L groups according to immunological competence from high to low. Heatmap shows ssGSEA enrichment value of each immune-related gene set of LUAD tissue in each group, and tumor purity, immune score, stromal score, and ESTIMATE score obtained by "ESTIMATE" analysis. **(B)** Violin plots of tumor purity, immune score, stromal score, and ESTIMATE score of three groups of LUAD tissue. **(C)** Boxplots of the expression of each HLA family gene in LUAD tissue in three groups. **(D)** Boxplots of the expression of CD274 gene in LUAD tissue in three groups. *** represents $p < 0.001$ (difference between groups is highly statistically significant).

expressing in many cancer cells and immune cells. Expression of PD-L1 in cancer cells binds to programmed death receptor-1 (PD-1) in cytotoxic T cells to inhibit proliferation, migration, and secretion of cytotoxic substances, so as to impair the activity of cytotoxic T cells and thus prevent the tumor cells from being killed (Herbst et al., 2014; Kim and Chen, 2016). HLA gene clusters encode the major histocompatibility complex (MHC) of human, which can be divided into HLA class I and HLA class II. HLA is involved in antigen processing and presentation. Its abnormal expression is thought to be related to malignant transformation and immune escape of cells. For example, loss of HLA class I molecule expression will affect the interaction between cancer cells and T cells, leading to the ability of cancer cells to evade immune surveillance, and patients with loss of HLA class I molecule expression tend to have a shorter survival time (Garrido et al., 2017). Among the HLA genes, HLA-G gene expression was found to be upregulated when cells develop malignant phenotype, and it is identified to be associated with cancer inflammation, expressed in both cancer cells and immune cells, especially in CD68+ macrophages and CD8+ T cells infiltrating cancer tissue (Rouas-Freiss et al., 2003). There is an important relationship between immune infiltration and prognosis, suggesting that immune-related molecules can be used as prognostic markers in cancer patients.

Then, after screening and obtaining immune-related DElncRNAs in this study, univariate COX regression analysis combined with multivariate COX regression analysis was used to construct an eight-immune-related-lncRNA prognostic signature. Among the eight risk factors of the risk score model, AL365181.2, AC012213.4, MRGPRG-AS1, AC092168.2, and LINC02412 have been rarely studied, and a few studies were conducted on the remaining three lncRNAs. By constructing ceRNA networks in hepatocellular carcinoma, AP002478.1 was screened for many times as a potential prognostic marker (Liao X. et al., 2019; Yue et al., 2019). In hepatitis virus-positive hepatocellular carcinoma (Huang et al., 2019) and *Helicobacter*

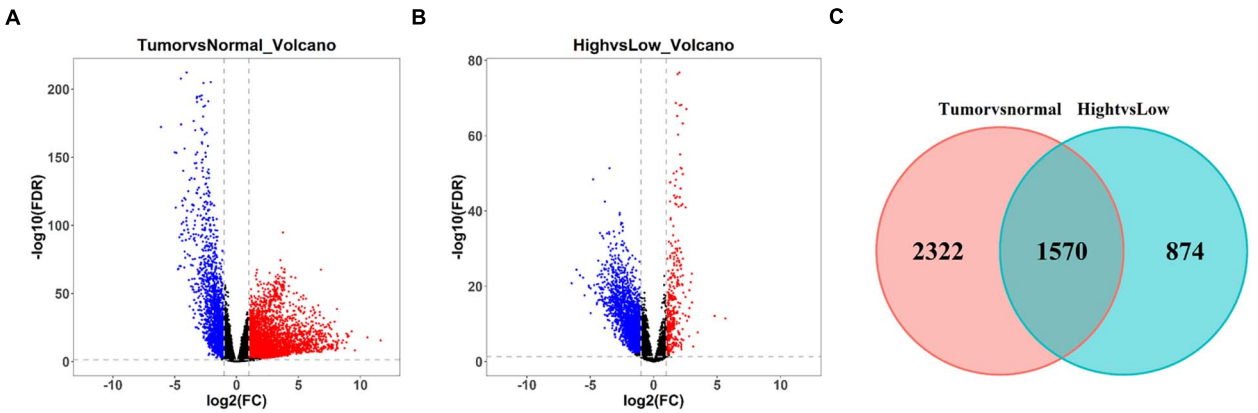


FIGURE 4 | Screening of differentially expressed lncRNAs related to immune infiltration in LUAD. **(A)** Volcano plot of differential expression analysis of lncRNAs in LUAD tissue and normal tissue. Blue dots represent downregulated lncRNAs and red dots represent upregulated lncRNAs (| logFC| > 1, FDR < 0.01). **(B)** Volcano plot of differential expression analysis of lncRNAs in LUAD tissue in Immunity_H and Immunity_L groups. Blue dots represent downregulated lncRNAs and red dots represent upregulated lncRNAs (| logFC| > 1, FDR < 0.01). **(C)** Venn diagram of differential lncRNAs related to LUAD occurrence in panel **(A)** (red circle) and differential lncRNAs related to immune infiltration in panel **(B)** (green circle).

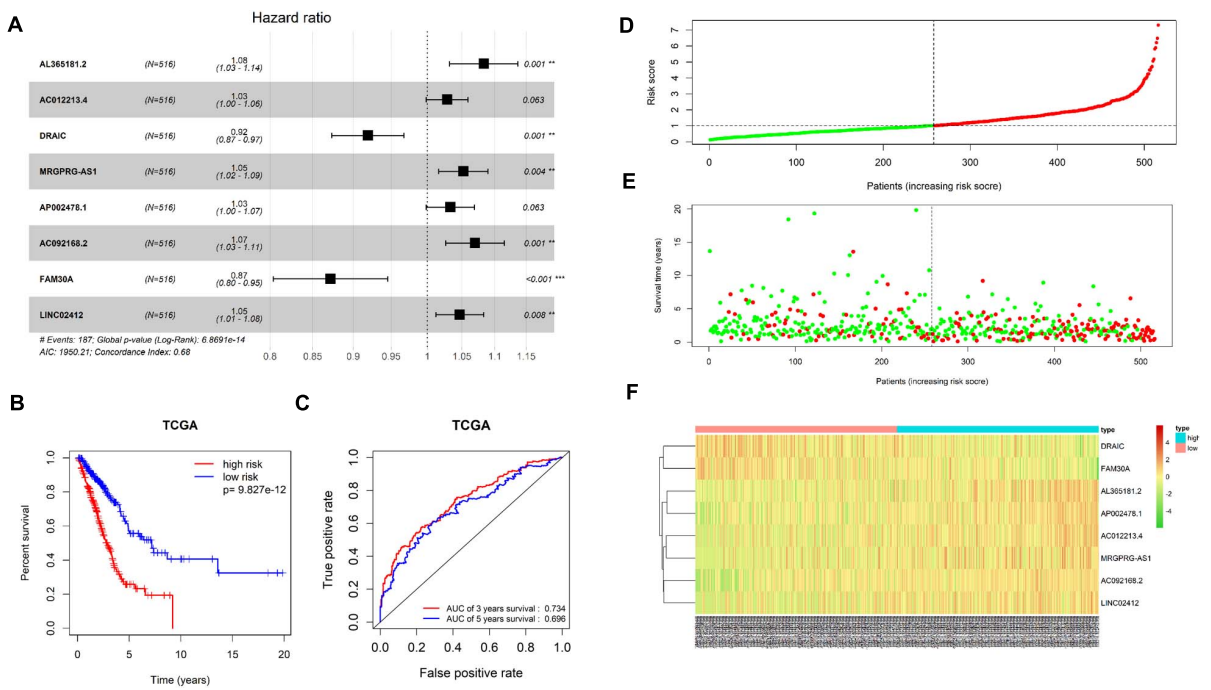


FIGURE 5 | Construction and validation of the eight-immune-related-lncRNA prognostic signature. **(A)** Screening and construction of an eight-immune-related-lncRNA prognostic signature by multivariate COX regression analysis. Forest plot shows hazard ratio (HR) and *p*-value of each lncRNA in the model. ** represents $p < 0.01$, *** represents $p < 0.001$. **(B)** Survival curves obtained by Kaplan–Meier survival analysis in high-risk and low-risk groups divided by the median risk score. **(C)** ROC curves and corresponding AUCs of 3-year survival and 5-year survival predicted by the prognostic model. **(D)** Distribution diagram of risk score of LUAD patients in TCGA-LUAD. The horizontal dotted line indicates the median risk score of all patients. **(E)** Distribution diagram of survival status of LUAD patients. Patient's risk score is presented from low to high by left to right. The vertical dotted line represents the patients with a median risk score. Green dots represent survival and red dots represent death. **(F)** Heatmap of the expression of eight immune-related lncRNAs in cancer tissue of LUAD patients in high- and low-risk groups.

pylori-positive gastric cancer (Liu et al., 2019), AP002478.1 was also screened and found to be a prognostic marker to participate in the construction of prognostic models. In addition, AP002478.1 in the models constructed in the above studies is a negative prognostic factor, which is consistent with the results of this study. However, the role of AP002478.1 in LUAD has not been discussed in relevant studies, and the specific mechanism of its influence on cancer prognosis is still unknown.

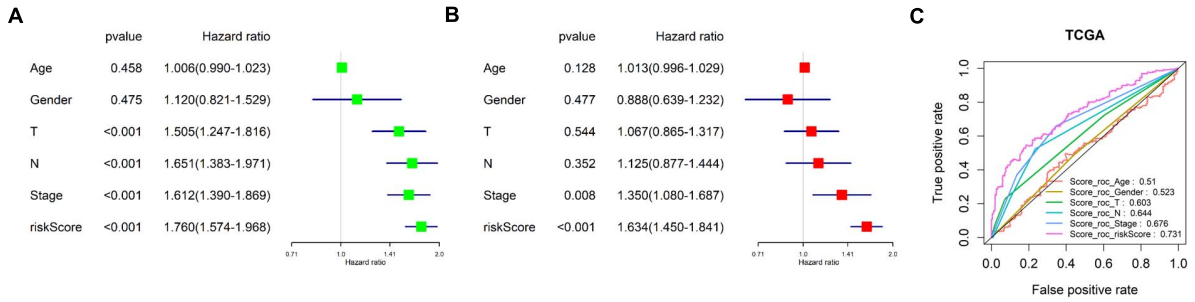


FIGURE 6 | Examination of the independence of the risk score predicted by the prognostic model. **(A)** Univariate COX regression analysis was conducted on clinical traits (age, gender, T stage, N stage, and clinical stage) and the risk score. The forest plot shows HR and *p*-value of each prognostic factor. **(B)** Multivariate COX regression analysis was performed on clinical traits and the risk score. **(C)** ROC curves and AUCs of 3-year survival of each prognostic factor.

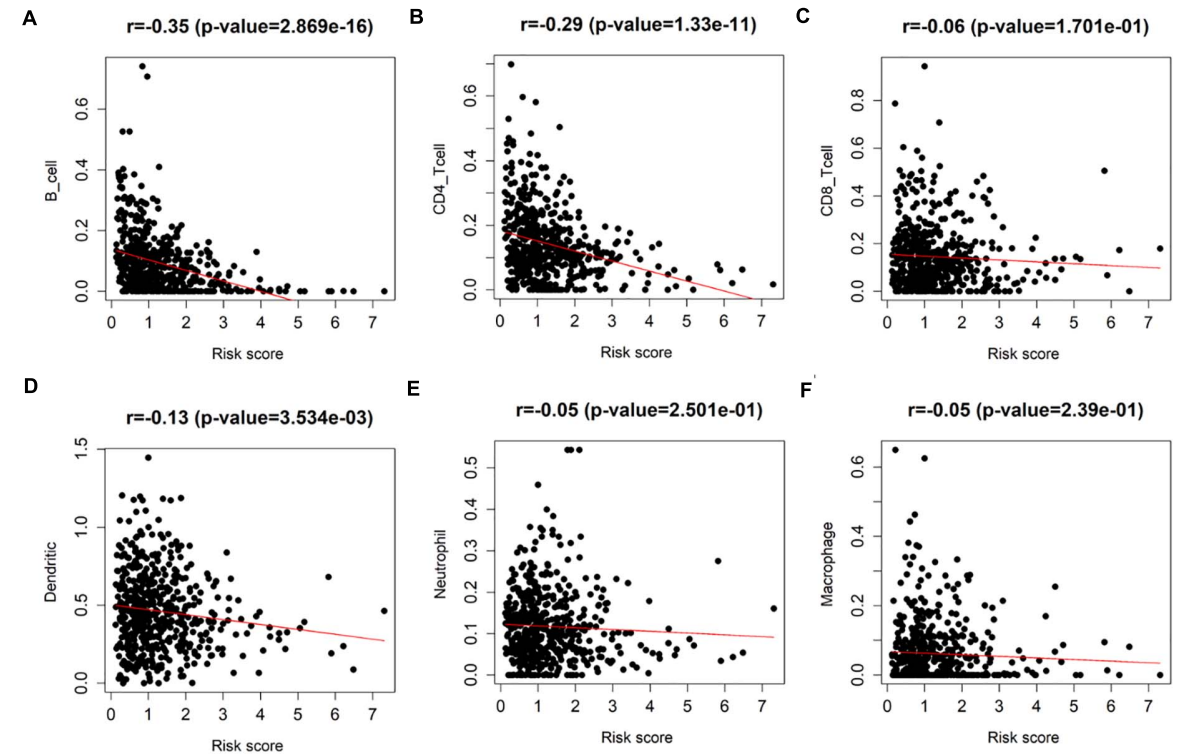


FIGURE 7 | Detection of the correlation between the eight-immune-related-lncRNA prognostic signature and each immune cell subgroup. Pearson correlation analysis was undertaken to detect the correlation between the risk score of LUAD patients predicted by the model and the infiltration of **(A)** immune B cells, **(B)** CD4+ T cells, **(C)** CD8+ T cells, **(D)** dendritic cells, **(E)** neutrophils, and **(F)** macrophages in patient's cancer tissue.

Long non-coding RNA FAM30A was considered as a positive prognostic factor for LUAD in this study. In past studies, FAM30A was found to be correlated with immunity. For example, the expression level of FAM30A in periodontitis was positively related to the proportion of plasma cells in inflammatory tissue (Wu et al., 2020). In a study on the body's response to vaccination and the protective mechanism induced by vaccines, the expression level of FAM30A was positively correlated with the antibody titer level after vaccination. Compared with other immune cells, the expression level of FAM30A was higher in B cells (de Lima et al., 2019). Therefore, FAM30A is closely immune-related. In the end of this study, the correlation between prognostic risk score and infiltration of multiple immune cell subgroups was analyzed. The eight-immune-related-lncRNA prognostic signature was significantly negatively correlated with B cells, CD4+ T cells, and dendritic cells, expanding subsequent research directions for FAM30A, AP002478.1, and other lncRNAs.

Studies on lncRNA DRAIC are more abundant, and there have been relevant functional studies. For example, it was found in gastric cancer cells that DRAIC can promote the degradation

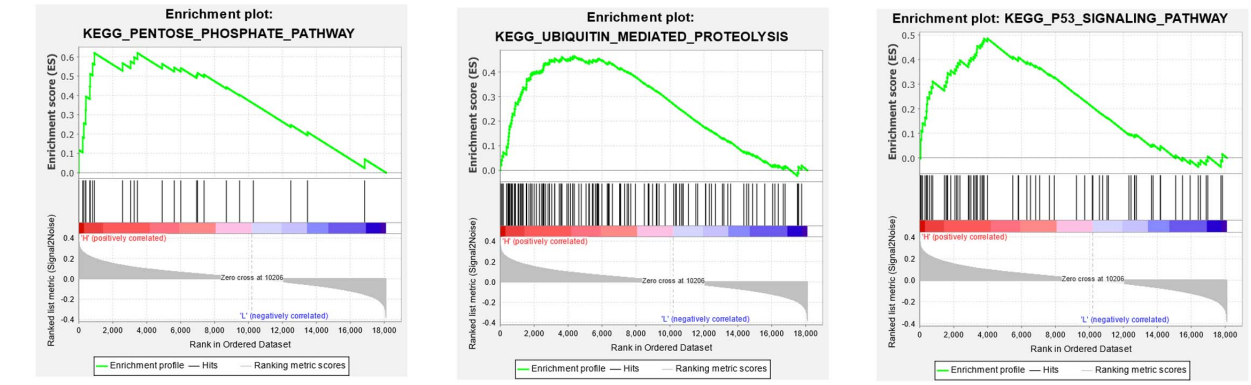


FIGURE 8 | GSEA enrichment analysis in the high- and low-risk groups. **(A–C)** Patients in the high- and low-risk groups showed marked differences in pentose phosphate pathway, ubiquitin mediated proteolysis, and P53 signaling pathway signaling pathways, respectively.

of NFRKB and inhibit the proliferation and metastasis of gastric cancer cells by blocking NFRKB de-ubiquitination mediated by UCHL5 (Zhang et al., 2020). In prostate cancer, DRAIC binds to IKK subunits to block interunit binding and inhibit NF-κB activation, thereby inhibiting invasion and proliferation of cancer cells (Saha et al., 2020). The above results indicate that DRAIC is a prognostic factor for cancer patients, which is the same as the analysis results in this study. In nasopharyngeal carcinoma studies, it was found that DRAIC can sponge miR-122 to promote the expression of SATB1, thus promoting the proliferation, migration, and invasion of nasopharyngeal carcinoma cells (Liao B. et al., 2019). DRAIC has diverse functions in different cancers. Another study also showed that DRAIC has the function of regulating autophagy (Tiessen et al., 2019). The above studies and this study can guide follow-up research directions of the screened eight prognostic lncRNAs, so as to conduct in-depth research on corresponding specific regulatory mechanism of the immune microenvironment of LUAD.

Viewed *in toto*, in this study, the mRNA expression profile of TCGA-LUAD was analyzed by ssGSEA and the immunological competence of the samples was scored to divide samples into three groups by consensus clustering analysis. Then, the differential expression analysis on the lncRNA profile of the samples in the high- and low-immune infiltration groups was performed. DElncRNAs obtained were intersected with the differential lncRNAs screened from the differential expression analysis in the normal samples and LUAD samples to obtain immune-related DElncRNAs in LUAD. Eventually, an eight-immune-related-lncRNA prognostic signature was obtained and a risk score model was established. Additionally, the risk score evaluated by this model was significantly negatively correlated with infiltration of B cells, CD4+ T cells, and dendritic cells. GSEA enrichment analysis discovered that patients in the high- and low-risk groups had marked differences in cancer-related signaling pathways. The risk score can reflect the patient's immune microenvironment to some extent, thereby helping patients choose appropriate treatment. However, the accuracy of the risk score model on the prediction of 5-year survival rate is not high; therefore, more data and experimental basis are still needed to improve the model. Besides, this study did not use independent validation sets to test the accuracy of the risk score model; hence, the universal applicability of the model has not been verified.

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/s.

AUTHOR CONTRIBUTIONS

GW contributed to the study design. CZ conducted the literature search, revised the article, and gave the final approval of the version to be submitted. QW and TZ acquired the data. LF and ZL wrote the article. YW performed the data analysis. ZL drafted the manuscript.

FUNDING

This work was funded by the Zhejiang Medical and Health Research Fund Project (2020KY321) and Zhejiang Science and Technology Agency Experimental Animal Project (2018C37104).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.681277/full#supplementary-material

REFERENCES

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39, 782–795. doi: 10.1016/j.immuni.2013.10.003

Catalano, V., Turdo, A., Franco, S. D., Dieli, F., Todaro, M., and Stassi, G. (2013). Tumor and its microenvironment: a synergistic interplay. *Semin. Cancer Biol.* 23, 522–532. doi: 10.1016/j.semcancer.2013.08.007

de Lima, D. S., Cardozo, L. E., Maracaja-Coutinho, V., Suhrbier, A., Mane, K., Jeffries, D., et al. (2019). Long noncoding RNAs are involved in multiple immunological pathways in response to vaccination. *Proc. Natl. Acad. Sci. U.S.A.* 116, 17121–17126. doi: 10.1073/pnas.1822046116

Garrido, F., Perea, F., Bernal, M., Sánchez-Palencia, A., Aptsiauri, N., and Ruiz-Cabello, F. (2017). The escape of cancer from T cell-mediated immune surveillance: HLA class I loss and tumor tissue architecture. *Vaccines (Basel)* 5:7. doi: 10.3390/vaccines5010007

Geisler, S., and Coller, J. (2013). RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat. Rev. Mol. Cell Biol.* 14, 699–712. doi: 10.1038/nrm3679

Giacomini, I., Ragazzi, E., Pasut, G., and Montopoli, M. (2020). The pentose phosphate pathway and its involvement in cisplatin resistance. *Int. J. Mol. Sci.* 21:937. doi: 10.3390/ijms21030937

Herbst, R. S., Soria, J. C., Kowanetz, M., Fine, G. D., Hamid, O., Gordon, M. S., et al. (2014). Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* 515, 563–567. doi: 10.1038/nature14011

Huang, Z. L., Li, W., Chen, Q. F., Wu, P. H., and Shen, L. J. (2019). Eight key long non-coding RNAs predict hepatitis virus positive hepatocellular carcinoma as prognostic targets. *World J. Gastrointest. Oncol.* 11, 983–997. doi: 10.4251/wjgo.v11.i11.983

Imielinski, M., Berger, A. H., Hammerman, P. S., Hernandez, B., Pugh, T. J., Hodis, E., et al. (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* 150, 1107–1120. doi: 10.1016/j.cell.2012.08.029

Kim, J. M., and Chen, D. S. (2016). Immune escape to PD-L1/PD-1 blockade: seven steps to success (or failure). *Ann. Oncol.* 27, 1492–1504. doi: 10.1093/annonc/mdw217

Li, X., and Song, Y. (2020). Proteolysis-targeting chimera (PROTAC) for targeted protein degradation and cancer therapy. *J. Hematol. Oncol.* 13:50. doi: 10.1186/s13045-020-00885-3

Liao, B., Wang, Z., Zhu, Y., Wang, M., and Liu, Y. (2019). Long noncoding RNA DRAIC acts as a microRNA-122 sponge to facilitate nasopharyngeal carcinoma cell proliferation, migration and invasion via regulating SATB1. *Artif. Cells Nanomed. Biotechnol.* 47, 3585–3597. doi: 10.1080/21691401.2019.1656638

Liao, X., Wang, X., Huang, K., Han, C., Deng, J., Yu, T., et al. (2019). Integrated analysis of competing endogenous RNA network revealing potential prognostic biomarkers of hepatocellular carcinoma. *J. Cancer* 10, 3267–3283. doi: 10.7150/jca.29986

Liu, Y., Lv, J., Liu, J., Liang, X., Jin, X., Xie, J., et al. (2018). STAT3/p53 pathway activation disrupts IFN-beta-induced dormancy in tumor-repopulating cells. *J. Clin. Invest.* 128, 1057–1073. doi: 10.1172/JCI96329

Liu, Y., Zhu, J., Ma, X., Han, S., Xiao, D., Jia, Y., et al. (2019). ceRNA network construction and comparison of gastric cancer with or without *Helicobacter* pylori infection. *J. Cell. Physiol.* 234, 7128–7140. doi: 10.1002/jcp.27467

Ma, F., Lei, Y., Ding, M., Luo, L., Xie, Y., and Liu, X. (2020). LncRNA NEAT1 interacted with DNMT1 to regulate malignant phenotype of cancer cell and cytotoxic T cell infiltration via epigenetic inhibition of p53, cGAS, and STING in lung cancer. *Front. Genet.* 11:250. doi: 10.3389/fgene.2020.00250

Mlecnik, B., Tosolini, M., Kirilovsky, A., Berger, A., Bindea, G., Meatchi, T., et al. (2011). Histopathologic-based prognostic factors of colorectal cancers are associated with the state of the local immune reaction. *J. Clin. Oncol.* 29, 610–618. doi: 10.1200/JCO.2010.30.5425

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34, 267–273. doi: 10.1038/ng1180

Rouas-Freiss, N., Moreau, P., Menier, C., and Carosella, E. D. (2003). HLA-G in cancer: a way to turn off the immune system. *Semin. Cancer Biol.* 13, 325–336. doi: 10.1016/s1044-579x(03)00023-3

Saha, S., Kiran, M., Kuscu, C., Chatrath, A., Wotton, D., Mayo, M. W., et al. (2020). Long noncoding RNA DRAIC inhibits prostate cancer progression by interacting with IKK to Inhibit NF-kappaB activation. *Cancer Res.* 80, 950–963. doi: 10.1158/0008-5472.CAN-19-3460

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590

Subramanian, A., Tamayo, P., Mootha, V., Mukherjee, S., Ebert, B., Gillette, M., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102

The Lancet (2018). GLOBOCAN 2018: counting the toll of cancer. *Lancet* 392:985. doi: 10.1016/S0140-6736(18)32252-9

Tiessen, I., Abildgaard, M. H., Lubas, M., Gylling, H. M., Steinhauer, C., Pietras, E. J., et al. (2019). A high-throughput screen identifies the long non-coding RNA DRAIC as a regulator of autophagy. *Oncogene* 38, 5127–5141. doi: 10.1038/s41388-019-0783-9

Wu, D., Zhou, P., Cao, F., Lin, Z., Liang, D., Zhou, X., et al. (2020). Expression profiling and cell type classification analysis in periodontitis reveal dysregulation of multiple lncRNAs in plasma cells. *Front. Genet.* 11:382. doi: 10.3389/fgene.2020.00382

Xu, M., Xu, X., Pan, B., Chen, X., Lin, A. K., Zeng, K., et al. (2019). LncRNA SATB2-AS1 inhibits tumor metastasis and affects the tumor immune cell microenvironment in colorectal cancer by regulating SATB2. *Mol. Cancer* 18:135. doi: 10.1186/s12943-019-1063-6

Yue, C., Ren, Y., Ge, H., Liang, C., Xu, Y., Li, G., et al. (2019). Comprehensive analysis of potential prognostic genes for the construction of a competing endogenous RNA regulatory network in hepatocellular carcinoma. *Onco Targets Ther.* 12, 561–576. doi: 10.2147/OTT.S188913

Zhang, Z., Hu, X., Kuang, J., Liao, J., and Yuan, Q. (2020). LncRNA DRAIC inhibits proliferation and metastasis of gastric cancer cells through interfering with NFRKB deubiquitination mediated by UCHL5. *Cell. Mol. Biol. Lett.* 25:29. doi: 10.1186/s11658-020-00221-0

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 07 July 2021
doi: 10.3389/fgene.2021.673501

A Novel circRNA–miRNA–mRNA Hub Regulatory Network in Lung Adenocarcinoma

Haiwei Zuo[1†], Xia Li[2,3†], Xixi Zheng[1], Qiuwen Sun[4], Qianqian Yang[5] and Yong Xin[2,3]*

[1] School of Medical Information and Engineering, Xuzhou Medical University, Xuzhou, China, [2] Department of Radiation Oncology, Affiliated Hospital of Xuzhou Medical University, Xuzhou, China, [3] Cancer Institute of Xuzhou Medical University, Xuzhou, China, [4] School of Medical Imaging, Xuzhou Medical University, Xuzhou, China, [5] Neonatal Intensive Care Unit, Affiliated Hospital of Xuzhou Medical University, Xuzhou, China

The growing evidence suggests that circular RNAs (circRNAs) have significant associations with tumor occurrence and progression, yet the regulatory mechanism of circRNAs in lung adenocarcinoma (LUAD) remains unclear. This study clarified the potentially regulatory network and functional mechanism of circRNAs in LUAD. The expression data of circRNAs, microRNAs (miRNAs), and messenger RNAs (mRNAs) were obtained from the Gene Expression Omnibus (GEO) database. Relying on GSE101586, GSE101684, and GSE112214, we identified differentially expressed circRNAs (DEcircRNAs). Depending on GSE135918 and GSE32863, we screened out differentially expressed miRNAs (DEmiRNAs) and mRNAs (DEmRNAs), respectively. Then, a novel competing endogenous RNA (ceRNA) regulatory network related to LUAD was constructed. We also revealed biological processes and signal pathways regulated by these DEcircRNAs. Based on gene expression data and survival information of LUAD patients in The Cancer Genome Atlas (TCGA) and GEO, we implemented survival analysis to select DEmRNAs related to prognosis and build a novel circRNA–miRNA–mRNA hub regulatory network. Meanwhile, quantitative real-time PCR (qRT-PCR) was utilized to validate DEcircRNAs in the ceRNA hub regulatory network. As a result, a total of 8 DEcircRNAs, 19 DEmiRNAs, and 85 DEmRNAs were identified. The novel ceRNA regulatory network included 5 circRNAs, 8 miRNAs, and 22 mRNAs. The final ceRNA hub regulatory network contained two circRNAs, two miRNAs, and two mRNAs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses indicated that the five DEcircRNAs may affect LUAD onset and progression through Wnt signaling pathway and Hippo signaling pathway. All in all, this study revealed the regulatory network and functional mechanism of circRNA-related ceRNAs in LUAD.

Keywords: lung adenocarcinoma, circRNA, ceRNA, GEO, TCGA, qRT-PCR

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Kang He,
Zhejiang University, China
Qingyuan Wang,
Xuzhou Central Hospital, China

***Correspondence:**
Haiwei Zuo
zuohaiwei123@126.com

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *27 February 2021*
Accepted: *09 June 2021*
Published: *07 July 2021*

Citation:
Zuo H, Li X, Zheng X, Sun Q, Yang Q and Xin Y (2021) A Novel circRNA–miRNA–mRNA Hub Regulatory Network in Lung Adenocarcinoma. Front. Genet. 12:673501. doi: 10.3389/fgene.2021.673501

INTRODUCTION

Among all cancers, lung cancer has become the most commonly diagnosed cancer (11.6% of the total cases) and a major cause of cancer deaths (18.4% of the total cancer deaths) (Bray et al., 2018). Besides, most lung cancer patients suffered from lung adenocarcinoma (LUAD). The American Cancer Society (ACS) estimates that, in 2021, there will be 1,898,160 new cancer cases and 608,570 new cancer deaths in the United States (Siegel et al., 2021). A large quantity of evidence show that cancer has become one of the thorny public health problems across the world.

In recent years, in spite of the great advances in surgical techniques, chemotherapy, radiotherapy, and immunotherapy, the overall survival (OS) and prognosis of LUAD patients remain poor (Lou et al., 2021). Therefore, it is essential for the development of diagnostic and prognostic biomarkers and effective therapeutic targets to elucidate the molecular network and regulation mechanism of LUAD.

Circular RNAs (circRNAs), a kind of RNA with a more stable structure, have been reported to act as a sponge for microRNAs (miRNAs) and interact with miRNAs, reducing the ability of miRNAs to target genes, thereby regulating gene expression (Li et al., 2015). In view of this, the role and regulatory mechanism of circRNAs in tumorigenesis and tumor progression have arose wide attention among related researchers. Recent studies have revealed the function and mechanism of some circRNAs in regulating lung cancer. First, Li et al. (2020) found that hsa_circ_0030998 could affect the function of miR-558, thereby inhibiting the proliferation, migration, invasion, and taxol resistance of lung tumor cells. In non-small cell lung cancer (NSCLC), Lu M. et al. (2020) presented that circRACGAP1 could affect the tumorigenesis and tumor progression of NSCLC by means of the circRACGAP1/miR-144-5p/CDKL1 axis. Similarly, Zhang Y. et al. (2020) illustrated that circDENND2A promoted NSCLC progression through miR-34a/CCNE1 regulatory axis. Later, He et al. (2020) declared that the expression level of plasma exo-hsa_circRNA_0056616 was higher in LUAD patients with lymph node metastasis than that in LUAD patients without lymph node metastasis. This suggested that plasma exo-hsa_circRNA_0056616 might be a biomarker to predict lymph node metastasis in LUAD patients. Although some progress has been made in the research for circRNAs, studies on circRNAs in regulating LUAD occurrence and progression remain few, and there is a lot of room for improvement.

In this study, we first collected the expression profiles of circRNAs, miRNAs, and messenger RNAs (mRNAs) in LUAD from the TCGA and GEO databases, respectively. Following, based on the expression profile data, DEcircRNAs, DEmiRNAs, and DEmRNAs were identified. At the same time, we obtained circRNA–miRNA pairs and miRNA–mRNA pairs. Then, we constructed a novel circRNA–miRNA–mRNA regulatory network related to LUAD through visualizing circRNA–miRNA pairs and miRNA–mRNA pairs. In addition, we explored the biological processes and signal pathways involved in these mRNAs in the novel regulatory network via utilizing GO and KEGG enrichment analysis. For the sake of further clarifying the regulatory mechanism of circRNAs on LUAD, we downloaded comprehensive data of LUAD patients from the TCGA and GEO databases. On basis of the comprehensive data, we screened mRNAs related to the prognosis of LUAD from the mRNAs in the novel regulatory network and constructed a novel circRNA–miRNA–mRNA hub regulatory network. Finally, based on fresh frozen tissue specimen cohort, we performed quantitative real-time PCR (qRT-PCR) to validate the DEcircRNAs in the novel competing endogenous RNA (ceRNA) hub regulatory network. On the one hand, our study is beneficial to exploring diagnostic and prognostic biomarkers of LUAD. On the other hand, this study puts forward novel insights on the occurrence and progression mechanism of LUAD. The flowchart of this study is shown in **Figure 1**.

MATERIALS AND METHODS

Data Acquisition From Publicly Available Databases

Six microarray datasets (GSE101586, GSE101684, GSE112214, GSE135918, GSE32863, and GSE26939) were achieved from the GEO database[1]. Specifically, GSE101586, GSE101684, and GSE112214 contain expression data of circRNAs between 12 LUAD samples and 12 normal samples. GSE135918 contains expression data of miRNAs between five LUAD samples and five normal samples. Expression data of mRNAs are included in GSE32863 between 58 LUAD samples and 58 normal samples. GSE26939 contains expression data of mRNAs and clinical information associated with 116 LUAD patients. Besides, we also acquired a total of 1,038 LUAD samples with expression profiles and complete clinical information from the TCGA database[2]. Because all data were downloaded from publicly available databases (GEO and TCGA), the approval by the Ethics Committee was not necessary.

Fresh Frozen Tissue Specimen Cohort

The 39 pairs of fresh frozen LUAD tissues and adjacent non-tumor tissues were collected at the Affiliated Hospital of Xuzhou Medical University. These samples were preserved at −80°C for qRT-PCR. The project was granted approval by the Ethics Committee of the Affiliated Hospital of Xuzhou Medical University.

Identification of Differentially Expressed circRNAs, miRNAs, and mRNAs

First, we adopted the normalization and logarithmic method to preprocess the raw microarray data (GSE101586, GSE101684, GSE112214, GSE135918, and GSE32863). Then, based on the Bioconductor Limma package, we integrated preprocessed microarray data (GSE101586, GSE101684, and GSE112214) to identify DEcircRNAs with thresholds of $|\log2(\text{fold change})| > 1$ and $p < 0.05$. Simultaneously, we utilized robust rank aggregation method to rank all of the DEcircRNAs. On the basis of preprocessed microarray data (GSE135918 and GSE32863), we also obtained DEmiRNAs and DEmRNAs, respectively. In addition, the ggplot2 package was applied to visualize DEcircRNAs, DEmiRNAs, and DEmRNAs.

Prediction of miRNA Binding Sites and miRNA Target Genes

In this study, we predicted miRNA binding sites (MREs) through integrating the Circular RNA Interactome (CircInteractome) database[3] (Lai et al., 2020) and the Cancer-Specific CircRNA

[1]http://www.ncbi.nlm.nih.gov/geo

[2]https://tcga-data.nci.nih.gov/tcga/

[3]https://circinteractome.nia.nih.gov/

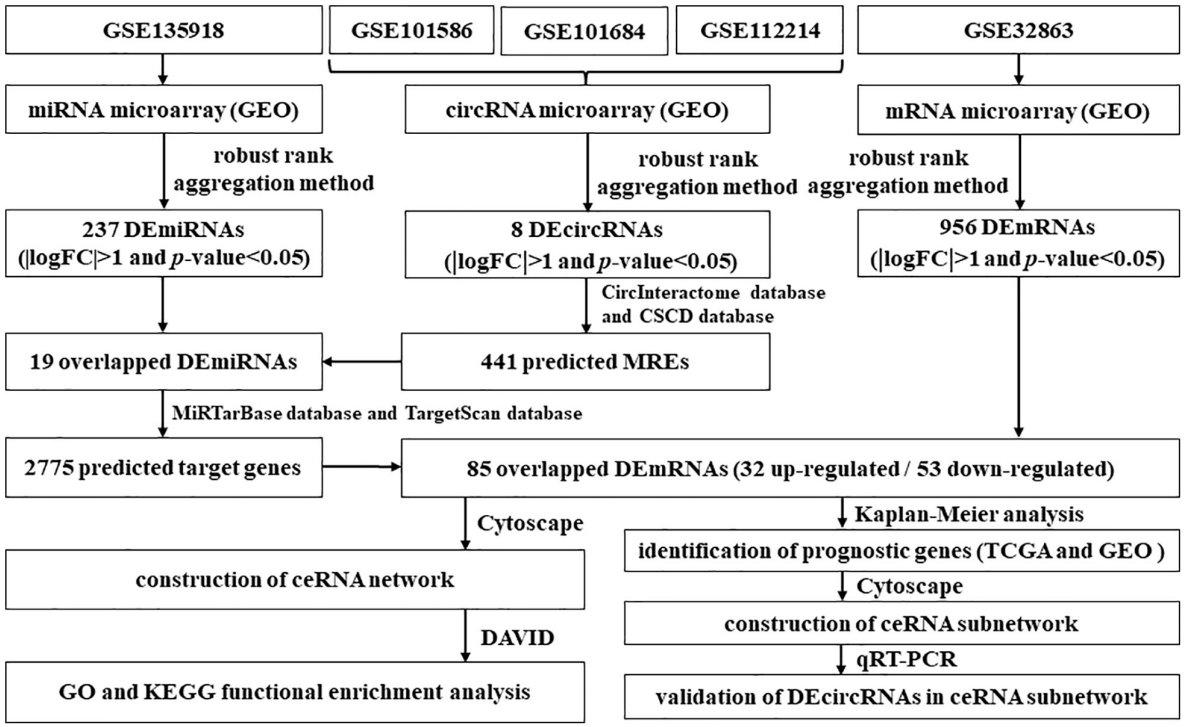


FIGURE 1 | Flowchart of the comprehensive analysis process and all methods utilized in present study.

(CSCD) database[4] (Abramowicz and Story, 2020). In detail, we treated the overlapped miRNAs between the CircInteractome database and the CSCD database as potential target miRNAs of DEcircRNAs from (GSE101586, GSE101684, and GSE112214). After that, these predicted target miRNAs were further screened by DEmiRNA from GSE135918. Furthermore, based on the TargetScan database[5] (Wong and Wang, 2015), the miRTarBase database[6] (Chou et al., 2016), and the miRDB database[7] (Fromm et al., 2015), we predicted the interactions between miRNAs and mRNAs. Specifically, these mRNAs identified by the three databases were regarded as candidate mRNAs. The candidate mRNAs were further screened by DEmRNAs from GSE32863.

[4]https://gb.whu.edu.cn/CSCD/

[5]http://www.targetscan.org/vert_72/

[6]http://mirtarbase.cuhk.edu.cn/php/index.php

[7]http://mirdb.org/

TABLE 1 | Basic information of five microarray datasets from the GEO database.

Data source	Series	Platform	Author	Year	Area	Sample size (T/N)	No. of RNAs
circRNA	GSE101586	GPL19978	Qiu	2019	China	5/5	4,425
circRNA	GSE101684	GPL21825	Xu	2019	China	4/4	9,114
circRNA	GSE112214	GPL19978	Dong	2019	China	3/3	4,407
miRNA	GSE135918	GPL18058	Jiang	2020	China	5/5	3,551
mRNA	GSE32863	GPL6884	Selamat	2019	United States	58/58	48,803

Construction of the Novel ceRNA Network Related to LUAD

Relying on the identified DEcircRNAs, the predicted target miRNAs, and the predicted target mRNAs, we could determine the number and types of circRNA–miRNA pairs as well as the miRNA–mRNA pairs. Then, based on these circRNA–miRNA pairs and miRNA–mRNA pairs, the novel circRNA–miRNA–mRNA regulatory network related to LUAD was constructed and visualized via utilizing Cytoscape 3.6.1 software.

GO and KEGG Functional Enrichment Analysis

For the purpose of exploring the molecular mechanism of genes in the novel ceRNA regulatory network related to

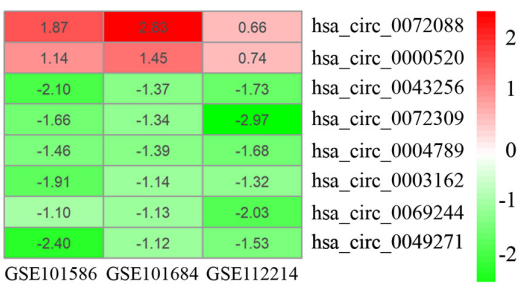


FIGURE 2 | Heatmap of eight DEcircRNAs in three microarray datasets.

TABLE 2 | Basic characteristics of eight DEcircRNAs.

circRNA ID	Position	Genomic length	Strand	Best transcript	Gene symbol	Regulation
hsa_circ_0072088	chr5:32379220-32388780	9,560	–	NM_016107	ZFR	Up
hsa_circ_0000520	chr14:20811436-20811559	123	–	NR_002312	RPPH1	Up
hsa_circ_0043256	chr17:35604934-35609962	5,028	–	NM_198839	ACACA	Down
hsa_circ_0072309	chr5:38523520-38530768	7,248	–	NM_001127671	LIFR	Down
hsa_circ_0004789	chr17:62587201-62594608	7,407	–	NM_022739	SMURF2	Down
hsa_circ_0003162	chr7:33185853-33217203	31,350	+	NM_198428	BBS9	Down
hsa_circ_0069244	chr4:16587544-16760883	173,339	–	NM_001130834	LDB2	Down
hsa_circ_0049271	chr19:10610070-10610756	686	–	NM_203500	KEAP1	Down

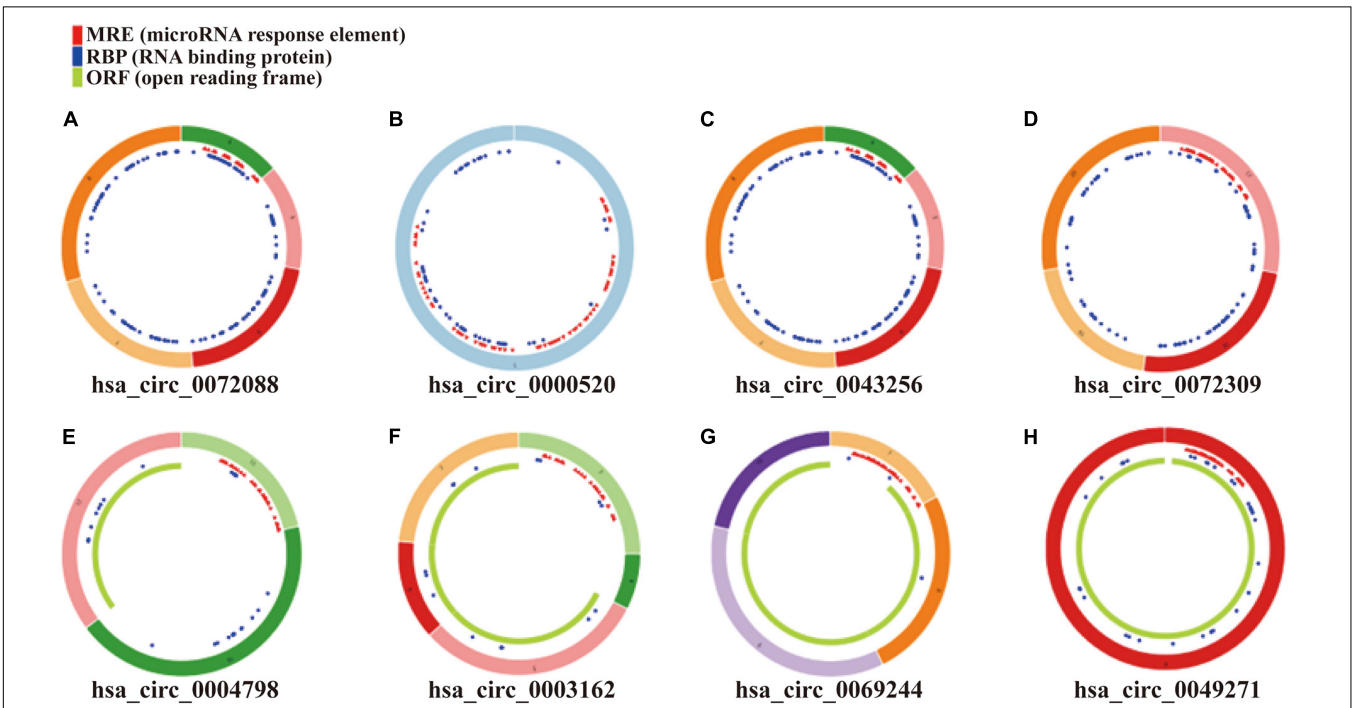


FIGURE 3 | Structural patterns of eight DEcircRNAs. **(A)** hsa_circ_0072088, **(B)** hsa_circ_0000520, **(C)** hsa_circ_0043256, **(D)** hsa_circ_0072309, **(E)** hsa_circ_0004789, **(F)** hsa_circ_0003162, **(G)** hsa_circ_0069244, and **(H)** hsa_circ_0049271.

LUAD, we implemented GO and KEGG functional enrichment analysis on basis of the Database for Annotation, Visualization and Integrated Discovery (DAVID)[8] v6.8. In GO and KEGG functional enrichment analysis, $p < 0.05$ was given to imply a statistically significant difference. What is more, GO enrichment analysis consists of cellular component (CC), molecular function (MF), and biological process (BP), which, respectively, demonstrate the biological functions of genes at different levels (Zhang L. et al., 2019; Sun et al., 2020). KEGG pathway enrichment analysis is applied to assess the enrichment degree of genes in different pathways. DAVID is an online biological information database that integrates biological data and analysis tools and provides gene annotation information and protein data.

[8]http://david.abcc.ncifcrf.gov/

Survival Analysis, RNA Extraction, and qRT-PCR

In order to select mRNAs associated with LUAD prognosis, we downloaded the mRNA expression profiles and the corresponding LUAD patient clinical information from the TCGA database and GEO database. Then, based on the survival algorithm package implemented by the R language, we conducted overall survival (OS) analysis experiments to verify each gene in the novel ceRNA regulatory network and plot Kaplan–Meier curves. The log-rank test was applied to achieve statistical verification ($p < 0.05$). The novel ceRNA hub regulatory network was constructed on the basis of these mRNAs related to LUAD prognosis. In addition, we utilized qRT-PCR to further validate the DEcircRNAs in the novel ceRNA hub regulatory network. According to instructions, we utilized Trizol reagent (Invitrogen, Carlsbad, CA, United States) to extract total RNA from fresh frozen LUAD tissues and adjacent

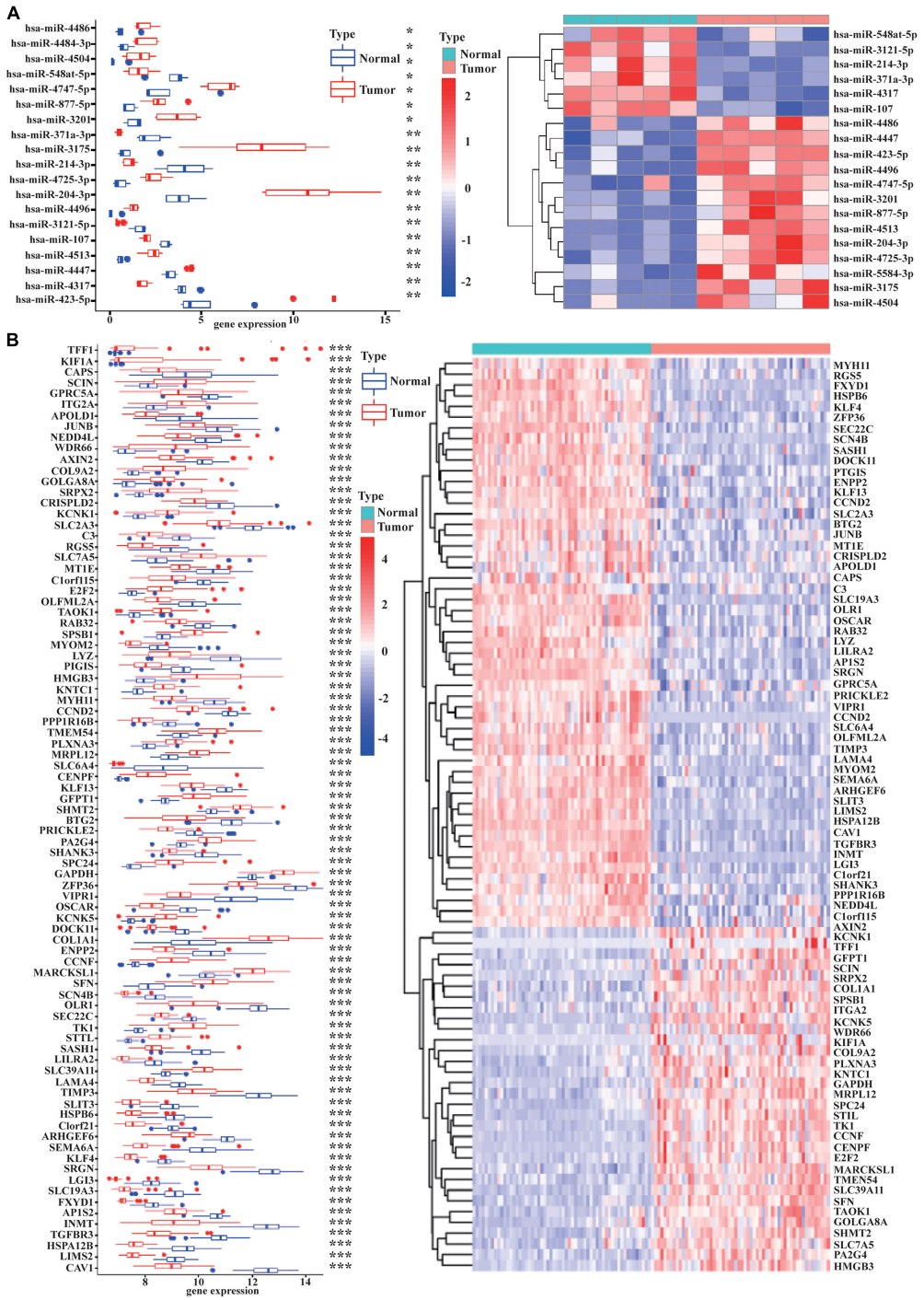


FIGURE 4 | The 19 overlapped DEmiRNAs and the 85 overlapped DEmRNAs. The boxplot and heatmap of the 19 overlapped DEmiRNAs are shown in **(A)**. The 19 DEmiRNAs were selected through overlapping the 441 miRNAs binding to the 8 DEcircRNAs and the 237 DEmiRNAs identified in GSE135918. The expression levels of the 19 overlapped DEmiRNAs are described in **(A)**. The boxplot and heatmap of the 85 overlapped DEmRNAs are shown in **(B)**. The 85 DEmiRNAs were obtained by overlapping the 2,775 target genes binding to the 19 DEmiRNAs and the 956 DEmRNAs identified in GSE32863. The expression levels of the 85 overlapped DEmRNAs are displayed in **(B)**. [| log2(fold change)| > 1 and $p < 0.05$ were the cutoff criteria]. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

non-tumor tissues. Then, Trans Script one-step guide DNA removal and complementary DNA synthesis super mix were used for the reverse transcription reaction. The primer sequences for PCR amplification were as follows: hsa_circ_0049271, forward, 5′-GCCCGGGAGTACATCTACAT-3′, reverse, 5′-TC AGTGGAGGCGTACATCAC-3′; hsa_circ_0004789, forward,

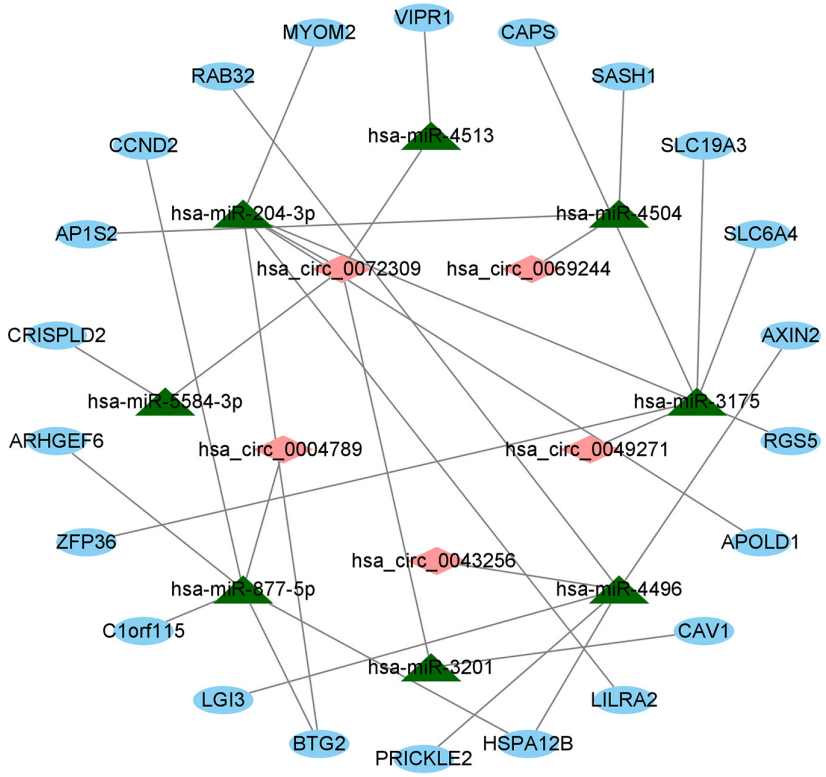


FIGURE 5 | The ceRNA regulatory network on basis of circRNA–miRNA pairs and miRNA–mRNA pairs in LUAD. Red diamonds indicate circRNAs, green triangles indicate miRNAs, and blue ovals indicate mRNAs.

5′-TCTCGGTTGTGTTCGTCTTC-3′, reverse, 5′-CCATCAACCACCACCTTAGC-3′; and hsa_circ_0043256, forward, 5′-TGTGGTGATCATGAATGGCTC-3′, reverse, 5′-TCACCCCGAATAGACAGCTC-3′. After qRT-PCR verification, we obtained the final ceRNA hub regulatory network associated with LUAD.

RESULTS

Identification of DEcircRNAs, DEmiRNAs, and DEmRNAs in LUAD

As displayed in **Table 1**, we have listed basic information of five microarray datasets (GSE101586, GSE101684, GSE112214, GSE135918, and GSE32863) regarding circRNAs, miRNAs, and mRNAs. Through integrating the data in GSE101586, GSE101684, and GSE112214, we identified eight DEcircRNAs (two upregulated circRNAs and six downregulated circRNAs, **Figure 2**). At the same time, the basic information and structural patterns of the eight DEcircRNAs are shown in **Table 2** and **Figure 3**, respectively. Based on the CircInteractome and CSCD databases, we predicted 441 potential target miRNAs associated with the eight DEcircRNAs. Besides, a total of 237 DEmiRNAs (116 upregulated miRNAs and 121 downregulated miRNAs) were identified in GSE135918. Then, 19 overlapped DEmiRNAs (13 upregulated miRNAs and 6 downregulated miRNAs, **Figure 4A**) were acquired. Moreover, through taking advantages of TargetScan, miRTarBase, and miRDB databases, we obtained 2,775 target genes associated with the 19 overlapped DEmiRNAs. Simultaneously, a total of 956 DEmRNAs (339 upregulated mRNAs and 617 downregulated mRNAs) were identified in GSE32863. Eventually, we gained 85 overlapped DEmRNAs (32 upregulated mRNAs and 53 downregulated mRNAs, **Figure 4B**) in this research.

Construction of the ceRNA Regulatory Network in LUAD

For the sake of further clarifying the mechanism by which circRNAs and miRNAs affect the occurrence and progression of LUAD, we established a ceRNA regulatory network and visualized the network. First, based on the predicted circRNA–miRNA pairs and miRNA–mRNA pairs, we identified eight miRNAs that have two types of relationships at the same time. Then, on the basis of eight DEcircRNAs, eight identified miRNAs and the predicted circRNA-miRNA pairs, a total of five circRNAs were obtained. Simultaneously, we received 22 mRNAs involved in the ceRNA regulatory network according to the 85 overlapped DEmRNAs, the 8 identified miRNAs, and the predicted miRNA–mRNA pairs. Furthermore, we also gained eight circRNA–miRNA pairs between the five identified circRNAs and the eight identified miRNAs, and 24 miRNA–mRNA pairs between the eight identified miRNAs and the 22 identified mRNAs. Ultimately, as described in **Figure 5**, a ceRNA regulatory network

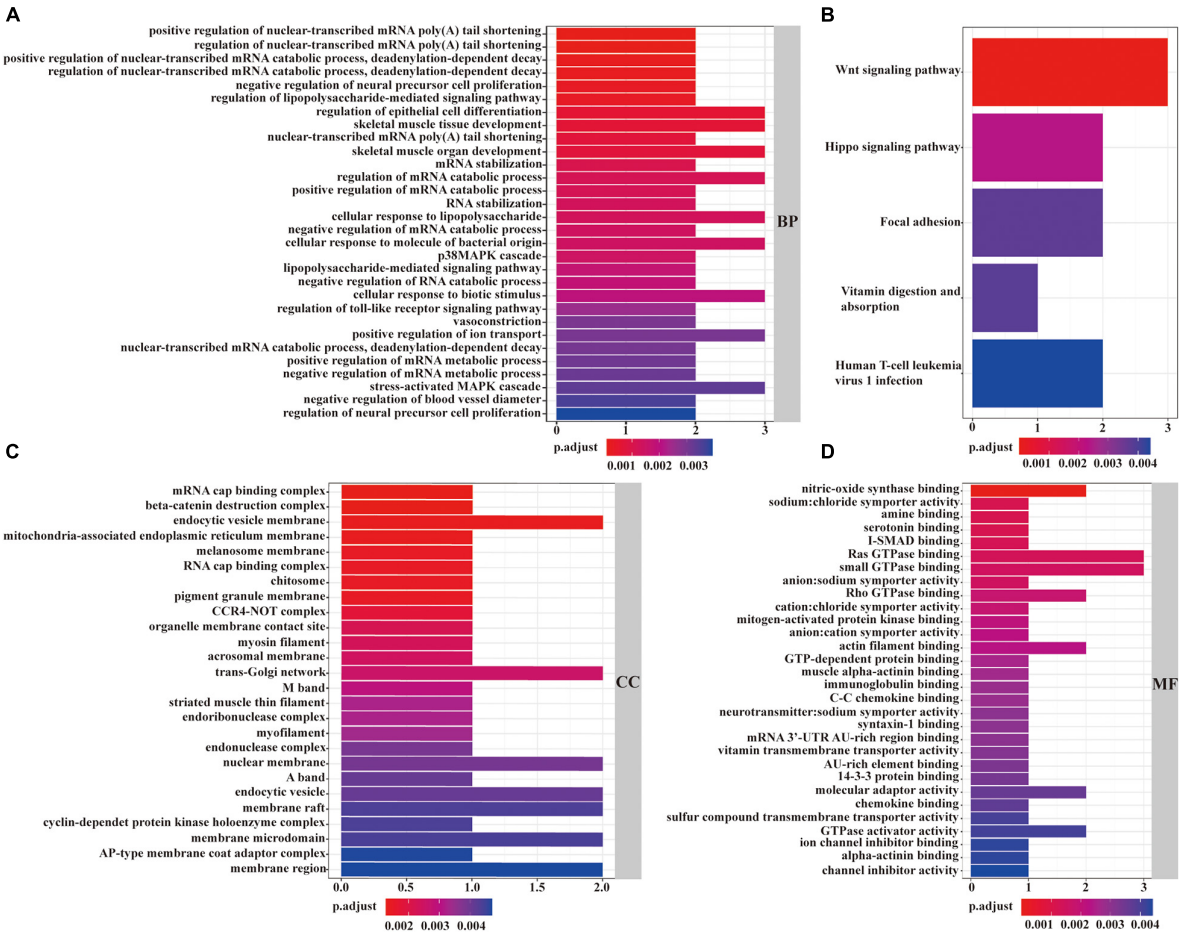


FIGURE 6 | GO and KEGG functional enrichment analyses of mRNAs in the ceRNA network. The enrichment of biological processes is shown in **(A)**. The enrichment of cellular components was described in **(C)**. The enrichment of molecular functions is displayed in **(D)**. The enrichment analysis of KEGG pathways is shown in **(B)**. A $p < 0.05$ was considered to indicate a statistically significant difference.

with the identified 32 edges (the 8 circRNA–miRNA edges and the 24 miRNA–mRNA edges) and the 35 identified nodes (5 circRNAs, 8 miRNAs, and 22 mRNAs) was constructed through utilizing the Cytoscape 3.8.0 software.

Functional and Pathway Enrichment Analysis

As described in **Figure 6A**, in the view of BPs, the identified mRNAs in the ceRNA network were mainly enriched in "regulation of epithelial cell differentiation," "skeletal muscle tissue development," "regulation of mRNA catabolic process," "cellular response to lipopolysaccharide," and "cellular response to lipopolysaccharide." According to CCs, these mRNAs were mainly enriched in "endocytic vesicle membrane," "trans-Golgi network," "mRNA cap binding complex," "pigment granule membrane," and "organelle membrane contact site" (**Figure 6C**). From the point of view of MFs, the mRNAs in the ceRNA network were mainly enriched in "nitric-oxide synthase binding," "Ras GTPase binding," "small GTPase binding," "actin filament binding," and "molecular adaptor activity" (**Figure 6D**). Besides, KEGG pathway analysis revealed the strong enrichment in "Wnt signaling pathway," "Hippo signaling pathway," "focal adhesion," "vitamin digestion and absorption," and "human T-cell leukemia virus 1 infection" (**Figure 6B**).

Survival Analysis, qRT-PCR Validation, and the Construction of a ceRNA Hub Regulatory Network

First, we downloaded 1,038 LUAD samples with expression profiles and complete clinical information from the TCGA database. After filtering out samples without survival information, we kept 999 samples. According to the median expression value, we divided the samples into the high expression group (499 samples) and the low expression group (500 samples). Then, the survival analysis was carried out for each mRNA (a total of 22 mRNAs in circRNA–miRNA–mRNA regulatory network). We observed that five mRNAs (C1orf115, CAPS, CCND2, PRICKLE2, and VIPR1) were associated with prognosis in LUAD patients (**Figures 7A–E**). To further investigate the significance of the five mRNAs in prognosis prediction of

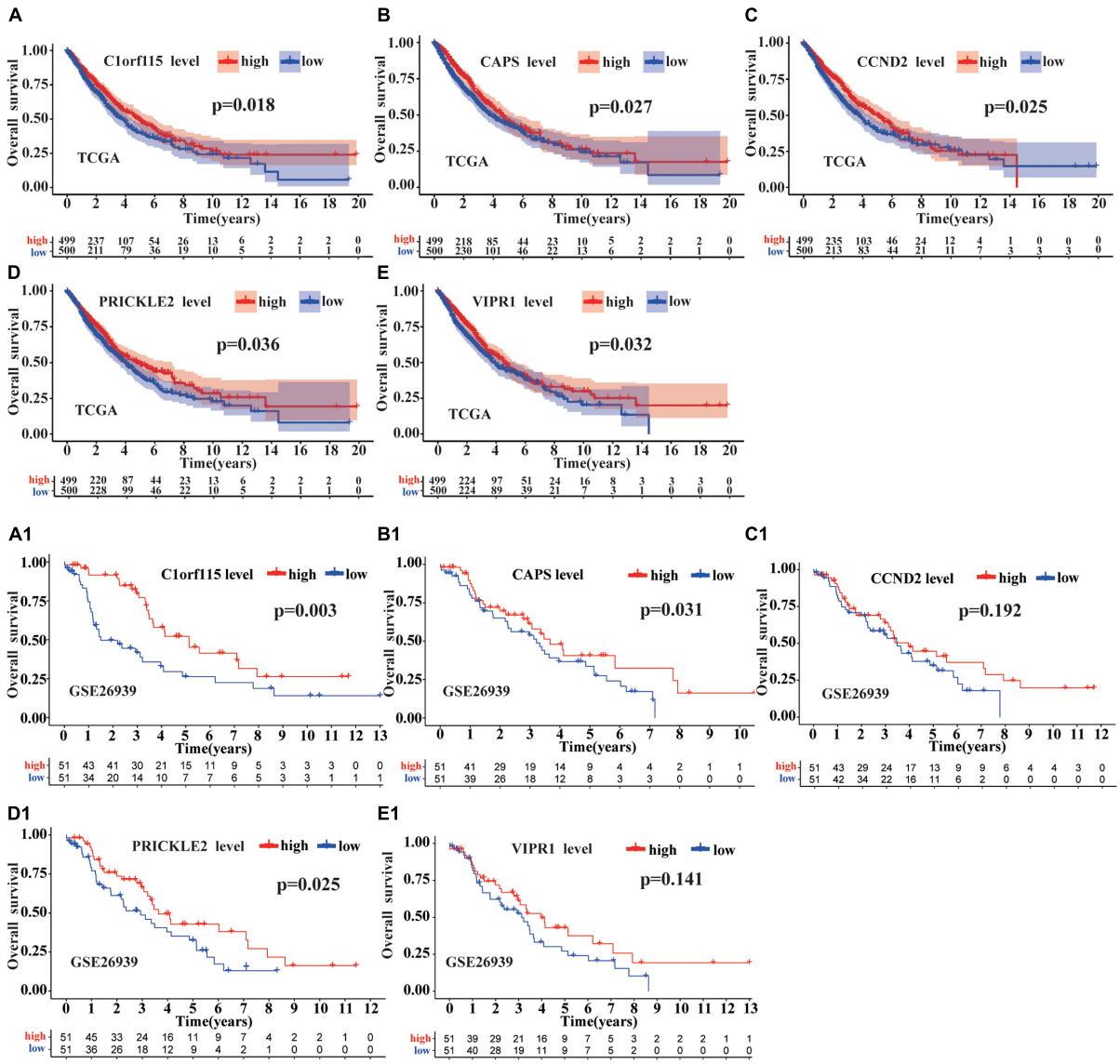


FIGURE 7 | Survival analysis of mRNAs in ceRNA regulatory network. The survival prognosis of 22 mRNAs was analyzed based on the TCGA database. As shown in **(A)**, compared with the group with low C1orf115 expression, the group with high C1orf115 expression had a better prognosis in LUAD (p = 0.018). As displayed in Panels **(B–E)**, we have reached the same conclusion for CAPS, CCND2, PRICKLE2, and VIPR1 (p = 0.027, p = 0.025, p = 0.036, and p = 0.032). As depicted in **(A1), (B1), (C1), (D1),** and **(E1)**, based on GSE26939 array from the GEO database, only C1orf115, CAPS, and PRICKLE2 (p = 0.003, p = 0.031, and p = 0.025) have been verified to be consistent with the above conclusions.

LUAD, we obtained GSE26939 array from the GEO database (51 samples in the high expression group and 51 samples in the low expression group) and conducted the survival analysis experiment. As a result, three mRNAs (C1orf115, CAPS, and PRICKLE2) were identified to be related to the prognosis prediction of LUAD (**Figures 7A1–E1**). The above analysis showed that C1orf115 was significantly associated with prognosis in LUAD patients (p = 0.018 in TCGA and p = 0.003 in GSE26939). Besides, from the statistical perspective, CAPS and PRICKLE2 had certain relationships with prognosis prediction of LUAD, but the observed differences in median survival between the high expression group and the low expression group were not prominent. Considering the impact of the sample size on statistical analysis, and the significant differences in expression levels of CAPS and PRICKLE2 between normal tissues and LUAD tissues, we accepted that CAPS and PRICKLE2 were related to the prognosis prediction of LUAD from the biological perspective. On basis of the three overlapped mRNAs (C1orf115, CAPS, and PRICKLE2), we constructed a circRNA–miRNA–mRNA hub regulatory network (**Figure 8A**). As shown in **Figures 8B–D**, qRT-PCR result confirmed that hsa_circ_0049271 was not significantly differentially expressed between LUAD tissues and adjacent non-tumor tissues (p = 0.6936). Thus, we held the opinion that hsa_circ_0049271/hsa-miR-3175/CAPS

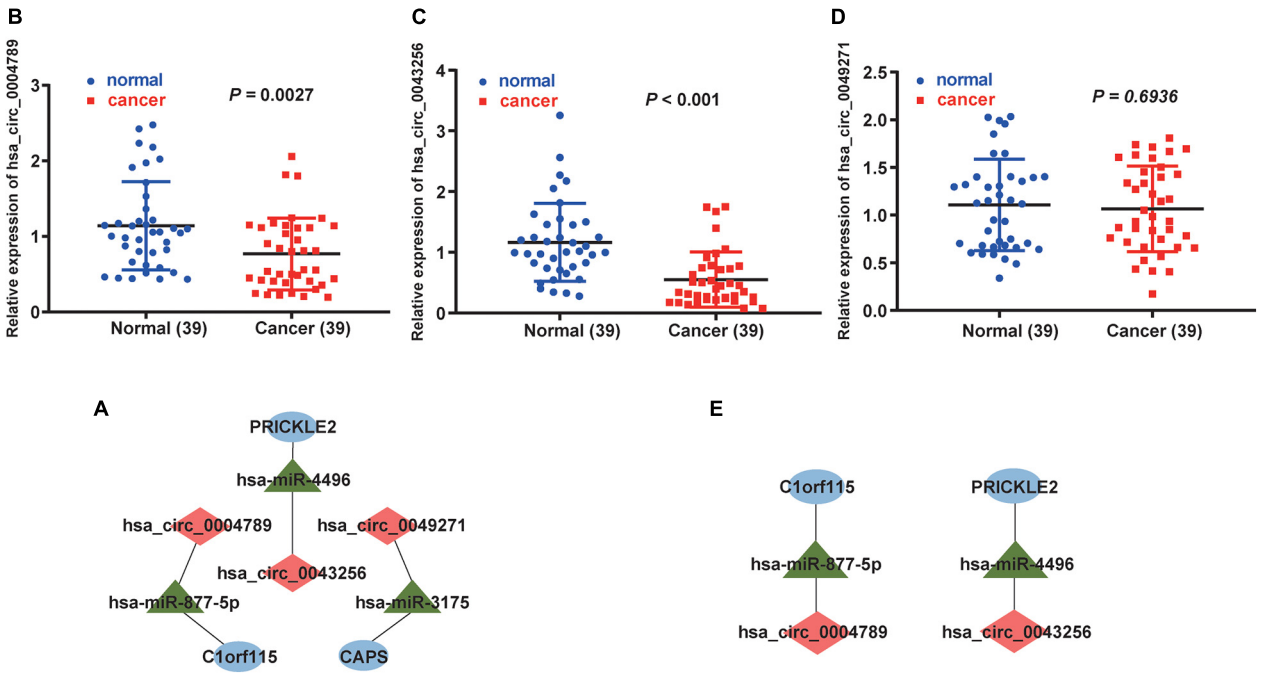


FIGURE 8 | Construction of the final ceRNA hub regulatory network. **(A)** Construction of the ceRNA hub regulatory network based on the mRNAs associated with LUAD prognosis. As shown in **(B–D)**, expression of hsa_circ_0004789, hsa_circ_0043256, and hsa_circ_0049271 in 39 LUAD tissues and corresponding non-tumor tissues were detected by qRT-PCR, respectively ($p = 0.0027$, $p < 0.001$, and $p = 0.6936$). **(E)** After qRT-PCR verification, the final ceRNA hub regulatory network was constructed. In the ceRNA hub regulatory network, red diamonds indicate circRNAs, green triangles indicate miRNAs, and blue ovals indicate mRNAs.

axis might not regulate the LUAD occurrence and progression. The independent regulatory mechanism of hsa_circ_0049271 downstream network needs further verification. In this study, we removed hsa_circ_0049271/hsa-miR-3175/CAPS axis and kept hsa_circ_0004789/hsa-miR-877-5p/C1orf115 and hsa_circ_0043256/hsa-miR-4496/PRICKLE2 axes verified by qRT-PCR to obtain the final ceRNA hub regulatory network (**Figure 8E**).

DISCUSSION

CircRNAs are a special type of non-coding RNAs with a circular biological structure (Zhang L. et al., 2020). Due to the closed loop structure, circRNAs are not easily affected by the RNA exonuclease and have the high stability in the human body (Cao et al., 2021). CircRNAs have been observed in various cancers and can affect the occurrence and progression of different cancers, such as colorectal cancer (Chen J. et al., 2020), gastric cancer (Chen H. et al., 2020), and liver cancer (Tang et al., 2020). Evidence has shown that circRNAs could indirectly regulate protein-coding genes through reducing the expression level of related miRNAs. The latest research announced circRNAs might play a decisive role in the occurrence and progression of tumors, and some circRNAs could become important biomarkers of cancer prognosis.

Lung adenocarcinoma, a type of NSCLC, is the most common kind of lung cancer, accounting for approximately 40% in all types of lung cancers (Xu J.Y. et al., 2020). LUAD is frequently detected at the metastatic stage with prevalence in the brain, bones, and respiratory system (Kim et al., 2020). The current treatments for LUAD are mainly surgery, radiotherapy, chemotherapy, and immunotherapy, but the prognosis of patients is still unsatisfactory. Due to their stable structure, circRNAs are promised to become a new strategy for LUAD-targeted therapy. Xu et al. observed that HMGA2 could regulate circASPH to promote tumor growth in patients with LUAD (Xu L. et al., 2020). The study of Lu G.J. et al. (2020) revealed that hsa_circ_0001715 was upregulated in LUAD and had expectation to become a novel diagnostic and prognostic marker. Besides, Wang et al. (2020) demonstrated that based on the miR-138/Sirt1 pathway, hsa_circ_0006571 could promote the migration and invasion of LUAD cells. Although these studies indicate that circRNAs are abnormally expressed in LUAD and have associations with LUAD occurrence and progression, the specific functional pathway and molecular mechanism of circRNAs in LUAD are still unclear.

To explore the role of circRNAs in LUAD, we constructed a ceRNA regulatory network based on data mining, bioinformatics methods, and public databases. GO analysis displayed that genes in our ceRNA regulatory network were involved in the regulation of beta-catenin destruction complex, trans-Golgi network, and small GTPase binding. Some studies declared that non-coding RNAs could regulate the progression cancer cells via regulating the Wnt/beta-catenin signal. Not only that, the Wnt/beta-catenin pathway could also promote cancer stem cell self-renewal, metastasis, and chemoresistance in all types

of epithelial ovarian cancer (Nguyen et al., 2019). Bao et al. (2015) proved that via interfering with trans-Golgi network trafficking, small molecule schweinfurthins selectively inhibited cancer cell proliferation and mTOR/AKT signaling. Additionally, in view of the intimate crosstalk of Rho/Rac1/Rac1b and TGF-beta signaling in tumor cell responses, targeting specific Rho GTPases might allow for selective interference with prooncogenic TGF-beta responses to aid in anticancer treatments (Ungefroren et al., 2018). KEGG analysis indicated that mRNAs in this ceRNA regulatory network was enriched in Wnt signaling pathway and Hippo signaling pathway, which have been confirmed by many studies to participate in cancer progression (Ansari et al., 2019; Flanagan et al., 2019). Besides, Namani et al. revealed that NRF2 could regulate the occurrence and progression of NSCLC by regulating genes involved in the focal adhesion pathway (Namani et al., 2019). It has not been reported that "vitamin digestion and absorption" and "human T-cell leukemia virus 1 infection" can regulate cancer progression, and further investigations are needed. All in all, the above evidence suggest that these circRNAs from this ceRNA regulatory network may act as a key role in LUAD occurrence and progression.

Based on the TCGA and GEO databases, we analyzed the correlation between expression levels of 22 mRNAs in the ceRNA regulatory network and OS of LUAD patients. As a result, three mRNAs (C1orf115, CAPS, and PRICKLE2) were identified to be associated with the prognosis of LUAD patients. Then, based on mRNAs related to LUAD prognosis, we constructed a novel ceRNA hub regulatory network. Hsa_circ_0004789 and hsa_circ_0043256 were verified by qRT-PCR to be differentially expressed between LUAD tissues and adjacent non-tumor tissues. Two miRNAs (hsa-miR-877-5p and hsa-miR-4496) were also included in final ceRNA hub regulatory network. It has been reported that the loss of C1orf115 can lead to the resistance to five anticancer drugs, and the low expression of C1orf115 is related to the poor prognosis of many cancers (Lau et al., 2020). Besides, Senchenko et al. (2013) proved that PRICKLE2 was significantly related to the prognosis of various cancers. What is more, the specific functional pathway and molecular mechanism of the two identified genes in final novel ceRNA hub regulatory network need to be further studied.

The significance of this study lies in not only constructing a novel ceRNA hub regulatory network through utilizing GEO database but also screening out mRNAs associated with LUAD prognosis via integrating clinical data in TCGA and GEO databases. However, this research also has some limitations because it mainly relies on analysis of sequencing data and public clinical data and requires further experimental explorations.

In conclusion, based on publicly available sequencing and clinical data, we identified differentially expressed circRNAs, miRNAs, and mRNAs and constructed a circRNA-related regulatory network. Through utilizing TCGA and GEO databases, we identified three mRNAs related to LUAD prognosis and built a novel ceRNA hub regulatory network. Two circRNAs in final ceRNA hub regulatory network were validated by qRT-PCR to be differentially expressed between LUAD tissues and adjacent non-tumor tissues. This study may provide new insights into the pathogenesis of LUAD and reveal potential therapeutic targets.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

ETHICS STATEMENT

The project was granted approval by the Ethics Committee of the Affiliated Hospital of Xuzhou Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HZ and XL designed the overall idea of this study, conceived the experiments, analyzed the data, prepared the figures and tables, and authored the drafts of the manuscript. XZ and QS collected the data from the TCGA and GEO dataset and performed the experiments. QY and YX supervised this study and reviewed the drafts of the manuscript. All authors read and approved the final draft.

FUNDING

This study was supported by Talented Scientific Research Foundation of Xuzhou Medical University (No. D2018018).

REFERENCES

Abramowicz, A., and Story, M. D. (2020). The Long and Short of It: the Emerging Roles of Non-Coding RNA in Small Extracellular Vesicles. *Cancers* 12:1445. doi: 10.3390/cancers12061445

Ansari, D., Ohlsson, H., Althini, C., Bauden, M., Zhou, Q., Hu, D., et al. (2019). The Hippo Signaling Pathway in Pancreatic Cancer. *Anticancer Res.* 39, 3317–3321. doi: 10.21873/anticanres.13474

Bao, X., Zheng, W., Hata Sugi, N., Agarwala, K. L., Xu, Q., Wang, Z., et al. (2015). Small molecule schweinfurthins selectively inhibit cancer cell proliferation and mTOR/AKT signaling by interfering with trans-Golgi-network trafficking. *Cancer Biol. Ther.* 16, 589–601. doi: 10.1080/15384047.2015.1019184

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492

Cao, Y. Z., Sun, J. Y., Chen, Y. X., Wen, C. C., and Wei, L. (2021). The roles of circRNAs in cancers: perspectives from molecular functions. *Gene* 767:145182. doi: 10.1016/j.gene.2020.145182

Chen, H., Liang, C., Wang, X., Liu, Y., Yang, Z., Shen, M., et al. (2020). The prognostic value of circRNAs for gastric cancer: a systematic

review and meta-analysis. *Cancer Med.* 9, 9096–9106. doi: 10.1002/cam4.3497

Chen, J., Yang, X., Liu, R., Wen, C., Wang, H., Huang, L., et al. (2020). Circular RNA GLIS2 promotes colorectal cancer cell motility via activation of the NF-κB pathway. *Cell Death Dis.* 11:788. doi: 10.1038/s41419-020-02989-7

Chou, C. H., Chang, N. W., Shrestha, S., Hsu, S. D., Lin, Y. L., Lee, W. H., et al. (2016). miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res.* 44, D239–D247. doi: 10.1093/nar/gkv1258

Flanagan, D. J., Vincan, E., and Phesse, T. J. (2019). Wnt Signaling in Cancer: not a Binary ON:OFF Switch. *Cancer Res.* 79, 5901–5906. doi: 10.1158/0008-5472.Can-19-1362

Fromm, B., Billipp, T., Peck, L. E., Johansen, M., Tarver, J. E., King, B. L., et al. (2015). A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annu. Rev. Genet.* 49, 213–242. doi: 10.1146/annurev-genet-120213-092023

He, F., Zhong, X., Lin, Z., Lin, J., Qiu, M., Li, X., et al. (2020). Plasma exo-hsa_circRNA_0056616: a potential biomarker for lymph node metastasis in lung adenocarcinoma. *J. Cancer* 11, 4037–4046. doi: 10.7150/jca.30360

Kim, N., Kim, H. K., Lee, K., Hong, Y., Cho, J. H., Choi, J. W., et al. (2020). Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* 11:2285. doi: 10.1038/s41467-020-16164-1

Lai, J., Xin, J., Fu, C., and Zhang, W. (2020). CircHIPK3 promotes proliferation and metastasis and inhibits apoptosis of renal cancer cells by inhibiting MiR-485-3p. *Cancer Cell Int.* 20:248. doi: 10.1186/s12935-020-01319-3

Lau, M. T., Ghazanfar, S., Parkin, A., Chou, A., Rouaen, J. R., Littleboy, J. B., et al. (2020). Systematic functional identification of cancer multi-drug resistance genes. *Genome Biol.* 21:27. doi: 10.1186/s13059-020-1940-8

Li, J., Yang, J., Zhou, P., Le, Y., Zhou, C., Wang, S., et al. (2015). Circular RNAs in cancer: novel insights into origins, properties, functions and implications. *Am. J. Cancer Res.* 5, 472–480.

Li, X., Feng, Y., Yang, B., Xiao, T., Ren, H., Yu, X., et al. (2020). A novel circular RNA, hsa_circ_0030998 suppresses lung cancer tumorigenesis and Taxol resistance by sponging miR-558. *Mol. Oncol.* doi: 10.1002/1878-0261.12852 [Online ahead of print]

Lou, Y., Xu, J., Zhang, Y., Zhang, W., Zhang, X., Gu, P., et al. (2021). Akt kinase LANCL2 functions as a key driver in EGFR-mutant lung adenocarcinoma tumorigenesis. *Cell Death Dis.* 12:170. doi: 10.1038/s41419-021-03439-8

Lu, G. J., Cui, J., Qian, Q., Hou, Z. B., Xie, H. Y., Hu, W., et al. (2020). Overexpression of hsa_circ_0001715 is a Potential Diagnostic and Prognostic Biomarker in Lung Adenocarcinoma. *Onco Targets Ther.* 13, 10775–10783. doi: 10.2147/ott.S274932

Lu, M., Xiong, H., Xia, Z. K., Liu, B., Wu, F., Zhang, H. X., et al. (2020). circRACGAP1 promotes non-small cell lung cancer proliferation by regulating miR-144-5p/CDKL1 signaling pathway. *Cancer Gene Ther.* 28, 197–211. doi: 10.1038/s41417-020-00209-0

Namani, A., Liu, K., Wang, S., Zhou, X., Liao, Y., Wang, H., et al. (2019). Genome-wide global identification of NRF2 binding sites in A549 non-small cell lung cancer cells by ChIP-Seq reveals NRF2 regulation of genes involved in focal adhesion pathways. *Aging* 11, 12600–12623. doi: 10.18632/aging.102590

Nguyen, V. H. L., Hough, R., Bernaudo, S., and Peng, C. (2019). Wnt/β-catenin signalling in ovarian cancer: insights into its hyperactivation and function in tumorigenesis. *J. Ovarian Res.* 12:122. doi: 10.1186/s13048-019-0596-z

Senchenko, V. N., Kisseljova, N. P., Ivanova, T. A., Dmitriev, A. A., Krasnov, G. S., Kudryavtseva, A. V., et al. (2013). Novel tumor suppressor candidates on chromosome 3 revealed by NotI-microarrays in cervical cancer. *Epigenetics* 8, 409–420. doi: 10.4161/epi.24233

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics, 2021. *CA Cancer J. Clin.* 71, 7–33. doi: 10.3322/caac.21654

Sun, Q., Li, X., Xu, M., Zhang, L., Zuo, H., Xin, Y., et al. (2020). Differential Expression and Bioinformatics Analysis of circRNA in Non-small Cell Lung Cancer. *Front. Genet.* 11:586814. doi: 10.3389/fgene.2020.586814

Tang, Y., Jiang, M., Jiang, H. M., Ye, Z. J., Huang, Y. S., Li, X. S., et al. (2020). The Roles of circRNAs in Liver Cancer Immunity. *Front. Oncol.* 10:598464. doi: 10.3389/fonc.2020.598464

Ungefroren, H., Witte, D., and Lehnert, H. (2018). The role of small GTPases of the Rho/Rac family in TGF-β-induced EMT and cell motility in cancer. *Dev. Dyn.* 247, 451–461. doi: 10.1002/dvdy.24505

Wang, H. L., Wang, H. R., Liang, Y., Hu, A. N., Enguita, F. J., Zhou, X. G., et al. (2020). Hsa_circ_0006571 promotes spinal metastasis through sponging microRNA-138 to regulate sirtuin 1 expression in lung adenocarcinoma. *Transl. Lung Cancer Res.* 9, 2411–2427. doi: 10.21037/tlcr-20-1250

Wong, N., and Wang, X. (2015). miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* 43, D146–D152. doi: 10.1093/nar/gku1104

Xu, J. Y., Zhang, C., Wang, X., Zhai, L., Ma, Y., Mao, Y., et al. (2020). Integrative Proteomic Characterization of Human Lung Adenocarcinoma. *Cell* 182, 245–261.e17. doi: 10.1016/j.cell.2020.05.043

Xu, L., Ma, Y., Zhang, H., Lu, Q. J., Yang, L., Jiang, G. N., et al. (2020). HMGA2 regulates circular RNA ASPH to promote tumor growth in lung adenocarcinoma. *Cell Death Dis.* 11:593. doi: 10.1038/s41419-020-2726-3

Zhang, L., Chen, X., and Yin, J. (2019). Prediction of Potential miRNA-Disease Associations Through a Novel Unsupervised Deep Learning Framework with Variational Autoencoder. *Cells* 8:1040. doi: 10.3390/cells8091040

Zhang, L., Hou, C., Chen, C., Guo, Y., Yuan, W., Yin, D., et al. (2020). The role of N(6)-methyladenosine (m(6)A) modification in the regulation of circRNAs. *Mol. Cancer* 19:105. doi: 10.1186/s12943-020-01224-3

Zhang, Y., Shan, C., Chen, Y., Sun, S., Liu, D., Zhang, X., et al. (2020). CircDENND2A Promotes Non-small Cell Lung Cancer Progression via Regulating MiR-34a/CCNE1 Signaling. *Front. Genet.* 11:987. doi: 10.3389/fgene.2020.00987

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 13 September 2021
doi: 10.3389/fgene.2021.702695

Identification and Elucidation of the Protective isomiRs in Lung Cancer Patient Prognosis

*Fu-Mei Hsieh, Su-Ting Lai, Ming-Fong Wu and Chen-Ching Lin**

Institute of Biomedical Informatics, National Yang Ming Chiao Tung University, Taipei, Taiwan

MicroRNAs (miRNAs) are approximately 20–22 nucleotides in length, which are well known to participate in the post-transcriptional modification. The mature miRNAs were observed to be varied on 5′ or 3′ that raise another term—the isoforms of mature miRNAs (isomiRs), which have been proven not the artifacts and discussed widely recently. In our research, we focused on studying the 5′ isomiRs in lung adenocarcinoma (LUAD) in The Cancer Genome Atlas (TCGA). We characterized 75 isomiRs significantly associated with better prognosis and 43 isomiRs with poor prognosis. The 75 protective isomiRs can successfully distinguish tumors from normal samples and are expressed differently between patients of early and late stages. We also found that most of the protective isomiRs tend to be with downstream shift and upregulated compared with those with upstream shift, implying that a possible selection occurs during cancer development. Among these protective isomiRs, we observed a highly positive and significant correlation, as well as in harmful isomiRs, suggesting cooperation within the group. However, between protective and harmful, there is no such a concordance but conversely more negative correlation, suggesting the possible antagonistic effect between protective and harmful isomiRs. We also identified that two isomiRs miR-181a-3p|-3 and miR-181a-3p|2, respectively, belong to the harmful and protective groups, suggesting a bidirectional regulation of their originated archetype—miR-181a-3p. Additionally, we found that the protective isomiRs of miR-21-5p, which is an oncomiR, may be evolved as the tumor suppressors through producing isomiRs to hinder metastasis. In summary, these results displayed the characteristics of the protective isomiRs and their potential for developing the treatment of lung cancer.

Keywords: miRNA, isomiR, lung adenocarcinoma, survival analysis, miR-181a-3p, miR-21-5p

OPEN ACCESS

Edited by:
Yadong Zheng,
Zhejiang Agriculture and Forestry University, China

Reviewed by:
Kehinde Ross,
Liverpool John Moores University, United Kingdom
Qi Xue,
Peking Union Medical College, China

***Correspondence:**
Chen-Ching Lin
chenching.lin@nycu.edu.tw

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *29 April 2021*
Accepted: *16 August 2021*
Published: *13 September 2021*

Citation:
Hsieh F-M, Lai S-T, Wu M-F and Lin C-C (2021) Identification and Elucidation of the Protective isomiRs in Lung Cancer Patient Prognosis. Front. Genet. 12:702695. doi: 10.3389/fgene.2021.702695

INTRODUCTION

MicroRNAs (miRNAs) are a class of small non-coding RNAs (ncRNA) of approximately 22 nts in length encoded in the genome of plants and animals, which have been found to regulate gene expression in the post-transcriptional level (Lee et al., 1993; Reinhart et al., 2000). MiRNAs were found to bind on the 3′ untranslated region (UTR) of target mRNA, which was called miRNA response element (MRE), and usually repress the target mRNA expression (Lai, 2002). Typically, miRNA is first transcribed to primary miRNA (pri-miRNA) by RNA polymerase II and folded to a hairpin structure. Then, the Drosha-DGCR8 complex binds on the pri-miRNA to prune the stem

and produce miRNA precursor (pre-miRNA), which is approximately 60 nts. The pre-miRNAs are further exported into the cytoplasm. In the cytoplasm, the hairpin loop of pre-miRNA is trimmed by Dicer to generate a paired miRNA duplex, each approximately 22 nts. Either strand of the paired miRNA duplex can further bind to the Argonaute protein (AGO) (Olina et al., 2018), form a silencing complex (RISC), and then regulate, usually repress, the target gene expression (Bartel, 2018).

"Lung cancer has been noticed as one of the leading causes of death in cancers in the United States and around the world" (Dela Cruz et al., 2011). Besides, lung cancer has a poor prognosis and thus is worth noticing its progression in treatment, biomarkers, etc. The most common type is non-small cell lung cancer (NSCLC), which takes 85% of lung cancer. NSCLC can be further classified into three subtypes: adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. In this study, we mainly focused on lung adenocarcinoma (LUAD), which comprises 40% of all lung cancer (Zappa and Mousa, 2016). MiRNAs are involved in many mechanisms in cells, such as immune responses (Sun et al., 2013; Xu et al., 2014) and tumor progression. They can be tumor suppressors and oncogenes (Esquela-Kerscher and Slack, 2006). More importantly, because of their ubiquitous characteristic, miRNAs are widely detected and found to contribute to cancer treatment (Berindan-Neagoe et al., 2014; Shah et al., 2016). Recent studies have provided insight into the role of miRNAs in tumorigenesis and predicted miRNAs as biomarkers in lung cancer (Azizi et al., 2021; Zhong et al., 2021). Additionally, the functions of miRNA in lung cancer have been widely reported (Iqbal et al., 2019; Wu et al., 2019).

With the development of sequencing technology, the mature miRNAs were observed to be varied on 5′ or 3′ that raise the isoforms of mature miRNAs (isomiRs) (Morin et al., 2008; Guo et al., 2016). Notably, these isomiRs were indicated to be mainly generated from imprecise cleavage or shifting by Drosha and Dicer rather than the degradation during sample preparation for sequencing (Lee et al., 2010; Neilsen et al., 2012). Additionally, some isomiRs can be generated by non-template nucleotide addition, such as adenylation at the 3′-end (Lu et al., 2009). Accordingly, the identification of isomiRs extends the functional significance and variance of miRNA regulation (Mercey et al., 2017). IsomiRs can generally be categorized into three types: 3′ isomiR, 5′ isomiR, and polymorphic isomiR. MiRNA can interact with its mRNA target by Watson–Crick pairing typically through the strong connection of "seed region," which is located on 2-7mer (or 8mer) of 5′ miRNA and guided mRNA targeting (Bartel, 2009; Tan et al., 2014). Therefore, the generation of 5′ isomiRs could alternate the seed region and affect the downstream regulation and functional roles of miRNAs. Recently, the isomiRs have been discovered to show different expression levels among disease subtypes, gender differences (Loher et al., 2014; Guo et al., 2016). The existence of isomiRs may also help to distinguish the cancer type (Telonis et al., 2017). Also, isomiRs may play the dominant miRNA in certain tissues or even become the canonical miRNAs through evolution (Tan et al., 2014). The miRNA regulation relies on miRNA–mRNA complementary binding. The presence or even domination of isomiR indicates the concern of nucleotides alternation in miRNA regulation, therefore elevating the urgency of clarifying the variance of miRNA regulation derived from isomiRs.

MATERIALS AND METHODS

Collection and Preprocessing of miRNA and mRNA Expression Profiles

The miRNA expression profiles "isoforms.quantification.txt" were obtained from The Cancer Genome Atlas (TCGA)-GDC portal[1], which recorded the coordinate information and expression level miRNAs and their isomiRs. We initially downloaded 567 LUAD samples. For data cleaning, 87 were removed as they contain "annotation.txt," which may be problematic. Also, three duplicated files from the same participant (TCGA-44-6775) and two "Recurrent Solid Tumor" files were skipped. Finally, 475 LUAD miRNA expression files, which contain 442 primary solid tumors and 33 solid tissue normal, were kept. Meanwhile, the RNA-seq data "htseq.counts" were also downloaded. After data cleaning as the same criteria with miRNA expression profiles, 480 RNA-Seq samples (46 normal, 434 tumor) remained. Finally, we kept a total of 444 RNA-Seq samples (14 normal, 430 tumor) that match the isomiR expression profiles for the analysis of miRNA-regulated target genes.

The Nomenclature of isomiRs

In our research, the isomiRs are defined by the relative 5′ start site of archetype miRNA. Since the miRNA annotation of TCGA data was according to the miRBase V21, and the archetype miRNA coordinate "hsa.gff3" was initially downloaded from miRBase (V21 from http://www.mirbase.org/) (Kozomara and Griffiths-Jones, 2014; Kozomara et al., 2019). The nomenclature of isomiRs was defined by subtracting the 5′ start site coordinates of the archetype miRNA, which are recorded in "hsa.gff3," and each corresponding isomiR, which is recorded in "isoforms.quantification.txt," followed by previous studies (Telonis et al., 2015, 2017). For example, "MIMAT0000062|-2" and "MIMAT0000062|4" are isomiRs of the archetype "MIMAT0000062" with respectively two upstream and four downstream nucleotides shifting of the archetype. Additionally, we denoted the archetypes as the miRNA with zero shift; for example, the archetype of MIMAT0000062 is named as MIMAT0000062|0. Positive or negative strands and location on chromosomes were all considered in different subtraction cases. Herein, we use miRNAs to represent the set of archetype miRNA and isomiRs. Furthermore, the miRBase has been updated to V22 (Kozomara et al., 2019); we then updated the annotation of miRNAs and isomiRs in TCGA dataset for the analysis in this study. The overall workflow of this study is depicted in **Figure 1A**.

Identification of Survival Influential and Differentially Expressed isomiRs

In this study, we applied the Cox regression model to discover survival influential miRNAs in LUAD. The Cox regression model

[1]https://portal.gdc.cancer.gov/

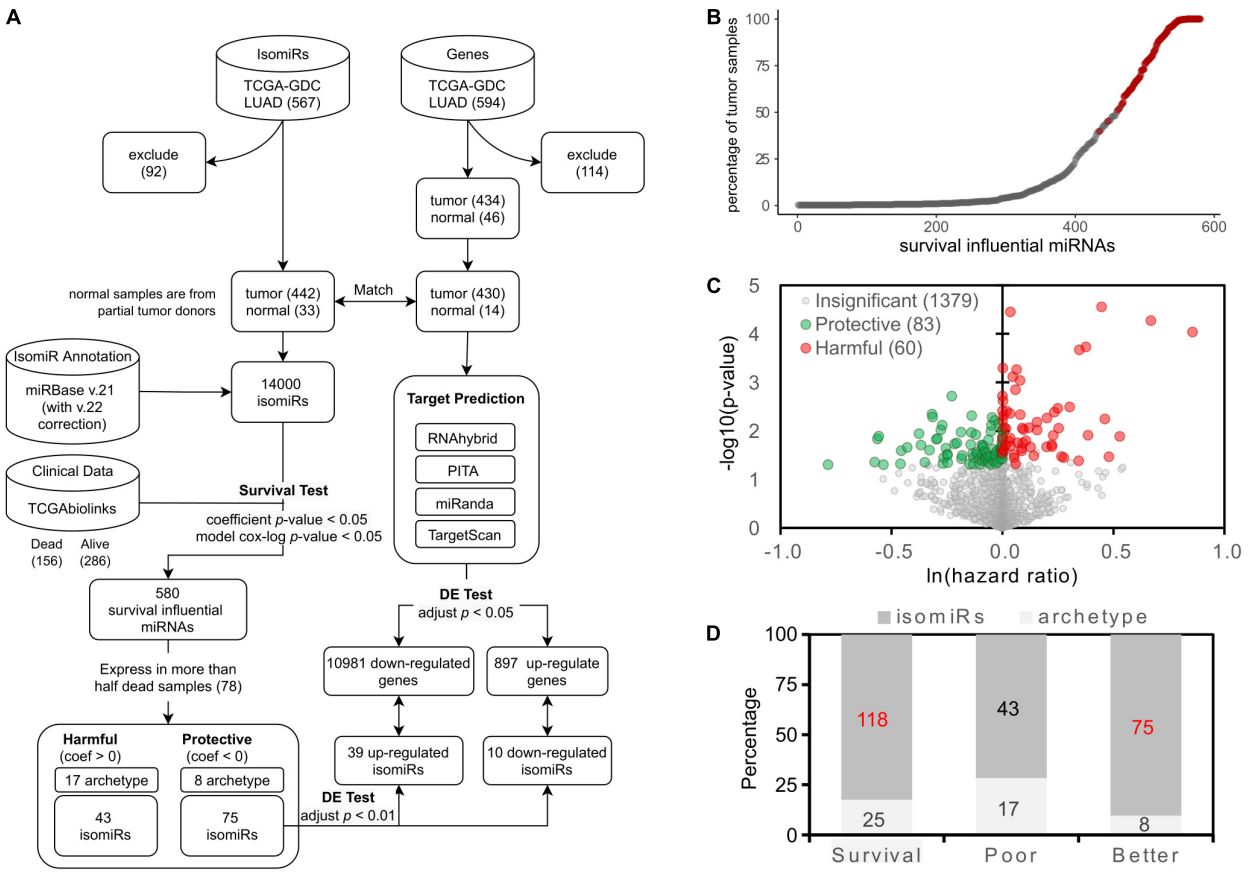


FIGURE 1 | Identification of the survival influential miRNAs. **(A)** The overall workflow of identifying the survival influential miRNAs. **(B)** Percentage of samples in which the survival influential miRNAs are expressed. The *x*-axis lists the miRNAs, which are significantly associated with patients' prognoses, ordered by the number of samples in which they are expressed. The red dots are those survival influential miRNAs expressed in at least 50% of dead samples. **(C)** Volcano plot for survival influential miRNAs. The *x*-axis is the natural log of the hazard ratio of miRNAs; the *y*-axis shows the corresponding significance. The miRNAs with significant influence on patient prognosis are colored by green (protective) and red (harmful). **(D)** The prevalence of survival influential isomiRs. The survival category indicates all the survival influential miRNAs, i.e., contains poor and better prognoses. Numbers labeled red indicate a significantly overrepresented proportion of isomiRs.

was performed using the coxph function of the Survival R package (Therneau, 2015). The clinical data of LUAD patients were retrieved by *TCGAbiolinks* (v.2.18.0) (Mounir et al., 2019) in *Bioconductor* (v3.12) release. We then performed the multivariate Cox regression model—the covariates are the expression level of the tested miRNA and the confounding clinical factors—to assess the influence of a miRNA on patient survival. Then, incorporated confounding factors are "age_at_index," "cigarettes_per_day," and "ajcc_pathologic_t." In this study, we denoted the miRNAs with positive and negative coefficients in the model as harmful and protective to patient survival, respectively. The significance of the coefficient β is determined by comparing with the null model, which hypothesizes that the changes of tested miRNA expression level do not affect patient survival, and tested by a likelihood ratio test and the Wald test (Andersen and Gill, 1982; Therneau et al., 2000). The Wald test examines whether the observed regression coefficient statistically differs from zero and reported a z-score (standard score) and a *p*-value estimated by z-score for the observed regression coefficient. The miRNAs with the coefficient of *p*-value < 0.05 were recognized as the survival influential miRNAs, that is, they have a significant impact on patient survival. Consequently, 580 miRNAs were identified as survival influential miRNAs. However, we observed that approximately 68% identified survival influential miRNAs were expressed in less than 20% of patients (**Figure 1B**). In addition, previous studies (Concato et al., 1995; Peduzzi et al., 1995; Harrell et al., 1996) suggested that 10-20 events per predictor variable (EPV) could be appropriate for the construction of Cox regression model. In our Cox regression model, we used four predictor variables, which are expression level of the tested miRNA, and three confounding factors; thus, the proper number of events (dead samples) could be ranged from 30 to 80. Accordingly, to preserve enough coverage of patients in which the tested miRNA expressed and met the proper EPV, we decided to keep the survival influential miRNAs expressed in more than 78 dead samples, which also can cover approximately 50% samples (**Figure 1B**, red circles). Finally, 143 survival influential miRNAs are used in the following analysis.

Next, we used the *limma* package (v.3.46.0) (Ritchie et al., 2015) and *edgeR* (v.3.32.1) (McCarthy et al., 2012) to identify the

differentially expressed (DE) isomiRs and their mRNA targets. To be consistent with the survival analysis, we used 1,522 isomiRs, which are expressed in more than 50% dead tumor samples and cover at least 41% samples, to perform differential expression analysis. The isomiR expression profiles were normalized by the Upper Quartile method and modeled with *limFit* function and *eBayes* function.

To speculate the possible functions of these isomiRs, we used the RNA-seq data of the LUAD samples as target gene expression profiles. The 3′-UTR sequences of mRNAs were also retrieved to predict the potential isomiRs' binding targets. The RNA-seq data were first normalized with TMM, a method for estimating relative RNA production levels (Robinson and Oshlack, 2010), and using *voom* function and *eBayes* function for modeling. We found 8,557 DE genes (adjusted p-value = 0.01), with 3,484 upregulated and 5,073 downregulated.

Prediction of isomiR-regulated Target Genes

We mainly use R (v4.0.3) and its packages in this research. The *Biopython* (Cock et al., 2009), which is a Python package, was used to get sequences of isomiRs and the transcript sequences. Since we only focused on the 5′ variation of isomiRs, the sequence length to where 3′ should stop was defined by the longest one for each isomiR in the whole LUAD dataset.

With the whole gene expression (444 samples) and isomiRs expression (475 samples), there are four tools used for isomiRs' target prediction, which are: *miRanda* (v.3.3a) (Enright et al., 2003; John et al., 2004), *RNAhybrid* (v.2.1.2) (Rehmsmeier et al., 2004), *PITA* (version 6, Aug-31-08) (Kertesz et al., 2007), and TargetScan (v.7.2) (Agarwal et al., 2015); the predicted isomiR-target pairs numbers were 106,335,587, 106,439,448, 9,231,767, and 77,795,459, respectively. These algorithms follow different principles, such as the seed match, sequence complementary, the secondary structure, and free energy, thus none of them could predict the isomiR-target interactions comprehensively. Thus, we set a stringent cut-off for pairs presented in three of the four prediction tools, where only 13,091,963 pairs were kept. Finally, combined with the DE isomiR, DE gene results, and survival analysis, the protective isomiR-target pairs were classified to 10,981 (39 upregulated isomiR vs. downregulated gene) and 897 (10 dowregulated isomiR vs. upregulated gene).

Functional Analysis of isomiRs in the Protective and Harmful Groups

The isomiR-regulated network for the functional test was selected by Spearman's correlation, estimated by *stats:cor.test*, and visualized by *psych:cor.plo*t. Because there are many zero counts in sample-wise dimension, we compute every isomiR–isomiR correlation by intersecting non-zero value. Thus, not every correlation has the same sample number. Then, the correlation was selected by a positive coefficient (rho $>$ 0) and top 5% by ranking both p-value and rho; also, only the nodes with degree $>$ 1 were selected. For analyzing biological processes of Gene Ontology, we used *clusterProfiler:enrichGO* function (v.3.18.1), where except for the qvalueCutoff (set as 1), the other parameters were set as default. The discovered functional modules were further summarized by REVIGO (Supek et al., 2011) algorithm with a similarity $\geq$ 0.9 that is calculated from the Resnik algorithm (Resnik, 1999) and visualized by CirGO package (Kuznetsova et al., 2019).

RESULTS

Identification of Survival Influential isomiRs in Lung Cancer

In this study, we performed the multivariate Cox regression model to identify survival influential miRNAs. The workflow is depicted in **Figure 1A**. To meet the appropriate EPV (Concato et al., 1995; Peduzzi et al., 1995; Harrell et al., 1996) and preserve the sufficient coverage of patients in which the tested miRNA was expressed (**Figure 1B**), only the identified survival influential miRNAs expressed in more than 50% dead patients (78) were kept for the following analysis. Consequently, 143 miRNAs are identified—60 and 83 miRNAs are harmful and protective, respectively (**Figure 1C**). Notably, 118 out of 143 survival influential miRNAs are isoforms (isomiRs) (**Figure 1D**). This proportion (83%) is significantly overrepresented (p-value = $3.21^{*}10^{-5}$, Fisher's exact test). In other words, the archetype miRNAs are significantly underrepresented in these 143 survival influential miRNAs. Moreover, isomiRs associated with better prognosis of patients are significantly enriched in these 118 survival influential isomiRs (75/118, p-value = $8.39^{*}10^{-7}$, Fisher's exact test). However, both archetype miRNAs and isomiRs are neither significantly overrepresented nor underrepresented in miRNAs associated with poor prognosis of patients. These results suggested that the survival influential isomiRs might tend to be protective and, thus, play roles as tumor suppressors and putative therapeutics for patients with lung cancer.

Characteristics of the Identified Protective isomiRs

To investigate the characteristics of these protective isomiRs, we studied their expression profiles in LUAD patients. **Figure 2A** shows the fold change and adj. p-value of these 75 protective isomiRs. Among the 75 protective isomiRs, 39 and 10 isomiRs are significantly up- and downregulated, respectively. The proportion of upregulated isomiRs (52%) is overrepresented, but the significance is moderate (right-tailed p-value = 0.09, Fisher's exact test), and the proportion of downregulated isomiRs (13%) is close to randomly expected (right-tailed p-value = 0.44, Fisher's exact test). This observation suggested that the identified protective isomiRs might tend to be upregulated in lung cancer. Additionally, we observed that the 75 protective isomiRs can separate tumor from normal samples significantly in the principal component analysis (PCA) (**Figure 2B**), but the expression profiles of completely 14,000 isomiRs and 1,522 selected isomiRs cannot (**Supplementary Figure 1**). Besides, using the principal components (PC1 and PC2) of 75 protective isomiRs with logistic regression model

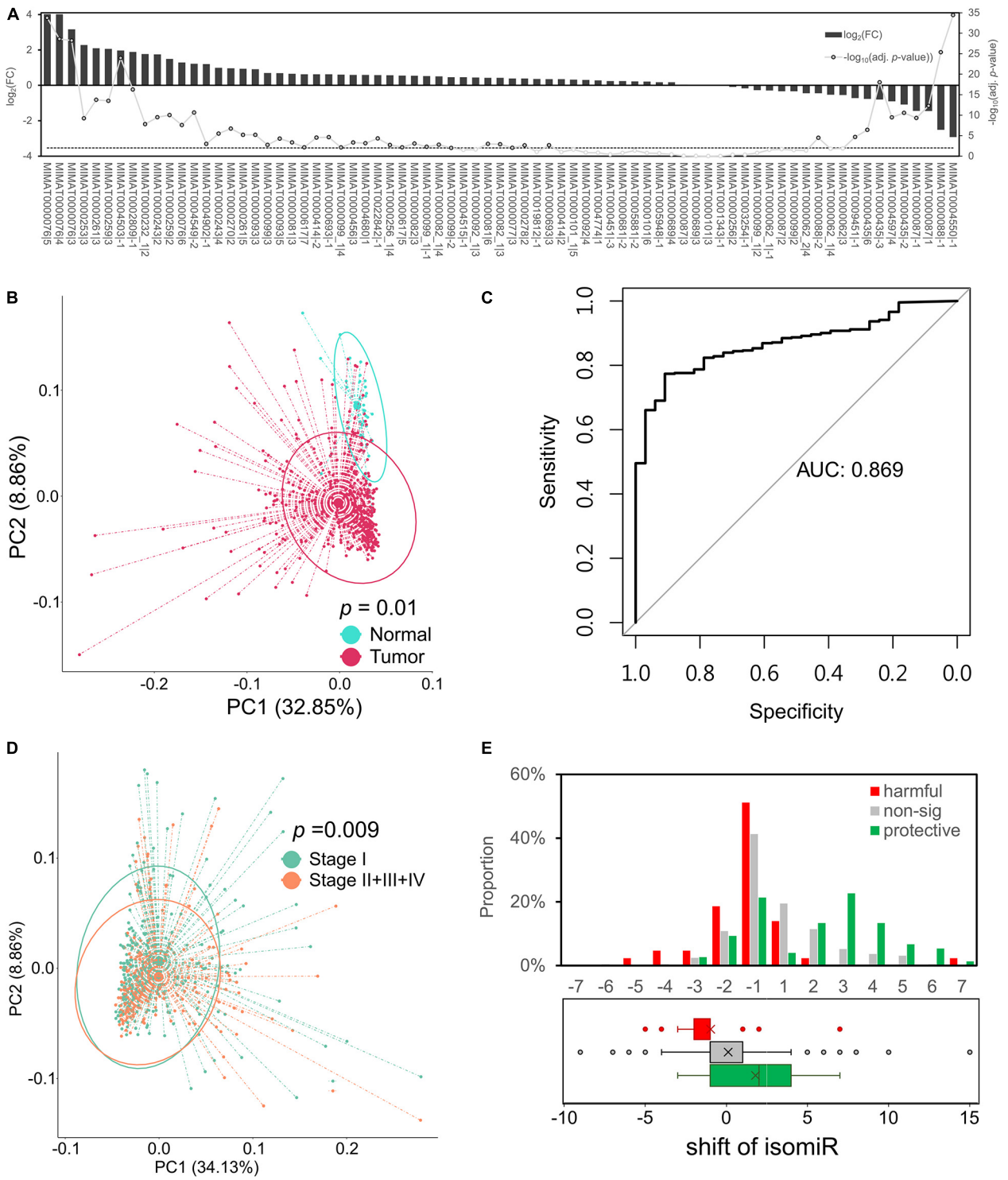


FIGURE 2 | Protective isomiRs are capable to separate LUAD samples. **(A)** The differential expression pattern of the 75 identified protective isomiRs. The bars (related to the left *y*-axis) show the differential expression level of the corresponding isomiR at the *x*-axis; the line chart indicates the significance level (related to the right *y*-axis). The dashed line marks the significance level with *p*-value < 0.01. **(B)** Principal component analysis (PCA) using the 75 protective isomiRs classifies tumor and normal samples. The *p*-value is derived from permutational analysis of variance (PERMANOVA). **(C)** The ROC curve of 75 protective isomiRs by its log-transformed PC1 and PC2 to predict the total tumor samples. **(D)** PCA using the 75 protective isomiRs classifies patients with early and late stage tumors. The *p*-value is derived by PERMANOVA. **(E)** The shift distribution of isomiRs. The upper panel shows the shift distribution of protective, harmful, and insignificant isomiRs. The lower panel displays the tendency of the shift preference of the three categories of isomiRs by a boxplot.

(by *glm* function in R) to predict the samples, the area under curve (AUC) reaches 0.86 (**Figure 2C**). These results suggested that the 75 protective isomiRs can classify the LUAD samples. The expression profiles of the 75 protective isomiRs are also distinguishable between patients of the early (I) and late (II + III + IV) stage (**Figure 2D**). The main difference between early and late stages is metastasis or not. Therefore, these 75 protective isomiRs might be involved in the regulation of tumor metastasis. Among these 75 protective isomiRs, 50 are positively shifted (downstream relative to archetype), and the remaining 25 are negatively shifted (upstream relative to archetype). Interestingly, we found that the protective isomiRs tend to possess positive shifts (p-value = $8.39*10^{-5}$, Fisher's exact test) and the harmful isomiRs negative shifts (p-value = $4.38*10^{-4}$, Fisher's exact test) (**Figure 2E**). Moreover, we observed that the protective isomiRs with downstream shift tend to be more upregulated than those with upstream shift. The average fold changes of upstream and downstream isomiRs are 1.56 and 1.01, respectively (p-value = 0.0564, Wilcoxon rank sum test, **Supplementary Figure 2**). Accordingly, these above results imply that the protective isomiRs with downstream shift might be selected to express against tumorigenesis during cancer development.

Positive Correlation Between the Protective isomiRs

In this study, we observed the prevalent positive correlation of expression profiles between the protective or harmful isomiRs—within-group (**Figures 3A,B**). Additionally, among the protective isomiRs, approximately 80% of pairs are significantly and positively correlated (z-score of Spearman's correlation coefficient > 2). This proportion is significantly overrepresented (p-value $< 10^{-271}$, Fisher's exact test). Among the harmful isomiRs, approximately 44% of pairs are significantly and positively correlated. This proportion is overrepresented, but the significance is moderate (p-value = 0.16, Fisher's exact test). These observations suggest the possible cooperativity between isomiRs with a similar influence on patient prognosis. Furthermore, the cooperativity between protective isomiRs might be stronger than that between harmful isomiRs. On the other hand, we found that the significant anti-correlations among either protective or harmful isomiRs are significantly underrepresented (protective: 2%, p-value = $5.68*10^{-40}$; harmful: 3%, p-value = $1.1*10^{-6}$, Fisher's exact). This observation further suggests that the isomiRs with similar influence on patient prognosis might not tend to be against each other.

On the other hand, we did not observe the concordance expression correlation between protective and harmful isomiRs (**Figure 3C**), suggesting that the cooperativities between protective and harmful isomiRs might not be prevalent. Furthermore, we found that the significantly negative correlations of expression profiles between protective and harmful isomiRs are significantly overrepresented (p-value = $1.38*10^{-5}$, Fisher's exact test), demonstrating the potential antagonistic relationship between protective and harmful isomiRs.

To further investigate the cooperative regulation of activated protective isomiRs, we focused on significantly 39 upregulated protective isomiRs. The functional enrichment analysis for the significantly downregulated target genes of these 39 protective isomiRs is shown in **Figure 3D**. We found that these target genes are significantly overrepresented in the cell migration-related functions, suggesting their roles in tumor metastasis. This observation further recapitulates the protective isomiRs' characteristic that can classify patients with early stage tumors from late stage very well. More importantly, the above results imply that the identified upregulated protective isomiRs might repress the genes involved in tumor metastasis to benefit patient prognosis.

The Bidirectional Regulation of Patient Prognosis

The interaction between isomiRs and their corresponding archetype miRNAs has been observed in lung cancer (Chan et al., 2013; Liang et al., 2020). More specifically, the isomiRs may disturb the classical regulatory network of their corresponding archetype miRNAs, even cause conflicting drug responses in cancers. In this study, we noticed, among the identified survival influential isomiRs, that MIMAT0000270|2 (hsa-miR-181a-3p|2) and MIMAT0000270|-3 (hsa-miR-181a-3p|-3) are originated from the same archetype—MIMAT0000270 (hsa-miR-181a-3p), but hsa-miR-181a-3p|2 is protective and hsa-miR-181a-3p|-3 is harmful. Interestingly, a previous study has observed that the hsa-miR-181a could function as an oncomiR or tumor suppressor in acute myeloid leukemia (Qiang et al., 2020). These observations suggest that hsa-miR-181a-3p could possess bidirectional regulation to patient prognosis. Moreover, the proportion of overlapped targets between these two isomiRs is small (0.02, Jaccard Index), emphasizing that these two isomiRs might regulate distinct downstream pathways. Additionally, these two isomiRs are upregulated in tumors (**Figure 4A**), suggesting that their regulatory downstream pathways might be repressed. We then performed functional enrichment analysis on their downregulated target genes separately (**Figures 4B,C**). We observed that the downregulated target genes of the protective hsa-miR-181a-3p|2 were significantly enriched in muscle cell or myotube functions (**Figure 4B**); the downregulated target genes of the harmful hsa-miR-181a-3p|-3 tend to be involved in fat cell differentiation and positive regulation of neuronal differentiation (**Figure 4C**). Previous studies have reported that muscle loss could be a significant predictor of patient mortality (Shiroyama et al., 2019; Rosa-Caldwell et al., 2020), supporting the protective role of hsa-miR-181a-3p|2 in lung cancer. In addition, we observed that the downregulated target genes of these two isomiRs are significantly enriched in the epithelial–mesenchymal transition (EMT)-associated functions (**Figures 4B,C**, marked by red). EMT is a reversible cell transition process, where cells could progressively lose the polarity, cell–cell junction, fixation, and finally turn into the mesenchymal morphology (Dongre and Weinberg, 2019). Evidence has shown that EMT was an epigenetic process, independent of DNA sequence alterations (Tam and Weinberg, 2013), and could

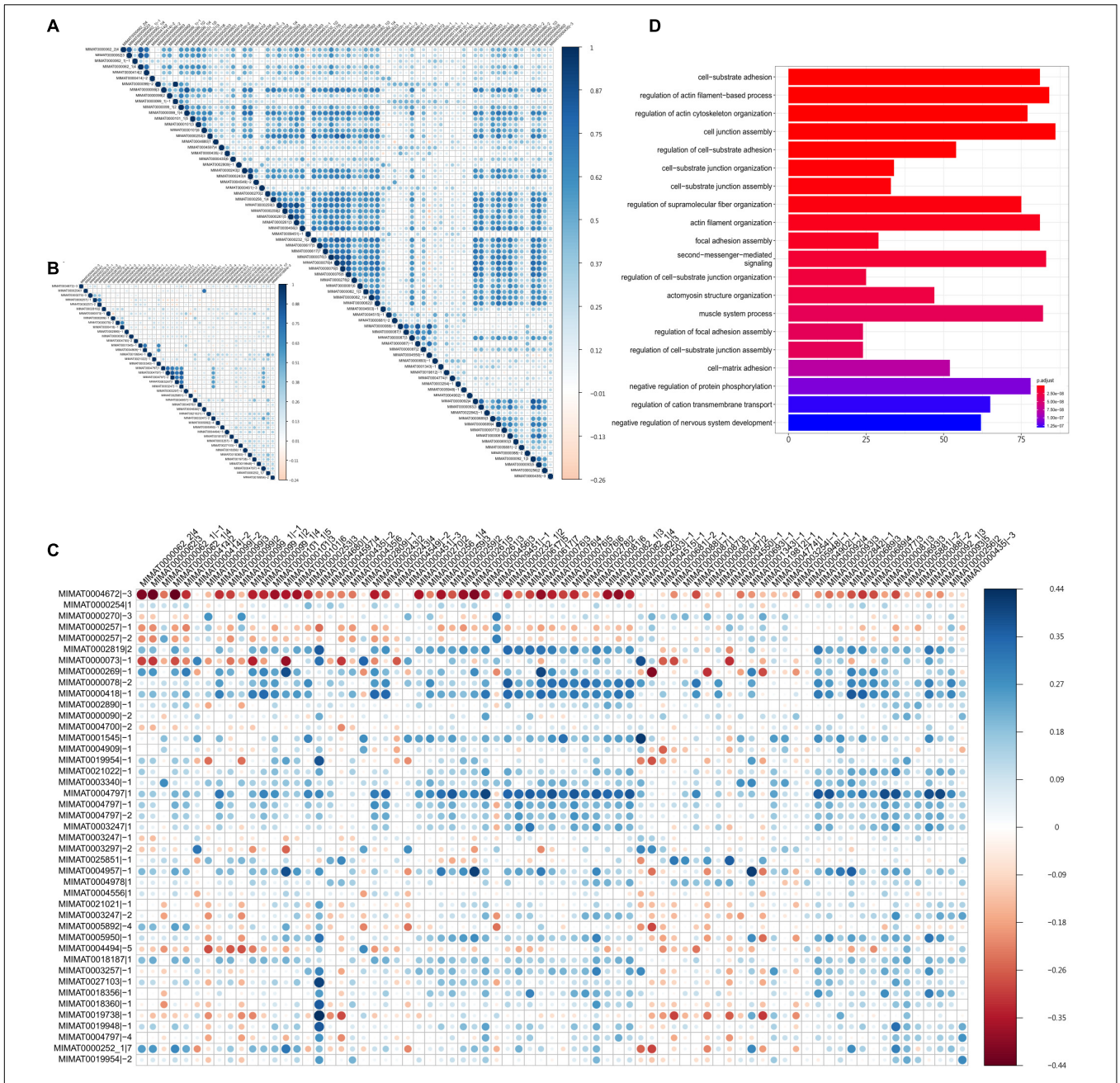


FIGURE 3 | The cooperativity between survival influential isomiRs. The Spearman's correlation between **(A)** the 75 protective isomiRs, **(B)** the 43 harmful isomiRs, and **(C)** the 75 protective and the 43 harmful isomiRs. **(D)** The top 20 significantly enriched functions regulated by the 39 highly co-expressed and upregulated protective isomiRs.

be significant during neoplasia, contributing to the malignant progression (Mani et al., 2008; Singh and Settleman, 2010), as well as in lung cancer (O'Leary et al., 2018). For example, the harmful isomiR—hsa-miR-181a-3p|-3—may regulate "regulation of actin cytoskeleton organization" and "protein localization to plasma membrane," pointing out that the dysregulation of its targeted genes might lose the function in fixation. On the other hand, the protective isomiR—hsa-miR-181a-3p|2—may also take part in EMT. The hsa-miR-181a-3p|2 may repress the function of the glucocorticoid metabolic process that could promote tumor cell invasion and lung metastasis (Shi et al., 2019). The other functions, such as "regulation of establishment of cell polarity" and "skeletal muscle fiber development," may also indicate the role. The above observations suggest the potential bidirectional regulation of the hsa-miR-181a-3p. That is, these two isomiRs are derived from the same arm of pre-miRNA but demonstrate the opposite way in regulating tumor metastasis, and further patient prognosis.

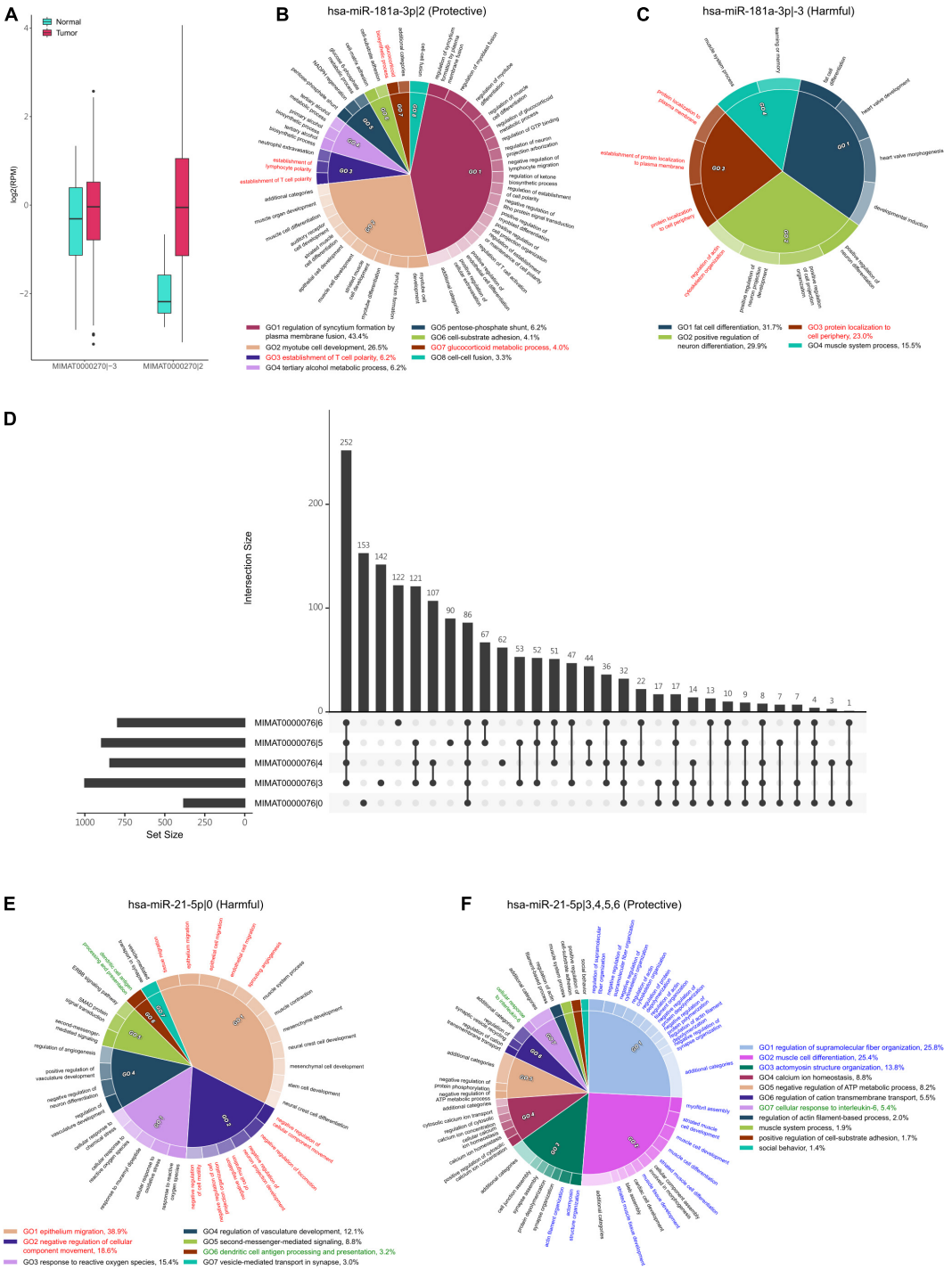


FIGURE 4 | The regulatory function of hsa-miR-181a-3p|-3 and hsa-miR-181a-3p|2. **(A)** The expression of (hsa-miR-181a-3p|-3) and (hsa-miR-181a-3p|2) in tumor and normal samples. Both are upregulated in tumor samples. **(B,C)** The significantly enriched functions in which the downregulated target genes of hsa-miR-181a-3p|2 **(B)** and hsa-miR-181a-3p|-3 **(C)** are involved. The percentage is the relative significance of one functional module to others. The functional modules associated with EMT are marked by red. **(D)** The intersection of downregulated target genes between the archetype of hsa-miR-21-5p and its four protective isomiRs. The horizontal bars on the left side show the numbers of downregulated target genes of corresponding isomiRs of hsa-miR-21-5p, while the vertical bars show the intersection number. The dot represents whether the set is part of the intersection. **(E,F)** The significantly enriched functions in which the downregulated target genes of hsa-miR-21-5p|0 **(E)** and hsa-miR-21-5p|3, 4, 5, and 6 **(F)** are involved. The percentage is the relative significance of one functional module to others. The functional modules associated with metastasis, regulation of immune response, and regulation of muscle development are marked by red, green, and blue, respectively.

In addition, miR-21 was found to be an oncogene in several cancers (Bica-Pop et al., 2018). In NSCLC, it was upregulated in patients (Yu et al., 2010; Yang et al., 2015), which also resembles our study. We observed that the archetype hsa-miR-21-5p (MIMAT0000076|0) is harmful to patients' prognoses. However, the hsa-miR-21-5p|3, hsa-miR-21-5p|4, hsa-miR-21-5p|5, and hsa-miR-21-5p|6 are protective. Moreover, these four isomiRs are the top four upregulated isomiRs in the identified protective miRNAs. Interestingly, we found that the protective isomiRs, hsa-miR-21-5p|3, hsa-miR-21-5p|4, hsa-miR-21-5p|5, and hsa-miR-21-5p|6, have the highest intersection of target genes, while the second-highest target gene set was exclusively presented in the archetype hsa-miR-21-5p|0 (**Figure 4D**). This observation suggests the possible cooperativity between these four isomiRs and divergent downstream regulation from the archetype—hsa-miR-21-5p. As expected, we observed that the archetype hsa-miR-21-5p|0 is largely involved in regulating tumor metastasis-associated functions (**Figure 4E**, marked as red). More specifically, the downregulated target genes of hsa-miR-21-5p|0 are significantly enriched in the negative regulation of cell migration and the negative regulation of cell motility. This observation suggests that the hsa-miR-21-5p|0 may promote tumor metastasis by repressing the negative regulation of cell migration or motility. Moreover, the hsa-miR-21-5p|0 might repress the function of dendritic cell antigen processing and presentation to help tumor progression (**Figure 4E**, marked as green). Interestingly, the intersection of downregulated target genes between the four isomiRs, hsa-miR-21-5p|3, 4, 5, and 6, is largely significantly enriched in the function of regulation of muscle development (**Figure 4F**, marked as blue). This observation shows the protective roles of these four isomiRs that might co-regulate the muscle development to control muscle loss and prolong patient survival (Shiroyama et al., 2019; Rosa-Caldwell et al., 2020). In addition, we found that these four isomiRs regulate the cellular response to interleukin-6 (IL-6) (**Figure 4F**, marked as green). Previous studies have demonstrated the IL-6 can promote lung cancer metastasis (Gross et al., 2018; Jiang et al., 2018; Liu et al., 2020). Accordingly, these four isomiRs might prevent the patients from tumor metastasis by repressing the function of IL-6. Briefly, these above results provide a possible insight in bidirectional regulation between archetype miRNA and its isomiRs: these four isomiRs might be against the regulation of archetype hsa-miR-21-5p to repress tumor metastasis.

DISCUSSION

Herein, we performed a standard pipeline to identify the survival influential miRNAs to patient prognosis. We found that the isomiRs are significantly enriched in the identified survival influential miRNAs, especially for the isomiRs associated with the better prognosis of patients, that is, they are protective. These protective isomiRs might be potential therapeutics for lung cancer. Previous studies have reported that hsa-miR-21-5p was upregulated consistently across cancers (Ferracin et al., 2015) and recognized as oncomiR in lung cancer (Li and Wu, 2018; Yuan et al., 2018). Interestingly, in this study, the isomiRs originated from hsa-miR-21-5p are upregulated and associated with a better prognosis of patients: they might be tumor suppressors. IsomiRs possess different seed regions from the archetypes, suggesting that the downstream regulation could vary from an archetype to its isomiRs. Accordingly, the possible scenario is that these four isomiRs (hsa-miR-21-5p|3, hsa-miR-21-5p|4, hsa-miR-21-5p|5, and hsa-miR-21-5p|6) might evolve tumor suppressor function during cancer development. Additionally, we noticed that the protective isomiRs tend to be provided with downstream shift. Moreover, these protective isomiRs with downstream shift are more likely to be upregulated than those with upstream shift. These results suggest that the isomiRs with downstream shift might be selected during carcinogenesis to prolong patient mortality. However, further experiments are required to validate this hypothesis.

On the other hand, we observed strong cooperative interactions within isomiRs with the same prognostic effect to patients, but antagonistic interactions between protective and harmful isomiRs. Furthermore, this cooperativity can be observed from the isomiRs originated from the same precursor and shared a large proportion of regulatory target genes. For example, hsa-miR-181b-5p|−1 and −2 are both harmful isomiRs, originate from hsa-miR-181b-5p, and share significantly overrepresented target genes (Jaccard Index = 0.45; p-value < 0.001, Fisher's exact test); hsa-miR-182-5p|2 and 3 are both protective isomiRs, originate from hsa-miR-182-5p, and share significantly overrepresented target genes (Jaccard Index = 0.4; p-value < 0.001, Fisher's exact test). On the other hand, the antagonistic interaction can also be observed from the small shared proportion of target genes between isomiRs, for example, hsa-miR-181a-3p|2 and hsa-miR-181a-3p|-3. This observation might further imply the synergistic effect between the isomiRs exerting a similar prognostic influence on patients. That is, the combination of multiple survival influential isomiRs could benefit or worsen patient prognosis more.

In addition, we identified that hsa-miR-21-5p|0 and its four isomiRs—hsa-miR-21-5p|3, 4, 5, and 6—could play critical roles in regulating tumor metastasis. More importantly, these four isomiRs are associated with a better prognosis of patients and might be against the harmful archetype: hsa-miR-21-5p. Among the downregulated target genes of these four protective isomiRs, *HER2*, *AKT*, and *DDR2* have been observed to be aberrantly expressed in NSCLC (Tsakonas and Ekman, 2018). Moreover, *DDR2* (Discoidin domain receptor tyrosine kinase 2) plays an important role in cellular connectivity, survival, migration, and cell proliferation (Bargal et al., 2009; Fathi et al., 2018), which contributed to therapeutic target in squamous cell lung cancer (Hammerman et al., 2011). Interestingly, we found that *DDR2* is the common downregulated target gene of hsa-miR-21-5p|3, 4, and 5, but not the archetype. This exclusive regulation of oncogene might support the proposed scenario, and these four isomiRs are against the regulation of archetype to repress cancer metastasis.

In summary, we identified survival influential miRNAs for lung cancer patients. We noticed that isomiRs are more associated with patients' mortality, especially for a better

prognosis than archetype miRNAs. Moreover, the expression profiles of the protective isomiR set can be a good predictor for classifying tumor and normal samples, as well as early and late stage patients. On the other hand, we observed the cooperativity between isomiRs with similar prognostic effects to patients and possible antagonistic interaction between protective and harmful isomiRs. In the end, we gave two examples—hsa-miR-181-3p and hsa-miR-21-5p—to demonstrate that the production of isomiRs could exert a distinct downstream regulatory effect, even opposite phenotypic change: better and poor prognoses.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The LUAD isoform miRNA expression profile as well as transcriptome expression profile ("isoforms.quantification.txt", "htseq.counts") for this study can be found in the TCGA-GDC porta (https://portal.gdc.cancer.gov/). The isomiR nomenclature was followed by miRBase (http://www.mirbase.org/), where the coordinate archetype miRNA ("hsa.gff3") downloaded from.

AUTHOR CONTRIBUTIONS

C-CL conceived, designed, and coordinated the study and revised the manuscript. F-MH, S-TL, M-FW, and C-CL implemented the computational method and carried out the analysis. F-MH and C-CL drafted the manuscript. All authors read and approved the final manuscript.

FUNDING

This research was funded by the Ministry of Science and Technology in Taiwan (MOST 107-2221-E-010-016-MY2).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.702695/full#supplementary-material

REFERENCES

Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4:e05005. doi: 10.7554/eLife.05005

Andersen, P. K., and Gill, R. D. (1982). Cox's Regression Model for Counting Processes: a Large Sample Study. *Ann. Statist.* 10, 1100–1120. doi: 10.1214/aos/1176345976

Azizi, M. I. H. N., Othman, I., and Naidu, R. (2021). The Role of MicroRNAs in Lung Cancer Metabolism. *Cancers* 13:1716. doi: 10.3390/cancers13071716

Bargal, R., Cormier-Daire, V., Ben-Neriah, Z., Le Merrer, M., Sosna, J., Melki, J., et al. (2009). Mutations in DDR2 gene cause SMED with short limbs and abnormal calcifications. *Am. J. Hum. Genet.* 84, 80–84. doi: 10.1016/j.ajhg.2008.12.004

Bartel, D. P. (2009). MicroRNAs: target Recognition and Regulatory Functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002

Bartel, D. P. (2018). Metazoan MicroRNAs. *Cell* 173, 20–51. doi: 10.1016/j.cell.2018.03.006

Berindan-Neagoe, I., Monroig Pdel, C., Pasculli, B., and Calin, G. A. (2014). MicroRNAome genome: a treasure for cancer diagnosis and therapy. *CA Cancer J. Clin.* 64, 311–336. doi: 10.3322/caac.21244

Bica-Pop, C., Cojocneanu-Petric, R., Magdo, L., Raduly, L., Gulei, D., and Berindan-Neagoe, I. (2018). Overview upon miR-21 in lung cancer: focus on NSCLC. *Cell. Mol. Life Sci.* 75, 3539–3551. doi: 10.1007/s00018-018-2877-x

Chan, Y. T., Lin, Y. C., Lin, R. J., Kuo, H. H., Thang, W. C., Chiu, K. P., et al. (2013). Concordant and discordant regulation of target genes by miR-31 and its isoforms. *PLoS One* 8:e58169. doi: 10.1371/journal.pone.0058169

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423. doi: 10.1093/bioinformatics/btp163

Concato, J., Peduzzi, P., Holford, T. R., and Feinstein, A. R. (1995). Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. *J. Clin. Epidemiol.* 48, 1495–1501. doi: 10.1016/0895-4356(95)00510-2

Dela Cruz, C. S., Tanoue, L. T., and Matthay, R. A. (2011). Lung cancer: epidemiology, etiology, and prevention. *Clin. Chest Med.* 32, 605–644. doi: 10.1016/j.ccm.2011.09.001

Dongre, A., and Weinberg, R. A. (2019). New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* 20, 69–84. doi: 10.1038/s41580-018-0080-4

Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D. S. (2003). MicroRNA targets in Drosophila. *Genome Biol.* 5:R1. doi: 10.1186/gb-2003-5-1-r1

Esquela-Kerscher, A., and Slack, F. J. (2006). Oncomirs — microRNAs with a role in cancer. *Nat. Rev. Cancer* 6, 259–269. doi: 10.1038/nrc1840

Fathi, Z., Mousavi, S. A. J., Roudi, R., and Ghazi, F. (2018). Distribution of KRAS, DDR2, and TP53 gene mutations in lung cancer: an analysis of Iranian patients. *PLoS One* 13:e0200633. doi: 10.1371/journal.pone.0200633

Ferracin, M., Lupini, L., Salamon, I., Saccenti, E., Zanzi, M. V., Rocchi, A., et al. (2015). Absolute quantification of cell-free microRNAs in cancer patients. *Oncotarget* 6, 14545–14555. doi: 10.18632/oncotarget.3859

Gross, A. C., Cam, H., Phelps, D. A., Saraf, A. J., Bid, H. K., Cam, M., et al. (2018). IL-6 and CXCL8 mediate osteosarcoma-lung interactions critical to metastasis. *JCI Insight* 3:e99791. doi: 10.1172/jci.insight.99791

Guo, L., Liang, T., Yu, J., and Zou, Q. (2016). A Comprehensive Analysis of miRNA/isomiR Expression with Gender Difference. *PLoS One* 11:e0154955. doi: 10.1371/journal.pone.0154955

Hammerman, P. S., Sos, M. L., Ramos, A. H., Xu, C., Dutt, A., Zhou, W., et al. (2011). Mutations in the DDR2 kinase gene identify a novel therapeutic target in squamous cell lung cancer. *Cancer Discov.* 1, 78–89. doi: 10.1158/2159-8274.CD-11-0005

Harrell, F. E. Jr., Lee, K. L., and Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* 15, 361–387. doi: 10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4

Iqbal, M. A., Arora, S., Prakasam, G., Calin, G. A., and Syed, M. A. (2019). MicroRNA in lung cancer: role, mechanisms, pathways and therapeutic relevance. *Mol. Aspects Med.* 70, 3–20. doi: 10.1016/j.mam.2018.07.003

Jiang, M., Wang, Y., Zhang, H., Ji, Y., Zhao, P., Sun, R., et al. (2018). IL-37 inhibits invasion and metastasis in non-small cell lung cancer by suppressing the IL-6/STAT3 signaling pathway. *Thorac. Cancer* 9, 621–629. doi: 10.1111/1759-7714.12628

John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., and Marks, D. S. (2004). Human MicroRNA targets. *PLoS Biol.* 2:e363. doi: 10.1371/journal.pbio.0020363

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat. Genet.* 39, 1278–1284. doi: 10.1038/ng2135

Kozomara, A., Birgaoanu, M., and Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. *Nucl. Acids Res.* 47, D155–D162. doi: 10.1093/nar/gky1141

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucl. Acids Res.* 42, D68–D73. doi: 10.1093/nar/gkt1181

Kuznetsova, I., Lugmayr, A., Siira, S. J., Rackham, O., and Filipovska, A. (2019). CirGO: an alternative circular way of visualising gene ontology terms. *BMC Bioinform.* 20:84. doi: 10.1186/s12859-019-2671-2

Lai, E. C. (2002). Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat. Genet.* 30, 363–364. doi: 10.1038/ng865

Lee, L. W., Zhang, S., Etheridge, A., Ma, L., Martin, D., Galas, D., et al. (2010). Complexity of the microRNA repertoire revealed by next-generation sequencing. *RNA* 16, 2170–2180. doi: 10.1261/rna.2225110

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843–854. doi: 10.1016/0092-8674(93)90529-y

Li, X., and Wu, X. (2018). MiR-21-5p promotes the progression of non-small-cell lung cancer by regulating the expression of SMAD7. *Onco. Targets Ther.* 11, 8445–8454. doi: 10.2147/ott.S172393

Liang, T., Han, L., and Guo, L. (2020). Rewired functional regulatory networks among miRNA isoforms (isomiRs) from let-7 and miR-10 gene families in cancer. *Comput. Struct. Biotechnol. J.* 18, 1238–1248. doi: 10.1016/j.csbj.2020.05.001

Liu, W., Wang, H., Bai, F., Ding, L., Huang, Y., Lu, C., et al. (2020). IL-6 promotes metastasis of non-small-cell lung cancer by up-regulating TIM-4 via NF-kappaB. *Cell Prolif.* 53:e12776. doi: 10.1111/cpr.12776

Loher, P., Londin, E. R., and Rigoutsos, I. (2014). IsomiR expression profiles in human lymphoblastoid cell lines exhibit population and gender dependencies. *Oncotarget* 5, 8790–8802. doi: 10.18632/oncotarget.2405

Lu, S., Sun, Y.-H., and Chiang, V. L. (2009). Adenylation of plant miRNAs. *Nucl. Acids Res.* 37, 1878–1885. doi: 10.1093/nar/gkp031

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715. doi: 10.1016/j.cell.2008.03.027

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. doi: 10.1093/nar/gks042

Mercey, O., Popa, A., Cavard, A., Paquet, A., Chevalier, B., Pons, N., et al. (2017). Characterizing isomiR variants within the microRNA-34/449 family. *FEBS Lett.* 591, 693–705. doi: 10.1002/1873-3468.12595

Morin, R. D., O'Connor, M. D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A. L., et al. (2008). Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* 18, 610–621. doi: 10.1101/gr.7179508

Mounir, M., Lucchetta, M., Silva, T. C., Olsen, C., Bontempi, G., Chen, X., et al. (2019). New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Comput. Biol.* 15:e1006701. doi: 10.1371/journal.pcbi.1006701

Neilsen, C. T., Goodall, G. J., and Bracken, C. P. (2012). IsomiRs–the overlooked repertoire in the dynamic microRNAome. *Trends Genet.* 28, 544–549. doi: 10.1016/j.tig.2012.07.005

O'Leary, K., Shia, A., and Schmid, P. (2018). Epigenetic Regulation of EMT in Non-Small Cell Lung Cancer. *Curr. Cancer Drug Targets* 18, 89–96. doi: 10.2174/1568009617666170203162556

Olina, A. V., Kulbachinskiy, A. V., Aravin, A. A., and Esyunina, D. M. (2018). Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes. *Biochemistry* 83, 483–497. doi: 10.1134/s0006297918050024

Peduzzi, P., Concato, J., Feinstein, A. R., and Holford, T. R. (1995). Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J. Clin. Epidemiol.* 48, 1503–1510. doi: 10.1016/0895-4356(95)00048-8

Qiang, P., Pan, Q., Fang, C., Fozza, C., Song, K., Dai, Y., et al. (2020). MicroRNA-181a-3p as a diagnostic and prognostic biomarker for acute myeloid leukemia. *Mediterr. J. Hematol. Infect. Dis.* 12:e2020012. doi: 10.4084/MJHID.2020.012

Rehmsmeier, M., Steffen, P., Hochsmann, M., and Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. *Rna* 10, 1507–1517. doi: 10.1261/rna.5248604

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* 403, 901–906. doi: 10.1038/35002607

Resnik, P. (1999). Semantic similarity in a taxonomy: an information-based measure and its application to problems of ambiguity in natural language. *J. Artif. Int. Res.* 11, 95–130. doi: 10.1613/jair.514

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11:R25. doi: 10.1186/gb-2010-11-3-r25

Rosa-Caldwell, M. E., Fix, D. K., Washington, T. A., and Greene, N. P. (2020). Muscle alterations in the development and progression of cancer-induced muscle atrophy: a review. *J. Appl. Physiol.* 128, 25–41. doi: 10.1152/japplphysiol.00622.2019

Shah, M. Y., Ferrajoli, A., Sood, A. K., Lopez-Berestein, G., and Calin, G. A. (2016). microRNA Therapeutics in Cancer - An Emerging Concept. *Ebiomedicine* 12, 34–42. doi: 10.1016/j.ebiom.2016.09.017

Shi, W., Wang, D., Yuan, X., Liu, Y., Guo, X., Li, J., et al. (2019). Glucocorticoid receptor-IRS-1 axis controls EMT and the metastasis of breast cancers. *J. Mol. Cell Biol.* 11, 1042–1055. doi: 10.1093/jmcb/mjz001

Shiroyama, T., Nagatomo, I., Koyama, S., Hirata, H., Nishida, S., Miyake, K., et al. (2019). Impact of sarcopenia in patients with advanced non-small cell lung cancer treated with PD-1 inhibitors: a preliminary retrospective study. *Sci. Rep.* 9:2447. doi: 10.1038/s41598-019-39120-6

Singh, A., and Settleman, J. (2010). EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* 29, 4741–4751. doi: 10.1038/onc.2010.215

Sun, Y. M., Lin, K. Y., and Chen, Y. Q. (2013). Diverse functions of miR-125 family in different cell contexts. *J. Hematol. Oncol.* 6:6. doi: 10.1186/1756-8722-6-6

Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6:e21800. doi: 10.1371/journal.pone.0021800

Tam, W. L., and Weinberg, R. A. (2013). The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat. Med.* 19, 1438–1449. doi: 10.1038/nm.3336

Tan, G. C., Chan, E., Molnar, A., Sarkar, R., Alexieva, D., Isa, I. M., et al. (2014). 5′ isomiR variation is of functional and evolutionary importance. *Nucleic Acids Res.* 42, 9424–9435. doi: 10.1093/nar/gku656

Telonis, A. G., Loher, P., Jing, Y., Londin, E., and Rigoutsos, I. (2015). Beyond the one-locus-one-miRNA paradigm: microRNA isoforms enable deeper insights into breast cancer heterogeneity. *Nucleic Acids Res.* 43, 9158–9175. doi: 10.1093/nar/gkv922

Telonis, A. G., Magee, R., Loher, P., Chervoneva, I., Londin, E., and Rigoutsos, I. (2017). Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. *Nucleic Acids Res.* 45, 2973–2985. doi: 10.1093/nar/gkx082

Therneau, T. (2015). *A Package for Survival Analysis in S. version 2*. 38. Available online at: https://CRAN.R-project.org/package=survival (accessed November 17, 2020).

Therneau, T. M., Grambsch, and Patricia, M. (2000). *Modeling Survival Data: extending the Cox Model*. New York: Springer-Verlag.

Tsakonas, G., and Ekman, S. (2018). Oncogene-addicted non-small cell lung cancer and immunotherapy. *J. Thorac. Dis.* 10, S1547–S1555. doi: 10.21037/jtd.2018.01.82

Wu, K.-L., Tsai, Y.-M., Lien, C.-T., Kuo, P.-L., Hung, J. L., and Jen, Y. (2019). The Roles of MicroRNA in Lung Cancer. *Int. J. Mol. Sci.* 20:1611. doi: 10.3390/ijms20071611

Xu, L. L., Shi, C. M., Xu, G. F., Chen, L., Zhu, L. L., Zhu, L., et al. (2014). TNF-α, IL-6, and leptin increase the expression of miR-378, an adipogenesis-related microRNA in human adipocytes. *Cell Biochem. Biophys.* 70, 771–776. doi: 10.1007/s12013-014-9980-x

Yang, J. S., Li, B. J., Lu, H. W., Chen, Y., Lu, C., Zhu, R. X., et al. (2015). Serum miR-152, miR-148a, miR-148b, and miR-21 as novel biomarkers in non-small

cell lung cancer screening. *Tumour Biol.* 36, 3035–3042. doi: 10.1007/s13277-014-2938-1

Yu, L., Todd, N. W., Xing, L., Xie, Y., Zhang, H., Liu, Z., et al. (2010). Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers. *Int. J. Cancer* 127, 2870–2878. doi: 10.1002/ijc.25289

Yuan, Y., Xu, X. Y., Zheng, H. G., and Hua, B. J. (2018). Elevated miR-21 is associated with poor prognosis in non-small cell lung cancer: a systematic review and meta-analysis. *Eur Rev Med Pharmacol. Sci.* 22, 4166–4180. doi: 10.26355/eurrev_201807_15410

Zappa, C., and Mousa, S. A. (2016). Non-small cell lung cancer: current treatment and future advances. *Transl. Lung Cancer Res.* 5, 288–300. doi: 10.21037/tlcr.2016.06.07

Zhong, S., Golpon, H., Zardo, P., and Borlak, J. (2021). miRNAs in lung cancer. A systematic review identifies predictive and prognostic miRNA candidates for precision medicine in lung cancer. *Transl. Res.* 230, 164–196. doi: 10.1016/j.trsl.2020.11.012

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 21 September 2021
doi: 10.3389/fgene.2021.743505

CircDTL Functions as an Oncogene and Regulates Both Apoptosis and Ferroptosis in Non-small Cell Lung Cancer Cells

*Wang Shanshan[1], Ma Hongying[1], Fang Jingjing[1], Yu Yiming[1], Ren Yu[2] and Yu Rui[3]**

[1] *The Affiliated Hospital of Medical School, Ningbo University, Ningbo, China,* [2] *Department of Urologic Surgery, Ningbo Urology and Nephrology Hospital, Ningbo, China,* [3] *Department of Biochemistry and Molecular Biology, Medical School of Ningbo University, Ningbo, China*

OPEN ACCESS

Edited by:
Wei Jiang,
Nanjing University of Aeronautics and Astronautics, China

Reviewed by:
Lihong Wang,
Southeast University, China
Xuexin Yu,
University of Texas Southwestern Medical Center, United States

****Correspondence:***
Yu Rui
yurui@nbu.edu.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *18 July 2021*
Accepted: *18 August 2021*
Published: *21 September 2021*

Citation:
Shanshan W, Hongying M, Jingjing F, Yiming Y, Yu R and Rui Y (2021) CircDTL Functions as an Oncogene and Regulates Both Apoptosis and Ferroptosis in Non-small Cell Lung Cancer Cells. Front. Genet. 12:743505. doi: 10.3389/fgene.2021.743505

Background: Circular RNAs (circRNA) play an essential role in the tumorigenesis of non-small cell lung cancer (NSCLC). CircDTL is a novel identified circRNA with little information regarding its biological role. However, the role of circDTL in NSCLC has not been investigated yet.

Method: In this study, the levels of circDTL in tissues and cells were measured by RT-PCR. Cell viability was measured by the CCK-8 assay. Cell migration and invasion were evaluated using the wound healing assay and transwell assay, respectively. Cell death was measured by the cell death ELISA kit. The levels of Fe^{2+}, ROS, MDA and GSH were measured using the commercial kits. The interactions between miR-1287-5p and circDTL/3′UTR GPX4 were verified by dual-luciferase activity assay. The effects of circDTL on tumor growth were evaluated *in vivo*.

Results: CircDTL was found to be upregulated and acted as an oncogene in NSCLC cells. Knockdown of circDTL promoted both apoptosis and ferroptosis of NSCLC cells. It was identified that circDTL exerts its oncogenic effects via the circDTL/miR-1287-5p/GPX4 axis and GPX4 inhibits both ferroptosis and apoptosis. Finally, this study showed that silencing of circDTL promoted the sensitivity of NSCLC cells to chemotherapeutic agents and inhibited the growth of tumors *in vivo*.

Conclusion: CircDTL acts as an oncogene and exerts its effects via the miR-1287-5p/GPX4 axis in NSCLC, providing a potential therapeutic target for NSCLC cancer therapy.

Keywords: apoptosis, circDTL, ferroptosis, non-small cell lung cancer, miRNA

Abbreviations: NSCLC, non-small cell lung cancer; circRNA, Circular RNA; miRNAs, MicroRNAs; ROS, reactive oxygen species; GPX4, Glutathione Peroxidase 4; MDA, Malondialdehyde; Fer-1, Ferrostatin-1; Lip-1, liproxstatin-1;RSL3, RAS-selective lethal 3;GSH, Glutathione.

INTRODUCTION

Lung cancer is the most prevalent diagnosed cancer and the leading cause of cancer-related death worldwide (Siegel et al., 2019). Non-small-cell lung carcinoma (NSCLC) accounts for nearly 85% of lung cancers and the prognosis of NSCLC is still dismal (Keith and Miller, 2013). Therefore, it is important to unveil the molecular mechanisms underlying the progression of NSCLC.

Circular RNA (circRNA) is a group of non-coding RNA featuring a closed-loop structure. Circular RNAs are more stable than linear RNAs due to their circular shape. CircRNAs have been shown to perform a variety of biological roles, including sponging miRNA, regulating RNA binding protein, and transcription in nuclear (Hsiao et al., 2017). Several studies have suggested that dysregulation of circRNAs is correlated to the development of NSCLC. CircDTL is a novel circRNA that has been reported to be dysregulated in medulloblastoma, but its role in NSCLC remains unknown, and further investigation is needed to understand its underlying mechanism (Lv et al., 2018).

MicroRNAs (miRNAs) are another form of small non-coding RNAs that regulate various biological processes such as differentiation, migration, metabolism, and programmed cell death (Hammond, 2015). Various miRNAs have been reported to participate in the tumorigenesis of cancers. miR-1287-5p has been found dysregulated and played essential roles in multiple cancers. For instance, miR-1287-5p inhibited the development of triple-negative breast cancer via inhibition of phosphoinositide-3-kinase CB (Schwarzenbacher et al., 2019). Another study found that miR-1287-5p could regulate the progression of cervical cancer via regulation of HOXA7 (Ji et al., 2020). However, there's little knowledge about the role of miR-1287-5p in NSCLC.

Ferroptosis is a novel type of regulated cell death mediated by reactive oxygen species (ROS) and lipid peroxidation (Cao and Dixon, 2016). Glutathione peroxidase 4 (GPX4), a glutathione peroxidase, belongs to the GPX family and plays an essential role in the process of ferroptosis (Yang et al., 2014). GPX4 is able to convert lipid hydroperoxides to lipid alcohols. Therefore, inhibition of GPX4 can lead to the accumulation of lipid peroxides and thereby promote the ferroptosis. Various studies have shown that induction of ferroptosis via targeting GPX4 might be a promising strategy for killing cancer cells including the NSCLC (Hassannia et al., 2019). For example, blockage of GPX4 could overcome resistance to Lapatinib via promoting ferroptosis in NSCLC (Ni et al., 2021). Another study also found that ammonium ferric citrate induced ferroptosis via blockage of GPX4 in NSCLC cells (Wu et al., 2021). Interestingly, several researchers have found that non-coding RNAs also play essential roles in regulating ferroptosis (Wang et al., 2019; Lu et al., 2020). However, there's still no investigation into the effects of circDTL on ferroptosis.

This study found that circDTL was upregulated and acted as an oncogene during the progression of NSCLC. Interestingly, silencing of circDTL induced ferroptosis of NSCLC cells and further investigation showed that circDTL regulated ferroptosis via the miR-1287-5p/GPX4 axis. These findings underlined the importance of circRNAs in the process of ferroptosis and introduced circDTL as a regulator of ferroptosis in NSCLC.

MATERIALS AND METHODS

Clinical Specimens

A total of 84 pairs of NSCLC tissues and adjacent normal tissues were obtained from the Affiliated Hospital of Ningbo University. Before the operation, no treatment was given to any of the patients. This study was authorized by The Ethics Committee of Ningbo University and conducted following the Declaration of Helsinki. Before the study, all participants signed written informed consents.

Cell Culture

Normal human lung epithelial cell line BEAS-2B and human NSCLC cell lines (H23, H522, PC9, and A549) were purchased from the Shanghai Bank of Cell Culture (Shanghai, China). All cells were cultured in RPMI1640 medium (Gibco, United States) supplemented with 10% fetal bovine serum (FBS, Gibco), 100 U/ml penicillin and 100 μg/ml streptomycin (Invitrogen, United States). At 37°C, the cells were kept in a humidity-controlled environment with 5% CO_2.

RNA Purification and RT-PCR

Total RNA was extracted using Trizol (Invitrogen, United States). RNA was reversely transcribed into cDNA using PrimeScript RT Reagent (Takara, China). The expression levels of circDTL and miRNAs were evaluated using the SYBR Premix Ex Taq (Takara, China). Internal controls were GAPDH and U6. The relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method and the experiments were performed in triplicate.

Cell Transfection

shRNAs against circDTL were sub-cloned into the GV248 (hU6-MCS-Ubiquitin-EGFP-IRES-puromycin) construct (Genechem, China). The full length of GPX4 was synthesized by GenePharm (China) and sub-cloned into the GV248 vector. The QuickChange site-directed mutagenesis kit (Agilent, United States) was used to create Mutant GRPX4. The plasmids were transfected into HEK-293 cells using the ViraPower Kit (Life Technologies), to generate the lentivirus. GenePharm supplied miR-1287-5p mimics/inhibitors and negative controls (NC-mimics/inhibitors), and transfection was performed using the Lipofectamine 2000 (Life Technologies) according to the manufacturer's instruction.

RNA Pull-Down Assay

The RNA pull-down assay was performed according to the protocol from GeneSeed (Guangzhou, China). Following formaldehyde fixation, the cells were sonicated. Then, the supernatant was incubated with the biotinylated circDTL or control probe (RioBio) and the magnetic streptavidin Dyna beads (Sigma). After total RNA extraction, qRT-PCR was used to assess the enrichment.

Cell Viability Assay

The viability of the cells was determined according to a previously published article (Yu et al., 2016). The cells were plated into 96-well plates at a density of 1×10^4 cells/well. After culturing for 24 h, cells were subjected to different treatments. 10 μl of CCK-8 solution was added to each well and incubated at 37°C for another 2 h. Then the absorbance was measured at 450 nm using a microplate reader (Biotek, United States).

Cellular Death Assay

The cellular death was measured by the Cell Death Detection ELISA^{plus} kit (Roche, Germany), following the manufacturer's guide.

Measurement of ROS, MDA, GSH, and Fe^{2+}

The levels of ROS, MDA, GSH, and Fe^{2+} were measured using the Cellular ROS Assay Kit (Abcam, United States), Lipid Peroxidation (MDA) Assay Kit (Abcam, United States), GSH Assay Kit (Abcam, United States), and Iron Assay Kit (Abcam, United States), respectively.

Sub-Cellular Fraction Assay

The location of RNAs was determined using the PARIS^{TM} kit (Invitrogen), according to the company's manual. Cells were suspended in cytoplasm lysis buffer and centrifuged at 1,500 rpm for 5 min. The cytoplasmic supernatant was collected, and the pellet was re-suspended in nucleus lysis buffer at 4°C for 1 h after centrifugation at 1,500 rpm for 10 min. The RNAs derived from cytoplasmic and nuclear extracts were purified by TRIzol. The expression levels of GAPDH (cytoplasm control), U6 (nucleus control) and circDTL in the nucleus and cytoplasm were assayed by qRT-PCR.

Dual-Luciferase Activity Assay

The wild-type (including miR-1287-5p binding sites) sequences of circDTL whole length or GPX4 3′UTR region were sub-cloned into the pmirGLO vector (Life Technologies). NSCLC cells were co-transfected with luciferase reporters along with miR-1287-5p mimics (wild type or mutants) or miR-NC using Lipofectamine 2000 according to the company's protocol. The Dual-Luciferase Reporter Assay System (System) was used to adjust relative luciferase activity to firefly luciferase activity 48 h after transfection (Promega Madison, WI, United States).

Cell Migration and Invasion Assay

The migration of cells was measured by the wound healing assay. After transfection for 24 h, cells were seeded into a 6-well plate at the density of 2×10^5 cells/well. 24 The monolayer was scratched with a sterile 20 l pipette tip 24 h later. Cell migration was observed 24 h later. Invasion of cells was assayed using the Transwell assay. 1×10^5 transfected cells were suspended in 200 μl of serum-free medium and seeded into the top chambers of Transwell (8 μm pore size, Corning, United States) coated with Matrigel (BD Bioscience). As an attractant, the bottom chamber was filled with the entire medium. After incubation for 24 h, non-invaded cells were gently removed and cells invaded were fixed with 4% paraformaldehyde (Sigma), stained with crystal violet solution (Beyotime) for 30 min, and visualized. All cells were counted in five microscopic areas selected at random.

Western Blot Assay

The RIPA lysis buffer was used to lyse the cells (Beyotime). SDS-PAGE was used to separate 20 g of total protein, which was then transferred to the PVDF membrane. The membranes were blocked with skimmed milk for 1 h at room temperature, and then the membrane was incubated with primary antibody overnight at 4°C. After that, the membrane was washed three times with PBS and incubated with corresponding HRP-conjugated secondary antibody at room temperature for 1 h. ECL Prime Western Blotting Kit was used to view the results (Beyotime). Cellular Signaling Technology, United States (CST) provided all of the main and secondary antibodies.

Animal Study

The Shanghai SLAC Animal Center (Shanghai, China) provided 4–6-week-old BALB/c male nude mice, which were kept according to the standards for the use and care of laboratory animals. A total of 1×10^7 NSCLC cells infected with shRNA were injected subcutaneously into the left flank of nude mice (3 per group). Every 3 days, the tumor volume was measured. After 30 days, all naked mice were killed, and the tumor tissues were analyzed by western blotting. These experiments gained permission from the committee on the Ethics of Animal Experiments of Affiliated Hospital of Ningbo University. The animal experiments were conducted according to the U.K Animals Act 1986 and associated guidelines.

Statistically Analysis

Statistical analyses were performed with SPSS12.0 (IBM, Chicago, IL, United States). Data are expressed as the mean ± SD. A one-way ANOVA analysis was used to determine the statistical difference between multiple groups. The statistical difference between the two groups was calculated using a *post hoc* test. P-value < 0.05 (two-tailed) was considered statistically significant.

RESULTS

CircDTL Was Upregulated and Acted as an Oncogene in NSCLC

The expression of circDTL was first assessed in 84 pairs of NSCLC tissues and surrounding normal tissues. The levels of circDTL were observed to be greater in NSCLC tissues than in normal tissues (**Figure 1A**). The levels of circDTL in NSCLC cells (H1299, NCI-H522, PC9, and A549) were also higher than human lung epithelial cell line BEAS-2B (**Figure 1B**). Two shRNAs targeting circDTL were utilized to knock down circDTL in NSCLC cells to study the biological role of circDTL in NSCLC (**Figure 1C**). Cell viability assays

showed that downregulation of circDTL significantly inhibited the viabilities of NSCLC cells compared with the control groups (**Figure 1D**). Moreover, wound healing and Transwell assays showed that downregulation of circDTL repressed the migration and invasion of NSCLC cells (**Figures 1E,F**). These findings showed that circDTL might have a role in NSCLC as an oncogene.

Inhibition of circDTL Induced Apoptosis and Ferroptosis in NSCLC Cells

The effects of knockdown of circDTL on cell death of NSCLC cells were evaluated. Silencing of the circDTL gene resulted in greater cell death than control groups (**Figure 2A**). To determine the cell death type caused by the knockdown of circDTL, various specific inhibitors were applied. Apoptosis inhibitor (z.VAD.FMK, 10 μM) and ferroptosis inhibitors (Fer-1 10 μM, Lip-1 20 μM) but not necrosis inhibitor (Necrosulfonamide, NEC) significantly attenuated the cell death caused by knockdown of circDTL in NSCLC cells (**Figure 2A**). Caspase-3 activity assays and western blots confirmed that knockdown of circDTL triggered activation of caspase-3 in NSCLC cells (**Figures 2B,C**). Simultaneously, it was shown that silencing circDTL resulted in the downregulation of Bcl-2 and Mcl-1, as well as the overexpression of Bax (**Figure 2C**). Silencing of circDTL also led to the release of mitochondrial proteins Smac/DIABLO and cytochrome c into the cytosol (**Figure 2C**). Hence, knocking down circDTL triggered apoptosis in NSCLC cells via the intrinsic apoptotic mechanism. To confirm whether ferroptosis was also triggered by silencing of circDTL in NSCLC cells, the levels of ROS, MDA, GSH, and Fe^{2+} were assayed. It was found that knockdown of circDTL led to the upregulation of cellular ROS, which could be counteracted by Fer-1 or Lip-1 (**Figure 2D**). Silencing of circDTL also increased the levels of MDA, Fe^{2+} and decreased GSH levels, and those effects could be reversed by Fer-1 and Lip-1 (**Figures 2E–G**). These findings showed that inhibiting circDTL caused NSCLC cells to undergo both apoptosis and ferroptosis.

circDTL Negatively Regulate the Expression of miR-1287-5p in NSCLC Cells

Several studies have suggested that circDTL acts as a "sponge" to regulate the expression of miRNAs. A sub-cellular fractionation experiment was used to confirm this, and it revealed that circDTL is mostly found in the cytoplasm of NSCLC cells (**Figure 3A**). Then, using bioinformatical tools (StarBase3.0, TargetScan), potential miRNAs that can bind to circDTL were predicted (**Supplementary File 1**). RNA pull-down assays showed that specific probes against circDTL could enrich circDTLRNAs over 30-fold changes compared with control (**Figure 3B**). We chose those predicted miRNAs with an Ago CLIP-seq Data $\geq$ 5 and have been previously reported involved in the tumorigenesis of NSCLC to further analysis. Among them, miR-1287-5p was found to be significantly elevated in the circDTL probe group, whereas other miRNAs showed minimal change (**Figure 3C**). To further test that miR-1287-5p can bind with circDTL, two mutant miR-1287-5p mimics (Mut1, Mut2) containing mutations within the binding site for circDTL and a wild-type (WT) miR-1287-5p mimics were transfected into NSCLC cells (**Figure 3D**, left). It was found that only the WT miR-1287-5p mimic but not the mutants, inhibited circDTL luciferase reporter activity (**Figure 3D**, right). To further, elucidate the correlation between miR-1287-5p and circDTL, the expression of miR-1287-5p after knockdown of circDTL was examined. The inhibition of circDTL led to the upregulation of miR-1287-5p (**Figure 3E**). In addition, the expression of miR-1287-5p in normal tissues was much higher than NSCLC tissues (**Figure 3F**). These data suggested that circDTL acts as a "sponge" to negatively regulate the expression of miR-1287-5p.

miR-1287-5p Targets GPX4 Which Regulates the Ferroptosis and Apoptosis of NSCLC Cells

In this study, it was attempted to identify the potential targets of miR-1287-5p. Various possible targets were predicted using bioinformatical techniques (TargetScan, miRDB), and GPX4 piqued the curiosity for this research for three reasons. Firstly, GPX4 was predicted in common by both analysis tools. Secondly, GPX4 was documented played a critical role in ferroptosis (Yang et al., 2014). Thirdly, GPX4 was also reported to be able to inhibit apoptosis *in vivo* (Ran et al., 2004). To verify whether miR-1287-5p can directly target GPX4, dual-luciferase reporter assays were conducted. Only the WT miR-1287-5p mimics, but not the mutants, inhibited GPX4 3′-UTR luciferase reporter activity (**Figure 4A**). Furthermore, transfection of the WT miR-1287-5p mimic into NSCLC cells inhibited the mRNA (**Figure 4B**) and protein levels (**Figure 4C**) of GPX4, whereas the mutants were ineffective (**Figures 4B,C**, right panel). RT-PCR results showed that mRNA levels of GPX4 in NSCLC tissues were much higher than adjacent normal tissues (**Figure 4D**). Moreover, GPX4 mRNA was shown to be concentrated in biotinylated WT miR-1287-5p but was unable to attach to mutant miR-1287-5p (Mut1, 2, 2, biotinylated) (**Figure 4E**). Similarly, it was observed that shRNAs against circDTL were also able to downregulate the GPX4 in NSCLC cells (**Figure 4F**). To investigate whether circDTL exerts its effects via the regulation of miR-1287-5p/GPX4 axis in NSCLC cells, a 3′UTR mutant GPX4 construct (mut GPX4) was transduced into shcircDTL stably expressing NSCLC cells. Mut GPX4 restored mRNA (**Figure 4G**) and protein (**Figure 4H**) expression of GPX4 in cells with knockdown of circDTL. According to cell viability assays, mutant GPX4 abolished knockdown of circDTL induced inhibition of cell viability (**Figure 4I**), migration (**Figure 4J**), and invasion (**Figure 4K**) in NSCLC cells. Meanwhile, increased cell death (**Figure 4L**) and release of mitochondrial proteins (**Figure 4M**) caused by the knockdown of circDTL were also abrogated by overexpression of mutant GPX4. Moreover, the effects of knockdown of circDTL on ROS, lipid peroxidation, GSH, and Fe^{2+} levels were also reversed by overexpression of mutant GPX4 (**Figure 4N**). These results suggested that

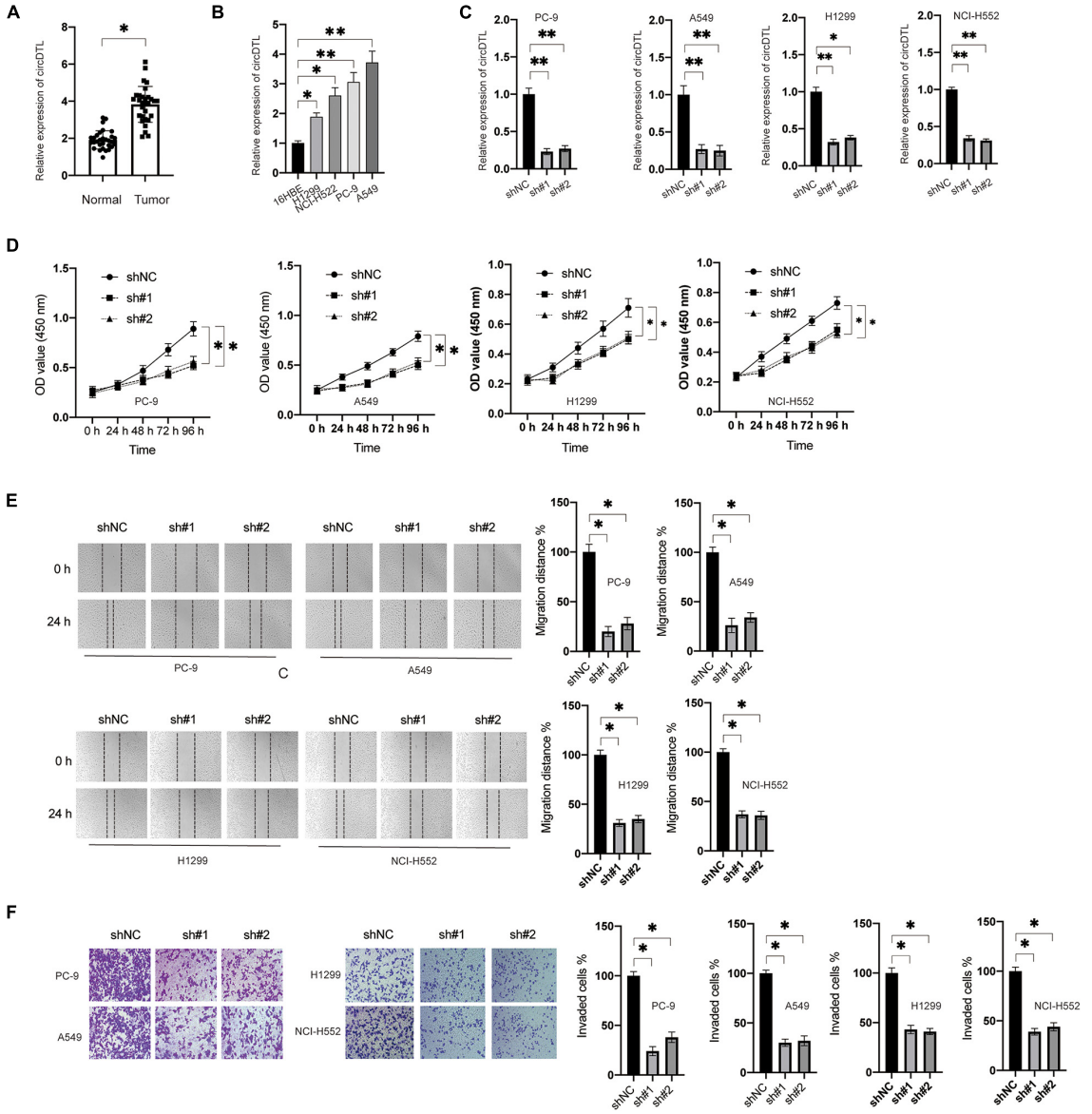


FIGURE 1 | circDTL was upregulated and acted as an oncogene in NSCLC. **(A)** Levels of circDTL in 84 pairs of NSCLC and adjacent tissues. **(B)** Levels of circDTL in normal cell line BEAS-2B and NSCLC cell lines. **(C)** NSCLC cells were transfected with shNC or shRNA against circDTL (sh#1, sh#2) for 24 h, levels of circDTL were measured. **(D)** NSCLC cells were transfected as indicated and cell viability was measured at different time points. **(E)** NSCLC cells were transfected as indicated, cell migration was assayed. **(F)** Cell invasion was measured. Data were presented as mean ± SD. Experiments were performed at least three times. $^{*}P < 0.05$; $^{**}P < 0.01$.

circDTL exerts its effect via modulation of miR-1287-5p/GPX4 axis in NSCLC cells.

Silencing of circDTL Increases Sensitivity to Chemotherapeutic/Ferroptosis Inducing Agents and Inhibits Tumor Growth *in vivo*

Furthermore, this study evaluated the effects of inhibition of circDTL on the chemosensitivity of NSCLC cells. In NSCLC cells, it was found that silencing circDTL enhanced cell death induced by different chemotherapeutic drugs (Cisplatin 10 M, Paclitaxel 10 M, Gefitinib 10 M, Docetaxel 10 M) (**Figure 5A**). Meanwhile, the cell death caused by those agents could be significantly inhibited by z.VAD (10 μM) (**Figure 5A**). Similarly, it was also observed that inhibition of circDTL increased sensitivity to ferroptosis inducers (Erastin 5 μM, RSL3 10 μM), and this effect could be blocked by ferroptosis inhibitors (Fer-1 10 μM, Lip-1 20 μM) (**Figure 5B**). The effects of downregulation of circDTL on the progression of NSCLC were then evaluated in nude mice xenograft models. The NSCLC cells transfected with shNC or shcircDTL were inoculated into nude mice,

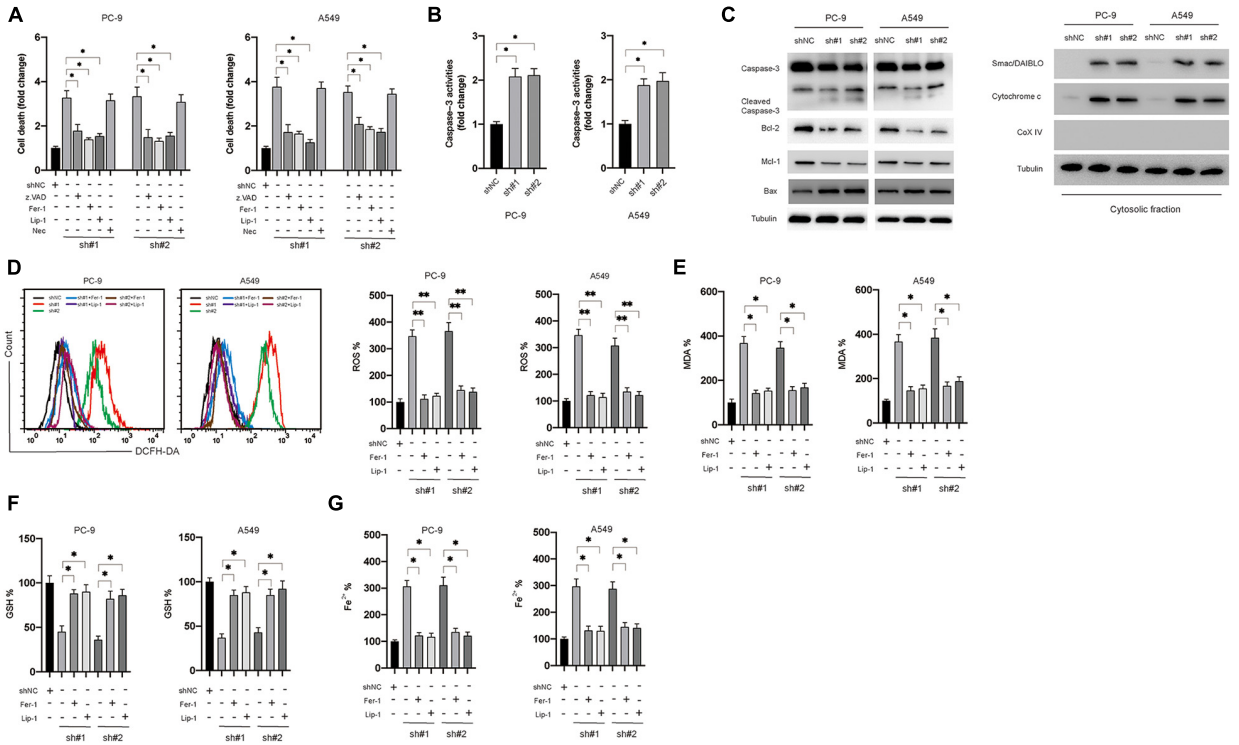


FIGURE 2 | Silencing of circDTL induced apoptosis and ferroptosis of NSCLC cells. **(A)** NSCLC cells were transfected as indicated with or without different inhibitors for 24 h, cellular death was assayed. **(B)** NSCLC cells were transfected as indicated 24 h, caspase-3 activities were measured. **(C)** Apoptosis-related proteins were measured by western blots. **(D)** NSCLC cells were transfected as indicated and treated with or without ferroptosis inhibitors (Fer-1, Lip-1) for 24 h, ROS levels were measured. **(E)** MDA levels. **(F)** GSH levels. **(G)** Fe^{2+} levels. Data were presented as mean ± SD. Experiments were performed at least three times. $^{*}P < 0.05$; $^{**}P < 0.01$.

which revealed that shcircDTL markedly repressed the growth of tumors *in vivo*, as tumor sizes were significantly decreased when compared with the shNC group (**Figure 5C**). Furthermore, following silencing of circDTL, xenograft tumor tissues were exposed to western blotting, and it was shown that caspase-3 was activated (**Figure 5D**). These data indicated that downregulation of circDTL not only enhanced the chemosensitivity of NSCLC cells but also inhibited the growth of NSCLC cells *in vivo*.

DISCUSSION

Various studies have suggested that ncRNAs played essential roles in the initiation and progression of NSCLC (Anastasiadou et al., 2018). In this study, it was found that circDTL acted as an oncogene in NSCLC. It was found that knockdown of circDTL inhibited proliferation, migration, and invasion, and promoted both apoptosis and ferroptosis of NSCLC cells via modulation of miR-1287-5p/GPX4 axis.

Apoptosis is a programmed cell death that is triggered via two pathways namely the extrinsic and intrinsic pathways (Yu et al., 2016). The intrinsic pathway is strictly subjected to the regulation of Bcl-2 family proteins. In this study, it was found that silencing of circDTL leads to downregulation of Bcl-2, Mcl-1, and upregulation of Bax. These findings suggested that inhibition of circDTL induced apoptosis via the intrinsic pathway. Considering that most chemotherapeutic agents primarily induce apoptosis via the intrinsic route (Wu et al., 2021), it might explain why the silencing of circDTL increased the chemosensitivity of NSCLC cells. Ferroptosis is a novel identified form of cell death. Till now, there are only a few studies regarding the role of circRNAs in ferroptosis. For instance, Liu et al. found that circRNA cIARS regulated ferroptosis of liver carcinoma cells by interacting with the RNA-binding protein ALKBH5 (Liu et al., 2020). Zhang et al. (2020) showed that circTTBK2 regulates ferroptosis via binding with miR-761 and promoting ITGB8 expression in glioma cells. In this study, it was reported that silencing of circDTL promoted ferroptosis via regulation of miR-1287-5p/GPX4 axis in NSCLC cells.

Many studies have suggested that circRNAs exert regulatory effects through acting as competing endogenous RNAs (Memczak et al., 2013). Based on bioinformatics analysis, miR-1287-5p was predicted to bind with circDTL. Moreover, this study also found that circDTL was mainly located in the cytoplasm, and dual-luciferase reporter assays further confirmed the interaction between circDTL and miR-1287-5p. These findings suggested that miR-1287-5p acted as a tumor suppressor in NSCLC, which is similar to previous studies that had reported that miR-1287-5p inhibited the growth of breast and cervical cancer (Schwarzenbacher et al., 2019; Ji et al., 2020).

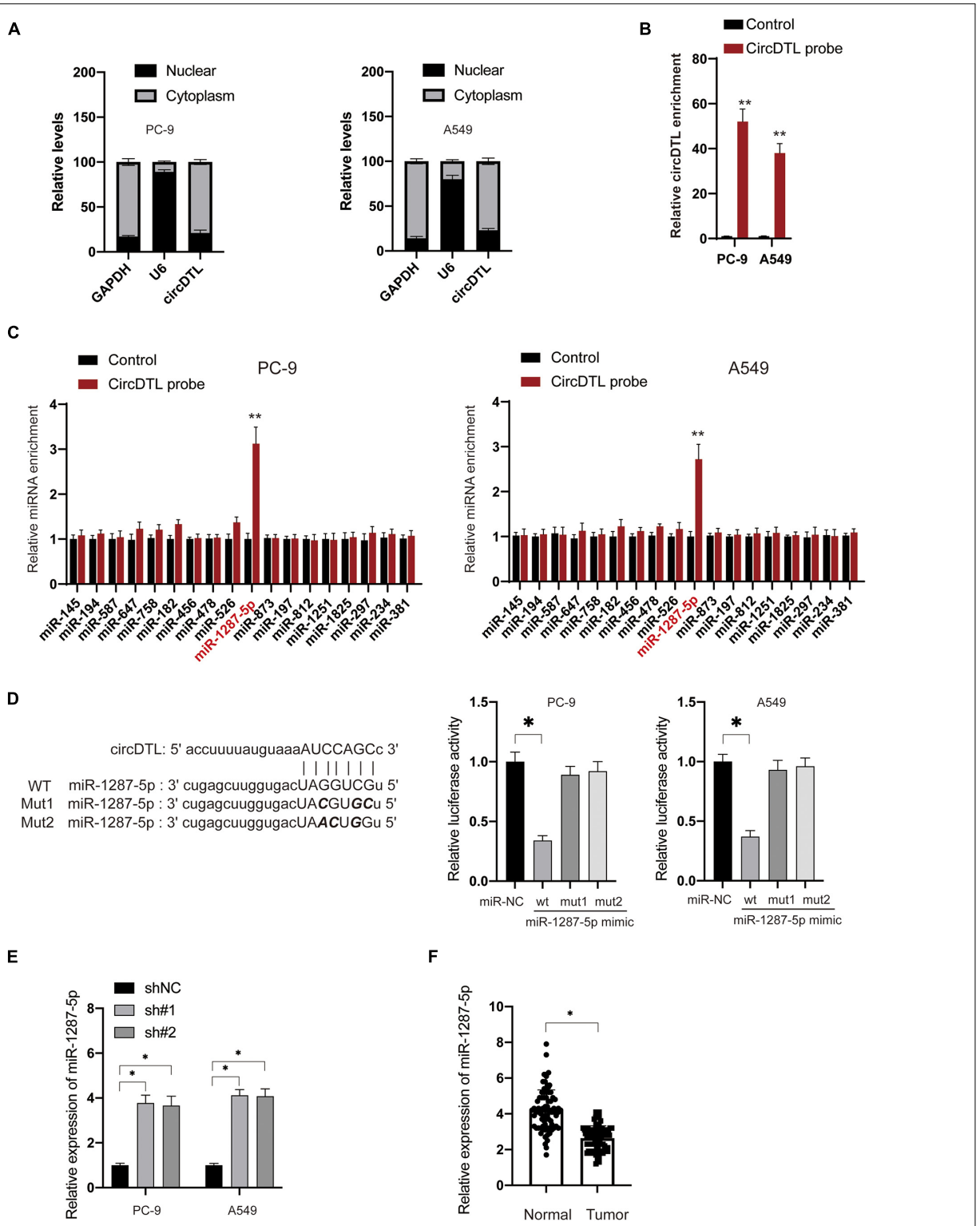


FIGURE 3 | circDTL negatively regulates the expression of miR-1287-5p in NSCLC cells. **(A)** Cellular location of GAPDH, U6, and circDTL was assayed in NSCLC cells. **(B)** RNA pull-down assay was used for the detection of circDTL. **(C)** RNA pull-down was performed to examine different miRNAs that bind with circDTL. **(D)** Predictive binding sites between circDTL and miR-1287-5p (left), the interaction between miR-1287-5p and circDTL was evaluated by dual-luciferase reporter assay (right). **(E)** NSCLC cells were transfected with shNC or shRNAs against circDTL (sh#1, sh#2) for 24 h, levels of miR-1287-5p were measured. **(F)** The expression of miR-1287-5p in 84 pairs of NSCLC and adjacent normal tissues. Data were presented as mean ± SD. Experiments were performed at least three times. $^{*}P < 0.05$; $^{**}P < 0.01$.

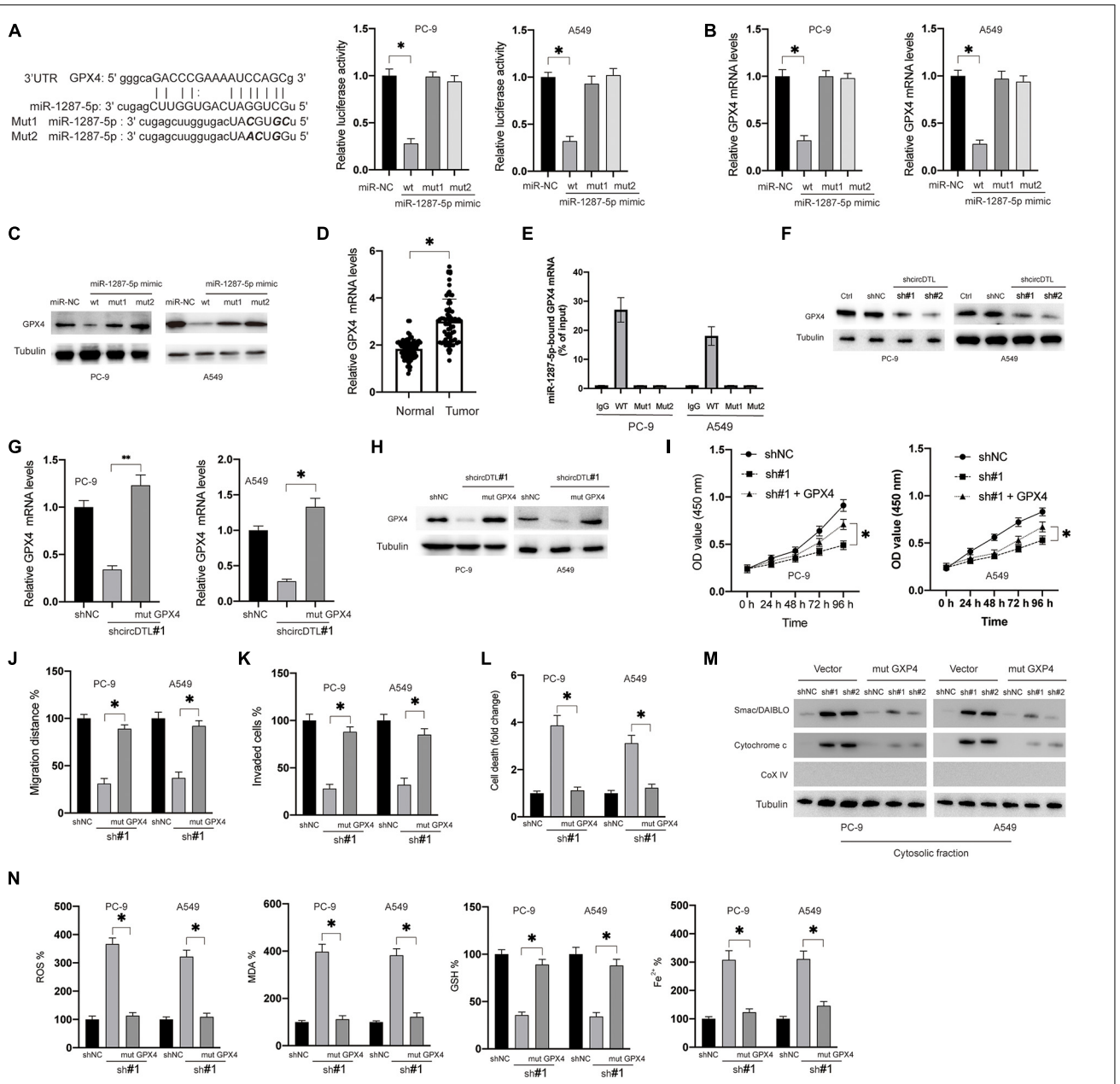


FIGURE 4 | miR-1287-5p targets GPX4 to regulate both apoptosis and ferroptosis of NSCLC cells. **(A)** Predictive binding sites between miR-1287-5p and 3′-UTR of GPX4 (left), the interaction between miR-1287-5p and 3′-UTR of GPX4 was evaluated by dual-luciferase reporter assay (right). **(B)** NSCLC cells were transfected as indicated for 24 h, mRNA levels of GPX4 were measured by RT-PCR. **(C)** Protein levels of GPX4 were measured. **(D)** mRNA levels of GPX4 were measured in NSCLC and adjacent normal tissues. **(E)** RNA pull-down was conducted to evaluate the GPX4 mRNA that binds with miR-1287-5p. **(F)** NSCLC cells were transfected with shNC or shRNAs against circDTL (sh#1, sh#2), protein levels of GPX4 were measured. **(G)** mRNA levels of GPX4 were measured. **(H)** Protein levels of GPX4 were measured. **(I)** Cell viabilities were measured at different time points. **(J)** cell migration distance was calculated. **(K)** Several invaded cells were measured. **(L)** Cell death was measured. **(M)** the release of mitochondrial proteins was measured. **(N)** Levels of ROS, MDA, GSH, and Fe^{2} + were measured. Data were presented as mean ± SD. Experiments were performed at least three times. $^{*}P < 0.05$; $^{**}P < 0.01$.

This study investigated the miR-1287-downstream 5p's targets as well. The data suggested that GPX4 was a target of miR-1287-5p. Rescue experiments also confirmed that upregulation of GPX4 might restore the effects of downregulation of circDTL in NSCLC cells. GPX4 is an intracellular antioxidant enzyme that capable of reducing intracellular peroxidized phospholipids (Imai and Nakagawa, 2003). In recent years, many studies indicated that GPX4 plays an essential role in the process of ferroptosis. A study reported that inhibition of GPX4 led to ferroptosis especially in drug-resistant tumor cells (Viswanathan et al., 2017). Many agents, such as RSL3, ML162, and ML210, have been designed to target GPX4 to induce ferroptosis

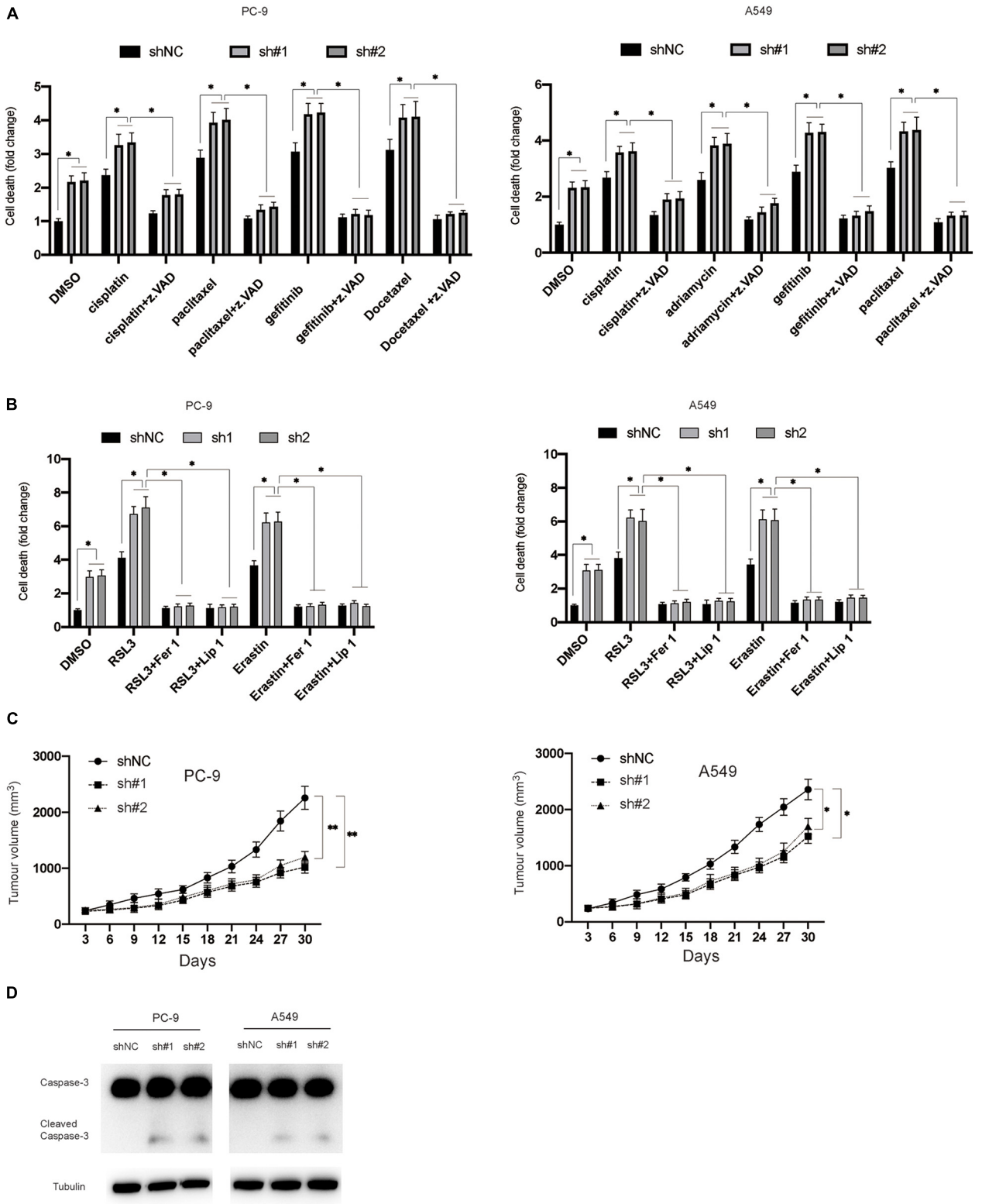


FIGURE 5 | Silencing of circDTL increased the chemosensitivity of NSCLC cells and inhibited the growth of NSCLC *in vivo*. **(A)** NSCLC cells were exposed to different chemotherapy agents with or without z.VAD for 24 h and cellular death was measured. **(B)** NSCLC cells were exposed to RSL3 or Erastin with or without Fer-1/Lip-1 for 24 h, cellular deaths were measured. **(C)** NSCLC cells stably transfected with shNC or shcircDTL and inoculated into nude mice, tumor volumes were measured at different time points. **(D)** Xenografts were subjected to western blotting analysis. Data were presented as mean ± SD. Experiments were performed at least three times. $^{*}P < 0.05$; $^{**}P < 0.01$.

(Cao and Dixon, 2016). Overexpression of GPX4 reduced the release of smacSmac/DIABLO and cytochrome c, indicating that GPX4 influences intrinsic apoptosis. This finding is in line with a previous study that also reported that GPX4 blocked the release of cytochrome c from mitochondria (Nomura et al., 2000). In this case, these findings showed that GPX4 affects both apoptosis and ferroptosis. Another study revealed that high levels of GPX4 were correlated with the poor prognosis of NSCLC patients (Liu et al., 2018). We also found that GPX4 was significantly upregulated in NSCLC tissues than adjacent normal tissues. Hence, targeting GPX4 as a potential treatment for NSCLC might be a promising strategy.

CONCLUSION

It can be concluded that this study is the first report that showed circDTL acted as an oncogene in NSCLC. Functional investigation showed that circDTL affected proliferation, migration, invasion, apoptosis, and ferroptosis of NSCLC cells. Mechanical examination showed that circDTL regulates both apoptosis and ferroptosis via miR-1287-5p/GPX4 axis. These findings have provided new insight into the development of cirRNA-based therapeutics against NSCLC.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Ningbo University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by The Animal Ethics Committee of Ningbo University.

AUTHOR CONTRIBUTIONS

WS conducted most of the experiments and draft the manuscript. MH repeated the experiments and verified the data. FJ conducted the statistical analysis. YY conducted the animal study. RY collected the clinical samples. YR designed and supervised this study. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We sincerely thank wenjun Shang for the helpful suggestion during this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.743505/full#supplementary-material

REFERENCES

Anastasiadou, E., Jacob, L. S., and Slack, F. J. (2018). Non-coding RNA networks in cancer. *Nat. Rev. Cancer* 18, 5–18. doi: 10.1038/nrc.2017.99

Cao, J. Y., and Dixon, S. J. (2016). Mechanisms of ferroptosis. *Cell. Mol. Life Sci.* 73, 2195–2209. doi: 10.1007/s00018-016-2194-1

Hammond, S. M. (2015). An overview of microRNAs. *Adv. Drug Deliv. Rev.* 87, 3–14. doi: 10.1016/j.addr.2015.05.001

Hassannia, B., Vandenabeele, P., and Vanden Berghe, T. (2019). Targeting ferroptosis to iron out cancer. *Cancer Cell* 35, 830–849. doi: 10.1016/j.ccell.2019.04.002

Hsiao, K. Y., Sun, H. S., and Tsai, S. J. (2017). Circular RNA–new member of noncoding RNA with novel functions. *Exp. Biol. Med. (Maywood)* 242, 1136–1141. doi: 10.1177/1535370217708978

Imai, H., and Nakagawa, Y. (2003). Biological significance of phospholipid hydroperoxide glutathione peroxidase (PHGPx, GPx4) in mammalian cells. *Free Radic. Biol. Med.* 34, 145–169. doi: 10.1016/s0891-5849(02)01197-8

Ji, F., Du, R., Chen, T., Zhang, M., Zhu, Y., Luo, X., et al. (2020). Circular RNA circSLC26A4 accelerates cervical cancer progression via miR-1287-5p/HOXA7 axis. *Mol. Ther. Nucleic Acids* 19, 413–420. doi: 10.1016/j.omtn.2019.11.032

Keith, R. L., and Miller, Y. E. (2013). Lung cancer chemoprevention: current status and future prospects. *Nat. Rev. Clin. Oncol.* 10, 334–343. doi: 10.1038/nrclinonc.2013.64

Liu, K., Jin, M., Xiao, L., Liu, H., and Wei, S. (2018). Distinct prognostic values of mRNA expression of glutathione peroxidases in non-small cell lung cancer. *Cancer Manag. Res.* 10, 2997–3005. doi: 10.2147/CMAR.S163432

Liu, Z., Wang, Q., Wang, X., Xu, Z., Wei, X., and Li, J. (2020). Circular RNA cIARS regulates ferroptosis in HCC cells through interacting with RNA binding protein ALKBH5. *Cell Death Discov.* 6:72. doi: 10.1038/s41420-020-00306-x

Lu, J., Xu, F., and Lu, H. (2020). LncRNA PVT1 regulates ferroptosis through miR-214-mediated TFR1 and p53. *Life Sci.* 260:118305. doi: 10.1016/j.lfs.2020.118305

Lv, T., Miao, Y. F., Jin, K., Han, S., Xu, T. Q., Qiu, Z. L., et al. (2018). Dysregulated circular RNAs in medulloblastoma regulate proliferation and growth of tumor cells via host genes. *Cancer Med.* 7, 6147–6157. doi: 10.1002/cam4.1613

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495, 333–338. doi: 10.1038/nature11928

Ni, J., Chen, K., Zhang, J., and Zhang, X. (2021). Inhibition of GPX4 or mTOR overcomes resistance to lapatinib via promoting ferroptosis in NSCLC cells. *Biochem. Biophys. Res. Commun.* 567, 154–160. doi: 10.1016/j.bbrc.2021.06.051

Nomura, K., Imai, H., Koumura, T., Kobayashi, T., and Nakagawa, Y. (2000). Mitochondrial phospholipid hydroperoxide glutathione peroxidase inhibits the release of cytochrome c from mitochondria by suppressing the peroxidation of cardiolipin in hypoglycaemia-induced apoptosis. *Biochem. J.* 351(Pt 1), 183–193. doi: 10.1042/0264-6021:3510183

Ran, Q., Liang, H., Gu, M., Qi, W., Walter, C. A., Roberts, L. J. II, et al. (2004). Transgenic mice overexpressing glutathione peroxidase 4 are protected against oxidative stress-induced apoptosis. *J. Biol. Chem.* 279, 55137–55146. doi: 10.1074/jbc.M410387200

Schwarzenbacher, D., Klec, C., Pasculli, B., Cerk, S., Rinner, B., Karbiener, M., et al. (2019). MiR-1287-5p inhibits triple negative breast cancer growth by interaction with phosphoinositide 3-kinase CB, thereby sensitizing cells

for PI3Kinase inhibitors. *Breast Cancer Res.* 21:20. doi: 10.1186/s13058-019-1104-5

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer statistics, 2019. *CA Cancer J. Clin.* 69, 7–34. doi: 10.3322/caac.21551

Viswanathan, V. S., Ryan, M. J., Dhruv, H. D., Gill, S., Eichhoff, O. M., Seashore-Ludlow, B., et al. (2017). Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. *Nature* 547, 453–457. doi: 10.1038/nature23007

Wang, M., Mao, C., Ouyang, L., Liu, Y., Lai, W., Liu, N., et al. (2019). Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA. *Cell Death Differ.* 26, 2329–2343. doi: 10.1038/s41418-019-0304-y

Wu, W., Geng, Z., Bai, H., Liu, T., and Zhang, B. (2021). Ammonium ferric citrate induced ferroptosis in non-small-cell lung carcinoma through the inhibition of GPX4-GSS/GSR-GGT axis activity. *Int. J. Med. Sci.* 18, 1899–1909. doi: 10.7150/ijms.54860

Yang, W. S., SriRamaratnam, R., Welsch, M. E., Shimada, K., Skouta, R., Viswanathan, V. S., et al. (2014). Regulation of ferroptotic cancer cell death by GPX4. *Cell* 156, 317–331. doi: 10.1016/j.cell.2013.12.010

Yu, R., Yu, B. X., Chen, J. F., Lv, X. Y., Yan, Z. J., Cheng, Y., et al. (2016). Anti-tumor effects of atractylenolide I on bladder cancer cells. *J. Exp. Clin. Cancer Res.* 35:40. doi: 10.1186/s13046-016-0312-4

Zhang, H. Y., Zhang, B. W., Zhang, Z. B., and Deng, Q. J. (2020). Circular RNA TTBK2 regulates cell proliferation, invasion and ferroptosis via miR-761/ITGB8 axis in glioma. *Eur. Rev. Med. Pharmacol. Sci.* 24, 2585–2600. doi: 10.26355/eurrev_202003_20528

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 28 September 2021
doi: 10.3389/fgene.2021.744113

Identification and Validation of Hypoxia-Related lncRNA Signature as a Prognostic Model for Hepatocellular Carcinoma

Chenghui Zhou[1,2], *Huajun Zhang*[3] *and Liqing Lu*[3]*

[1]Department of General Surgery, Xiangya Hospital, Central South University, Changsha, China, [2]Department of General, Visceral, Cancer and Transplantation Surgery, University Hospital Cologne, Cologne, Germany, [3]Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Xiangya Hospital, Central South University, Changsha, China

Hepatocellular carcinoma (HCC) is one of the most general malignant tumors. Hypoxia is a critical clinical characteristic and acts as a significant part in the development and cancers' prognosis. The prognostic value and biological functions of hypoxia-related lncRNAs in hepatocellular carcinoma is little known. Thus, we aim to establish a hypoxia-related lncRNA signature to predict the HCC patients' survival. First, we extracted the hypoxia-related genes and expression of lncRNAs from the MSigDB and TCGA database, respectively. The co-expression analysis among hypoxia-related mRNAs and lncRNAs was employed to identify hypoxia-related lncRNAs. Then, comprehensive analyses of lncRNAs expression level and survival data were applied to establish the signature. We built a prognostic signature on the foundation of the three differently expressed hypoxia-related lncRNAs. Kaplan-Meier curves indicated the low-risk group is associated with better survival. The 1–, 3–, and 5 years AUC values of the signature were 0.805, 0.672 and 0.63 respectively. The test set performed consistent outcomes. A nomogram was built grounded on the risk score and clinicopathological features. GSEA showed the immune-related pathways in high-risk group, while metabolism-related pathways in low-risk group. Besides, we found this model was correlated with the clinical features, tumor immune cell infiltration, immune checkpoints, and m6A-related genes. Finally, a novel signature based on hypoxia-related lncRNAs was established and validated for predicting HCC patients' survival and may offer some useful information for immunotherapies.

Keywords: hepatocellular carcinoma, hypoxia, lncRNA, prognosis, immune cells infiltration, immune checkpoint, m6A

OPEN ACCESS

Edited by:
Wei Jiang,
Nanjing University of Aeronautics and Astronautics, China

Reviewed by:
Huiyu Li,
University of Texas Southwestern Medical Center, United States
Xinyi Liu,
University of Illinois at Chicago, United States

***Correspondence:**
Liqing Lu
liqing1992@csu.edu.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *03 August 2021*
Accepted: *07 September 2021*
Published: *28 September 2021*

Citation:
Zhou C, Zhang H and Lu L (2021) Identification and Validation of Hypoxia-Related lncRNA Signature as a Prognostic Model for Hepatocellular Carcinoma. Front. Genet. 12:744113. doi: 10.3389/fgene.2021.744113

INTRODUCTION

Hepatocellular carcinoma (HCC) one of the most general subtype of liver cancer, ranks the fifth in terms of cancer-related deaths in the world (Bray et al., 2018). The surgical treatment, radiofrequency ablation, chemotherapy and as well liver transplant are the traditional treatments for hepatocellular carcinoma patients and the therapeutic effects are limited. The majority of patients are diagnosed with advanced stage and there are distant metastasis, tumor thrombosis of portal vein, and so on, leading to extremely poor prognosis (Galle et al., 2017; Vilgrain et al., 2017). Nowadays, in China, the survival rate of

5 years of HCC patients is only 14.1% (Allemani et al., 2018). Thus, novel predictive biomarkers and therapeutic options are critically required.

Hypoxia belongs to a public microenvironmental property of most solid tumors, because of the imbalance between the tumor proliferation and oxygen supply (Hill et al., 2015; Muz et al., 2015). Previous studies have proved that hypoxia plays critical roles in angiogenesis, differentiation, proliferation and apoptosis of tumor cells (Nishida and Kudo, 2013; Wu et al., 2007). As one of the organs most prone to hypoxia, hypoxia is correlated with the metastasis, prognosis, and resistance of HCC (Erler and Giaccia, 2006; Graham and Unger, 2018). Whereas lncRNA is a subset of RNA transcripts over 200 nt, which not only regulates the mRNA's expression but also mediates the development, occurrence, metastasis and prognosis of various cancers (Kang et al., 2019; Gao et al., 2021; Xiao et al., 2021). At present, multiple studies found that hypoxia-related lncRNAs are associated with overall survival in various tumors (Chen et al., 2020; Zhang et al., 2021). Nonetheless, whether the hypoxia-related lncRNAs are related to the prognosis of hepatocellular carcinoma is still largely not clear.

Herein, we mined the expression data of lncRNAs and the relevant clinical information of hepatocellular carcinoma samples, and hypoxia-related genes from the opening databases. Next, the prognostic signature of hypoxia-related lncRNA was built and then evaluated it based on the TCGA cohort. Lastly, we validated the predictive capability of the risk model and explored the role in diagnostic value, infiltrating immune cells, N6-methyladenosine (m6A) mRNA, and immune checkpoints.

TABLE 1 | The clinical information of hepatocellular carcinoma patients in the whole cohort.

Variable	Number of patients
Age at diagnosis	
≤65	216
>65	127
Gender	
Female	110
Male	233
Grade	
G1	53
G2	161
G3	112
G4	12
Unknown	5
Stage	
I	161
II	77
III	80
IV	3
Unknown	22

MATERIALS AND METHODS

Data Collection

The RNA-sequencing (FPKM) of 374 hepatocellular carcinoma samples was derived from the TCGA data portal. The corresponding data of clinical, for example, age, gender, clinical grade, clinical stage and survival information, was also downloaded from the TCGA database. The clinical information of these patients is displayed in **Table 1**. The Ensembl was applied to obtain the gene transfer annotation files to differentiate the messenger RNAs (mRNAs) from long non-coding lncRNAs (lncRNAs) for subsequent analyses. Besides, the merge language script Perl was used to merge the data files of RNA-sequencing into a matrix file. Then, the genes' Ensembl IDs were translated into an array of gene symbols by utilizing the script of the Perl language.

Additionally, the Molecular Signatures database (MSigDB) (M10508, M26925) was applied to download the hypoxia-related genes. The co-expression analysis was employed to identify hypoxia-related lncRNAs. The hypoxia-related genes with correlation coefficients > 0.5 and p-value < 0.001 were regarded as hypoxia-related lncRNAs. The lncRNAs with the average expression less than 0.5 were delated. The expression profiles of hypoxia-related genes/lncRNAs of the 50 adjacent normal tissue samples and 374 samples of hepatocellular carcinoma were subjected to DE-HRGs/DE-HRlncRNAs using the "limma" package in R. The GO and KEGG were applied to explore the potential function of the differentially expressed hypoxia-related genes (DE-HRGs) (FDR < 0.05 and $|\log2FC| \geq 1$) using the "clusterProfiler", "enrichplot" and "ggplot2" R packages. The samples of hepatocellular carcinoma were distributed into the train and test sets applying the "caret" package randomly. We built the prognostic signature based on the hypoxia-related lncRNAs' expression level in the train set and validated its prognostic value in the test set.

Construction of a Hypoxia-Related lncRNA Signature to Assess the Risk Score

To identify the hypoxia-related lncRNAs that related overall survival, we merged the hypoxia-related lncRNAs and survival information utilizing the package "limma". The survival time of hepatocellular carcinoma samples ranks from 0 to 3,675 days. To improve the quality analyses, samples that lack complete survival information and those whose overall survival was not over 30 days were excluded from this analysis. Then, the univariate Cox regression analysis was applied to get survival-related hypoxia-related lncRNAs from the hypoxia-related lncRNAs ($p < 0.05$). These survival-related hypoxia-related lncRNAs were used for further analysis.

The LASSO Cox analysis was employed to identify the prognostic hypoxia-related lncRNAs. To prevent overfitting, we applied 1000-round cross-validation for tuning parameter selection to fulfill the minimum criteria of the partial likelihood deviance. Finally, the multivariate Cox regression analysis was fulfilled on these prognostic hypoxia-related lncRNAs to build a prognostic hypoxia-related lncRNAs signature and compute the coefficients. The risk score for each sample was counted utilizing

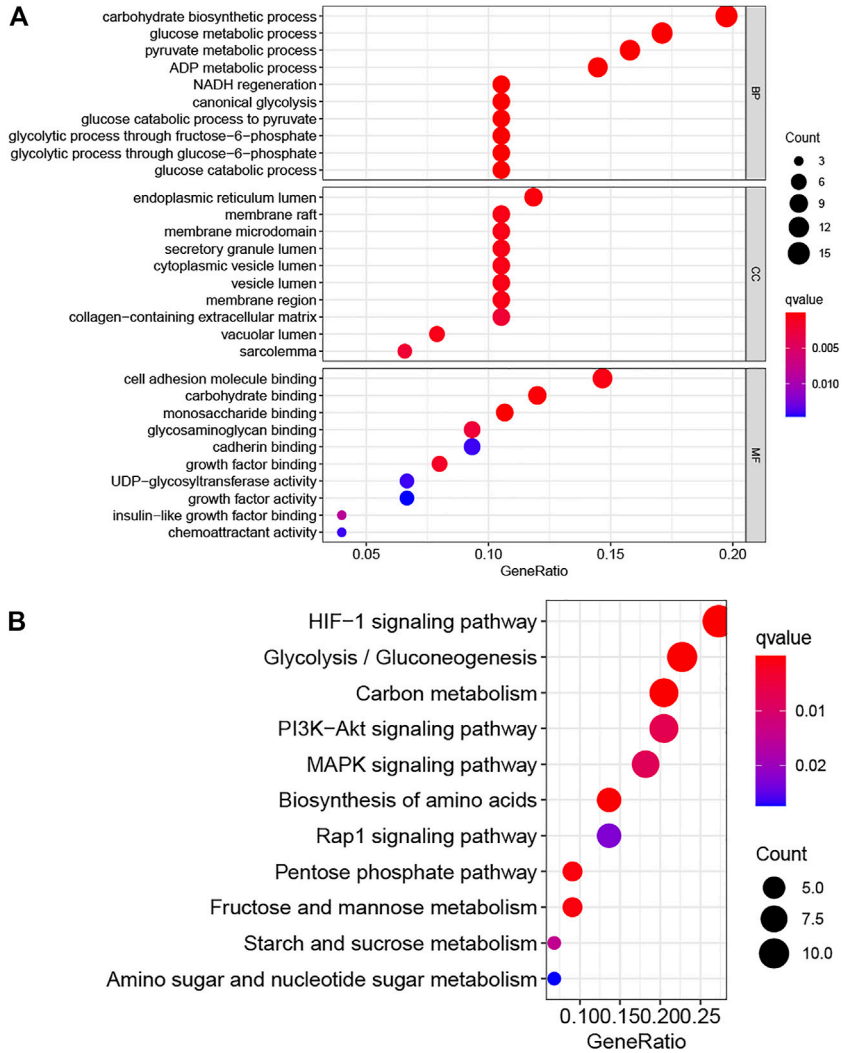


FIGURE 1 | GO **(A)** and KEGG **(B)** analysis for hypoxia-related differentially expressed genes.

the following formula: risk score = ∑icoefficient (hypoxia-related lncRNA) × expression of (hypoxia-related lncRNAi).

Validation of the Hypoxia-Related lncRNA Signature

The patients of train and test sets were divided into two subgroups (low-risk vs high-risk) on the foundation of the median value of the risk score of train cohort, respectively. We applied Kaplan-Meier survival analysis to explore the difference of survival between the two groups in the train and test sets. The time-dependent receiver operating characteristic (ROC) curves of 1-, 3- and 5 years were plotted using the "survivalROC" package in two sets. Besides, to explore whether the risk score can regard as an independent biomarker for the hepatocellular carcinoma patients' survival, the univariate and multivariate Cox regression analysis were used in the train and test sets. To prove the clinical application value of this signature, we applied the chi-square test to explore the relationship between the model and clinical variables. We compared the differences in the risk score among subgroups with different clinical features with a Wilcoxon signed-rank test.

Construction of Nomogram and Gene Set Enrichment Analysis

A nomogram was built on the foundation of the risk score and other clinical biomarkers for predicting the 1-, 3-, 5 years overall survival of hepatocellular carcinoma patients. We applied the calibration curves to assess the nomogram' accuracy. The Gene set enrichment analyses (GSEA) was performed to investigate the KEGG pathways between two risk (low vs high) groups in whole cohort of TCGA-LIHC using GSEA 4.0.1 software. $p < 0.05$ and FDR < 5% were considered as statistical significance.

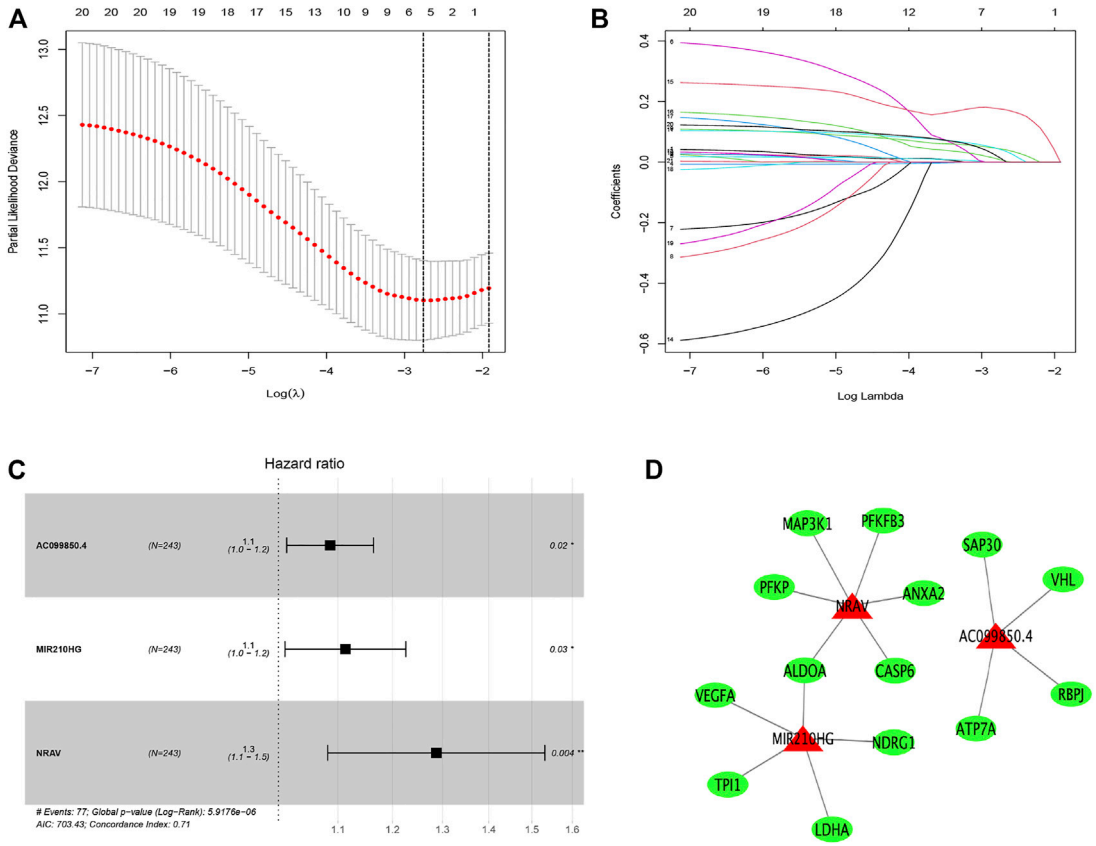


FIGURE 2 | Distinction of hypoxia-related lncRNAs signature: LASSO **(A,B)** and multivariate **(C)** Cox regression analysis. The relationship between the prognostic lncRNAs (red) and genes (green) **(D)**.

Association Between Risk Scores and Immune Cell Infiltration

To explore the association between the risk score and features of immune cell infiltration, we used the various acknowledged methods to count the tumor infiltrating immune cells levels among the patients from the whole cohort, including CIBERSORT-ABS, XCELL, EPIC, MCPcounter, TIMER, QUANTISEQ, and CIBERSORT. We used the Wilcoxon signed-rank test to explore the difference in immune cell infiltration level between two risk groups. To explore the association between two risk groups and immune status, we quantified the tumor-infiltrating cells related pathways with ssGSEA (Xiao et al., 2020), namely, APC_co_inhibition, APC_co_stimulation, CCR, check-point, cytolytic activity, HLA, inflammation promoting, MHC_class_I, parainflammation T cell co-inhibition, T cell co-stimulation, type I_IFN response, and type II_IFN response. Previous studies have proved that m6a-related genes were associated with prognosis in hepatocellular carcinoma (Zhao et al., 2018; Wu et al., 2020). What is more, m6a-related genes can mediate the expression of lncRNAs, which dyregulate the tumorigenesis, development, metastasis, as well as prognosis (Lan et al., 2019; Wu et al., 2021a). Thus, we evaluated the association between the risk score and N6-methyladenosine (m6A) mRNAs as well as immune checkpoints in the whole cohort. The m6A mRNAs and immune checkpoints were obtained from previous literature.

RESULTS

Enrichment Analysis of Hypoxia-Related Genes

We obtained 76 DE-HRGs including 58 upregulated and 18 downregulated genes (**Supplementary Table S1**). The most enriched GO terms were CC (cellular component) such as membrane raft, endoplasmic reticulum lumen, secretory granule lumen, membrane microdomain and cytoplasmic vesicle lumen; BP (biological process), for example, glucose catabolic process to pyruvate, NADH regeneration, carbohydrate biosynthetic process, glucose metabolic process and canonical glycolysis; and MF (molecular function) such as glycosaminoglycan binding, monosaccharide binding, carbohydrate binding, growth factor binding, and cell adhesion molecule binding (**Figure 1A**, **Supplementary Table S2**). The KEGG pathways of 76 DE-HRGs were predominantly related to glycolysis/gluconeogenesis, starch and sucrose metabolism,

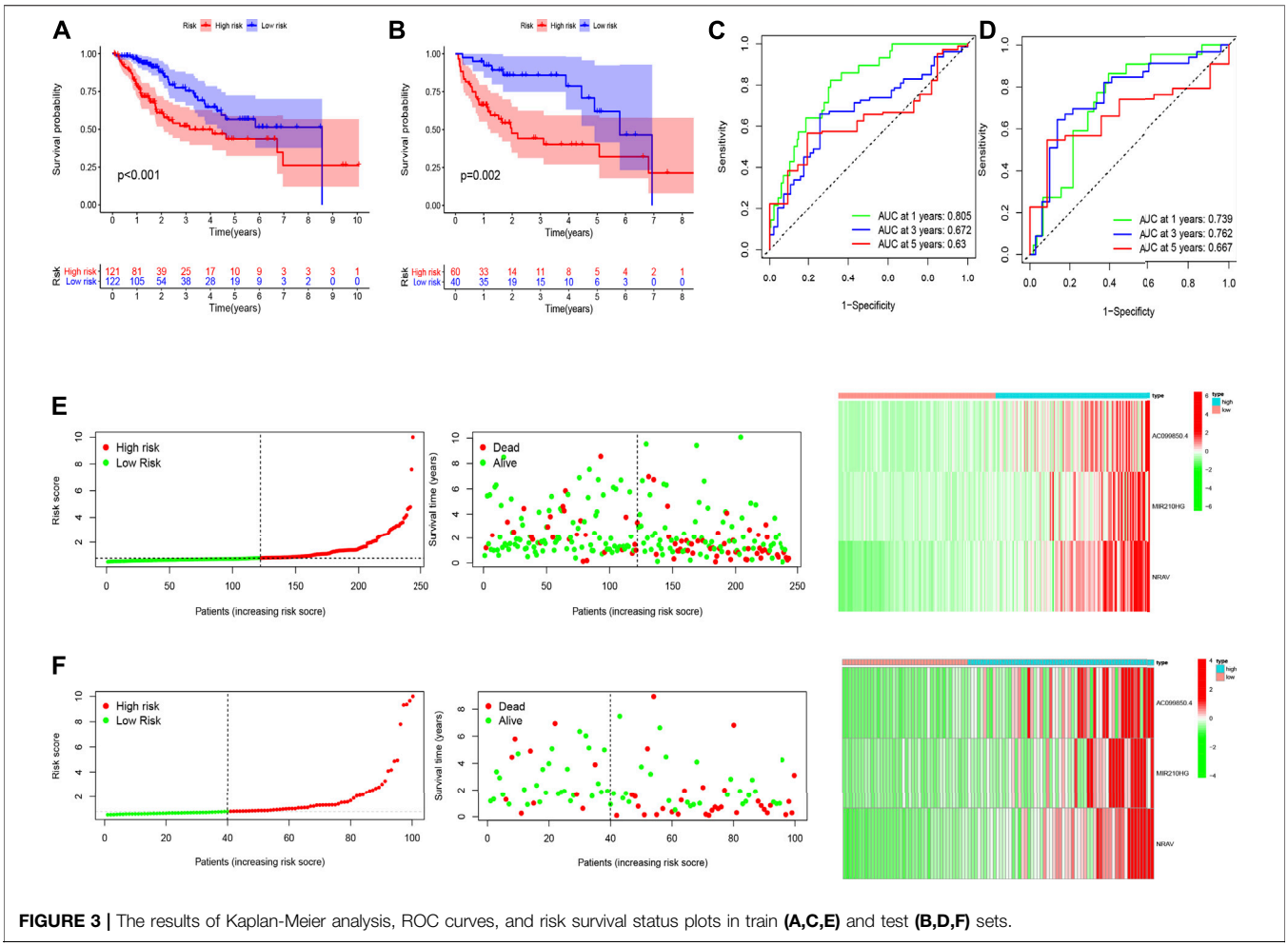


FIGURE 3 | The results of Kaplan-Meier analysis, ROC curves, and risk survival status plots in train **(A,C,E)** and test **(B,D,F)** sets.

fructose and mannose metabolism, as well as pentose phosphate pathway (**Figure 1B**, **Supplementary Table S2**).

Establishment of Hypoxia-Related lncRNA Signature

We obtained 200 hypoxia-related genes from the MSigDB database (**Supplementary Table S3**). Then, a total of 171 hypoxia-related lncRNAs were identified by applying co-expression analysis between hypoxia-related genes and lncRNAs. Among these hypoxia-related lncRNAs, 111 lncRNAs were differentially expressed hypoxia-related lncRNAs (**Supplementary Figures S1A,B**). Then, we added the hypoxia-related lncRNAs with survival data by merging complete survival material of samples and hypoxia-related lncRNAs. Next, 343 hepatocellular carcinoma samples were distributed into the train (243) and test (100) sets applying the "caret" package randomly. The samples in the train set were employed for constructing the prognostic model.

First, we identified 21 prognostic hypoxia-related lncRNAs utilizing univariate Cox regression analysis, as shown in **Supplementary Table S4**. Then, utilizing LASSO regression analysis, 16 hypoxia-related lncRNAs were removed (**Figures 2A,B**). Finally, three hypoxia-related lncRNAs (AC099850.4, MIR210HG, and NRAV) were obtained through using multivariate Cox regression analysis (**Figure 2C**). The hypoxia-related lncRNAs signature was constructed based on these three lncRNAs and corresponding coefficients: risk score = (0.08283797* expression of AC099850.4) + (0.10722211* expression of MIR210HG) + (0.25250738* expression of NRAV). Besides, the association between three hypoxia-related lncRNAs in the signature and hypoxia-related genes is displayed in **Figure 2D**.

Assessment of Clinical Performance of Risk Model

To explore the underlying prognostic capability of the hypoxia-related lncRNA signature, we assigned the patients in train and test sets into two groups (low-risk vs high-risk) on the foundation of the median of risk score. The Kaplan-Meier analysis demonstrated the high-risk group was correlated with worse survival ($p < 0.001$, **Figure 3A**) in the train set and the K-M analysis in the test set revealed the similar result ($p = 0.002$, **Figure 3B**). Meanwhile, the 1-, 3-, and 5 years AUC values of ROC curve were 0.805, 0.672 and 0.63, respectively, in the train

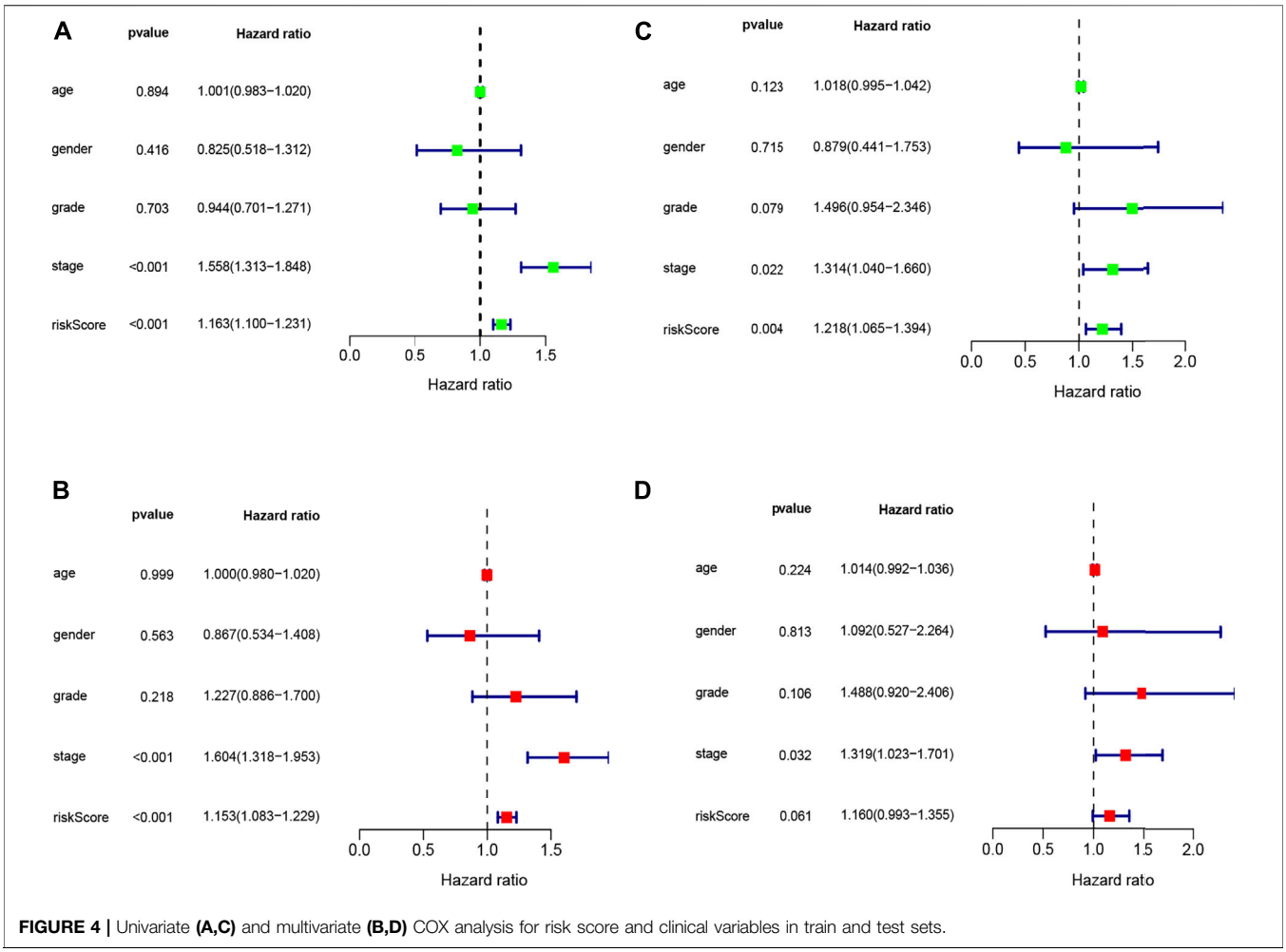


FIGURE 4 | Univariate **(A,C)** and multivariate **(B,D)** COX analysis for risk score and clinical variables in train and test sets.

set (**Figure 3C**). The corresponding values in the test set were 0.739, 0.762 and 0.667, respectively (**Figure 3D**). Utilizing the risk survival status plots of patients, we revealed that the risk scores of patients were negatively correlated with the hepatocellular carcinoma patients' survival in train and test sets. Besides, the heatmaps showed that these three lncRNAs were positively associated with the risk signature in two sets (**Figures 3E,F**).

The univariate Cox regression analysis revealed that risk score (HR = 1.163, 95% CI: 1.100–1.231, $p < 0.001$) and stage (HR = 1.558, 95% CI: 1.313–1.848, $p < 0.001$) were remarkably associated with overall survival in the train set (**Figure 4A**), and risk score (HR = 1.218, 95% CI: 1.065–1.394, $p = 0.004$) and stage (HR = 1.314, 95% CI: 1.040–1.660, $p = 0.022$) were also associated with overall survival in the test set (**Figure 4C**). Furthermore, the multivariate Cox regression analysis also demonstrated that risk score (HR = 1.153, 95% CI: 1.318–1.953, $p < 0.001$) and stage (HR = 1.604, 95% CI: 1.313–1.848, $p < 0.001$) were the independent factors to predict the hepatocellular carcinoma patients' prognosis in the train set (**Figure 4B**), and risk score (HR = 1.160, 95% CI: 0.993–1.355, $p = 0.061$) and stage (HR = 1.319, 95% CI: 1.023–1.701, $p = 0.032$) were also the independent factors in the test set (**Figure 4D**). Subsequently, we applied the ROC curve to appraise the accurateness of clinical variables in predicting the HCC patients' overall survival. As shown in **Supplementary Figure S1C**, the AUC value of clinical variables was lower than that of risk score. Besides, to further explore the signature's prognostic value in hepatocellular carcinoma patients arranged by clinical variables, the patients were divided into distinct subgroups according to clinical variables (age, sex, stage and clinical grade). **Figure 5** shows that the survival of patients in high-risk group was worse than those in the low-risk group in most of subgroups.

The nomogram based on the risk score and clinical features was constructed for predicting the 1-, 3- and 5 years overall survival of the hepatocellular carcinoma patients (**Figure 6A**). The 1-, 3-, and 5 years calibration curves showed stable and accurate performance (**Figures 6B–D**); therefore, they may be used for hepatocellular carcinoma patients' clinical management.

Gene Set Enrichment Analysis

GSEA demonstrated the prognostic hypoxia-related lncRNAs signature was enriched in immune-, metabolism- and tumor-related pathways. The high-risk group was enriched in B cell receptor signaling pathway, T cell receptor signaling pathway, P53 signaling pathway, TGF_BETA signaling pathway and WNT

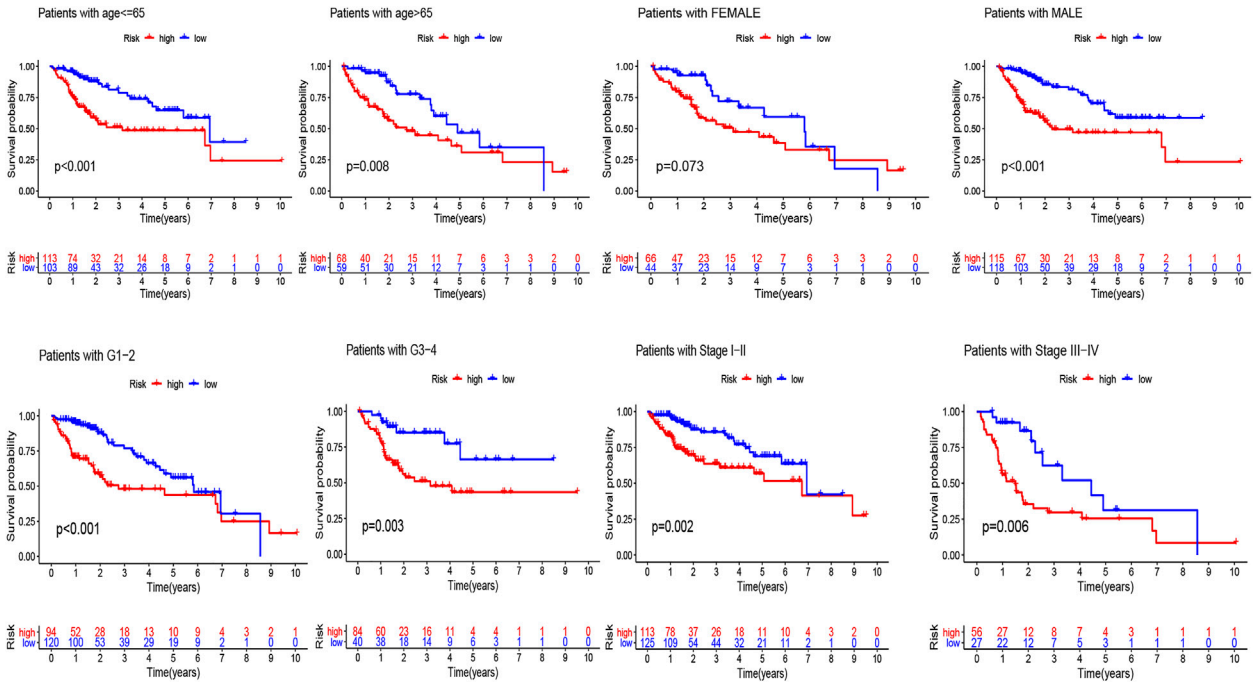


FIGURE 5 | Kaplan-Meier analyses for the two subgroups (high-risk vs low-risk) categorized by clinical variables, comprising age, gender, grade, and stage.

signaling pathway, while the low-risk group was enriched in butanoate metabolism, fatty acid metabolism, retinol metabolism, tryptophan metabolism and glycine serine and threonine metabolism (**Figure 7**).

Tumor Immune Cell Infiltration and Gene Expression

The heatmap of tumor immune cell infiltration based on TIMER, EPIC, CIBERSORT−ABS, QUANTISEQ, XCELL, MCP counter and CIBERSORT algorithms is displayed in **Figure 8** and **Supplementary Table S5**. The difference of immune cell related functions indicated that cytolytic activity, type II IFN response and MHC class I were remarkably different between the two risk groups applying the single-sample gene set enrichment analysis (ssGSEA) (**Figure 9A**). Considering the significance of immune checkpoints in immunotherapies, we next investigated the difference of immune checkpoints' expression between two risk groups. The boxplot showed most of these genes, for example, PDCD-1 (PD-1), CD86, LAIR1, CTLA4 and CD70, existed significantly different between the two groups (**Figure 9B**). Besides, most of m6A-related genes' expression also had significant difference between two risk groups except for FMR1 and ZC3H13 (**Figure 9C**).

DISCUSSION

Despite the development of surgical resection, liver transplantation, chemotherapy, immunotherapy, and targeted therapy in the therapy of hepatocellular carcinoma, the hepatocellular carcinoma patients' prognosis is still poor (Maluccio and Covey, 2012; Xu et al., 2020a). The studies showed that the treatment response and prognosis are influenced by tumor molecular characteristics, although patients own homogeneous clinical risk variables. Thus, it is urgent to identify novel molecular prognostic markers that distinguish from common clinical risk features. Previous studies have proved that hypoxia is associated with tumor angiogenesis, metabolic alteration, metastasis, and prognosis (Krock et al., 2011; Eales et al., 2016; Sun et al., 2020). Meanwhile, lncRNAs play a crucial role in regulating the hypoxia-related genes' expression. Therefore, we built a signature based on the hypoxia-related lncRNAs and explored tumor immune cell infiltration, immune checkpoints and m6A-related genes in the prognosis of hepatocellular carcinoma, which may offer underlying biomarkers and therapeutic targets for HCC.

First, we explored the function of 76 hypoxia related DEGs. The KEGG analysis indicated these genes mainly enriched in glycolysis/gluconeogenesis, fructose and mannose metabolism, pentose phosphate pathway, as well as starch and sucrose metabolism. Under hypoxia, cancer cells could retain survival and stimulate tumor proliferation, migration, and invasion *via* glycolysis (Huang and Zong, 2017). Wu et al. reported that OMA, an ATP-dependent zinc metalloprotease, promotes cancer development by reprogramming metabolic colorectal cancer under hypoxia (Wu et al., 2021b). A recent study revealed that hypoxia downregulates the enzymes of pentose phosphate pathway and upregulates the enzymes of glycolysis in tumor and then stimulate tumor migration while inhibiting tumor

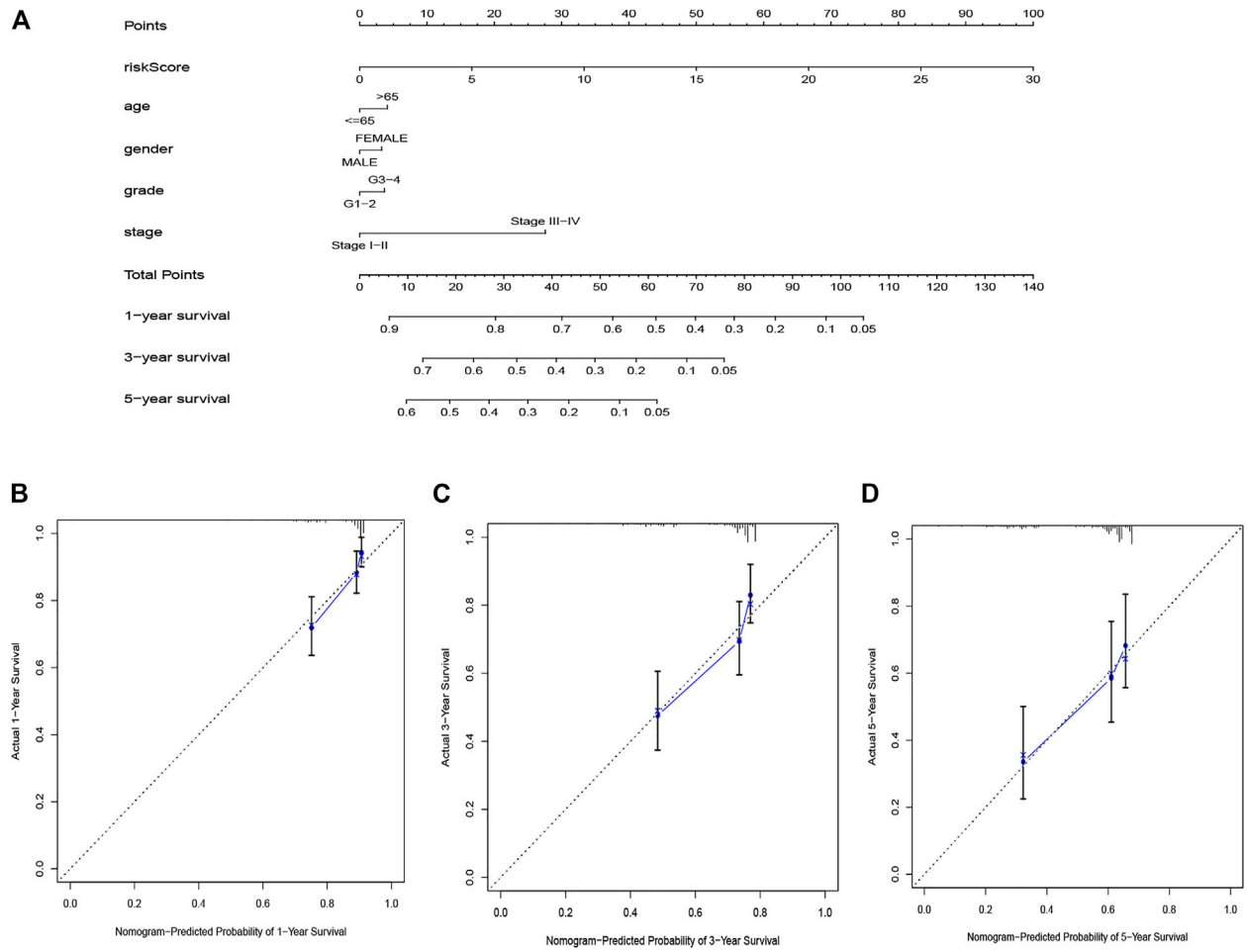


FIGURE 6 | The nomogram for predicting HCC patients' survival **(A)**. Calibration plots applied for predicting the 1-year **(B)**, 3-year **(C)** and 5-year **(D)** survival.

proliferation (Kathagen-Buhmann et al., 2016). Besides, found that hypoxia and the hypoxia inducible factors (HIFs) step in the metabolic crosstalk between tumor cells and their microenvironment (Mucaj et al., 2012). Then, we established a prognostic signature based on three hypoxia-related lncRNAs (AC099850.4, MIR210HG, and NRAV). Wang et al. revealed that high expression of MIR210HG in hepatocellular carcinoma samples and cells not only associated with clinicopathological characteristics, for example, stage, tumor size, vascular invasion, and histological differentiation but also with poor prognosis (Wang et al., 2019). Du et al. revealed MIR210HG is a hypoxia-related lncRNA in triple-negative breast cancer and promotes the Warburg effect by directly regulating HIF-1α (Du et al., 2020). Besides, overexpression of MIR210HG promotes the progression of colorectal adenocarcinoma *via* modulating hypoxia (He et al., 2019) and was associated with worse overall survival in colon cancer patients (Ruan et al., 2019). AC099850.4 is one of the top lncRNA in miRNA–lncRNA network and may be associated with tumor division, tumor proliferation, cell cycle and so on in ovarian cancer (Zhao et al., 2020). Higher expression of NRAV was associated with advanced clinical stage, poor prognosis and immunologic characteristics in hepatocellular carcinoma patients (Feng et al., 2021; Xu et al., 2021). However, to date, there are few studies on these lncRNAs in hepatocellular carcinoma under hypoxia. These findings may offer crucial insights into the cancers' control in the future.

Herein, we divided the samples into the two risk subgroups on the foundation of the hypoxia-related lncRNA signature to explore the underlying roles in HCC. Hypoxia acts as an important role in anti-tumor immune responses *via* mediating the expression levels of immune checkpoints (Multhoff and Vaupel, 2020). There are very little studies exploring the connection between immune checkpoint inhibitors and hypoxia at present. We compared the difference of immune checkpoints between two risk group and discovered that most of them exist remarkably difference, which may offer underlying therapeutic targets for hepatocellular carcinoma. More and more studies show that miRNA and lncRNA act as an important role in hypoxia. Hypoxia-inducible factor (HIF) is a critical to cellular

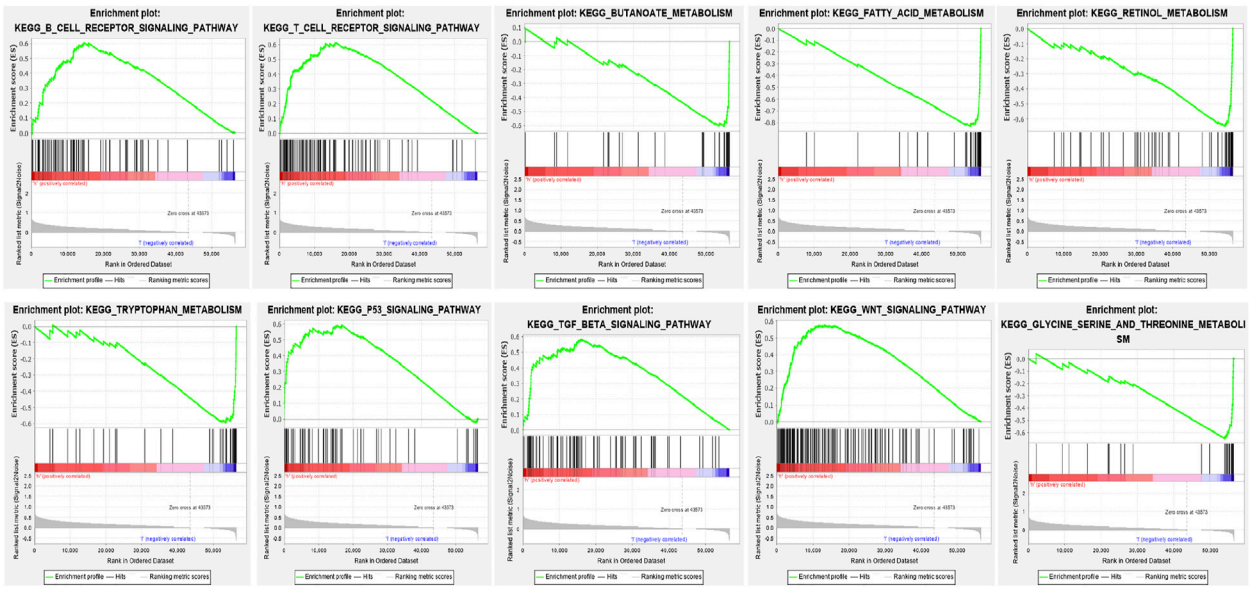


FIGURE 7 | GSEA for two risk groups based on the signature.

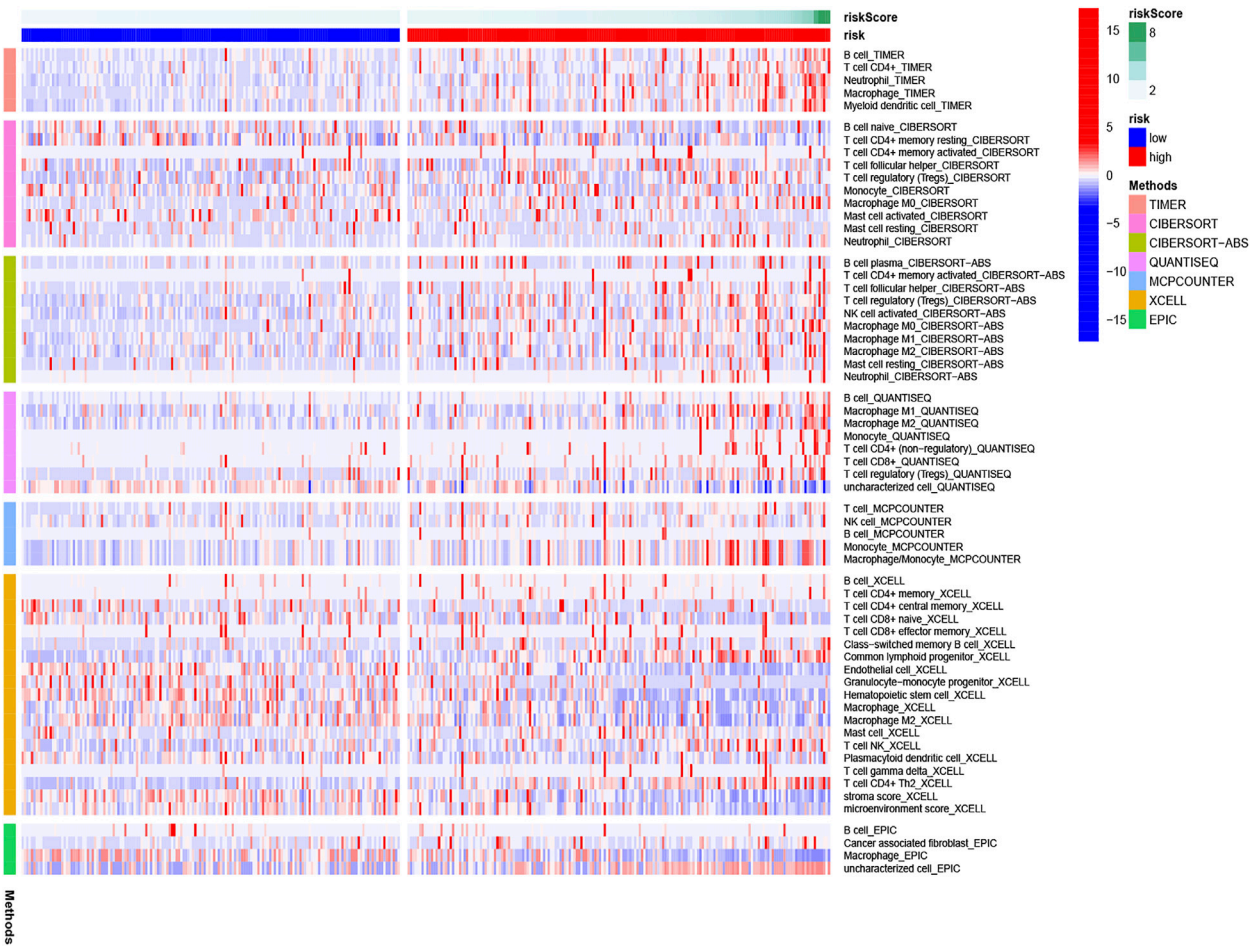


FIGURE 8 | Heatmap for immune cells based on EPIC, CIBERSORT–ABS, CIBERSORT, QUANTISEQ, XCELL, MCP counter, and TIMER algorithms between two risk groups.

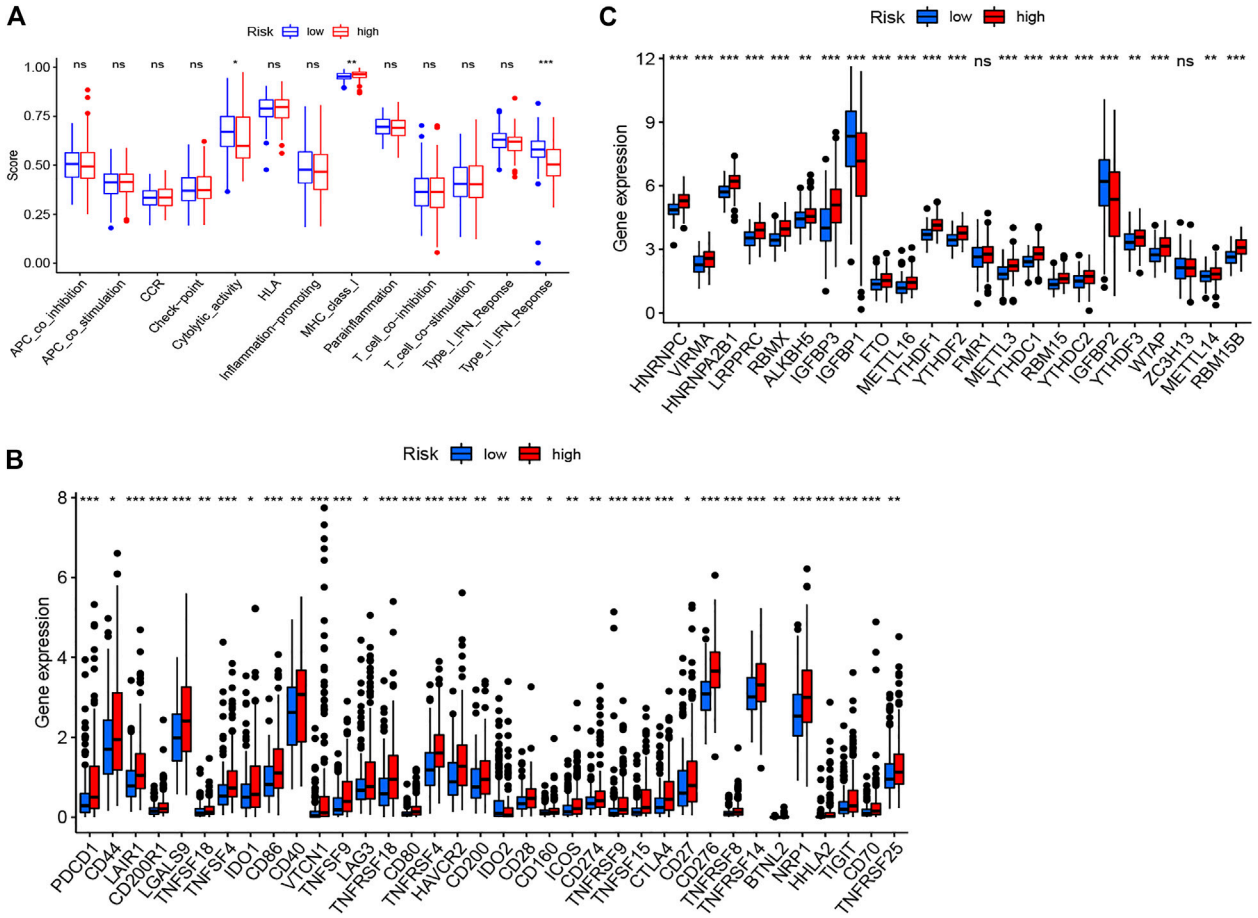


FIGURE 9 | The immune cells related functions **(A)**, immune checkpoints **(B)** and m6A-related genes **(C)** between two risk groups.

responses to hypoxic stress. LncRNA SNHG11 promotes invasion and metastasis of colorectal cancer cells by regulating HIF-1α (Xu et al., 2020b). Hypoxia activated lncRNA HABON enhances the proliferation and growth of hepatocarcinoma cells by mediating HIF-1α (Ma et al., 2021). Hypoxia-inducible miR-182 accelerates tumor angiogenesis and growth by regulating the expression level of HIF1α in prostate cancer cells (Li et al., 2015).

Besides, hypoxia can change the tumor cells' interplay and crosstalk with the tumor micro-environment, resulting to immune suppression and resistance, which contribute tumor cells to escape immune surveillance (Jing et al., 2019; Noman et al., 2019). To explore the relationship between our signature and tumor micro-environment, we fulfilled GSEA. The outcomes revealed that various immune-related pathways participated in the high-risk group, while the metabolism-related pathways participated in the low-risk group. We next plotted the heatmap of immune cell infiltration of HCC patients to explore the difference of immune microenvironment between two risk groups by using multiple algorithms. We found that T cell follicular helper, NK cell, macrophage M0, macrophages M2, neutrophil, and mast cells resting were highly infiltrated in HCC, associating with tumorigenesis, progression, and metastasis (Zhang et al., 2017; Aponte-López and Muñoz-Cruz, 2020; Ngabire et al., 2020). These results showed that this hypoxia-related lncRNA signature can partially reflect tumor immune cell infiltration and may offer some useful information for immunotherapies. Moreover, majority of m6a-related genes were highly expressed in high risk group, which is similar to previous studies. Chen et al. revealed that WTAP was extremely overexpressed and regarded as an independent prognostic biomarker for HCC patients (Chen et al., 2019). It was reported that FTO was over-expressed in HCC tissue and associated with poor prognosis of HCC patients as well as promoted cell proliferation by mediating the demethylation of PKM2 (Li et al., 2019). YTHDF1 was remarkably up-regulated and positively associated with pathology stage in HCC. K-M curve revealed that higher expression of YTHDF1 is correlated with worse survival (Zhao et al., 2018). What is more, most of immune checkpoints were over-expressed in high risk group, which may provide some underlying targets for immunotherapy. Some of them have proved correlated with poor prognosis in HCC, such as LAG3 (Yarchoan et al., 2017), NRP1 (Lin et al., 2018) and LAIR1 (Wu et al., 2019).

The data of this analysis was downloaded from TCGA database, including complete survival and clinicopathological information of most patients of HCC. Besides, the HCC samples were sufficient enough to divide into train set and test set. Thus, the prognostic signature was constructed and validated using the patients from the single database. However, there were some limitations. First, we did not validate the prognostic model in external database and lack of the validation in clinical practice without enough HCC samples. Second, we did not perform basic experiment to verify the lncRNAs in this study.

CONCLUSION

Specific hypoxia associated lncRNAs provide a novel prognostic signature for hepatocellular carcinoma patients and may improve the individualized treatment strategies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

CZ designed the project and wrote the manuscript; HZ support data analysis and discussion. LL guided the project and critically reviewed the manuscript. All authors have contributed to the preparation of the manuscript and reviewed/approved it in its final form.

FUNDING

CZ designed the project and wrote the manuscript; HZ support data analysis and discussion. LL guided the project and critically reviewed the manuscript. All authors have contributed to the preparation of the manuscript and reviewed/approved it in its final form. CZ was financially supported by CSC scholarship (The China Scholarship Council).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.744113/full#supplementary-material.

REFERENCES

Allemani, C., Matsuda, T., Di Carlo, V., Harewood, R., Matz, M., Nikšić, M., et al. (2018). Global Surveillance of Trends in Cancer Survival 2000-14 (CONCORD-3): Analysis of Individual Records for 37 513 025 Patients Diagnosed with One of 18 Cancers from 322 Population-Based Registries in 71 Countries. *Lancet* 391 (10125), 1023–1075. doi:10.1016/S0140-6736(17)33326-3

Aponte-López, A., and Muñoz-Cruz, S. (2020). Mast Cells in the Tumor Microenvironment. *Adv. Exp. Med. Biol.* 1273, 159–173. doi:10.1007/978-3-030-49270-0_9

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J. Clinicians* 68 (6), 394–424. doi:10.3322/caac.21492

Chen, Q., Hu, L., and Chen, K. (2020). Construction of a Nomogram Based on a Hypoxia-Related lncRNA Signature to Improve the Prediction of Gastric Cancer Prognosis. *Front. Genet.* 11, 570325. doi:10.3389/fgene.2020.570325

Chen, Y., Peng, C., Chen, J., Chen, D., Yang, B., He, B., et al. (2019). WTAP Facilitates Progression of Hepatocellular Carcinoma via m6A-HuR-dependent Epigenetic Silencing of ETS1. *Mol. Cancer* 18 (1), 127. doi:10.1186/s12943-019-1053-8

Du, Y., Wei, N., Ma, R., Jiang, S.-H., and Song, D. (2020). Long Noncoding RNA MIR210HG Promotes the Warburg Effect and Tumor Growth by Enhancing HIF-1α Translation in Triple-Negative Breast Cancer. *Front. Oncol.* 10, 580176. doi:10.3389/fonc.2020.580176

Eales, K. L., Hollinshead, K. E. R., and Tennant, D. A. (2016). Hypoxia and Metabolic Adaptation of Cancer Cells. *Oncogenesis* 5–e190, e190. doi:10.1038/oncsis.2015.50

Erler, J. T., and Giaccia, A. J. (2006). Lysyl Oxidase Mediates Hypoxic Control of Metastasis: Figure 1. *Cancer Res.* 66 (21), 10238–10241. doi:10.1158/0008-5472.CAN-06-3197

Feng, Y., Hu, X., Ma, K., Zhang, B., and Sun, C. (2021). Genome-Wide Screening Identifies Prognostic Long Noncoding RNAs in Hepatocellular Carcinoma. *Biomed. Res. Int.* 2021, 1–16. doi:10.1155/2021/6640652

Galle, P. R., Tovoli, F., Foerster, F., Wörns, M. A., Cucchetti, A., and Bolondi, L. (2017). The Treatment of Intermediate Stage Tumours beyond TACE: From Surgery to Systemic Therapy. *J. Hepatol.* 67 (1), 173–183. doi:10.1016/j.jhep.2017.03.007

Gao, K., Lv, C.-L., Zhang, T., Zhang, K. Z.-L., and Lv, C. L. (2021). LncRNA Airsci Increases the Inflammatory Response after Spinal Cord Injury in Rats through the Nuclear Factor Kappa B Signaling Pathway. *Neural Regen. Res.* 16 (4), 772–777. doi:10.4103/1673-5374.295335

Graham, K., and Unger, E. (2018). Overcoming Tumor Hypoxia as a Barrier to Radiotherapy, Chemotherapy and Immunotherapy in Cancer Treatment. *Ijn* 13, 6049–6058. doi:10.2147/IJN.S140462

He, Z., Dang, J., Song, A., Cui, X., Ma, Z., and Zhang, Z. (2019). Identification of LINC01234 and MIR210HG as Novel Prognostic Signature for Colorectal Adenocarcinoma. *J. Cel Physiol* 234 (5), 6769–6777. doi:10.1002/jcp.27424

Hill, R. P., Bristow, R. G., Fyles, A., Koritzinsky, M., Milosevic, M., and Wouters, B. G. (2015). Hypoxia and Predicting Radiation Response. *Semin. Radiat. Oncol.* 25 (4), 260–272. doi:10.1016/j.semradonc.2015.05.004

Huang, R., and Zong, X. (2017). Aberrant Cancer Metabolism in Epithelial-Mesenchymal Transition and Cancer Metastasis: Mechanisms in Cancer Progression. *Crit. Rev. Oncology/Hematology* 115, 13–22. doi:10.1016/j.critrevonc.2017.04.005

Jing, X., Yang, F., Shao, C., Wei, K., Xie, M., Shen, H., et al. (2019). Role of Hypoxia in Cancer Therapy by Regulating the Tumor Microenvironment. *Mol. Cancer* 18 (1), 157. doi:10.1186/s12943-019-1089-9

Kang, C. L., Qi, B., Cai, Q. Q., Fu, L. S., Yang, Y., Tang, C., et al. (2019). LncRNA AY Promotes Hepatocellular Carcinoma Metastasis by Stimulating ITGAV Transcription. *Theranostics* 9 (15), 4421–4436. doi:10.7150/thno.32854

Kathagen-Buhmann, A., Schulte, A., Weller, J., Holz, M., Herold-Mende, C., Glass, R., et al. (2016). Glycolysis and the Pentose Phosphate Pathway Are Differentially Associated with the Dichotomous Regulation of Glioblastoma Cell Migration versus Proliferation. *Neuonc* 18 (9), 1219–1229. doi:10.1093/neuonc/now024

Krock, B. L., Skuli, N., and Simon, M. C. (2011). Hypoxia-induced Angiogenesis: Good and Evil. *Genes & Cancer* 2 (12), 1117–1133. doi:10.1177/1947601911423654

Lan, T., Li, H., Zhang, D., Xu, L., Liu, H., Hao, X., et al. (2019). KIAA1429 Contributes to Liver Cancer Progression through N6-methyladenosine-dependent post-transcriptional Modification of GATA3. *Mol. Cancer* 18 (1), 186. doi:10.1186/s12943-019-1106-z

Li, J., Zhu, L., Shi, Y., Liu, J., Lin, L., and Chen, X. (2019). m6A Demethylase FTO Promotes Hepatocellular Carcinoma Tumorigenesis via Mediating PKM2 Demethylation. *Am. J. Transl Res.* 11 (9), 6084–6092.

Li, Y., Zhang, D., Wang, X., Yao, X., Ye, C., Zhang, S., et al. (2015). Hypoxia-inducible miR-182 Enhances HIF1α Signaling via Targeting PHD2 and FIH1 in Prostate Cancer. *Sci. Rep.* 5, 12495. doi:10.1038/srep12495

Lin, J., Zhang, Y., Wu, J., Li, L., Chen, N., Ni, P., et al. (2018). Neuropilin 1 (NRP1) Is a Novel Tumor Marker in Hepatocellular Carcinoma. *Clinica Chim. Acta* 485, 158–165. doi:10.1016/j.cca.2018.06.046

Ma, C.-N., Wo, L.-L., Wang, D.-F., Zhou, C.-X., Li, J.-C., Zhang, X., et al. (2021). Hypoxia Activated Long Non-coding RNA HABON Regulates the Growth and Proliferation of Hepatocarcinoma Cells by Binding to and Antagonizing HIF-1 Alpha. *RNA Biol.* 21 (1), 1–16. doi:10.1080/15476286.2020.1871215

Maluccio, M., and Covey, A. (2012). Recent Progress in Understanding, Diagnosing, and Treating Hepatocellular Carcinoma. *CA: A Cancer J. Clinicians* 62 (6), 394–399. doi:10.3322/caac.21161

Mucaj, V., Shay, J. E. S., and Simon, M. C. (2012). Effects of Hypoxia and HIFs on Cancer Metabolism. *Int. J. Hematol.* 95 (5), 464–470. doi:10.1007/s12185-012-1070-5

Multhoff, G., and Vaupel, P. (2020). Hypoxia Compromises Anti-cancer Immune Responses. *Adv. Exp. Med. Biol.* 1232, 131–143. doi:10.1007/978-3-030-34461-0_18

Muz, B., de la Puente, P., Azab, F., and Azab, A. K. (2015). The Role of Hypoxia in Cancer Progression, Angiogenesis, Metastasis, and Resistance to Therapy. *Hp* 3, 83–92. doi:10.2147/HP.S93413

Ngabire, D., Niyonizigiye, I., Patil, M. P., Seong, Y.-A., Seo, Y. B., and Kim, G.-D. (2020). M2 Macrophages Mediate the Resistance of Gastric Adenocarcinoma Cells to 5-Fluorouracil through the Expression of Integrin β3, Focal Adhesion Kinase, and Cofilin. *J. Immunol. Res.* 2020, 1–9. doi:10.1155/2020/1731457

Nishida, N., and Kudo, M. (2013). Oxidative Stress and Epigenetic Instability in Human Hepatocarcinogenesis. *Dig. Dis.* 31 (5-6), 447–453. doi:10.1159/000355243

Noman, M. Z., Hasmim, M., Lequeux, A., Xiao, M., Duhem, C., Chouaib, S., et al. (2019). Improving Cancer Immunotherapy by Targeting the Hypoxic Tumor Microenvironment: New Opportunities and Challenges. *Cells* 8 (9), 1083. doi:10.3390/cells8091083

Ruan, Z., Xu, Z., Li, Z., and Lv, Y. (2019). Integral Analyses of Survival-related L-ong N-on-coding RNA MIR210HG and its Prognostic Role in colon Cancer. *Oncol. Lett.* 18 (2), 1107–1116. doi:10.3892/ol.2019.10435

Sun, J., Zhao, T., Zhao, D., Qi, X., Bao, X., Shi, R., et al. (2020). Development and Validation of a Hypoxia-Related Gene Signature to Predict Overall Survival in Early-Stage Lung Adenocarcinoma Patients. *Ther. Adv. Med. Oncol.* 12, 175883592093790. doi:10.1177/1758835920937904

Vilgrain, V., Pereira, H., Assenat, E., Guiu, B., Ilonca, A. D., Pageaux, G.-P., et al. (2017). Efficacy and Safety of Selective Internal Radiotherapy with Yttrium-90 Resin Microspheres Compared with Sorafenib in Locally Advanced and Inoperable Hepatocellular Carcinoma (SARAH): an Open-Label Randomised Controlled Phase 3 Trial. *Lancet Oncol.* 18 (12), 1624–1636. doi:10.1016/S1470-2045(17)30683-6

Wang, Y., Li, W., Chen, X., Li, Y., Wen, P., and Xu, F. (2019). MIR210HG Predicts Poor Prognosis and Functions as an Oncogenic lncRNA in Hepatocellular Carcinoma. *Biomed. Pharmacother.* 111, 1297–1301. doi:10.1016/j.biopha.2018.12.134

Wu, J., Pang, R., Li, M., Chen, B., Huang, J., and Zhu, Y. (2021). m6A-Induced LncRNA MEG3 Suppresses the Proliferation, Migration and Invasion of Hepatocellular Carcinoma Cell through miR-544b/BTG2 Signaling. *Ott* 14, 3745–3755. doi:10.2147/OTT.S289198

Wu, X.-Z., Xie, G.-R., and Chen, D. (2007). Hypoxia and Hepatocellular Carcinoma: The Therapeutic Target for Hepatocellular Carcinoma. *J. Gastroenterol. Hepatol.* 22 (8), 1178–1182. doi:10.1111/j.1440-1746.2007.04997.x

Wu, X., Zhang, L., Zhou, J., Liu, L., Fu, Q., Fu, A., et al. (2019). Clinicopathologic Significance of LAIR-1 Expression in Hepatocellular Carcinoma. *Curr. Probl. Cancer* 43 (1), 18–26. doi:10.1016/j.currproblcancer.2018.04.005

Wu, X., Zhang, X., Tao, L., Dai, X., and Chen, P. (2020). Prognostic Value of an m6A RNA Methylation Regulator-Based Signature in Patients with Hepatocellular Carcinoma. *Biomed. Res. Int.* 2020, 1–11. doi:10.1155/2020/2053902

Wu, Z., Zuo, M., Zeng, L., Cui, K., Liu, B., Yan, C., et al. (2021). OMA1 Reprograms Metabolism under Hypoxia to Promote Colorectal Cancer Development. *EMBO Rep.* 22 (1), e50827. doi:10.15252/embr.202050827

Xiao, B., Liu, L., Chen, Z., Li, A., Wang, P., Xiang, C., et al. (2021). Identification of Epithelial-Mesenchymal Transition-Related Prognostic lncRNAs Biomarkers Associated with Melanoma Microenvironment. *Front. Cel Dev. Biol.* 9, 679133. doi:10.3389/fcell.2021.679133

Xiao, B., Liu, L., Li, A., Xiang, C., Wang, P., Li, H., et al. (2020). Identification and Verification of Immune-Related Gene Prognostic Signature Based on ssGSEA for Osteosarcoma. *Front. Oncol.* 10, 607622. doi:10.3389/fonc.2020.607622

Xu, L., Huan, L., Guo, T., Wu, Y., Liu, Y., Wang, Q., et al. (2020). LncRNA SNHG11 Facilitates Tumor Metastasis by Interacting with and Stabilizing HIF-1α. *Oncogene* 39 (46), 7005–7018. doi:10.1038/s41388-020-01512-8

Xu, Q., Wang, Y., and Huang, W. (2021). Identification of Immune-Related lncRNA Signature for Predicting Immune Checkpoint Blockade and Prognosis in Hepatocellular Carcinoma. *Int. Immunopharmacology* 92, 107333. doi:10.1016/j.intimp.2020.107333

Xu, Y., Sun, C., Han, B., Xi, Y., Zhang, M., Yang, J., et al. (2020). High KIAA1522 Expression Predicts a Poor Prognosis in Patients with Hepatocellular Carcinoma. *Oncol. Lett.* 20 (1), 509–516. doi:10.3892/ol.2020.11588

Yarchoan, M., Xing, D., Luan, L., Xu, H., Sharma, R. B., Popovic, A., et al. (2017). Characterization of the Immune Microenvironment in Hepatocellular Carcinoma. *Clin. Cancer Res.* 23 (23), 7333–7339. doi:10.1158/1078-0432.CCR-17-0950

Zhang, H., Qin, C., Liu, H. W., Guo, X., and Gan, H. (2021). An Effective Hypoxia-Related Long Non-coding RNAs Assessment Model for Prognosis of Clear Cell Renal Carcinoma. *Front. Oncol.* 11, 616722. doi:10.3389/fonc.2021.616722

Zhang, H., Yue, R., Zhao, P., Yu, X., Li, J., Ma, G., et al. (2017). Proinflammatory Follicular Helper T Cells Promote Immunoglobulin G Secretion, Suppress Regulatory B Cell Development, and Correlate with Worse Clinical Outcomes in Gastric Cancer. *Tumour Biol.* 39 (6), 101042831770574. doi:10.1177/1010428317705747

Zhao, J., Song, X., Xu, T., Yang, Q., Liu, J., Jiang, B., et al. (2020). Identification of Potential Prognostic Competing Triplets in High-Grade Serous Ovarian Cancer. *Front. Genet.* 11, 607722. doi:10.3389/fgene.2020.607722

Zhao, X., Chen, Y., Mao, Q., Jiang, X., Jiang, W., Chen, J., et al. (2018). Overexpression of YTHDF1 Is Associated with Poor Prognosis in Patients with Hepatocellular Carcinoma. *Cbm* 21 (4), 859–868. doi:10.3233/CBM-170791

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, orclaim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Genetics

ORIGINAL RESEARCH
published: 04 October 2021
doi: 10.3389/fgene.2021.704712

Targeting Long Non-Coding RNA TTN-AS1 Suppresses Bladder Cancer Progression

Huiyuan Xiao[1†], *Wen Huang*[1†], *Yanlei Li*[2], *Rongxin Zhang*[3] *and Long Yang*[1*]

[1]Department of Urology, Tianjin Medical University General Hospital, Tianjin, China, [2]Department of Pathology, Tianjin Medical University, Tianjin, China, [3]Department of Radiotherapy, Tianjin Medical University General Hospital, Tianjin, China

Background: To explore the biological and clinical effects of titin-antisense RNA1 (TTN-AS1) in bladder cancer (BC) and the association between TTN-AS1 and activating transcription factor 2 (ATF2) in BC.

Methods: The Kaplan–Meier method was performed to analyze the association between the expression of TTN-AS1 and prognosis of BC patients from TCGA data set and our institution. Quantitative real-time PCR (RT-PCR) was conducted to explore the expression of TTN-AS1 between the patients who underwent TURBT and Re-TURBT. MTT, colony formation, and tumor formation assays were conducted to evaluate the effect of TTN-AS1 on the ability of proliferation in BC cell lines. Transwell assay was performed to evaluate the effect of TTN-AS1 on the ability of invasion in BC cell lines. Bioinfomatics and immunohistochemical staining was used to identify the relationship between TTN-AS1 and ATF2.

Results: The higher expression of TTN-AS1 was related to poorer disease-free survival (DFS) in patients with BC. The expression of TTN-AS1 was higher in BC patients who underwent Re-TURBT compared with BC patients who underwent TURBT. Knocking down TTN-AS1 resulted in inhibiting the ability of proliferation and invasion of BC cells. ATF2 may serve as a downstream target of TTN-AS1 in BC, and the high expression of ATF2 is also related to adverse DFS.

Conclusion: Our study reveals that TTN-AS1 serves as an oncogene by activating ATF2 in BC. The findings suggest that TTN-AS1 may act as a novel therapeutic target for patients with BC.

Keywords: long non-coding RNA, bladder cancer, prognosis, oncogene, bioinformactics

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Yuchen Liu,
Shenzhen University, China
Yuan Zhou,
Peking University, China
Xu Li,
China Agricultural University, China

***Correspondence:**
Long Yang
yanglong7996@tmu.edu.cn

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *03 May 2021*
Accepted: *26 August 2021*
Published: *04 October 2021*

Citation:
Xiao H, Huang W, Li Y, Zhang R and Yang L (2021) Targeting Long Non-Coding RNA TTN-AS1 Suppresses Bladder Cancer Progression. Front. Genet. 12:704712. doi: 10.3389/fgene.2021.704712

INTRODUCTION

Bladder cancer (BC) is one of the most frequency urinary malignancies in males with high recurrence and poor prognosis (Torre et al., 2015; Siegel et al., 2019). As one of the lethal causes in patients, BC is still a major issue for public health. In patients with non-muscle-invasive BC (NMIBC), the tumor is confined to the mucosa or submucosa of the bladder wall. However, if the tumor invades the muscle layer, it is called muscular invasive BC (MIBC). If the patient with MIBC does not develop distant metastasis or invade adjacent tissues, the bladder and the tumor should be considered to be surgically removed (Hinsenveld et al., 2021). With deepening research, more and

more efforts have been made to find effective therapeutic targets against BC improving patient outcomes.

Long noncoding RNA (lncRNA), which is a subtype of RNA defined as transcripts >200 nt without protein-coding ability, is explored often and suggested to have a regulatory role in the development and progression of BC (Lv et al., 2017; Li et al., 2018; Quan et al., 2018; Zhang et al., 2019). A recent study shows that LncRNA NR2F1-AS1 could serve as an independent prognostic factor in BC, and the researchers established a genomic-related nomogram to predict the prognosis of BC (Peng et al., 2021). LncRNA titin-antisense RNA1 (TTN-AS1) transcripts are from the antisense strand of TTN and are suggested to be oncogenic lncRNA in numerous tumors (Chen et al., 2018; Lin et al., 2018; Cui et al., 2019; Dong et al., 2019). However, the clinical relevance and functional role of lncRNA TTN-AS1 has not been investigated in BC. In the present study, we find that highly expressed lncRNA TTN-AS1 is correlated with BC progression. Importantly, lncRNA TTN-AS1 could be upregulated in re-TURBT BC specimens. Then, growth and invasion suppression of BC cells *in vitro* and *in vivo* by targeting lncRNA TTN-AS1 supports its potential therapeutic effects.

Activating transcription factor 2 (ATF2), a member of the AP1 transcription factor family, is demonstrated to be an inducer for BC cells (with functional androgen receptor) proliferation as well as invasion (Inoue et al., 2018). In this study, we explore the relationship between ATF2 and lncRNA TTN-AS1 *in vitro* and *in vivo*. We also demonstrate that ATF2 and lncRNA TTN-AS1 plays a vital role in the progression of BC.

MATERIALS AND METHODS

Ethical Approval for the Study Protocol

This study was approved by the Ethics Committee of Tianjin Medical University General Hospital, Tianjin, China (No. KY2020K09). Written informed consent was obtained from all patients, and the study was conducted in accordance with the Declaration of Helsinki.

Human Samples

The bladder tissue specimens were surgical specimens from BC patients with complete clinicopathological data. BC tissues that were embedded in paraffin were acquired by transurethral resection of bladder tumor (TURBT) at the Tianjin Medical University General Hospital (n = 130). All patients did not receive chemotherapy and radiotherapy before surgery. The fresh paired BC tissues which were recurred (recurrence means that a patient who was diagnosed with BC for the first time has a new tumor in the bladder after undergoing TURBT surgery) were immediately stored in liquid nitrogen after TURBT (n = 6).

Antibodies

The antibodies used in this study are listed as follows: ATF2 (Abcam, ab239361, 1:250 dilution for IHC) and Ki67 (Abcam, ab15580, 1:1,000 dilution for IHC).

Cell Lines and Cell Culture

Human BC cell line T24 and 5,637 cells were purchased from the American Tissue Culture Collection (ATCC). T24 and 5,637 cells were incubated in RPMI 1640 (Gibco, Waltham, MA United States). All media were supplemented with 10% fetal bovine serum (Gibco, Waltham, MA United States). All experimental cell lines were cultured at 37°C containing 5% CO_2.

In vitro Transfection

Transfection was performed using cells in the logarithmic growth phase. The cells are seeded into six-well plates at a density of 5 × 105 cells/well. After the cells grow to 60% of the area of each well, transfection is performed. The specific procedure was carried out with Lipofectamine 3,000 according to the manufacturer's instructions (Invitrogen, Thermo Scientific, MA, United States). The shRNA targeting TTN-AS1 (shTTN-AS1) and the shRNA targeting scramble (shSCR) were obtained from Shanghai GenePharma Co., Ltd (Shanghai, China).

Polymerase Chain Reaction

Total RNA was extracted using Trizol and reverse transcribed into cDNA. Using the cDNA as a template, PCR was performed to detect the expression of TTN-AS1 using GAPDH as an internal reference. The primer sequences were as follows: TTN-AS1 forward primer, 5′- GCCAGGTAGAGTTGCAGGTT-3′; reverse primer, 5′- GAAGCTGCTGCGGATGAATG-3′; GAPDH forward primer, 5′-GATTCCACCCATGGCAAATT-3′; reverse primer, 5′-TCTCGCTCCTGGAAGATGGT-3′. Reaction bands were observed using gel electrophoresis.

MTT Assay

Weigh 0.5 g of MTT [3-(4,5)-dimethylthiahiazo (-z-y1)-3,5-di-phenytetrazoliumromide] order and dissolve it in 100 ml of PBS and pass the solution through a sterile filter membrane with a pore size of 0.22 μm to filter out bacteria in the solution. Add 100 μl of cell culture medium to each well of the 96-well plate and ensure that there are 2 × 103 cells in each well. Place the paved 96-well plate in a constant temperature cell incubator containing 5% CO_2 at 37°C for culture. After a certain period of time, discard the medium in the well plate, add 100 μl of MTT solution to each experimental well and then place it in a cell culture incubator for 4 h. The supernatant was discarded again, and 150 μl of DMSO was added to each experimental well. Place the 96-well plate on a plate shaker and shake it slowly, and then place the solution to be checked in the microplate reader, measure the optical density at a wavelength of 570 nm, and record the optical density value.

Colony Formation Assay

When the cells are in the logarithmic growth phase, the BC cells were collected and cultured by a conventional cell culture method. When macroscopic colonies appeared in the Petri dish, the culture was terminated. Discard the supernatant, add 4% paraformaldehyde to fix the cells, and then stain the cells with GIMSA application staining solution. Slowly wash away the staining solution with running water and let the plate dry completely. Finally calculate the number of clones.

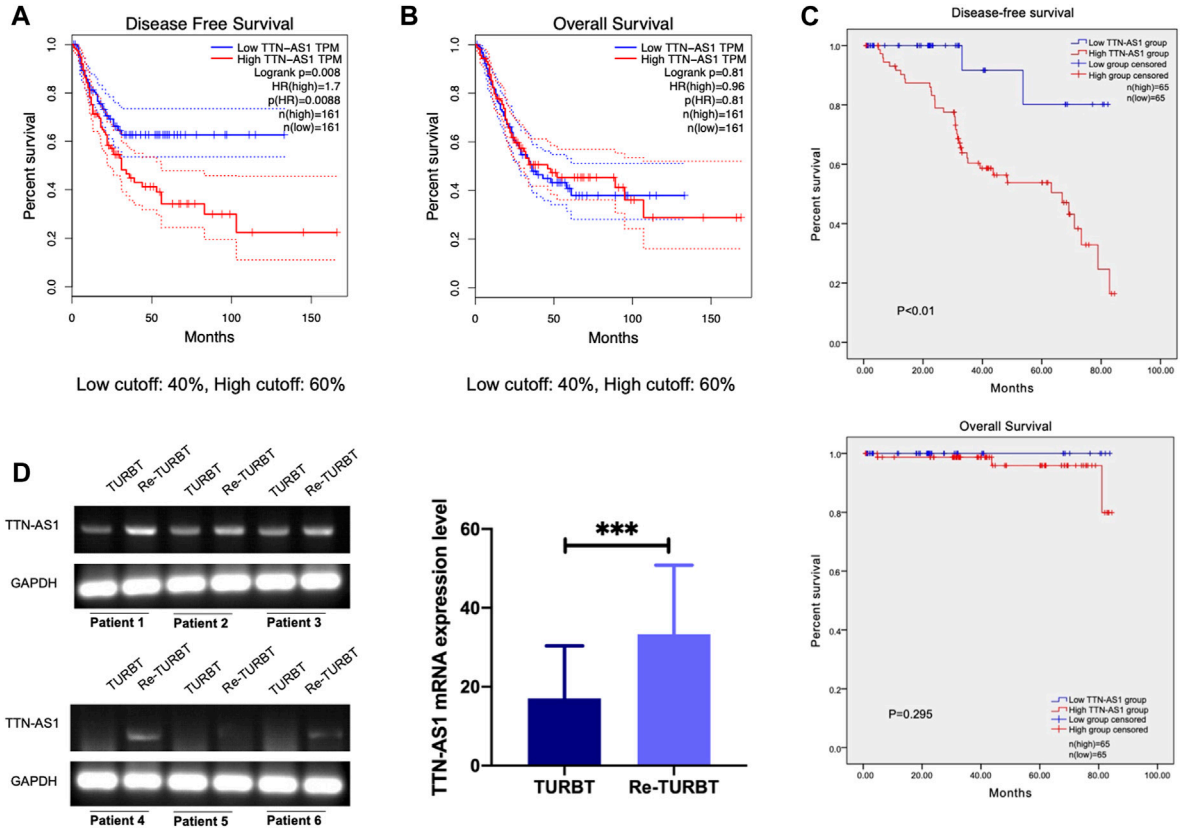


FIGURE 1 | LncRNA TTN-AS1 was upregulated in nonmuscular invasive BC who underwent TURBT and is related with adverse DFS **(A)**. The DFS of BC patients from TCGA database ($n = 322$, $p = 0.0088$) **(B)**. The OS of BC patients from TCGA database ($n = 322$, $p = 0.81$) **(C)** The DFS ($p < 0.01$) and OS ($p < 0.295$) of BC patients from our institution ($n = 130$) **(D)**. qRT-PCR was used to detect the expression level of lncRNA TTN-AS1 at mRNA level in surgical specimens of BC patients who underwent TURBT and Re-TURBT (*** means: $p < 0.001$).

Invasion Transwell Assay

Take 200 μl of the cell suspension and add it to the Transwell chamber to ensure that each well contains 1 × 104 cells. Perform cell culture in a conventional cell culture method. After 48 h, remove the chamber and discard the supernatant. Fix and stain the cells under the Transwell membrane; observe and count under a microscope.

In vivo Experiment

Male nude mice (8 weeks old, $n = 10$) were purchased from Beijing HFK Bioscience Co., Ltd (Beijing, China). Subcutaneous tumor growth assays were performed with T24 cell lines (1 × 106 T24 cells injected in each mouse). The tumor size was measured every 7 days with a Vernier calliper, and the volume change of the tumor was calculated. The tumors were harvested under standard, institutionally approved processes. The mass of each tumor was measured with an electronic balance. Tumor samples were paraffin-fixed and processed for immunohistochemical (IHC) analysis.

Immunohistochemistry

For IHC staining, paraffin-embedded tissue samples were deparaffinized, rehydrated, and incubated with primary antibodies (rabbit antihuman monoclonal antibody ATF2 antibody) overnight at 4°C. The secondary antibodies (goat antirabbit IgG) were incubated at room temperature for 1 h. The signals were examined under a microscope.

Statistical Analysis

Statistical significance was determined using two-tailed Student's *t*-test or ANOVA for functional analysis. Survival curves were plotted using Kaplan–Meier survival plots and a log-rank test was used to test significance. All statistical analyses were performed using SPSS software and GraphPad Prism7. $p < 0.05$ was considered statistically significant. $p < 0.05$ was marked as *, $p < 0.01$ was marked as **, $p < 0.001$ was marked as ***, and no significant difference was expressed as n.s.

RESULTS

Highly Expressed lncRNA TTN-AS1 Is Correlated With BC Progression

LncRNA TTN-AS1 has been stated and functionally confirmed to be involved in a variety of cancer progressions (Chang et al., 2020; Feng et al., 2020; Lin et al., 2020; Miao et al., 2020; Qi and Li, 2020). Although the role of TTN-AS1 in BC is still poorly

TABLE 1 | Clinicopathologic variables and TTN-AS1 expression in 130 bladder cancer patients.

Variables	All *n* = 130	TTN-AS1		*p* value[#]
		Low *n* = 35	High *n* = 95	
Age				
<65	61	16	45	0.87
≥65	69	19	50	
Sex				
Male	54	14	40	0.83
Female	76	21	55	
Tumor stage				
T2	60	25	35	≤0.01
T3/T4	70	10	60	
Tumor grade				
Low	60	22	38	0.02
High	70	13	57	
Lymph node metastasis				
No	50	20	30	≤0.01
Yes	80	15	65	
Vascular invasion				
No	70	25	45	0.01
Yes	60	10	50	

[a]#p *value was analyzed by Chi-square test; * indicates* p < *0.05 with statistical significance; iPSA means initial PSA.*

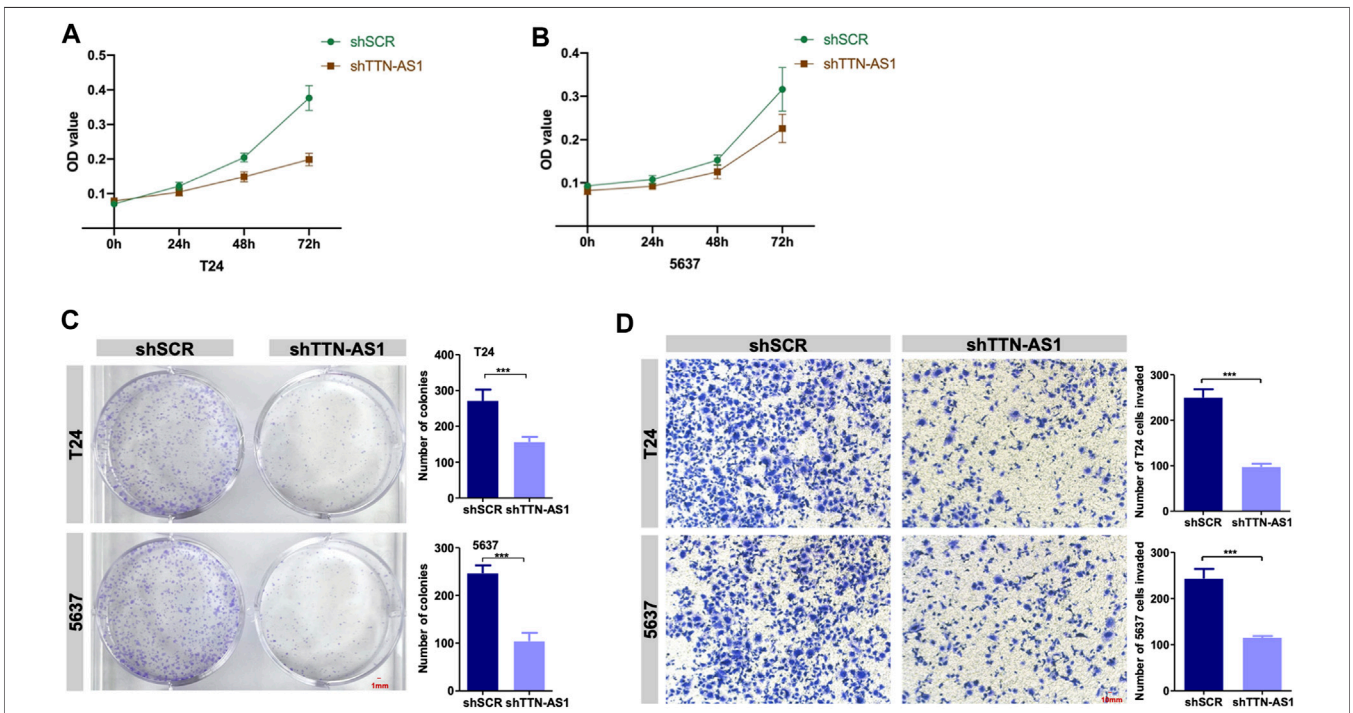


FIGURE 2 | Knockdown of lncRNA TTN-AS1 inhibits the proliferation and invasion of BC cells *in vitro*. **(A,B)** The MTT assay was used to explore the proliferation of two BC cell lines (T24 and 5,637) after the knockdown of lncRNA TTN-AS1. The absorbance value was detected at a wavelength of 490 nm (* means: $p < 0.05$, ** means: $p < 0.01$) **(C)** Colony formation assays was conducted to detected the abilities of proliferation in two BC cell lines (T24 and 5,637) after knocking down lncRNA TTN-AS1 and the number of colonies was counted and plotted on a graph (*** means: $p < 0.001$). **(D)** Transwell invasion assays were conducted to detect the invasiveness of the BC cell lines (T24 and 5,637) after knockdown of lncRNA TTN-AS1, and the number of invaded cells was counted and plotted on a graph (*** means: $p < 0.001$).

characterized. To investigate the clinical relevance of TTN-AS1 expression in BC, we retrieved public TCGA data sets from cBioPortal (Cerami et al., 2012; Gao et al., 2013) and evaluated TTN-AS1 expression in relation to disease-free survival (DFS) data as well as overall survival (OS) data on 322 BC patients. We found that BC patients with TTN-AS1 gene high expression had worse DFS (cutoff high expression: 60% of all patients; cutoff low expression: 40% of all patients)

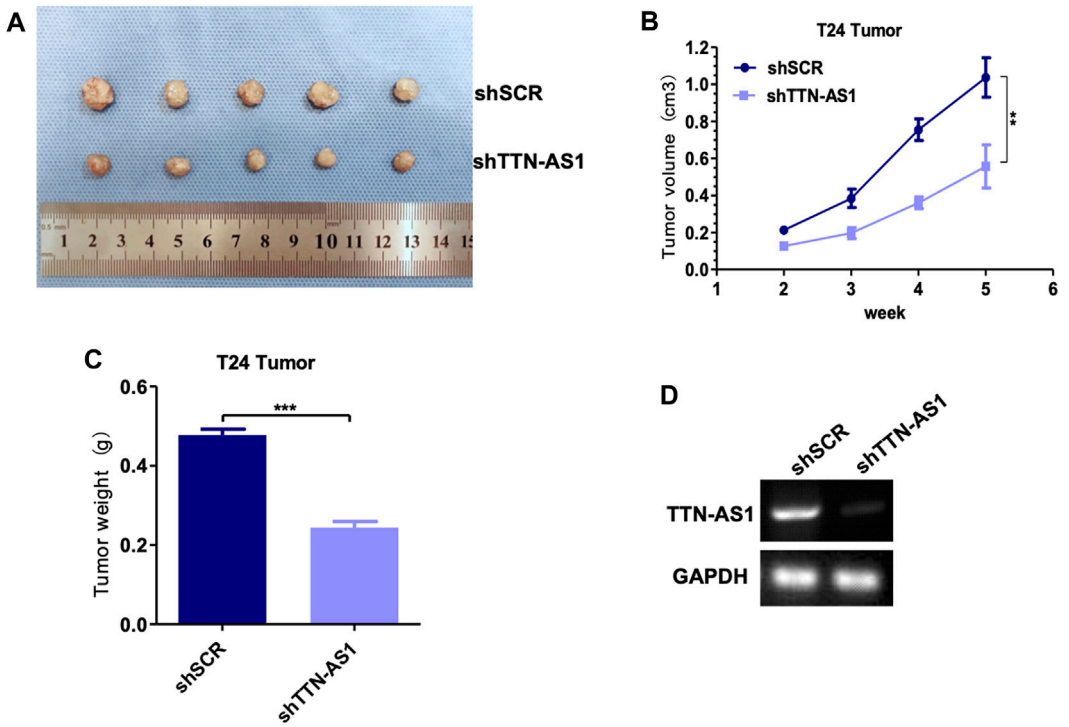


FIGURE 3 | Knocking down lncRNA TTN-AS1 inhibits the growth of bladder tumors *in vivo*. **(A)** The tumor size and morphology were compared between shSCR and shTTN-AS1 groups. **(B)** The tumor volume was analyzed in both shSCR and shTTN-AS1 groups. **(C)** The tumor weight was plotted on a graph (** means: $p < 0.001$). **(D)** qRT-PCR was used to detect the expression level of lncRNA TTN-AS1 at mRNA level from the tumor in both shSCR and shTTN-AS1 groups.

(**Figure 1A**), and its expression has no prediction of OS (**Figure 1B**). Except that the DFS of BC patients from our institution was also analyzed. As shown in **Figure 1C**, the TTN-AS1 high expression group had an adverse DFS, which was coincident with the TCGA database. The association between TTN-AS1 expression and clinical characters in BC patients was also analyzed by using chi-square test. As shown in **Table 1**, the high TTN-AS1 expression was related with advanced tumor stage ($p \leq 0.01$), high tumor grade ($p = 0.02$), lymph node metastasis ($p \leq 0.01$) and vascular invasion ($p = 0.01$).

Given the association with BC progression of highly expressed lncRNA TTN-AS1, we performed PCR detection on TURBT and paired re-TURBT specimens ($n = 6$) collected in our hospital. PCR results suggest patients received re-TURBT with recurrence may be correlated with local increased expression of TTN-AS1 (**Figure 1D**). These results led us to propose the clinical association and oncogenic role of TTN-AS1 in BC.

Silence of lncRNA TTN-AS1 Suppresses BC Cell Proliferation and Invasion

To investigate whether lncRNA TTN-AS1 participates in BC growth, a loss of function assay was performed to evaluate T24 and 5,637 cell growth upon TTN-AS1 silencing. Knockdown of TTN-AS1 in T24 and 5,637 cell lines resulted in significant inhibition of cell growth (**Figures 2A,B**). Consistent with the results in **Figures 2A,B**, colony assays confirmed the inhibition of cell proliferation in T24 and 5,637 cells after TTN-AS1 knockdown (**Figure 2C**). Transwell assays demonstrate the significantly decreased abilities of invasion (**Figure 2D**).

Targeting lncRNA TTN-AS1 Attenuates BC Growth *in vivo*

To explore the potential of targeting lncRNA TTN-AS1 in BC, we employed lentiviral injection in T24 tumor xenografts. T24 cells were xenografted into 8-week-old immunocompromised severe combined immunodeficiency (SCID) mice ($n = 10$). The scramble set (shSCR) ($n = 5$) was injected with lentiviruses carrying control scramble shRNA, and the treatment set (shTTN-AS1) ($n = 5$) was injected with lentiviruses carrying lncRNA TTN-AS1 shRNA. Injections were carried out every 3 days for 3 weeks lncRNA TTN-AS1 knocked down during this long-term treatment period led to a significant suppression in the growth of the T24 tumors (**Figures 3A–C**). The efficiency of silence for TTN-AS1 was validated by PCR in the scramble set and treatment set of tumors (**Figure 3D**). These results support the potential preclinical significance of targeting lncRNA TTN-AS1 in BC therapy.

ATF2 Acts as a Downstream Target of lncRNA TTN-AS1 in BC

To explore the downstream of lncRNA TTN-AS1 in BC, we first evaluated its similar expression pattern genes upon the TCGA BC

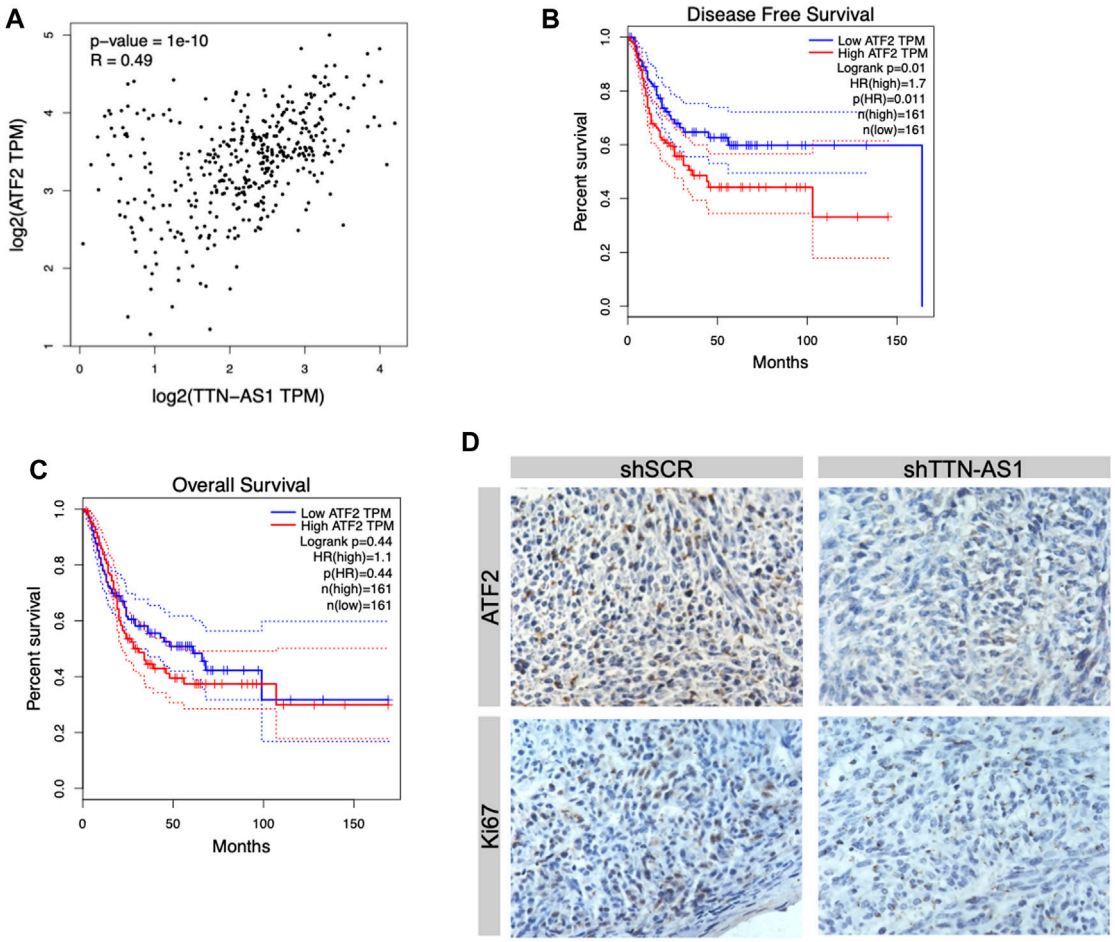


FIGURE 4 | ATF2 was the downstream factor of lncRNA TTN-AS1 and was related to poor prognosis in BC patients. **(A)** Bioinformatics was conducted to analyze the correlation between lncRNA TTN-AS1 and ATF2. **(B)** The DFS of BC patients from TCGA database (n = 322, p = 0.011) **(C)** The OS of BC patients from TCGA database (n = 322, p = 0.44) **(D)** IHC was performed to explore the expression of ATF2 in both shSCR and shTTN-AS1 groups.

data set by the GEPIA online tool (Tang et al., 2017). ATF2, a member of the AP1 transcription factor family, was listed as the top similar gene of TTN-AS1 (Hai and Curran, 1991; Shaulian and Karin, 2002). A previous study reveals that ATF2 protein is upregulated significantly in BC surgical specimens compared with normal urothelial tissues (Inoue et al., 2018). This study further shows that moderate/strong expression of ATF2 detected by IHC is correlated with poor prognosis of low-grade and muscle-invasive tumors (Inoue et al., 2018). To further investigate the expression relevance of lncRNA TTN-AS1 and ATF2 in BC specimens, we retrieved public TCGA data sets from cBioPortal (Cerami et al., 2012; Gao et al., 2013) and analyzed the two gene-expressed correlation coefficients. We found that these two gene-expressed levels were correlated significantly (**Figure 4A**). Then, we evaluated ATF2 expression in relation to DFS data as well as OS data on 322 BC patients. We found that BC patients with ATF2 gene high expression had worse DFS (cutoff high expression: 60% of all patients; cutoff low expression: 40% of all patients) (**Figure 4B**) although its expression has no prediction of OS (**Figure 4C**). To validate the downstream of lncRNA TTN-AS1, determination of ATF2 protein expression was performed by IHC in xenografts with silence of TTN-AS1 shown in **Figure 3A**. The results reveal that knockdown of lncRNA TTN-AS1 downregulated the expressed level of ATF2 and ki67 in our *in vivo* T24 tumor xenografts (**Figure 4D**).

DISCUSSION

LncRNA, the length of which was more than 200 nt, plays a crucial effect in modulating proliferation, invasion, and migration of various types of tumors, such BC (Li et al., 2020a), renal cancer (Liu et al., 2020), thyoroid cancer (Samimi et al., 2020), and so on. In recent years, more and more studies have found that lncRNA that is abnormally expressed in pan cancer plays an important role in different biological activities, such as autophagy (Cao et al., 2020), epithelial-mesenchymal transition (Wan et al., 2020), cell proliferation (Zhang et al., 2020) and apotosis (Lun et al., 2020). Numerous studies have pointed out that lncRNA plays a vital role in the carcinogenesis, drug resistance, and metastasis of BC. For

instance, lncRNA KCNQ1 opposite strand/antisense transcript 1 was found to be overexpressed in BC tissues compared with normal bladder tissues and can promote cell proliferation, migration, invasion, and apoptosis *via* modulating the miR-218–5p/HS3ST3B1 axis (Li et al., 2020a). Except that, Li et al. (2020b), demonstrate that overexpression of insulin-like growth factor binding protein 4-1 could activate the JAK/STAT pathway to promote BC cell proliferation, cell cycle, and cell apotosis. More and more studies have explored that lncRNA TTN-AS1 have participated in carcinogenesis and tumor development and served as a potential therapeutic target in many types of tumors. For example, TTN-AS1 regulates the TTN-AS1/miR-195/cyclin D1 axis through sponge adsorption so as to promote tumor cell proliferation and modulate tumor cell cycle activities in clear cell renal cell carcinoma (Lin et al., 2020). In breast cancer, the abnormally high expression of TTN-AS1 promotes the proliferation, migration, and invasion of tumor cells by inhibiting mir-524 and promoting the expression level of RRM2 (Feng et al., 2020).

So far, no relevant research has pointed out the role of TTN-AS1 in the occurrence and progression of nonmuscular invasive BC. In this study, we found, for the first time, that TTN-AS1 is abnormally highly expressed in BC and is related to the poor prognosis of BLACA patients. In terms of mechanism, TTN-AS1 can promote the proliferation and invasion of tumor cells by regulating the level of downstream factor ATF2. In short, we find that TTN-AS1 can exhibit the role of oncogenes and serve as a therapeutic target for BC through *in vitro* experiments, *in vivo* experiments, and clinical data analysis.

LncRNA regulates tumor mechanisms in a variety of ways, including affecting the transcription of the upstream promoter region of the coding protein gene, mediating chromatin remodeling, histone modification, and changing the cellular localization of proteins, etc., (Chen et al., 2020). ATF2 is reported to play two roles in different types of tumors during tumor progression, namely, tumor-promoting genes and tumor-suppressor factors (Lopez-Bergami et al., 2010). In Inoue's study, they found that ATF2 acted as an oncogene that promoted cell viability, migration, and invasion of BC *via* the AR signal pathway. Our study analyzes the survival information of BC patients in the TCGA database by bioinformatics and also finds that BC patients in the ATF2 high expression group had adverse DFS. Interestingly, we found that TTN-AS1 and ATF3 were correlated significantly via bioinformatics. Therefore, we propose the hypothesis that ATF2 may be a downstream factor of TTN-AS1 in BC. Furthermore, we used IHC to analyze the protein expression level of ATF2 in tumor tissues in both the knockdown TTN-AS1 and control groups. We find that, after knocking down TTN-AS1, the protein level of ATF2 was also significantly downregulated. Therefore, the oncogenic factor ATF2 of BC may be a potential downstream factor of TTN-AS1. TTN-AS1 promotes the proliferation and invasion of BC cells via inducing the expression of ATF2.

In summary, TTN-AS1 plays a role as a tumor-promoting factor in the progression of BC and can positively regulate the proliferation and invasion of BC cells by regulating the expression of ATF2. At the same time, TTN-AS1 is related to the poor prognosis of BC patients. Therefore, TTN-AS1 can be used as both a potential therapeutic target for Ponge and a predictor to evaluate the prognosis of patients with BC. In addition, this study contains the following shortcoming: the mechanism through which TTN-AS1 regulates the expression of its downstream factor ATF2 requires further experiments to explore.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Tianjin Medical University General Hospital. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by the Ethics Committee of Tianjin Medical University General Hospital.

AUTHOR CONTRIBUTIONS

LY: Project development and manuscript editing; HX: Manuscript writing; YL: Data collection and analysis; RZ: Reviewed the manuscript. All authors contributed to the article.

REFERENCES

Cao, Y., Pan, L., Zhang, X., Guo, W., and Huang, D. (2020). LncRNA SNHG3 Promotes Autophagy-Induced Neuronal Cell Apoptosis by Acting as a ceRNA for miR-485 to Up-Regulate ATG7 Expression. *Metab Brain Dis.* 35, 1361–1369.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discov.* 2 (5), 401–404. doi:10.1158/2159-8290.cd-12-0095

Chang, K., Wang, G., Lou, J., Hao, S., Lv, R., Duan, D., et al. (2020). lncRNA TTN-AS1 Upregulates RUNX1 to Enhance Glioma Progression via Sponging miR-27b-3p. *Oncol. Rep.* 44 (3), 1064–1074. doi:10.3892/or.2020.7684

Chen, P., Wang, R., Yue, Q., and Hao, M. (2018). Long Non-coding RNA TTN-AS1 Promotes Cell Growth and Metastasis in Cervical Cancer via miR-573/E2F3. *Biochem. Biophys. Res. Commun.* 503 (4), 2956–2962. doi:10.1016/j.bbrc.2018.08.077

Chen, W., Yang, J., Fang, H., Li, L., and Sun, J. (2020). Relevance Function of Linc-ROR in the Pathogenesis of Cancer. *Front. Cel Dev. Biol.* 8, 696. doi:10.3389/fcell.2020.00696

Cui, Z., Luo, Z., Lin, Z., Shi, L., Hong, Y., and Yan, C. (2019). Long Non-coding RNA TTN-AS1 Facilitates Tumorigenesis of Papillary Thyroid Cancer through Modulating the miR-153-3p/ZNRF2 axis. *J. Gene Med.* 21 (5), e3083. doi:10.1002/jgm.3083

Dong, M. M., Peng, S. J., Yuan, Y. N., and Luo, H. P. (2019). LncRNA TTN-AS1 Contributes to Gastric Cancer Progression by Acting as a Competing Endogenous RNA of miR-376b-3p. *Neoplasma* 66 (4), 564–575. doi:10.4149/neo_2018_180927N721

Feng, H., Wang, Q., Xiao, W., Zhang, B., Jin, Y., and Lu, H. (2020). LncRNA TTN-AS1 Regulates miR-524-5p and RRM2 to Promote Breast Cancer Progression. *Ott* 13, 4799–4811. doi:10.2147/ott.s243482

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* 6 (269), pl1. doi:10.1126/scisignal.2004088

Hai, T., and Curran, T. (1991). Cross-family Dimerization of Transcription Factors Fos/Jun and ATF/CREB Alters DNA Binding Specificity. *Proc. Natl. Acad. Sci.* 88 (9), 3720–3724. doi:10.1073/pnas.88.9.3720

Hinsenveld, F. J., Boormans, J. L., Van der poel, van der Schoot, D. K. E., Vis, A. N., and Aben, K. K. H., (2021). Intermediate-term Survival of Robot-Assisted versus Open Radical Cystectomy for Muscle-Invasive and High-Risk Non-muscle Invasive Bladder Cancer in The Netherlands.*Urol Oncol.* S1078-1439 (21), 00277–5. doi:10.1016/j.urolonc.2021.06.018

Inoue, S., Mizushima, T., Ide, H., Jiang, G., Goto, T., Nagata, Y., et al. (2018). ATF2 Promotes Urothelial Cancer Outgrowth via Cooperation with Androgen Receptor Signaling. *Endocr. Connect.* 7 (12), 1397–1408. doi:10.1530/ec-18-0364

Li, Z., Hong, S., and Liu, Z. (2018). LncRNA LINC00641 Predicts Prognosis and Inhibits Bladder Cancer Progression through miR-197-3p/KLF10/PTEN/PI3K/AKT cascade. *Biochem. Biophys. Res. Commun.* 503 (3), 1825–1829. doi:10.1016/j.bbrc.2018.07.120

Li, Y., Shi, B., Dong, F., Zhu, X., Liu, B., and Liu, Y. (2020). LncRNA KCNQ1OT1 Facilitates the Progression of Bladder Cancer by Targeting MiR-218-5p/HS3ST3B1 *Cancer Gene Ther.* 28 (3-4), 212–220. doi:10.1038/s41417-020-00211-6

Li, C., Cao, Y., Zhang, L., Li, J., Wu, H., Ling, F., et al. (2020). LncRNA IGFBP4-1 Promotes Tumor Development by Activating Janus Kinase-Signal Transducer and Activator of Transcription Pathway in Bladder Urothelial Carcinoma. *Int. J. Biol. Sci.* 16 (13), 2271–2282. doi:10.7150/ijbs.46986

Lin, C., Zhang, S., Wang, Y., Wang, Y., Nice, E., Guo, C., et al. (2018). Functional Role of a Novel Long Noncoding RNA TTN-AS1 in Esophageal Squamous Cell Carcinoma Progression and Metastasis. *Clin. Cancer Res.* 24 (2), 486–498. doi:10.1158/1078-0432.ccr-17-1851

Lin, K., Chen, H., Su, C., Zhu, H., Lai, C., and Shi, Y. (2020). Long Non-coding RNA TTN-AS1 Serves as a Competing Endogenous RNA of miR-195 to Facilitate Clear Cell Renal Cell Carcinoma Progression. *Cmar* Vol. 12, 3091–3097. doi:10.2147/cmar.s249456

Liu, Y., Li, X., Zhang, C., Zhang, H., and Huang, Y. (2020). Involves in Cancer Immune Suppression through Positive Regulation of Siglec-15 in ccRCC. *Cancer Sci.* 111 (10), 3693–3704. doi:10.1111/cas.14611

Lopez-Bergami, P., Lau, E., and Ronai, Z. e. (2010). Emerging Roles of ATF2 and the Dynamic AP1 Network in Cancer. *Nat. Rev. Cancer.* 10 (1), 65–76. doi:10.1038/nrc2681

Lun, Y.-Z., Pan, Z.-P., Liu, S.-A., Sun, J., Han, M., Liu, B., et al. (2020). The Peptide Encoded by a Novel Putative lncRNA HBVPTPAP Inducing the Apoptosis of Hepatocellular Carcinoma Cells by Modulating JAK/STAT Signaling Pathways. *Virus. Res.* 287, 198104. doi:10.1016/j.virusres.2020.198104

Lv, M., Zhong, Z., Huang, M., Tian, Q., Jiang, R., and Chen, J. (2017). lncRNA H19 Regulates Epithelial-Mesenchymal Transition and Metastasis of Bladder Cancer by miR-29b-3p as Competing Endogenous RNA. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1864 (10), 1887–1899. doi:10.1016/j.bbamcr.2017.08.001

Miao, S., Wang, J., Xuan, L., and Lu, X. (2020). LncRNA TTN-AS1 Acts as Sponge for miR-15b-5p to Regulate FBXW7 Expression in Ovarian Cancer. *Biofactors*, 46, 600–607. doi:10.1002/biof.1622

Peng, M., Cheng, X., Xiong, W., Yi, L., and Wang, Y. (2021). Integrated Analysis of a Competing Endogenous RNA Network Reveals a Prognostic lncRNA Signature in Bladder Cancer. *Front. Oncol.* 11, 684242. doi:10.3389/fonc.2021.684242

Qi, G., and Li, L. (2020). LncRNA TTN-AS1 Promotes Progression of Non-small Cell Lung Cancer via Regulating miR-491-5p/ZNF503 Axis. *Ott* 13, 6361–6371. doi:10.2147/ott.s238890

Quan, J., Pan, X., Zhao, L., Li, Z., Dai, K., Yan, F., et al. (2018). LncRNA as a Diagnostic and Prognostic Biomarker in Bladder Cancer: a Systematic Review and Meta-Analysis. *Ott* 11, 6415–6424. doi:10.2147/ott.s167853

Samimi, H., Sajjadi-Jazi, S. M., Seifirad, S., Atlasi, R., Mahmoodzadeh, H., Faghihi, M. A., et al. (2020). Molecular Mechanisms of Long Non-coding RNAs in Anaplastic Thyroid Cancer: a Systematic Review. *Cancer Cel Int.* 20, 352. doi:10.1186/s12935-020-01439-w

Shaulian, E., and Karin, M. (2002). AP-1 as a Regulator of Cell Life and Death. *Nat. Cel Biol.* 4 (5), E131–E136. doi:10.1038/ncb0502-e131

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer Statistics, 2019. *CA A. Cancer J. Clin.* 69 (1), 7–34. doi:10.3322/caac.21551

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: a Web Server for Cancer and normal Gene Expression Profiling and Interactive Analyses. *Nucleic Acids Res.* 45 (W1), W98–W102. doi:10.1093/nar/gkx247

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global Cancer Statistics, 2012. *CA Cancer J. Clin.* 65 (2), 87–108. doi:10.3322/caac.21262

Wan, Y., Liang, F., Wei, M., and Liu, Y. (2020). Long Non-coding RNA LINC00525 Regulates the Proliferation and Epithelial to Mesenchymal Transition of Human Glioma Cells by Sponging miR-338-3p. *AMB Expr.* 10 (1), 156. doi:10.1186/s13568-020-01094-4

Zhang, C., Wang, W., Lin, J., Xiao, J., and Tian, Y. (2019). lncRNA CCAT1 Promotes Bladder Cancer Cell Proliferation, Migration and Invasion. *Int. Braz. J Urol.* 45 (3), 549–559. doi:10.1590/s1677-5538.ibju.2018.0450

Zhang, Z., Yang, Li, Chen, X., Ma, F., Yang, J., et al. (2020). LncRNA MCF2L-AS1 Aggravates Proliferation, Invasion and Glycolysis of Colorectal Cancer Cells via the Crosstalk with miR-874-3p/FOXM1 Signaling axis. *Carcinogenesis* 42, 263–271. doi:10.1093/carcin/bgaa093

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 19 October 2021
doi: 10.3389/fgene.2021.710412

Prognostic and Immune-Infiltrate Significance of miR-222-3p and Its Target Genes in Thyroid Cancer

Taofeng Zhang[1†], Yihuan Chen[2†], Weixun Lin[1†], Jiehua Zheng[1,3], Yiyuan Liu[1,3], Juan Zou[1,3], Jiehui Cai[1,3], Yaokun Chen[1,3], Zhiyang Li[1] and Yexi Chen[1]**

[1]Department of General Surgery, The Second Affiliated Hospital of Shantou University Medical College, Shantou, China, [2]Department of Ultrasound in Obstetrics and Gynecology, The Second Affiliated Hospital of Shantou University Medical College, Shantou, China, [3]Department of Breast Disease Research Center, Medical Research Institute of Shantou Doctoral Association, Shantou, China

Thyroid cancer (THCA) is a common endocrine malignancy. With increasing incidence and low mortality, balancing the therapeutic approach is an inevitable issue. This study aimed to confirm the role of miR-222-3p and its target genes in THCA survival and immune infiltration. From different expression analyses based on the GEO and TCGA databases, we predicted and subsequently identified the key target genes of miR-222-3p. We then explored the expression, enrichment, pairwise correlation, protein expression, survival analysis, principal component analysis, and immune significance of the critical genes using bioinformatics analysis. The present study demonstrated that *NEGR1*, *NTNG1*, *XPNPEP2*, *NTNG2*, *CD109*, *OPCML*, and *PRND* are critical genes. The miR-222-3p was highly expressed, probably leading to low *NEGR1* and high *PRND* expression in THCA tissues. Low *NEGR1* expression indicated favorable prognosis in THCA patients, and high *PRND* expression indicated poor prognosis. Seven critical genes were significantly related to gender, age, race, tumor stage, and lymph node metastasis. In addition, the seven-gene biomarker exhibited a certain diagnostic value. Finally, *CD109* expression was closely correlated with immune cells, especially B cells and $CD4^+$ T cells. The miR-222-3p and its critical target genes could be promising biomarkers for the prognosis of THCA and may emerge as key regulators of immune infiltration in THCA.

Keywords: miR-222-3p, thyroid cancer, prognostic significance, immune infiltration, target gene, diagnostic significance

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Hsiuying Wang,
National Chiao Tung University,
Taiwan
Somnath Tagore,
Columbia University, United States

****Correspondence:***
Zhiyang Li
s_zyli4@stu.edu.cn
Yexi Chen
yxchen3@stu.edu.cn

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *16 May 2021*
Accepted: *04 October 2021*
Published: *19 October 2021*

Citation:
Zhang T, Chen Y, Lin W, Zheng J, Liu Y, Zou J, Cai J, Chen Y, Li Z and Chen Y (2021) Prognostic and Immune-Infiltrate Significance of miR-222-3p and Its Target Genes in Thyroid Cancer. Front. Genet. 12:710412. doi: 10.3389/fgene.2021.710412

INTRODUCTION

Thyroid cancer (THCA) is one of the most common endocrine malignancies and is a complex disease resulting from radiation exposure, abnormal hormone levels, abnormal iodine uptake, and genetic susceptibility(Cabanillas et al., 2016). Globally, among all types of cancers, the incidence of THCA ranks fifth worldwide, mostly due to the popularization of diagnostic imaging and surveillance(Cabanillas et al., 2016; Bray et al., 2018). Even though the prognosis of THCA is encouraging, avoiding overtreatment of low-risk patients and undertaking aggressive treatment for high-risk patients to improve prognosis can be challenging(Cabanillas et al., 2016). Indolent THCAs have low mortality, whereas other types may be aggressive, for instance anaplastic thyroid cancer. Therefore, the identification of new biomarkers for predicting the prognosis of THCA effectively is crucial for individual treatment.

MicroRNAs (miRNAs) are a class of endogenous non-coding RNA molecules, 19–25 nucleotides in length. Mature microRNAs can target the exact complementary sequence of the 3′-untranslated region (3′-UTR), resulting in the inhibition of translation and degradation of messenger ribonucleic acid (mRNA) (Bartel, 2004; Saliminejad et al., 2019). Many miRNAs are important regulators in multiple cancers, including THCA, can distinguish benign from malignant nodules(Di Leva et al., 2014; Dai et al., 2020). In recent years, studies have shown that miR-222 is highly expressed in THCA and is probably related to the occurrence, development, and prognosis(Song et al., 2017) (Lima et al., 2017; Zhang et al., 2017; Gomez-Perez et al., 2019). For example, Huang et al. found that miR-222 promotes tumor invasion and metastasis in papillary thyroid cancer by targeting *PPP2R2A*(Huang et al., 2018). However, as a familiar member of miRNA in THCA, there are few studies that have focused on the target genes of miR-222-3p, and the clinical and prognostic value of the target genes in THCA requires further study.

This study investigated the differential miR-222-3p expression in THCA tissues and non-cancerous thyroid tissues, acquired the differentially expressed genes (DEGs) in the THCA dataset of the Cancer Genome Atlas (TCGA), predicted the possible target genes of miR-222-3p using miRWalk3.0, and identified the common genes of DEGs and the possible target genes as the consensus genes. Subsequently, Gene Ontology (GO) annotation, Kyoto Encyclopedia of Genes and Genomes (KEGG) functional interpretation, Gene Set Enrichment Analysis (https://www.gsea-msigdb.org/gsea/index.jsp), and protein-protein interaction (PPI) network analyses were conducted to analyze the consensus genes. We then conducted expressed validation of critical genes and analyses of related clinical data using UALCAN (http://ualcan.path.uab.edu/analysis.html). The pairwise correlation and prognostic role of the critical target genes were validated based on the Gene Expression Profiling Interactive Analysis 2 (GEPIA2, http://gepia2.cancer-pku.cn/#index) and their protein expression were ascertained via the Human Protein Atlas (HPA, http://www.proteinatlas.org). In addition, the diagnostic value of critical target genes was estimated using multiple-gene comparison and principal component analysis. Finally, because the tumor immune microenvironment in THCA has attracted increasing attention, we mined the Cell Type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT, https://cibersort.stanford.edu) and Tumor Immune Estimation Resource (TIMER, https://cistrome.shinyapps.io/timer/) to demonstrate the relationship between miR-222-3p in THCA and immune infiltrate cells.

MATERIAL AND METHODS

Evaluation of Differential Expressisssson of miR-222-3p Expression

MiRNA expression data were downloaded from TCGA and GEO databases. The inclusion criteria for the datasets in the GEO database were as follows: 1) samples from Homo sapiens, and 2) sample size of each group ≥3. All data was normalized using the min-max normalization method. An unpaired *t*-test was performed, and the mean and standard deviations of expression levels were calculated to examine the differential miR-222-3p expression in THCA and non-cancerous tissues. Statistical significance was set at $P_{adj} < 0.05$.

Identification of Possible Target Gene

To predict the possible target genes of miR-222-3p, we utilized the online analysis website miRWalk3.0, including four prediction algorithms, TargetScan, miRDB, mirtarbase, and TarPmiR(Sticht et al., 2018). To improve the accuracy, the calculated mRNA expressions from the TCGA database were analyzed using R software package edgeR, and the DEGs were obtained. The overlapped genes of DEGs and the possible target genes were identified as the consensus genes using a Venn diagram.

Pathway Analysis

To elucidate the biological functions and signal transduction pathways of miR-222-3p and its target genes, GO annotation and KEGG functional interpretation were performed for the consensus genes using database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang et al., 2009; Kanehisa et al., 2017; The Gene Ontology, 2019). The GSEA analysis was performed to reveal the signature pathways and biological processes(Mootha et al., 2003; Subramanian et al., 2005). Statistical significance was set at $p < 0.05$.

Construction of PPI Network and Module Analysis

The PPI network and module analysis helped identify the critical genes of miR-222-3p and verified their interaction. We input the consensus genes into Cytoscape 3.7.2(Shannon et al., 2003), and imported the network data from the public database, Search Tool for the Retrieval of Interacting Genes (STRING), to construct the PPI network(Szklarczyk et al., 2019). Parameters of the construction of PPI network were as follows: species = Homo sapiens, confidence (score) cutoff = 0.40, and maximum additional interactors = 0. Module analysis was performed using the MCODE and cytoHubba plug-in. The MCODE plug-in identified the key module with the following truncation criteria: degree cutoff = 2, node score cutoff = 0.2, k-core = 2, max. depth = 100. The cluster network with genes with the highest MCODE score (6.0) was confirmed as the core network. The topological algorithm MCC in the cytoHubba plug-in was applied to explore important nodes in the PPI network.

External Validation of Critical Genes

Differential Expression of Critical Genes and Analyses of Related Clinical Data

Based on TCGA gene expression and clinical data, UALCAN was used to verify the expression of critical genes and perform the analyses of related clinical data in samples of THCA and normal thyroid (Chandrashekar et al., 2017). The 505 THCA samples and 59 normal thyroid tissue samples were involved in validation of

expression difference of critical genes and correlation analyses between critical genes and gender, age, race, tumor stage and lymph node metastasis.

Survival Analysis and Integrated Research

GEPIA2 could provide survival and integrated analysis, allowing researchers to select their genes and cancers of interest for evaluation of prognosis. We input the critical genes into the Survival Plots module, and the Kaplan-Meier curves were plotted with the log rank p value and hazard ratio (HR) with 95% confidence intervals (95% CI). Meanwhile, we conducted a multiple gene comparison and principal component analysis for the critical genes to explore their diagnostic value.

Correlation Analysis

To illustrate the association of miR-222-3p with the critical genes, we performed a Spearman's correlation analysis based on 500 thyroid carcinoma samples utilizing the LinkedOmics database(Vasaikar et al., 2018). The pairwise correlation among the critical genes was also verified using the database (Sample size = 501). Statistical significance was set at $p < 0.05$.

Protein Expression of Critical Genes

To validate the protein expression of the critical genes in THCA and adjacent normal tissues, the immunohistochemistry results of the critical genes were acquired from the Human Protein Atlas (HPA, http://www.proteinatlas.org), and the staining intensity of the antibody revealed the protein expression of critical genes(Uhlen et al., 2015).

Immune Infiltration Analysis

To explore the potential relationship between the gene and immune cells in the tumor immune environment of THCA, the mRNA expression matrix from TCGA was normalized and the CIBERSORT algorithm was utilized to estimate the composition of 22 human immune cells in tumor immune environment(Newman et al., 2015). Then, the composition of immune cells in THCA and normal tissue group was compared with an unpaired t-test. Moreover, we used the TIMER algorithm to illustrate the correlation between the expression of critical genes and the abundance of six tumor-infiltrating immune cells (CD4 + T cells, CD8 + T cells, B cells, neutrophils, dendritic cells, and macrophages) in THCA(Li et al., 2017). The correlation values were processed using purity-corrected partial Spearman's correlation. In general, genes upregulated in the microenvironment are negatively associated with tumor purity, whereas genes upregulated in tumor cells are positively associated with tumor purity. All p-values < 0.05 were considered significant.

RESULTS

Differential miR-222-3p Expression

From the TCGA and GEO database, a total of 603 THCA and 120 normal thyroid tissue samples were compared in the present study. The miR-222-3p expression in THCA samples was significantly upregulated compared to that in normal thyroid tissue ($p < 0.001$). The mean ± standard deviation were 0.25250 ± 0.18090 and 0.04267 ± 0.02269, respectively (**Figure 1A**). As shown in **Figure 1B–F**, the study included five datasets from the GEO database. Among them, three databases (gse73182, gse113629, and gse40807) shared the same result, consistent with that of TCGA ($p < 0.05$), whereas no statistical significance was found in the two remaining datasets (gse62054 and gse97070).

Possible Target Genes of miR-222-3p in THCA

Possible target genes were determined using the target gene prediction platform miRWalk3.0 and differential expression based on TCGA. The differential expression analysis involved 510 THCA and 58 matched group, in which 5,533 differently expressed genes ($|\log2FC| > 1$) were identified by edgeR. In addition, 2,105 possible target genes were predicted using the miRWalk3.0. Consequently, we regarded 303 overlapping genes as consensus target genes of miR-222-3p (**Figure 2**).

Pathway Analysis

GO and KEGG enrichment analyses of 303 consensus genes were performed using DAVID. GO annotation showed that 28, 26, and 11 clustered terms were statistically significant in biological process (BP), cellular component (CC), and molecular function (MF), respectively, and a total of six significant pathways were verified by KEGG analysis. These noticeably enriched GO terms and KEGG pathways may play an important role in the mechanism of thyroid carcinoma, which might contribute to appropriate tailoring of treatment for patients. The BP GO term "nervous system development," CC GO term "integral component of plasma membrane" and the MF GO term "calcium ion binding" had the smallest p-value, indicating the closest connection with the tumor (**Figures 3A–C**). For the KEGG pathway, nicotine addiction had the smallest p value, whereas neuroactive ligand-receptor interaction had the largest number of consensus genes (**Figure 3D**). The results demonstrated three significantly enriched gene sets from the following pathways: HALLMARK_COAGULATION, HALLMARK_P53_PATHWAY, and HALLMARK_APICAL_JUNCTION (**Figure 3E**).

Identification of Critical Genes

The PPI network consisted of 302 nodes and 369 edges generated by entering 303 consensus genes into the Cytoscape 3.7.2 (**Figure 4A**). The module analysis of the PPI network, conducted by the MCODE and cytoHubba plug-in, uncovered two core modules (**Figure 4B, C**). In the two above modules, except for *ATP2B3*, *SLC17A7*, and *SYP*, the remaining seven genes (*NEGR1*, *NTNG1*, *XPNPEP2*, *NTNG2*, *CD109*, *OPCML*, and *PRND*) were identified as critical genes, indicating that they may be important in the pathogenesis of THCA.

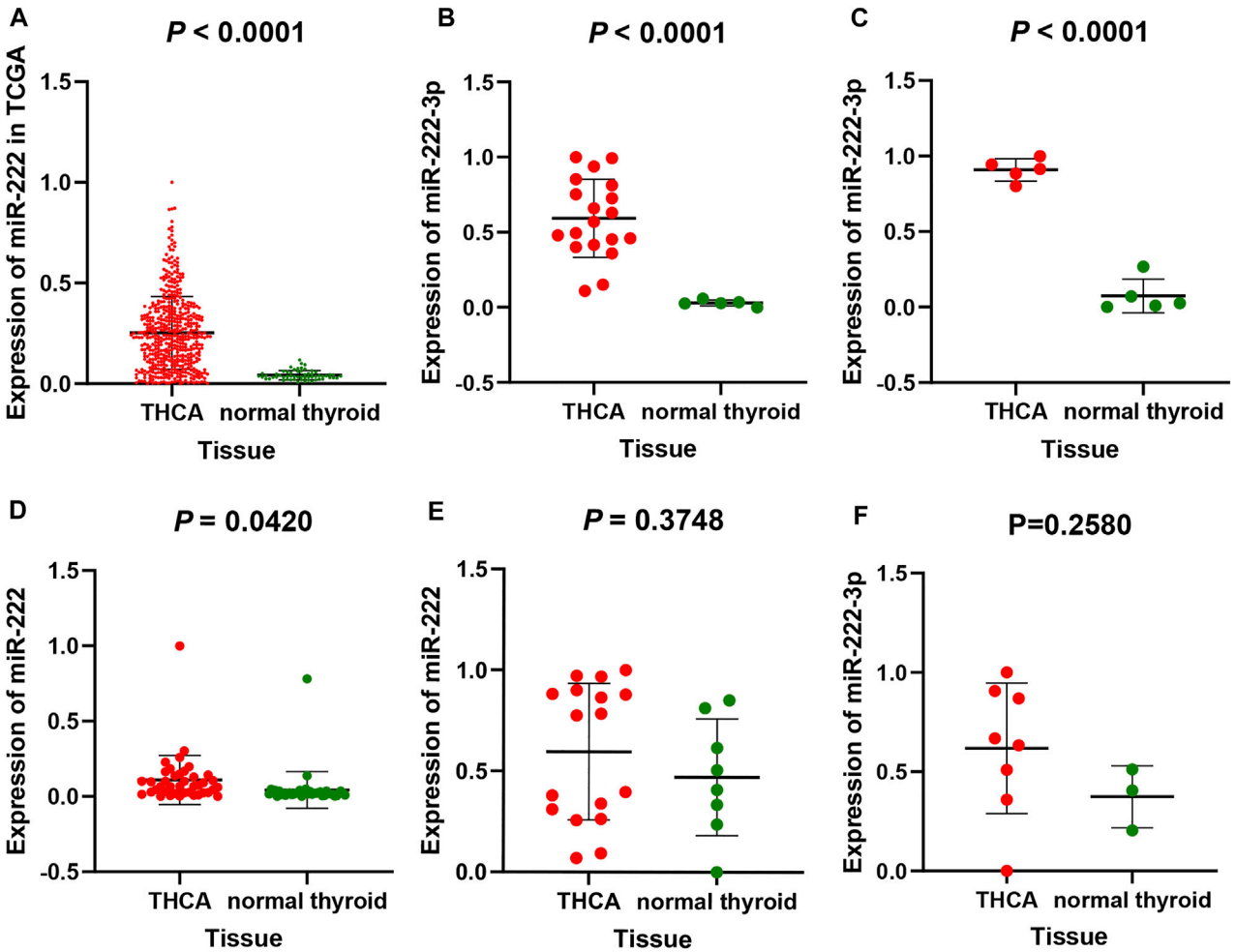


FIGURE 1 | miR-222-3p differential expression in THCA and non-cancerous tissues based on TCGA database **(A)** and GEO database, including gse73182 **(B)**, gse113629 **(C)**, gse40807 **(D)**, gse62054 **(E)** and gse97070 **(F)** dataset. FA stands for thyroid follicular adenoma.

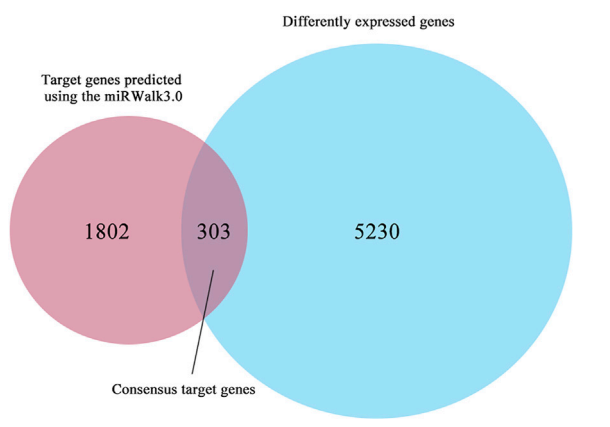


FIGURE 2 | The Venn diagram shows a total of 303 consensus genes.

External Validation of Critical Genes

Differential Expression of Critical Genes and Analyses of Related Clinical Data

Seven critical genes were identified in the module closest to the PPI network. Based on the UALCAN, we evaluated differential expression of the seven critical genes and preformed the analyses of related clinical data in samples of THCA and normal thyroid. In differential expression of critical genes (**Figure 5**), three of the seven genes (*NEGR1*, *NTNG1*, and *XPNPEP2*) showed low expression, and the remaining critical genes (*NTNG2*, *CD109*, *OPCML*, and *PRND*) were highly expressed in THCA samples, comparing to normal thyroid samples. In terms of gender (**Supplementary Figure S1**), differential expression of seven critical genes all had statistical significance in the groups of male and female. In terms of age (**Supplementary Figure S2**), differential expression of *NEGR1*, *CD109*, and *PRND* had statistical significance in all age groups. Differential expression of *NTNG1*, *XPNPEP2*, and *OPCML* had statistical significance in the groups of 21–40, 41–60, and 61–80 years, but not in the group of 81–100 years. And differential expression of *NTNG2* had

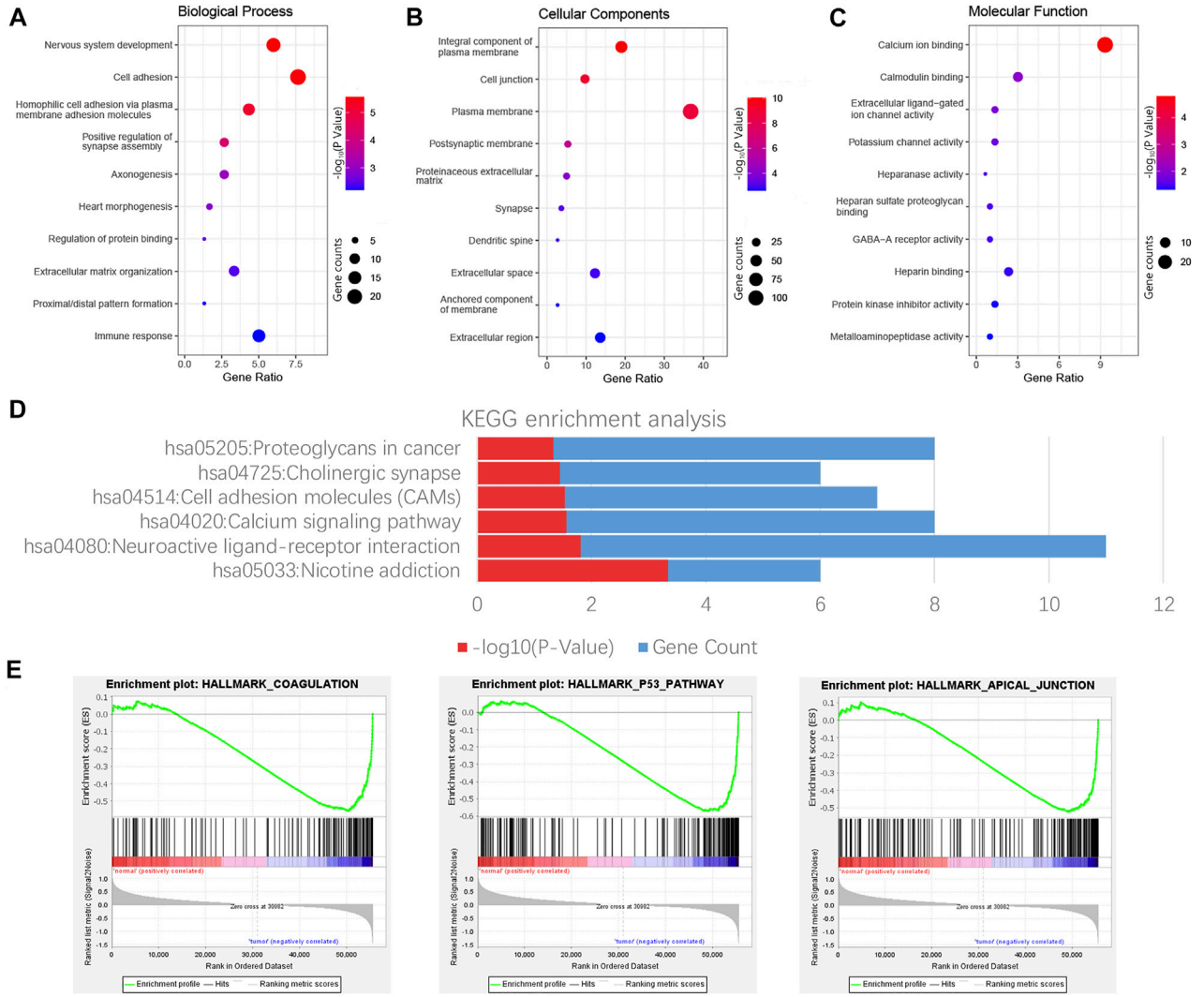


FIGURE 3 | Functional enrichment analysis of consensus genes. GO enrichment analysis results include Biological process (BP), cell composition (CC), and molecular function (MF), and each part displays 10 terms **(A–C)**. The result shown by KEGG is six possible signature pathways for enrichment of consensus genes **(D)**. Pathway enrichment plots of three gene sets by performing GSEA **(E)**.

statistical significance in the group of 21–40 years, but not in the group of 41–60, 61–80, or 81–100 years. In terms of race (**Supplementary Figure S3**), differential expression of *NTNG1*, *XPNPEP2*, *CD109*, *OPCML*, and *PRND* had statistical significance in all race groups. Differential expression of *NEGR1* had statistical significance in the groups of Caucasian and Asian, but not in the group of African-American. And differential expression of *NTNG2* had statistical significance in the group of Caucasian, but not in the groups of African-American or Asian. In terms of tumor stage (**Supplementary Figure S4**), differential expression of *NEGR1*, *NTNG1*, *XPNPEP2*, *CD109*, and *PRND* had statistical significance in all tumor stage groups. Differential expression of *OPCML* had statistical significance in the groups of stage1, 3, and 4, but not in the group of stage2. And differential expression of *NTNG2* had statistical significance in the group of stage1, but not in the groups of stage2, 3, or 4. In terms of lymph node metastasis (**Supplementary Figure S5**), differential expression of *NEGR1*, *NTNG1*, *XPNPEP2*, *CD109*, *OPCML*, and *PRND* had statistical significance in the groups of N0 and N1. And there was no statistical significance in differential expression of *NTNG2* in the groups of N0 or N1.

Prognostic Value of Critical Genes

A total of 505 THCA patients underwent survival analysis using the Kaplan-Meier (KM) curve. Interestingly, high *NEGR1* expression was significantly correlated with worse OS (HR = 5.3, p = 0.0097), whereas DFS was not significantly different (**Figure 6A**). On the contrary, low *PRND* expression was significantly correlated with better DFS (HR = 2, p = 0.03), whereas the OS results were not significantly different

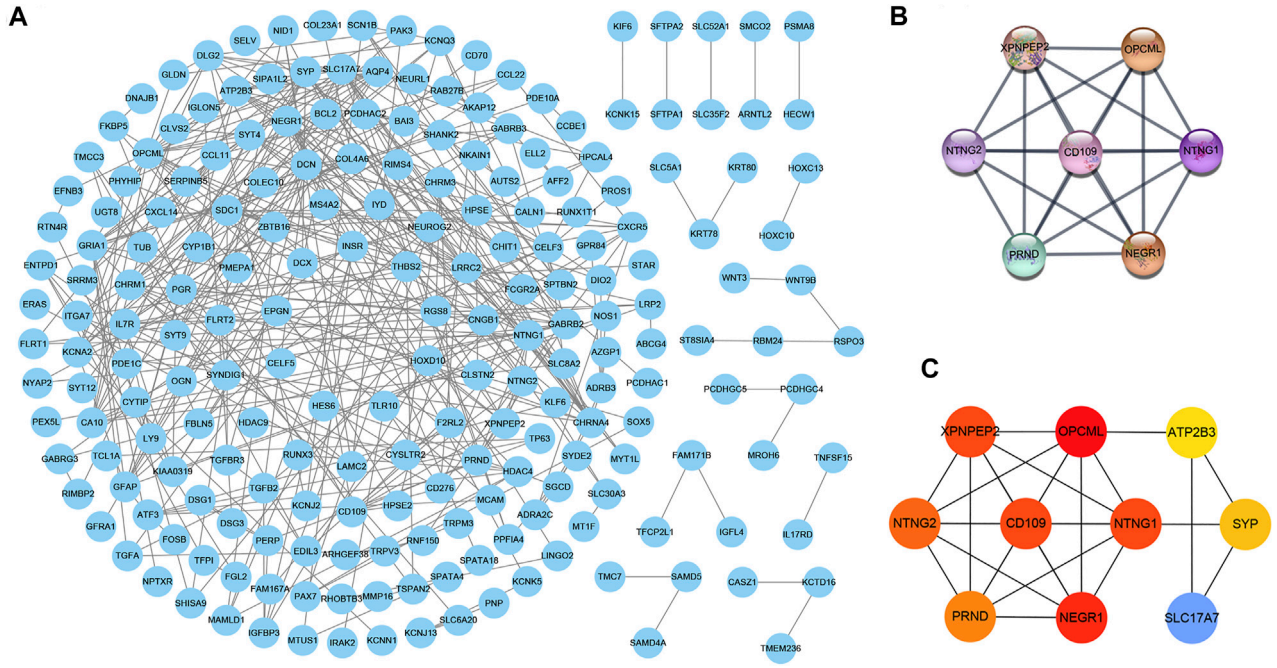


FIGURE 4 | The PPI network composed of 303 consensus genes, in which each node and edge represents each gene-encoded protein and link between proteins, respectively **(A) (B)** The top scored module calculated by the MCODE plug-in of Cytoscape software **(C)** The top 10 scored genes calculated by cytoHubba plug-in of Cytoscape software, and the **red** represents the score of genes. The deeper the **red,** the higher the score.

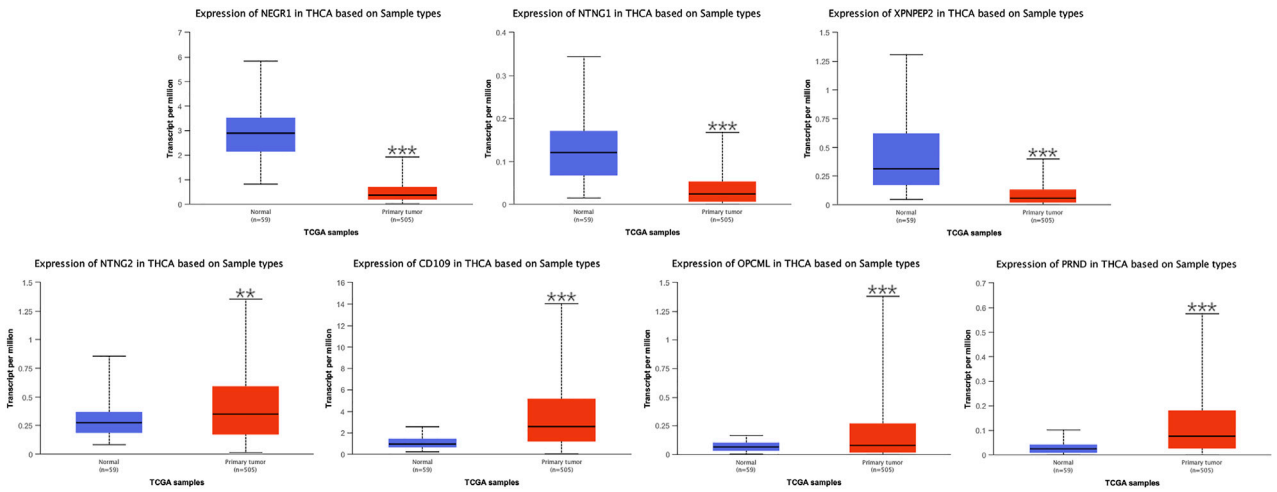


FIGURE 5 | Differential expression of *NEGR1*, *NTNG1*, *XPNPEP2*, *NTNG2*, *CD109*, *OPCML*, and *PRND* based on UALCAN (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

(**Figure 6B**). For the other critical genes, results of the survival analysis were not statistically significant.

Diagnostic Value of Seven-Gene Biomarker

The two analyzed modules on GEPIA2, multiple-gene comparison and principal component analysis, were utilized to estimate the diagnostic value of the seven-gene biomarker. Among them, *CD109* exhibited the highest expression level in THCA samples via multiple-gene comparison, followed by *NTNG2*, *NEGR1*, *OPCML*, *XPNPEP2*, *PRND*, and *NTNG1*(**Figure 7A**). Notably, results of principal component analysis with TCGA tumor data, TCGA normal data, and GTEx thyroid data revealed that the seven-gene biomarker could contribute to effectively differentiate THCA from normal thyroid samples (**Figure 7B**), suggesting its potential value in the diagnosis of THCA.

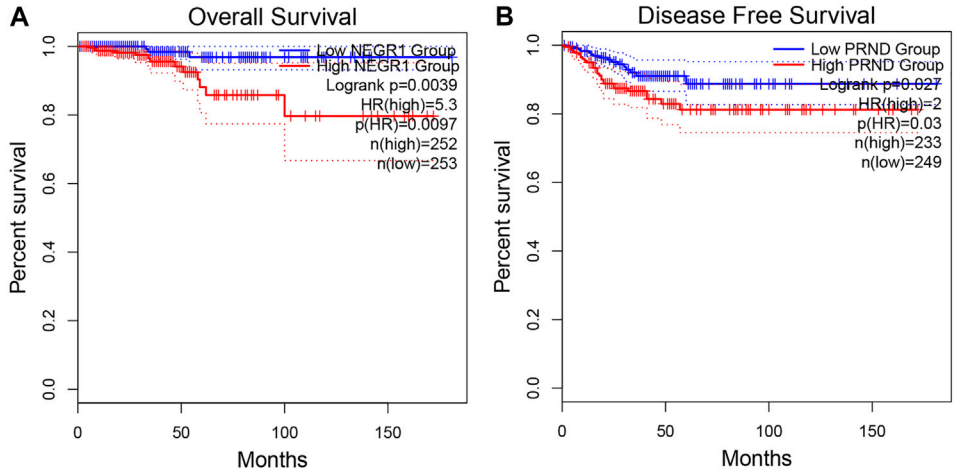


FIGURE 6 | Overall survival analysis of *NEGR1* **(A)** and disease free survival analysis of *PRND* **(B)** based on GEPIA2.

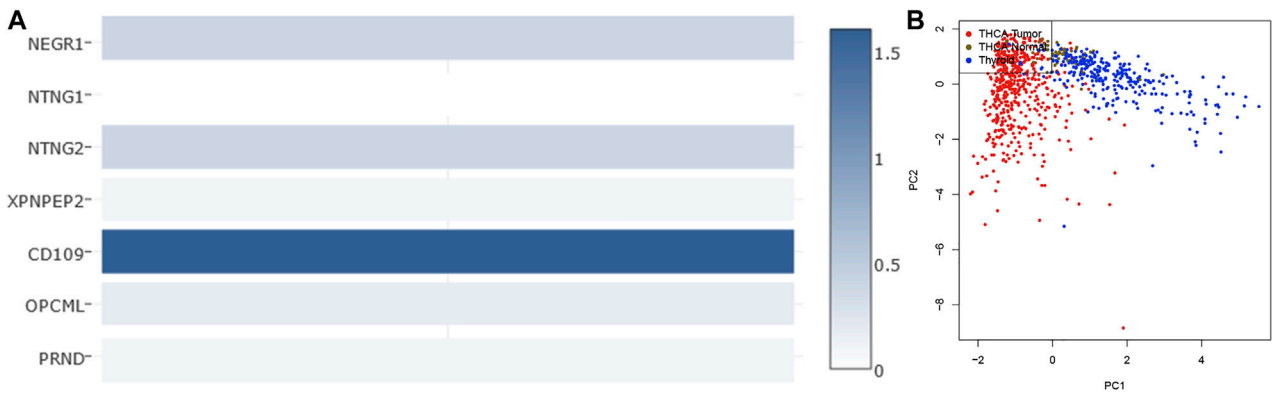


FIGURE 7 | Multiple-gene comparison analysis **(A)** and principal component analysis **(B)** of the seven-gene biomarker in THCA.

Correlation Analysis

Spearman's correlation analysis revealed that five of the seven critical genes were significantly correlated with miR-222-3p expression. As shown in **Figure 8A**, *CD109* (r = 0.482, $p < 0.001$), *NTNG2* (r = 0.169, $p < 0.001$), and *OPCML* (r = 0.176, $p < 0.001$) were positively correlated with miR-222-3p, whereas *NEGR1* (r = −0.285, $p < 0.001$) and *PRND* (r = −0.122, $p < 0.01$) were negatively correlated. However, the correlation of *NTNG1* ($p = 0.29$) and *XPNPEP2* ($p = 0.15$) with miR-222-3p was not statistically significant. Interestingly, a negative correlation (cor = -0.19) only existed between *NEGR1* and *CD109*, while the other groups exhibited positive correlations (**Figure 8B**).

Protein Expression of Critical Genes

The immunohistochemistry results of critical genes of THCA in the HPA database were investigated to ascertain their protein expression. It is worth noting that only *NEGR1* had a marked negative protein expression in carcinoma tissues compared to normal gland tissues, consistent with the above differential expression of critical genes (**Figure 9**), whereas *CD109*, *NTNG2*, *PRND*, and *XPNPEP1* showed no difference in protein expression between normal and tumor tissues. The immunohistochemistry results of *NTNG1* and *OPCML* were not available for HPA.

Immune Infiltration Analysis

The composition of 22 immune cells in tumor immune environment was estimated. In TCGA databases, the THCA and normal tissue group showed a significant difference for the composition of 21 immune cells (naive B cells, memory B cells, plasma cells, $CD8^+$ T cells, naive $CD4^+$ T cells, memory-resting $CD4^+$ T cells, memory-activated $CD4^+$ T cells, follicular helper T cells, regulatory T cells (Tregs), gamma delta T cells, resting NK cells, activated NK cells, M0 macrophages, M1 macrophages, M2 macrophages, resting dendritic cells, activated dendritic cells, resting mast cells, activated mast cells, eosinophils, neutrophils) (**Figure 10A**). Contributing to providing insight for immunotherapeutic treatment in THCA, the tumor immune infiltrates of THCA, including B cells, $CD4^+$ T cells, $CD8^+$

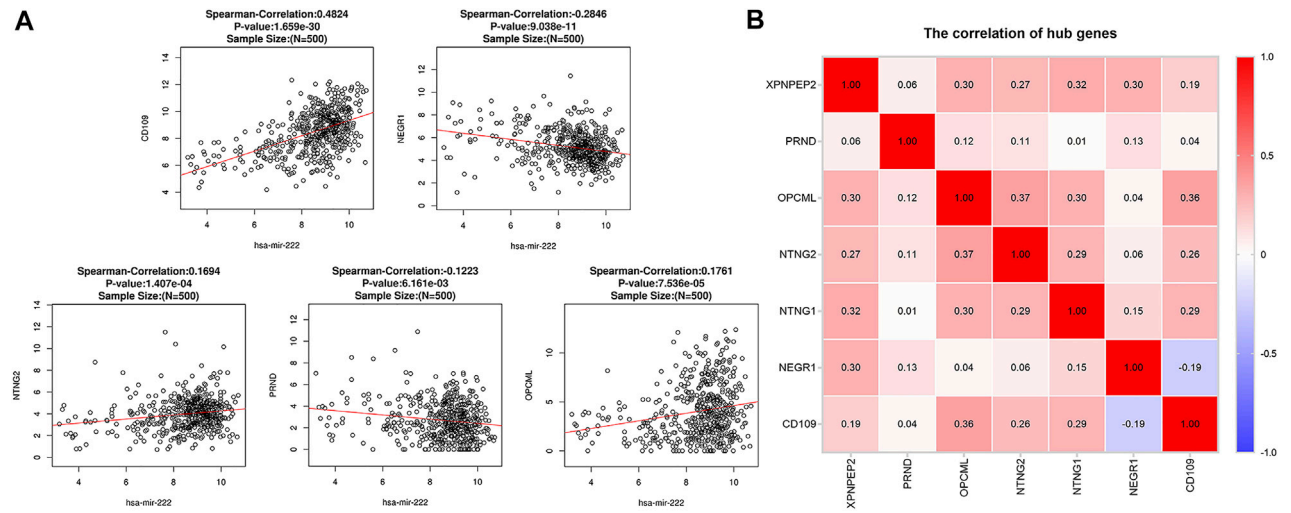


FIGURE 8 | Correlation analysis between miR-222-3p and *CD109*, *NTNG2*, *OPCML*, *NEGR1* and *PRND* in THCA **(A)**. Pairwise correlation analysis of the expression of critical genes in THCA **(B)**. **Red:** positive correlation; **blue:** negative correlation.

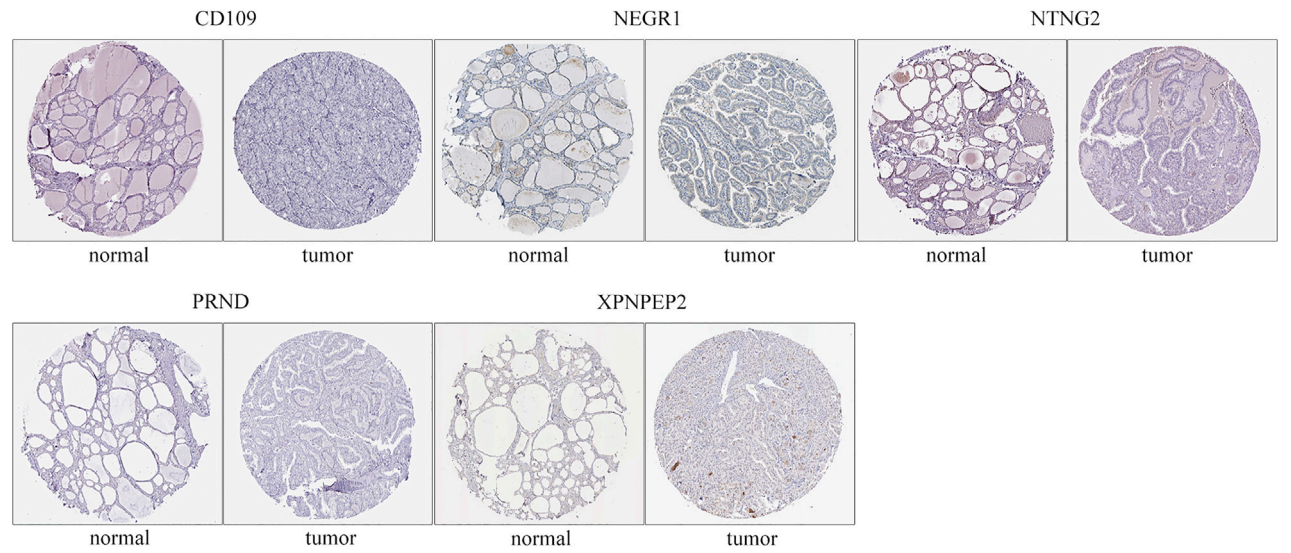


FIGURE 9 | Protein expression of critical genes. Images were obtained from the HPA database.

T cells, neutrophils, macrophages and dendritic cells, were analyzed via the TIMER platform. *CD109* demonstrated the highest positive correlation with B cells (cor = 0.58, $p < 0.001$), $CD4^+$ T cells (cor = 0.66, $p < 0.001$), macrophages (cor = 0.48, $p < 0.001$), neutrophils (cor = 0.60, $p < 0.001$), and dendritic cells (cor = 0.53, $p < 0.001$), and the highest negative correlation with $CD8^+$ T cells (cor = −0.24, $p < 0.001$) (**Figure 10B**). Moreover, *NTNG1* and *XPNPEP2* were also closely related to several types of immune-infiltrate cells (**Figure 10B**). In addition, an insignificant correlation was found between all types of immune-infiltrate cells and *NEGR1* and *PRND* (**Figure 10B**).

DISCUSSION

Based on the TCGA and GEO databases, we confirmed miR-222-3p overexpression in THCA, consistent with previous studies(Paskas et al., 2015; Borrelli et al., 2017). Moreover, elevated miR-222-3p expression was associated with high-risk characteristics, including tumor invasion, lymph node metastasis, and recurrence that may be closely related to the poor prognosis of THCA patients (Aragon Han et al., 2015; Lima et al., 2017; Gomez-Perez et al., 2019). Visone et al. reported that miR-221 and miR-222 regulate the cell cycle by targeting $p27^{Kip1}$ in human thyroid papillary carcinomas, while abnormal cell cycle leads to

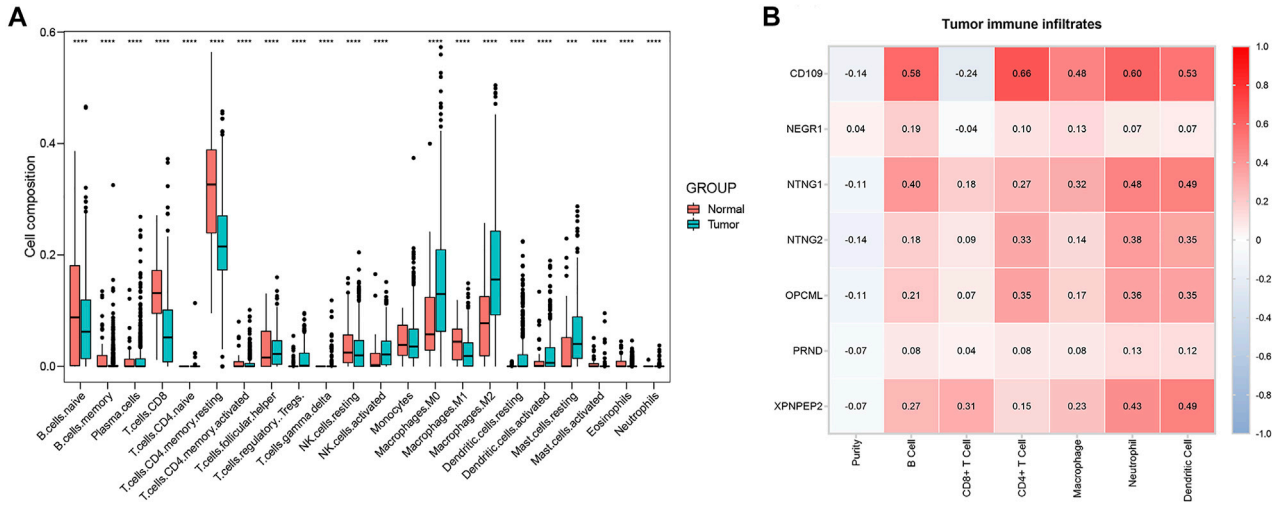


FIGURE 10 | Boxplot of 22 immune cells composition in THCA and normal tissue group **(A)**. The correlation of critical genes and tumor immune infiltrates analyzed via the Tumor Immune Estimation Resource (TIMER) platform **(B)**. **Red:** positive correlation; **blue:** negative correlation.

malignant transformation. Moreover, Huang et al. indicated that miR-222 promotes tumor invasion and metastasis in papillary thyroid cancer by targeting *PPP2R2A*(Huang et al., 2018). However, few researchers have focused on the prognostic and diagnostic value of miR-222-3p and its target genes and the correlation between target genes and immune invasion.

In the present study, we aimed to elucidate the biological functions of miR-222-3p and its target genes in THCA. Differential expression analysis of miR-222-3p based on the GEO and TCGA database indicated the upregulation of miR-222-3p expression in THCA; 303 consensus genes were obtained by overlapping the target genes predicted by miRWalk3.0 and DEGs from the TCGA database. To search for enriched pathways of miR-222-3p and its target genes, we acquired three possible signature pathways using the GSEA analysis. Moreover, using the STRING public platform, the PPI network was constructed, and seven genes (*NEGR1*, *NTNG1*, *XPNPEP2*, *NTNG2*, *CD109*, *OPCML*, and *PRND*) were identified as critical genes through module analysis using the Cytoscape software.

We discovered close relationships between critical genes and prognosis of THCA. Results confirmed that *NEGR1* correlated with the overall survival (OS) of THCA, while *PRND* was associated with disease-free survival (DFS) and THCA. Neuronal growth regulator 1 (*NEGR1*) is a member of the immunoglobulin LON (IgLON) family, which pertains to GPI-anchored cell adhesion molecules (CAMs), consistent with the pathways of KEGG enrichment analysis(Kubick et al., 2018). Previous studies revealed that *NEGR1* is commonly downregulated as a tumor suppressor gene in many cancers, including neuroblastoma, breast, colon, and kidney cancers(Takita et al., 2011; Hiemer et al., 2014; Huang et al., 2020). By investigating the genetic landscape of anaplastic thyroid cancer, Woodward et al. confirmed that the *NEGR1* frequent deletion were associated with the tumorigenesis of anaplastic thyroid carcinoma(Woodward et al., 2017). This study showed that *NEGR1* was related to the overall survival of THCA patients, and a low *NEGR1* expression may drive tumor progression by abnormal cell adhesion. However, further studies are needed to verify this hypothesis. Prion-like protein doppel (*PRND*), a paralog of the prion (PrP) protein, was upregulated in many malignant diseases, including astrocytomas, osteosarcoma, acute myeloid leukemia, and myelodysplastic syndromes(Travaglino et al., 2005; Comincini et al., 2007; Sollazzo et al., 2012; Al-Hilal et al., 2016). To date, no study has elucidated the regulatory mechanism of *PRND* in thyroid cancer. In our study, *PRND* had a high level expression in THCA, indicating poor DFS. Nevertheless, more research is needed to explore the target signature pathway.

Even though *NEGR1* and *PRND* exhibited prognostic significance in THCA, their pairwise correlations indicated that the remaining critical genes, *CD109*, *NTNG1*, *NTNG2*, *XPNPEP2*, and *OPCML*, also play important roles in the regulation of biological processes. The analyses of related clinical data revealed that critical genes were all significantly related to tumor stage and lymph node metastasis of THCA. In addition, multiple-gene comparison and principal component analysis of critical genes revealed that they had potential diagnostic value in THCA, which indicated that their conjoint analysis offered a method for effectively distinguishing THCA from normal thyroid tissue. Therefore, seven critical genes were suggested as biomarkers for the precise diagnosis of THCA. These results provide insights into the role of multiple genes and interactive networks in the multilevel regulation of the molecular mechanism of THCA.

Previous studies revealed that immune cells in the tumor microenvironment play an essential role in the occurrence and development of THCA(Zhuang et al., 2020). As an important part of the complex tumor microenvironment, immune infiltrate cells are correlated with the biological processes of cancers and the survival of cancer patients. In this study, the analysis of tumor

immune infiltrates revealed that gene signature of THCA had a close relationship with 21 types of immune cells in tumor immune environment. And *CD109* had the highest positive correlation with B cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells, and the highest negative correlation with $CD8^+$ T cells. *NTNG1* and *XPNPEP2* also exhibited a high correlation with a variety of immune cells. Overall, the results of immune infiltrates illustrated that miR-222-3p was pertinent to the regulation of the immune microenvironment, mainly through targeting *CD109*, *NTNG1*, and *XPNPEP2* in THCA.

Inevitably, there were certain limitations to our study. As the datasets were obtained from different databases and platforms, an uncertain systematic bias might exist. Furthermore, although enriched analysis was performed to preliminarily investigate the regulatory function of miR-222-3p and its target genes in THCA, the particular mechanism between miR-222-3p and critical genes requires further *in vitro* and *in vivo* experiments. Nevertheless, the results of this study are significant as they highlight on promising biomarkers of prognosis, diagnosis, and immune infiltrates.

CONCLUSION

In summary, using integrated bioinformatics analysis, we demonstrated that miR-222-3p was upregulated in THCA. The elevated miR-222-3p expression may result in downregulated *NEGR1* expression and elevated *PRND* expression. Downregulated *NEGR1* expression indicating a positive prognosis of THCA, whereas elevated *PRND* expression indicated poor prognosis. *CD109* may be capable of regulating B cells, CD4+T cells, macrophages, neutrophils, and dendritic cells in immune microenvironment of THCA. Hence, the critical target genes of miR-222-3p, which indicate potential prognostic value, play an essential role in the infiltration of immune cells and could act as potential targets for the treatment of THCA patients.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The Cancer Genome Atlas (https://portal.gdc.cancer.gov/repository).

AUTHOR CONTRIBUTIONS

TZ, YHC, WL, JZh, YL, JZo, JC, YKC, ZL, and YXC: conception. TZ, YHC, WL, and JZh: data collection and statistics. YL, JZo, JC, and YKC: literature search. TZ and YHC: Drafting, review, and revision of the manuscript. YXC and ZL: Direction.

FUNDING

This work was supported by the Special Fund Project of Guangdong Science and Technology (210728156901524, 210728156901519), Medical Scientific Research Foundation of Guangdong Province, China (grant number A2021432, B2021448), Shantou Medical Science and Technology Planning Project (grant number 210521236491457, 210625106490696), and the Undergraduate Innovation Training Project of Shantou University (grant number 31/38/47/54).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.710412/full#supplementary-material.

REFERENCES

Al-Hilal, T. A., Chung, S. W., Choi, J. U., Alam, F., Park, J., Kim, S. W., et al. (2016). Targeting Prion-Like Protein Doppel Selectively Suppresses Tumor Angiogenesis. *J. Clin. Invest.* 126 (4), 1251–1266. doi:10.1172/JCI83427

Aragon Han, P., Weng, C.-H., Khawaja, H. T., Nagarajan, N., Schneider, E. B., Umbricht, C. B., et al. (2015). MicroRNA Expression and Association With Clinicopathologic Features in Papillary Thyroid Cancer: A Systematic Review. *Thyroid.* 25 (12), 1322–1329. doi:10.1089/thy.2015.0193

Bartel, D. P. (2004). MicroRNAs. *Cell.* 116 (2), 281–297. doi:10.1016/s0092-8674(04)00045-5

Borrelli, N., Denaro, M., Ugolini, C., Poma, A. M., Miccoli, M., Vitti, P., et al. (2017). miRNA Expression Profiling of 'Noninvasive Follicular Thyroid Neoplasms With Papillary-Like Nuclear Features' Compared With Adenomas and Infiltrative Follicular Variants of Papillary Thyroid Carcinomas. *Mod. Pathol.* 30 (1), 39–51. doi:10.1038/modpathol.2016.157

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J. Clinicians.* 68 (6), 394–424. doi:10.3322/caac.21492

Cabanillas, M. E., McFadden, D. G., and Durante, C. (2016). Thyroid Cancer. *The Lancet.* 388 (10061), 2783–2795. doi:10.1016/s0140-6736(16)30172-6

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B. V. S. K., et al. (2017). UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia.* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Comincini, S., Ferrara, V., Arias, A., Malovini, A., Azzalin, A., Ferretti, L., et al. (2007). Diagnostic Value of PRND Gene Expression Profiles in Astrocytomas: Relationship to Tumor Grades of Malignancy. *Oncol. Rep.* 17 (5), 989–996. doi:10.3892/or.17.5.989

Dai, D., Tan, Y., Guo, L., Tang, A., and Zhao, Y. (2020). Identification of Exosomal miRNA Biomarkers for Diagnosis of Papillary Thyroid Cancer by Small RNA Sequencing. *Eur. J. Endocrinol.* 182 (1), 111–121. doi:10.1530/EJE-19-0524

Di Leva, G., Garofalo, M., and Croce, C. M. (2014). MicroRNAs in Cancer. *Annu. Rev. Pathol. Mech. Dis.* 9, 287–314. doi:10.1146/annurev-pathol-012513-104715

Gómez-Pérez, A. M., Cornejo Pareja, I. M., García Alemán, J., Coín Aragüez, L., Sebastián Ochoa, A., Alcaide Torres, J., et al. (2019). New Molecular Biomarkers in Differentiated Thyroid Carcinoma: Impact of miR-146, miR-221 and miR-222 Levels in the Evolution of the Disease. *Clin. Endocrinol.* 91 (1), 187–194. doi:10.1111/cen.13972

Hiemer, S. E., Szymaniak, A. D., and Varelas, X. (2014). The Transcriptional Regulators TAZ and YAP Direct Transforming Growth Factor β-Induced Tumorigenic Phenotypes in Breast Cancer Cells. *J. Biol. Chem.* 289 (19), 13461–13474. doi:10.1074/jbc.M113.529115

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat. Protoc.* 4 (1), 44–57. doi:10.1038/nprot.2008.211

Huang, T., Huang, X., Nie, Y., Shi, X., and Shu, C. (2020). A Combined Effect of Expression Levels of Obesity-Related Genes and Clinical Factors on Cancer Survival Rate. *Biomed. Res. Int.* 2020, 1–20. doi:10.1155/2020/8838676

Huang, Y., Yu, S., Cao, S., Yin, Y., Hong, S., Guan, H., et al. (2018). MicroRNA-222 Promotes Invasion and Metastasis of Papillary Thyroid Cancer Through Targeting Protein Phosphatase 2 Regulatory Subunit B Alpha Expression. *Thyroid.* 28 (9), 1162–1173. doi:10.1089/thy.2017.0665

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* 45 (D1), D353–D361. doi:10.1093/nar/gkw1092

Kubick, N., Brösamle, D., and Mickael, M.-E. (2018). Molecular Evolution and Functional Divergence of the IgLON Family. *Evol. Bioinform Online.* 14, 117693431877508. doi:10.1177/1176934318775081

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* 77 (21), e108–e110. doi:10.1158/0008-5472.CAN-17-0307

Lima, C. R., Gomes, C. C., and Santos, M. F. (2017). Role of microRNAs in Endocrine Cancer Metastasis. *Mol. Cell Endocrinol.* 456, 62–75. doi:10.1016/j.mce.2017.03.015

Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1α-responsive Genes Involved in Oxidative Phosphorylation Are Coordinately Downregulated in Human Diabetes. *Nat. Genet.* 34 (3), 267–273. doi:10.1038/ng1180

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets From Tissue Expression Profiles. *Nat. Methods.* 12 (5), 453–457. doi:10.1038/nmeth.3337

Paskaš, S., Janković, J., Živaljević, V., Tatić, S., Božić, V., Nikolić, A., et al. (2015). Malignant Risk Stratification of Thyroid FNA Specimens with Indeterminate Cytology Based on Molecular Testing. *Cancer Cytopathology.* 123 (8), 471–479. doi:10.1002/cncy.21554

Saliminejad, K., Khorram Khorshid, H. R., Soleymani Fard, S., and Ghaffari, S. H. (2019). An Overview of microRNAs: Biology, Functions, Therapeutics, and Analysis Methods. *J. Cell Physiol.* 234 (5), 5451–5465. doi:10.1002/jcp.27486

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Sollazzo, V., Galasso, M., Volinia, S., and Carinci, F. (2012). Prion Proteins (PRNP and PRND) Are Over-Expressed in Osteosarcoma. *J. Orthop. Res.* 30 (6), 1004–1012. doi:10.1002/jor.22034

Song, J., Ouyang, Y., Che, J., Li, X., Zhao, Y., Yang, K., et al. (2017). Potential Value of miR-221/222 as Diagnostic, Prognostic, and Therapeutic Biomarkers for Diseases. *Front. Immunol.* 8, 56. doi:10.3389/fimmu.2017.00056

Sticht, C., De La Torre, C., Parveen, A., and Gretz, N. (2018). miRWalk: An Online Resource for Prediction of MicroRNA Binding Sites. *PLoS One.* 13 (10), e0206239. doi:10.1371/journal.pone.0206239

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc. Natl. Acad. Sci.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING V11: Protein-Protein Association Networks With Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res.* 47 (D1), D607–D613. doi:10.1093/nar/gky1131

Takita, J., Chen, Y., Okubo, J., Sanada, M., Adachi, M., Ohki, K., et al. (2011). Aberrations of NEGR1 on 1p31 and MYEOV on 11q13 in Neuroblastoma. *Cancer Sci.* 102 (9), 1645–1650. doi:10.1111/j.1349-7006.2011.01995.x

The Gene Ontology, C. (2019). The Gene Ontology Resource: 20 Years and Still GOing Strong. *Nucleic Acids Res.* 47 (D1), D330–D338. doi:10.1093/nar/gky1055

Travaglino, E., Comincini, S., Benatti, C., Azzalin, A., Nano, R., Rosti, V., et al. (2005). Overexpression of the Doppel Protein in Acute Myeloid Leukaemias and Myelodysplastic Syndromes. *Br. J. Haematol.* 128 (6), 877–884. doi:10.1111/j.1365-2141.2005.05386.x

Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Tissue-based Map of the Human Proteome. *Science.* 347 (6220), 1260419. doi:10.1126/science.1260419

Vasaikar, S. V., Straub, P., Wang, J., and Zhang, B. (2018). LinkedOmics: Analyzing Multi-Omics Data Within and Across 32 Cancer Types. *Nucleic Acids Res.* 46 (D1), D956–D963. doi:10.1093/nar/gkx1090

Woodward, E. L., Biloglav, A., Ravi, N., Yang, M., Ekblad, L., Wennerberg, J., et al. (2017). Genomic Complexity and Targeted Genes in Anaplastic Thyroid Cancer Cell Lines. *Endocr. Relat. Cancer.* 24 (5), 209–220. doi:10.1530/ERC-16-0522

Zhang, Y., Xu, D., Pan, J., Yang, Z., Chen, M., Han, J., et al. (2017). Dynamic Monitoring of Circulating microRNAs as a Predictive Biomarker for the Diagnosis and Recurrence of Papillary Thyroid Carcinoma. *Oncol. Lett.* 13 (6), 4252–4266. doi:10.3892/ol.2017.6028

Zhuang, G., Zeng, Y., Tang, Q., He, Q., and Luo, G. (2020). Identifying M1 Macrophage-Related Genes Through a Co-expression Network to Construct a Four-Gene Risk-Scoring Model for Predicting Thyroid Cancer Prognosis. *Front. Genet.* 11, 591079. doi:10.3389/fgene.2020.591079

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 26 October 2021
doi: 10.3389/fgene.2021.690196

miR-25 Regulates Gastric Cancer Cell Growth and Apoptosis by Targeting EGR2

*Liuqing Yang[1], Lina Li[2], Pan Chang[1], Ming Wei[3], Jianting Chen[1], Chaofan Zhu[1] and Jing Jia[1]**

[1]Second Affiliated Hospital of Xi'an Medical University, Xi' an, China, [2]First Department of Medical Oncology, Affiliated Shaanxi Provincial Cancer Hospital, Xi'an, China, [3]Department of Pharmacology, Xi'an Medical University, Xi'an, China

Gastric cancer is one of the most common malignancies harmful to human health. The search for effective drugs or gene therapy has aroused the attention of scientists. So far, microRNAs, as small non-coding RNAs, have the potential to be therapeutic targets for cancer. Herein, we found a highly expressed miR-25 in gastric cancer cell. However, the function of miR-25 for gastric cancer cell growth and apoptosis was unknown. Functionally, we used RT-qPCR, western blot, CCK-8, and flow cytometry to detect gastric cancer cell growth and apoptosis. The results indicated that miR-25 promoted gastric cancer cell growth and inhibited their apoptosis. Mechanistically, we found that a gene EGR2 was a potential target gene of miR-25. Further dual-luciferase results supported this prediction. Moreover, knockdown of EGR2 promoted gastric cancer cell growth and inhibited their apoptosis by flow cytometry detection. Altogether, these findings revealed miR-25 as a regulator of gastric cancer cell growth and apoptosis through targeting EGR2.

Keywords: gastric cancer, MiR-25, cell growth, apoptosis, Egr2

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Lihong Wang,
Southeast University, China
Somnath Tagore,
Columbia University, United States

****Correspondence:***
Jing Jia
jiajing0810@126.com

Specialty section:
This article was submitted to RNA,
a section of the journal Frontiers in Genetics

Received: *02 April 2021*
Accepted: *01 October 2021*
Published: *26 October 2021*

Citation:
Yang L, Li L, Chang P, Wei M, Chen J, Zhu C and Jia J (2021) miR-25 Regulates Gastric Cancer Cell Growth and Apoptosis by Targeting EGR2. Front. Genet. 12:690196. doi: 10.3389/fgene.2021.690196

INTRODUCTION

Gastric cancer is one of the most common malignancies of the digestive system, which endangers human life and health (Bertuccio et al., 2009; Ferlay et al., 2015). In the last few decades, many achievements have been made in gastric cancer -related proto-oncogenes, suppressor genes, and related signal transduction pathways, but the exact mechanism of tumorigenesis has not been fully elucidated. The emergence of a new class of non-protein coding microRNA (miRNA) provides a new idea for gastric cancer research. MiRNAs play an important role in the pathogenesis of tumors, which provides a potential new strategy for understanding the clinical therapeutics of gastric cancer (Li et al., 2011). MiRNAs generally bind to their target mRNAs through complementary pairing, thus regulating the expression of target genes at the post-transcriptional level (Brennecke et al., 2003).

Recent studies have shown that miRNA plays an important role in the pathogenesis and progression of gastric cancer, which provides us with a potential direction for the study of therapeutic targets for this major disease (Huang et al., 2016; Zhang et al., 2019). Studies have shown that abnormal expression of miRNAs plays a crucial role in the pathogenesis and malignant biological phenotypes of gastric cancer, such as proliferation, apoptosis, invasion, and migration of gastric cancer cells (Wang et al., 2016; Zhang et al., 2016; Zhang et al., 2017a; Zhang et al., 2017b). MiR-25 is widely expressed in various types of tumors, and previous studies have shown that its expression level is related to the clinical characteristics and

TABLE 1 | The gene primer sequence for RT-qPCR or PCR.

Gene name	Forward primer	Reverse primer
miR-25 RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCAGAC	
miR-25	AGTCTGGCTCTGTTCACG	CCAGTGCAGGGTCCGAGGTA
Bax	CCCGAGAGGTCTTTTTCCGAG	CCAGCCCATGATGGTTCTGAT
p53	CAGCACATGACGGAGGTTGT	TCATCCAAATACTCCACACGC
Bcl-2	GGTGGGGTCATGTGTGTGG	CGGTTCAGGTACTCAGTCATCC
EGR2	TCAACATTGACATGACTGGAGAG	AGTGAAGGTCTGGTTTCTAGGT
EGR2 (3′UTR wild)	GCTCGAGGACCCTGGATGTCAGAGTTG	GGCGGCCGCCCTCTGAGATCATAAATGC
EGR2 (3′UTR mutant)	CAGAGAACAGAAGGTCCATGTGATGGG	CCCATCACATGGACCTTCTGTTCTCTG

prognosis of tumors (Xu et al., 2014; Mingjun, 2015; Peng et al., 2015; Wang et al., 2015). However, the role of miR-25 in different tumors is consistent but the target genes are different (Caiazza and Mallardo, 2018). Studies have reported that miR-25 plays an oncogene role in lung cancer (Wu et al., 2015), liver cancer (Wang et al., 2015; Tian and Yao, 2016; Avencia et al., 2019), and gastric cancer (Li et al., 2015). MiRNA-25 had been investigated in gastric cancer via targeting RECK (Zhao et al., 2014), but we found that there might be other targets. Therefore, in this study, we explored the biological effect and significance of miR-25 in the tumorigenesis of gastric cancer, looking for new targets and hoping to provide a new experimental basis for understanding the pathogenesis and therapeutics of gastric cancer.

In this study, we detected the expression feature of miR-25 in normal gastric epithelial cells and gastric cancer cells by RT-qPCR. The result revealed that the expression level of miR-25 in gastric cancer cells was significantly higher than that in normal cells. Overexpression of miR-25 promoted gastric cancer cells' growth and inhibited their apoptosis. Mechanically, we demonstrated that *EGR2* was a target gene of miR-25 by dual-luciferase reporter gene assay in gastric cancer cells.

MATERIAL AND METHOD

Cell Culture

The cell line GES-1 cells (ATCC) and BGC823 cells (ATCC) were preserved in this laboratory. The cells were placed in DMEM medium containing 10% fetal bovine serum, 1×10^5 IU/L penicillin, and 1×10^5 IU/L streptomycin conventional culture in an incubator at 37°C and 5% CO_2.

Cell Transfection

One day before transfection, cells were transferred to a suitable density so that the cells were in a logarithmic growth phase at the time of transfection. Firstly, DMEM was used to culture cells for about 4 h excluding serum and antibiotics. Then, negative control, miRNA mimic or inhibitor, and siRNA were transfected into cell by Lipofectamine 2000 (Invitrogen) according to operation instructions. After transfection for 6 h, the fresh medium was replaced and cultured for a continuous 48 h for further experiments.

Real Time Quantitative PCR

The total RNA was extracted with Trizol kit, and the concentration of RNA was determined by Nucleic acid analyzer. The first strand cDNA of miRNA was synthesized by reverse transcription reaction of stem-loop primers. The SYBR GREEN reagent was used for PCR amplification according to kit instructions. The real-time quantitative PCR (RT-qPCR) reaction was performed on an ABI PCR instrument and the results were analyzed. All primer sequences were listed in **Table 1**.

Western Blot

Proteins of each group were extracted and quantified with BCA protein quantitative kit. After SDS-PAGE, the proteins were transferred to PVDF membrane by electrol transfer. The product was sealed with 5% skim milk at room temperature for 60 min, separately added with corresponding primary antibody, incubated at 4°C overnight, and washed with TBST three times. Then, the secondary antibody was added at room temperature for 2 h incubation, and then washed with TBST three times. After exposure in the darkroom, the results were processed by Quantity One gel analysis software. Primary antibody Bax (21kDa, source: mouse), Bcl-2(26kDa, source: mouse) and β-actin (42kDa, source: mouse) were purchased from Proteintech. β-actin acts as the internal gene.

CCK-8 Assay

About 5×10^4/ml cell suspensions were prepared. 100 μl cell suspension was added to the 96-well plate and continued to be cultured in an incubator at 37°C. At different time points, 10 μl CCK-8 solution was added to each well and incubated at 37°C for 1–4 h. The absorbance value at 450 nm was determined.

Cell Cycle Assay

To detect the cell cycle assay, gastric cancer cells were fixed with 70% ethanol for about 24 h at −20°C. The cells were labeled with propidium iodide (PI) for 30 min at 37°C and detected in the BD FACSVerse flow cytometer. The results of cell cycle were analyzed using the cell cycle analysis software. The detailed experimental process follows the manufacturer's instructions (Multisciences, Hangzhou, China).

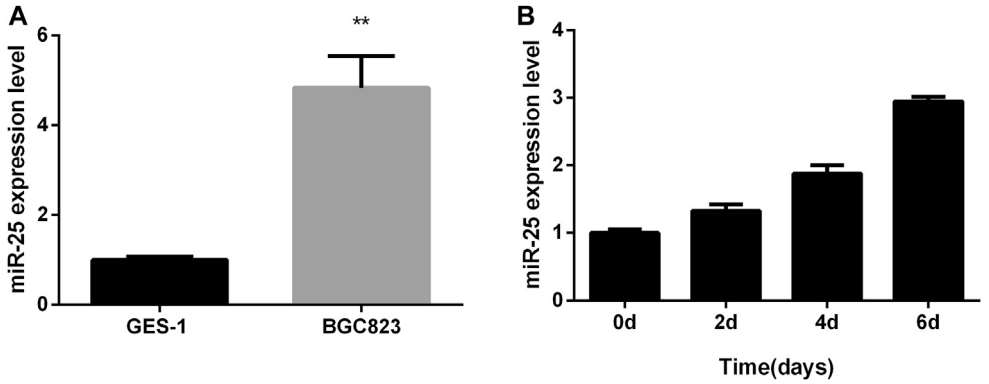


FIGURE 1 | The expression profile of miR-25 in BGC823 cell. **(A)** The expression level of miR-25 was detected in BGC823 cell and GES-1 cell by RT-qPCR. **(B)** In different stages (0, 2, 4, 6 days) of culture, miR-25 expression level was detected by RT-qPCR. Value in graphs represents means ± SD. At least three independent experiments were carried out. $^{**}p < 0.01$.

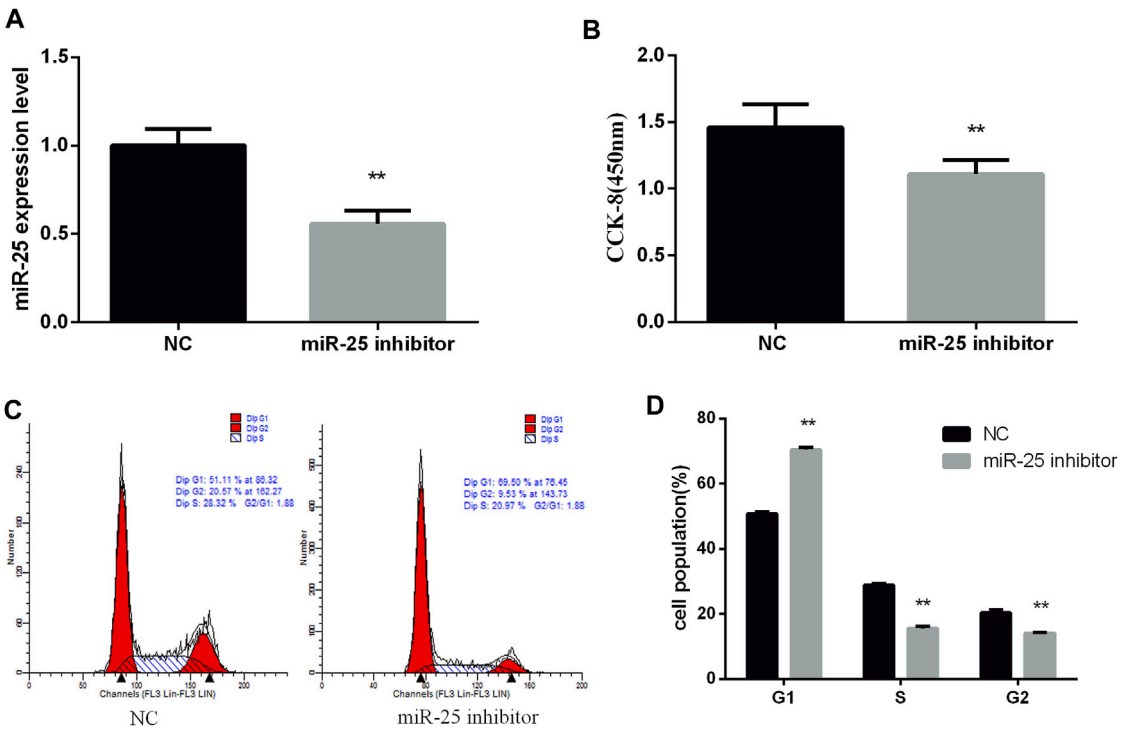


FIGURE 2 | Inhibition of miR-25 inhibits BGC823 cell growth. **(A)** After transfection of miR-25 inhibitor into BGC823 cell, the expression level of miR-25 was detected by RT-qPCR. **(B)** After loss of miR-25, CCK-8 reagent was used to detect cell proliferation activity. **(C)** Cell cycle were analyzed by flow cytometry after being transfected with miR-25 inhibitor. **(D)** Statistical results of flow cytometry. Value in graphs represents means ± SD. At least three independent experiments were carried out. $^{**}p < 0.01$.

Cell Apoptosis Assay

The cells were collected directly into a 10 ml centrifuge tube. The rinse solution was washed once, and centrifuged at 800 r/min for 5 min. The cells were resuspended in solution and incubated in the dark for 15 min at room temperature. The precipitated cells were centrifuged at 800 r/min for 5 min and incubated with buffer solution once. The fluorescent dye solution was added and incubated at 4°C for 20 min to avoid light. The wavelength of excitation light in the flow cytometer was 488 nm, the wavelength of FITC was detected by a 515 nm passband filter, and the other wave length was greater than 560 nm to detect PI. On the scatter diagram of bivariate flow cytometry, living cells are shown in the lower left quadrant, which is (FITC-/PI-). The upper right quadrant is the non-living cells, i.e., dead cells, which are (FITC +/PI +). The lower right quadrant for apoptotic cells is (FITC +/PI-).

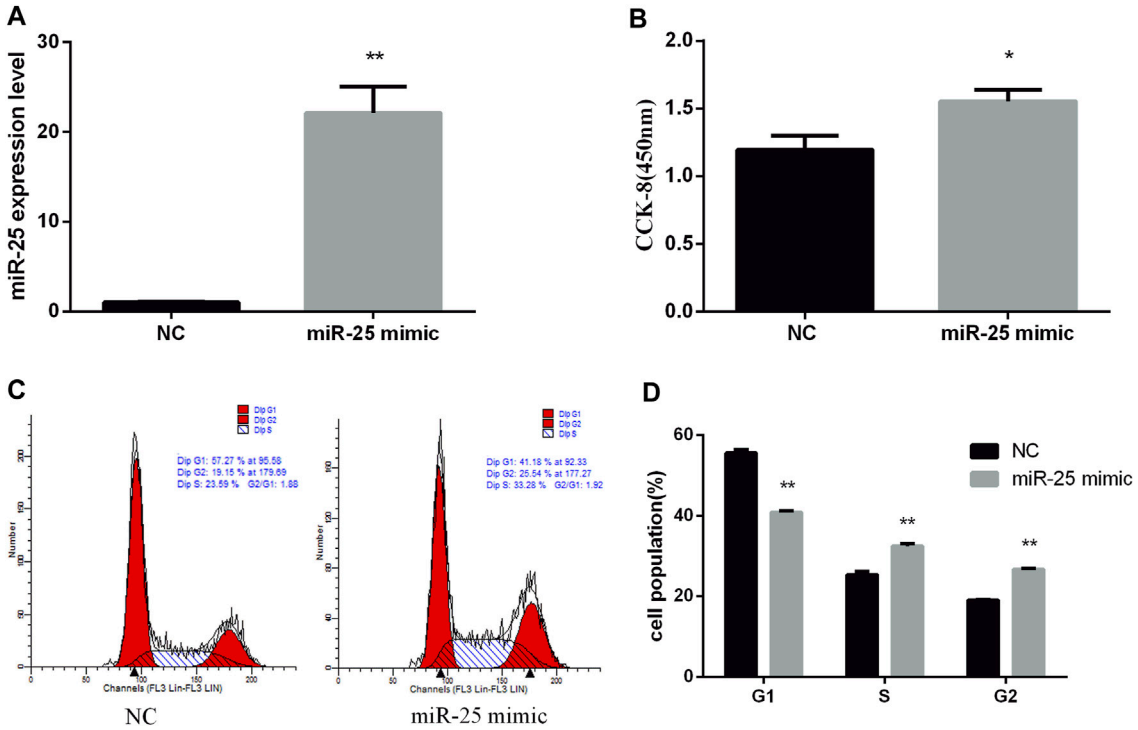


FIGURE 3 | Overexpression of miR-25 promotes BGC823 cell growth. **(A)** After transfection of miR-25 mimic into BGC823 cell, the expression level of miR-25 was detected by RT-qPCR. **(B)** After gain of miR-25, CCK-8 reagent was used to detect cell proliferation activity. **(C)** Cell cycle were analyzed by flow cytometry after being transfected with miR-25 mimic. **(D)** Cell number statistics at each stage of cell cycle. Value in graphs represents means ± SD. At least three independent experiments were carried out. $^{**}p < 0.01$; $^{*}p < 0.05$.

Luciferase Report Assay

The target genes predictions were revealed by bioinformatics software Targetscan 7.2 (http://www.targetscan.org/vert_72/) online. The wild-type 3′UTR or mutant-type 3′UTR of EGR2 mRNA sequence were inserted into luciferase reporter plasmids psi-CHECK2™ (Promega). EGR2 wild-type or mutant-type luciferase reporter vector and miR-25 mimic or negative control were co-transfected into 293T cells. Following the manufacturer's instruction of Dual Luciferase Assay System (Promega), luciferase activities were measured.

Statistics Analysis

All data were expressed as mean ± standard deviation. SPSS18.0 software was used for statistical analysis, and student' t-test was used for comparison between groups analysis of variance. $^{*}p < 0.05$, $^{**}p < 0.01$.

RESULT

Expression of miR-25 in Normal Gastric Epithelial Cells and Gastric Cancer Cells

Total RNA from *GES-1* cells and *BGC823* cells were extracted at the growth stage, respectively. Real-time quantitative PCR (RT-qPCR) result showed that the expression level of miR-25 in *BGC823* cells was significantly higher than that in *GES-1* cell (**Figure 1A**). Further RT-qPCR results showed that the expression of miR-25 in *BGC823* gastric cancer cells at the growth stage was found to have an upward trend (**Figure 1B**). All these results indicated that the expression characteristic of miR-25 might be related to the occurrence of gastric cancer.

miR-25 Promotes Gastric Cancer Cells Growth

To reveal the function of endogenous miR-25 for BGC823 gastric cancer cells' growth, we transfected miR-25 inhibitor into cells to affect its expression level. RT-qPCR result showed that the expression level of miR-25 was significantly decreased compared to negative control group (**Figure 2A**). The proliferation activity of BGC823 cells was significantly attenuated by CCK-8 detection (**Figure 2B**). Flow cytometry detection of cell cycle showed that G_1 phase cell number increased and S phase cell number decreased (**Figures 2C,D**). Meanwhile, we also revealed the effect of exogenous miR-25 on BGC823 cell proliferation. MiR-25 mimic was transfected into BGC823 cell, which could enhance the expression level of miR-25 (**Figure 3A**). Compared to the control group, overexpression of miR-25 enhanced the BGC823 cell activity by CCK-8 assay (**Figure 3B**). In addition, overexpression of miR-25 resulted in G_1 phase cell population reducing and S phase cell population increasing (**Figures 3C,D**). Together, these

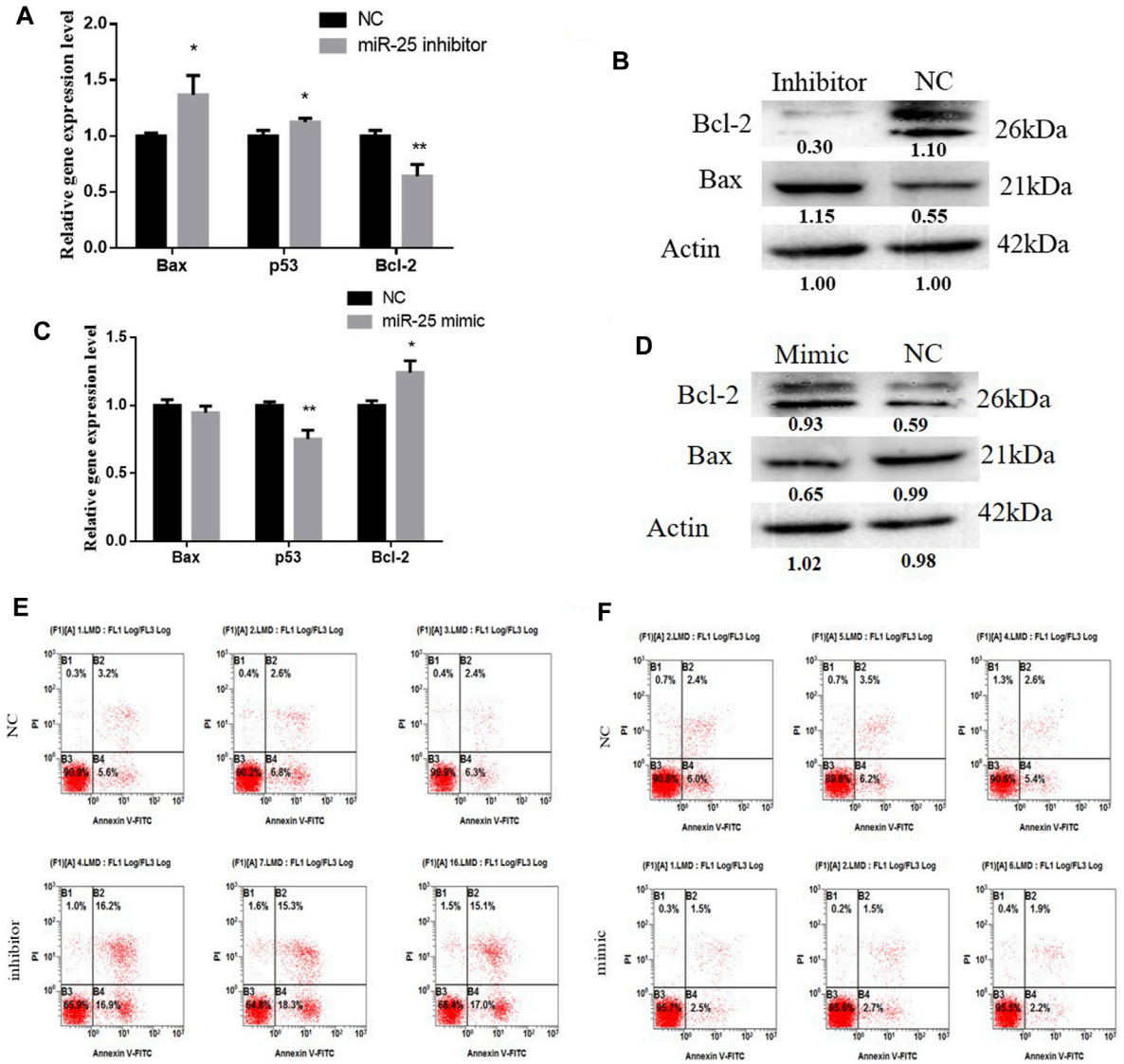


FIGURE 4 | The function of miR-25 for BGC823 cell apoptosis. **(A)** After loss of miR-25, genes (Bax, p53, Bcl-2) associated with apoptosis were detected by RT-qPCR. **(B)** After loss of miR-25, Bax and Bcl-2 protein level were detected by western blot. **(C)** After gain of miR-25, genes (Bax, p53, Bcl-2) associated with apoptosis were detected by RT-qPCR. **(D)** After gain of miR-25, Bax and Bcl-2 protein level were detected by western blot. **(E)** After loss of miR-25, cell apoptosis was detected by Annexin V-FITC reagent. **(F)** After gain of miR-25, cell apoptosis was detected by Annexin V-FITC reagent. Value in graphs represents means ± SD. At least three independent experiments were carried out. $^{**}p < 0.01$; $^{*}p < 0.05$.

data revealed that miR-25 could promote BGC823 gastric cancer cells' growth.

miR-25 Inhibits Gastric Cancer Cells Apoptosis

The function of miRNAs in the apoptosis of gastric cancer cells might provide an effective way to treat tumorigenesis. To elucidate the role of miR-25 for BGC823 cell apoptosis, we used molecular biology techniques and cell biology techniques to detect its effects. Firstly, after transfection of miR-25 mimic or inhibitor into BGC823 cell, we used RT-qPCR to detect the mRNA expression level of Bax, Bcl-2, and p53. Compared to the NC group, knockdown of miR-25 enhanced Bax and p53 expression level, but impeded Bcl-2 expression level (**Figure 4A**). The western blot analysis results indicated that knockdown of miR-25 enhanced Bax expression level and reduced Bcl-2 expression (**Figure 4B**). By contrast, overexpression of miR-25 reduced the Bax and p53 expression level, but induced Bcl-2 expression level (**Figure 4C**). Bax protein level was decreased and Bcl-2 protein level was increased (**Figure 4D**). Then, we used

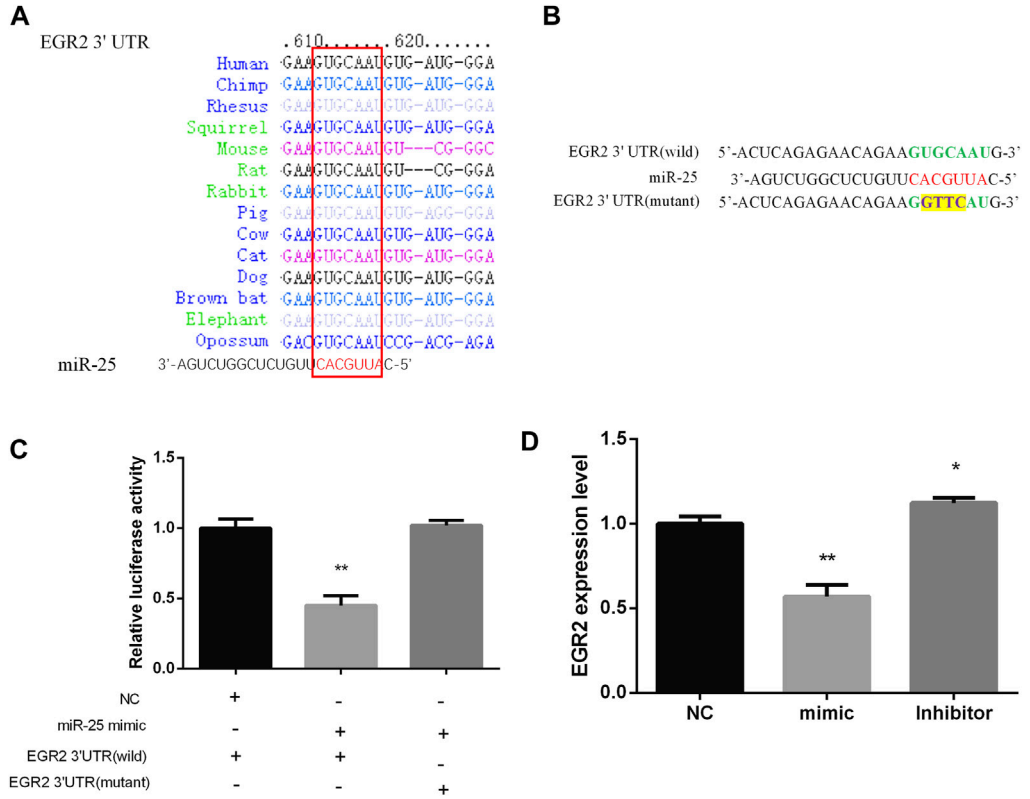


FIGURE 5 | EGR2 is a target gene of miR-25. **(A)** miR-25 is highly conserved in various species. **(B)** Dual-luciferase reporter vectors containing the 3′UTR sequence of EGR2 gene (wild type and mutant type), respectively. **(C)** Luciferase activity analysis was detected after co-transfection miR-25 and dual-luciferase reporter vectors into cell. **(D)** EGR2 mRNA level was detected after gain or loss miR-25. Value in graphs represents means ± SD. At least three independent experiments were carried out. $^{**}p < 0.01$; $^{*}p < 0.05$.

Annexin V-FITC/PI dual-staining apoptosis detection kit to detect gastric cancer cells' apoptosis after gain or loss of miR-25. The result revealed that knockdown of miR-25 induced the number of gastric cancer cells' apoptosis (**Figure 4E**). Meanwhile, overexpression of miR-25 reduced the number of apoptotic cells (**Figure 4F**). Together, these results demonstrated that miR-25 could inhibit gastric cancer cells' apoptosis.

EGR2 is a Target Gene of miR-25 in Gastric Cancer Cells

Bioinformatics methods and a series of experiments were used to screen and verify target genes of miR-25. EGR2 caught our attention by Targetscan software prediction, the 3 ′UTR sequence of EGR2 is completely complementary to the seed region sequence of miR-25, and is also highly conserved in various species (**Figure 5A**). In order to identify the targeting relationship between miR-25 and EGR2, we constructed dual-luciferase reporter vectors containing the 3′UTR sequence of EGR2 gene (wild type and mutant type) respectively (**Figure 5B**). Firstly, miR-25 mimic was co-transfected with a dual-luciferase reporter vector containing the 3′UTR sequence of EGR2 gene (wild type or mutant type). The result revealed that overexpression of miR-25 reduced the luciferase activity of EGR2 gene 3 ′UTR wild-type reporter vector, but it did not affect the mutant vector activity (**Figure 5C**). Moreover, we also detected EGR2 mRNA level by RT-qPCR after gain or loss of miR-25. The results indicated that overexpression of miR-25 reduced EGR2 mRNA expression level, but knockdown of miR-25 enhanced EGR2 mRNA expression level (**Figure 5D**). These results revealed that EGR2 was a target gene of miR-25 in gastric cancer cells.

Knockdown of EGR2 Counteracts the Role of miR-25 Inhibitor

To investigate the function of EGR2 in gastric cancer cells' growth and apoptosis, we used EGR2 siRNA to reduce the expression of EGR2 (**Figure 6A**). First, the cell cycle assay results indicated that knockdown of EGR2 reduced the number of G1 phase and elevated the S phase cell population (**Figures 6B,C**). Meanwhile, we also detected the number of apoptosis cells' changes by Annexin V-FITC/PI dual-staining apoptosis detection kit. The result showed that knockdown of EGR2 could inhibit gastric cancer cells' apoptosis (**Figure 6D**). Together, these

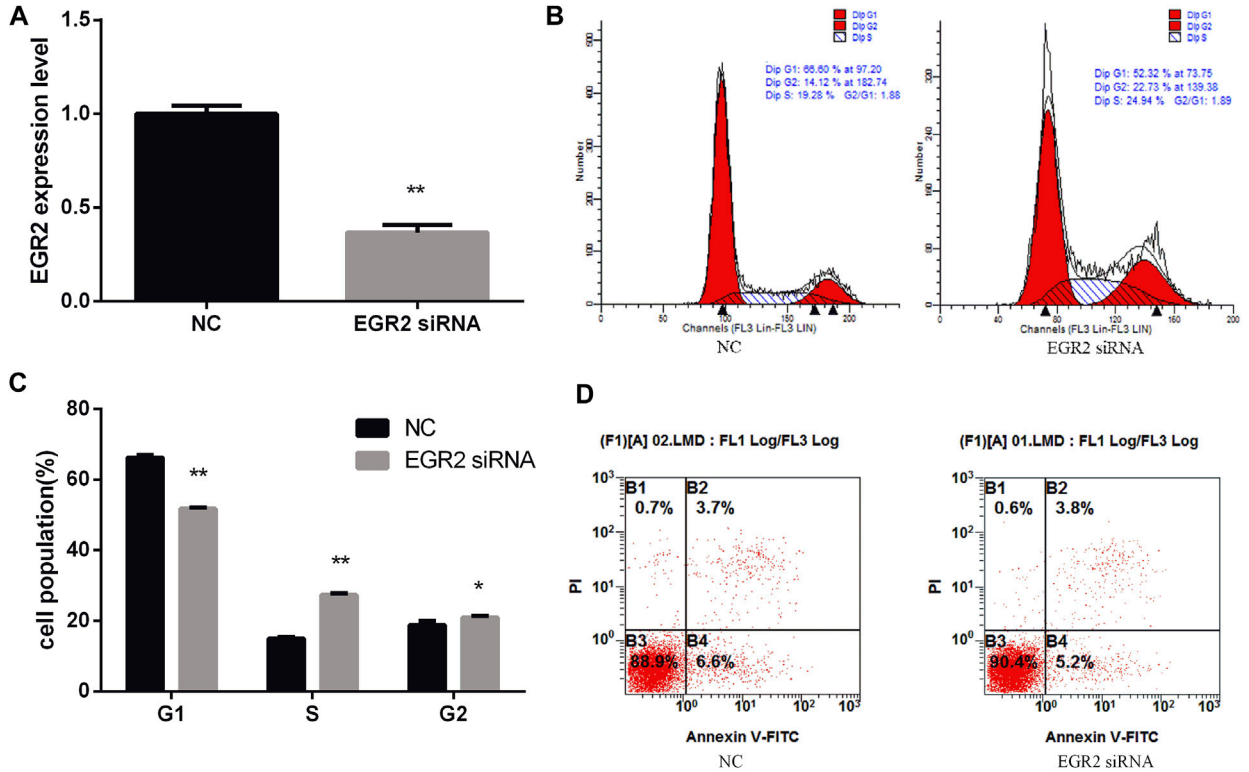


FIGURE 6 | Knockdown of EGR2 promotes BGC823 cell growth and inhibits cell apoptosis. **(A)** After knockdown of EGR2, the expression level of EGR2 was detected by RT-qPCR. **(B)** Cell cycle were analyzed by flow cytometry after being transfected with EGR2 siRNA. **(C)** Cell number statistics at each stage of cell cycle. **(D)** After knockdown of EGR2, cell apoptosis was detected by Annexin V-FITC reagent. Value in graphs represents means ± SD. At least three independent experiments were carried out. $^{**}p < 0.01$; $^{*}p < 0.05$.

results demonstrated that knockdown of EGR2 promotes gastric cancer cells' growth and inhibits their apoptosis. To further reveal the relationship of miR-25 and EGR2, an antagonistic experiment was designed. miR-25 inhibitor and EGR2 siRNA were co-transfected into gastric cancer cells. The cell cycle assay results revealed that there was no significant change in the number of cells at each stage (**Figures 7A,B**). Compared to the control group, there was also no significant difference in the number of apoptosis cells (**Figure 7C**). These data demonstrated that Knockdown of EGR2 counteracts the role of miR-25 inhibitor in gastric cancer cells.

DISCUSSION

In the present study, we found a differentially expressed miRNA in BGC823 gastric cancer cells and normal gastric mucosa cell (GES-1 cells and RGM-1 cells): miR-25. Overexpression of miR-25 promoted BGC823 gastric cancer cells and inhibited their apoptosis. EGR2 was a target gene of miR-25. Knockdown of EGR2 could promote gastric cancer cells' growth and inhibit their apoptosis. Therefore, miR-25 might be used as a potential biomarker or therapeutic target for gastric cancer.

MiRNAs are a class of widely distributed endogenous non-coding RNA, with a size of 20–25 nucleotide, which are highly conserved genetically and can negatively or positively regulate the expression of their target genes (Gambari et al., 2011). Studies have found that miRNAs can be used as oncogenes or tumor suppressor genes to participate in the regulation of cell growth, apoptosis, and cycle (Voorhoeve et al., 2006; Manikandan et al., 2008). Some miRNAs are directly involved in the formation of human tumors (such as lung cancer, breast cancer, craniocerebral tumor, liver cancer, colorectal cancer, and lymphoma). MiRNAs can be used as both oncogenes and tumor suppressor genes to participate in multiple signaling pathways of human tumor formation (Esquela-Kerscher and Slack, 2006). Therefore, the study of specific miRNA function provides a new direction for tumor treatment and prevention.

The role of miR-25 in the pathogenesis of malignant tumors has been studied frequently. In human ovarian cancer, up-regulation of miR-25 could enhance cell proliferation and down-regulation of miR-25 could enhance the expression of pro-apoptotic proteins such as Bax and caspase-3 (Zhang et al., 2012). In non-small cell lung cancer (NSCLC), the expression level of miR-25 was up-regulated. Over-expression of miR-25 could significantly increase NSCLC cells' proliferation, migration, and invasion, but down-regulation of miR-25 remarkably reduced cell proliferation, migration, and invasion

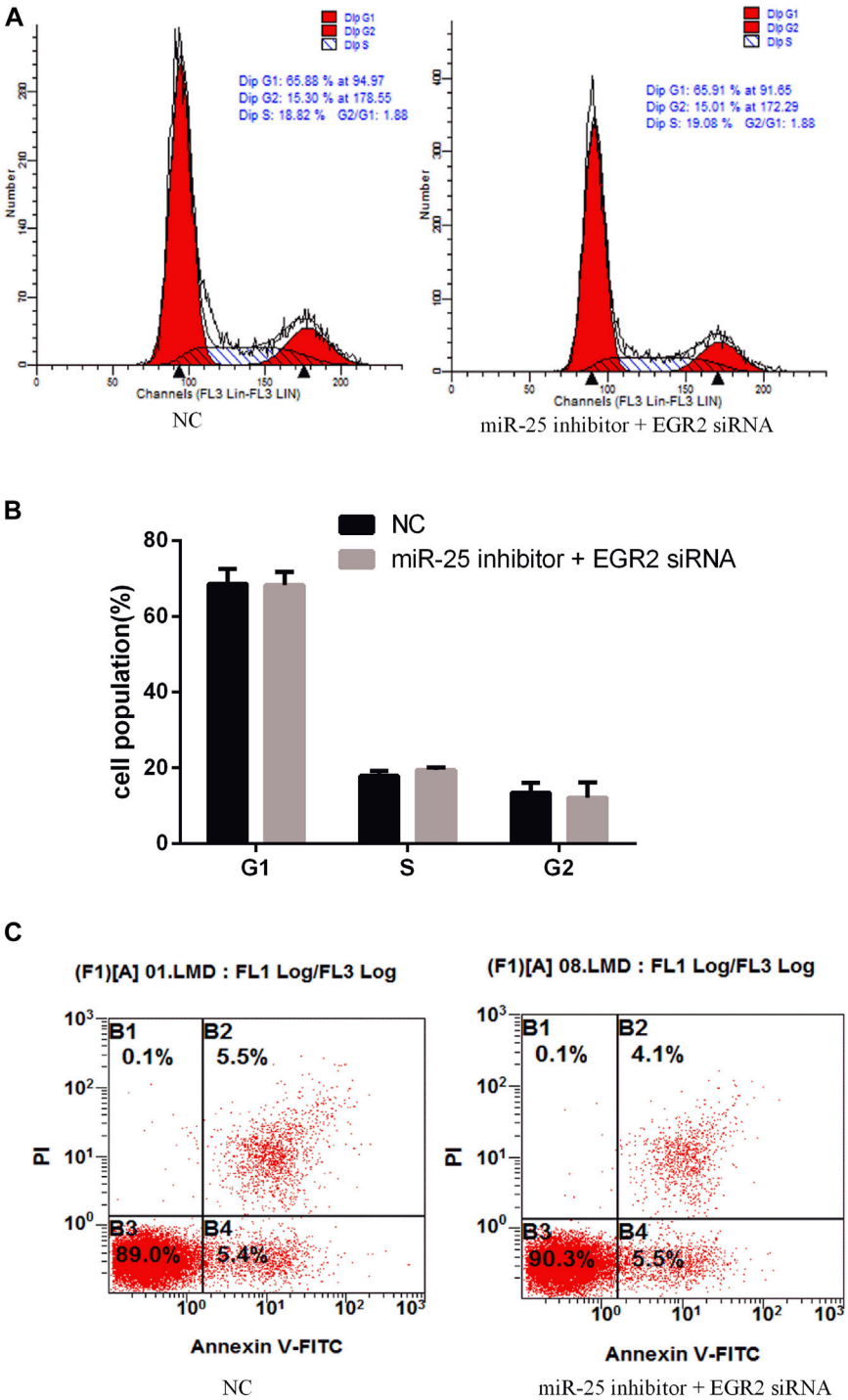


FIGURE 7 | Knockdown of EGR2 counteracts the role of miR-25 inhibitor. **(A)** Cell cycle were analyzed by flow cytometry after being co-transfected with miR-25 inhibitor and EGR2 siRNA. **(B)** Cell number statistics at each stage of cell cycle. **(C)** After being co-transfected miR-25 inhibitor and EGR2 siRNA, cell apoptosis was detected by Annexin V-FITC reagent.

in NSCLC cells. Further results demonstrated that FBXW7 was a direct target gene of miR-25 (Xiang et al., 2015). Moreover, in gastric cancer, overexpression of miR-25 could promote gastric cancer cell proliferation, migration, and invasion via targeting RECK (Zhao et al., 2014). This is consistent with our current findings.

In conclusion, we found that a high expression of miR-25 in BGC823 gastric cancer cells can enhance cell growth and reduce cell apoptosis via targeting gene *EGR2*. The result revealed that miR-25 might be a potential therapeutic target for gastric cancer.

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.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This research was supported by the Education Department Service Local Special Scientific Research Projects of Shaanxi Province, China (20JC030) and the Key scientific research projects of Shaanxi Provincial Education Department in 2020(20JS135).

REFERENCES

Avencia, S. M., Kwon, J., Hong Chew, X., Siemens, A., Sohn, H. S., Jing, G., et al. (2019). A Novel SOCS5/miR-18/miR-25 Axis Promotes Tumorigenesis in Liver Cancer. *Int. J. Cancer* 144 (2), 311–321. doi:10.1002/ijc.31857

Bertuccio, P., Chatenoud, L., Levi, F., Praud, D., Ferlay, J., Negri, E., et al. (2009). Recent Patterns in Gastric Cancer: A Global Overview. *Int. J. Cancer* 125 (3), 666–673. doi:10.1002/ijc.24290

Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., and Cohen, S. M. (2003). Bantam Encodes a Developmentally Regulated microRNA that Controls Cell Proliferation and Regulates the Proapoptotic Gene Hid in Drosophila. *Cell* 113, 25–36. doi:10.1016/s0092-8674(03)00231-9

Caiazza, C., and Mallardo, M. (2018). The Roles of miR-25 and its Targeted Genes in Development of Human Cancer. *Microrna* 5 (2), 113–119. doi:10.2174/2211536605666160905093429

Esquela-Kerscher, A., and Slack, F. J. (2006). Oncomirs - microRNAs with a Role in Cancer. *Nat. Rev. Cancer* 6 (4), 259–269. doi:10.1038/nrc1840

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer Incidence and Mortality Worldwide: Sources, Methods and Major Patterns in GLOBOCAN 2012. *Int. J. Cancer* 136 (5), E359–E386. doi:10.1002/ijc.29210

Gambari, R., Fabbri, E., Borgatti, M., Lampronti, I., Finotti, A., Brognara, E., et al. (2011). Targeting microRNAs Involved in Human Diseases: A Novel Approach for Modification of Gene Expression and Drug Development. *Biochem. Pharmacol.* 82 (10), 1416–1429. doi:10.1016/j.bcp.2011.08.007

Huang, Z., Zhu, D., Wu, L., He, M., Zhou, X., Zhang, L., et al. (2016). Six Serum-Based miRNAs as Potential Diagnostic Biomarkers for Gastric Cancer. *Cancer Epidemiol. Biomarkers Prev.* 26, 188–196. doi:10.1158/1055-9965.epi-16-0607

Li, X., Zhang, Y., Zhang, H., Liu, X., Gong, T., Li, M., et al. (2011). miRNA-223 Promotes Gastric Cancer Invasion and Metastasis by Targeting Tumor Suppressor EPB41L3. *Mol. Cancer Res.* 9 (7), 824–833. doi:10.1158/1541-7786.mcr-10-0529

Li, B.-S., Zuo, Q.-F., Zhao, Y.-L., Xiao, B., Zhuang, Y., Mao, X.-H., et al. (2015). MicroRNA-25 Promotes Gastric Cancer Migration, Invasion and Proliferation by Directly Targeting Transducer of ERBB2, 1 and Correlates with Poor Survival. *Oncogene* 34 (20), 2556–2565. doi:10.1038/onc.2014.214

Manikandan, J., Aarthi, J. J., Kumar, S. D., and Pushparaj, P. N. (2008). Oncomirs: The Potential Role of Non-Coding microRNAs in Understanding Cancer. *Bioinformation* 2 (8), 330–334. doi:10.6026/97320630002330

Mingjun, Z. (2015). miR-107 and miR-25 Simultaneously Target LATS2 and Regulate Proliferation and Invasion of Gastric Adenocarcinoma (GAC) Cells. *Biochem. Biophysical Res. Commun.* 460 (3), 806–812. doi:10.1016/j.bbrc.2015.03.110

Peng, G., Yuan, X., Yuan, J., Liu, Q., Dai, M., Shen, C., et al. (2015). miR-25 Promotes Glioblastoma Cell Proliferation and Invasion by Directly Targeting NEFL. *Mol. Cell Biochem.* 409 (1-2), 103–111. doi:10.1007/s11010-015-2516-x

Tian, C., and Yao, S. (2016). *MicroRNA-25 Promotes Growth and Migration of Liver Cancer Cells by Regulating Klf4 Gene Expression*. Shanghai, China: Tumor.

Voorhoeve, P. M., le Sage, C., Schrier, M., Gillis, A. J. M., Stoop, H., Nagel, R., et al. (2006). A Genetic Screen Implicates miRNA-372 and miRNA-373 as Oncogenes in Testicular Germ Cell Tumors. *Cell* 124 (6), 1169–1181. doi:10.1016/j.cell.2006.02.037

Wang, C., Wang, X., Su, Z., Fei, H., Liu, X., and Pan, Q. (2015). miR-25 Promotes Hepatocellular Carcinoma Cell Growth, Migration and Invasion by Inhibiting RhoGDI1. *Oncotarget* 6 (34), 36231–36244. doi:10.18632/oncotarget.4740

Wang, L. L., Wang, L., Wang, X. Y., Shang, D., Yin, S. J., and Sun, L. L. (2016). MicroRNA-218 Inhibits the Proliferation, Migration, and Invasion and Promotes Apoptosis of Gastric Cancer Cells by Targeting LASP1. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* 37, 15241–15252. doi:10.1007/s13277-016-5388-0

Wu, T., Chen, W., Kong, D., Li, X., H. LuLiu, S., et al. (2015). miR-25 Targets the Modulator of Apoptosis 1 Gene in Lung Cancer. *Carcinogenesis* 36 (8), 925–935. doi:10.1093/carcin/bgv068

Xiang, J., Hang, J. B., Che, J. M., and Li, H. C. (2015). miR-25 Is Up-Regulated in Non-Small Cell Lung Cancer and Promotes Cell Proliferation and Motility by Targeting FBXW7. *Int. J. Clin. Exp. Pathol.* 8 (8), 9147–9153.

Xu, F.-X., Su, Y.-L., Zhang, H., Kong, J.-Y., Yu, H., and Qian, B.-Y. (2014). Prognostic Implications for High Expression of MiR-25 in Lung Adenocarcinomas of Female Non-Smokers. *Asian Pac. J. Cancer Prev.* 15 (3), 1197–1203. doi:10.7314/apjcp.2014.15.3.1197

Zhang, H., Zuo, Z., Lu, X., Wang, L., Wang, H., and Zhu, Z. (2012). MiR-25 Regulates Apoptosis by Targeting Bim in Human Ovarian Cancer. *Oncol. Rep.* 27 (2), 594–598. doi:10.3892/or.2011.1530

Zhang, R., Li, F., Wang, W., Wang, X., Li, S., and Liu, J. (2016). The Effect of Antisense Inhibitor of miRNA 106b~25 on the Proliferation, Invasion, Migration, and Apoptosis of Gastric Cancer Cell. *Tumor Biol.* 37, 10507–10515. doi:10.1007/s13277-016-4937-x

Zhang, Y. Q., Sun, Q., Lou, X., Jin, C., and Zhang, Y. (2017). Molecular Mechanisms of *Helicobacter pylori* Infection Related miRNA-148b in Proliferation and Invasion of Gastric Cancer Cells. *Chin. J. Nosocomiology* 2177 –2182. doi:10.11816/cn.ni.2017-163263

Zhang, X. T., Zhang, X.-Y., Huang, S.-Y., Wang, Q., and Liu, W. (2017). miRNA-101-3p Inhibits Proliferation and Migration of Gastric Cancer Cells by Targeting EZH2. *Chin. J. Pathophysiology* 33, 2143–2150. doi:10.3969/j.issn.1000-4718.2017.12.006

Zhang, Y., Sun, B., Zhao, L., Liu, Z., Xu, Z., Tian, Y., et al. (2019). Up-Regulation of miRNA-148a Inhibits Proliferation, Invasion, and Migration while Promoting Apoptosis of Cervical Cancer Cells by Down-Regulating RRS1. *Biosci. Rep.* 39 (5), BSR20181815. doi:10.1042/BSR20181815

Zhao, H., Wang, Y., Yang, L., Jiang, R., and Li, W. (2014). MiR-25 Promotes Gastric Cancer Cells Growth and Motility by Targeting RECK. *Mol. Cel Biochem.* 385 (1-2), 207–213. doi:10.1007/s11010-013-1829-x

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 03 December 2021
doi: 10.3389/fgene.2021.778742

AC093797.1 as a Potential Biomarker to Indicate the Prognosis of Hepatocellular Carcinoma and Inhibits Cell Proliferation, Invasion, and Migration by Reprogramming Cell Metabolism and Extracellular Matrix Dynamics

Xiaoling Liu[1,2,3†], *Chenyu Wang*[1,2,3†], *Qing Yang*[4], *Yue Yuan*[1,2,3], *Yunjian Sheng*[1,2,3], *Decheng Li*[1,2,3], *Suvash Chandra Ojha*[1,2,3], *Changfeng Sun*[1,2,3*] *and Cunliang Deng*[1,2,3*]

[1]*The Department of Infectious Diseases, The Affiliated Hospital of Southwest Medical University, Luzhou, China,* [2]*The Department of Tuberculosis, The Affiliated Hospital of Southwest Medical University, Luzhou, China,* [3]*Laboratory of Infection and Immunity, The Affiliated Hospital of Southwest Medical University, Luzhou, China,* [4]*The Department of Gastroenterology, The Second People's Hospital of Neijiang, Neijiang, China*

Purpose: The risk signature composed of four lncRNA (AC093797.1, POLR2J4, AL121748.1, and AL162231.4.) can be used to predict the overall survival (OS) of patients with hepatocellular carcinoma (HCC). However, the clinical significance and biological function of AC093797.1 are still unexplored in HCC or other malignant tumors. In this study, we aimed to investigate the biological function of AC093797.1 in HCC and screen the candidate hub genes and pathways related to hepatocarcinogenesis.

Methods: RT-qPCR was employed to detect AC093797.1 in HCC tissues and cell lines. The role of AC093797.1 in HCC was evaluated *via* the cell-counting kit-8, transwell, and wound healing assays. The effects of AC093797.1 on tumor growth *in vivo* were clarified by nude mice tumor formation experiments. Then, RNA-sequencing and bioinformatics analysis based on subcutaneous tumor tissue was performed to identify the hub genes and pathways associated with HCC.

Results: The expression of AC093797.1 decreased in HCC tissues and cell lines, and patients with low expressed AC093797.1 had poor overall survival (OS). AC093797.1 overexpression impeded HCC cell proliferation, invasion, and migration *in vitro* and suppressed tumor growth *in vivo*. Compared with the control group, 710 differentially expressed genes (243 upregulated genes and 467 downregulated genes) were filtered *via* RNA-sequencing, which mainly enriched in amino acid metabolism, extracellular matrix structure constituents, cell adhesion molecules cams, signaling to Ras, and signaling to ERKs.

OPEN ACCESS

Edited by:
Wei Jiang,
Nanjing University of Aeronautics and Astronautics, China

Reviewed by:
Chang Zeng,
Northwestern University, United States
Xuexin Yu,
University of Texas Southwestern Medical Center, United States

***Correspondence:**
Changfeng Sun
sun_cf88@swmu.edu.cn
Cunliang Deng
dengcunl@swmu.edu.cn

[†]*These authors share first authorship*

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *17 September 2021*
Accepted: *28 October 2021*
Published: *03 December 2021*

Citation:
Liu X, Wang C, Yang Q, Yuan Y, Sheng Y, Li D, Ojha SC, Sun C and Deng C (2021) AC093797.1 as a Potential Biomarker to Indicate the Prognosis of Hepatocellular Carcinoma and Inhibits Cell Proliferation, Invasion, and Migration by Reprogramming Cell Metabolism and Extracellular Matrix Dynamics. Front. Genet. 12:778742. doi: 10.3389/fgene.2021.778742

Conclusion: AC093797.1 may inhibit cell proliferation, invasion, and migration in HCC by reprograming cell metabolism or regulating several pathways, suggesting that AC093797.1 might be a potential therapeutic and prognostic marker for HCC patients.

Keywords: biomarker, HCC, AC093797.1, biological function, long-non-coding RNA

INTRODUCTION

Hepatocellular carcinoma (HCC) causes a considerable number of deaths around the world every year. This has been a challenge in recent years, and it will affect approximately one million people each year by 2025 (Villanueva, 2019; Llovet et al., 2021; Sung et al., 2021). China is one of the countries with the highest incidence and mortality of liver cancer, especially HCC (Chen et al., 2016; Feng et al., 2019). Although many studies have clarified the risk factors of HCC and improved the diagnosis and treatment methods, the 5-year survival rate is still very poor (Greten et al., 2015). The mechanism of tumorigenesis is not yet fully understood, which is a huge obstacle to the treatment of tumors.

Long non-coding RNA (lncRNA) is comprised of a type of non-coding RNA with a length of more than 200 nucleotides, and it plays an important regulatory role in many cellular processes such as epigenetics, cell cycle, and cell differentiation. With deepening research, increasing evidence has confirmed that lncRNAs play a vital role in the biological process of various cancers, including HCC (Zhang et al., 2015; Yang et al., 2020).

In the previous study, we found that LncRNA AC093797.1 (Ensemble ID: ENSG00000233110, also known as: RP11-301L8.2) combined with three lncRNA (POLR2J4, AL121748.1, and AL162231.4) can be used to predict the prognosis of HCC patients by mining the GEO database and analyzing the gene microarray of HCC tissues (Ma and Deng, 2019). However, the biological role of AC093797.1 in liver cancer or any other malignant tumor has not been documented. In this study, we aimed to investigate the biological function of AC093797.1 and predict the key genes or signal pathways on which it is dependent. Finally, we found that the expression of AC093797.1 was decreased in HCC tissues and HCC cells, and patients with higher expression of AC093797.1 tend to have better 5-year survival. Combined with age, BMI, TNM stage, grade, and family-related history, AC093797.1 has good predictive power for the 5-year survival of patients with HCC. In addition, overexpression of AC093797.1 can significantly inhibit the proliferation, invasion, and migration of HCC cells *in vitro* and inhibit tumor growth *in vivo*. AC093797.1 can be involved in the disease process of HCC through a variety of pathways, and that might be a potential diagnostic and therapeutic target.

MATERIAL AND METHODS

Specimens and Data Collection

The tumor tissues and paired paracancerous tissues used in this study were obtained from 16 HCC patients who underwent hepatectomy in the department of Hepatobiliary Surgery of the Affiliated Hospital of Southwest Medical University. The specimens were frozen in liquid nitrogen immediately after removal. The histopathological diagnosis of HCC specimens was diagnosed by two pathologists independently, and both were diagnosed as HCC. This study was approved by the Institutional Review Board of the Affiliated Hospital of Southwest Medical University (Luzhou, China). Further, the RNA-seq data and clinical data with survival information were downloaded from UCSC Xena (https://xenabrowser.net/datapages/) to evaluate the predictive power of AC093797.1 and analyze the correlation with clinical indicators.

Cell Culture and Transfection

HCC cell lines, including HCCLM3, MHCC97-H, Huh7, and HepG2, and human normal liver cells LO2 were purchased from Procell Life Science and Technology Co., Ltd. (Wuhan, China). The overexpression vector of AC093797.1 (OE-AC093797.1) and empty vector (pcDNA3.1) were purchased from Shanghai Generay Biotech Co, Ltd. (Shanghai, China). HCCLM3, MHCC97-H, and Huh7 cells were cultured in DMEM medium (Gibco), LO2 cells were cultured in RPMI 1640 (Gibco), and HepG2 cells were cultured in MEM medium (Gibco). Above media were supplemented with 10% fetal bovine serum (FBS). All the cells were incubated in an incubator with 5% CO_2 at 37°C.

For the transfection, HCC cells were incubated in a 24-well polypropylene plate and transfected with the empty vector pcDNA3.1 or OE-AC093797.1 according to the instructions of Lipofectamine™ 2000 (Thermo, USA) when the cell confluence reaches about 70–90%. To obtain the stable transfected cell lines, the transfected HCC cells were screened by G418 (1 mg/ml) for 2 weeks.

RNA Isolation and qPCR

Total RNA of liver tissues or cells was isolated using Trizol reagent. RT first Strand cDNA Synthesis Kit was used to reverse transcribe 2 μg of total RNA into cDNA, and SYBR Green qPCR Master Mix (High ROX) was used for qRT-PCR. All the above reagents were purchased from Servicebio Biotechnology Co., Ltd. (Wuhan, China). The housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as the reference gene, and the relative expression of AC093797.1 was calculated *via* the $2^{-\triangle\triangle Ct}$ method. The primers used in this study were as follows: GAPDH forward primer: 5′-GGACCTGACCTGCCGTCTAG-3′; GAPDH reverse primer: 5′-GTAGCCCAGGATGCCCTTGA-3′; AC093797.1 forward primer: 5′-TGCCGCAAGGAGGAGGCTATTGTT-3′; AC093797.1 reverse primer: 5′-TGGGAAGGCTTATTCATG GACCTA-3′.

CCK-8 Assay

Cell-counting Kit-8 (CCK8) reagent was used to determine cell proliferation. Briefly, post-transfection for 24 h, the HCCLM3 cells were seeded into 96-well plates at 3×10^3 cells/well. At 0, 24, 48, 72, and 96 h, 10 μl CCK8 reagent was added to the wells and incubated at 37°C for an extra 4 h, then to detect the absorbance value at 450 nm by a microplate reader.

Wound-Healing Assay

After 24 h of transfection, the transfected cells were plated into a 24-well plate (2.5×10^5 cells/well). A sterile 200 μl plastic pipette tip was used to make scratches, the image of scratches at 0 and 48 h was captured by an inverted microscope at 40× magnification, and Image J was used to calculate the healing area of scratches.

Transwell Assays

The 24-well transwell chambers with 8 μm pore size were employed to assess the motility potential of HCC cells. For the invasion experiment, the bottom membrane of the chamber was pre-coated with 50 μl of Matrigel (3 mg/ml). After 24 h of transfection, the transfected cells were harvested and suspended in the serum-free medium at a density of 4×10^5 cells/ml, then 100 μl of cell suspension were dispensed into the upper chamber. Complete medium (600 μl) was added to the lower chamber. After 24 h, adherent cells on the membrane were fixed with paraformaldehyde fixative and stained with 0.1% crystal violet. A cotton swab was used to remove non-migrated cells, the images of 10 random fields at 200× magnification were captured by a microscope, and the number of cells was counted.

Nude Mice Tumor Formation Experiment

The 5-week-old female BALB/c nude mice weighing 15–20 g were maintained under specific-pathogen-free conditions and randomly divided into control group and OE-AC093797.1 group (n = 5 per group). In the nude mice tumor formation experiment, stably transfected HCCLM3 cells (5×10^7 cells in 100 μl) were subcutaneously injected into each nude mice. The mice with tumors were observed, and the volume of the tumor was calculated by the following formula: $V_{tumor} = length \times width^2 \times \pi/6$. When the tumor grows to about 2000 mm^3, the mice were sacrificed after anesthesia and tumor tissues were isolated. All procedures for the nude mice experiment were approved by the Animal Care Committee of Southwest Medical University (Luzhou, China).

Immunohistochemical Assay

Part of the tumor tissue isolated from the nude mice was fixed with 4% paraformaldehyde and used to detect the expression of Ki67 in the tumor tissue according to the instructions of the Ki-67 detection kit (Immunohistochemistry, Sangon Biotech Co., Ltd., Shanghai, China).

RNA Transcriptome Sequencing and Bioinformatics Analysis

The total RNA was extracted from subcutaneous tumor tissue of nude mice in the OE-AC093797.1 and control group and used to prepare a common transcriptome library; then, the RNA sequencing was performed on the Illumina NovaSeq 6000 platform. The "DNSeq2" package in R 4.0 was used to filter the differentially expressed genes between two groups. $|\log2foldchange| > 1.5$ and adjust p-value (FDR) < 0.05 were set as the criterion. Then, gene ontology (GO), KEGG signaling pathway enrichment analysis, and gene set enrichment analysis (GSEA, including c2.cp.kegg v7.4.symbols.gmt and c2.cp.reactome.v7.4.symbols.gmt) were performed. Adjusted p-value (FDR) < 0.05 indicates that the gene or the signaling pathway was significantly enriched in the corresponding category. Furthermore, the protein-protein interaction (PPI) network was constructed by the online tools STRING (https://string-db.org/) and visualized by Cytoscape 3.8. Then, the function modules and hub genes were identified by the plunge-in Molecular Complex Detection (MCODE) and cytoHubba using the default parameters and MCC algorithm, respectively.

Statistical Analysis

All data were expressed as mean ± standard deviation (SD), and the standard two-tailed t-test or one-way analysis of variance (ANOVA) was used to compare the difference between two groups. The correlation between the expression of AC093797.1 and clinical variables was assessed by chi-square test. Univariate and multivariate Cox regression models were used to screen the independent prognostic factor of HCC patients in the software SPSS 26. The subgroup analysis between the expression of AC093797.1 and clinical variables was performed *via* the package "forestplot" in R, and the survival analysis between the expression of AC093797.1 and survival time was performed *via* the package "survival" and "survminer" in R. $P < 0.05$ indicates the differences with statistical significance.

RESULTS

LncRNA AC093797.1 was Decreased in HCC and Associated With the Poor Survival of Patients

The GEPIA database integrates the data from the TCGA and the GTEx database, which recomputed the RNA sequencing raw data based on a standard processing pipeline to minimize differences from distinct sources to make data from different sources more compatible. The GEPIA database contains RNA expression profiles of 369 HCC tissues and 160 normal tissues. Based on the expression profiles of GEPIA, AC093797.1 was significantly low-expressed in HCC tissues than normal liver tissues ($p < 0.05$, **Figure 1A**). In this study, the expression of AC093797.1 was also confirmed in paired pathological tissues of HCC patients and liver cancer cell lines. As shown in **Figures 1B,C**, the expression of AC093797.1 in the 16 HCC tissues and four liver cancer cell lines (including HCCLM3, Huh7, HepG2, and MHCC97-H) was significantly decreased than adjacent non-tumor specimens and normal liver cells LO2. Further, the HCC patients with lower expression of AC093797.1 have poorer 5-year survival, but the expression level of the lncRNA was not related to the 5-year disease-

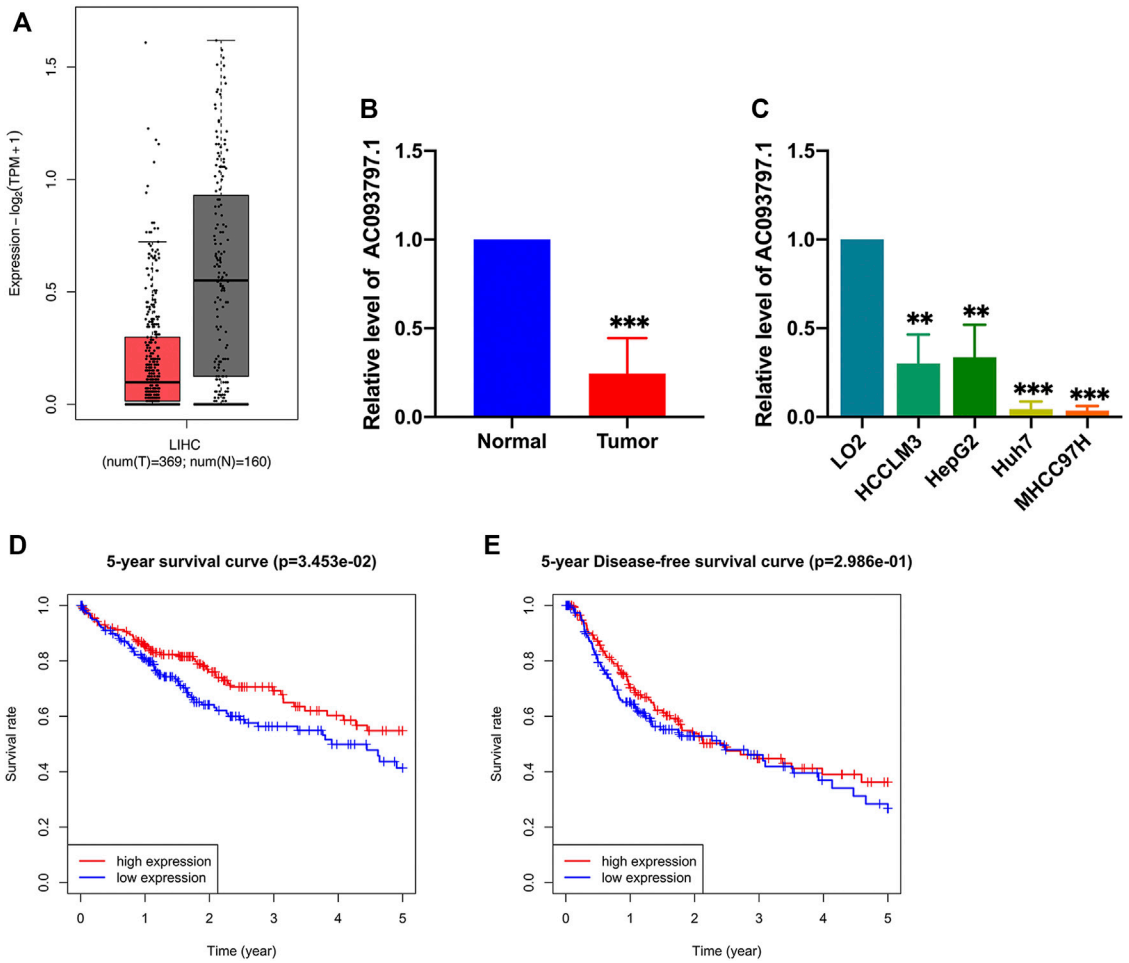


FIGURE 1 | The expression and prognosis value of AC093797.1. **(A)** The expression AC093797.1 in HCC tissues in GEPIA database (red: HCC tissues; grey: normal tissues). **(B)** The expression of AC093797.1 in the 16 HCC tissues collected in this study. **(C)** The expression of AC093797.1 in the HCC cell line HCCLM3, HepG2, Huh7, and MHCC97H. **(D–E)** The 5-year survival and DFS plot based on the expression of AC093797.1. $^{**}p < 0.01$, $^{***}p < 0.001$.

free survival (DFS) (**Figures 1D,E**), indicating that the expression of AC093797.1 may be a biomarker of 5-year survival of HCC patients.

The Relationship Between AC093709.1 and Clinical Indicators

To explore the correlation between the expression of AC093709.1 and clinical indicators, 365 HCC patients with survival information in the TCGA database were divided into low- and high-expression groups according to the median of AC093797.1. Combined with the 5-year overall survival information, the BMI (>23.9 vs. ≤23.9) and living status of HCC patients were significantly different between the two groups (**Table 1**), suggesting that the expression level of AC093797.1 may be related to the BMI and 5-year OS of HCC patients. While combined with the 5-year DFS information, the BMI (>23.9 vs. ≤23.9) and hepatic inflammation of HCC patients was significantly different between the two groups (**Table 1**).

Further, univariate Cox regression and multivariate Cox regression analyses were performed to identify the risk factors that affect the prognosis of HCC patients. As shown in **Figure 2A**, the BMI ≤23.9, TNM stage III-IV, with tumor, and the low expression of AC093797.1 were the risk factors of 5-year survival. The low expression of AC093797.1 was not acting as a risk factor of 5-year DFS, and the risk factors of 5-year DFS of HCC patients were the TNM stage III-IV, with tumor and vascular invasion (**Supplementary Figures S1A, B**). However, the expression level of AC093797.1 was not an independent risk factor that affects the 5-year OS (**Figure 2B**), implying that the expression of AC093797.1 may interact with other clinical indicators to affect the 5-year survival of HCC patients.

The results of subgroup analysis showed that the expression of AC093797.1 was an independent risk factor affecting 5-year survival in patients with age ≤60, BMI ≤23.9, related family history, TNM stage III-IV, or grade III-IV (**Figure 2C**). But the expression of AC093797.1 was not an independent risk factor affecting 5-year DFS in any subgroup (**Supplementary Figure S2C**).

Besides, the expression of AC093797.1 combined with age, BMI, TNM stage, grade, and related family history can better

TABLE 1 | The clinical characters of HCC patients in the AC093797.1 low-/high-expression group.

Clinical features		OS			DFS		
		AC093797.1 expression		*p* value	AC093797.1 expression		*p* value
		High	Low		High	Low	
Age	≤60	81	92	0.161	69	83	0.230
	>60	104	88		84	77	
Gender	Female	66	53	0.119	97	111	0.263
	Male	115	113		56	49	
BMI	≤23.9	66	86	0.045	58	79	0.038
	>23.9	98	82		86	72	
TNM stage	I-II	128	126	0.279	111	116	0.632
	III-IV	38	49		31	37	
Grade	G1/G2	115	115	0.575	95	97	0.570
	G3/G4	61	69		54	63	
Tumor status	Tumor free	80	81	0.550	71	73	0.570
	With tumor	67	55		53	47	
Family history	None	94	110	0.115	80	99	0.080
	Yes	62	45		53	42	
Vascular invasion	None	106	99	0.360	88	87	0.533
	Yes	49	57		44	51	
AFP	≤300 ng/ml	110	102	0.357	112	107	0.291
	>300 ng/ml	29	35		27	35	
Hepatic inflammation	None	70	47	0.067	65	44	0.049
	Yes	55	60		56	48	
Living status	Alive	125	109	0.042	108	100	0.109
	Dead	55	75		44	60	

predict the 5-year survival in patients with age ≤ 60, BMI ≤ 23.9, related family history, TNM stage III-IV, or grade III-IV (**Figures 2D–M**).

It was speculated that the expression of AC093797.1 may have a significant clinical correlation with clinical indicators in specific HCC patients.

AC093797.1 Inhibited Cell Proliferation, Migration, and Invasion of the HCC Cells

To further clarify the biological role of AC093797.1 in HCC, the overexpression vector OE-AC093797.1 was transfected into HCCLM3 cells. The expression of AC093797.1 in the transfected HCCLM3 was detected by qPCR. As shown in **Figure 3A**, the expression of AC093797.1 was significantly increased in the cells transfected with OE-AC093797.1. Through CCK8, transwell, and wound healing assays, it was found that the overexpressed AC093797.1 in the HCCLM3 cells could inhibit cell proliferation (**Figure 3B**), migration (**Figures 3C,D, G–H**), and invasion (**Figures 3E,F**).

AC093797.1 Overexpression Inhibits the Tumor Growth *in vivo*

Via the nude mice tumor formation experiment, it was found that the volume of the tumor isolated from the OE-AC093797.1 group was significantly smaller than the control group on the 45th day after subcutaneous injection (**Figures 4A–D**), which suggested that high expression of AC093797.1 can inhibit the growth of tumors formed by HCCLM3 *in vivo*. Ki67 protein is a nuclear protein strictly associated with cell proliferation, and it is generally believed that the higher the expression level of Ki67, the faster the tumor growth, the lower the degree of differentiation, and the worse the prognosis of tumor patients. The immunohistochemical results of ki67 showed that the positive staining area in the tumor tissue of the OE-AC097997.1 group was significantly lower compared with the control group (**Figures 4E–G**), indicating that AC093797.1 significantly weakened the cell division of HCC cells *in vivo*.

Bioinformatics Analysis Based on Subcutaneous Tumor Specimens

In this study, subcutaneous tumor tissues of three nude mice in the OE-AC093797.1 and control group were used to extract the total RNA and performed RNA sequencing independently. Then, the differentially expressed genes of the OE-AC097997.1 group and control group were filtered by the package "DESeq2" in R. The result showed that 710 differentially expressed genes were identified, including 243 upregulated genes and 467 downregulated genes (**Figures 5A,B**, **Supplementary Table S1**).

To explore the involved biological processes and functions of these differentially expressed genes, the GO, KEGG signal pathway enrichment analyses, and GSEA were conducted. The GO analysis contained three terms, including biological process (BP), cellular component (CC), and Molecular Function (MF). For the BP, the differentially expressed genes were significantly enriched in the extracellular matrix organization, extracellular structure organization, etc.; for the CC, the differentially expressed genes were significantly enriched in the collagen-

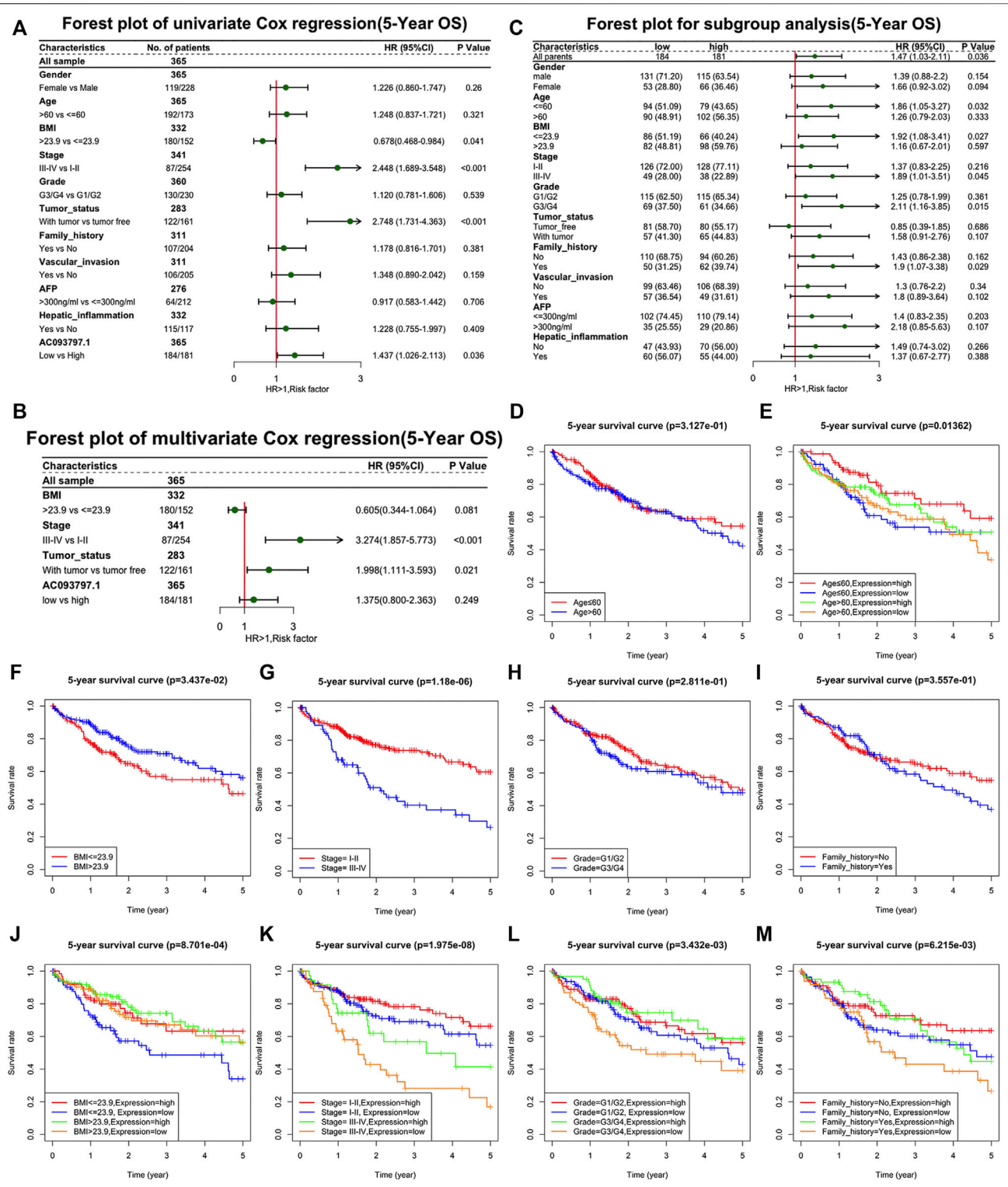


FIGURE 2 | The relationship between AC093709.1 and clinical indicators. **(A)** Forest plot of univariate Cox regression (5-year overall survival, OS). **(B)** Forest plot of multivariate Cox regression (5-year OS). **(C)** Forest plot for subgroup analysis (5-year OS). **(D–M)** The survival plot of HCC patients is based on age, BMI, TNM stage, grade, family-related history, and the expression of AC093797.1.

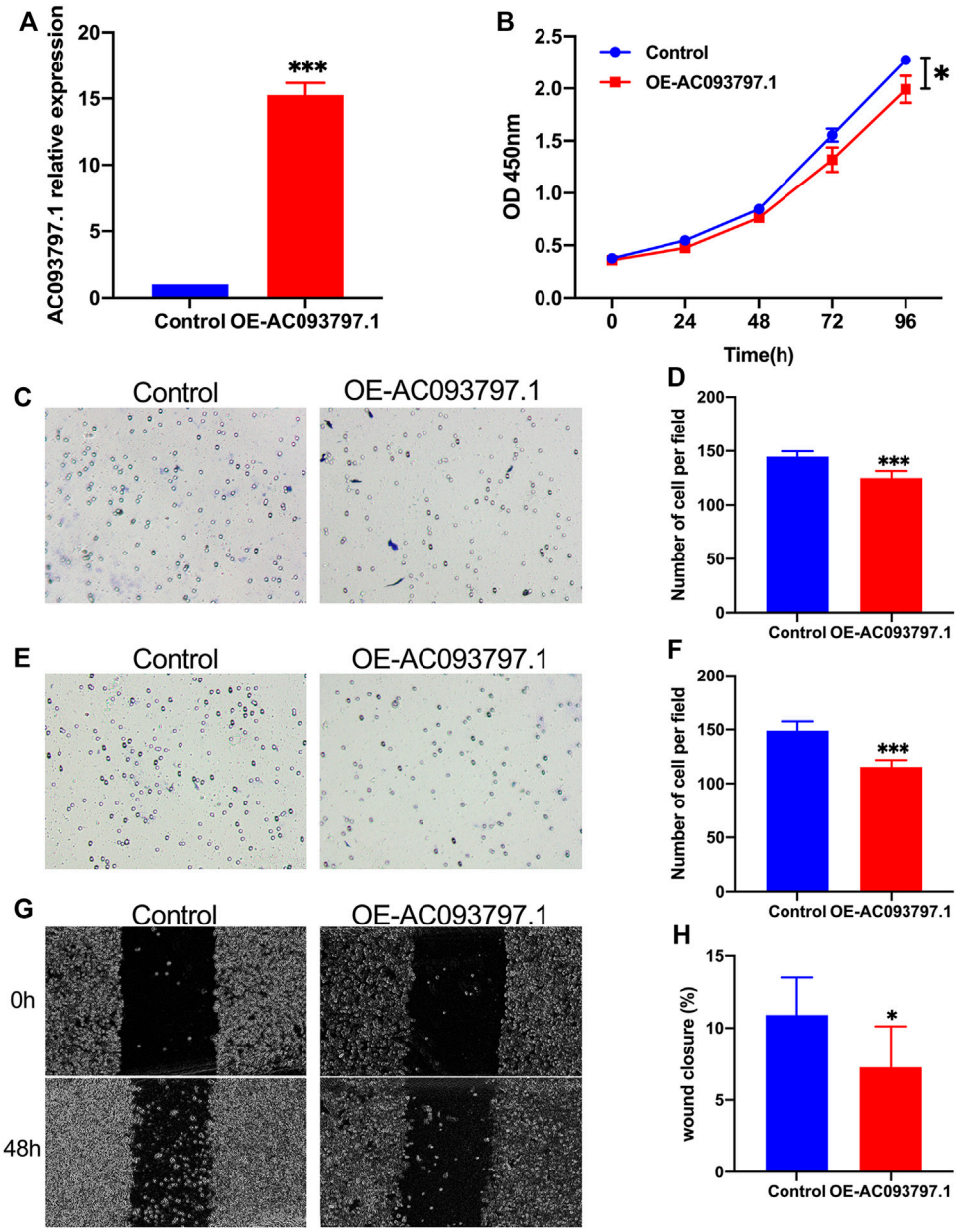


FIGURE 3 | AC093797.1 inhibited cell proliferation, migration, and invasion of the HCC cells. **(A)** The expression of AC093797.1 in the HCC cells HCCLM3 transfected with OE-AC093797.1 or empty vector. **(B)** CCK8 assay detect the proliferation of HCCLM3 transfected with OE-AC093797.1 or empty vector. **(C, D)** The invasion assay of HCCLM3 transfected with OE-AC093797.1 or empty vector. **(E, F)** The migration assay of HCCLM3 transfected with OE-AC093797.1 or empty vector. **(G–H)** The wound-healing assay of HCCLM3 transfected with OE-AC093797.1 or empty vector. *$p < 0.05$, ***$p < 0.001$.

containing extracellular matrix, endoplasmic reticulum lumen, nucleosome, etc.; for the MF, the differentially expressed genes were significantly enriched in peptidase regulator activity, extracellular matrix structure constituent, etc. (**Figure 5C**). *Via* the KEGG enrichment analysis, the differentially expressed genes were significantly enriched in retinal metabolism, alcoholism, drug metabolism-cytochrome P450, steroid hormone biosynthesis, metabolism of xenobiotics by cytochrome P450, etc. (**Figure 5B**). The result of GSEA showed that the KEGG pathway gene set of the metabolism of glycolipid metabolism, tyrosine metabolism, starch and sucrose metabolism, arginine and proline metabolism, valine leucine and isoleucine degradation, cell adhesion molecules cams, pentose and glucuronate interaction, amyotrophic lateral sclerosis (**Figure 6A**), and the Reactome pathway gene set of the NLRP3 inflammasome, signaling to Ras, and signaling to ERKs, and other three pathway was significantly enriched in the control group (**Figure 6B**).

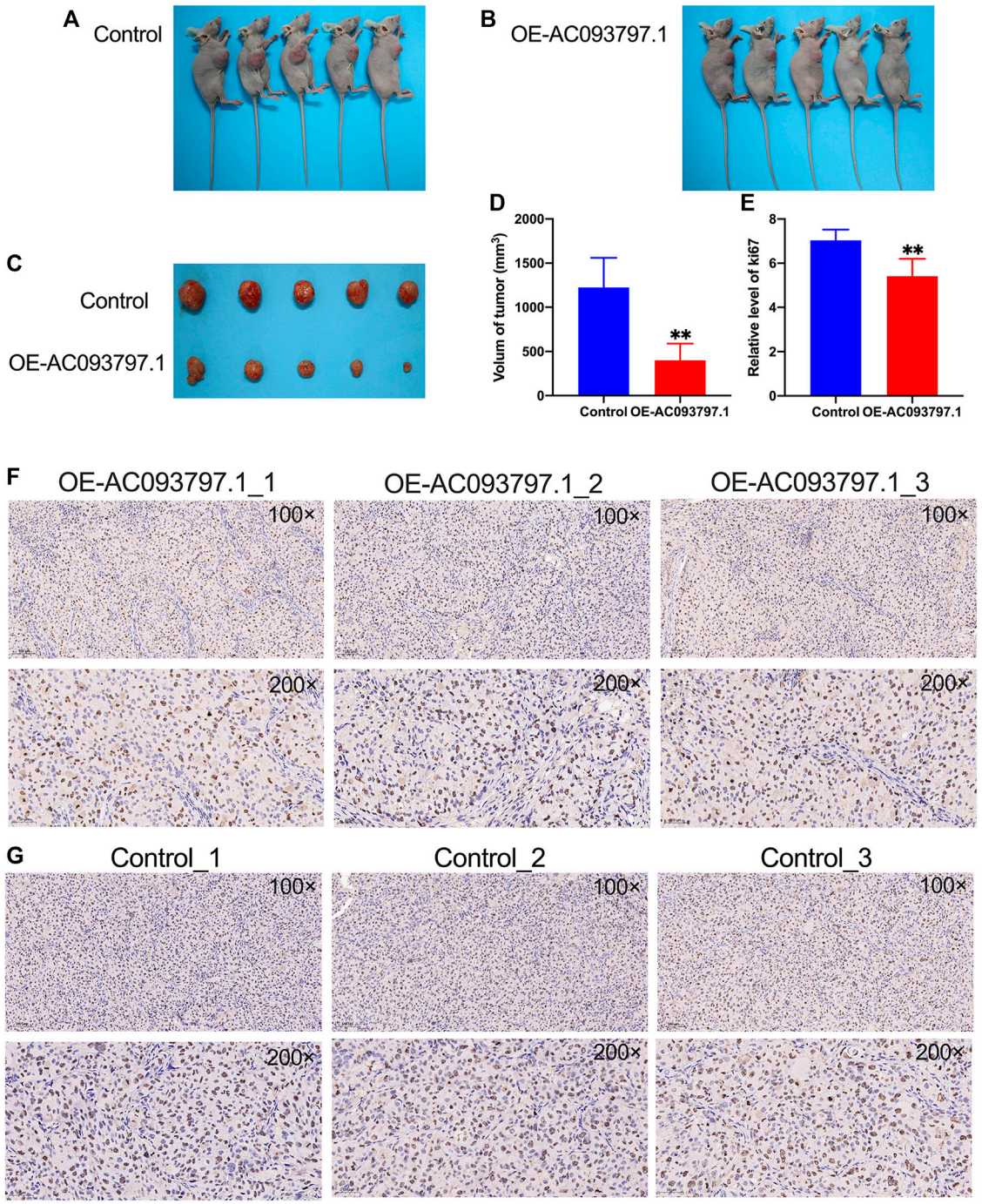


FIGURE 4 | AC093797.1 can inhibit the growth of tumors in nude mice. **(A, B)** The nude mice with tumors in the OE-AC093797.1 group and control group. **(C)** The tumors isolated from the mice in the OE-AC093797.1 group and control group. **(D)** The volume of tumors isolated from the mice in the OE-AC093797.1 group and control group. **(E)** The percentage of positive straining area of Ki67 in the OE-AC093797.1 and control group. **(F, G)** The expression of Ki67 in the tumors isolated from OE-AC093797.1 and control group detected by immunohistochemical assay. $^{**}p < 0.01$.

Furthermore, a PPI network was constructed in the STRING database and visualized by the Cytoscape. By the MCODE plug-in in Cytoscape, three function modules of the PPI network were parsed out (**Figures 5E–G**), and the top 20 core genes were extracted from the PPI network *via* the plug-in CytoHubba using the MCC algorithms (**Figure 5H**). Combining the above two algorithms, 13 histone family members HIST1H2BJ, HIST1H2BK, HIST1H2AC, HIST2H2AA3, HIST1H2AI, HIST1H2BG, HIST2H4B, HIST2H4A, HIST1H2BC, HIST1H1C, HIST1H3H, HIST2H3D, HIST2H2BF, and TTR, A2M, PTPRC, and FGG were selected as hub genes finally.

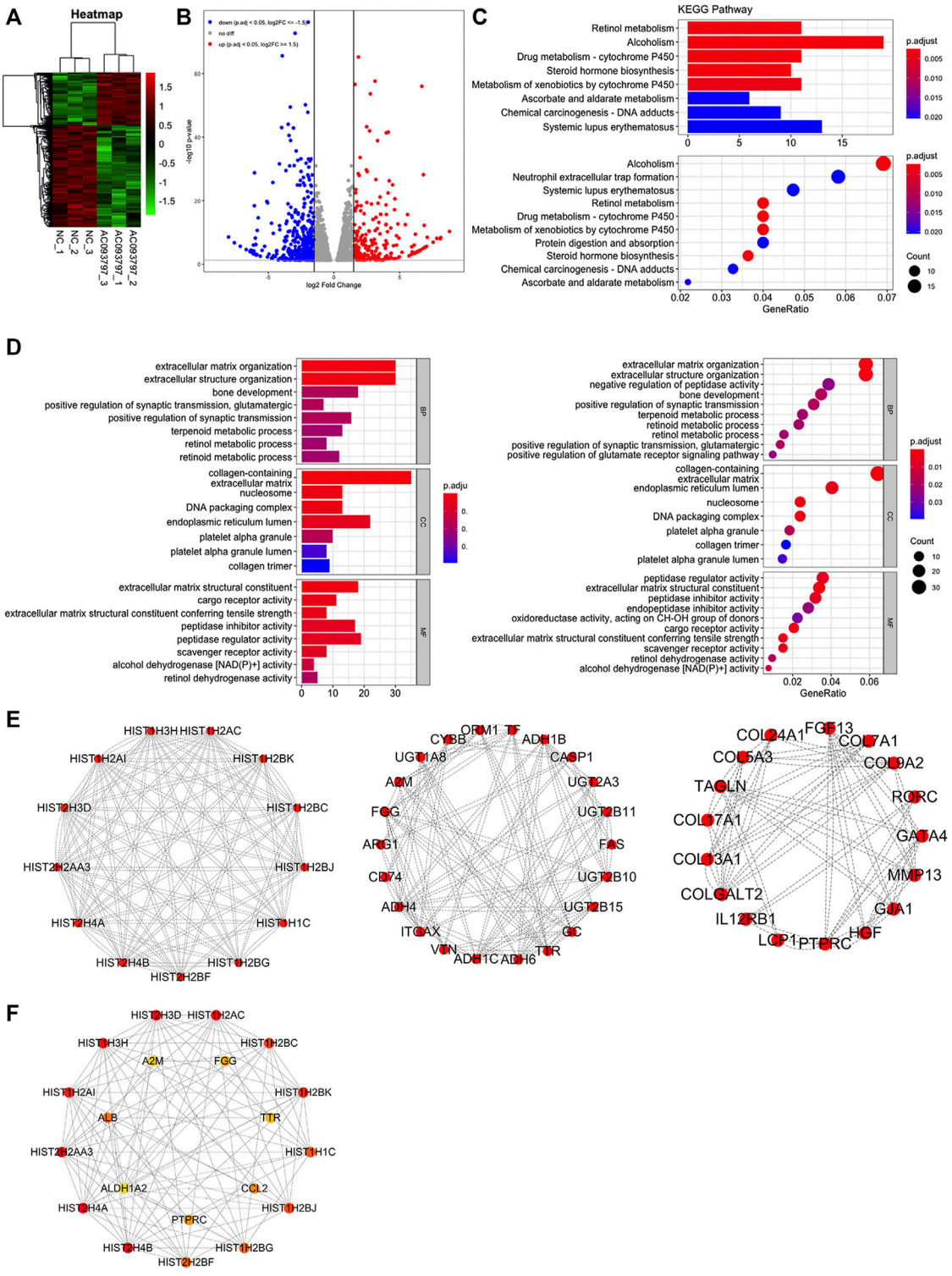


FIGURE 5 | Bioinformatics analysis based on subcutaneous tumor specimens. **(A, B)** The heatmap and volcano plot of the differentially expressed genes between OE-AC093797.1 group and control group. **(C)** The bar and dot figures of KEGG of the differentially expressed genes. **(D)** The bar and dot figures of GO of the differentially expressed genes. **(E)** Three function module of protein-protein interaction (PPI) network. **(F)** The top20 core genes of PPI network.

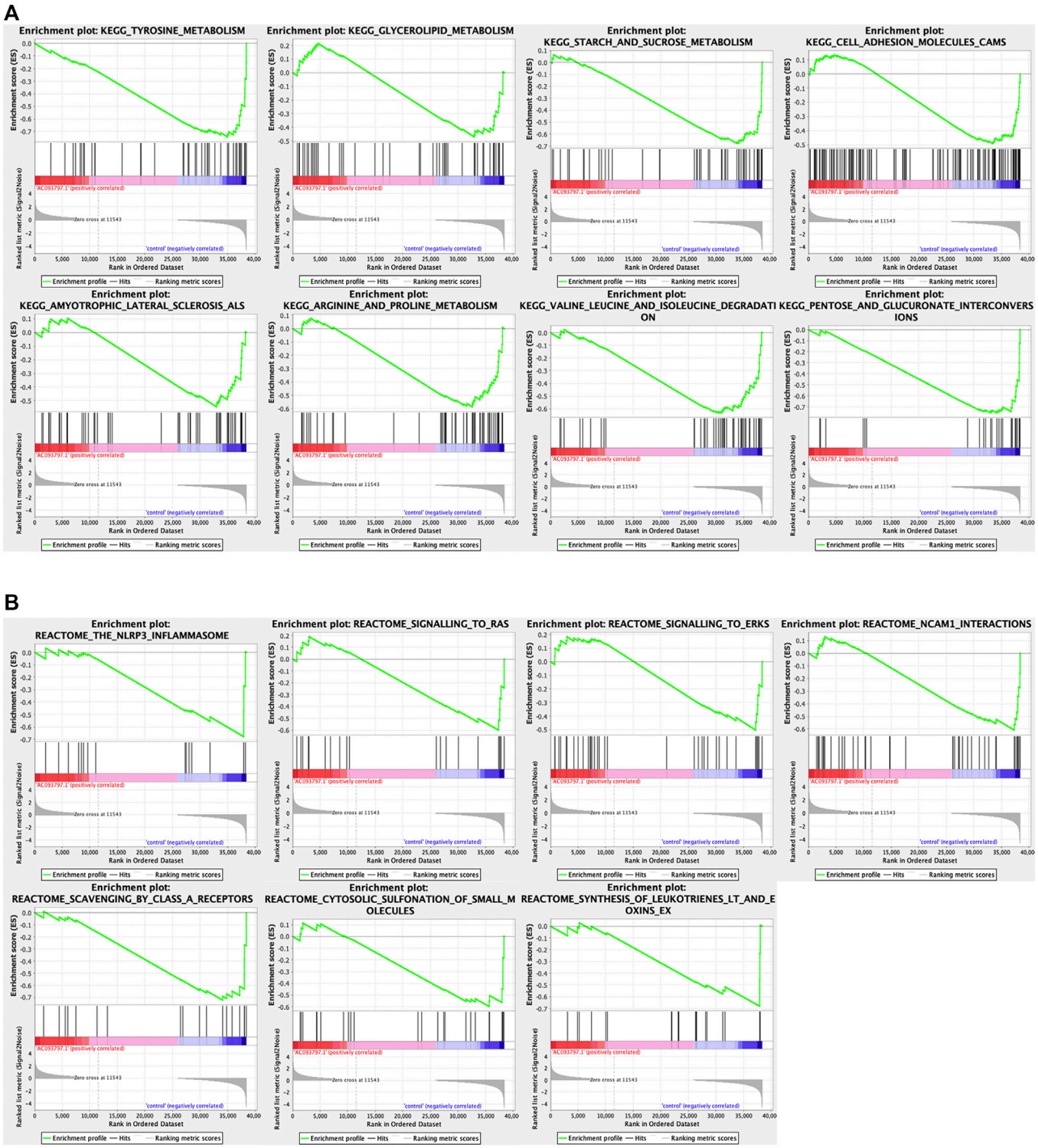


FIGURE 6 | The GSEA analysis between the OE-AC093797.1 and control group. **(A)** GSEA analysis based on the "c2.cp.kegg v7.4.symbols.gmt". **(B)** GSEA analysis based on the "c2.cp.reactome.v7.4.symbols.gmt".

DISCUSSION

According to the newest report from World Health Organization (WHO), among all malignant tumors, the incidence rate of liver cancer patients ranks sixth and the mortality rate ranks third (Valery et al., 2018; Sung et al., 2021). The stages of the tumor determine the prognosis of liver cancer generally, and the 5-year survival of patients with early-stage liver cancer can be more than 70%. In contrast, patients with advanced symptoms who need systemic therapy only had a median survival of 1–1.5 years

(Villanueva, 2019). At present, the main treatment method for liver cancer is surgery, and this treatment also makes the early-stage patients obtain the maximum benefit. Consistent with clinical observation, we found that the TNM stage was one of the independent risk factors that affect the 5-year OS and DFS of HCC patients, and patients at stage III-IV have a worse prognosis.

Nonetheless, other clinical indicators, including age, gender, grade, related family history, etc., are also factors that cannot be ignored in clinical practice that affect the prognosis of HCC patients. However, those indicators were not risk factors to the 5-year OS and DFS based on the survival information of HCC patients in the TCGA database. The lncRNA AC093797.1 was downregulated in HCC, and the HCC patients with lower expression of the lncRNA have a poorer prognosis. *Via* subgroup analysis, we found that the expression of AC093797.1 was an independent risk factor affecting 5-year survival in patients with age ≤ 60, BMI ≤ 23.9, related family history, TNM stage III-IV, or grade III-IV, and AC093797.1 combined with age, BMI, TNM stage, grade, and related family history can better predict the 5-year survival in these specific patients, indicating that AC093797.1 may have a close correlation with these clinical indicators and more suitable as a prognostic marker in those specific HCC patients.

Until now, the biological function of AC093797.1 has not been reported in liver cancer or any other malignant tumors. To explore the role of AC093797.1 in the HCC, we transformed the overexpression vector OE-AC093797.1 into HCCLM3, a liver cancer cell line with high spontaneous lung metastasis. In the CCK8, transwell, and wound healing assays, overexpressed AC093997.1 showed inhibitory effects on cell proliferation, invasion, and migration. This inhibitive tumorigenicity role of AC093797.1 was also confirmed by the nude mice tumor formation experiment, the result showed that the volume of tumors in the OE-AC093797.1 group was smaller than the control group, and the decreased Ki67 also revealed the weakened cell proliferation of tumors in the OE-AC093797.1 group. However, no lung metastases were seen in all the mice examined.

Considering that the research on AC093797.1 is very limited, we performed RNA sequencing on tumor tissues isolated from nude mice to gain a broader understanding of the regulatory role of AC093797.1 in HCC. Finally, we obtained 710 differentially expressed genes between the cells transfected with OE-AC093797.1 or empty vector, including 243 upregulated and 467 downregulated genes. *Via* the GO, KEGG, and GSEA enrichment analysis, we found that those differential genes were mainly significant enriched in the following KEGG signaling pathway: the retinal metabolism, alcoholism, drug metabolism-cytochrome P450, steroid hormone biosynthesis, metabolism of xenobiotics by cytochrome P450, etc., and mainly involved in the extracellular matrix organization, extracellular structure organization, and other biology process or molecular function. Extracellular matrix (EMC) is the major component of the local microenvironment or niche of cancer. The importance of ECM to cancer progression is now well recognized (Poltavets et al., 2018). The abnormal ECM dynamics are a hallmark of cancer (Lu et al., 2012) and lead to the development of cancer (Walker et al., 2018). The upregulated matrix metallopeptidase, including MMP-1, −3, −7, −10, −11, −13, −14, −16, −26, and −28, are favoring the invasion and metastasis (Scheau et al., 2019), while MMP1, MMP13, and MMP28 were significantly downregulated in the AC093797.1 overexpressed HCC cell HCCLM3, suggesting that this may be one of the reasons AC093797.1 inhibited the migration and invasion of HCC cells.

It was worth noting that the GSEA enrichment analysis using the expression matrix found that KEGG pathway gene set of the metabolism of glycolipid, tyrosine, starch and sucrose, arginine and proline, valine leucine and isoleucine degradation, cell adhesion molecules cams, pentose and glucuronate interaction, and amyotrophic lateral sclerosis were significantly enriched in the control group.

The rapid growth of cancer cells usually runs out of glucose and must use other raw materials as substitution, such as fats and amino acids (Nguyen et al., 2020). Tyrosine is one of the preferred raw materials for cancer, but it is rarely used by normal healthy cells (Nguyen et al., 2020). The catabolism of all three essential amino acids valine, isoleucine, and leucine can also yield NADH and FADH2 which can be utilized for ATP generation. The process of arginine metabolism can generate two derivatives (glutamine and proline) with different functions in cancer development (Zou et al., 2019). The rapid tumor growth has a high demand for glutamine, although it is a non-essential amino acid (Zou et al., 2019). Cancer cells prefer to rely on aerobic glycolysis rather than the oxidation of pyruvate to meet the rapid proliferation of the high energy demand (Zhou et al., 2021). Zhou et al. found that the oroxyloside (OAG)-induced glycolipid metabolic switch could increase ROS levels and lead to G1 cell cycle arrest and growth inhibition of HCC cells (Zhou et al., 2021). Pentose and glucuronate interconversions were deregulated in 100% cancer (Becker et al., 2012). The deregulated pentose and glucuronate interconversions were not enriched in the HCC cells transfected with OE-AC093797.1. It seems that AC073797.1 might alleviate the effect of this pathway on HCC cells. The above results reveal that cells transfected with OE-AC093797.1 seem to have less energy and biomaterial supply than the cells in the control group or reprogramed the metabolic process, and result in weaker proliferation ability of the cells transfected with OE-AC093797.1.

In addition, the Reactome pathway gene set of the NLRP3 inflammasome, signaling to Ras, and signaling to ERKs and other three pathways was shown. It has been reported that NLRP3 inflammasome has different effects in different malignancies and is considered a double-edged sword against cancer (Ju et al., 2021). But the role of NLRP3 inflammasome activation in HCC remains unclear. The Ras/MAPK pathway is activated in 50–100% of human HCC and is associated with poor prognosis (Delire and Stärkel, 2015). Ras is the first intracellular effector of the ERK1/2 pathway. Various extracellular stimuli can trigger the transformation of RAS from an inactive form to an active form and then through multiple effectors to activate the pathways implicated in cell growth, survival, differentiation, and migration (Delire and Stärkel, 2015). Ras/MAPK is usually activated in more than half of HCC patients and is associated with a poor prognosis

(Delire and Stärkel, 2015). Ras/MAPK pathway effectors are considered potential targets for the treatment of HCC.

Besides, we also analyzed the interaction between the differentially expressed genes *via* Cytoscape and identified 13 histones and TTR, A2M, PTPRC, and FGG as the hub genes. Although these histones and other hub genes are mostly used as prognostic markers of tumor patients, a few studies have shown that histones may be related to the malignancy of tumor cells. In low-grade glioma (LGG), the upregulated HIST1H2BK was an indicator of poor prognosis and may be a promising biomarker for the treatment of LGG (Liu et al., 2020). HIST1H2AI might involve in nucleosome assembly and DNA packaging (Han et al., 2014). Upregulated HIST1HABF can enhance the cancer stem cell phenotype, malignancy, and liver metastasis through the activation of Notch signaling in colorectal carcinoma (Qiu et al., 2021). In adrenocortical carcinoma, the hub gene HIST1H1C was associated with poor overall survival (Alshabi et al., 2019). For other hub genes, transthyretin (TTR) has the ability to stimulate tumor growth through regulation of tumor, immune, and endothelial cells (Lee et al., 2019); FGG is capable of promoting migration and invasion in hepatocellular carcinoma cells through activating epithelial to mesenchymal transition (EMT) (Zhang et al., 2019); and PTPRC can interact with CXCR4 in a putative molecular network constructed from microarray data, which is closely related to colon cancer metastasis (Chu et al., 2017).

CONCLUSION

In this study, we found that AC093797.1 was downregulated in the HCC tissues and four HCC cell lines, and low expression of AC093797.1 in HCC patients was associated with a poor prognosis. AC093797.1 overexpression may inhibit tumor growth in nude mice and inhibits cell proliferation, invasion, and migration *in vitro*. The differentially expressed genes identified by RNA-sequencing are mostly involved in the cell division or metastatic pathways. In summary, our research suggests that AC093797.1 may be a promising diagnostic and therapeutic target for hepatocellular cancer.

DATA AVAILABILITY STATEMENT

The RNA-sequencing data presented in the study were deposited in the Gene Expression Omnibus (GEO) repository, accession number GSE186933 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE186933).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Independent Ethics Committee of the affiliated Hospital of Southwest Medical university. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by the Animal Care Committee of the Southwest Medical University.

AUTHOR CONTRIBUTIONS

CD and CS contributed to study design and supervision of the manuscript for important intellectual content. XL and CW performed data analysis and drafted the manuscript. QY, DL, and YY contributed to data collection. SO and YS contributed to revision of the manuscript. 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/fgene.2021.778742/full#supplementary-material

REFERENCES

Alshabi, A. M., Vastrad, B., Shaikh, I. A., and Vastrad, C. (2019). Identification of Important Invasion and Proliferation Related Genes in Adrenocortical Carcinoma. *Med. Oncol.* 36, 73. doi:10.1007/s12032-019-1296-7

Becker, D., Sfakianakis, I., Krupp, M., Staib, F., Gerhold-Ay, A., Victor, A., et al. (2012). Genetic Signatures Shared in Embryonic Liver Development and Liver Cancer Define Prognostically Relevant Subgroups in HCC. *Mol. Cancer* 11, 55. doi:10.1186/1476-4598-11-55

Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., Bray, F., et al. (2016). Cancer Statistics in China, 2015. *CA: a Cancer J. clinicians* 66, 115–132. doi:10.3322/caac.21338

Chu, S., Wang, H., and Yu, M. (2017). A Putative Molecular Network Associated with colon Cancer Metastasis Constructed from Microarray Data. *World J. Surg. Onc* 15, 115. doi:10.1186/s12957-017-1181-9

Delire, B., and Stärkel, P. (2015). The Ras/MAPK Pathway and Hepatocarcinoma: Pathogenesis and Therapeutic Implications. *Eur. J. Clin. Invest.* 45, 609–623. doi:10.1111/eci.12441

Feng, R.-M., Zong, Y.-N., Cao, S.-M., and Xu, R.-H. (2019). Current Cancer Situation in China: Good or Bad News from the 2018 Global Cancer Statistics. *Cancer Commun.* 39, 22. doi:10.1186/s40880-019-0368-6

Greten, T. F., Wang, X. W., and Korangy, F. (2015). Current Concepts of Immune Based Treatments for Patients with HCC: from Basic Science to Novel Treatment Approaches. *Gut* 64, 842–848. doi:10.1136/gutjnl-2014-307990

Han, J. A., Kim, J.-Y., and Kim, J.-I. (2014). Analysis of Gene Expression in Cyclooxygenase-2-Overexpressed Human Osteosarcoma Cell Lines. *Genomics Inform.* 12, 247–253. doi:10.5808/gi.2014.12.4.247

Ju, M., Bi, J., Wei, Q., Jiang, L., Guan, Q., Zhang, M., et al. (2021). Pan-cancer Analysis of NLRP3 Inflammasome with Potential Implications in Prognosis and Immunotherapy in Human Cancer. *Brief Bioinform* 22, bbaa345. doi:10.1093/bib/bbaa345

Lee, C.-C., Ding, X., Zhao, T., Wu, L., Perkins, S., Du, H., et al. (2019). Transthyretin Stimulates Tumor Growth through Regulation of Tumor, Immune, and Endothelial Cells. *J. Immunol.* 202, 991–1002. doi:10.4049/jimmunol.1800736

Liu, W., Xu, Z., Zhou, J., Xing, S., Li, Z., Gao, X., et al. (2020). High Levels of HIST1H2BK in Low-Grade Glioma Predicts Poor Prognosis: A Study Using

CGGA and TCGA Data. *Front. Oncol.* 10, 627. doi:10.3389/fonc.2020.00627

Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular Carcinoma. *Nat. Rev. Dis. Primers* 7, 6. doi:10.1038/s41572-020-00240-3

Lu, P., Weaver, V. M., and Werb, Z. (2012). The Extracellular Matrix: a Dynamic Niche in Cancer Progression. *J. Cel Biol* 196, 395–406. doi:10.1083/jcb.201102147

Ma, L., and Deng, C. (2019). Identification of a Novel Four-lncRNA Signature as a Prognostic Indicator in Cirrhotic Hepatocellular Carcinoma. *PeerJ* 7, e7413. doi:10.7717/peerj.7413

Nguyen, T. N., Nguyen, H. Q., and Le, D.-H. (2020). Unveiling Prognostics Biomarkers of Tyrosine Metabolism Reprogramming in Liver Cancer by Cross-Platform Gene Expression Analyses. *PLoS One* 15, e0229276. doi:10.1371/journal.pone.0229276

Poltavets, V., Kochetkova, M., Pitson, S. M., and Samuel, M. S. (2018). The Role of the Extracellular Matrix and its Molecular and Cellular Regulators in Cancer Cell Plasticity. *Front. Oncol.* 8, 431. doi:10.3389/fonc.2018.00431

Qiu, L., Yang, X., Wu, J., Huang, C., Miao, Y., and Fu, Z. (2021). HIST2H2BF Potentiates the Propagation of Cancer Stem Cells via Notch Signaling to Promote Malignancy and Liver Metastasis in Colorectal Carcinoma. *Front. Oncol.* 11, 677646. doi:10.3389/fonc.2021.677646

Scheau, C., Badarau, I. A., Costache, R., Caruntu, C., Mihai, G. L., Didilescu, A. C., et al. (2019). The Role of Matrix Metalloproteinases in the Epithelial-Mesenchymal Transition of Hepatocellular Carcinoma. *Anal. Cel Pathol (Amst)* 2019, 9423907. doi:10.1155/2019/9423907

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Valery, P. C., Laversanne, M., Clark, P. J., Petrick, J. L., McGlynn, K. A., and Bray, F. (2018). Projections of Primary Liver Cancer to 2030 in 30 Countries Worldwide. *Hepatology* 67, 600–611. doi:10.1002/hep.29498

Villanueva, A. (2019). Hepatocellular Carcinoma. *N. Engl. J. Med.* 380, 1450–1462. doi:10.1056/nejmra1713263

Walker, C., Mojares, E., and Del Río Hernández, A. (2018). Role of Extracellular Matrix in Development and Cancer Progression. *Int. J. Mol. Sci.* 19, 3028. doi:10.3390/ijms19103028

Yang, J. J., Yang, Y., Zhang, C., Li, J., and Yang, Y. (2020). Epigenetic Silencing of LncRNA ANRIL Enhances Liver Fibrosis and HSC Activation through Activating AMPK Pathway. *J. Cel Mol Med* 24, 2677–2687. doi:10.1111/jcmm.14987

Zhang, X., Wang, F., Huang, Y., Ke, K., Zhao, B., Chen, L., et al. (2019). FGG Promotes Migration and Invasion in Hepatocellular Carcinoma Cells through Activating Epithelial to Mesenchymal Transition. *Cmar* 11, 1653–1665. doi:10.2147/cmar.s188248

Zhang, Y., Li, Z., Zhang, Y., Zhong, Q., Chen, Q., and Zhang, L. (2015). Molecular Mechanism of HEIH and HULC in the Proliferation and Invasion of Hepatoma Cells. *Int. J. Clin. Exp. Med.* 8, 12956–12962.

Zhou, Y., Guo, Y., Zhu, Y., Sun, Y., Li, W., Li, Z., et al. (2021). Dual PPARγ/□ Agonist Oroxyloside Suppresses Cell Cycle Progression by Glycolipid Metabolism Switch-Mediated Increase of Reactive Oxygen Species Levels. *Free Radic. Biol. Med.* 167, 205–217. doi:10.1016/j.freeradbiomed.2021.02.032

Zou, S., Wang, X., Liu, P., Ke, C., and Xu, S. (2019). Arginine Metabolism and Deprivation in Cancer Therapy. *Biomed. Pharmacother.* 118, 109210. doi:10.1016/j.biopha.2019.109210

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 22 December 2021
doi: 10.3389/fgene.2021.775006

Identification of Potential Hub Genes and miRNA-mRNA Pairs Related to the Progression and Prognosis of Cervical Cancer Through Integrated Bioinformatics Analysis

Mingxu Fu[1†], *Yongyan Pei*[2†], *Fang Lu*[1], *Huici Jiang*[1], *Yingying Bi*[1], *Jiajing Cheng*[1] *and Jinlong Qin*[1*]

[1]Department of Obstetrics and Gynecology, Shanghai Fourth People 's Hospital, School of Medicine, Tongji University, Shanghai, China, [2]School of Medicine and Chemical Engineering, Guangdong Pharmaceutical University, Guangzhou, China

In recent years, the incidence and mortality of cervical cancer have increased worldwide. At the same time, increasing data have confirmed that miRNA-mRNA plays a positive or negative regulatory role in many cancers. This study attempted to screen effective miRNA-mRNA in the progression of cervical cancer, and to study the mechanism of miRNA-mRNA in the progression of cervical cancer. The expression profile data of GSE7410, GSE 63514, GSE 86100 and TCGA-CESC were downloaded, and 34 overlapping differentially expressed genes (22 up-regulated and 12 down-regulated) and 166 miRNAs (74 down-regulated and 92 up-regulated) were screened through limma package. Then, miR-197-3p/TYMS pairs were obtained by PPI, functional enrichment, Kaplan-Meier plotter analysis, Cox univariate and multivariate analysis, risk modeling, WGCNA, qPCR and dual-luciferase experiments. The results showed that TYMS was an independent prognostic factor of cervical cancer, and its expression level was negatively correlated with cervical cancer tissue grade (TMN), tumor grade, age, microsatellite stability and tumor mutation load, and positively correlated with methyl expression in DNMT1, DNMT2, DNMT3A and DNMT3B. Functional experiments showed that TYMS knockout could promote the proliferation, migration and invasion of HeLa cells and reduce apoptosis. Overexpression of TYMS showed the opposite trend, miR-197-3p was negatively correlated with the expression of TYMS. MiR-197-3p inhibitor reversed the effect of si-TYMS on the proliferation of HeLa cells. In conclusion, these results reveal that TYMS plays a very important role in the prognosis and progression of cervical cancer, and has the potential to be thought of as cervical cancer biomarkers. At the same time, miR-197-3p/TYMS axis can regulate the deterioration of cervical cancer cells, which lays a foundation for the molecular diagnosis and treatment of cervical cancer.

Keywords: cervical cancer, miR-197-3p/TYMS, proliferation, apoptosis, invasion and migration

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Nana Jin,
Codex Genetics Limited, Hong Kong SAR, China
Nicola Mosca,
UOC di Neurologia, Fondazione Policlinico Universitario A. Gemelli IRCCS, Italy
Somnath Tagore,
Columbia University, United States

****Correspondence:***
Jinlong Qin
jinqinlong@yeah.net

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *13 September 2021*
Accepted: *29 November 2021*
Published: *22 December 2021*

Citation:
Fu M, Pei Y, Lu F, Jiang H, Bi Y, Cheng J and Qin J (2021) Identification of Potential Hub Genes and miRNA-mRNA Pairs Related to the Progression and Prognosis of Cervical Cancer Through Integrated Bioinformatics Analysis. Front. Genet. 12:775006. doi: 10.3389/fgene.2021.775006

INTRODUCTION

Cervical cancer is the most common female malignancy worldwide, and directly causes high incidence and mortality rates in women (Wang et al.,2020). According to the database of the International Agency for Research on Cancer, there were more than 500,000 new cases of cervical cancer and 311,000 deaths in 2018 (Stelzle et al.,2021). To date, surgery and radiotherapy are still the main treatment methods for cervical cancer, but recurrence, metastasis and drug resistance often occur after treatment (Sharma et al.,2020; Shin et al.,2020). The mechanism of cervical cancer is complex. Different genes, RNAs and signaling pathways are related to the tumorigenicity of cervical cancer (Huang et al., 2020; Rasmi and Sakthivel, 2020). Therefore, it is worthwhile to find new methods to study the basic mechanism of cervical cancer to improve treatment.

RNA plays an important role in the regulation of several major biological processes affecting tumorigenesis and progression, and has always been at the forefront of tumor molecular mechanism research (Goodall and Wickramasinghe, 2021; Barbieri & Kouzarides, 2020). The combination of high-throughput technology and bioinformatics analysis can provide researchers with valuable data available in the form of public datasets to search for biomarkers and therapeutic targets (Liu et al.,2020; Chen et al., 2008). Gene Expression Omnibus (GEO) facilitates the submission, storage, and retrieval of heterogeneous datasets from high-throughput gene expression and genomic experiments (Clough and Barrett, 2016; Yang et al., 2020). The Cancer Genome Atlas (TCGA) is a comprehensive dataset that provides a unified data analysis pipeline for further exploration of oncogene signaling changes and their associated significance in cancer patient outcomes (Hutter and Zenklusen, 2018; Linehan and Ricketts, 2019). Therefore, the combination of GEO and TCGA data sets may provide an important perspective for the study of new biomarkers. There have been many reports on screening tumor biomarkers based on GEO and TCGA data, and a series of markers with high specificity and sensitivity have been found (Ren et al., 2020; Xu et al.,2020). Compared with traditional screening methods, bioinformation-based analysis of high-throughput data enables researchers to obtain stable and reliable biomarkers in a large number of clinical samples.

In this study, we combined cervical cancer in GEO database and TCGA databases to screen differentially expressed genes and differentially expressed miRNAs. Based on the differentially expressed genes, the candidate gene TYMS was screened by functional annotation, pathway analysis, protein interaction (PPI) network construction, prognosis analysis, risk assessment model and WGCNA. Then, potential candidate pairs (miR-197-3p/TYMS) were screened by dual-luciferase experiment and qPCR. Subsequently, the effects of TYMS and miR-197-3p/TYMS on the progression of cervical cancer were analyzed by gain-of-function analysis. We hope to provide more extensive data support for the clinical application of TYMS and the miR-197-3p/TYMS axis in cervical cancer.

MATERIALS AND METHODS

Data Collection and Preprocessing

The original datasets were downloaded from geo database to compare the mRNA expression and miRNA expression between cervical cancer and normal tissues. The expression profile data of GSE7410, GSE63514, GSE86100 and GSE9750 were downloaded based on different platforms (**Table 1**). The mRNA expression data of cervical cancer were downloaded from TCGA database for preprocessing. Clinical samples related to cervical cancer were selected. The data set includes 307 cervical cancer samples, 307 normal samples and corresponding clinical data. Then, Sangerbox tool 2.0 (http://sangerbox.com/Index) corrects and normalizes the original expression data background of each GEO dataset, and eliminates batch effects and other irrelevant variables.

Differentially Expressed mRNAs and miRNAs Screening

The differentially expressed mRNAs and miRNAs between cervical cancer and normal tissues were identified by R software limma software package. The GEO dataset filtering criteria are: log2|Fold Change|>1, adjusted p-value< 0.05 and false discovery rate (FDR) < 0.05. The screening criteria of the TCGA dataset were: log2|Fold Change|>2, adjusted p-value< 0.05 and FDR<0.05. The ggplot2 and pheatmap packages were used to draw differential mRNAs and miRNAs volcano maps and heatmaps respectively. Then, a Venn diagram was used to find overlapping differentially expressed mRNAs.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Pathway Enrichment Analyses

To further analyze differentially expressed mRNAs, GO and KEGG enrichment analyses were performed using the DAVID online tool 6.8 (https://david.ncifcrf.gov/home.jsp) and Metascape (http://metascape.org/gp/index.html). GO enrichment analysis mainly annotates the biological process (BP), cytological component (CC) and molecular function (MF) of genes. KEGG enrichment analysis mainly predicts the signal pathways that may be involved. $p < 0.05$ was considered statistically significant.

PPI and Key Module Selection

PPI network interactions were identified and constructed using the STRING online database 11.5 (http://string-db.org). An interaction confidence≥0.4 was considered significant. Then, the data were imported into Cytoscape software 3.8.0 for reprocessing. Subsequently, the molecular complex detection (MCODE) plug-in was used to filter hub modules in the PPI network. Then, GO and KEGG analyses of mRNAs in hub modules were performed using Metascape (http://metascape.org/gp/index.html).

TABLE 1 | Information of GEO datasets.

Dataset		Platform	Tumor	Normal	References
GSE7410	GPL1708	Agilent-012391 Whole Human Genome Oligo Microarray G4112A	40	5	Biewenga et al. (2008)
GSE63514	GPL570	Affymetrix Human Genome U133 Plus 2.0 Array	28	24	den Boon et al. (2015)
GSE86100	GPL19730	Agilent-046064 Unrestricted Human miRNA V19.0 Microarray	8	4	Gao et al.2016
GSE9750	GPL96	Affymetrix Human Genome U133A Array	43	23	Scotto et al., 2008

Prognostic Analysis

The clinical characteristic data of cervical cancer patients from TCGA were downloaded. First, we subtyped the downloaded cervical cancer samples, and then analyzed the survival of each subtype and each pathological parameter. Then, univariate Cox proportional hazards regression analysis was used to identify potential genes highly related to overall survival and to screen mRNAs related to the prognosis of cervical cancer ($p < 0.05$). Then, the data were further screened by multivariate Cox proportional hazards regression. A risk assessment model was constructed with key prognostic mRNAs as dependent variables to evaluate the clinical value of key prognostic mRNAs in predicting patient survival. The risk assessment model formula is: risk score = expression of mRNA1 $\times \beta_{mRNA1}$ + expression of mRNA2 $\times \beta_{mRNA2}$ +. . . . . + expression of mRNAn $\times \beta_{mRNAn}$. β represents the multivariate Cox regression coefficient, which is obtained from the multivariable Cox proportional hazards regression model of each gene. According to the median risk score, patients with cervical cancer were divided into high and low-risk groups. Then the survival rates of the high and low-risk groups were analyzed. At the same time, a receiver operating characteristic (ROC) curve was constructed.

Weighted Gene Coexpression Network Analysis

First, we preprocessed the GSE9750 data, removed the samples with incomplete clinical data, removed the outliers, and performed WGCNA in the Sangerbox tool 2.0 (http://sangerbox.com/Index). The soft threshold was set as $\beta = 1$, and a scale-free network was constructed. Then, we transformed the adjacency matrix into a topological overlap matrix (TOM). Next, we performed hierarchical clustering to identify modules, and each module contained at least 20 genes (min module size = 20). Subsequently, characteristic genes were calculated, hierarchical clustering modules were used, and similar modules were merged (abline = 0.25). Finally, we calculated the correlation between the module and clinical data to determine meaningful clinical modules.

Gene Set Enrichment Analysis

We performed GSEA using normalized RNA SEQ data obtained by TCGA. GO and KEGG analyses of candidate genes were performed with Sangerbox tool 2.0 (http://sangerbox.com/Index) to study the possible biological functions of candidate genes. Adjusted p value < 0.05 and FDR <0.05 were considered to be statistically significant.

Cell Culture and Transfection

HeLa cells were purchased from Beijing Beina Chuang lian Biotechnology Research Institute (Beijing, China). The cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Invitrogen, United States). At the same time, 10% fetal bovine serum (Invitrogen, United States), 100 U/ml penicillin (Sigma, United States) and 100 g/ml streptomycin (Sigma, United States) were added. The cells were cultured in 5% CO_2 at 37 C in a constant-temperature incubator. The eighth passage HeLa cells were used in this study. According to the instructions for HeLa cells, Lipofectamine 3,000 reagent (Invitrogen, United States) was used to transfect 50 nm siRNA (si-TYMS; RiboBio, China), pCDH-TYMS (RiboBio, China), 50 nm miR-197-3p mimic (mir100000 227-one to five, RiboBio, China), 100 nm miR-197-3p inhibitor (mir200000 227-one to five, RiboBio, China) and their NC control for 24 h.

RNA Extraction and Real-Time Quantitative PCR

A total of 30 cervical cancer tissues and 30 adjacent normal tissues were collected (**Supplementary Material S1**). This study was approved by the hospital ethics committee and was carried out in accordance with the Declaration of Helsinki. In addition, each patient provided written informed consent. The cervical cancer tissue was cut off during surgery and immediately stored at 80°C until use.

Total RNA was extracted from tissues and cells using RNeasy Mini Kit (Qiagen, Germany). Then, the concentration of extracted RNA was determined with Nanodrop 2000 (Invitrogen, United States). The same amount of RNA was reverse transcribed into cDNA with Goldstar™ RT6 cDNA synthesis kit ver2 (TsingKe, China). Master qPCR mix (SYBR Green 1) (Beyotime, China) was used for qPCR detection. The expression of GAPDH was used as the endogenous control. Primer 5.0 software was used to design primers 5.0 (TYMS forward primer 5'-3': ACTTGTGCAGATTATTCAGGAC, TYMS reverse primer 5'-3': ATTCTTCTGTCGTCAGGGT; GAPDH forward primer 5'-3':CATTTCCTGGTATGACAA CGA, GAPDH reverse primer 5'-3': GGGTCTTACTCCTTG GAGG). The primer sequences were sent to Sangon (China) for synthesis. For miRNA detection, U6 was used as the endogenous control, and the primers were purchased from RiboBiology (miR-197-3p: mqps0000765-1-100, RiboBio, China). QPCR detection was performed using bulge loop miRNA QRT PCR Starter Kit (RiboBio, China). MiRNA and mRNA expression levels were calculated by $2^{-\Delta\Delta Ct}$.qRT-PCR data

were analyzed by GraphPad Prism, and each reaction was repeated 3 times. A *t* test was used in both groups ($p < 0.05$), and one-way ANOVA was used in more than two groups ($p < 0.05$).

CCK8 Assays

Logarithmic growth HeLa cells were inoculated into 96-well plates at 4,000 cells/well per well. 24 h after transfection, CCK-8 reagent (10 μl per well, Beyotime, China) was added to all cells, and the cells were incubated for 3 h. Then, a multi-function microplate (Thermo, United States) was used to measure the absorbance at 450 nm. Each experiment was conducted three times.

Wound-Healing Assay

HeLa cells at the logarithmic growth stage were inoculated into 96-well plates at 1×10^5 cells per well. 24 h after transfection, the fused cell monolayer was scratched with 200 μl sterile pipette tip and then added to serum-free DMEM. The scratch was recorded by microscopy at 0 and 24 h, and the scratch closure rate was evaluated by ImageJ software.

Transwell Assay

HeLa cells transfected for 24 h were collected for single-cell suspension, and 1×10^3 cells were added to each upper chamber (Corning, United States) containing matrix gel. DMEM with 20% fetal bovine serum was added to lower chamber. After 24 h, the cells were fixed with 4% paraformaldehyde (Beyotime, China) and stained with 1% crystal violet (Beyotime, China). Cells were observed and counted under a light microscope.

Cell Apoptosis

24 h after transfection, the HeLa cells were digested with trypsin to prepare single cell suspension. Then, the cells were gently washed with precooled PBS 3 times, and the cell concentration was maintained at 3×10^5 cell/ml. The cells were then treated with annexin V-FITC & PI apoptosis detection kit (Sangon, China). After treatment, the number of apoptotic cells analyzed by Beckman flow cytometry (Beckman, United States).

Luciferase Reporter Assay

HeLa cells were inoculated into 96-well plates and cotransfected with 100 ng of the dual-luciferase reporter constructs pmiR-RB-Report-TYMS 3'UTR (RiboBio, China) and miR-197-3p mimic (RiboBio, China). After incubation for 48 h, the supernatants of the cells were collected and the fluorescence value was determined using dual-luciferase reporter kit (Promega, United States).

Western Blotting

48 h after the transfection of HeLa cells, the culture medium was removed. After PBS rinsing, the cells were lysed with protein lysate (Beyotime, China), and the total proteins were collected. Then the protein concentration was detected by BCA (Beyotime, China) protein quantitative kit. After adding proper amount of SDS loading buffer (Beyotime, China), denaturation was carried out in boiling water at 100°C for 5 min. Then 12% SDS-PAGE electrophoresis (Beyotime, China) was performed. Protein bands were transferred to PVDF membranes (Beyotime, China) by western transmembrane system. Subsequently, the PVDF membrane was sealed in 5% skim milk powder (Beyotime, China) and incubated for 4 h. After washing with TBST (Beyotime, China), the PVDF membrane was incubated with primary antibody at 4°C overnight. After washing, the membrane was incubated with the secondary antibody at room temperature for 60 min. Finally, the PVDF membrane was stained with western TMB substrate (Beyotime, China; P0211) and scanned. The antibodies used were β-actin (Beyotime, China; AF5003), TYMS (Abcam, United States, ab108995) and horseradish peroxidase labeled goat anti-rabbit IgG (Beyotime, China; A0208).

Statistical Analysis

All data were processed by SPSS 25 and GraphPad Prism 8. The measurement data are expressed as mean ± standard deviation (SD). A *t* test and one-way ANOVA were performed between two groups and multiple groups, respectively. $p < 0.05$ was considered statistically significant.

RESULTS

Identification of Differentially Expressed miRNAs and mRNAs in Cervical Cancer

Three cervical cancer expression profile datasets GSE7410, GSE63514 and GSE86100 were downloaded from the GEO database. **Table 1** shows the details of the three GEO datasets. GSE7410 and GSE63514 screened for differentially expressed mRNAs between cervical cancer and normal tissues, and GSE86100 screened for differentially expressed miRNAs. Based on screening criteria, GSE7410 and GSE63514 obtained 1,138 mRNAs (709 down-regulated and 429 up-regulated) and 1,009 mRNAs (641 down-regulated and 368 up-regulated), respectively (**Figure 1A**, **Figures 1B–D**, **Figures 1A,B**, **Figure 1B**, **Supplementary Material S2** and **Supplementary Material S3**). GSE86100 obtained 166 miRNAs (74 down-regulated and 92 up-regulated) (**Figures 1A–C**, **Figures 1B,C** and **Supplementary Material S4**). In addition, 822 mRNAs (567 down-regulated and 255 up-regulated) were obtained (**Figures 1A–D**, **Figures 1A,B** and **Supplementary Material S5**) from TCGA-CESC. Subsequently, Venn diagram was used to compare mRNAs screened by GSE7410, GSE63514 and TCGA-CESC, and 34 overlapping differentially expressed mRNAs were obtained, including 22 up-regulated genes and 12 down-regulated genes (**Figure 1C**, **Supplementary Table S1**).

Functional Enrichment Analysis of Overlapping Genes

To explore the function and pathway of overlapping genes, we performed GO and KEGG pathway analyses. GO analysis showed that, in terms of biological progression, overlapping genes were mainly enriched in the regulation of mitotic nuclear division, T cell chemotaxis, regulation of nuclear division, leukocyte

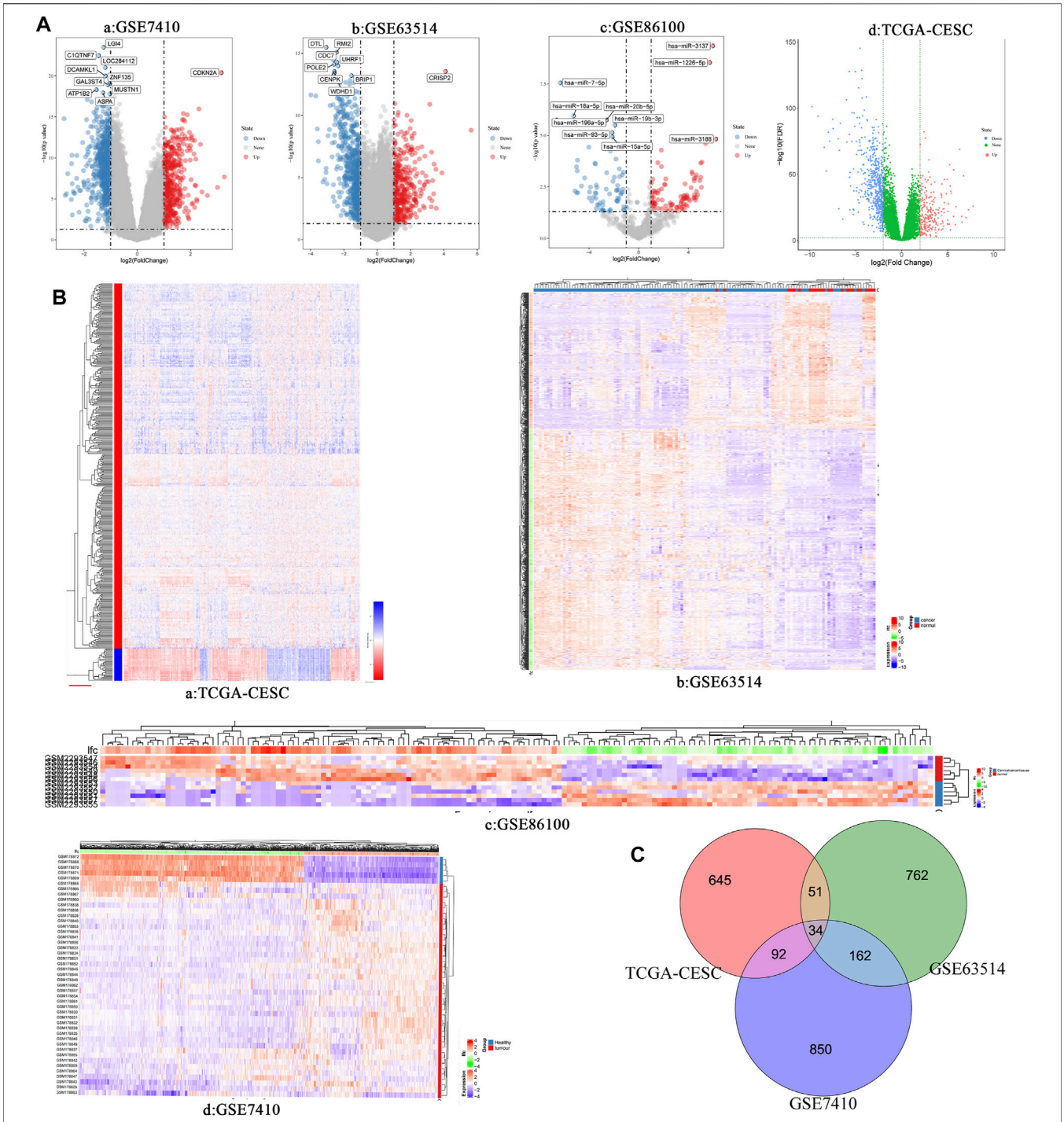


FIGURE 1 | (A) Volcano plots of differentially expressed RNAs (mRNA and miRNA) in different data set of cervical cancer. a: GSE7410; b: GSE63514; c: GSE86100; d: TCGA-CESC. **(B)** Heatmap of differentially expressed RNAs (miRNA and mRNA) in different data set of cervical cancer. a: TCGA-CESC; b: GSE63514; c: GSE86100; d: GSE7410. **(C)** Venn diagram show the intersection results of differentially expressed genes in three different data sets of cervical cancer.

chemotaxis, positive regulation of the release of sequestered calcium into the cytosol, lymphocyte chemotaxis and mitotic nuclear division (**Figure 2A**). In terms of molecular function, overlapping genes were mainly enriched in CXCR chemokine receptor binding, chemokine receptor binding, chemokine activity, receptor regulator activity, G protein-coupled receptor binding, receptor ligand activity and cytokine activity (**Figures 2A,B**). Pathway enrichment analysis showed that overlapping genes interact with the Toll-like receptor signaling pathway, the p53 signaling pathway, viral protein interaction with cytokine

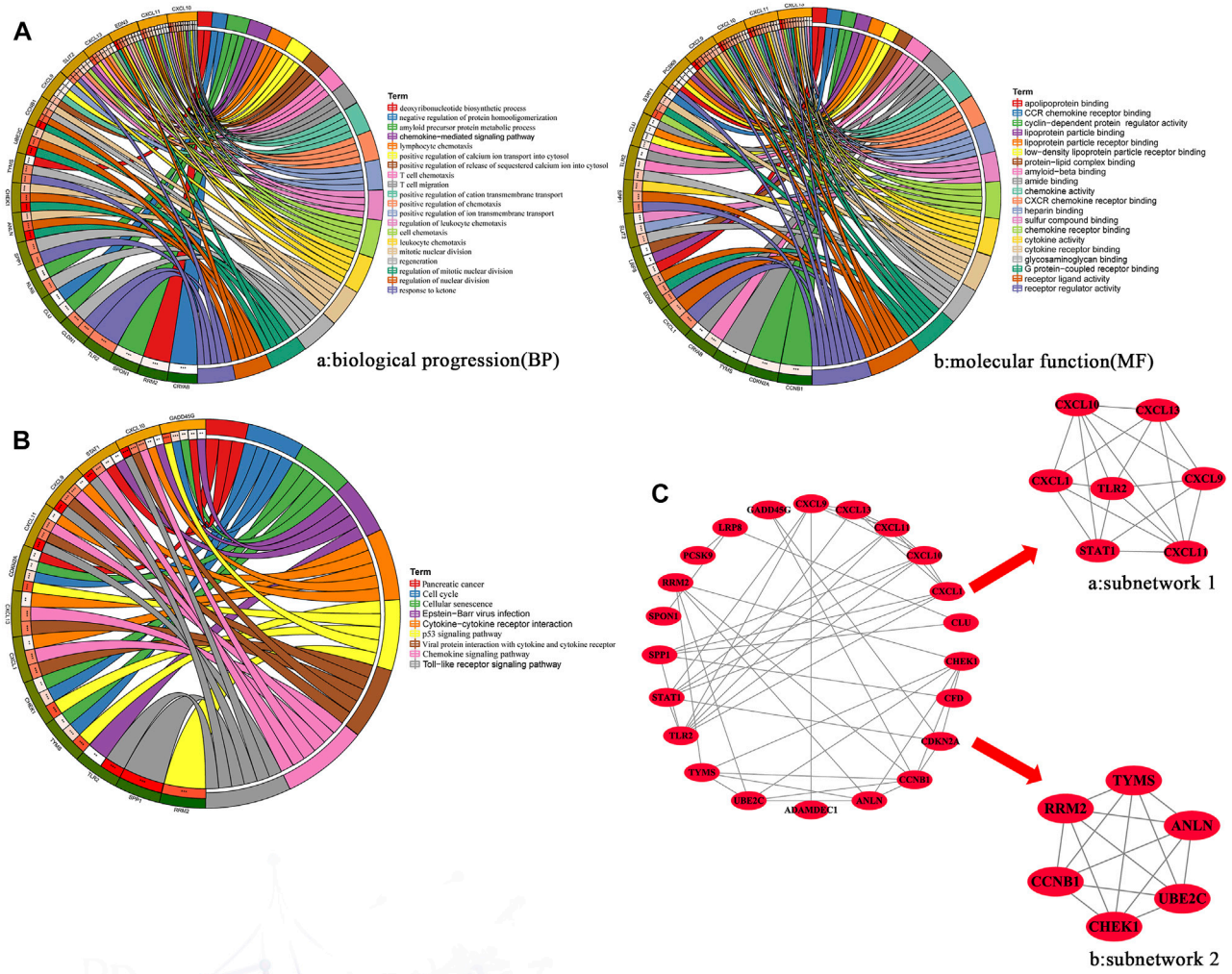


FIGURE 2 | (A) GO enrichment analysis of hub overlapping genes were performed using DAVID online tool 6.8. MF: Molecular function: Biological processes. **(B)** KEGG enrichment analysis of hub overlapping genes were performed using DAVID online tool 6.8. **(C)** PPI network of the hub overlapping genes. The subnetwork 1(a) and subnetwork (b) of hub genes was screened by MCODE in cytoscape.

and cytokine receptors, chemokine signaling pathways, the cell cycle and cellular senescence were related (**Figure 2B**).

Construction and Analysis of PPI Networks for Overlapping Genes

We constructed PPI networks for 34 overlapping genes using the SRTING online database 11.5. The PPI network contains 22 nodes and 48 edges (**Figure 2C**). Then, two hub subnetworks were filtered through the MCODE plug-in from the PPI network (**Figures 2A–C** and **Figures 2B,C**). Subnetwork one contains the genes CXCL1, CXCL9, CXCL10, CXCL13, STAT1 and TLR2. Subnetwork two contains the genes ANLN, CCNB1, CHEK1, RRM2, TYMS and UBE2C. Based on the TCGA-CESC dataset, we mapped the expression profiles of 12 hub genes in cervical and normal tissues (**Supplementary Figure S1**). GO and KEGG analysis results showed that genes of subnetwork one were closely related to immunity, stress chemotaxis, angiogenesis, Toll-like receptor signaling pathway and chemokine signaling pathway (**Supplementary Figure S2A**). Genes in subnetwork two were mainly enriched in mitosis, the cell cycle and the p53 signaling pathway (**Supplementary Figure S2B**).

Cox Regression and Proportional Risk Model Analysis of RRM2 and TYMS

First, overall survival (OS) data of cervical cancer patients were obtained from TCGA database. The survival curves of various subtypes and pathological parameters of cervical cancer were analyzed (**Supplementary Figure S3**). Subsequently, three genes were identified by univariate cox proportional risk regression model ($p < 0.05$) (**Supplementary Figure S4**, **Table 2**). Two further prognostic genes: RRM2 and TYMS, were selected by a multivariate Cox proportional risk regression model (**Table 2**).

TABLE 2 | Prognostic value of 12 hub genes in the cervical cancer patients of the TCGA cohort.

Hub genes	Univariate cox analysis		Multivariate cox analysis		Coefficient
	Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value	
CXCL1	2.29 (1.40–3.76)	0.033*	1.13 (1.00–1.27)	0.174	0.1181
CXCL9	0.42 (0.22–0.80)	0.077	0.91 (0.83–1.01)	0.071	-0.0908
CXCL10	0.63 (0.39–1.01)	0.160	0.95 (0.86–1.04)	0.250	-0.0560
CXCL13	0.55 (0.35–0.88)	0.360	0.94 (0.86–1.02)	0.129	-0.0667
STAT1	0.65 (0.40–1.03)	0.330	0.90 (0.7–1.13)	0.370	-0.1043
TLR2	0.58 (0.35–0.96)	0.180	0.85 (0.69–1.06)	0.160	-0.1578
ANLN	1.77 (1.10–2.83)	0.087	1.46 (1.03–2.07)	0.341	0.3790
CCNB1	0.59 (0.37–0.96)	0.320	0.87 (0.57–1.31)	0.493	-0.1450
CHEK1	0.80 (0.49–1.30)	0.910	1.08 (0.71–1.65)	0.715	0.0784
RRM2	1.40 (0.87–2.26)	0.029*	1.22 (0.81–1.83)	0.034*	0.2560
TYMS	0.47 (0.29–0.77)	0.041*	0.76 (0.52–1.12)	0.048*	-0.3332
UBE2C	0.75 (0.45–1.26)	0.510	1.00 (0.67–1.49)	0.997	8.00E-04

"" represents* p < *0.05.*

These two genes were also independent prognostic genes ($p < 0.05$). Risk prognosis model was established based on multivariate Cox screening results. The risk score was calculated as follows: mRNA risk score = (0.256)×expression (RRM2)+(-0.333)×expression (TYMS). Patients in TCGA-CESC were divided into high and low-risk groups based on the median risk score. The risk score of the TCGA-CESC dataset is shown in **Figure 3**. **Figure 3A** shows the details of the risk score. Kaplan-Meier survival analysis showed that the prognosis of high-risk group was worse than that of low-risk group (**Figure 3C**). The AUCs at 1 year, 3 and 5 years were 0.75, 0.72 and 0.69 respectively (**Figure 3B**). These results show that prognostic gene markers (RRM2 and TYMS) have good survival prediction ability, indicating that the model can effectively predict the prognosis of cervical cancer patients.

Co-Expression Network Construction

The coexpression network was constructed by (WGCNA) to further identify genes strongly associated with cervical cancer. First, the optimal soft threshold $\beta = 1$ was obtained through analysis, and the scale-free network was constructed (**Figure 4A**). Finally, two gene modules were obtained for further analysis (**Figure 4B**). From the heatmap of the topological overlap map, it can be seen that there was a high correlation between the genes in the module (**Figure 4C**). As shown in **Figure 4D**, the turquoise module was identified as the most specific module with a correlation coefficient of 0.84 (p = 2E-18). Therefore, this module was selected as clinically important for further analysis. By comparing the hub genes screened by multivariate Cox analysis with the genes in the turquoise module, we obtained a common key gene TYMS (**Supplementary Table S2**). This suggests that TYMS may play a very important role in the progression of cervical cancer.

Low Expression of TYMS is Associated With Clinical Progression in Patients With Cervical Cancer

To investigate the expression of TYMS in cervical cancer, qRT-PCR was used to detect the expression levels of TYMS in cervical cancer tissue samples and adjacent tissues. The results showed that TYMS was significantly down-regulated in cervical cancer (**Figure 4E**), which was consistent with GEO database analysis results. Subsequently, we downloaded the expression data of TYMS and the clinical characteristics of cervical cancer patients from TCGA-CESC, and evaluated the relationship between the expression level of TYMS and various clinicopathological parameters. The analysis showed that down-regulation of TYMS expression was significantly correlated with histological grading ($p < 0.001$), clinical stage ($p < 0.001$) and patient age ($p < 0.001$) (**Figure 4F**). These results suggest that cervical cancer patients with low expression of TYMS are more likely to have more advanced tumor status, grade and stage than those with high expression of TYMS.

GSEA of TYMS and its Expression in Relation to Immune Neoantigens, Tumor Mutation Load, Microsatellite Instability and Methylase Expression

To further analyze and explore the potential biological functions of TYMS, we performed GSEA on TYMS. As shown in **Figure 5A**, angiogenesis, myogenesis, inflammatory response and protein levels were significantly positively correlated with TYMS expression at GO concentrations. DNA repair, G2M checkpoint and apoptosis were negatively correlated with TYMS expression (**Figures 5A,B**). KEGG pathway analysis showed that the pathways with a significant positive correlation with TYMS expression included fatty acid metabolism, the PPAR signaling pathway, fatty acid metabolism, arachidonic acid metabolism and histidine metabolism. The p53 signaling pathway, cell cycle and pyrimidine metabolism were significantly negatively correlated with TYMS expression (**Figures 5A,B** and **Figure 5B**). These results suggest that essential cell cycle control, amino acid metabolic pathways and regeneration processes are closely related to TYMS expression in patients with cervical cancer. In addition, we

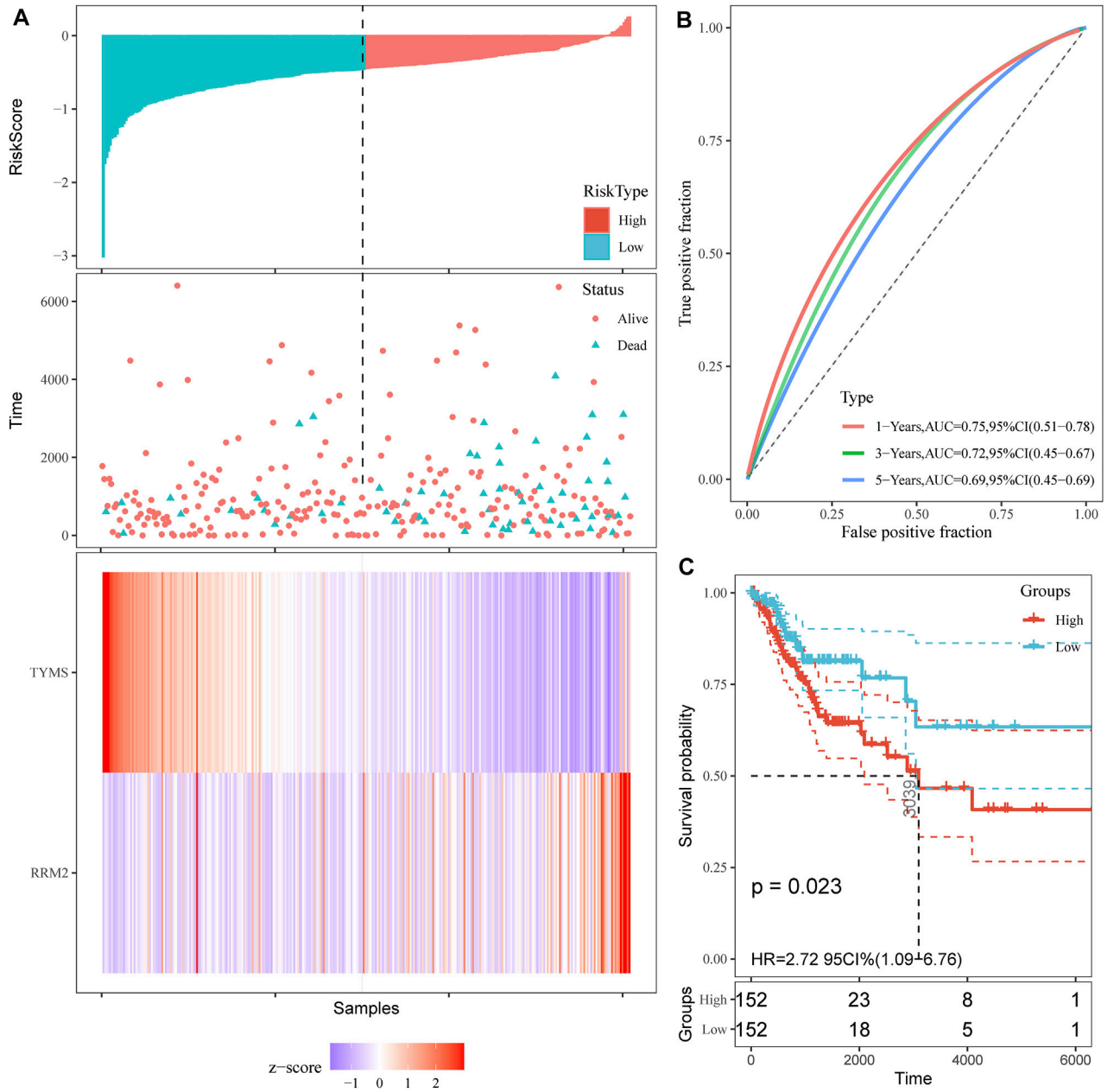


FIGURE 3 | Prognostic risk score model analysis of TYMS and RRM2 in cervical cancer. **(A)** Risk score and cervical cancer patient of survival status distribution, and heatmap of TYMS and RRM2 expression by risk score. **(B)** ROC curves for predicting survival in cervical cancer patients by the risk score. **(C)** Kaplan–Meier curves for high and low-risk groups of cervical cancer patient by risk score.

also analyzed the relationships between TYMS expression and immune neoantigens, tumor mutational burden, microsatellite instability and methylase expression. The results showed that the expression of TYMS was negatively correlated with tumor mutational burden (**Figure 5C**) and microsatellite instability (**Figure 5D**), and positively correlated with the expression of four methyltransferases (**Figure 5E**, DNMT1: red, DNMT2: blue, Dnmt3A: green and DNMT3B: purple). These results also show that low-level TYMS may easily leads to cell mutation, resulting in tumor formation or deterioration.

Interference With TYMS Expression can Regulate HeLa cell Proliferation, Migration, Apoptosis and Invasion

To further verify the analysis results of the biological function of TYMS, gain-of-function analysis was performed on HeLa

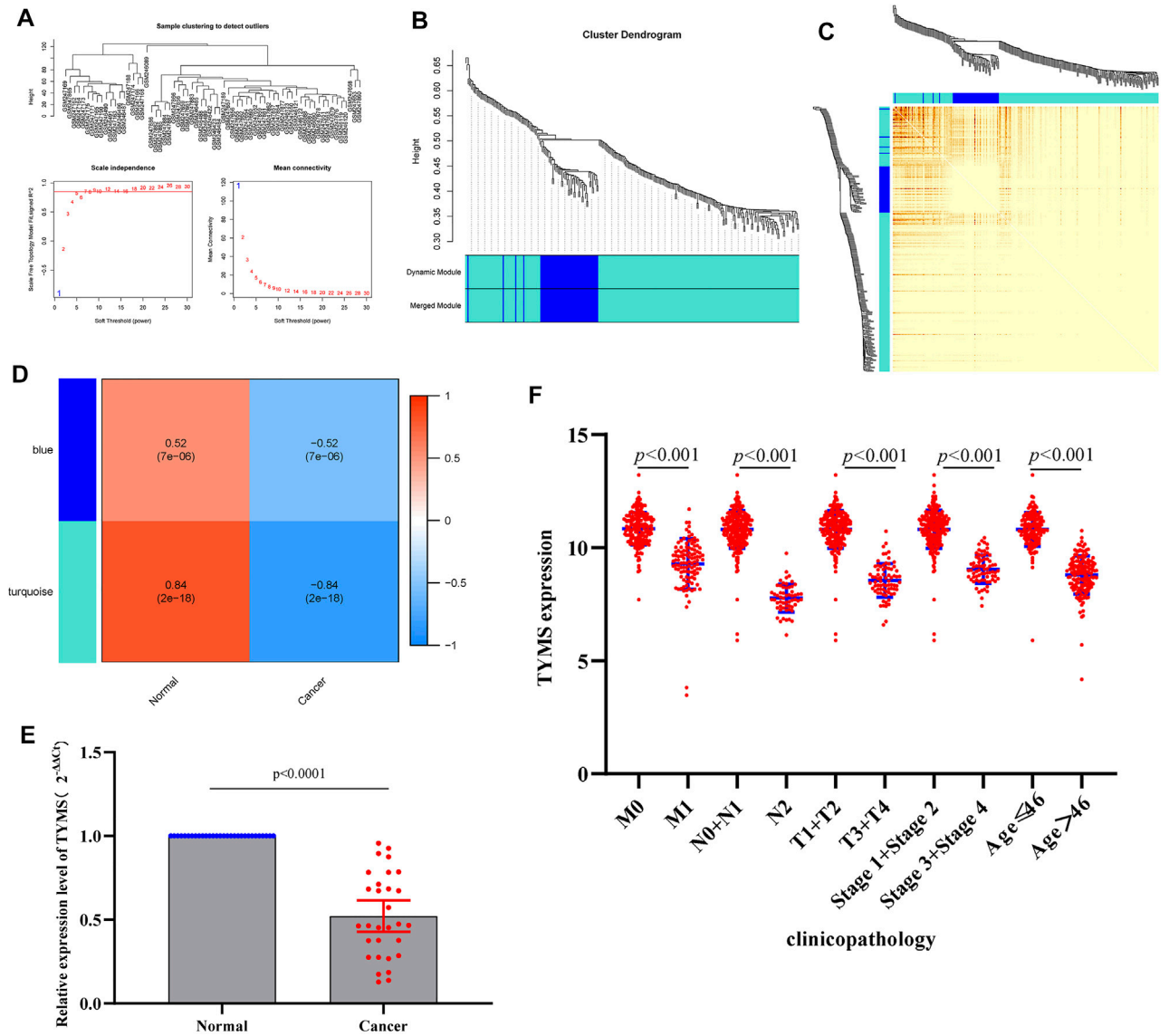


FIGURE 4 | (A) a scale-free network of samples in GSE9750 data. **(B)** Module Dendrogram of genes in GSE9750 data. **(C)** Heatmap of the topological overlap map. **(D)** Module–trait relationships (normal sample vs cervical cancer sample). **(E)** RT-qPCR results of the expression levels of TYMS in 30 paired cervical cancer tissues and their corresponding adjacent normal tissues. The data were shown as mean ± standard deviations. **(F)** Correlation analysis of the expression of TYMS with clinicopathological parameters.

cells transfected with the TYMS vector. QRT-PCR results showed that TYMS was effectively expressed in HeLa cells after transfection (**Figure 6A**). Then, the cell proliferation was detected by CCK-8 assay. The results showed that compared with the control group, HeLa cells transfected with si-TYMS showed significant growth promotion. In contrast, TYMS overexpression showed the opposite effect (**Figure 6B**). The results of wound healing and transwell assay showed that silencing TYMS promoted cell migration and invasion, and overexpression of TYMS inhibited cell migration and invasion (**Figure 6C**, **Figure 6D**, **Figure 6E** and **Figure 6F**). However, flow cytometry showed that overexpression of TYMS significantly promoted the amount of apoptosis in HeLa cells, while the apoptosis rate was relatively low when TYMS was knocked out (**Figure 6G** and **Figure 6H**). These results suggest that TYMS may be involved in the development of cervical cancer.

TYMS is a Direct Target Gene of miR-197-3p in Cervical Cancer

To determine the potential mechanism of TYMS in cervical cancer, we analyzed the potential miRNAs targeted by TYMS using miRWalk and starbase online tools (**Supplementary**

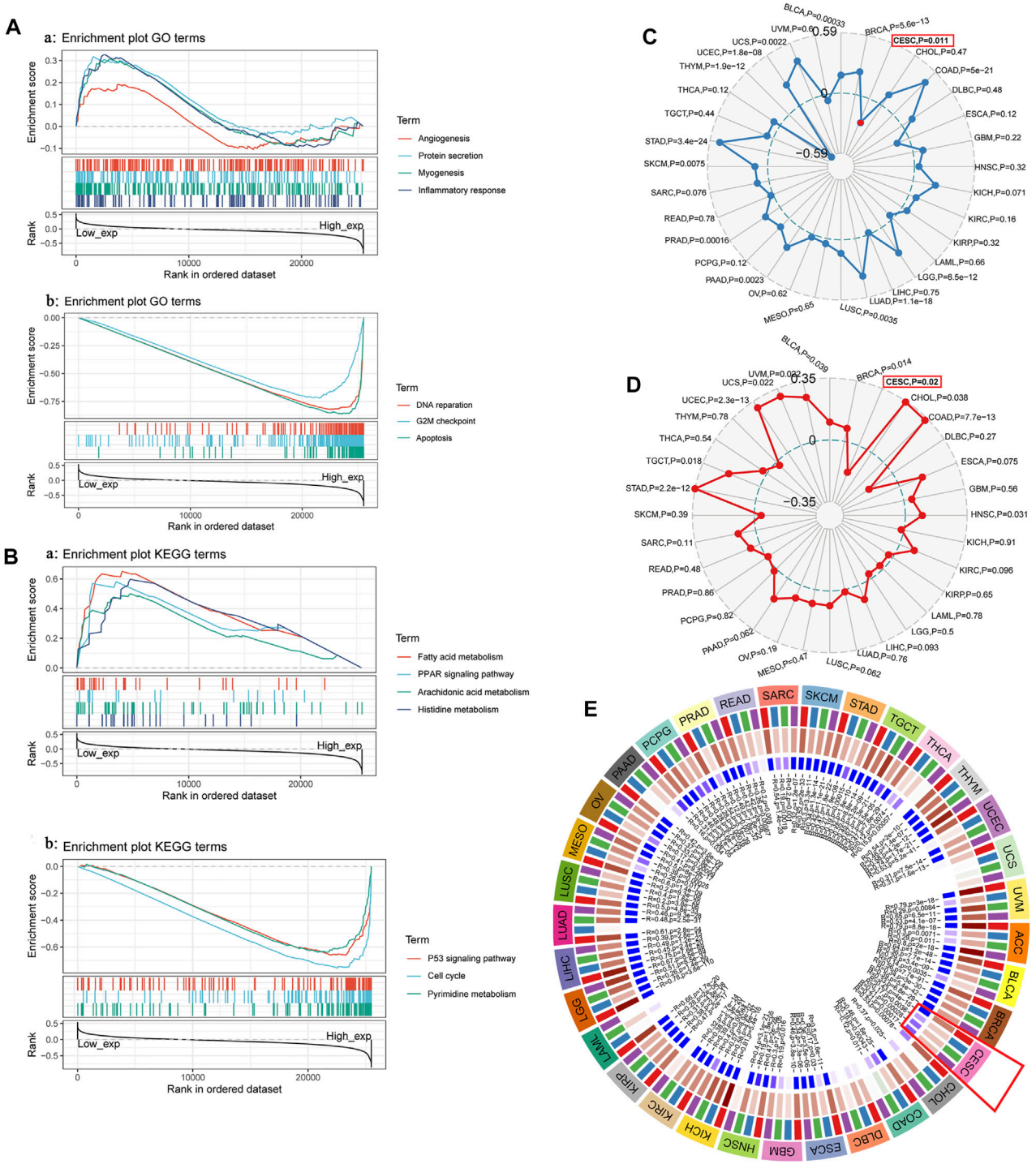


FIGURE 5 | (A) and **(B)** GSEA investigation of TYMS. **(C)** The radar diagram shown the correlation between TYMS gene expression and tumor mutational burden (TMB). The red dots represent the correlation between TYMA and TMB in cervical cancer. The dotted blue circle in the middle represents the dividing line between positive and negative correlations. **(D)**The radar diagram shown the correlation between TYMS gene expression and microsatellite instability (MSI). The red box represents the association between TYMS and MSI in cervical cancer. The dotted circle in the middle represents the dividing line between positive and negative correlations. **(E)** Correlation between TYMS gene expression and four methyltransferases (DNMT1: red, DNMT2: blue, DNMT3A: green, DNMT3B: purple) expression. The red box represents the association between TYMS and methyltransferases in cervical cancer.

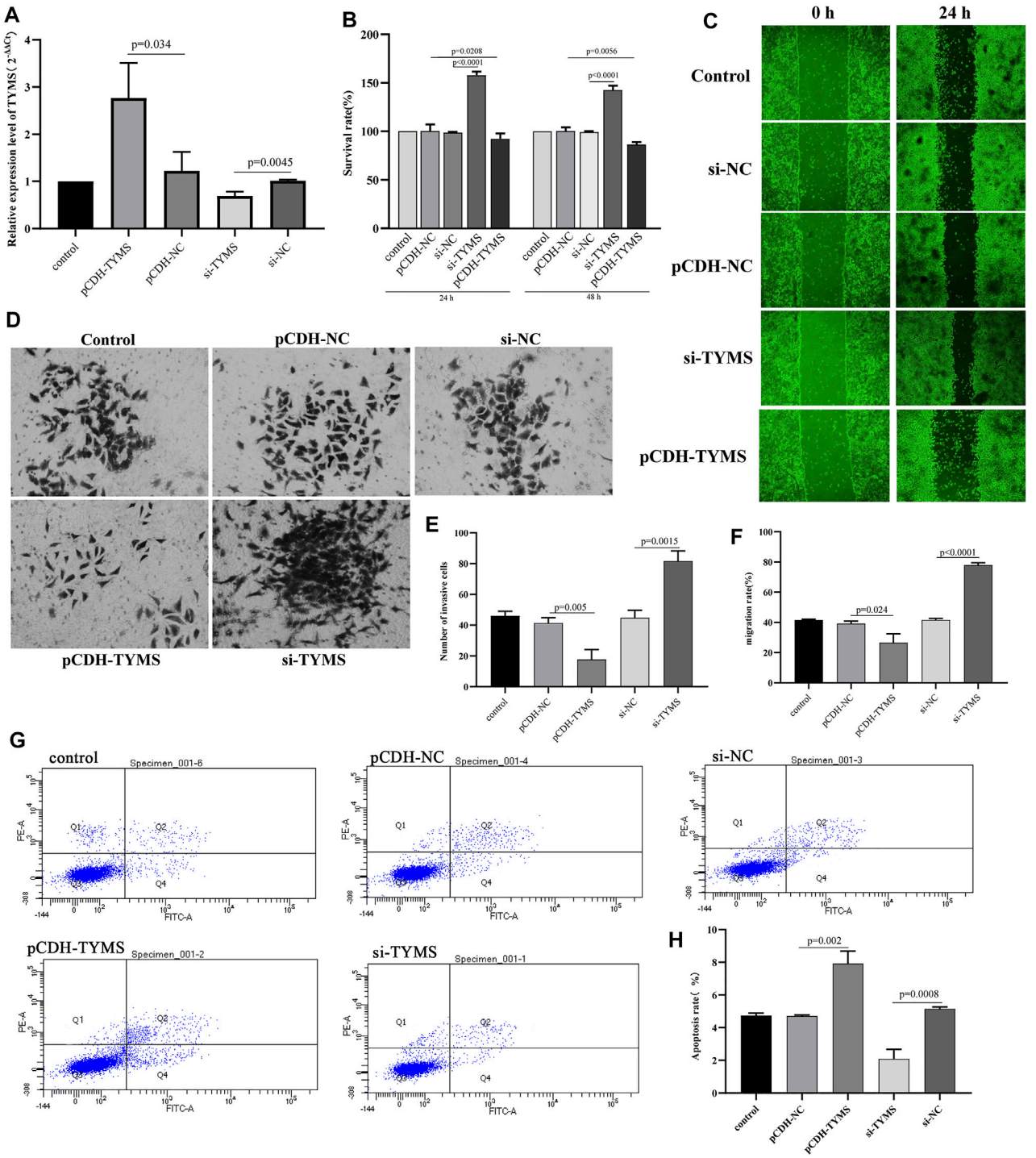


FIGURE 6 | (A) QPCR results of the transfection efficiency in HeLa cell with TYMS overexpressed plasmid and si-TYMS. **(B)** The effects of si-TYMS and TYMS overexpressed on the proliferation activity of HeLa cells measured using the CCK-8. **(C)** and **(F)** The effects of si-TYMS and TYMS overexpressed on the migration of HeLa cells measured using wound-healing assay. **(D)** and **(E)** The effects of si-TYMS and TYMS overexpressed on the invasion of HeLa cells measured using *trans*-well assay. **(G)** and **(H)** The effects of si-TYMS and TYMS overexpressed on the apoptosis of HeLa cells measured using flow cytometry.

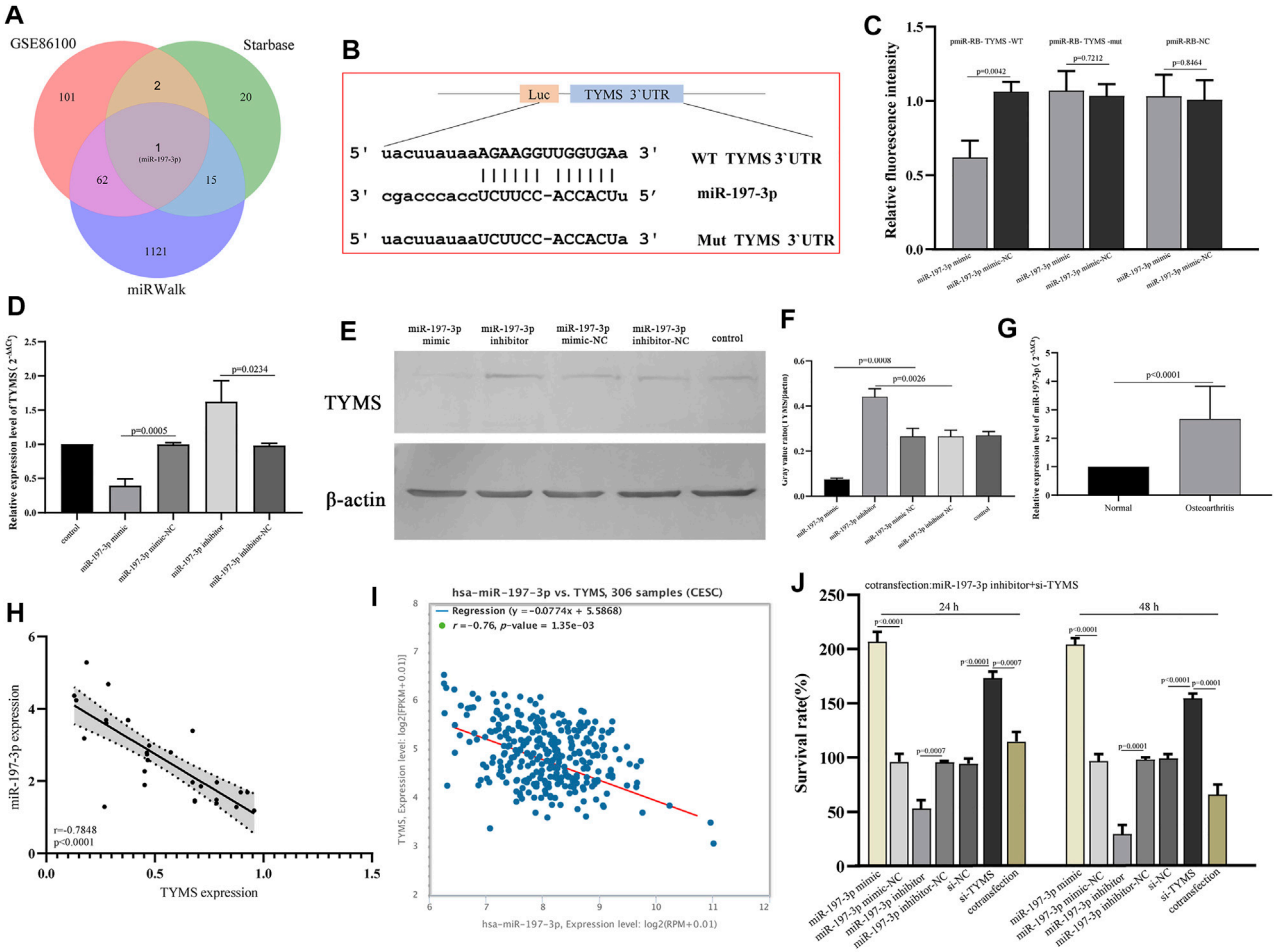


FIGURE 7 | (A) Venn diagram show the intersection results of miRNAs from starbase, miRwalk and GSE86100. **(B)** The 3′UTR binding site sequence of miR-197-3p and TYMS. **(C)** Results of dual luciferase assay for miR-197-3p and TYMS. **(D)** QPCR results of expression changes of TYMS in HeLa cells after transfection with miR-197-3p mimic and inhibitor. **(E)** and **(F)** Western blot of expression changes of TYMS in HeLa cells after transfection with miR-197-3p mimic and inhibitor. **(G)** RT-qPCR results of the expression levels of miR-197-3p in 30 paired cervical cancer tissues and their corresponding adjacent normal tissues. The data were shown as mean ± standard deviations. **(H)** The correlation between the expression of TYMS and the expression of miR-197-3p in cervical cancer tissues and their corresponding adjacent normal tissues. **(I)** The correlation between the expression of TYMS and the expression of miR-197-3p in cervical cancer by starbase. **(J)** The effects of si-TYMS, pCDH-TYMS, miR-197-3p mimic, miR-197-3p inhibitor and co-transfection (si-TYMS and miR-197-3p inhibitor) on the proliferation activity of HeLa cells measured using the CCK-8.

Material S6 and **Supplementary Material S7**). The analysis results were compared with the screened differentially expressed miRNAs, and a hub miRNA miR-197-3p was obtained (**Figure 7A**). The 3'UTR binding site sequences of miR-197-3p and TYMS are shown in **Figure 7B**. To verify the correlation between miR-197-3p and TYMS, luciferase reporter gene detection was performed. The results showed that miR-197-3p overexpression caused a significant decrease in luciferase activity in the wt-TYMS group. While, overexpression of miR-197-3p failed to inhibit luciferase activity in the mut-TYMS group (**Figure 7C**). In addition, qRT-PCR and western blot results showed that overexpression of miR-197-3p could reduce the expression of TYMS in HeLa cells. Similarly, the expression of TYMS was increased in HeLa cells after transfection with miR-197-3p inhibitor (**Figure 7D**, **Figure 7E** and **Figure 7F**). These results suggest a direct targeting relationship between TYMS and miR-197-3p.

Overexpression of miR-197-3p Enhanced the Effect of TYMS on the Proliferation of HeLa Cells

To further confirm the regulatory relationship between miR-197-3p and TYMS, we detected the mRNA level of miR-197-3p in cervical cancer tissues. Detection results showed that miR-197-3p was up-regulated in cervical cancer tissues (**Figure 7G**). Then, correlation analysis was conducted on

the expression levels of miR-197-3p and TYMS, and the results showed a significant negative correlation between miR-197-3p and TYMS (**Figure 7H**), which was consistent with the results of bioinformatics analysis (**Figure 7I**). The results of proliferation assay showed that overexpression of miR-197-3p significantly enhanced the proliferation activity of HeLa cells. In contrast, the miR-197-3p inhibitor inhibited HeLa cell proliferation (**Figure 7J**). In HeLa cells cotransfected with miR-197-3p inhibitor and si-TYMS, cell proliferation activity was significantly lower than that in the TYMS knockout group (**Figure 7J**). These results suggested that down-regulation of miR-197-3p could reverse the effect of the TYMS inhibitor on HeLa cell viability and suggest that miR-197-3p directly targets TYMS and negatively regulates its expression.

DISCUSSION

Worldwide, cervical cancer is the second most common malignancy in women and a leading cause of morbidity and mortality (Jedy-Agba et al.,2020; Shrestha et al.,2020). The greatest progress has been made in reducing cervical cancer deaths with the advent and implementation of screening programs (Silver and Kobri, 2020). Despite some advances in diagnostic accuracy, the prognosis for cervical cancer patients remains unsatisfactory. In previous studies, it has been found that mRNA can regulate the carcinogenesis of cervical cancer by binding miRNA (Ren et al., 2020). confirmed that miR-587 can promote the occurrence of cervical cancer by inhibiting interferon regulatory factor 6. MiR-125 inhibits the progression of cervical cancer by regulating VEGF and PI3K/AKT signaling pathways (Fu et al.,2020). In addition, Sun et al. (2020) proposed that miR-4429 sensitized cervical cancer cells to radiation by targeting RAD51. Similarly, our results suggest that TYMS is a protective factor against cervical cancer and is down-regulated in cervical cancer tissues. However, miR-197-3p directly targeted by TYMS is present a high level in cervical cancer tissues, and up-regulation of TYMS can inhibit the proliferation, invasion and migration of HeLa cells, and promote the apoptosis of HeLa cells. Up-regulation of miR-197-3p reversed the inhibitory effect of TYMS overexpression on HeLa cell proliferation. This evidence suggests that miR-197-3p/TYMS plays an important role in the progression of cervical cancer.

In this study, based on the expression profile data of cervical cancer in GEO and TCGA, through PPI, functional enrichment, Kaplan-Meier plotter analysis, Cox univariate and multivariate analysis, WGCNA and other bioinformatics analysis methods, we screened the factor, TYMS, which was significantly correlated with the prognosis of cervical cancer. TYMS is located on chromosome 18p and encodes thymine synthetase (TS), which is an essential enzyme involved in DNA replication and repair, and plays an important role in the biosynthesis of dTMP (Gallegos-Arreola et al.,2018; Fu et al.,2019). Its role in tumor cell proliferation has been reported. TYMS overexpression is associated with poor disease-specific survival and local recurrence-free survival of various solid tumors such as lung cancer (Ding et al.,2020), gastric cancer (Cao et al.,2017) and prostate cancer (Burdelski et al.,2015). TYMS is more frequently expressed in patients with clinically advanced prostate cancer than in patients with advanced prostate cancer (Burdelski et al.,2015). In colorectal cancer, knockdown of TYMS can inhibit tumor cell proliferation and promote cell apoptosis (Varghese et al.,2019). We also found that the expression level of TYMS was correlated with the tissue grade, tumor grade and age of cervical cancer patients, and TYMS with low expression was more likely to have advanced tumor state, grade and staging. GSEA further found that TYMS may be involved in cell cycle, apoptosis, amino acid metabolism and other biological processes in cervical cancer. The expression level of TYMS was negatively correlated with microsatellite instability (MSI) and Tumor mutation burden (TMB), and positively correlated with methylase expression in DNMT1, DNMT2, DNMT3A and DNMT3B. MSI and TMB are predictive biomarkers of immune checkpoint inhibitors (ICIs), reflecting the number of mutations in somatic cells (Salem et al.,2020). Cancer is a disease characterized by abnormal and uncontrolled cell growth mainly caused by gene mutations, and its mutation forms affect the steady-state development of a series of key cell functions (Martínez-Jiménez et al., 2019), which indicates that high MSI and high TMB may lead to tumorigenesis or affect tumor progression. In addition, the precise control of chromatin kinetics depends on the coordination of DNA methylation and covalent histone modification, and the enzymes and chromatin factors involved in these processes are also regulated by various RNA. The loss of control of these processes will directly lead to the occurrence of various pathological processes, such as tumors (Köhler F and Rodríguez-Paredes, 2020). The expression of TYMS is related to the content of methylase, which indicates that TYMS may be involved in the cellular genetic epigenetic regulation process, and its abnormal expression may lead to the process out of control. These analyses need more experimental data to verify. However, the study of TYMS in cervical cancer has not been reported, and its regulatory mechanism on the progression of cervical cancer is still unknown.

Through dual-luciferase reporter gene and functional experiments, it was clarified that TYMS may be the target gene of miR-197-3p in cervical cancer. Moreover, interference with the miR-197-3p mimic reduced the expression level of TYMS. After transfection with the miR-197-3p inhibitor, the expression levels of TYMS at the mRNA and protein levels were enhanced. *In vitro* functional experiments showed that inhibition of miR-197-3p could reverse the effects of TYMS overexpression on the proliferation activity of HeLa cells. These observations suggest that miR-197-3p can promote the progression of cervical cancer by targeting TYMS. MiR-197-3p has been

reported to be involved in several human cancers (Ni et al., 2019; Huang et al., 2020). MiR-197-3p is significantly down-regulated in ovarian cancer, and overexpression can inhibit the growth of ovarian cancer cells (Xie et al.,2020). MiRNAs generally exert their regulatory role by selectively targeting protein-coding genes. Similarly, miR-197-3p has been shown to target multiple genes in human cells, such as KLF10 (You et al.,2021), ABCA7(Xie et al.,2020) and EHD2 (Fan et al.,2020). However, TYMS as a target of miR-197-3p in any type of cancer cell has not been studied. In summary, miR-197-3p, as an upstream regulator of TYMS, affects the progression of cervical cancer through the posttranscriptional effects of TYMS, which is crucial for the tumorigenicity of cervical cancer. The results point toward the therapeutic implications of miR-197-3p in human cervical cancer cells. However, studies in different cell lines and *in vivo* conditions need further confirmation.

This study did not further evaluate the potential signaling pathway of miR-197-3p/TYMS in cervical cancer. These signaling pathways are responsible for tumor cell progression behavior. However, in the functional enrichment and GSEA data of this study, TYMS was mainly related to p53, the cell cycle and other signaling pathways. Therefore, future research will focus on the potential signaling pathway mechanism and the *in vivo* regulation of miR-197-3p/TYMS in cervical cancer.

CONCLUSION

In conclusion, our data show that TYMS is an independent prognostic gene of cervical cancer, that affects the proliferation, migration, invasion and apoptosis of cervical cancer cells. Therefore, it has a potential role as a biomarker for cervical cancer. MiR-197-3p/TYMS is an effective pair of regulators of cervical cancer, and can provide a good data basis for the treatment of cervical cancer.

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 studies involving human participants were reviewed and approved by This study was approved by the Ethical Committee of Shanghai Fourth People's Hospital Affiliated to Tongji University School of Medicine. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MF, YP and JQ conducted data analysis and wrote manuscripts. FL, HJ, JC and YB carried out data analysis and manuscript revision. JQ designed the study. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank all TCGA and GEO data builders and data contributors, as well as the team that built the Sangerbox and starbase online analysis page, and thank to Shanghai Ordovician Biotechnology Co., LTD. for providing the platform of biological information analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.775006/full#supplementary-material

REFERENCES

Barbieri, I., and Kouzarides, T. (2020). Role of RNA Modifications in Cancer. *Nat. Rev. Cancer* 20 (6), 303–322. doi:10.1038/s41568-020-0253-2

Biewenga, P., Buist, M. R., Moerland, P. D., van Themaat, E. V. L., van Kampen, A. H. C., ten Kate, F. J. W., et al. (2008). Gene Expression in Early Stage Cervical Cancer. *Gynecol. Oncol.* 108 (3), 520–526. doi:10.1016/j.ygyno.2007.11.024

Burdelski, C., Strauss, C., Tsourlakis, M. C., Kluth, M., Hube-Magg, C., Melling, N., et al. (2015). Overexpression of Thymidylate Synthase (TYMS) Is Associated with Aggressive Tumor Features and Early PSA Recurrence in Prostate Cancer. *Oncotarget* 6 (10), 8377–8387. doi:10.18632/oncotarget.3107

Cao, Y., Zhang, G., Wang, P., Zhou, J., Gan, W., Song, Y., et al. (2017). Clinical Significance of UGT1A1 Polymorphism and Expression of ERCC1, BRCA1, TYMS, RRM1, TUBB3, STMN1 and TOP2A in Gastric Cancer. *BMC Gastroenterol.* 17 (1), 2. doi:10.1186/s12876-016-0561-x

Chen, W., Gao, C., Liu, Y., Wen, Y., Hong, X., and Huang, Z. (2008). Bioinformatics Analysis of Prognostic miRNA Signature and Potential Critical Genes in Colon Cancer. *Front. Genet.* 11, 478. doi:10.3389/fgene.2020.00478

Clough, E., and Barrett, T. (2016). The Gene Expression Omnibus Database. *Methods Mol. Biol.* 1418, 93–110. doi:10.1007/978-1-4939-3578-9_5

den Boon, J. A., Pyeon, D., Wang, S. S., Horswill, M., Schiffman, M., Sherman, M., et al. (2015). Molecular Transitions from Papillomavirus Infection to Cervical Precancer and Cancer: Role of Stromal Estrogen Receptor Signaling. *Proc. Natl. Acad. Sci. USA* 112 (25), E3255–E3264. doi:10.1073/pnas.1509322112

Ding, X., Gu, Y., Jin, M., Guo, X., Xue, S., Tan, C., et al. (2020). The Deubiquitinating Enzyme UCHL1 Promotes Resistance to Pemetrexed in Non-small Cell Lung Cancer by Upregulating Thymidylate Synthase. *Theranostics* 10 (13), 6048–6060. doi:10.7150/thno.42096

Fan, H., Liu, T., Tian, H., and Zhang, S. (2020). TUSC8 Inhibits the Development of Osteosarcoma by Sponging miR-197-3p and T-argeting EHD2. *Int. J. Mol. Med.* 46 (4), 1311–1320. doi:10.3892/ijmm.2020.4684

Fu, K., Zhang, L., Liu, R., Shi, Q., Li, X., and Wang, M. (2020). MiR-125 Inhibited Cervical Cancer Progression by Regulating VEGF and PI3K/AKT Signaling Pathway. *World J. Surg. Onc* 18 (1), 115. doi:10.1186/s12957-020-01881-0

Fu, Z., Jiao, Y., Li, Y., Ji, B., Jia, B., and Liu, B. (2019). TYMS Presents a Novel Biomarker for Diagnosis and Prognosis in Patients with Pancreatic Cancer. *Medicine (Baltimore)* 98 (51), e18487. doi:10.1097/MD.0000000000018487

Gallegos-Arreola, M. P., Zúñiga-González, G. M., Sánchez-López, J. Y., Naranjo-Cruz, A. Y., Peralta-Leal, V., Figuera, L. E., et al. (2018). TYMS 2R3R Polymorphism and DPYD [IVS]14+1G>A Mutation Genes in Mexican Colorectal Cancer Patients. *Acta Biochim. Pol.* 65 (2), 227–234. doi:10.18388/abp.2017_2338

Gao, D., Zhang, Y., Zhu, M., Liu, S., and Wang, X. (20162016). miRNA Expression Profiles of HPV-Infected Patients with Cervical Cancer in the Uyghur Population in China. *PLoS One* 11 (10), e0164701. doi:10.1371/journal.pone.0164701

Goodall, G. J., and Wickramasinghe, V. O. (2021). RNA in Cancer. *Nat. Rev. Cancer* 21 (1), 22–36. doi:10.1038/s41568-020-00306-0

Huang, H., Chen, Y.-F., Du, X., and Zhang, C. (2020). Identification and Characterization of Tumorigenic Circular RNAs in Cervical Cancer. *Cell Signal.* 73, 109669. doi:10.1016/j.cellsig.2020.109669

Huang, Q., Ma, B., Su, Y., Chan, K., Qu, H., Huang, J., et al. (2020). miR-197-3p Represses the Proliferation of Prostate Cancer by Regulating the VDAC1/AKT/β-catenin Signaling Axis. *Int. J. Biol. Sci.* 16 (8), 1417–1426. doi:10.7150/ijbs.42019

Hutter, C., and Zenklusen, J. C. (2018). The Cancer Genome Atlas: Creating Lasting Value beyond its Data. *Cell* 173 (2), 283–285. doi:10.1016/j.cell.2018.03.042

Jedy-Agba, E., Joko, W. Y., Liu, B., Buziba, N. G., Borok, M., Korir, A., et al. (2020). Trends in Cervical Cancer Incidence in Sub-saharan Africa. *Br. J. Cancer* 123 (1), 148–154. doi:10.1038/s41416-020-0831-9

Köhler, F., and Rodríguez-Paredes, M. (2020). DNA Methylation in Epidermal Differentiation, Aging, and Cancer. *J. Invest. Dermatol.* 140 (1), 38–47. doi:10.1016/j.jid.2019.05.011

Linehan, W. M., and Ricketts, C. J. (2019). The Cancer Genome Atlas of Renal Cell Carcinoma: Findings and Clinical Implications. *Nat. Rev. Urol.* 16 (9), 539–552. doi:10.1038/s41585-019-0211-5

Liu, S., Wang, W., Zhao, Y., Liang, K., and Huang, Y. (2020). Identification of Potential Key Genes for Pathogenesis and Prognosis in Prostate Cancer by Integrated Analysis of Gene Expression Profiles and the Cancer Genome Atlas. *Front. Oncol.* 10, 809. doi:10.3389/fonc.2020.00809

Martínez-Jiménez, F., Muiños, F., Sentís, I., Deu-Pons, J., Reyes-Salazar, I., Arnedo-Pac, C., et al. (2020). A Compendium of Mutational Cancer Driver Genes. *Nat. Rev. Cancer* 20 (10), 555–572. doi:10.1038/s41568-020-0290-x

Ni, J. S., Zheng, H., Huang, Z. P., Hong, Y. G., Ou, Y. L., Tao, Y. P., et al. (2019). MicroRNA-197-3p A-cts as a P-rognostic M-arker and I-nhibits C-ell I-nvasion in H-epatocellular C-arcinoma. *Oncol. Lett.* 17 (2), 2317–2327. doi:10.3892/ol.2018.9848

Rasmi, R. R., Sakthivel, K. M., and Guruvayoorappan, C. (2020). NF-κB Inhibitors in Treatment and Prevention of Lung Cancer. *Biomed. Pharmacother.* 130, 110569. doi:10.1016/j.biopha.2020.110569

Ren, N., Liang, B., and Li, Y. (2020). Identification of Prognosis-Related Genes in the Tumor Microenvironment of Stomach Adenocarcinoma by TCGA and GEO Datasets. *Biosci. Rep.* 40 (10), BSR20200980. doi:10.1042/BSR20200980

Ren, Y., Dong, J., He, P., Liang, Y., Wu, L., Wang, J., et al. (2020). miR-587 Promotes Cervical Cancer by Repressing Interferon Regulatory Factor 6. *J. Gene Med.* 22 (11), e3257. doi:10.1002/jgm.3257

Salem, M. E., Bodor, J. N., Puccini, A., Xiu, J., Goldberg, R. M., Grothey, A., et al. (2020). Relationship between MLH1 , PMS2 , MSH2 and MSH6 Gene-specific Alterations and Tumor Mutational burden in 1057 Microsatellite Instability-high Solid Tumors. *Int. J. Cancer* 147 (10), 2948–2956. doi:10.1002/ijc.33115

Scotto, L., Narayan, G., Nandula, S. V., Arias-Pulido, H., Subramaniyam, S., Schneider, A., et al. (2008). Identification of Copy Number Gain and Overexpressed Genes on Chromosome Arm 20q by an Integrative Genomic Approach in Cervical Cancer: Potential Role in Progression. *Genes Chromosom. Cancer* 47 (9), 755–765. doi:10.1002/gcc.20577

Sharma, S., Deep, A., and Sharma, A. K. (2020). Current Treatment for Cervical Cancer: An Update. *Acamc* 20 (15), 1768–1779. doi:10.2174/1871520620666200224093301

Shin, H. J., Han, J. M., Choi, Y. S., and Jung, H. J. (2020). Pterostilbene Suppresses Both Cancer Cells and Cancer Stem-like Cells in Cervical Cancer with Superior Bioavailability to Resveratrol. *Molecules* 25 (1), 228. doi:10.3390/molecules25010228

Shrestha, G., Mulmi, R., Phuyal, P., Thakur, R. K., and Siwakoti, B. (2020). Experiences of Cervical Cancer Survivors in Chitwan, Nepal: A Qualitative Study. *PLoS One* 15 (11), e0234834. doi:10.1371/journal.pone.0234834

Silver, M. I., and Kobrin, S. (2020). Exacerbating Disparities?: Cervical Cancer Screening and HPV Vaccination. *Prev. Med.* 130, 105902. doi:10.1016/j.ypmed.2019.105902

Stelzle, D., Tanaka, L. F., Lee, K. K., Ibrahim Khalil, A., Baussano, I., Shah, A. S. V., et al. (2021). Estimates of the Global burden of Cervical Cancer Associated with HIV. *Lancet Glob. Health* 9 (2), e161–e169. doi:10.1016/s2214-109x(20)30459-9

Sun, H., Fan, G., Deng, C., and Wu, L. (2020). miR-4429 Sensitized Cervical Cancer Cells to Irradiation by Targeting RAD51. *J. Cel Physiol* 235 (1), 185–193. doi:10.1002/jcp.28957

Varghese, V., Magnani, L., Harada-Shoji, N., Mauri, F., Szydlo, R. M., Yao, S., et al. (2019). FOXM1 Modulates 5-FU Resistance in Colorectal Cancer through Regulating TYMS Expression. *Sci. Rep.* 9 (1), 1505. doi:10.1038/s41598-018-38017-0

Wang, R., Pan, W., Jin, L., Huang, W., Li, Y., Wu, D., et al. (2020). Human Papillomavirus Vaccine against Cervical Cancer: Opportunity and challenge. *Cancer Lett.* 471, 88–102. doi:10.1016/j.canlet.2019.11.039

Xie, W., Shui, C., Fang, X., Peng, Y., and Qin, L. (2020). miR-197-3p Reduces Epithelial-Mesenchymal Transition by Targeting ABCA7 in Ovarian Cancer Cells. *3 Biotech.* 10 (8), 375. doi:10.1007/s13205-020-02362-7

Xu, Y., Wu, G., Li, J., Li, J., Ruan, N., Ma, L., et al. (2020). Screening and Identification of Key Biomarkers for Bladder Cancer: A Study Based on TCGA and GEO Data. *Biomed. Res. Int.* 2020, 1–20. doi:10.1155/2020/8283401

Yang, L., Zeng, W., Sun, H., Huang, F., Yang, C., Cai, X., et al. (2020). Bioinformatical Analysis of Gene Expression Omnibus Database Associates TAF7/CCNB1, TAF7/CCNA2, and GTF2E2/CDC20 Pathways with Glioblastoma Development and Prognosis. *World Neurosurg.* 138, e492–e514. doi:10.1016/j.wneu.2020.02.159

You, M., Zhang, L., Zhang, X., Fu, Y., and Dong, X. (2021). MicroRNA-197-3p Inhibits the Osteogenic Differentiation in Osteoporosis by Down-Regulating KLF 10. *Cia* 16, 107–117. doi:10.2147/CIA.S269171

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 12 January 2022
doi: 10.3389/fgene.2021.765275

Integrated Dissection of lncRNA-miRNA-mRNA Pairs and Potential Regulatory Role of lncRNA PCAT19 in Lung Adenocarcinoma

Xiaomei Tang[1], *Xiaoyan Hua*[2], *Xujin Peng*[2], *Yongyan Pei*[3]* and *Zhigang Chen*[4]*

[1]Jiangxi Chest Hospital, Nanchang, China, [2]Department of Oncology, Wannian County Hospital of Traditional Chinese Medicine, Shangrao, China, [3]School of Medicine and Chemical Engineering, Guangdong Pharmaceutical University, Guangzhou, China, [4]Department of Oncology, Shangrao People's Hospital, Shangrao, China

Lung adenocarcinoma (LUAD) is the main cause of cancer-related deaths worldwide. Long noncoding RNAs have been reported to play an important role in various cancers due to their special functions. Therefore, identifying the lncRNAs involved in LUAD tumorigenesis and development can help improve therapeutic strategies. The TCGA-LUAD RNA expression profile was downloaded from The Cancer Genome Atlas, and a total of 49 differential lncRNAs, 112 differential miRNAs, and 2,953 differential mRNAs were screened. Through Kaplan–Meier curves, interaction networks, hub RNAs (lncRNAs, miRNAs, and mRNAs) were obtained. These hub genes are mainly involved in cell proliferation, cell cycle, lung development, and tumor-related signaling pathways. Two lncRNAs (SMIM25 and PCAT19) more significantly related to the prognosis of LUAD were screened by univariate Cox analysis, multivariate Cox analysis, and risk model analysis. The qPCR results showed that the expression levels of SMIM25 and PCAT19 were downregulated in clinical tissues, A549 and SPC-A1 cells, which were consistent with the bioinformatics analysis results. Subsequently, the PCAT19/miR-143-3p pairs were screened through the weighted gene co-expression network analysis and miRNA-lncRNA regulatory network. Dual luciferase detection confirmed that miR-143-3p directly targets PCAT19, and qPCR results indicated that the expression of the two is positively correlated. Cell function tests showed that overexpression of PCAT19 could significantly inhibit the proliferation, migration, and invasion of A549 and SPC-A1 cells. In contrast, knockout of PCAT19 can better promote the proliferation and migration of A549 and SPC-A1 cells. The expression of PCAT19 was negatively correlated with tumor grade, histological grade, and tumor mutation load in LUAD. In addition, co-transfection experiments confirmed that the miR-143-3p mimic could partially reverse the effect of PCAT19 knockout on the proliferation of A549 and SPC-A1 cells. In summary, PCAT19 is an independent prognostic factor in patients with LUAD that can regulate the proliferation, migration, and invasion of LUAD cells and may be a potential biomarker for the diagnosis of LUAD. PCAT19/miR-143-3p plays a very important regulatory role in the occurrence and development of LUAD.

Keywords: lung adenocarcinoma, lncRNA, proliferation, migration, invasion

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

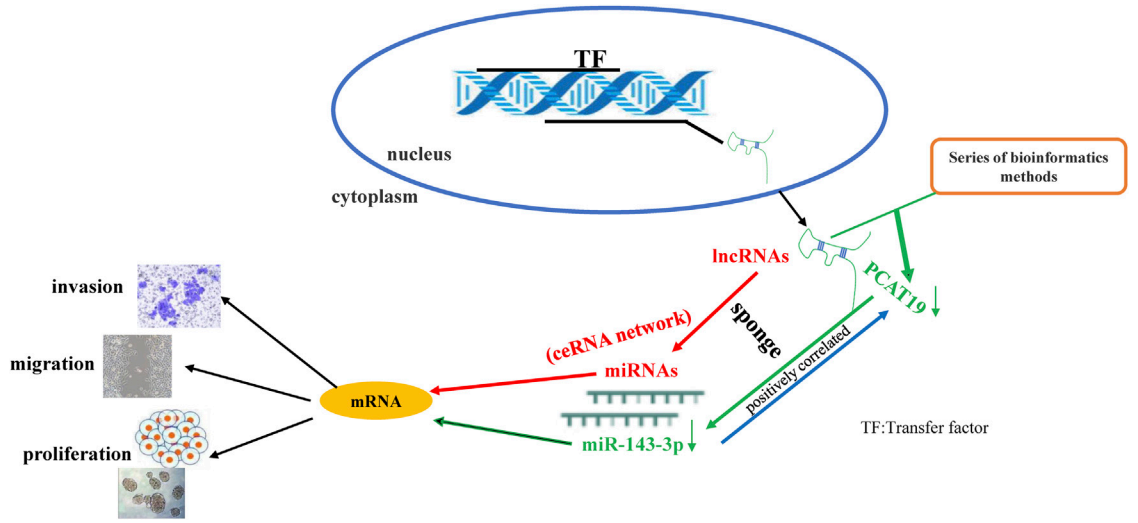
Reviewed by:
Andy T. Y. Lau,
Shantou University, China
Asaf Shilo,
National Institutes of Health (NIH),
United States

***Correspondence:**
Zhigang Chen
czg15979378686@163.com
Yongyan Pei
peiyongyan@gdpu.edu.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *26 August 2021*
Accepted: *13 December 2021*
Published: *12 January 2022*

Citation:
Tang X, Hua X, Peng X, Pei Y and Chen Z (2022) Integrated Dissection of lncRNA-miRNA-mRNA Pairs and Potential Regulatory Role of lncRNA PCAT19 in Lung Adenocarcinoma. Front. Genet. 12:765275. doi: 10.3389/fgene.2021.765275



GRAPHICAL ABSTRACT | Our data revealed that PCAT19 regulates the proliferation, migration, and invasion of lung cancer cells by targeting the miR-301b-3p. The red arrow, red line, and red font indicate that the progression of lung cancer cells can be regulated through an lncRNA/miRNA/mRNA regulatory network. The green arrow, green line, and green font represent the PCAT19/miR-143-3p/mRNA regulatory network we screened and analyzed in this project. Black arrows, black lines, and black fonts represent the common regulation process of the two situations in the figure. TF, transfer factor.

INTRODUCTION

Lung cancer is one of the most common malignant tumors in the world, of which non–small cell lung cancer (NSCLC) is the main subtype (Gandhi et al., 2018; Ferlay et al., 2019; Cai and Liu, 2021). Furthermore, lung adenocarcinoma (LUAD) is the most common NSCLC and accounts for a large proportion of cancer-related deaths worldwide (Liu et al., 2019; Siegel et al., 2019). Despite great efforts to improve the treatments and prognosis of patients with LUAD, the average 5-year survival rate is still under 20% (Kim et al., 2015; Lin et al., 2016). The leading causes of this phenomenon are cancer cell metastasis and chemoresistance (Alvarado-Luna and Morales-Espinosa, 2016; Chen et al., 2019). Therefore, an intensive study on the molecular mechanism of LUAD progression is required to identify new biomarkers and targeted therapies.

Long noncoding RNAs (lncRNAs) are noncoding RNAs that are more than 200 nucleotides in length (Peng et al., 2017). Emerging studies have confirmed that lncRNAs are abnormally expressed in the biological processes of various cancers, including breast cancer (Liang et al., 2020), lung cancer (Lai et al., 2020), esophageal cancer (Zhang et al., 2020), and gastric cancer (He et al., 2019). MicroRNAs (miRNAs) are small endogenous noncoding RNAs that play an important role in regulating gene expression, and their regulatory networks are involved in many biological processes (Ali Syeda et al., 2020). Salmena et al. (2011) proposed a competing endogenous RNA (ceRNA) hypothesis, the core concept of which is that ceRNAs interact with target miRNAs through miRNA response elements to control the transcriptome on a large scale. The cross-talk between ceRNAs is achieved by lncRNA-miRNA-mRNA networks. This delicate and complex regulatory network may contribute to a more precise understanding of the disease process (Wang et al., 2019; Zhou et al., 2019; Ma et al., 2020). However, lncRNA-miRNA-mRNA networks have not been sufficiently studied in human cancers, including LUAD. Over the past few years, an increasing number of studies have shown the regulation of lncRNAs in LUAD (Dong et al., 2018; Guo et al., 2019; Jia et al., 2019; Ying et al., 2020). These lncRNAs participate in multiple biological processes, such as cell proliferation, cell cycle, cell migration, and drug resistance by different regulatory mechanisms (Cheng et al., 2019; Zhu et al., 2019). Therefore, studies on lncRNAs may be of enormous value in understanding the occurrence and development of tumors. Although many lncRNAs have been reported in numerous tumors, still many lncRNAs are to be investigated.

In this study, we aimed to identify new therapeutic targets and prognostic biomarkers for LUAD. The RNA expression profile of LUAD in The Cancer Genome Atlas (TCGA) was analyzed to identify the differentially expressed lncRNAs, miRNAs, and mRNAs, and Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), Kaplan–Meier, Cox regression, gene set enrichment analysis (GSEA), ceRNA, protein–protein interaction (PPI), and other bioinformatics analyses were used to determine the functions of differentially expressed hub lncRNAs. At the same time, we detected the expression of hub lncRNAs in LUAD tissues and A549 cells and explored the effects of hub lncRNAs and their coaction on A549 cells through functional loss and gain experiments, dual luciferase experiments, and co-transfection experiments.

MATERIALS AND METHODS

Screening of Differentially Expressed RNAs

The RNA expression profile of LUAD was derived from TCGA, containing sequencing data of 585 samples (**Table 1**). Clinical information (age, race, survival time, survival status, tumor grade,

TABLE 1 | Clinical pathological characteristics of 585 patients with lung adenocarcinoma.

Parameter	Subtype	Patients (%)
Age (years)	33 to ≤ 66	287 (49.06%)
	66 to ≤ 89	298 (50.94%)
Gender	Male	316 (54.02%)
	Female	269 (45.98%)
Pathologic stage	stage i	316 (54.02%)
	stage ii	135 (23.08%)
	stage iii	97 (16.58%)
	stage iv	37 (6.32%)
Pathologic T	T1	191 (32.65%)
	T2	321 (54.87%)
	T3	50 (8.55%)
	T4	23 (3.93%)
Pathologic M	M0	424 (72.48%)
	M1	161 (27.52%)
Pathologic N	N0	376 (64.27%)
	N1	113 (19.32%)
	N2	91 (15.56%)
	N3	5 (0.85%)
Vital status	Alive	365 (62.39%)
	Dead	220 (37.61%)

and histological grade) of all samples was also downloaded. The limma package of R was used to analyze the differential expression profile. The screening criteria were $p < 0.01$, false discovery rates (FDR) < 0.05, and $|\log_2$ fold change $|>1$.

Kaplan–Meier Survival Analysis of lncRNAs

To explore whether the expression level of differentially expressed lncRNAs has an effect on the prognostic survival of patients with LUAD, GEPIA 2.0 (http://gepia2.cancer-pku.cn) was used to perform Kaplan–Meier survival analysis on all differentially expressed lncRNAs. $p < 0.05$ was judged as statistically significant. Then, Starbase 2.0 (http://starbase.sysu.edu.cn/) software was used to verify the data obtained from the analysis.

Constructing the lncRNA-miRNA-mRNA ceRNA Network

To construct the ceRNA network, we used Starbase 2.0 software to verify the regulatory correlation between miRNA-mRNA and miRNA-lncRNA. Then, the differentially expressed RNAs (lncRNAs, miRNAs, and mRNAs) were compared to obtain overlapping RNAs (lncRNAs, miRNAs, and mRNAs). The network was constructed and observed using Cytoscape.

Protein–Protein Interaction Network Construction of mRNAs in the ceRNA Network

To identify relevant hub genes, the search tool for the Retrieval of Interacting Genes/Proteins (STRING 11.5, https://string-db.org/) was utilized. The PPI was analyzed using the STRING database, and a combined score>0.4 was used as the cutoff criterion. Additionally, Cytoscape 3.9 was used to construct and visualize the PPI network.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Analysis of mRNAs in the ceRNA Network

The functional enrichment analysis of mRNAs in the ceRNA network was performed using Metascape (http://metascape.org/). All statistically enriched terms (GO and KEGG) were identified based on accumulative hypergeometric *p*-values.

Cox Proportional Regression Model Based on the Expression of Hub lncRNAs

To analyze the independent influence of individual lncRNAs on the overall survival of patients with LUAD, we used an online tool (Sangerbox tools 3.0, http://sangerbox.com/Tool) to perform univariate and multivariate Cox proportional regression analysis. Based on the analysis results, we established a risk regression model to further verify the clinical value of lncRNAs. The risk model formula is as follows: a × exp (lncRNA1) + a × exp (lncRNA2) + ... + a × exp (lncRNAn), where a represents the multivariate Cox regression coefficient and exp () represents the expression level of lncRNA. Then, based on the risk value, the patients were divided into high- and low-risk groups, the risk curves of the two groups were calculated, and their 1-, 3-, and 5-year ROC (receiver operating characteristic) curves were drawn to test the predictive ability of the model.

Weighted Gene Co-Expression Network Analysis of miRNAs in GSE74190

The LUAD dataset GSE74190 in the GEO database was used to perform the weighted gene co-expression network analysis (WGCNA). We first remove sample outliers based on the thresholding power value and built a sample tree. The Pearson correlation coefficient was used to calculate the correlation between gene expression, and the correlation matrix of gene expression data was constructed. Then, the matrix was converted to a connection matrix to calculate the topological overlap matrix (TOM). Subsequently, gene modules were established, the minimum module size of each module was set to 30, and the correlation between the module and the trait was calculated based on the data of each module. Finally, TOM-based dissimilarity was used for cluster analysis and the Pearson correlation coefficient was used to analyze the relationship between LUAD in each module and normal tissues.

Gene Set Enrichment Analysis

We used the standardized expression profile data obtained from the LUAD dataset in TCGA for GSEA. The number of permutations was set to 1,000. Using GSEA, we analyzed the GO and KEGG pathways to study the possible biological

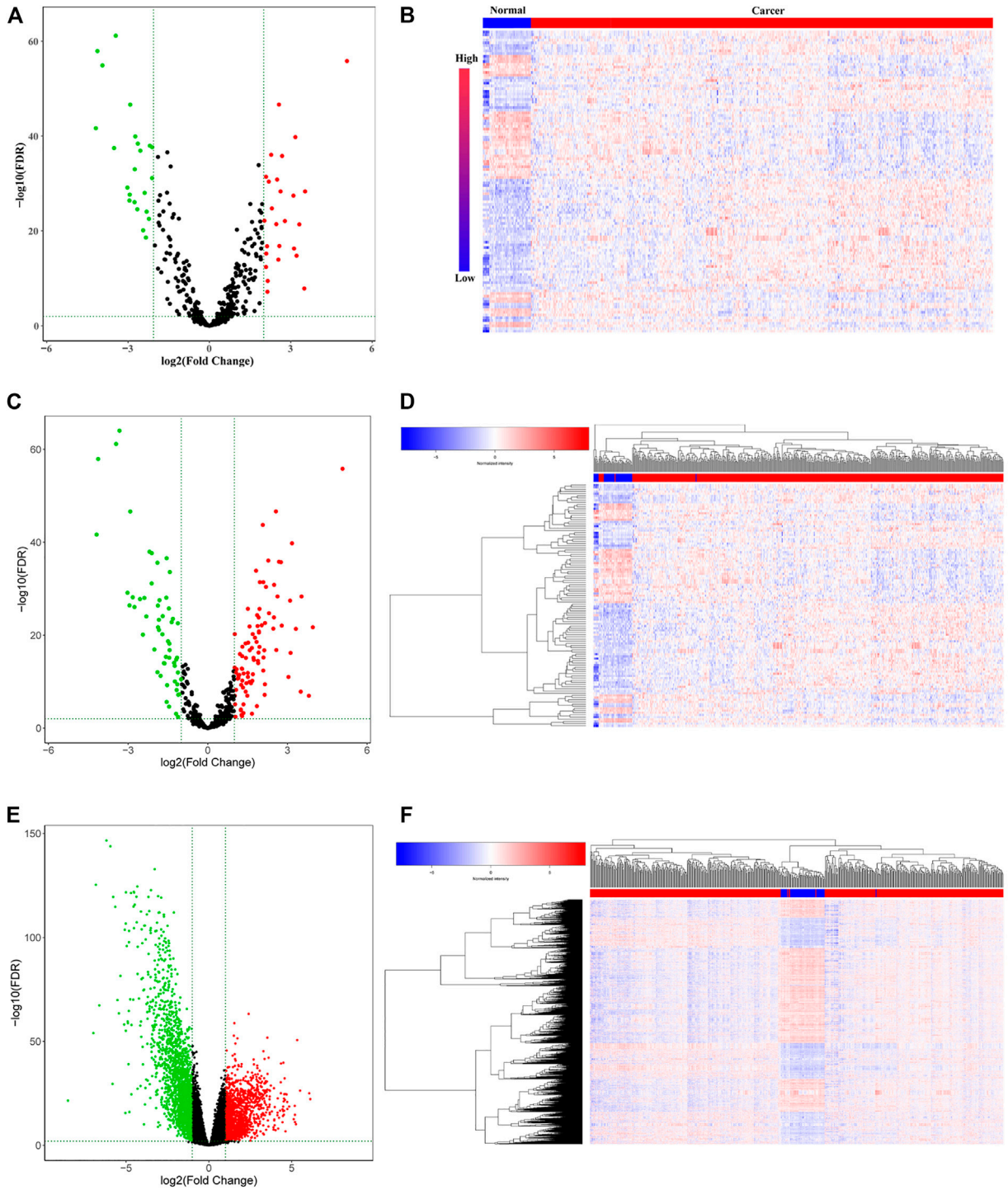


FIGURE 1 | Differentially expressed RNAs (lncRNAs, miRNAs, and mRNAs) in the RNA expression profile of lung adenocarcinoma in TCGA. **(A)** Volcano plot and **(B)** heatmap for the differentially expressed lncRNAs. **(C)** Volcano plot and **(D)** heatmap for the differentially expressed miRNAs. **(E)** Volcano plot and **(F)** heatmap for the differentially expressed miRNAs. The red dots represent upregulated genes, and the green dots represent downregulated genes. The black dots represent genes with no significant difference.

functions of PCAT19. $p < 0.05$ indicates that the analysis result is meaningful.

Cell Lines

A549, SPC-A1, and 16HBE cells used in this study were acquired from the Institute of Beina Chuanglian Biotechnology Research, Beijing, China. A549 and 16HBE cells were cultured in RPMI-1640 medium (Invitrogen, United States), and 10% fetal bovine serum (FBS, Gibco, United States) was used to supplement the medium with 100 mg ml^{-1} streptomycin and 100 U ml^{-1} penicillin. SPC-A1 cells were cultured in Dulbecco-modified Eagle medium (Invitrogen, United States), and 10% fetal bovine serum (FBS, Gibco, United States) was used to supplement the medium with 100 mg ml^{-1} streptomycin and 100 U ml^{-1} penicillin. The medium was kept in incubators at 37°C with 5% CO_2.

RNA Extraction and qRT–PCR

A total of 30 LUAD and 30 adjacent tissues were collected (**Supplementary Material S1**). This study was approved by the hospital ethics committee and was carried out in accordance with the Declaration of Helsinki. In addition, each patient provided written informed consent. All the tissues were stored at 80°C immediately after surgical resection.

Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Germany) according to the manufacturer's instructions. The concentration of the extracted RNA was measured with Nanodrop. The same amount of RNA was reverse transcribed into cDNA using the Golden star™ RT6 cDNA synthesis kit (TsingKe, Beijing). Then, the Master qPCR Mix (SYBR GREEN 1) (Qingdao, Beijing) was used to perform qRT-PCR to detect the expression of RNAs. The expression of GAPDH was used as a reference for lncRNA, and U6 was used as a reference for miRNA. The formula $2^{-\Delta\Delta Ct}$ was used to calculate the expression levels of lncRNAs and miRNAs. Primer 5.0 software (SMIM25: Forward (5′-3′)-TCTCTGGGTGGAATGTCAC, Reverse (5′–3′)-TTTACTGGG CACTTGTCCT; PCAT19: Forward (5′-3′)-CCAATGACATCC AATGGAGG, Reverse (5′–3′)-TCCTGGTGGTTGTTTAATCAC) was used for the lncRNA primer design, and miRNA primers (miR-143-3p: MQPS0000635-1-100) were purchased from Guangzhou RiboBio. In addition, the qRT-PCR data were analyzed by Excel. The two groups were tested by *t*-test ($p < 0.05$).

Cell Transfection

Si-PCAT19 and the corresponding siRNA control (si-NC) were purchased from Guangzhou RiboBio. The PCAT19 overexpression plasmid (pCDH-GFP-PCAT19) and the corresponding control plasmid (NC) were also purchased from Guangzhou RiboBio. All plasmids were tested according to the manufacturer's instructions and were transfected into lung cancer cells (A549 and SPC-A1 cells) using the Lipofectamine 3000 reagent (Invitrogen, United States).

Cell Counting Kit-8 Assay

The proliferation of lung cancer cells (A549 and SPC-A1 cells) was evaluated by the cell counting kit-8 (CCK-8) assay (Beyotime, Beijing) according to the manufacturer's instructions. Lung cancer cells were collected and then seeded in 96 wells at 1×10^4 cells/well with culture medium at 37°C with 5% CO_2. Cells were then subjected to culture for 24 and 48 h before the addition of 10 μl of CCK-8 (5 mg/ml) to each well. After 3 h of incubation, a multimode reader (Thermo, United States) was utilized to detect the absorbance at 450 nm of each well. Each experiment was carried out three times.

TABLE 2 | Prognosis values of differentially expressed long noncoding RNAs in lung adenocarcinoma.

lncRNA	log_2 FC	P Adj	Overall survival (*p*-value)
FAM83A-AS1	5.33	5.31E-53	3.90E-05[a]
AFAP1-AS1	5.25	1.91E-11	5.60E-01
Z98257.1	4.73	1.63E-22	6.00E-01
FEZF1-AS1	4.54	3.36E-24	3.70E-01
MNX1-AS1	4.36	1.52E-41	1.30E-01
ZFPM2-AS1	4.35	6.81E-27	9.60E-02
AL391056.1	3.39	4.72E-27	4.20E-01
LINC00511	3.05	1.41E-26	6.30E-02
AC025580.1	3.03	3.97E-21	2.90E-01
AL354719.2	3.01	1.18E-20	3.10E-01
AL365181.3	2.96	1.76E-11	1.20E-02[a]
LINC01270	2.54	6.39E-31	7.00E-01
LUCAT1	2.52	1.92E-12	5.70E-01
AC004816.1	2.52	1.12E-34	3.30E-01
VPS9D1-AS1	2.50	6.26E-24	6.80E-03[a]
BLACAT1	2.48	6.84E-16	2.20E-01
U62317.2	2.47	1.18E-25	3.20E-01
PVT1	2.44	5.19E-28	6.10E-01
AL354707.1	2.38	8.11E-19	8.70E-01
PCAT6	2.35	2.83E-26	4.50E-01
LINC01426	2.26	4.86E-18	3.80E-01
LINC02362	2.16	3.51E-17	6.80E-02
AL157838.1	2.10	4.93E-28	7.50E-02
AL355312.3	2.04	6.22E-13	5.20E-01
LINC00857	2.04	3.84E-26	6.10E-02
AL445524.1	2.01	2.63E-16	1.20E-01
AC011511.5	−2.02	4.65E-24	3.70E-01
AC144831.1	−2.06	1.21E-42	5.80E-01
TBX5-AS1	−2.06	5.63E-47	1.70E-02[a]
AF131215.5	−2.18	2.96E-38	8.10E-03
MBNL1-AS1	−2.24	2.70E-69	5.60E-02
AC125807.2	−2.27	4.84E-72	3.20E-01
AC010329.1	−2.35	6.31E-16	8.00E-04[a]
AC116407.1	−2.53	6.69E-52	5.30E-02
AL162511.1	−2.55	8.17E-22	2.50E-03[a]
LINC00261	−2.60	2.46E-19	3.70E-02[a]
AC011899.2	−2.62	5.48E-81	1.60E-01
PCAT19	−2.70	1.52E-101	1.60E-04[a]
SMIM25	−2.76	6.05E-79	3.30E-02[a]
RHOXF1-AS1	−2.77	4.41E-34	3.80E-02[a]
HHIP-AS1	−2.78	1.36E-33	8.40E-02
LINC01936	−2.83073824	1.72E-63	1.40E-01
C8orf34-AS1	−2.83225284	6.53E-25	2.80E-02[a]
SFTA1P	−3.231128408	2.64E-42	1.90E-01
AC079630.1	−3.496028269	5.49E-51	1.70E-02[a]
AC093110.1	−3.601261229	2.22E-116	2.7E-01
LHFPL3-AS2	−4.083740091	1.30E-54	4.80E-03[a]
FENDRR	−4.550706947	1.41E-118	2.60E-03[a]
AC008268.1	−5.015890223	1.37E-57	2.00E-01

[a]*Represents prognostic lncRNAs.*

Wound-Healing Assay

Lung cancer cells (A549 and SPC-A1 cells) were collected and then seeded in 96 wells at 5×10^5 cells/well with culture medium at 37°C with 5% CO_2. Cells were cultured to a confluent monolayer, which was scratched by a sterile pipette tip (200 μl). Then, the cells were flushed twice with PBS and subjected to reduced serum medium, and photos were taken after 0, 12, 24, and 48 h. The scratched areas were recorded by a microscope (Olympus, Japan), and ImageJ software was utilized to evaluate the percentage of closure.

Transwell Assay

After 48 h of transfection, lung cancer cells (A549 and SPC-A1 cells) were collected to prepare a single cell suspension. The lung cancer cell suspension (3×10^3 cells/well) was added to the upper chamber (Corning, United States), and the medium (20% fetal bovine serum) was added to the lower chamber. Matrigel was added to the upper chamber. After 24 h, the cells were fixed with 4% paraformaldehyde (Sigma, United States) and stained with 1% crystal violet (Sigma, United States). Cell observation and counting were performed under a light microscope (Olympus, Japan).

Soft Agar Colony Formation Assay

A549 and SPC-A1 cells were seeded on a six-well culture plate at 3×10^5 cells per well, and the plasmid was transfected at 70% cell confluence. After 24 h of transfection, the cells of each group were collected for viable cell count, and the cell density was adjusted to 2000 cells/ml with medium containing 20% FBS and then maintained in a medium containing 10% FBS, 100 μg/ml streptomycin, 100 units/ml penicillin, 0.7% soft agar, and 1.2% soft agar. The medium was incubated at 37°C with 5% CO_2 in an incubator. After 14 days, the colonies were fixed with 4% paraformaldehyde at room temperature and then stained with 0.005% crystal violet. The number of clones larger than 0.05 mm and the number of total clones in the field of view were manually counted under the microscope: the clone formation rate = the number of clones larger than 0.05 mm/the number of all clones × 100%.

Luciferase Reporter Assay

A549 cells were seeded into 96-well plates and were co-transfected with 60 ng of the dual luciferase reporter construct pmiR-RB-Report™ -PCAT19 3′UTR (RiboBio, Guangdong) and miR-143-3p mimic. After 48 h of incubation, the fluorescence value was measured by the dual luciferase reporter kit (Promega, United States).

Statistical Analysis

GraphPad Prism 8 and SPSS 25.0 (California, CA) were utilized to perform statistical analysis. The differences were deemed statistically significant at $p < 0.05$.

RESULT

Aberrantly Expressed lncRNAs in Lung Adenocarcinoma

To identify the hub lncRNAs involved in the progression of LUAD, we downloaded the LUAD expression profile data from TCGA and performed differential analysis. A total of 26 upregulated lncRNAs and 23 downregulated lncRNAs were identified (**Figures 1A,B**; **Table 2**). Then, the screened differentially expressed lncRNAs were analyzed comprehensively. The Kaplan–Meier curve analysis revealed 14 lncRNAs associated with the prognosis of LUAD patients (**Supplementary Figure S1**; **Table 2**), indicating that these lncRNAs may play an important role in the occurrence of LUAD. However, the function of most lncRNAs in LUAD requires further study.

Construction of the lncRNA-miRNA-mRNA Network in Lung Adenocarcinoma

The limma package was used to analyze the expression profile data of LUAD from TCGA and to analyze the differentially expressed miRNAs and mRNAs according to the adjusted $p < 0.01$, FDR<0.05, and $|\log_2$ fold change $|>1$ screening criteria. A total of 112 differentially expressed miRNAs (70 upregulated and 42 downregulated) (**Figures 1C,D**; **Supplementary Table S1**) and 2,953 differentially expressed mRNAs (1,097 upregulated and 1,856 downregulated) (**Figures 1E,F**; **Supplementary Table S2**) were identified. Then, the Starbase tool was used to analyze the targeted miRNAs of the screened prognostic lncRNAs and to identify the intersection with the differentially expressed miRNAs. Seven common miRNAs were obtained (**Figure 2A**). Similarly, the Starbase tool was used to analyze the targeted mRNAs of seven common miRNAs and identify the intersection with the differentially expressed mRNAs. Twenty-seven common mRNAs were obtained (**Figure 2B**). The screened lncRNAs, miRNAs, and mRNAs were used to construct a ceRNA co-expression network, which consisted of 40 nodes and 80 edges (**Figure 2C**; **Table 3**).

PPI Network Construction and Functional Enrichment Analysis of lncRNA-Related mRNAs in Lung Adenocarcinoma

The co-expression analysis was carried out to construct PPI networks to disclose the potential roles and mechanisms of hub lncRNAs in LUAD. Twenty-seven DEmRNAs in the ceRNA network were considered to be research objects. The co-expression analysis of lncRNA-mRNAs was performed using Starbase and GEPIA. LncRNA-mRNAs pairs with $p < 0.05$ were considered reliable (**Supplementary Table S3**). Then, the STRING database was used to construct a PPI network with lncRNA-associated mRNAs. In total, 38 edges and 13 nodes were involved in the PPI network (**Figure 2D**). One significant module containing 7 edges and 15 nodes was selected from the PPI network using the MCODE plug-in in Cytoscape (**Figure 2E**).

GO and KEGG analyses of mRNAs in the PPI network were performed using Metascape. The results showed that these lncRNA-associated mRNAs were obviously associated with cell senescence, cell division, cell cycle, angiogenesis,

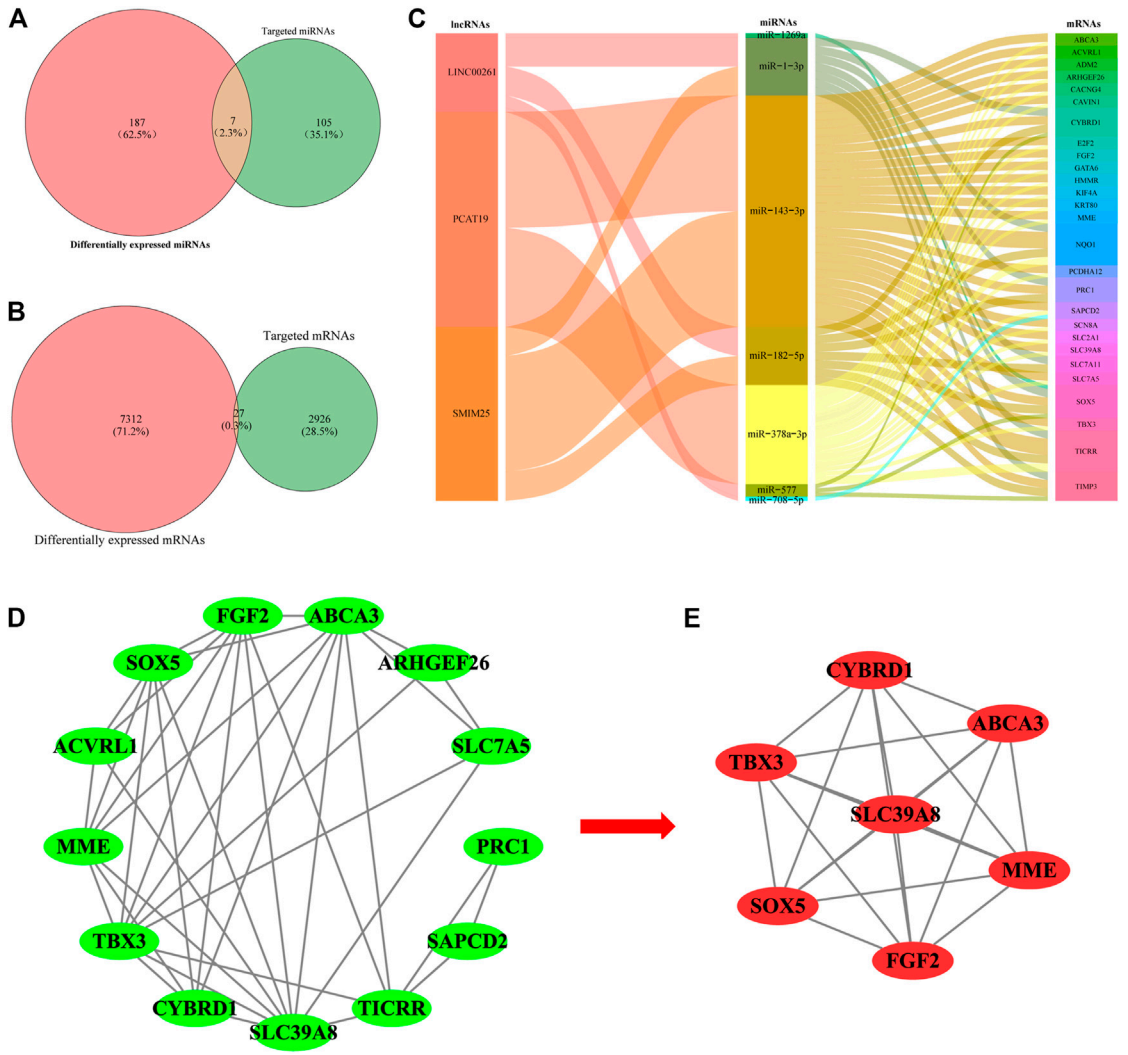


FIGURE 2 | Venn diagram of the common **(A)** miRNAs and **(B)** mRNAs between differentially expressed miRNAs/mRNAs and target miRNAs/mRNAs. **(C)** Sankey diagram of the ceRNA network in lung adenocarcinoma. Each rectangle represents a gene, and the degree of connectivity of each gene is indicated by the size of the rectangle. **(D)** PPI network of the lncRNA-correlated mRNAs in the ceRNA network. **(E)** Most significant module in the PPI network of the lncRNA-correlated mRNAs by the MCODE plug-in.

TABLE 3 | Key lncRNAs, miRNAs, and mRNAs in the ceRNA network of lung adenocarcinoma.

lncRNAs	Binding miRNAs	Associated mRNAs
SMIM25	hsa-miR-182-5p, hsa-miR-1-3p, hsa-miR-143-3p	*TICRR*, *SLC7A11*, *NQ O 1*, *PRC1*, *CYBRD1*, *SOX5*, *TIMP3*
PCAT19	hsa-miR-378a-3p, hsa-miR-143-3p	*KIF4A*, *SAPCD2*, *ADM2*, *SLC2A1*, *HMMR*, *TICRR*, *NQ O 1*, *SCN8A*, *SLC7A5*, *E2F2*, *KRT80*, *PCDHA12*, *CAVIN1*, *GATA6*, *TBX3*, *CACNG4*, *FGF2*, *ACVRL1*, *SLC39A8*, *ABCA3*, *ARHGEF26*, *MME*
LINC00261	hsa-miR-577, hsa-miR-1269a, hsa-miR-708-5p, hsa-miR-182-5p, hsa-miR-1-3p	*SLC7A11*

transmembrane signal transduction, and lung development (**Figure 3A**). For KEGG pathways, pathways in cancer were the most significantly enriched (**Figure 3B**). All identified processes and pathways of lncRNA-associated mRNAs were connected. To further explore the relationships, a network plot of enriched terms was selected and shown (**Figures 3C,D**). The enriched items were

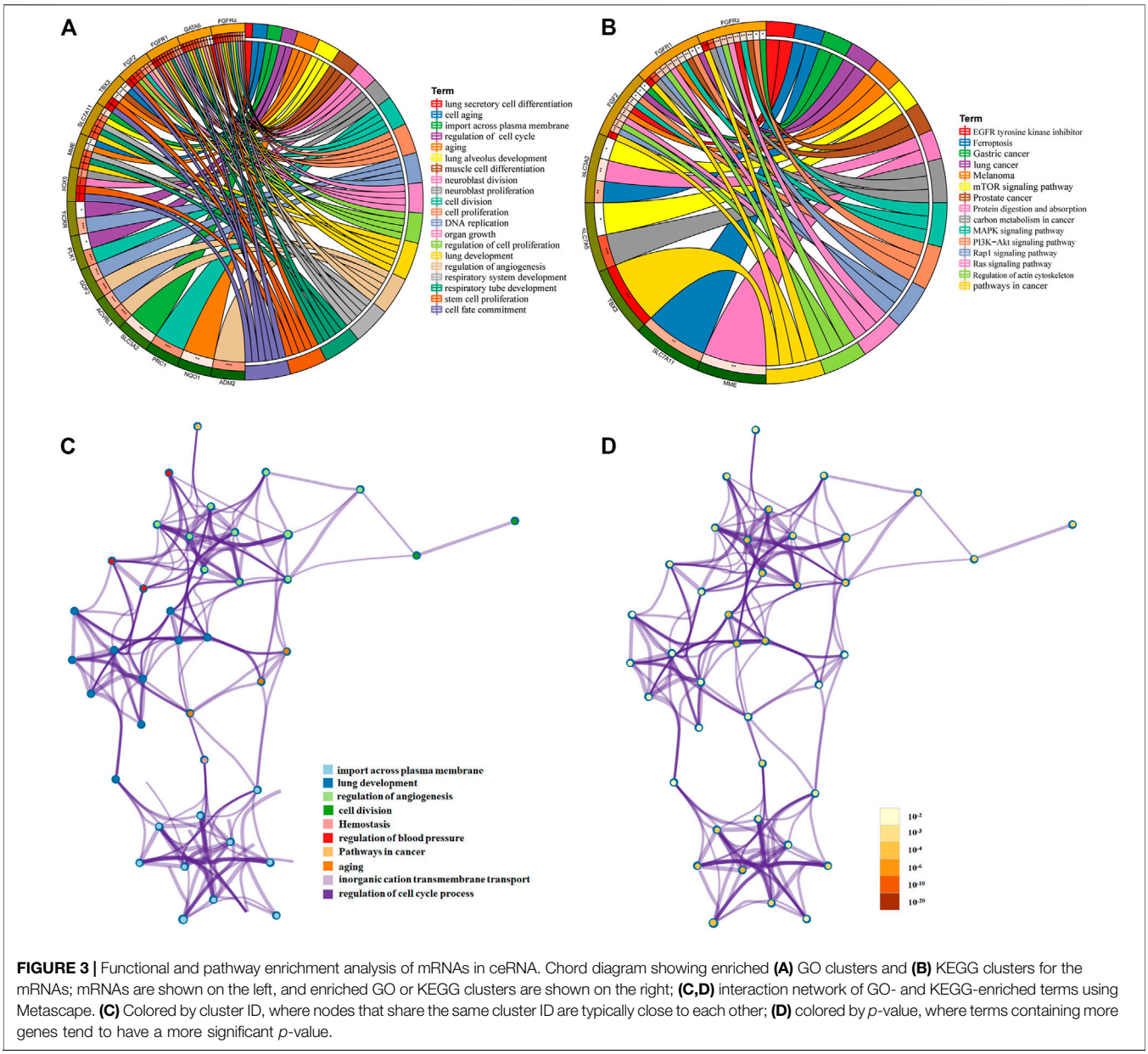


FIGURE 3 | Functional and pathway enrichment analysis of mRNAs in ceRNA. Chord diagram showing enriched **(A)** GO clusters and **(B)** KEGG clusters for the mRNAs; mRNAs are shown on the left, and enriched GO or KEGG clusters are shown on the right; **(C,D)** interaction network of GO- and KEGG-enriched terms using Metascape. **(C)** Colored by cluster ID, where nodes that share the same cluster ID are typically close to each other; **(D)** colored by *p*-value, where terms containing more genes tend to have a more significant *p*-value.

distributed in concentration, and the interaction between the enriched items was conspicuous.

Cox Analysis of lncRNAs and Validation of Hub lncRNA Expression in Clinical Samples of Lung Adenocarcinoma and Lung Cancer Cells (A549 and SPC-A1 Cells)

To further identify the lncRNAs related to the prognosis of LUAD patients, we conducted the univariable Cox regression analysis. As a result, we found that two lncRNAs (SMIM25 and PCAT19) were more significantly related to the prognosis of patients with LUAD (**Figure 4A**). Subsequently, the multivariate Cox analysis showed that SMIM25 and PCAT19 may be independent prognostic factors for LUAD (**Table 4B**). Then, we constructed a risk scoring model based on the expression levels of SMIM25 and PCAT19 in LUAD expression profiles from TCGA. The risk score calculation formula is as follows: risk score = (0.0156) × Exp (SMIM25) + (−0.068) × Exp (PCAT19). Based on the risk score, patients with LUAD were divided into high- and low-risk groups. The risk curve, scatter plot, and Kaplan–Meier analysis showed that the overall survival (OS) of the high-risk group was poor (**Figures 4B–E**). The ROC curve analysis showed that the AUCs at 1, 3, and 5 years were 0.61, 0.66, and 0.72, respectively (**Figure 4F**). All results show that the risk scoring model based on the expression of SMIM25 and PCAT19 has a

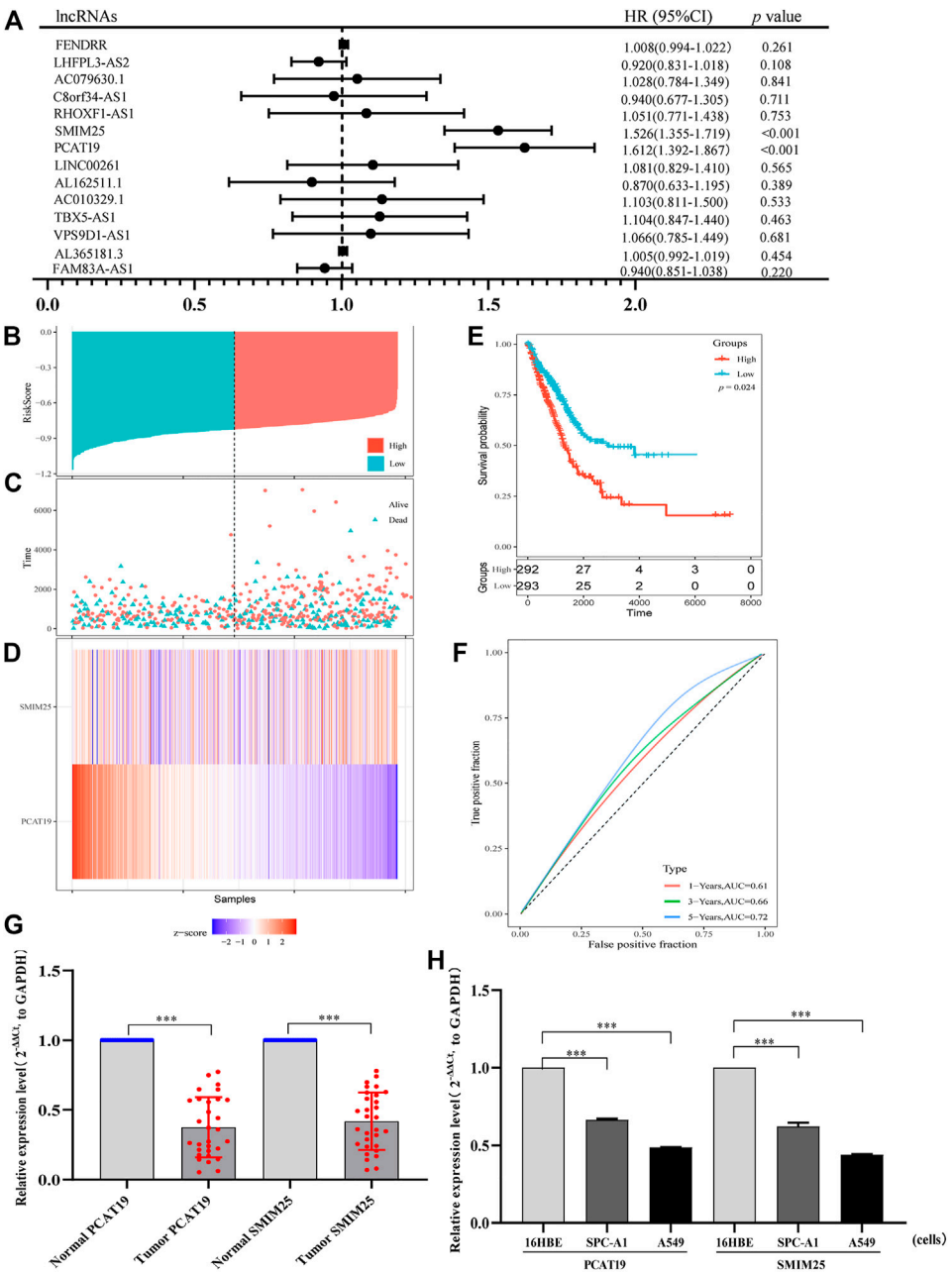


FIGURE 4 | (A) Univariate Cox proportional hazard regression analysis of 14 lncRNAs related to the prognosis of patients with lung adenocarcinoma. **(B)** Risk score distribution, **(C)** patient survival status distribution, and **(D)** heatmap of PCAT19 and SMIM25 expression by the risk score. **(E)** Kaplan–Meier curves for high-risk and low-risk groups. **(F)** ROC curves for predicting the survival in lung adenocarcinoma patients by the risk score. **(G,H)** RT-qPCR analysis was conducted to assess the expression levels of PCAT19 and SMIM25 in **(G)** lung adenocarcinoma tissues (the control was advanced normal tissues) and **(H)** A549 and SPC-A1 cells (the control was 16HBE cells). The expression level of GAPDH was used as an internal reference, and the data are shown as the mean ± standard deviation. ***$p < 0.001$.

predictive value. Therefore, the above results confirm that SMIM25 and PCAT19 might be useful candidates for LUAD patient outcome prediction.

Furthermore, we validated two prognostic lncRNAs (SMIM25 and PCAT19) in LUAD clinical tissues (normal adjacent tissues as the control) and cell line (16HBE as the control). Thirty LUAD and 30 adjacent normal tissues were detected by using qRT-PCR. As depicted in **Figure 4G**, SMIM25 and PCAT19 transcript levels were significantly lower in LUAD than normal tissues ($p < 0.05$). The same results were shown in A549 and SPC-A1 cells ($p < 0.05$) (**Figure 4H**). The expression difference of PCAT19 in the

TABLE 4 | (A) Correlation between overall survival and multivariable characteristics *via* multivariate survival analysis of PCAT19 and **(B)** multivariate survival analysis of SMIM25 and PCAT19.

Clinical characteristics	HR	Lower	Upper	*P*
A				
PCAT19	0.920	0.831	0.918	0.011
Age	1.008	0.994	1.022	0.261
M	0.940	0.677	1.305	0.711
N	1.266	1.048	1.529	0.014
T	1.236	1.035	1.477	0.019
Gender	1.028	0.784	1.349	0.841
Stage	1.328	1.115	1.583	0.002
LncRNAs	**HR**	**Lower**	**Upper**	***P***
B				
SMIM25	1.032	0.937	1.382	0.027
PCAT19	1.124	0.729	1.029	0.011

two lncRNAs was more significant; therefore, PCAT19 was selected as the follow-up research object.

Correlation Analysis of PCAT19 Expression Level and Clinicopathological Parameters

We evaluated the relationship between the expression level of PCAT19 and various clinicopathological parameters in patients with LUAD using data from TCGA. The low expression of PCAT19 was significantly correlated with the tumor histological grade (MNT), age and tumor stage, showing a significant negative correlation (**Figure 5A**, $p < 0.05$). Then, the multivariate logistic analysis using the above three parameters demonstrated that the expression level of PCAT19 (HR = 0.920, 95% CI: 0.831–1.018, $p = 0.011$), tumor histological grade (NT) (HR = 1.266, 95% CI: 1.048–1.529, $p = 0.014$, and HR = 1.236, 95% CI: 1.035–1.477, $p = 0.019$), and tumor stage (HR = 1.328, 95% CI: 1.115–1.583, $p = 0.002$) were extremely significantly associated with the overall survival of LUAD patients (**Table 4A**). In addition, we analyzed the relationship among PCAT19 expression, immune neoantigens, tumor mutational burden, and microsatellite instability in LUAD. The results showed that PCAT19 expression and tumor mutation burden were also significantly negatively correlated, indicating that the lower the PCAT19 expression in LUAD tumor cells, the higher the number of mutations contained (**Figure 5B**, $p < 0.05$). These results suggested that LUAD patients with low expression of PCAT19 were more likely to develop more advanced tumors than those with high expression of PCAT19.

In addition, we conducted a gene set enrichment analysis (GSEA) on PCAT19 to further explore the biological functions in which PCAT19 may participate. As shown in **Figures 5C–F**, the biological processes and molecular functions closely related to PCAT19 were mainly pneumotocele and lung development, cell circulation, and DNA repair. The KEGG pathway analysis showed that four pathways had the strongest positive correlation with PCAT19 expression: DNA repair, cell cycle, pentose phosphate pathway, and RNA polymerase. The three most negatively correlated pathways were the MAPK signaling pathway, JAK signaling pathway, and cell adsorption process.

Effects of PCAT19 on the Proliferation, Migration, and Invasion of Lung Cancer Cells (A549 and SPC-A1 Cells)

To further study the role and effect of PCAT19 in A549 and SPC-A1 cells, we transfected cells with PCAT19 overexpression plasmid and siRNA to perform functional gain experiments. Fluorescence microscopy analysis of PCAT19 confirmed successful transfection of the overexpression plasmid (**Figure 6A**). Subsequently, the expression level of PCAT19 in cells was detected by qPCR. PCAT19 expression was significantly increased in cells transfected with PCAT19 overexpression plasmid. The cells transfected with si-PCAT19 showed an opposite trend (**Figure 6B**). Then, we conducted CCK-8, soft agar, wound healing, and transwell assays. CCK-8 detection results showed that PCAT19 knockdown significantly promoted the proliferation of A549 and SPC-A1 cells, and overexpression of PCAT19 inhibited the proliferation of A549 and SPC-A1 cells (**Figures 6C,D**). The soft agar colony formation assay showed the same trend. Overexpression of PCAT19 inhibited the formation of clones, and the number of clones in the si-PCAT19 group increased significantly (**Figures 6E–G**). In addition, wound healing and transwell assays showed that si-PCAT19 significantly promoted A549 and SPC-A1 cell migration and invasion compared with the control group ($p < 0.05$) (**Figures 6H–L**). In addition, the PCAT19 overexpression group had a poor migration distance and invasion rate than the control group. Similarly, the PCAT19 knockout group showed a greater distance and invasion rate than the control group (**Figures 6I,L**). These results further showed that downregulation of PCAT19 significantly promoted the migration and invasion of A549 and SPC-A1 cells, and overexpression of PCAT19 decreased the migration and invasion of A549 and SPC-A1 cells. Collectively, these results suggested that PCAT19 regulates the proliferation, migration, and invasion of A549 and SPC-A1 cells.

MiR-143-3p Reverses the Effect of PCAT19 on the Proliferation of A549 and SPC-A1 Cells

To further study the miRNAs targeted by PCAT19, we used WGCNA to construct a co-expression network to identify hub miRNAs related to LUAD. First, we downloaded the LUAD miRNA expression profile dataset GSE74190 from the GEO database and performed WGCNA based on these data. When soft-power β was set to 2, the scale-free topology was suitable for indices exceeding 0.85 (**Figure 7A**). β = 2 was used to construct a hierarchical clustering tree with different colors representing different modules in LUAD samples, and nine gene modules were obtained (**Figures 7B,C**). As shown in **Figure 7D**, the turquoise module was considered to be the hub module with a

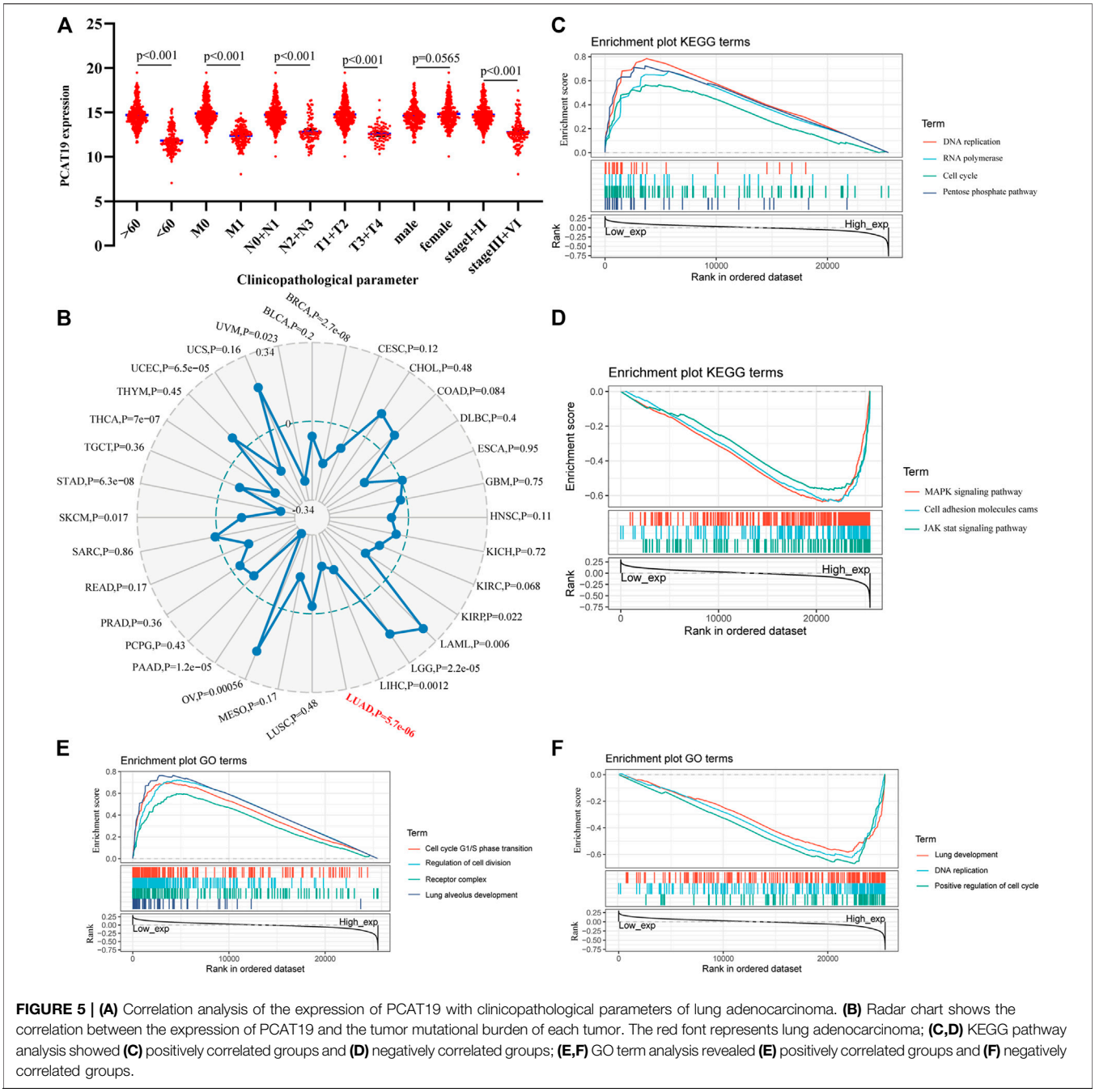


FIGURE 5 | (A) Correlation analysis of the expression of PCAT19 with clinicopathological parameters of lung adenocarcinoma. **(B)** Radar chart shows the correlation between the expression of PCAT19 and the tumor mutational burden of each tumor. The red font represents lung adenocarcinoma; **(C,D)** KEGG pathway analysis showed **(C)** positively correlated groups and **(D)** negatively correlated groups; **(E,F)** GO term analysis revealed **(E)** positively correlated groups and **(F)** negatively correlated groups.

correlation coefficient of 0.81 (p = 2e-17). Twelve hub miRNAs were further screened from the turquoise module and then intersected with the miRNAs targeted by PCAT19 in the ceRNA network. Finally, miR-143-3p was obtained (**Supplementary Table S4**). The dual luciferase experiment further verified that PCAT19 and miR-143-3p are directly targeted (**Figures 8A,B**). Correlation analysis results showed that PCAT19 and miR-143-3p were positively correlated (**Figure 8C**). The qPCR results showed that the expression level of miR-143-3p was increased in A549 and SPC-A1 cells overexpressing of PCAT19. Conversely, the expression level of miR-143-3p in A549 and SPC-A1 cells with PCAT19 knockout was also decreased (**Figure 8D**). In addition, co-transfected with si-PCAT19 and miR-143-3p mimic in A549 and SPC-A1 cells, the CCK-8 results showed that the proliferation activity of A549 and SPC-A1 cells was inhibited compared with that of the PCAT19 knockout group (**Figure 8E**), indicating that the miR-143-3p mimic might reverse the proliferative effect of si-PCAT19 on A549 and SPC-A1 cells.

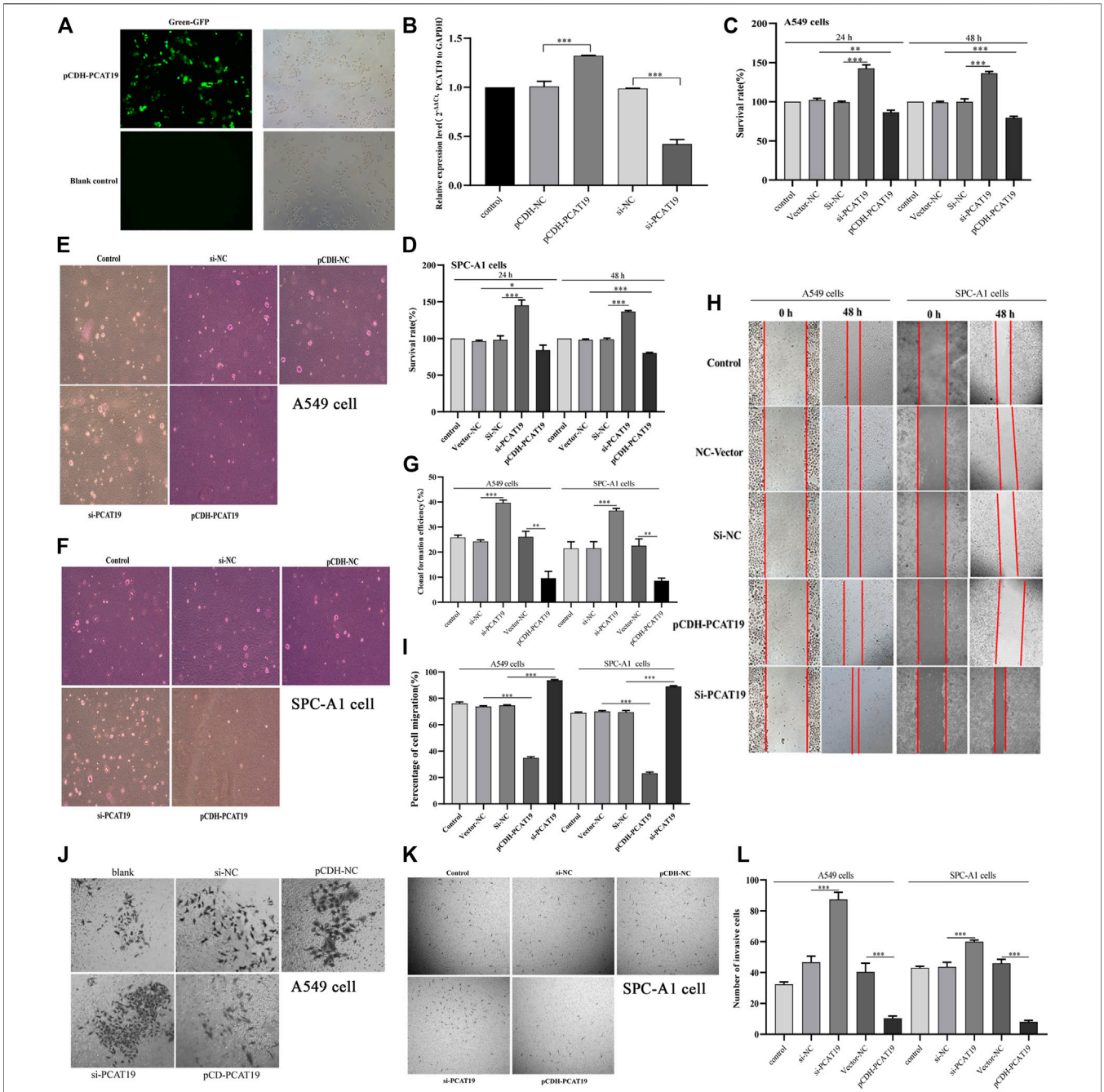


FIGURE 6 | (A) Overexpression of PCAT19 in A549 cells analyzed by fluorescence observation; **(B)** transfection efficiency of pCDH-PCAT19 and si-PCAT19 was detected by qPCR; **(C)** effects of PCAT19 knockdown or overexpression on the vitality of A549 cells measured using the CCK-8 assay; **(D)** effects of PCAT19 knockdown or overexpression on the vitality of SPC-A1 cells measured using the CCK-8 assay; **(E–G)** effect of PCAT19 knockdown or overexpression on A549 and SPC-A1 cell proliferation measured using the soft agar assay. **(H,I)** effects of PCAT19 knockdown or overexpression on A549 and SPC-A1 cell migration measured using the migration assay. **(J–L)** effects of PCAT19 knockdown or overexpression on A549 and SPC-A1 cells invasion using transwell assays. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

DISCUSSION

LUAD is one of the leading causes of cancer-related deaths worldwide. Abnormal expression of lncRNAs can affect biological functions of cells such as tumor cell proliferation, migration, invasion, and apoptosis (Murillo-Maldonado and Riesgo-Escovar, 2019; Ginn et al., 2020). Recent studies have shown that lncRNAs play an important role in the progression of

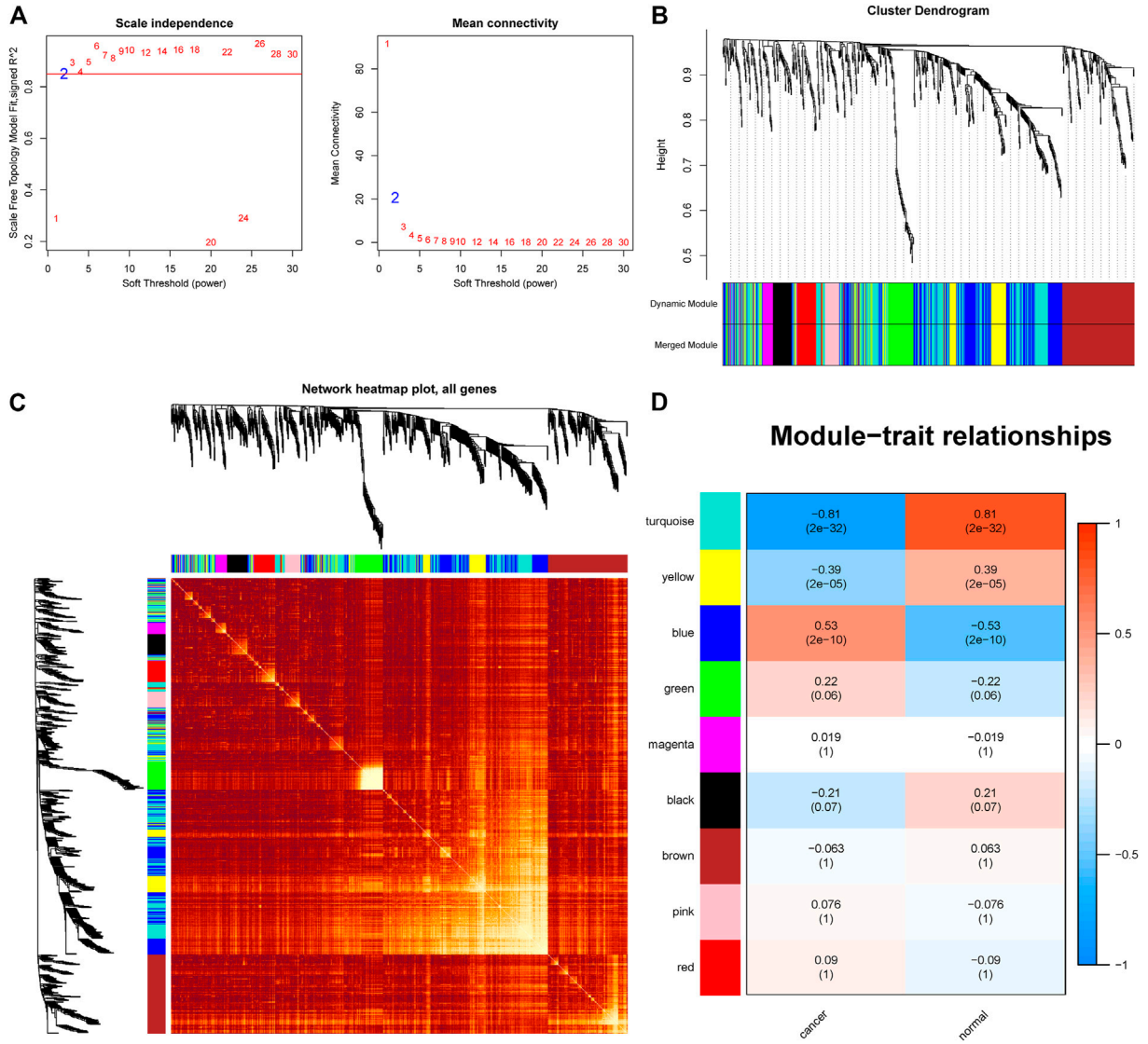


FIGURE 7 | WGCNA identification of lung adenocarcinoma associated miRNA modules in GSE74190. **(A)** Soft-threshold power versus scale-free topology model fit index and mean connectivity. **(B)** Cluster dendrogram of the co-expression network modules. **(C)** Cluster dendrogram of the co-expression network modules. **(D)** Analysis of the miRNA relationship between lung adenocarcinoma and normal sample modules.

LUAD. For example, lncRNA DGCR5 promotes the proliferation of LUAD cells (Dong et al., 2018), and lncRNA TTN-AS1 promotes the migration and invasion of LUAD cells (Jia et al., 2019). lncRNA HAGLR inhibits tumor growth of LUAD by silencing E2F1 (Guo et al., 2019). lncRNA ACTA2-AS1 inhibits the progression of LUAD by increasing the expression of SOX7 (Ying et al., 2020). Existing studies have shown that the role of lncRNAs in lung cancer (including LUAD) is largely unclear, which will help understand the molecular mechanisms involved in the development of lung cancer and find therapeutic strategies. However, the reported expression pattern and molecular function are only 1%, and there are fewer characterizations of lncRNAs (Yang et al., 2014). Therefore, in-depth research on the function of unknown lncRNAs is still needed.

At present, public databases of tumor RNA expression profiles, such as TCGA and GEO, show a continuous growth trend, which indicates that these databases will become key tools for verifying and researching clinical problems. In this study, we downloaded LUAD RNA expression profile data from the TCGA database. A total of 49 differentially expressed lncRNAs were screened, and 112 differentially expressed miRNAs and 2,953 differentially expressed mRNAs were obtained. Through Kaplan–Meier curve analysis, 16 lncRNAs related to the prognosis of LUAD patients were obtained. Then, based on the differentially expressed lncRNA-miRNA co-expression

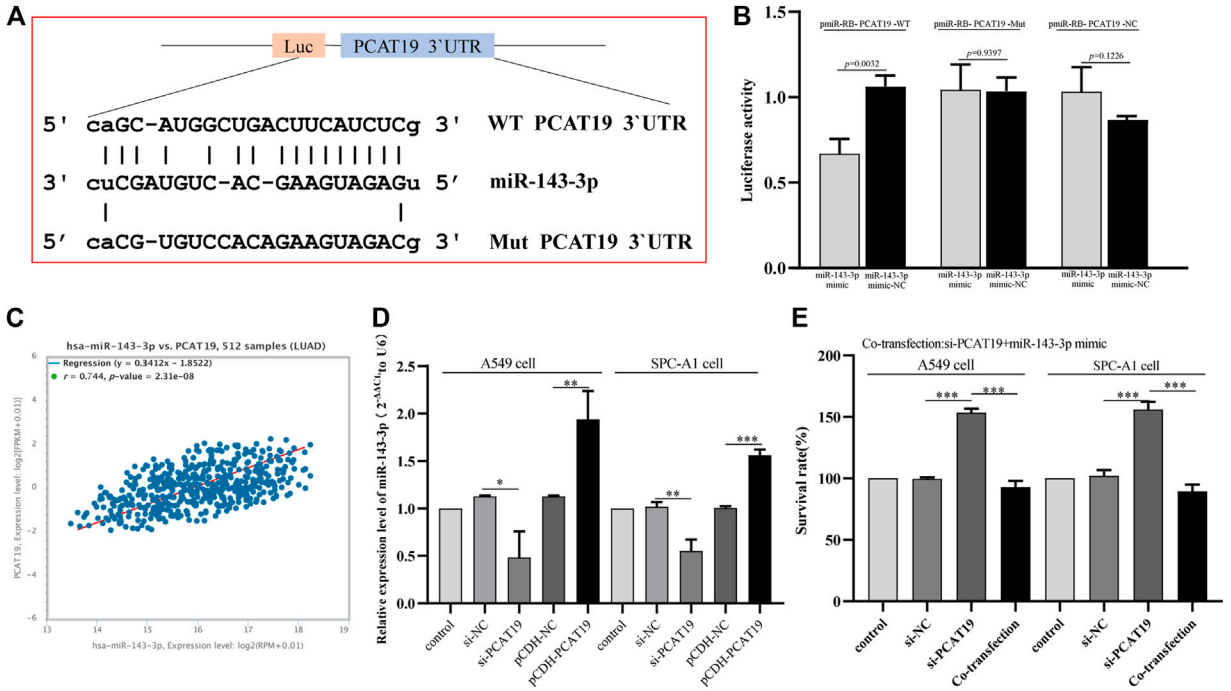


FIGURE 8 | (A) Schematic diagram of the binding site between miR-143-3p and WT or Mut PCAT19. **(B)** Luciferase activity of WT or Mut PCAT19 after co-transfection of miR-143-3p mimic and WT or Mut PCAT19 dual fluorescent vector into A549 cells. **(C)** Correlation between the expression levels of PCAT19 and miR-143-3p. **(D)** Expression change of miR-143-3p after PCAT19 knockdown or overexpression in A549 and SPC-A1 cells. The expression level of U6 was used as internal reference. **(E)** CCK-8 analysis of the proliferation activity of A549 and SPC-A1 cells co-transfected with si-PCAT19 and miR-143-3p mimic. The data shown are the mean ± standard deviation of three experiments. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

relationship, the differentially expressed miRNA-mRNA co-expression relationship and targeted RNAs (lncRNAs, miRNAs, and miRNAs), a ceRNA coaction network was constructed, including 3 lncRNAs (SMIM25, PCAT19, and LINC00261), 7 miRNAs, and 27 mRNAs. Through the comprehensive analysis of lncRNA-related genes, we can better grasp the molecular functions of abnormal lncRNA expression (Ma et al., 2012; Signal et al., 2016). Therefore, we analyzed 27 mRNAs related to three lncRNAs (SMIM25, PCAT19, and LINC00261) in the ceRNA network. GO and KEGG analyses of mRNAs were mainly enriched in cell proliferation, cell cycle, transmembrane signal transduction, and lung development. The pathway in cancer was the most significantly enriched KEGG. These data indicate that three lncRNAs (SMIM25, PCAT19, and LINC00261) were involved in the important tumor-related progression of LUAD.

In addition, to further study the impact of the three selected lncRNAs on the progression of LUAD, we used univariate Cox, multivariate Cox, risk prediction models, and qPCR. The results showed that the expression of PCAT19 was more significant in LUAD tissues and was more correlated with the prognosis of LUAD patients. Low expression of PCAT19 was an independent prognostic factor of poor survival in patients with LUAD. Further analysis of PCAT19 found that the expression of PCAT19 was significantly correlated with the histological grade (TMN), age, and tumor grade of LUAD, showing a significant negative correlation. In addition, the expression of PCAT19 was also negatively correlated with the tumor mutational burden of LUAD. Studies have reported that the high expression of PCAT19 in patients with laryngeal cancer is associated with poor overall survival of patients with laryngeal cancer (Xu et al., 2019). The high expression of PCAT19 in NSCLC significantly reduces the survival rate of patients (Zhang et al., 2019). All these results indicate that PCAT19 has a certain influence on tumor progression. To further understand the influence of PCAT19 on the progression of LUAD, we first performed GSEA and found that PCAT19 was mainly involved in the process of cell function. Subsequently, functional loss-gain experiments confirmed that the low expression of PCAT19 can promote the proliferation, migration, and invasion of A549 and SPC-A1 cells. In contrast, the high expression of PCAT19 had the opposite effect. These results indicate that PCAT19 has an important regulatory effect on the progression of LUAD. Previous studies also confirmed that PCAT19 could regulate the progression of other cancers. For example, lncRNA PCAT19 promotes the proliferation of laryngocarcinoma cells *via* modulation of the miR-182/PDK4 axis (Xu et al., 2019). PCAT19 interacts with HNRNPAB to activate a subset of cell-cycle genes associated with prostate cancer progression, thereby promoting prostate cancer tumor growth and metastasis (Hua et al., 2018). Additionally, PCAT19 negatively regulates the p53 tumor-suppression pathway, promoting cancer cell proliferation in patients with NSCLC (Zhang et al., 2019).

Moreover, to further understand the regulatory process of PCAT19, we obtained the coacting pairs (PCAT19/miR-143-3p) of PCAT19 through miRNAs in WGCNA and ceRNA. A dual luciferase assay confirmed the direct targeting relationship between PCAT19 and miR-143-3p, and the expression levels of PCAT19 and miR-143-3p were positively correlated, suggesting that PCAT19 may play a sponge role of miR-143-3p in LUAD. miR-143-3p has been proven to play an important role in many tumors. For example, miR-143-3p inhibits the tumorigenesis of pancreatic ductal adenocarcinoma by targeting KRAS (Xie et al., 2019); MiR-143-3p targets MAPK7 to inhibit the proliferation, migration, and invasion of osteosarcoma cells (Hou et al., 2019); LINC00667/miR-143-3p regulates RRM2 to affect the progression of small-cell lung cancer cells (Yang et al., 2019). The results of co-transfection experiments show that overexpression of miR-143-3p can reverse the effect of PCAT19 knockout on the proliferation of A549 and SPC-A1 cells, which further illustrates the regulatory relationship between PCAT19 and miR-143-3p, indicating that PCAT19/miR-143-3p plays an important role in the progression of LUAD.

However, limitations still existed in this study. For example, because we examined only 30 patients, our results may be accidental and need to be confirmed in large LUAD clinical cases. In addition, the mechanism of the PCAT19 regulation axis has not been further studied in our research. Even so, the trend should be analogous. Taken together, our study proves that PCAT19 may stimulate the development of LUAD and possesses great potential to be a prognostic biomarker and therapeutic target for LUAD.

CONCLUSION

In summary, using TCGA RNA sequencing data of LUAD, combined with ceRNA, PPI, GO and KEGG analyses, Cox regression analysis, risk model, and qPCR, the independent prognostic factor lncRNA PCAT19 for LUAD patients was screened. Cell function experiments confirmed that PCAT19 can affect proliferation, migration, and invasion of LUAD cells. In addition, the direct targeting RNA miR-143-3p of PCAT19 was screened in WGCNA and dual luciferase experiments. Overexpression of miR-143-3p reversed the effect of PCAT19 knockout on the proliferation of LUAD cells. In short, PCAT19 can be used as a potential biomarker for predicting LUAD.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, and further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of Jiangxi Chest Hospital and was conducted in accordance with the Declaration of Helsinki. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XT and YP conducted data analysis and wrote the manuscript. XH and XP carried out data analysis and manuscript revision. ZC designed the study. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Health Commission Project of Jiangxi Province (No. 20202BABL206090).

ACKNOWLEDGMENTS

The authors thank all TCGA and GEO data builders and data contributors as well as the team that built the GEPIA and Starbase online analysis page. They also thank Shanghai Ordovician Biotechnology Co., Ltd., for providing the platform of biological information analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.765275/full#supplementary-material

REFERENCES

Ali Syeda, Z., Langden, S. S. S., Munkhzul, C., Lee, M., and Song, S. J. (2020). Regulatory Mechanism of MicroRNA Expression in Cancer. *Ijms* 21 (5), 1723. doi:10.3390/ijms21051723

Alvarado-Luna, G., and Morales-Espinosa, D. (2016). Treatment for Small Cell Lung Cancer, where Are We Now? A Review. *Transl Lung Cancer Res.* 5 (1), 26–38. doi:10.3978/j.issn.2218-6751.2016.01.13

Cai, Z., and Liu, Q. (2021). Understanding the Global Cancer Statistics 2018: Implications for Cancer Control. *Sci. China Life Sci.* 64 (6), 1017–1020. doi:10.1007/s11427-019-9816-1

Chen, J., Liu, X., Xu, Y., Zhang, K., Huang, J., Pan, B., et al. (2019). TFAP2C-Activated MALAT1 Modulates the Chemoresistance of Docetaxel-Resistant Lung Adenocarcinoma Cells. *Mol. Ther. - Nucleic Acids* 14, 567–582. doi:10.1016/j.omtn.2019.01.005

Cheng, J.-T., Wang, L., Wang, H., Tang, F.-R., Cai, W.-Q., Sethi, G., et al. (2019). Insights into Biological Role of LncRNAs in Epithelial-Mesenchymal Transition. *Cells* 8 (10), 1178. Published 2019 Sep 30. doi:10.3390/cells8101178

Dong, H. X., Wang, R., Jin, X. Y., Zeng, J., and Pan, J. (2018). LncRNA DGCR5 Promotes Lung Adenocarcinoma (LUAD) Progression via Inhibiting Hsa-mir-22-3p. *J. Cel. Physiol.* 233 (5), 4126–4136. doi:10.1002/jcp.26215

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., et al. (2019). Estimating the Global Cancer Incidence and Mortality in 2018: GLOBOCAN Sources and Methods. *Int. J. Cancer* 144 (8), 1941–1953. doi:10.1002/ijc.31937

Gandhi, L., Rodríguez-Abreu, D., Gadgeel, S., Esteban, E., Felip, E., De Angelis, F., et al. (2018). Pembrolizumab Plus Chemotherapy in Metastatic Non-small-cell Lung Cancer. *N. Engl. J. Med.* 378 (22), 2078–2092. doi:10.1056/NEJMoa1801005

Ginn, L., Shi, L., La Montagna, M., and Garofalo, M. (2020). LncRNAs in Non-small-cell Lung Cancer. *ncRNA* 6 (3), 25. doi:10.3390/ncrna6030025

Guo, X., Chen, Z., Zhao, L., Cheng, D., Song, W., and Zhang, X. (2019). Long Non-coding RNA-HAGLR Suppressed Tumor Growth of Lung Adenocarcinoma through Epigenetically Silencing E2F1. *Exp. Cel. Res.* 382 (1), 111461. doi:10.1016/j.yexcr.2019.06.006

He, W., Liang, B., Wang, C., Li, S., Zhao, Y., Huang, Q., et al. (2019). MSC-regulated lncRNA MACC1-AS1 Promotes Stemness and Chemoresistance through Fatty Acid Oxidation in Gastric Cancer. *Oncogene* 38 (23), 4637–4654. doi:10.1038/s41388-019-0747-0

Hou, Y., Feng, H., Jiao, J., Qian, L., Sun, B., Chen, P., et al. (2019). Mechanism of miR-143-3p Inhibiting Proliferation, Migration and Invasion of Osteosarcoma Cells by Targeting MAPK7. *Artif. Cell Nanomedicine, Biotechnol.* 47 (1), 2065–2071. doi:10.1080/21691401.2019.1620252

Hua, J. T., Guo, H., Zhang, Y., Chen, S., Soares, F., Lu, J., et al. (2018). Risk SNP-Mediated Promoter-Enhancer Switching Drives Prostate Cancer through lncRNA PCAT19. *Cell* 174 (3), 564. doi:10.1016/j.cell.2018.06.014

Jia, Y., Duan, Y., Liu, T., Wang, X., Lv, W., Wang, M., et al. (2019). LncRNA TTN-AS1 Promotes Migration, Invasion, and Epithelial Mesenchymal Transition of Lung Adenocarcinoma via Sponging miR-142-5p to Regulate CDK5. *Cell Death Dis* 10 (8), 573. doi:10.1038/s41419-019-1811-y

Kim, M. H., Kim, Y. K., Shin, D. H., Lee, H. J., Shin, N., Kim, A., et al. (2015). Yes Associated Protein Is a Poor Prognostic Factor in Well-Differentiated Lung Adenocarcinoma. *Int. J. Clin. Exp. Pathol.* 8 (12), 15933–15939.

Lai, X.-N., Li, J., Tang, L.-B., Chen, W.-T., Zhang, L., and Xiong, L.-X. (2020). MiRNAs and LncRNAs: Dual Roles in TGF-β Signaling-Regulated Metastasis in Lung Cancer. *Ijms* 21 (4), 1193. doi:10.3390/ijms21041193

Liang, Y., Song, X., Li, Y., Chen, B., Zhao, W., Wang, L., et al. (2020). LncRNA BCRT1 Promotes Breast Cancer Progression by Targeting miR-1303/PTBP3 axis. *Mol. Cancer* 19 (1), 85. doi:10.1186/s12943-020-01206-5

Lin, J. J., Cardarella, S., Lydon, C. A., Dahlberg, S. E., Jackman, D. M., Jänne, P. A., et al. (2016). Five-Year Survival in EGFR -Mutant Metastatic Lung Adenocarcinoma Treated with EGFR-TKIs. *J. Thorac. Oncol.* 11 (4), 556–565. doi:10.1016/j.jtho.2015.12.103

Liu, J., Cui, J., Liu, F., Yuan, Y., Guo, F., and Zhang, G. (2019). Multi-subtype Classification Model for Non-small Cell Lung Cancer Based on Radiomics: SLS Model. *Med. Phys.* 46 (7), 3091–3100. doi:10.1002/mp.13551

Ma, H., Hao, Y., Dong, X., Gong, Q., Chen, J., Zhang, J., et al. (2012). Molecular Mechanisms and Function Prediction of Long Noncoding RNA. *Scientific World J.* 2012, 1–11. doi:10.1100/2012/541786

Ma, N., Tie, C., Yu, B., Zhang, W., and Wan, J. (2020). Identifying lncRNA-miRNA-mRNA Networks to Investigate Alzheimer's Disease Pathogenesis and Therapy Strategy. *Aging* 12 (3), 2897–2920. doi:10.18632/aging.102785

Murillo-Maldonado, J. M., and Riesgo-Escovar, J. R. (2019). The Various and Shared Roles of lncRNAs during Development. *Develop. Dyn.* 248 (11), 1059–1069. doi:10.1002/dvdy.108

Peng, W.-X., Koirala, P., and Mo, Y.-Y. (2017). LncRNA-mediated Regulation of Cell Signaling in Cancer. *Oncogene* 36 (41), 5661–5667. doi:10.1038/onc.2017.184

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA Hypothesis: the Rosetta Stone of a Hidden RNA Language? *Cell* 146 (3), 353–358. doi:10.1016/j.cell.2011.07.014

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer Statistics, 2019. *CA A. Cancer J. Clin.* 69 (1), 7–34. doi:10.3322/caac.21551

Signal, B., Gloss, B. S., and Dinger, M. E. (2016). Computational Approaches for Functional Prediction and Characterisation of Long Noncoding RNAs. *Trends Genet.* 32 (10), 620–637. doi:10.1016/j.tig.2016.08.004

Wang, W., Lou, W., Ding, B., Yang, B., Lu, H., Kong, Q., et al. (2019). A Novel mRNA-miRNA-lncRNA Competing Endogenous RNA Triple Sub-network Associated with Prognosis of Pancreatic Cancer. *Aging* 11 (9), 2610–2627. doi:10.18632/aging.101933

Xie, F., Li, C., Zhang, X., Peng, W., and Wen, T. (2019). MiR-143-3p Suppresses Tumorigenesis in Pancreatic Ductal Adenocarcinoma by Targeting KRAS. *Biomed. Pharmacother.* 119, 109424. doi:10.1016/j.biopha.2019.109424

Xu, S., Guo, J., and Zhang, W. (2019). lncRNA PCAT19 Promotes the Proliferation of Laryngocarcinoma Cells via Modulation of the miR-182/PDK4 axis. *J. Cel. Biochem* 120 (8), 12810–12821. doi:10.1002/jcb.28552

Yang, G., Lu, X., and Yuan, L. (2014). LncRNA: a Link between RNA and Cancer. *Biochim. Biophys. Acta (Bba) - Gene Regul. Mech.* 1839 (11), 1097–1109. doi:10.1016/j.bbagrm.2014.08.012

Yang, Y., Li, S., Cao, J., Li, Y., Hu, H., and Wu, Z. (2019). RRM2 Regulated by LINC00667/miR-143-3p Signal Is Responsible for Non-small Cell Lung Cancer Cell Progression. *Ott* 12, 9927–9939. doi:10.2147/OTT.S221339

Ying, K., Wang, L., Long, G., Lian, C., Chen, Z., and Lin, W. (2020). ACTA2-AS1 Suppresses Lung Adenocarcinoma Progression via Sequestering miR-378a-3p and miR-4428 to Elevate SOX7 Expression. *Cell Biol. Int.* 44 (12), 2438–2449. doi:10.1002/cbin.11451

Zhang, S., Zheng, F., Zhang, L., Huang, Z., Huang, X., Pan, Z., et al. (2020). LncRNA HOTAIR-Mediated MTHFR Methylation Inhibits 5-fluorouracil Sensitivity in Esophageal Cancer Cells. *J. Exp. Clin. Cancer Res.* 39 (1), 131. doi:10.1186/s13046-020-01610-1

Zhang, X., Wang, Q., Xu, Y., Wang, B., Jia, C., Wang, L., et al. (2019). lncRNA PCAT19 Negatively Regulates P53 in Non-small C-ell L-ung C-ancer. *Oncol. Lett.* 18 (6), 6795–6800. doi:10.3892/ol.2019.11041

Zhou, R.-S., Zhang, E.-X., Sun, Q.-F., Ye, Z.-J., Liu, J.-W., Zhou, D.-H., et al. (2019). Integrated Analysis of lncRNA-miRNA-mRNA ceRNA Network in Squamous Cell Carcinoma of Tongue. *BMC Cancer* 19 (1), 779. doi:10.1186/s12885-019-5983-8

Zhu, J., Yu, W., Wang, Y., Xia, K., Huang, Y., Xu, A., et al. (2019). lncRNAs: Function and Mechanism in Cartilage Development, Degeneration, and Regeneration. *Stem Cel Res Ther* 10 (1), 344. Published 2019 Nov 21. doi:10.1186/s13287-019-1458-8

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 January 2022
doi: 10.3389/fgene.2021.791919

LncRNA–miRNA–mRNA Networks of Gastrointestinal Cancers Representing Common and Specific LncRNAs and mRNAs

Hassan Dastsooz[1,2,3†], Ahad Alizadeh[4†], Parham Habibzadeh[5], Ali Nariman[6], Asieh Hosseini[7], Yaser Mansoori[8,9] and Hamed Haghi-Aminjan[10]*

[1]Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy, [2]Candiolo, C/o IRCCS, IIGM-Italian Institute for Genomic Medicine, Turin, Italy, [3]Candiolo Cancer (IT), FPO-IRCCS, Candiolo Cancer Institute, Turin, Italy, [4]Medical Microbiology Research Center, Qazvin University of Medical Sciences, Qazvin, Iran, [5]Research Center for Health Sciences, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran, [6]Genetics and Molecular Biology Department, Isfahan University of Medical Sciences, Isfahan, Iran, [7]Razi Drug Research Center, Iran University of Medical Sciences, Tehran, Iran, [8]Department of Medical Genetics, Fasa University of Medical Sciences, Fasa, Iran, [9]Noncommunicable Diseases Research Center, Fasa University of Medical Sciences, Fasa, Iran, [10]Pharmaceutical Sciences Research Center, Ardabil University of Medical Sciences, Ardabil, Iran

Gastrointestinal (GI) cancers are responsible for approximately half of cancer-related deaths, highlighting the need for the identification of distinct and common features in their clinicopathological characteristics. Long ncRNA (lncRNAs), which are involved in competitive endogenous RNA (ceRNA) networks with critical roles in biological processes, constitute a substantial number of non-coding RNAs. Therefore, our study aimed to investigate the similarities and differences in the ceRNA networks of The Cancer Genome Atlas (TCGA)-GI cancers. We performed a comprehensive bioinformatics analysis of ceRNA networks for TCGA-GI cancers in terms of the deferential mRNA, lncRNA, and miRNA expression levels, ceRNA networks, overall survival analysis, correlation analysis, pathological cancer stages, and gene set enrichment analysis. Our study revealed several common and distinct mRNAs and lncRNAs with prognostic values in these networks. It was specifically noteworthy that *MAGI2-AS3* lncRNA was found to be shared in almost all GI cancers. Moreover, the most common shared mRNAs between GI cancers were *MEIS1, PPP1R3C, ADAMTSL3, RIPOR2,* and *MYLK.* For each cancer ceRNA network, we found that the expression level of a number of lncRNAs and mRNAs was specific. Furthermore, our study provided compelling evidence that several genes, most notably *KDELC1*, can act as novel proto-oncogenes in cancers. This, in turn, can highlight their role as new prognostic and therapeutic targets. Moreover, we found cell cycle and extracellular matrix structural constituent as the top shared KEGG and molecular function, respectively, among GI cancers. Our study revealed several known lncRNAs and known and unknown mRNAs in GI cancers with diagnostic and prognostic values.

Keywords: tumor biomarkers, gastrointestinal cancers, long-non-coding RNA, The Cancer Genome Atlas (TCGA), competitive endogenous RNA (ceRNA)

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Sayed Haidar Abbas Raza,
Northwest A. and F. University, China
Lei Wang,
Changsha University, China

****Correspondence:***
Hamed Haghi-Aminjan
hamedhaghi.a@gmail.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *09 October 2021*
Accepted: *19 November 2021*
Published: *24 January 2022*

Citation:
Dastsooz H, Alizadeh A, Habibzadeh P, Nariman A, Hosseini A, Mansoori Y and Haghi-Aminjan H (2022) LncRNA–miRNA–mRNA Networks of Gastrointestinal Cancers Representing Common and Specific LncRNAs and mRNAs. Front. Genet. 12:791919. doi: 10.3389/fgene.2021.791919

1 INTRODUCTION

The gastrointestinal (GI) system consists of the substantial cellular mass in the human body, and its cancers are among the most common malignancies in different populations, accounting for 35% of the total deaths due to cancers (Arnold et al., 2020). Over the past decade or two, tremendous efforts have been invested into the identification of the molecular and biological processes responsible for the development of these cancers, mainly for colorectal cancer, hepatic cancer, gastric cancer, and head and neck squamous cell carcinomas (Sharma et al., 2018). In particular, identification of distinct and common features in their molecular and biological processes and also their clinical presentations can help shed light on the development and identification of diagnostic and therapeutic biomarkers.

More than 90% of the mammalian genome is transcribed into non-coding RNA (ncRNA), with a considerable number of them consisting of long ncRNAs (lncRNAs), transcripts over 200 nucleotides (nt) long (Sulayman et al., 2019). Increasing evidence has shown that lncRNAs are involved in competitive endogenous RNA (ceRNA) networks playing critical biological functions including the regulation of major cellular processes, namely, proliferation, differentiation, apoptosis, and stress response (Niu et al., 2020b; Li et al., 2021). Therefore, our study aimed to investigate the similarities and differences in the ceRNA networks of TCGA-GI cancers with more focus on lncRNAs. We included GI cancers with 11 and more than 11 matched normal tissues. These cancers were as follows: colon adenocarcinoma (COAD), rectal adenocarcinoma (READ), esophageal carcinoma (ESCA), stomach adenocarcinoma (STAD), head and neck squamous cell carcinoma (HNSC), and liver hepatocellular carcinoma (LIHC). Since HNSC is involved in the mucosa of the aero-digestive tract and also occurs in the oral cavity and salivary glands, we considered this cancer in our study as well.

Here, we report different and common shared lncRNAs and mRNAs involved in the ceRNA networks of these cancers. Moreover, our study shows the common and specific pathways in these cancers. Our finding also shows that a combination of lncRNAs and mRNA with prognostic values can promote their use for diagnostic and therapeutic aims. Furthermore, we propose a number of mRNAs that can function as new proto-oncogenes or tumor suppressors in their corresponding cancers, highlighting their diagnostic and therapeutic aspects.

Since our study reveals common shared and distinct mRNAs, lncRNAs, miRNAs, KEGG, GO-MF, GO-BP, and GO-CC among GI cancers, the identified results can help other researchers to use the data for further functional studies and differential diagnostic and prognostic strategies.

2 MATERIALS AND METHODS

2.1 Data Collection and the Differential Expression Levels of lncRNAs, mRNAs, and miRNAs

The Cancer Genome Atlas (TCGA) (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga) contains raw data of genomic, epigenomic, transcriptomic, and proteomic experiments from over 20,000 primary tumor tissues and their matched normal counterparts from 33 cancer types. In this study, we used an R/Bioconductor package, GDCRNATool developed by Ruidong Li et al. (Li et al., 2018) to download, organize and analyze lncRNAs, mRNAs, and miRNAs and clinical data of patients with GI cancers (HNSC, ESCA, STAD, LIHC, COAD, and READ) from TCGA.

Based on the following advantages, we used GDCRNATool for our analysis: One—GDCRNATools is an easy-to-use package for researchers with little coding experience. Two—It includes all data needed for performing the entire analysis smoothly in a very convenient way to download, organize, and perform comprehensive RNA expression analysis of TCGA data, with main focus on construction of the lncRNA–mRNA–miRNA-related ceRNA networks in cancer. Three—It gives not only genes involved in ceRNA networks but also all differentially expressed protein-coding genes, non-coding genes, and their GO terms and pathways. Therefore, it also has a flexibility to do other analyses that are not in the scopes of ceRNA network analysis. As an example, the differential expression data extracted from GDCRNATool can be used for considering the functional studies. Four—Moreover, to show that our differential expression data have an acceptable performance comparable to the state-of-the-art methods, we validated our expression data by looking for the expression of one gene (*UBE2C*) that was shown to be upregulated in all cancers (Dastsooz et al., 2019). Our differential expression data extracted from GDCRNATool also showed its higher expression in all these GI cancers. So, this tool can validate the differential expression data.

To validate our results, we also performed all analyses for the example that the developer of the tool did in their paper (for CHOL cancer). After confirmation of consistency between our results and theirs, we performed our analyses for GI cancers. So, we confirmed that this method of analysis is reproducible for all TCGA cancers. In our study, several analyses were carried out using GDCRNATools, which included differential gene expression analysis, univariate survival analysis, competing endogenous RNA network analysis, and functional enrichment analysis.

The appropriate data of these GI tumors and their matched adjacent non-tumor tissues were extracted. We started our analysis with the following sample numbers of RNA sequencing, miRNA, and clinical data for each cancer: COAD: 521 mRNA data, 465 miRNA data, and 459 clinical data; ESCA: 173 mRNA data, 200 miRNA data, and 185 clinical data; HNSC: 546 mRNA data, 569 miRNA data, and 528 clinical data; LIHC: 424 mRNA data, 425 miRNA data, and 377 clinical data; READ: 177 mRNA data, 165 miRNA data, and 171 clinical data; STAD: 407 mRNA data, 491 miRNA data, and 443 clinical data.

In this package, we downloaded RNA sequencing, mature miRNA, and clinical data from TCGA. Then, we parsed RNA sequencing metadata, filtered duplicated samples in this RNA metadata, and filtered non-primary tumor and non-solid tissue normal samples in them. Then, we parsed miRNAs metadata, filtered duplicated samples in these miRNAs data, and filtered non-primary tumor and non-solid tissue normal samples in the

miRNAs metadata. Next, we separately merged each RNA sequencing, miRNAs, and clinical data. After that, we normalized RNA sequencing data and miRNAs data (using gdcVoomNormalization). We then used the gdcDEAnalysis function to have a whole differential expression and the gdcDEReport function for differentiation expression of each mirRNA, lncRNA, and mRNA. Using the gdcCEAnalysis function considering data of differential expression of lncRNA and protein-coding genes, and extracting data of Starbas tool, we could decipher miRNA–lncRNA and miRNA–mRNA interactions and ceRNA networks of each GI cancer.

We used this package to have the gdcParseMetadata function for the efficient organization of RNA and clinical data. With this tool, we applied the gdcFilterDuplicate function to remove duplicated samples and gdcFilterSampleType to filter out samples without primary tumors or matched normal counterparts. Raw counts data were normalized using the gdcVoomNormalization function, which applied the TMM method in edgeR and the voom method in limma. Low-expression genes with logcpm less than one were excluded from the analysis. In this R package, the gdcDEAnalysis function, which included limma, edgeR, and DESeq2, was used for the identification of differentially expressed genes (DEGs) and miRNAs between primary GI tumors and their matched normal tissues. Visualization methods were volcano, scatter, and bubble plots, and also three simple shiny apps were used in GDCRNATools. All the figures were plotted using the ggplot2 package included in this tool.

2.2 Construction of the ceRNA Network

Using GDCRNATool, we investigated ceRNA networks in GI cancers. This tool with its gdcCEAnalysis function considers hypergeometric test (significantly commonly shared miRNAs between lncRNA and mRNA are detected), Pearson correlation analysis (lncRNA and mRNA with positive correlations are picked up), regulation similarity analysis (commonly shared miRNAs with similar function for regulation of the lncRNA and mRNA are considered), and sensitivity Pearson partial correlation to construct ceRNA networks. Using this tool, we first extracted several miRNAs that were correlated among both the lncRNAs and mRNAs. Then, this tool identified lncRNAs and mRNAs with positive expression correlation (Pearson correlation coefficient). Finally, we gained miRNAs with similar regulative effects on lncRNAs and mRNAs. In our analysis, we applied lncTarget (to get miRNA–lncRNA interactions) and pcTarget (to consider miRNA–mRNA interactions) data along with gdcCEAnalysis function to extract miRNA–target interactions predicted from several datasets. GdcCEAnalysis function considers data of miRNA–mRNA and miRNA–lncRNA interaction databases, for example, the interaction between miRNA and mRNA and also mRNA predicted by online datasets such as Starbas. In our analysis, lncRNA–miRNA–mRNA interactions extracted as edges and nodes were visualized in Cytoscape 3.7.2.

2.3 mRNA–lncRNA Correlation Analysis

To find the main mRNA in each ceRNA network of GI cancers, using GDCRNA package with gdcCEAnalysis function, we also investigated lncRNAs–mRNAs correlation.

2.4 Survival Analysis

We performed univariate survival analysis using the gdcSurvivalAnalysis function in the GDCRNA package with the selection of Kaplan–Meier (KM) analysis. In this analysis, the patients were divided into high- and low-expression groups according to the median.

2.5 Pathological Tumor Stages

We looked for expression of lncRNAs and mRNAs identified in ceRNA networks of GI cancers across their pathological tumor stages using GEPIA2 web server (http://gepia2.cancer-pku.cn/#index) (Tang et al., 2019).

2.6 Protein–Protein Interaction Network

We used the STRING database (https://string-db.org) (Szklarczyk et al., 2019) to analyze the possible protein–protein interactions between possible novel biomarkers identified in ceRNA networks of these cancers.

2.7 Functional Enrichment Analysis

Using *GDCRNATool,* we investigated the role of the genes that were found to be involved in the GI cancers, in biological processes (BP), molecular functions (MF), and cellular components (CC). Using gdcEnrichAnalysis, we also carried out Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG).

3 RESULTS

3.1 Differential Expression Profile of ceRNA Network Components in GI Cancers

In the current study, we found different numbers of up- and downregulated mRNAs, lncRNAs, and miRNAs in HNSC, ESCA, STAD, COAD, READ, and LIHC given in **Table 1**. Regarding COAD, these deferentially expression numbers and the lncRNAs were also previously listed (Poursheikhani et al., 2020). Regarding STAD and LIHC, some deferentially expression numbers have been reported by Zhang et al. (2020a) and Zheng and Yu (2021), respectively.

3.2 ceRNA Networks in GI Cancers

Our study revealed the following numbers of differentially expressed mRNAs (DEmRNAs), lncRNA (DElncRNA), and miRNAs (DemiRNAs) maintained in the ceRNA networks of each cancer (**Supplementary Table S1**): HNSC (6 lncRNAs, 30 miRNAs, and 64 mRNAs), ESCA (4 lncRNAs, 17 miRNAs, and 15 mRNAs), COAD [9 lncRNAs, 37 miRNAs, and 71 mRNAs, the ceRNA network for COAD cancer was previously shown in a different way (Poursheikhani et al., 2020)], STAD (4 lncRNAs, 18miRNAs, and 29 mRNAs), READ (9 lncRNAs, 21miRNAs, and 45 mRNAs), and LIHC [10 lncRNAs, 30 miRNAs, and 64 mRNAs, its ceRNA network was shown in a different figure by Zheng and Yu (2021)] (**Supplementary Table S1**).

Our findings showed that LIHC, COAD, and READ had more ceRNA networks. It is worth noting that our study revealed *MAGI2-AS3* lncRNA to be shared among almost all ceRNA networks of GI

TABLE 1 | Number of differentially expressed mRNAs, lncRNAs, and miRNAs in GI cancers.

Tumor type	COAD		READ		STAD		HNSC		LIHC		ESCA	
RNA type	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
mRNA	1,094	1,901	1,169	1,790	935	1,192	940	1,103	715	1,448	610	722
lncRNA	128	77	181	53	119	51	76	32	68	80	49	49
miRNA	170	160	165	114	67	59	88	81	59	71	46	33

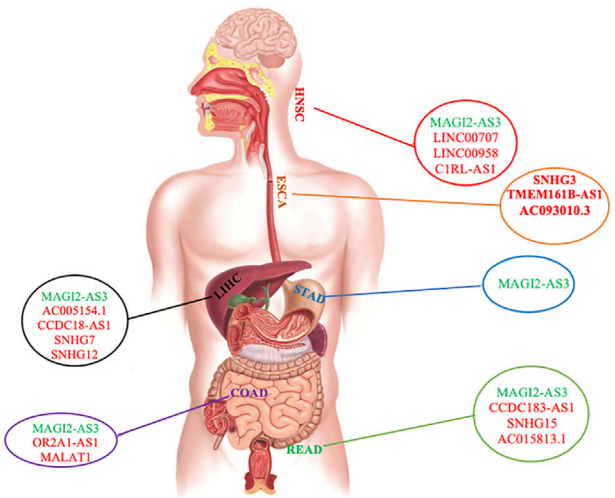


FIGURE 1 | Distribution of common and specific lncRNAs among different GI cancers. lncRNA in green color is the common one among GI cancers. lncRNAs in red are specific in each GI cancer. The gastrointestinal tract shown in this figure was reproduced from an image published under creative commons attribution license (Iriondo-Dehond et al., 2020).

cancers, except for ESCA, indicating the involvement of the same pathways in GI cancers (**Supplementary Table S2**, **Supplementary Figure S1**, **Figure 1**). Furthermore, *PVT1* was common between LIHC, ESCA, and STAD; *GAS5*, *SNHG1*, and *SNHG20* were common between LIHC, COAD, and READ; *KCNQ1OT1* was common between STAD, HNSC, COAD, and READ; *H19* was found to be common between LIHC, HNSC, and COAD; and finally, *MIR17HG* was common between STAD, COAD, and READ. We also investigated commonly shared lncRNAs between COAD and READ (CRCs) and found the presence of *MAGI2-AS3*, *GAS5*, *SNHG1*, *KCNQ1OT1*, *SNHG20*, and *MIR17HG* in both cancers. We were also able to distinguish COAD and READ using differential expression of *H19*, *OR2A1-AS1*, and *MALAT1* (specific for COAD) and *SNHG15*, *AC015813.1*, and *CCDC183-AS1* (specific for READ) (**Supplementary Table S2** and **Supplementary Figure S1**).

For each cancer ceRNA network, we found that the expression level of a number of lncRNAs was specific to each cancer: *SNHG12*, *AC005154.1*, *CCDC18-AS1*, and *SNHG7* were specific for LIHC; *SNHG3*, *TMEM161B-AS1*, and *AC093010.3* were specific for ESCA; *SNHG15*, *AC015813.1*, and *CCDC183-AS1* were specific for READ; *C1RL-AS1*, *LINC00707*, and *LINC00958* were specific for HNSC; *OR2A1-AS1* and *MALAT1* were specific for COAD (**Figure 1** and **Supplementary Table S2**). There was no specific lncRNA for ceRNA network of STAD. However, in comparison with other GI cancers, it was possible to distinguish it from other cancers (**Figure 1**, **Supplementary Table S2**, and **Supplementary Figure S1**).

Our study revealed that ceRNA networks with involvements of the following lncRNAs consist of a number of important mRNAs, indicating the presence of important ceRNA networks in their corresponding cancers: *MAGI2-AS3, SNHG15, KCNQ1OT1, SNHG1,* and *MIR17HG* in READ; *MAGI2-AS3, KCNQ1OT1, H19, MIR17HG, SNHG1,* and *MALAT1* in COAD; *MAGI2-AS3, KCNQ1OT1,* and *MIR17HG* in STAD; *MAGI2-AS3*, *PVT1*, *AC005154.1*, *H19*, *CCDC18*, and *SNHG1* in LIHC; *MAGI2-AS3, KCNQ1OT1, H19, CIRL-AS1, LINC00707,* and *LINC00958* in HNSC; and finally *TMEM161B-AS1* and *AC093010.3* in ESCA (**Supplementary Figure S1**).

We mentioned that *MAGI2-AS3* is shared in approximately all GI cancers. Then, we looked for proteins involved in its network and its significant interaction with them. We found several important ones shown in **Supplementary Table S3**. In all *MAGI2-AS3* networks, it has interacted with the following miRNAs: *has-miR-374a-5p* and *has-miR-374b-5p*. Its main mRNA partners that have positive correlations are *MEIS1, PPP1R3C, ADAMTSL3, RIPOR2,* and *MYLK*. It seems that the function of *MAGI2-AS3* is mainly dependent on its interaction with these five positive correlated genes in COAD, READ, and STAD. However, in LIHC, it is dependent on *ADAMTSL3, RIPOR2*, and *MYLK*, and in HNSC, it is dependent on the interaction with *MEIS1* and *PPP1R3C* (**Supplementary Table S3** and **Supplementary Figure S1**).

To find shared mRNAs among ceRNA networks of GI cancers, we compared the networks and identified four mRNAs, which were common in most of them as follows: *MEIS1* and *PPP1R3C* among CC-1 (COAD, READ, STAD, and HNSC), and *ADAMTSL3* and *MYLK* among CC-2 (COAD, READ, STAD, and LIHC) (**Table 2** and **Supplementary Figure S1**). Moreover, we found several important shared mRNAs between ceRNA networks of COAD and READ (CRC, cancer cluster 3:CC-3) (**Table 2**), which included *APC*, *EDIL3*, *FGFR2*, *HOMER1*, *MIER3*, *PI15*, *RBM28*, *RRS1*, *SCML1*, *SGPP1*, *TOMM34*, *TRAF5*, *WNT5A*, and *ZNF655*. For other cancer clusters, other important shared mRNAs as shown in **Table 2** were identified. These data indicated the common pathways for ceRNA networks of each cancer cluster. However, our data also showed specific protein-coding genes for each cancer presented in **Supplementary Table S2**, demonstrating specific ceRNA pathways for each cancer.

TABLE 2 | Commonly shared mRNAs between ceRNA networks of GI cancers.

CC-1	*MEIS1* and *PPP1R3C*
CC-2	*ADAMTSL3* and *MYLK*
CC-3	*APC, EDIL3, FGFR2, HOMER1, MIER3, PI15, RBM28, RRS1, SCML1, SGPP1, TOMM34, TRAF5, WNT5A,* and *ZNF655*
CC-4	*ATP8B2, L1CAM, LPAR1, NTN1, PCDH7, TTLL7, UST,* and *VAMP2*
CC-5	*ANKRD13B* and *RIPOR2*
CC-6	*CDC7, SOX12,* and *SRPX*
CC-7	*BUB1* and *LMNB2*
CC-8	*COL5A2, DI O 2, KDELC1, SPARC,* and *STX1A*
CC-9	*NFIX* and *PDZD2*
CC-10	*SPRY2*
CC-11	*RPS6KA5* (ESCA and READ) and *DAAM2* (ESCA and COAD)

CC-1: mRNAs shared between ceRNA networks of COAD, READ, HNSC, and STAD; CC-2: mRNA shared between ceRNA networks of COAD, READ, LIHC, and STAD; CC-3: mRNAs shared between ceRNA networks of COAD, and READ (CRCs); CC-4: mRNAs shared between ceRNA networks of STAD, and CRCs; CC-5: mRNA shared between ceRNA networks of LIHC, and CRCs, CC-6: mRNAs shared between ceRNA networks of LIHC, and COAD; CC-7: mRNAs shared between ceRNA networks of LIHC, and READ; CC-8: mRNAs shared between ceRNA networks of HNSC, and COAD; CC-9: mRNAs shared between ceRNA networks of HNSC, and STAD; CC-10: mRNAs shared between ceRNA networks of HNSC, and LIHC; C11: mRNAs shared between ceRNA networks of ESCA, and other GI, cancers.

3.3 OS Analysis

In the GI-ceRNA networks, we identified that the overexpression of *LINC00958* and *LINC00707* in HNSC and overexpression of *SNHG20* and *SNHG12* in LIHC were correlated with reduced OS and worse prognosis. However, overexpression of *SNHG20* in READ and *PVT1* in STAD and low expression of *MAGI2-AS3* in STAD were correlated with good prognosis (**Figure 2**). Moreover, OS analysis for mRNAs involved in ceRNA networks of these cancers showed prognostic values of several mRNAs given in **Table 3** and **Supplementary Figure S2**. Among them, we identified some new prognostic biomarkers (reduced OS with worse prognosis) that have not been reported or supported by functional studies in their corresponding tumors as follows: upregulation of *KDELC1* (*POGLUT2*) in HNSC, downregulation of *RIPOR2* and *SMOC1*, and upregulation of *TMEM164*, *B3GNT5*, *ZNF607*, and *ANKRD13B* in LIHC. These genes showed reduced OS with worse prognosis. The downregulated genes with reduced OS and worse prognosis may function as tumor suppressors, but those overexpressed mRNAs with reduced OS and worse prognosis can act as proto-oncogenes.

It worth noting that for LIHC, combinations of prognostic values of *SNHG20* lncRNA and *BUB1* mRNA, and also *SNHG12* lncRNA and *TMEM164* mRNA (associated with reduced OS with worse prognosis), could strengthen the predictive value of these markers for this cancer. Each of the lncRNA–mRNA combinations was in the same ceRNA network.

3.4 Pathological GI Tumor Stages

We found that the expression of some lncRNAs with worse prognosis in ceRNA networks of GI cancers had significant alteration across cancer stages, indicating their roles in cancer progression and invasion. These lncRNAs were *SNHG12* and *SNHG20* in LIHC (**Figure 3**).

In addition to the lncRNAs, we found that several mRNA with worse prognosis in their corresponding networks showed significant altered expression across tumor stages, indicating their roles in cancer progression and invasion. These genes were as follows: *LIMK1* ($p = 0.000934$), *PLXNA1* ($p = 0.00176$), *E2F8* ($p = 0.0155$), *EFNA3* ($p = 3.74e{-}07$), *SLC7A11* ($p = 0.0497$), *TMEM164* ($p = 0.00469$), *H2AFZ* ($p = 5.27e{-}06$), *BUB1* ($p = 5.82e{-}06$), *CBX2* ($p = 0.00033$), *EXO1* ($p = 4.88e{-}06$), *FANCE* ($p = 0.000301$), *LMNB2* ($p = 0.000311$), *SOX12* ($p = 0.00467$), *JPT1* ($p = 4.08e{-}07$), *MAFG* ($p = 0.00305$), *STK39* ($p = 0.0428$), *ANKRD13B* ($p = 0.000412$) (**Supplementary Figure S3**).

3.5 Providing Evidence That Highlights the Need for Functional Study on the Role of *KDELC1* in Cancers

The role of *RIPOR2, SMOC1, B3GNT5, ANKRD13B*, and *TMEM164* has not been clearly demonstrated in LIHC. Moreover, the role of *ZNF607* and mainly *KDELC1* has not been clearly delineated in any cancer [*in-silico* analysis by Zhou et al. (2021)]. However, our findings in this study highlight the need for functional study of *KDELC1* in cancers as follows: One—Our data showed its over-expression with worse prognosis (reduced OS) in HNSC. Two—It showed overexpression in several tumors such COAD, kidney renal clear cell carcinoma (KIRC), and kidney renal papillary cell carcinoma (KIRP) and lower expression in Kidney Chromophobe (KICH) (**Table 4**). Three—It is mainly localized in the nucleus (confidence score: 5) and endoplasmic reticulum (confidence score: 5), followed by cytosol (confidence score: 4) (https://www.genecards.org and https://www.proteinatlas.org). Four—It specifically targets extracellular EGF repeats of important proteins such as NOTCH1 and NOTCH3 (Takeuchi et al., 2018), indicating its possible roles in the Notch signaling pathway. Five—Using STRING database, we found that it had interactions with several proteins involved in different cancers such as *TXNDC9* (Feng et al., 2020b), *GXYLT2* (Cui et al., 2019), *POFUT1* (Chabanais et al., 2018), *XXYLT1*(Zeng et al., 2021), *FLNA* (Guo et al., 2018), and *ZC3H12C* (Suk et al., 2018) (**Figure 4**). These evidences strengthen the possible important role of *KDELC1* in various malignancies.

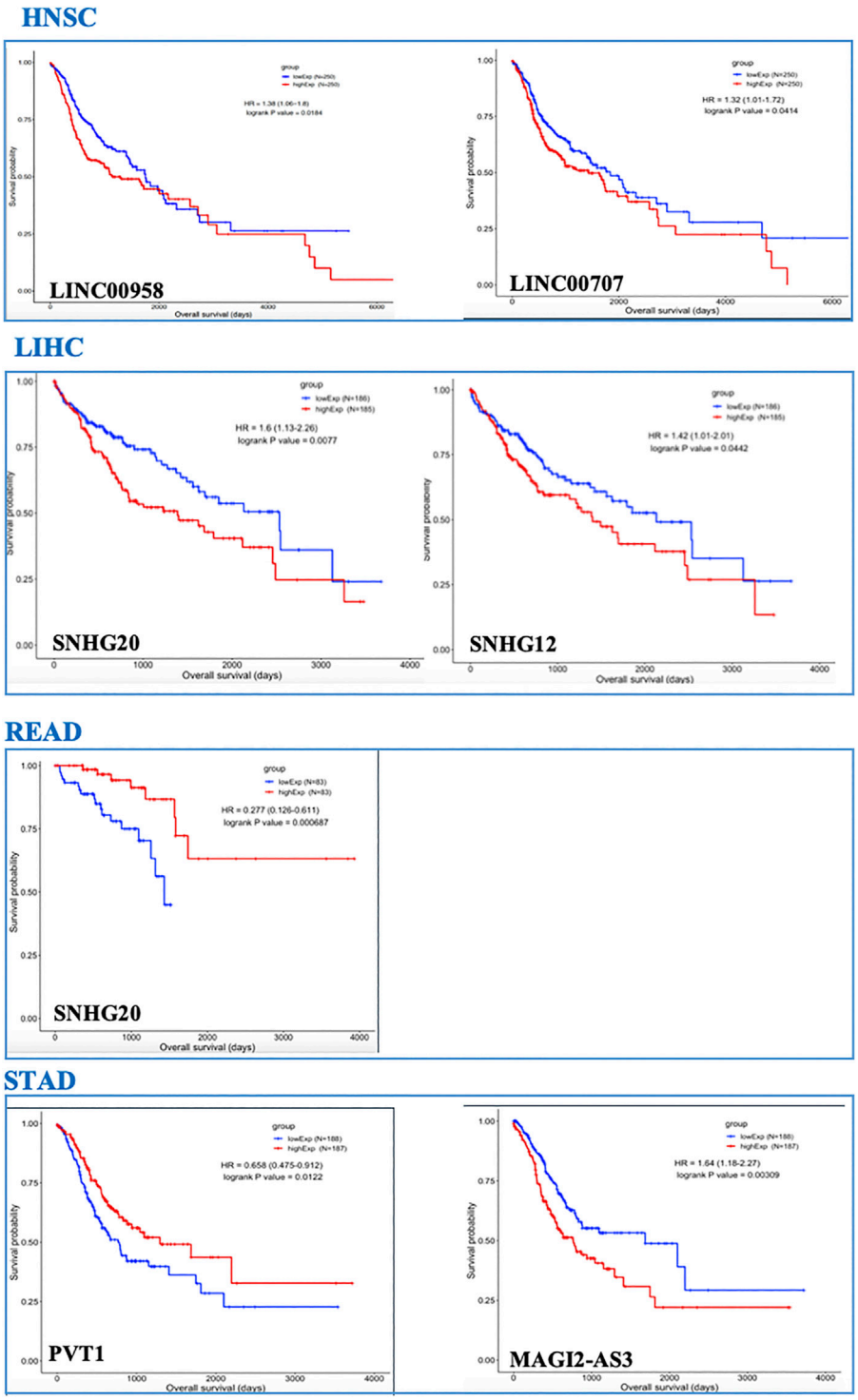


FIGURE 2 | OS analysis for lncRNAs in TCGA-GI cancers. Only lncRNAs with prognostic values are shown.

3.6 Functional Enrichment Analysis

In our study, firstly, we extracted the 10 top GO terms for BP, CC, and MF in the TCGA-GI cancers. Then, we compared these GO terms between GI cancers (**Supplementary Figure S4** and **Supplementary Table S4**). Our study revealed some common and specific GO terms for these cancers given in **Supplementary Tables S5,6**, respectively. These terms were previously shown in COAD cancer (Poursheikhani et al., 2020);

TABLE 3 | mRNAs as prognostic biomarkers identified in ceRNA networks of TCGA-GI cancers.

TCGA cancer type	Gene	Worse prognosis	Good prognosis
COAD	*HOMER1*		Over-expression
	NRG1	Low expression	
HNSC	*KDELC1*, *GNA12*, *ITGA5*, and *CDCA4*	Over-expression	
	MEIS1	Low expression	
LIHC	*RUNX3*, *CD69*, *RIPOR2 (FAM65B)*, *GADD45A*, *SYBU*, *ADAMTSL3*, *AUTS2*, *SMOC1*	Low expression	
	SCML2, *LIMK1*, *FANCE*, *PLXNA1*, *E2F8*, *EFNA3*, *SLC7A11*, *ENAH*, *TMEM164*, *H2AFZ*, *BUB1*, *CBX2*, *EXO1*, *LPL*, *B3GNT5*, *LMNB2*, *SOX12*, *JPT1*, *MAFG*, *ZNF607*, *STK39*, and *ANKRD13B*	Over-expression	
	CDC7 and *PFKFB3*		Low expression
	DNM3 and *COL15A1*		Over-expression
READ	*ATP8B2*, *NTN1*, *NOVA1*, and *PCDH7*		Low expression
STAD	*MAP7*		Over-expression

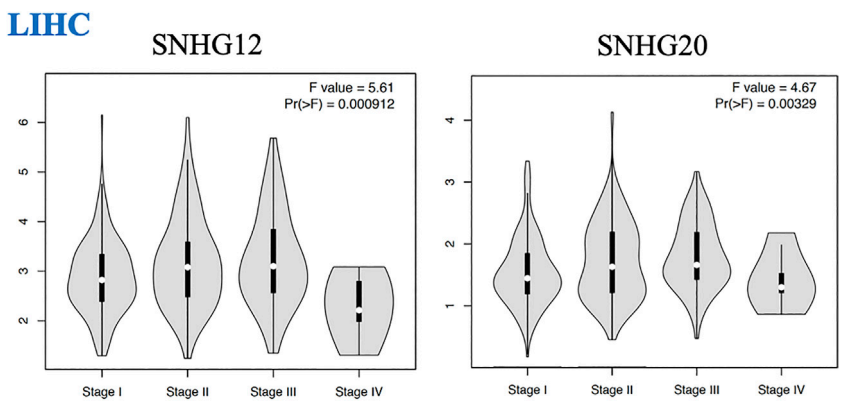


FIGURE 3 | Pathological tumor stages of LIHC and lncRNAs with significant altered expression across different stages of this malignancy.

TABLE 4 | *KDELC1* over-expression in TCGA cancers.

Cancer type	logFC	Average expression	*p*-value	Fdr
HNSC	1.4	2.62	7.05E-19	1.25E-17
COAD	1.1	2.38	1.79E-16	8.36E-16
KIRC	1.1	3.72	3.40E-24	1.61E-23
KIRP	1.95	4.35	5.58E-14	5.14E-13
KICH	-1.6	1.5	1.27E-16	2.18E-15

however, in our study, we looked for common and specific terms in GI cancers. For MF, our study revealed extracellular matrix structural constituent (GO:0,005,201) as the top shared term among all GI cancers (**Supplementary Table S5**). In relation to the CC, we identified collagen-containing extracellular matrix (GO:0062023), which was shared between all GI cancers (**Supplementary Table S5**). For BP, the most significantly shared GO terms in these cancers were extracellular matrix organization (GO:0030198), which was shared among all studied malignancies except LIHC. In addition, regulation of chromosome segregation (GO: 0051983) and mitotic nuclear division (GO:0140014) were shared among all investigated cancers excepting READ (**Supplementary Table S5**). Regarding KEGG, we found Cell cycle (hsa04110) as the top shared one among all GI cancers (**Supplementary Table S5**). However, we found that some BP, CC, and MF terms were exclusively specific for each cancer (**Supplementary Table S6**).

4 DISCUSSION

Our study revealed differential expression of mRNAs, lncRNAs, and miRNAs in ceRNA networks of GI cancers. We highlighted the important function of several lncRNAs and mRNAs specific for each cancer or commonly shared between GI cancers, which might be used for diagnostic and therapeutic purposes.

Our study showed that *MAGI2-AS3* lncRNAs were common between most of these cancers, representing possible involvement of the same pathways in different GI cancers. In ceRNA networks of each cancer, we also found that deferential expression of some lncRNAs was specific to each cancer.

We found several lncRNAs as prognostic biomarkers including *LINC00958* and *LINC00707* in HNSC; *SNHG20* and *SNHG12* in LIHC (correlated with reduced OS with worse prognosis); *SNHG20* in READ; and *PVT1* and *MAGI2-AS3* in STAD (correlated with good prognosis). We found that over-expression of *SNHG12* and *SNHG20* in LIHC (**Figure 3**) had significant alteration across cancer

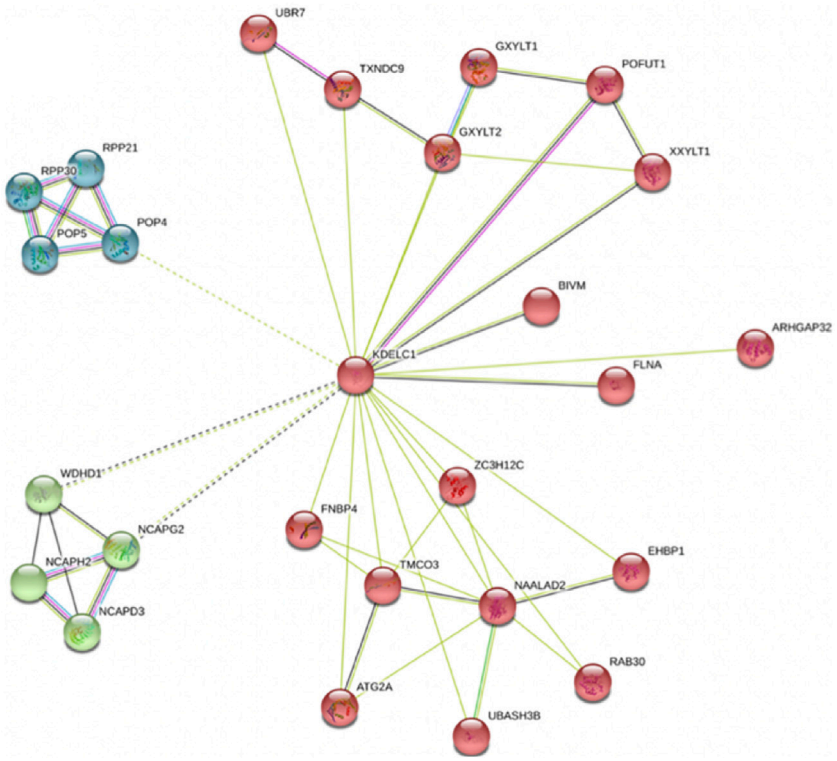


FIGURE 4 | KDELC1 protein network extracted from STRING webserver.

stages, representing their roles in cancer progression and invasion. These lncRNAs identified in our study were previously reported in their corresponding cancers (Yang et al., 2015; Zhang et al., 2016; Li et al., 2019; Tamang et al., 2019; Feng et al., 2020a; Niu et al., 2020a; Martinez-Barriocanal et al., 2020; Shen et al., 2020; Zeng et al., 2020; Zhang et al., 2020b; Zheng et al., 2020) but a comprehensive investigation of distinct or shared lncRNAs in these cancers has not yet been conducted.

Notably, we found a number of mRNAs with prognostic values (reduced OS with worse prognosis), which were not previously reported or supported by functional studies in their corresponding human tumors as follows: over-expression of *KDELC1* in HNSC; over-expression of *TMEM164*, *B3GNT5*, *ZNF607*, and *ANKRD13B*; and under-expression of *RIPOR2* and *SMOC1* in LIHC. The role of *ZNF607* and mainly *KDELC1* has not been clearly elucidated in cancers. Among these mRNAs, *TMEM164, ZNF607,* and *ANKRD13B* showed significant altered expression across tumor stages, indicating their role in cancer progression and invasion. Therefore, functional studies focused on these prognostic biomarkers in their corresponding cancers are suggested. Moreover, among GI cancers, four mRNAs were more common in most of them, which included *MEIS1* and *PPP1R3C* among COAD, READ, STAD, and HNSC; and *ADAMTSL3* and *MYLK* among COAD, READ, STAD, and LIHC. Furthermore, we found several important shared mRNAs between ceRNA networks of COAD and READ (**Table 2**, CRC, cancer cluster 3:CC-3).

According to the findings of the current study, we strongly proposed a novel proto-oncogene, *KDELC1* (*POGLUT2*), in cancers since it is mainly localized in nucleus and endoplasmic reticulum, it specifically targets extracellular EGF repeats of important proteins involved in Notch signaling pathways, it has interactions with several main proteins involved in cancers, and it is upregulated in several TCGA cancers. Moreover, our study revealed several mRNAs associated with worse prognosis in cancer and showed their significant altered expression across tumor stages.

Our study found that in ceRNA networks of LIHC, combination of prognostic values of *SNHG20* lncRNA and *BUB1* mRNA, and also *SNHG12* lncRNA and *TMEM164* mRNA (associated with reduced OS with worse prognosis), could be used as collective prognostic values for this cancer.

In this study, we found extracellular matrix structural constituent as the top shared MF term, collagen-containing extracellular matrix as the top shared CC, and Cell cycle (hsa04110) as the top shared KEGG among all these GI cancers. However, extracellular matrix organization, except for LIHC, and regulation of chromosome segregation and mitotic nuclear division, both except for READ, were found to be the most significantly shared BP terms among the GI cancers.

In conclusion, we found several common and distinct prognostic mRNAs and lncRNAs in GI cancers. The most important finding was the involvement of *MAGI2-AS3* lncRNA in almost all GI cancers. Another interesting result was the identification of the most common shared mRNAs in these cancers. Moreover, our study revealed a number of specific lncRNAs and mRNAs in GI cancers. Furthermore, our study reported several genes, most notably *KDELC1*, that can have proto-oncogenic roles in

cancers. In addition to the abovementioned findings, we reported the top shared KEGG, GO-MF, GO-BP, and GO-CC among GI cancers. Overall, our study showed that bioinformatic studies regarding mRNA-miRNA–lncRNA networks of TCGA data can help identify prognostic biomarkers. The identified common shared and distinct mRNAs, lncRNAs, miRNAs, KEGG, and GO-terms among GI cancers can be used by other researchers for further functional experiments and also for diagnostic and prognostic approaches.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study include TCGA, GEPIA2, STRING, Genecards, and Human Protein Atlas. The raw data can be found here: https://tcga-data.nci.nih.gov/tcga/, http://gepia2.cancer-pku.cn/#index, https://string-db.org, https://www.genecards.org, and https://www.proteinatlas.org.

AUTHOR CONTRIBUTIONS

HD conceptualized and designed the study, performed bioinformatic analysis, wrote, reviewed and edited the manuscript. AA performed bioinformatic analysis, reviewed, and edited the manuscript. PH reviewed and edited the manuscript. AN reviewed and edited the manuscript. AH reviewed and edited the manuscript. YM reviewed and edited the manuscript. HH-A conceptualized the study, performed systematic search for genes in article databases, reviewed and edited the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.791919/full#supplementary-material

Supplementary Figure S1 | ceRNA networks of LIHC, HNSC, READ, and ESCA.

Supplementary Figure S2 | OS analysis of protein-coding genes in patients with TCGA-GI cancers (HNSC, LIHC, READ, STAD, COAD).

Supplementary Figure S3 | mRNAs with significant altered expression across tumor stages in ceRNA networks of GI cancers.

Supplementary Figure S4 | GSEA of GI cancers (COAD, ESCA, HNSC, LIHC, READ, STAD).

Supplementary Table S1 | Differentially expressed mRNAs, lncRNA, and miRNAs in the ceRNA networks of GI cancers.

Supplementary Table S2 | Common and distinct lncRNAs and specific mRNAs in ceRNA networks of GI cancers.

Supplementary Table S3 | Protein-coding genes involved in MAGI2-AS3 ceRNA network of GI cancers.

Supplementary Table S4 | The GO terms (BP, CC, and MF) and KEGG pathways involved in the TCGA-GI cancers.

Supplementary Table S5 | Common GO terms and KEGG pathways of the GI cancers.

Supplementary Table S6 | Specific and distinct GO terms and KEGG pathways of the GI cancers.

REFERENCES

Arnold, M., Abnet, C. C., Neale, R. E., Vignat, J., Giovannucci, E. L., Mcglynn, K. A., et al. (2020). Global Burden of 5 Major Types of Gastrointestinal Cancer. *Gastroenterology* 159, 335–349. e315. doi:10.1053/j.gastro.2020.02.068

Chabanais, J., Labrousse, F., Chaunavel, A., Germot, A., and Maftah, A. (2018). POFUT1 as a Promising Novel Biomarker of Colorectal Cancer. *Cancers (Basel)* 10, 1. doi:10.3390/cancers10110411

Cui, Q., Xing, J., Gu, Y., Nan, X., Ma, W., Chen, Y., et al. (2019). GXYLT2 Accelerates Cell Growth and Migration by Regulating the Notch Pathway in Human Cancer Cells. *Exp. Cel Res.* 376, 1–10. doi:10.1016/j.yexcr.2019.01.023

Dastsooz, H., Cereda, M., Donna, D., and Oliviero, S. (2019). A Comprehensive Bioinformatics Analysis of UBE2C in Cancers. *Int. J. Mol. Sci.* 20, 1. doi:10.3390/ijms20092228

Feng, L., Li, H., Li, F., Bei, S., and Zhang, X. (2020a). LncRNA KCNQ1OT1 Regulates microRNA-9-Lmx1a Expression and Inhibits Gastric Cancer Cell Progression. *Aging* 12, 707–717. doi:10.18632/aging.102651

Feng, T., Zhao, R., Sun, F., Lu, Q., Wang, X., Hu, J., et al. (2020b). TXNDC9 Regulates Oxidative Stress-Induced Androgen Receptor Signaling to Promote Prostate Cancer Progression. *Oncogene* 39, 356–367. doi:10.1038/s41388-019-0991-3

Guo, Y., Li, M., Bai, G., Li, X., Sun, Z., Yang, J., et al. (2018). Filamin A Inhibits Tumor Progression through Regulating BRCA1 Expression in Human Breast Cancer. *Oncol. Lett.* 16, 6261–6266. doi:10.3892/ol.2018.9473

Iriondo-Dehond, A., Uranga, J. A., Del Castillo, M. D., and Abalo, R. (2020). Effects of Coffee and its Components on the Gastrointestinal Tract and the Brain-Gut Axis. *Nutrients* 13, 1. doi:10.3390/nu13010088

Li, R., Qu, H., Wang, S., Wei, J., Zhang, L., Ma, R., et al. (2018). GDCRNATools: an R/Bioconductor Package for Integrative Analysis of lncRNA, miRNA and mRNA Data in GDC. *Bioinformatics* 34, 2515–2517. doi:10.1093/bioinformatics/bty124

Li, S. J., Wang, L., Sun, Z. X., Sun, S. J., Gao, J., and Ma, R. L. (2019). LncRNA SNHG1 Promotes Liver Cancer Development through Inhibiting P53 Expression via Binding to DNMT1. *Eur. Rev. Med. Pharmacol. Sci.* 23, 2768–2776. doi:10.26355/eurrev_201904_17550

Li, Z., Wang, F., Zhu, Y., Guo, T., and Lin, M. (2021). Long Noncoding RNAs Regulate the Radioresistance of Breast Cancer. *Anal. Cel Pathol (Amst)* 2021, 9005073. doi:10.1155/2021/9005073

Martínez-Barriocanal, Á., Arango, D., and Dopeso, H. (2020). PVT1 Long Non-coding RNA in Gastrointestinal Cancer. *Front. Oncol.* 10, 38. doi:10.3389/fonc.2020.00038

Niu, J., Song, X., and Zhang, X. (2020a). Regulation of lncRNA PVT1 on miR-125 in Metastasis of Gastric Cancer Cells. *Oncol. Lett.* 19, 1261–1266. doi:10.3892/ol.2019.11195

Niu, Z.-S., Wang, W.-H., Dong, X.-N., and Tian, L.-M. -L. (2020b). Role of Long Noncoding RNA-Mediated Competing Endogenous RNA Regulatory Network in Hepatocellular Carcinoma. *Wjg* 26, 4240–4260. doi:10.3748/wjg.v26.i29.4240

Poursheikhani, A., Abbaszadegan, M. R., Nokhandani, N., and Kerachian, M. A. (2020). Integration Analysis of Long Non-coding RNA (lncRNA) Role in Tumorigenesis of colon Adenocarcinoma. *BMC Med. Genomics* 13, 108. doi:10.1186/s12920-020-00757-2

Sharma, K. L., Bhatia, V., Agarwal, P., and Kumar, A. (2018). Gastrointestinal Cancers: Molecular Genetics and Biomarkers. *Can. J. Gastroenterol. Hepatol.* 2018, 4513860. doi:10.1155/2018/4513860

Shen, A., Ma, J., Hu, X., and Cui, X. (2020). High Expression of lncRNA-SNHG7 Is Associated with Poor Prognosis in Hepatocellular Carcinoma. *Oncol. Lett.* 19, 3959–3963. doi:10.3892/ol.2020.11490

Suk, F. M., Chang, C. C., Lin, R. J., Lin, S. Y., Chen, Y. T., and Liang, Y. C. (2018). MCPIP3 as a Potential Metastasis Suppressor Gene in Human Colorectal Cancer. *Int. J. Mol. Sci.* 19, 1. doi:10.3390/ijms19051350

Sulayman, A., Tian, K., Huang, X., Tian, Y., Xu, X., Fu, X., et al. (2019). Genome-wide Identification and Characterization of Long Non-coding RNAs Expressed during Sheep Fetal and Postnatal Hair Follicle Development. *Sci. Rep.* 9, 8501. doi:10.1038/s41598-019-44600-w

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING V11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-wide Experimental Datasets. *Nucleic Acids Res.* 47, D607–D613. doi:10.1093/nar/gky1131

Takeuchi, H., Schneider, M., Williamson, D. B., Ito, A., Takeuchi, M., Handford, P. A., et al. (2018). Two Novel Protein O-Glucosyltransferases that Modify Sites Distinct from POGLUT1 and Affect Notch Trafficking and Signaling. *Proc. Natl. Acad. Sci. USA* 115, E8395–E8402. doi:10.1073/pnas.1804005115

Tamang, S., Acharya, V., Roy, D., Sharma, R., Aryaa, A., Sharma, U., et al. (2019). SNHG12: An LncRNA as a Potential Therapeutic Target and Biomarker for Human Cancer. *Front. Oncol.* 9, 901. doi:10.3389/fonc.2019.00901

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Res.* 47, W556–W560. doi:10.1093/nar/gkz430

Yang, M.-H., Hu, Z.-Y., Xu, C., Xie, L.-Y., Wang, X.-Y., Chen, S.-Y., et al. (2015). MALAT1 Promotes Colorectal Cancer Cell Proliferation/migration/invasion via PRKA Kinase Anchor Protein 9. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1852, 166–174. doi:10.1016/j.bbadis.2014.11.013

Zeng, H., Wang, Y., Wang, Y., and Zhang, Y. (2021). XXYLT1 Methylation Contributes to the Occurrence of Lung Adenocarcinoma. *Medicine (Baltimore)* 100, e24150. doi:10.1097/md.0000000000024150

Zeng, J., Liu, Z., Zhang, C., Hong, T., Zeng, F., Guan, J., et al. (2020). Prognostic Value of Long Non-coding RNA SNHG20 in Cancer. *Medicine (Baltimore)* 99, e19204. doi:10.1097/md.0000000000019204

Zhang, D., Cao, C., Liu, L., and Wu, D. (2016). Up-regulation of LncRNA SNHG20 Predicts Poor Prognosis in Hepatocellular Carcinoma. *J. Cancer* 7, 608–617. doi:10.7150/jca.13822

Zhang, X., Zheng, P., Li, Z., Gao, S., and Liu, G. (2020a). The Somatic Mutation Landscape and RNA Prognostic Markers in Stomach Adenocarcinoma. *Onco Targets Ther.* 13, 7735–7746. doi:10.2147/OTT.S263733

Zhang, Y., Huang, W., Yuan, Y., Li, J., Wu, J., Yu, J., et al. (2020b). Long Non-coding RNA H19 Promotes Colorectal Cancer Metastasis via Binding to hnRNPA2B1. *J. Exp. Clin. Cancer Res.* 39, 141. doi:10.1186/s13046-020-01619-6

Zheng, X., Wang, X., Zheng, L., Zhao, H., Li, W., Wang, B., et al. (2020). Construction and Analysis of the Tumor-specific mRNA-miRNA-lncRNA Network in Gastric Cancer. *Front. Pharmacol.* 11, 1112. doi:10.3389/fphar.2020.01112

Zheng, C., and Yu, S. (2021). Expression and Gene Regulatory Network of SNHG1 in Hepatocellular Carcinoma. *BMC Med Genomics* 14 (1), 28. doi:10.1186/s12920-021-00878-2

Zhou, H., He, Y., Li, L., Wu, C., and Hu, G. (2021). Identification Novel Prognostic Signatures for Head and Neck Squamous Cell Carcinoma Based on ceRNA Network Construction and Immune Infiltration Analysis. *Int. J. Med. Sci.* 18 (5), 1297–1311. doi:10.7150/ijms.53531

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 08 February 2022
doi: 10.3389/fgene.2022.804848

circEPS15 Overexpression in Hepatocellular Carcinoma Modulates Tumor Invasion and Migration

*Bin Jiang†, Maolin Tian†, Gang Li†, Abuduhaibaier Sadula, Dianrong Xiu *, Chunhui Yuan * and Yuntao Bing **

Department of General Surgery, Peking University Third Hospital, Beijing, China

Hepatocellular carcinoma (HCC) is one of the leading causes of cancer-related deaths worldwide. Recent evidence has shown that circular RNAs (circRNAs) play important roles in tissue development, gene transcription, signal regulation and tumorigenesis. However, whether circRNAs are involved in HCC progression and encode functional proteins remains largely unknown. In the present study, we aimed to explore the function and molecular mechanism of circRNAs in HCC. First, many circRNAs were found to be differentially expressed in HCC samples and paired adjacent normal liver tissues. The validation of dysregulated circRNAs by qRT-PCR revealed that circEPS15 expression was downregulated in HCC tissues, and the survival curves showed that low circEPS15 levels were associated with poor overall survival in HCC patients. Then, the overexpression of circEPS15 suppressed tumor cell invasion and migration by inhibiting the TJP1/CDH2/VIM signaling pathway and retarded cell cycle progression, which was confirmed by the Transwell culture system, wound healing assays, flow cytometry and western blot assays. After that, the spanning junction open reading frame in circEPS15 driven by IRES was shown to encode a novel protein, which was verified by western blotting with full-length, mutated, and truncated sequences of circEPS15 with a FLAG tag. Moreover, ceRNA analysis and qRT-PCR results suggest a possible circRNA (circEPS15)-miRNA-mRNA network in HCC. Collectively, our study reveals that endogenous circEPS15 plays a novel role in repressing HCC through the ceRNA network and encodes a functional protein.

Keywords: circular RNA, hepatocellular carcinoma, eps15, microarray, ceRNA

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Sayed Haidar Abbas Raza,
Northwest A&F University, China
Piyush Khandelia,
Birla Institute of Technology and Science, India

****Correspondence:***
Dianrong Xiu
xiudianrong@163.com
Chunhui Yuan
ychdoctor@163.com
Yuntao Bing
bingyuntao@bjmu.edu.cn

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *29 October 2021*
Accepted: *17 January 2022*
Published: *08 February 2022*

Citation:
Jiang B, Tian M, Li G, Sadula A, Xiu D, Yuan C and Bing Y (2022) circEPS15 Overexpression in Hepatocellular Carcinoma Modulates Tumor Invasion and Migration. Front. Genet. 13:804848. doi: 10.3389/fgene.2022.804848

BACKGROUND

Hepatocellular carcinoma (HCC) is one of the most common malignant tumors, ranking fifth in the incidence of malignant tumors and third in mortality (Olofson et al., 2018). Intrahepatic metastasis and postoperative recurrence are the main causes of death in patients with HCC. Epithelial-mesenchymal transition (EMT) is the process of transforming epithelial cells into mesenchymal cells with specific physiological or pathological phenotypes (Tam and Weinberg, 2013). In recent years,

Abbreviations: HCC, Hepatocellular carcinoma; circRNAs, Circular RNAs; IRES, Internal ribosomal entry site; TCGA, The Cancer Genome Atlas; RT-PCR, Real-time PCR; EMT, Epithelial mesenchymal transition; siRNA, Small interfering RNA; lncRNA, Long non-coding RNA; HCV, Hepatitis C virus; EPS15, Epidermal growth factor receptor pathway substrate 15; FC, Fold-change; cDNAs, Complementary DNAs; ROC, Receiver Operating Characteristic; AUC, Area Under Curve; NC, Negative control; EGF, Epidermal growth factor; EDH2, Eps15 homology domain-containing 2.

EMT has been found to be closely associated with the recurrence and metastasis of HCC and has attracted increased attention. Elucidating the molecular regulatory mechanism underlying EMT and the relationship between EMT and malignant tumors would assist in the search for new molecular targets (Tam and Weinberg, 2013; Xie and Diehl, 2013).

Circular RNA (circRNA) is different from traditional linear RNA. It has a 5′- 3′ polar covalent closed loop structure and is relatively stable, abundant, and conserved in the eukaryotic transcriptome. Recent studies have shown that the expression of circRNAs is tissue- and cell-specific (Salzman et al., 2012; Jeck et al., 2013; Lu et al., 2015). Evidence shows that circRNAs are associated with the occurrence, development, prognosis, and prognosis of cancer (Li et al., 2015; Starke et al., 2015; Chen et al., 2016; Rajappa et al., 2020; Kristensen et al., 2021); however, the role and molecular mechanism of circRNAs in HCC have not been elucidated. Therefore, elucidation of the role of circRNAs in the occurrence and development of HCC and identification of an effective target for HCC diagnosis and treatment are urgently needed.

HCC is a complex molecular process driven by gene mutations accompanied by epigenetic modifications (Bing et al., 2014). At present, there is no effective diagnostic or treatment strategy; thus, identification of potential effective diagnostic markers and therapeutic targets is urgently needed. Long noncoding RNAs (lncRNAs) have been found to be scrambled in patients with HCC, possibly having potential clinical value (He et al., 2014). However, the expression and function of circRNAs in HCC have not been elucidated.

Studies have shown that circRNA imbalance is associated with several human diseases and may be associated with the onset and progression of systemic lupus erythematosus, diabetes, and Alzheimer's disease (Zhao et al., 2016; Li C. et al., 2018; Li H. et al., 2018). Other studies have shown that circRNAs are involved in the occurrence and development of tumors, including colon, gastric, and cervical cancers (Gao et al., 2017; Hsiao et al., 2017; Huang et al., 2017; Weng et al., 2017). Although circRNAs have received increased attention, the biological functions and mechanisms of several circRNAs have yet to be elucidated, and it remains unknown whether these circRNAs are transcribed.

The molecular mechanism of circRNA biological function primarily includes miRNA absorption through the sponge effect, which affects the biological function of the miRNA (Zheng et al., 2016; Weng et al., 2017). circRNA is also a biological regulator that binds to proteins (Zeng et al., 2017) and can translate functional proteins, thereby influencing glioma tumorigenesis (Yang et al., 2018). However, whether circRNAs play a biological role in HCC and the underlying mechanism have not been elucidated.

miRNAs in circulation can be used as potential biomarkers for pancreatic cancer, Barrett's esophagus and esophageal adenocarcinoma (Gablo et al., 2019; Lv et al., 2020). However, although many clinical studies have been conducted on HBV-related HCC and hepatitis C virus (HCV)-related HCC (Li et al., 2019; Morishita et al., 2020), the instability of miRNAs and difficulty in reproducing results under different experimental conditions make miRNAs poor biomarkers for clinical practice. In contrast, circRNA is more stable due to its closed-loop structure and tolerance to several RNA enzymes present in the environment. Moreover, due to its tissue-specific expression (Melo et al., 2015), circRNA can be an ideal biomarker for predicting the occurrence, progression, and prognosis of disease.

Further investigations into the mechanism of circRNA may elucidate its potential therapeutic use in the future. Additionally, due to its stability, circRNA can be used as a continuous cell regulator. For example, a specific circRNA can be synthesized through genetic engineering to play a specific biological role by binding to a specific miRNA, competing for binding to ribosomal proteins, or translating a small polypeptide. One study showed that an overexpressed synthesized circRNA could absorb miRNA-122 through sponge-like interactions, thereby affecting the HCV replication cycle by inhibiting its protein synthesis (Jost et al., 2018). Therefore, synthetic circRNAs have broad applications in molecular medicine.

In this study, we investigated the cancer inhibitory role of circRNA-100226 in HCC using paired human HCC and paracarcinoma tissues. Using the human reference genome, we further estimated that the circRNA located at chr1: 51868106–51874004 is derived from EPS15 (epidermal growth factor receptor pathway substrate 15), which is located on chromosome 1p32.3. Thus, we referred to circRNA-100226 as "***circESP15***". We also overexpressed circEPS15 to further investigate the mechanisms of tumor growth.

MATERIALS AND METHODS

Ethics Statement

All clinical and experimental procedures were reviewed and approved by the Ethics Committee of Peking University Third Hospital. Written consent was obtained from all patients prior to surgery. All protocols were performed in accordance with the Declaration of Helsinki (General Assembly of the World Medical Association, 2014).

Patients and Specimens

A total of 71 HCC tissues and matched adjacent nontumor tissues were obtained from patients who underwent surgery in the Third Hospital, Peking University. None of the patients had received chemotherapy, radiotherapy, or any other anticancer therapy before surgery. The diagnoses were histologically confirmed by experienced pathologists. Paired tissue specimens (tumor and adjacent normal tissues) were collected from the patients, immediately frozen in liquid nitrogen, and stored at −80°C until use.

CircRNA Microarray Hybridization and Data Analysis

Total RNA was digested with RNase R (20 U/μL, Epicentre, Madison, WI, United States) to remove linear RNAs and enrich circular RNAs. The enriched circular RNAs were amplified and transcribed into fluorescent cRNA utilizing a random priming method (Super RNA Labeling Kit; Arraystar, Rockville, MD, United States). The labeled cRNAs were

hybridized onto the Human circRNA Array (8 × 15K, Arraystar). The slides were incubated for 17 h at 65°C in a hybridization oven (Agilent Technologies, Santa Clara, CA, United States). CircRNAs with significant differential expression between the HCC and paired normal tissues (fold change (FC) ≥ 2 and $p \leq 0.05$) were identified by volcano plot filtering. Hierarchical clustering was performed to determine the distinguishable expression pattern of circRNAs among the samples. The raw microarray data were uploaded to NCBI, and the GEO access number is GSE164803.

GO and KEGG Pathway Analysis

GO and KEGG pathway analyses were performed using standard techniques. GO analysis was performed to explore the functional roles of DE mRNAs and miRNAs in terms of "biological processes (BPs)," "cellular components (CCs)" and "molecular functions (MFs)" (http://www.geneontology.org). Pathway analysis was performed using the KEGG (http://www.genome.jp/kegg) database.

Prediction of Target miRNAs and Target mRNAs for circEPS15 and Biological Pathway Enrichment Analysis

CircRNAs act as miRNA sponges by binding miRNA response elements (MREs) in a competitive manner, preventing the inhibitory effects of miRNAs on target genes and increasing target gene amounts (Salmena et al., 2011; Phelps et al., 2016). To identify potential targets of miRNAs, we predicted the target/miRNAs with homemade miRNA target prediction software based on TargetScan and miRanda (Garcia et al., 2011; Li et al., 2014). First, the circRNA/miRNA interaction was predicted. Next, databases were used to identify miRNA-mRNA pairs. Finally, circRNA-miRNA-mRNA networks were constructed. GO and KEGG analyses were performed for targeted mRNAs.

Cell Culture

Human HCC cell lines (BEL-7402, BEL-7404, Huh7, Hep3B, and HepG2) and the normal hepatocyte cell line HL-7702 were purchased from the Shanghai Institute of Cell Biology (Shanghai, China) with a certificate of authenticity for each cell line. The samples were maintained in DMEM/RPMI 1640 (Gibco BRL, Grand Island, NY, United States) with 10% heat-inactivated fetal calf serum, 2 mM L-glutamine, and 100 U/mL penicillin–streptomycin mixture (Gibco BRL, Grand Island, NY, United States) at 37°C in 5% CO_2. After the cells were cultured for several months, all cell lines were evaluated for genetic identification prior to the experiments. Based on the appraisal reports provided by CoBioer Biosciences Co., Ltd. (Nanjing, China), no mutations or contaminations were found in our cell lines.

Quantitative Reverse Transcription Polymerase Chain Reaction

The cDNAs for circRNA, mRNA, or miRNA measurement were synthesized using random, oligo (dT)18, or stem-loop primers, respectively, using a RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). qRT-PCR was performed in triplicate using Maxima SYBR Green qPCR Master Mix (Thermo Fisher Scientific) on the CFX connect real-time system (Bio-Rad, Hercules, CA, USA). All primers used for qRT-PCR were designed using Primer 5.0 software (Premier, Canada) and synthesized by TsingKe Biotech (Chengdu, China). The primer sequences for the related qPCR validation are shown in **Supplementary Table S2**.

circEPS15-Overexpressing Vector Construction and Transduction

The circEPS15-overexpressing plasmid was synthesized and cloned into the adenoviral vector GV486 (GeneChem, Shanghai, China) with a Flag tag. Full-length circEPS15 was serially diluted with X-tremeGENE siRNA transfection reagent (Roche, Mannheim, Germany). For western blot analysis, cell proteins were prepared 72 h post-transfection.

Wound Healing Assay

The cells were seeded in 6-well plates and transfected following the manufacturer's instructions. When the cell confluence reached approximately 80% at 48 h post-transfection, the cells were serum-starved for 12 h and scratched using the fine end of a 100-μL pipette tip. Wound healing was observed at different time points within the scratch line. Every 24 h, scratch lines were imaged using an inverted Leica phase-contrast microscope (DFC 300 FX, Wetzlar, Germany) under a 20× objective lens. Image processing and analysis in Java was performed to obtain the wound-healing assay results. Duplicate wells were examined for each condition, and each experiment was repeated three times.

Transwell Assay

A cell invasion assay was performed in a Transwell chamber (24-well type, 8 mm pore size, Corning, Inc., Corning, NY, United States). BD Matrigel Basement Membrane Matrix (BD Biosciences, Franklin Lakes, NJ, United States) was used according to the manufacturer's instructions. Next, 0.5 ml of serum-free medium was added to the upper chambers; DMEM containing 10% FBS was added to the bottom chambers. Equal numbers (1×10^5) of cells were plated in the upper chambers of the quadruplicate wells and incubated at 37°C for 72 h. The cells were then fixed with paraformaldehyde and stained with crystal violet to visualize the nuclei. Average values were calculated from the results of three independent experiments. Image processing and analysis in Java was performed to obtain the Transwell assay results.

Cell Cycle and Apoptosis Assay

For analysis of the cell cycle and apoptosis, 3×10^5 treated cells were seeded in 6-well plates and cultured for 48 h at 37°C. For cell cycle analysis, the cells were digested using trypsin (HyClone, Logan, UT, United States), washed twice with phosphate-buffered saline (PBS), and fixed in 70% ethanol overnight at 4°C. The cells were centrifuged at 500 ×*g* for 5 min, washed twice with cold PBS, and centrifuged. After the cells were treated with RNase A (0.1 mg/ml) and propidium iodide (PI, 0.05 mg/ml, 4A Biotech, Beijing, China) for 30 min at 37°C, cell cycle analysis was performed using fluorescence-activated cell sorting flow cytometry (Beckman Coulter, Palo Alto, CA, USA).

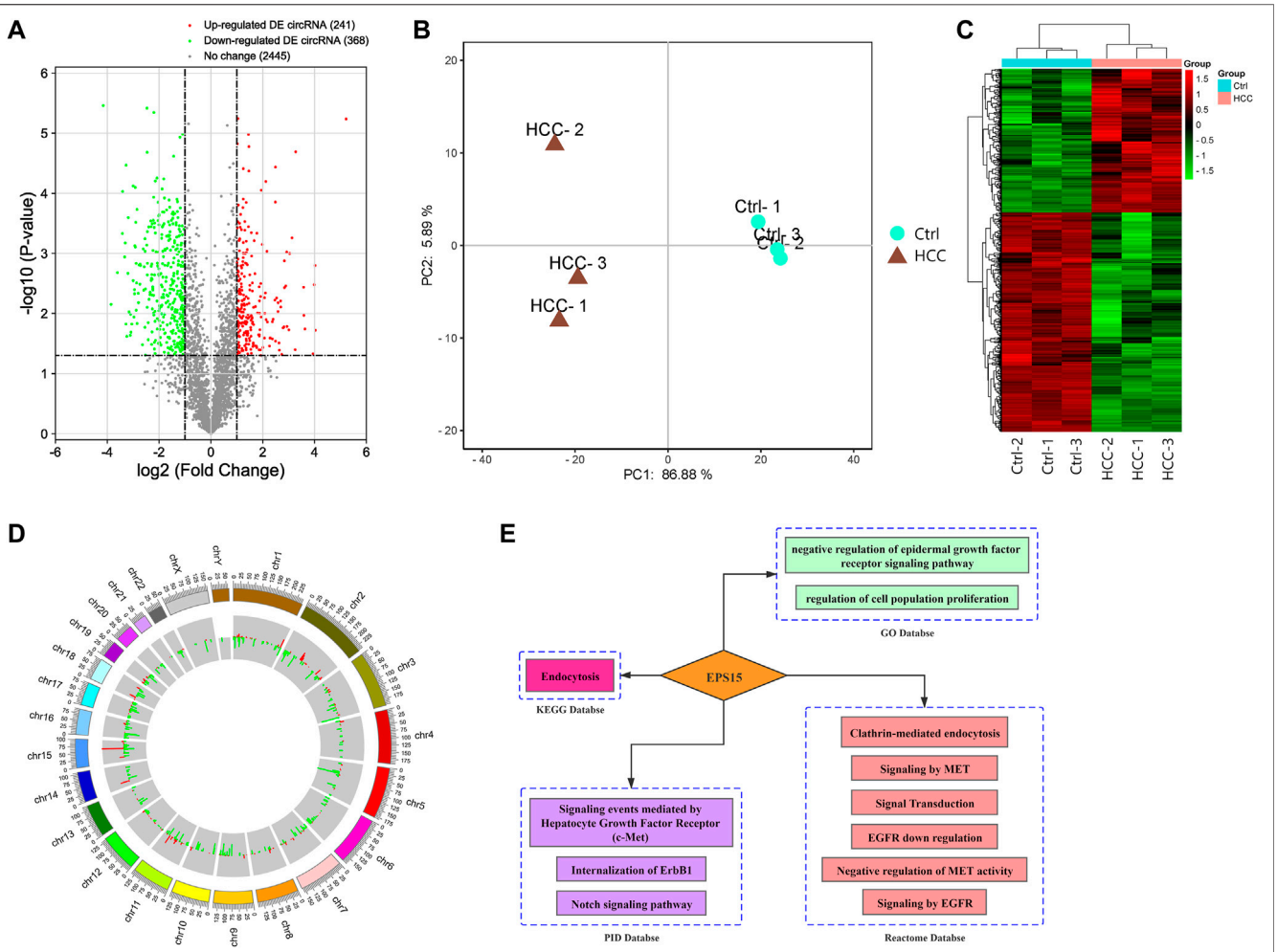


FIGURE 1 | (A) Volcano plot of differential expression of circRNAs between the HCC tissues and the control tissues. The *X*-axis and *Y*-axis of the volcano plot represent the log of the fold change (FC) so that changes in both directions (up and down) appear equidistant from the center and a negative log of the *p* value. **(B)** Principal component analysis (PCA) indicated that there was a distinct classification between the HCC group and the control group, and the biological repeatability of samples in the group was relatively good. **(C)** Heatmap of differentially expressed (DE) circRNAs: the DE circRNAs in the HCC tissues compared with the control tissues ($n = 3$). The FC cutoff was 2. **(D)** Circle plot of DE circRNA distribution in different chromosomes. **(E)** Biological function analysis of EPS15. GO database, KEGG database PID database and Reactome database enrichment analysis indicated that the host gene EPS15 was involved in the MET pathway.

For analysis of apoptosis, the cells were trypsinized followed by two PBS washes. The cells were stained using the Annexin V/PI detection kit (4A Biotech) for 5 min at 25°C. Apoptotic cells were measured using flow cytometry (Beckman Coulter). All experiments were repeated at least three times.

Western Blot Analysis

Exosomes were lysed with RIPA Lysis Buffer I (Sangon Biotech, Cat: C500005) to obtain the total protein. The protein concentration of each sample was determined using a NanoDrop™ spectrophotometer (Thermo Fisher Scientific, Cat: A30221). Protein (100 μg) from each sample was subjected to SDS-PAGE (4% stacking and 10% separating gels) and then transferred overnight onto polyvinylidene fluoride membranes (Millipore, Billerica, MA, United States). The primary antibodies used here included anti-human TJP1 antibody (Abcam, Cat: AB216880), anti-CDH2 antibody (Abcam, Cat: AB76057), anti-VIM antibody (Abcam, Cat: AB 137321), anti-EPS15 antibody (Abcam, Cat: AB224811), anti-GAPDH antibody (Abcam, Cat: AB8245), and anti-Flag antibody (Abcam, Cat: AB1162).

IRES Prediction

IRESfinder (Zhao et al., 2018) was used to identify the IRES sequence of circRNAs. Each sequence of circRNA was evaluated by the sliding window approach, with a window size of 174 nt and a step size of 50 nt. The region with the highest score was considered the IRES sequence of circRNA.

Statistical Analyses

All statistical analyses were performed using SPSS version 21.0 (SPSS, Inc., Chicago, IL, United States) and Prism version 5.0 (GraphPad Software, La Jolla, CA, United States) software. Categorical variables

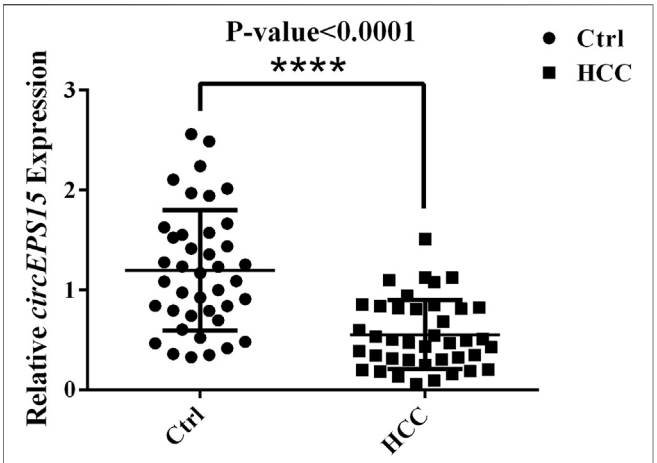


FIGURE 2 | Expression levels of circEPS15 in 40 randomly selected paired hepatoma and adjacent tissue samples detected with junction primers ($p < 0.01$).

are expressed as a count or percentage and were tested using $\chi 2$ or Fisher's exact test, as deemed appropriate. Continuous data are shown as the mean ± standard deviation (SD) and were compared using Student's t test, one-way analysis of variance, or Mann-Whitney test as deemed appropriate. Correlations were calculated using Pearson's correlation analysis. The median expression of target genes was used as the cutoff value to stratify patients into high and low expression groups. Survival curves were plotted using the Kaplan-Meier method and compared using the log-rank test. All tests were 2-sided, and results with $p < 0.05$ were considered statistically significant.

RESULTS

circRNA Expression Profiles in HCC

A microarray was performed to assess differentially expressed circular RNAs in the HCC vs. control groups ($n = 3$). circRNAs were considered differentially expressed with FC > 2 and p value <0.05. Six hundred and nine differentially expressed circRNAs were identified: 241 with upregulated and 368 with downregulated expression (**Figure 1A**). Principal component analysis was performed comparing the HCC and ctrl groups. Biological repeatability among the samples was relatively good (**Figure 1B**), as revealed by cluster analysis (**Figure 1C**). The chromosomal distribution of differentially expressed circRNAs is shown in **Figure 1D**. Raw intensity was set to >1,000 in each of the HCC and control groups, and the differentially expressed circRNAs up- and downregulated expression with a relatively high abundance and the top 10 multiples of differences were selected, as shown in **Supplementary Table S1**.

With the KOBAS V3.0 database (http://kobas.cbi.pku.edu.cn/), host circRNA genes with raw intensity >1,000 and with the top 10-fold up- and downregulation scores were annotated. We found that circRNA-100226, corresponding to endothelial growth factor receptor pathway substrate 15, is involved in the MET pathway and has related regulatory functions, as shown in **Figure 1E**. Based on the finding that the host gene EPS15 is involved in the MET pathway, we chose circEPS15 (circRNA-100226) as a candidate molecule to verify the related functional mechanisms.

qRT-PCR Validation of circEPS15

To verify the expression of circEPS15, we selected 40 random hepatoma samples and their adjacent tissues for detection by qRT-PCR with splice junction-specific primers. The primers are listed in **Supplementary Table S2**. circEPS15 had downregulated expression in HCC tissues, which was consistent with our microarray analysis (**Figure 2**).

Prognostic and Diagnostic Value of circEPS15

Seventy-one patients completed the clinical follow-up and had available survival information from which receiver operating

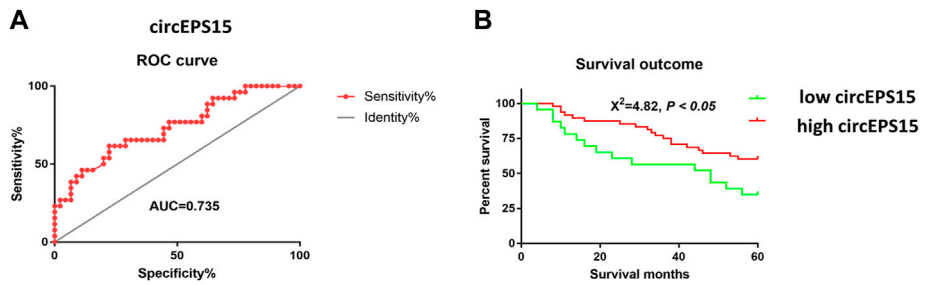


FIGURE 3 | (A) Combined with the prognostic results from these patients, an ROC curve was constructed, and an AUC of 0.735 for circEPS15 was obtained, indicating that the expression of circEPS15 has predictive value for prognosis of HCC patients. **(B)** Kaplan–Meier survival curves depict the correlation between circEPS15 and the overall survival of HCC patients. The patients were stratified into two groups according to a cutoff value of 0.8796 for circEPS15. Patients with low circEPS15 expression had a shorter overall survival than patients with high circEPS15 expression. A total of 71 patients were included in each experiment, and values indicate the mean ± SD of three independent experiments. Student's two-tailed paired t test and log-rank test were used for statistical analysis of these data.

TABLE 1 | Correlation between *circEPS15* expression and HCC clinical characteristics.

	Total	*circEPS15*		*p*-value
		Low (≤0.8976)	High (>0.8976)	
Age				
<60	48	32	16	0.515
≥60	23	16	7	
Gender				
Male	51	33	18	0.295
Female	20	15	5	
HBV				
<10^3	43	31	12	0.228
>10^3	12	17	11	
AFP				
<400	17	11	6	0.494
≥400	54	37	17	
ALB				
<35 g/L	13	10	3	0.329
≥35 g/L	58	38	20	
Tbil				
<23.3	55	38	17	0.416
≥23.3	16	10	6	
PT				
<12.8	47	34	13	0.177
≥12.8	24	14	10	
Tumor diameter				
<5 cm	37	27	10	0.225
≥5 cm	34	21	13	
Tumor differentiation				
M + L	32	17	15	0.017
H + M	39	31	18	
Tumor number				
1	57	37	20	0.260
2	14	11	3	
Intravascular tumor thrombus				
No	28	14	14	0.011
Yes	43	34	9	

characteristic (ROC) analysis was performed to determine whether circEPS15 had diagnostic significance. The results showed that the area under the curve (AUC) of circEPS15 was 0.735, which confirmed that circEPS15 was diagnostically significant for HCC (**Figure 3A**). Kaplan-Meier analysis of the correlation between circEPS15 and patient survival showed that low circEPS15 levels were significantly associated with poor overall survival (**Figure 3B**).

Correlation Between circEPS15 Expression and HCC Patients' Clinical Characteristics

To further confirm the role of circEPS15 in HCC, we analyzed the relationship between circEPS15 expression in HCC tissues and the clinicopathological characteristics of HCC patients. As shown in **Table 1**, there was a significant correlation between the low expression level of circEPS15 in HCC tissues and tumor differentiation (p = 0.017) and vascular invasion (p = 0.011). In addition, other characteristics of HCC, including the AFP level, tumor number, tumor size and BCLC stage, were not associated with circEPS15 expression. This correlation study may imply that circEPS15 is involved in the tumorigenesis and intrahepatic metastasis of HCC.

Effects of circEPS15 on HCC Tumor Cell Proliferation and Cell Cycle Progression

To investigate the specific function of circEPS15 in HCC formation and development, we confirmed that the tumor cell line BEL-7402 had low expression of circEPS15 compared to the normal cell line HL-7702 among 6 cell lines associated with HCC (**Supplementary Figures S1A**). We constructed an adenovirus vector to overexpress circEPS15 in the BEL-7402 cell line (**Supplementary Figures S1B**). Following transfection, both the fluorescence intensity of GFP and the expression of circEPS15 increased in BEL-7402 cells, indicating successful transfection and overexpression of circEPS15 in BEL-7402 cells (**Supplementary Figures S1C,D**); therefore, we used BEL-7402 cells for further study. We investigated the effect of circEPS15 overexpression on the cell cycle by measuring the percentage of HCC cells in different cell cycle stages (e.g., G1/G0, S, and G2/M) and found that circEPS15 overexpression increased the percentage of cells in G2/M compared to that of the vector control (**Figures 4A,B**). In addition, no significant difference in apoptosis was observed between the transfected cells and the negative controls (**Figures 4C,D**). These results indicated that circEPS15 elicited cell cycle arrest.

circEPS15 Overexpression Suppressed HCC Cell Migration and Invasion

A correlation study might show that circEPS15 is involved in HCC tumorigenesis and intrahepatic metastasis. Wound healing and Transwell assays indicated that circEPS15 significantly decreased HCC cell migration and invasion compared to the negative control (NC) (**Figures 4E–H**).

Through the Gene Ontology (GO, http://geneontology.org/), Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.genome.jp/kegg/), and Reactome (https://reactome.org/) databases, we identified the TJP1, CDH2, and VIM proteins involved in epithelial cell differentiation, cell junctions, cell adhesion molecules, the Rho GTPase cycle, and other pathways (**Supplementary Figures S2A**). Using the HCCDB database (http://lifeome.net/database/hccdb/home.html), we found that TJP1 was highly expressed in HCC compared with adjacent tissues, and a good prognosis was found with high expression (**Supplementary Figures S2B,C**). CDH2 was expressed at a low level, but high expression could be observed in individual datasets, and a good prognosis was linked to low expression (**Supplementary Figures S2E,D**). The expression trend of VIM in HCC was unstable, and the prognosis could not be accurately determined (**Supplementary Figures S2F,G**). We examined the TJP1, CDH2, and VIM expression levels using western blotting to further confirm that circEPS15 affected cell migration and invasion. The results showed that the TJP1 and CDH2 expression levels increased in cells

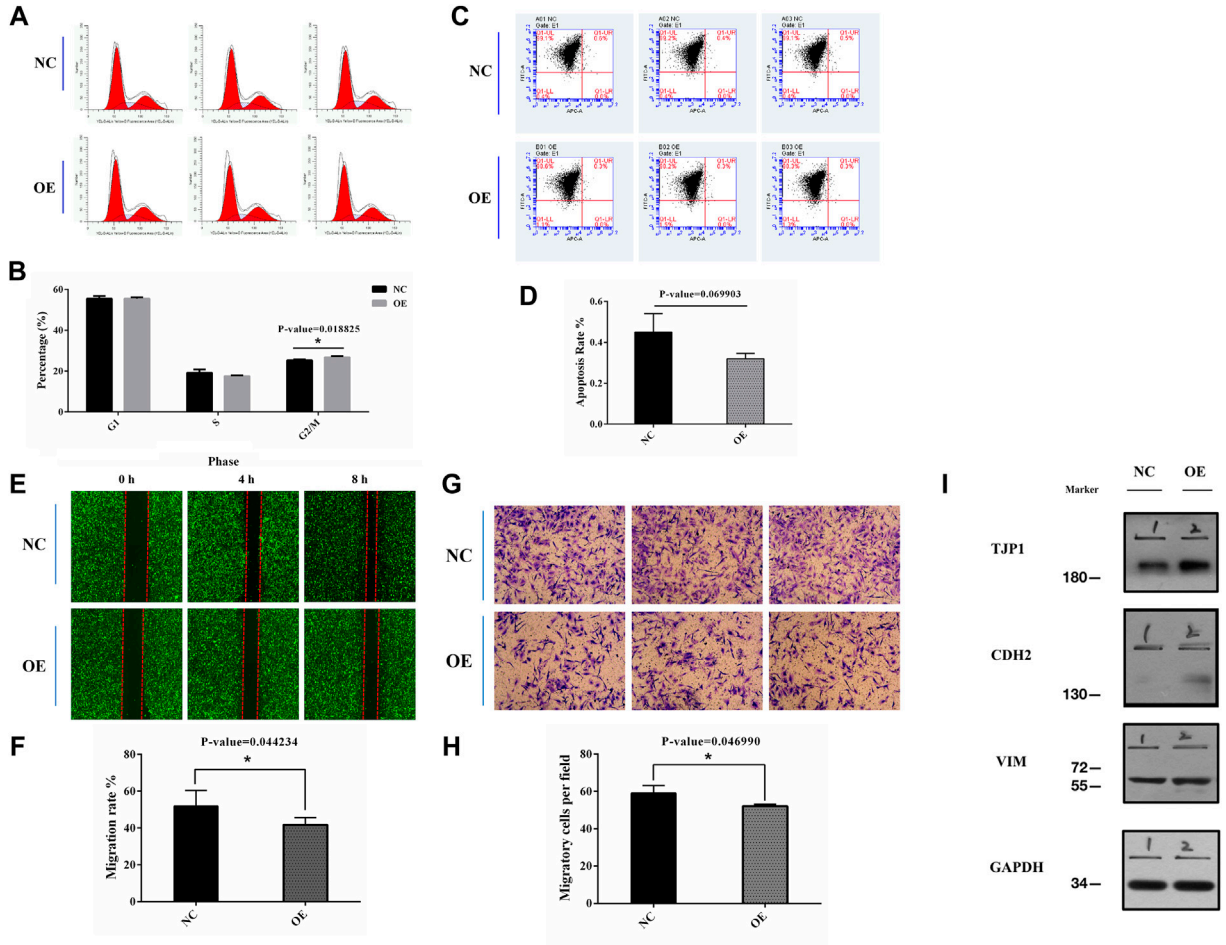


FIGURE 4 | (A,B) Flow cytometry was used to detect the cell cycle in transfected BEL-7402 cells (p = 0.018825). **(C,D)** Flow cytometry was used to detect cell apoptotic progression in transfected BEL-7402 cells (p = 0.069903). **(E,F)** Migration of cells overexpressing circEPS15 or the negative control was measured in a wounding assay at 0 and 48 h. The bar chart illustrates the relative migratory distance of cells into the wounded region. The data are the mean ± SEM of three independent experiments (p = 0.044234, compared to 0 h). **(G,H)** A Transwell assay was used to evaluate the invasion of cells overexpressing circEPS15 or the empty vector control after 48 h. The bar chart illustrates the relative percentage of migrating cells. The data are the mean ± SEM of three independent experiments (p = 0.046990, compared to the mock control). **(I)** Representative western blotting images show that TJP1 expression in BEL-7402 cells decreased and CDH2 expression increased in circEPS15-overexpressing cells compared with those transfected with the empty vector. However, there was no significant change in VIM expression.

overexpressing circEPS15, while that of VIM did not change significantly (**Figure 4I**).

circRNA-miRNA Interactions and Competing Endogenous RNA Regulatory Network Analysis

circRNAs act as microRNA sponges associated with related miRNAs, and together, they make up the circRNA-miRNA axis involved in disease pathogenesis. To determine the function of circEPS15, we constructed a circEPS15-miRNA-mRNA network to identify the potential targets of circEPS15 (**Figure 5A**). Detailed information on circEPS15-miRNA-mRNAs is shown in **Supplementary Table S3**. GO and KEGG analysis of the target mRNAs of circEPS15 was performed. The GO analysis indicated that the BPs were protein destabilization, regulation of DNA-templated transcription, elongation and circadian rhythm. The top three GO terms in the molecular function (MF) category were G-protein coupled peptide receptor activity, peptide receptor activity and C-C chemokine receptor activity. The top three GO terms in the cellular component (CC) category were cytoplasm, microtubule organizing center part and intracellular part (**Figure 5B**). Additionally, KEGG analysis indicated the related pathways and associated functions. Four KEGG pathways for the target genes of circEPS15 were glutathione metabolism, the IL-17 signaling pathway, the TNF signaling pathway and the Notch signaling pathway, indicating the role of circEPS15 in HCC (**Figure 5C**). In the prediction, the top five miRNAs for circEPS15 sponging were hsa-miR-24-3p, hsa-miR-138-5p, hsa-miR-145-3p, hsa-miR-620 and hsa-miR-875-3p (**Supplementary Figures S3A**). Additionally, DIANA-miRPath analysis revealed that these five miRNAs were related to ECM-receptor interactions, the cell cycle and protein binding transcription factor activity (**Supplementary Figures S3B,C**).

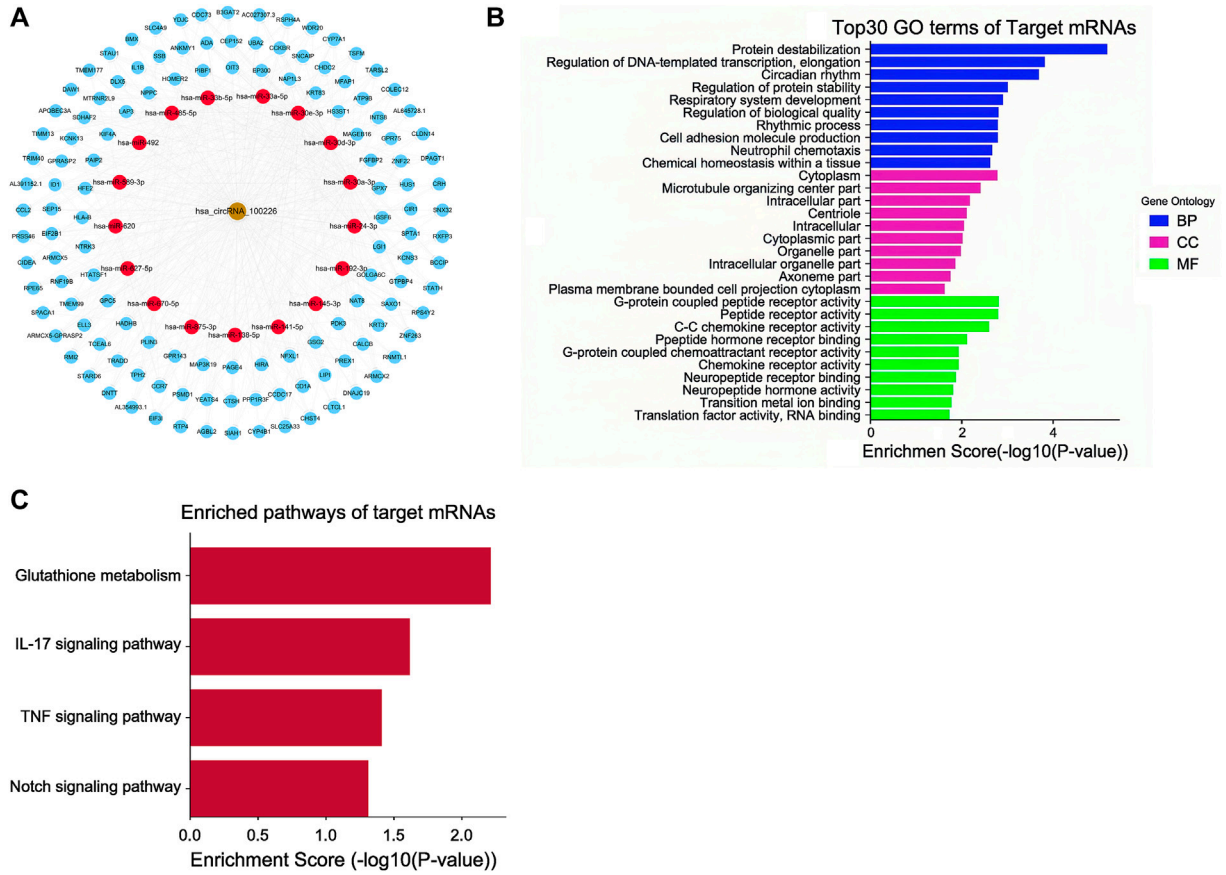


FIGURE 5 | ceRNA analysis of circEPS15. **(A)** CeRNA network plot of circEPS15: the gold circle represents circEPS15 (circRNA-100226), red circle represents miRNA, and blue circle represents mRNA. A ceRNA network including 1 circRNA, 17 miRNAs, and 127 mRNAs was constructed by integrated analysis. **(B)** GO analysis of ceRNA target genes. The top 30 GO terms of target mRNAs and their biological process (BP), cellular component (CC), and molecular function (MF) categories identified by DAVID. **(C)** KEGG analysis of ceRNA target genes. The statistically significant KEGG pathways of target mRNAs identified by DAVID.

Evaluation of the circEPS15 Protein-Coding Ability

The circEPS15 sequence was matched in circRNADb; therefore, we next analyzed the putative open reading frame (ORF) in circEPS15 (**Figure 6A**). There was a potential spanning junction ORF in circEPS15 that could encode a 150-amino acid protein. Conservation analysis showed that this ORF was highly conserved among different species, suggesting that this ORF was translatable. To test the putative ORF transcriptional ability, we cloned wild-type (ATG) or mutant-type (AAG) circEPS15 and inserted it into the GV486 vector (**Figure 6B**). To test the putative IRES activity in circEPS15, we also cloned a full-length or truncated putative circEPS15 IRES into the GV486 vector. To add a FLAG tag to the gene of interest, we moved the junction to the stop codon, with side flanking sequences (circEPS15-FLAG) (**Figure 6B**). We transfected these plasmids into 293T cells and detected their potential translated products. As shown in **Figure 6C**, the FLAG antibody only detected an ~25 kDa protein in the wild-type circEPS15-FLAG (OE5) and circEPS15 (ins_FLAG)-IRES1del (OE4)-transfected cells but not in the mutated (AAG), IRES2 truncated, or IRES1 and IRES2 truncated circEPS15-FLAG transfected cells. The scores for the IRES prediction in these five sequences were consistent with the western blotting results, wherein OE4 and OE5 had the highest scores (0.75) by IRESFinder (**Table 2**). These results indicated that circEPS15 can encode a protein, the ORF is translatable, and ATG and IRES2 are essential for the 5′-cap-independent translation of circEPS15.

DISCUSSION

Because various studies have shown that many circRNAs can be stably expressed in eukaryotic cells, several researchers have investigated the role of circRNAs in various diseases. The role(s) of circRNAs in tumors is also becoming a hot research topic.

At present, research on the role of circRNAs in tumors is mainly focused on the discovery and identification of new target

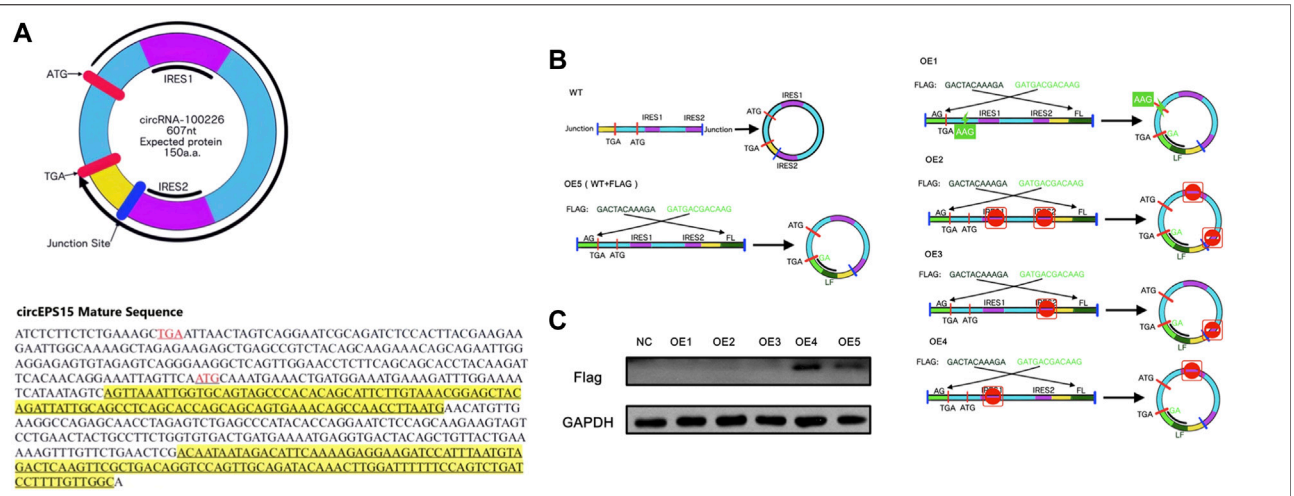


FIGURE 6 | Evaluation of the protein coding ability of circEPS15. **(A)** Upper panel: The putative open reading frame (ORF) in circEPS15. Note that the circEPS15 junction is inside the ORF. Lower panel: The sequences of the putative ORF are shown in black, internal ribosomal entrance site (IRES) sequences are shown in yellow, and start (ATG) and stop (TGA) codons are shown in red. **(B)** WT: Illustration showing how endogenous circEPS15 is formed; the circular junction is inside the ORF and forms unique sequences (yellow). OE5: A FLAG tag was divided on both sides (light and dark green). Circularization of this vector formed the same circular RNA as endogenous circEPS15, except with a FLAG tag added behind the ORF. OE1: The start codon was mutated from ATG to AAG to clarify the role of the start codon. OE2: IRES sequences (IRES1 and IRES2) in circEPS15 were truncated. OE3: IRES sequences (IRES2) in circEPS15 were truncated. OE4: IRES sequences (IRES1) in circEPS15 were truncated. **(C)** A FLAG tag antibody was used to detect circEPS15 expression in 293T cells transfected with the vectors described above.

TABLE 2 | The IRESfinder was used to identify the IRES sequence in different circRNAs. The region with the highest score was considered to be the IRES sequence for the circRNA.

		ORF (n)	Score (max)	Score (mean)	Score (sd)
OE1	circEPS15(ins_FLAG)-AAG	—	—	—	—
OE2	circEPS15(ins_FLAG)-IRES1/2del	4	0.60	0.55	0.03
OE3	circEPS15(ins_FLAG)-IRESdel2	7	0.72	0.59	0.07
OE4	circEPS15(ins_FLAG)-IRES1del	6	0.75	0.59	0.08
OE5	circEPS15(ins_FLAG)	9	0.75	0.6	0.08

circRNAs, their influence on tumor biological behavior, and the molecular mechanism of circRNAs.

A variety of new circRNAs have been found due to the popularization and application of circRNA chips. Through RNA enzyme digestion and sequencing technology, we can identify the circular structure of newly found circRNAs and confirm them as circRNAs.

The role of circRNAs in tumor biological behavior is becoming a new research hotspot. In hepatocellular carcinoma, some circRNAs have been found and confirmed to be related to the biological behavior and prognosis of hepatocellular carcinoma. For example, circFBLIM1 can competitively bind to miR-346 through a ceRNA mechanism, thus promoting the progression of HCC. This study indicates that circRNAs exist in HCC and play an important role in the occurrence and development of HCC. As a new type of RNA, circRNA should be investigated to confirm its role in liver cancer, as it may increase detection of liver cancer and provide a novel way(s) to treat the disease.

The molecular mechanism of circRNA has always been an interesting and difficult area in this field. Interestingly, artificial circRNA has been proven to be translatable in eukaryotic cells (Chen and Sarnow, 1995), and the current evidence also shows that other types of noncoding RNAs can initiate protein synthesis, raising the question of whether endogenous circRNA can encode proteins in mammalian cells (Magny et al., 2013; Lauressergues et al., 2015; Nelson et al., 2016). At present, the molecular mechanisms of circRNAs are miRNA sponge adsorption, ceRNA mechanisms, and recently discovered and confirmed coding polypeptide effects in some circRNAs and lncRNAs. For example, circFBXW7, found in glioma, can encode small peptides and inhibit the growth and replication of glioma, and peptides encoded by a segment of lncRNA are related to SERCA activity in muscle. These findings greatly expand the possible mechanism of circRNAs and overturn the belief that circRNAs cannot be translated as noncoding RNAs. Protein, as the basic component of life, is an important form of biological function.

In this study, we detected the expression level of circEPS15, statistically analyzed the overall survival of corresponding liver cancer patients and selected the effective cutoff value, divided circEPS15 into a high expression group and a low expression group, and found that this molecule from eps15 was related to the overall survival of patients. This conclusion suggests that

circEPS15 can be used as a potential biomarker for prognostic analysis and monitoring of HCC patients.

In this study, we used microarray and qRT-PCR analyses, combined with *in vivo* experiments, to determine the relationship between the downregulation of circEPS15 expression in HCC and the prognosis of HCC. The downregulation of circEPS15 expression in HCC tissues was significantly correlated with tumor differentiation and intravascular tumor thrombus, suggesting that circEPS15 has an antitumor effect and may play an important role in the epithelial-mesenchymal transition of HCC. This hypothesis has been verified *in vivo* and *in vitro*. In addition, we demonstrated that circEPS15 can regulate mRNA through the ceRNA regulatory network and has the ability to encode proteins.

CircEPS15 is derived from the parent gene eps15. Therefore, we need to understand eps15 and its functions before elucidating the functions of circEPS15. Eps15 is transcribed from ch38. p13, which encodes a component of the EGFR pathway (Niehof and Borlak, 2008), exists in the pits of the reticular protein coating, and participates in the receptor-mediated endocytosis of epidermal growth factor (EGF) (Morgan et al., 2003). In addition, alternative splicing leads to multiple transcriptional variants that encode different subtypes (Kieken et al., 2010). In addition to affecting cell migration and cancer-related signal transduction, eps15 is also associated with poor prognosis in various tumor types, including breast cancer, esophageal squamous cell carcinoma, and glioma (Smith et al., 2000; Li et al., 2013; Shi et al., 2015). Previous studies have shown that edh2 (eps15 homologous domain contains 2) may inhibit the migration and invasion of hepatocellular carcinoma through interactions with E-cadherin (Liu et al., 2016). However, the underlying mechanism by which eps15 inhibits tumors remains unclear. Consistent with previous studies, the mRNA expression of eps15 was significantly downregulated in HCC tissues compared to matched nontumor tissues, while the eps15 mRNA level in HCC tissues was positively correlated with that of circEPS15. These observations suggest that the expression of circEPS15 in HCC may be mainly regulated by host gene amplification. Bai et al. (2018) and He et al. (2017) reported that circFBLIM1 and circGRFA1 can regulate the corresponding linear fblim1 and grfa1 mRNA transcriptional expression through a ceRNA mechanism. Therefore, circEPS15 may regulate the transcriptional expression of the corresponding linear eps15 mRNA through the abovementioned ceRNA mechanism. This hypothesis will need to be further tested, proven and discussed in future research.

In this study, we found that circEPS15 could encode a protein *in vitro* by matching the RNA deep sequencing data with circaDB. Given the key role of the eps15 gene in tumorigenesis, we hypothesized that the eps15 subtype with the same function and the shorter and more stable translated protein of circEPS15 may be key regulators of cell proliferation. As this study is limited to cultured cells, the clinical importance of circEPS15 needs to be studied in a larger clinical cohort. In addition, this study confirmed the translatability of circEPS15, but the structure and function of the peptide encoded by circEPS15 need to be further verified by synthesizing specific antibodies, which are expected in subsequent studies. As an important biological regulator, circRNAs have been found to encode polypeptides in gliomas. However, no translatable circRNAs or their products have been reported in HCC.

CONCLUSION

A novel circRNA named circEPS15 was found in HCC, and its expression level was related to the overall survival of HCC patients. Our observations of the expression and activity of circEPS15 suggest that circEPS15 and its translatability may play some role in the progression of HCC and the clinical prognosis of patients.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: GSE189043.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Peking University Third Hospital Medical Science Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YB and CY performed the described studies, analyzed data, and prepared the manuscript. MT, GL and BJ provided clinical and experiments suggestions. MT, GL, BJ, and AS provided materials and assisted in experiments. DX and CY advised on study design and prepared the manuscript. All authors have read and approved the manuscript.

FUNDING

The research was supported by grants from the National Natural Science Foundation of China (No. 81702855). Special research project of Wu Jieping Medical Foundation (No.320.6750.2021-11-13) and Youth incubation project of the Peking University Third Hospital (No.BYSYFY2021040).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.804848/full#supplementary-material

REFERENCES

Bai, N., Peng, E., Qiu, X., Lyu, N., Zhang, Z., Tao, Y., et al. (2018). circFBLIM1 Act as a ceRNA to Promote Hepatocellular Cancer Progression by Sponging miR-346. *J. Exp. Clin. Cancer Res.* 37, 172. doi:10.1186/s13046-018-0838-8

Bing, Y., Zhu, S., Yu, G., Li, T., Liu, W., Li, C., et al. (2014). Glucocorticoid-induced S-Adenosylmethionine Enhances the Interferon Signaling Pathway by Restoring STAT1 Protein Methylation in Hepatitis B Virus-Infected Cells. *J. Biol. Chem.* 289 (47), 32639–32655. doi:10.1074/jbc.m114.589689

Chen, C.-Y., and Sarnow, P. (1995). Initiation of Protein Synthesis by the Eukaryotic Translational Apparatus on Circular RNAs. *Science* 268 (5209), 415–417. doi:10.1126/science.7536344

Chen, X., Fan, S., and Song, E. (2016). Noncoding RNAs: New Players in Cancers. *Adv. Exp. Med. Biol.* 927, 1–47. doi:10.1007/978-981-10-1498-7_1

Gablo, N. A., Prochazka, V., Kala, Z., Slaby, O., and Kiss, I. (2019). Cell-free microRNAs as Non-invasive Diagnostic and Prognostic Bio- Markers in Pancreatic Cancer. *Curr. Genomics* 20 (8), 569–580. doi:10.2174/1389202921666191217095017

Gao, Y.-L., Zhang, M.-Y., Xu, B., Han, L.-J., Lan, S.-F., Chen, J., et al. (2017). Circular RNA Expression Profiles Reveal that Hsa_circ_0018289 Is Up-Regulated in Cervical Cancer and Promotes the Tumorigenesis. *Oncotarget* 8 (49), 86625–86633. doi:10.18632/oncotarget.21257

Garcia, D. M., Baek, D., Shin, C., Bell, G. W., Grimson, A., and Bartel, D. P. (2011). Weak Seed-Pairing Stability and High Target-Site Abundance Decrease the Proficiency of Lsy-6 and Other microRNAs. *Nat. Struct. Mol. Biol.* 18 (10), 1139–1146. doi:10.1038/nsmb.2115

General Assembly of the World Medical Association (2014). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *J. Am. Coll. Den* 81 (3), 14–18.

He, R. F., Liu, P., Xie, X., Zhou, Y., Liao, Q., Xiong, W., et al. (2017). circGFRA1 and GFRA1 Act as ceRNAs in Triple Negative Breast Cancer by Regulating miR-34a. *J. Exp. Clin. Cancer Res.* 36, 145. doi:10.1186/s13046-017-0614-1

He, Y., Meng, X.-M., Huang, C., Wu, B.-M., Zhang, L., Lv, X.-W., et al. (2014). Long Noncoding RNAs: Novel Insights into Hepatocelluar Carcinoma. *Cancer Lett.* 344 (1), 20–27. doi:10.1016/j.canlet.2013.10.021

Hsiao, K.-Y., Lin, Y.-C., Gupta, S. K., Chang, N., Yen, L., Sun, H. S., et al. (2017). Noncoding Effects of Circular RNA CCDC66 Promote Colon Cancer Growth and Metastasis. *Cancer Res.* 77 (9), 2339–2350. doi:10.1158/0008-5472.can-16-1883

Huang, M., He, Y.-R., Liang, L.-C., Huang, Q., and Zhu, Z.-Q. (2017). Circular RNA Hsa_circ_0000745 May Serve as a Diagnostic Marker for Gastric Cancer. *World. J. Gastroenterol.* 23 (34), 6330–6338. doi:10.3748/wjg.v23.i34.6330

Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., et al. (2013). Circular RNAs Are Abundant, Conserved, and Associated with ALU Repeats. *RNA* 19 (2), 141–157. doi:10.1261/rna.035667.112

Jost, I., Shalamova, L. A., Gerresheim, G. K., Niepmann, M., Bindereif, A., and Rossbach, O. (2018). Functional Sequestration of microRNA-122 from Hepatitis C Virus by Circular RNA Sponges. *RNA Biol.* 15 (8), 1032–1039. doi:10.1080/15476286.2018.1435248

Kieken, F., Sharma, M., Jović, M., Giridharan, S. S. P., Naslavsky, N., Caplan, S., et al. (2010). Mechanism for the Selective Interaction of C-Terminal Eps15 Homology Domain Proteins with Specific Asn-Pro-Phe-Containing Partners. *J. Biol. Chem.* 285 (12), 8687–8694. doi:10.1074/jbc.m109.045666

Kristensen, L. S., Jakobsen, T., Hager, H., and Kjems, J. (2021). The Emerging Roles of circRNAs in Cancer and Oncology. *Nat. Rev. Clin. Oncol.* Epub ahead of print. doi:10.1038/s41571-021-00585-y

Lauressergues, D., Couzigou, J.-M., Clemente, H. S., Martinez, Y., Dunand, C., Bécard, G., et al. (2015). Primary Transcripts of microRNAs Encode Regulatory Peptides. *Nature* 520 (7545), 90–93. doi:10.1038/nature14346

Li, C., Zhao, L., Jiang, W., Che, L., and Xu, Y. (2018). Correct Microarray Analysis Approaches in 'Hsa-circRNA11783-2 in Peripheral Blood Is Correlated with Coronary Artery Disease and Type 2 Diabetes Mellitus'. *Diabetes Vasc. Dis. Res.* 15 (1), 92–93. doi:10.1177/1479164117739435

Li, F., Zhang, L., Li, W., Deng, J., Zheng, J., An, M., et al. (2015). Circular RNA ITCH Has Inhibitory Effect on ESCC by Suppressing the Wnt/β-Catenin Pathway. *Oncotarget* 6 (8), 6001–6013. doi:10.18632/oncotarget.3469

Li, H., Li, K., Lai, W., Li, X., Wang, H., Yang, J., et al. (2018). Comprehensive Circular RNA Profiles in Plasma Reveals that Circular RNAs Can Be Used as Novel Biomarkers for Systemic Lupus Erythematosus. *Clinica Chim. Acta* 480, 17–25. doi:10.1016/j.cca.2018.01.026

Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and Protein-RNA Interaction Networks from Large-Scale CLIP-Seq Data. *Nucleic Acids Res.* 42, D92–D97. doi:10.1093/nar/gkt1248

Li, J., Jin, B., Wang, T., Li, W., Wang, Z., Zhang, H., et al. (2019). Serum microRNA Expression Profiling Identifies Serum Biomarkers for HCV-Related Hepatocellular Carcinoma. *Cancer Biomark.* 26 (4), 501–512. doi:10.3233/cbm-181970

Li, M., Yang, X., Zhang, J., Shi, H., Hang, Q., Huang, X., et al. (2013). Effects of EHD2 Interference on Migration of Esophageal Squamous Cell Carcinoma. *Med. Oncol.* 30 (1), 396. doi:10.1007/s12032-012-0396-4

Liu, J., Ni, W., Qu, L., Cui, X., Lin, Z., Liu, Q., et al. (2016). Decreased Expression of EHD2 Promotes Tumor Metastasis and Indicates Poor Prognosis in Hepatocellular Carcinoma. *Dig. Dis. Sci.* 61 (9), 2554–2567. doi:10.1007/s10620-016-4202-6

Lu, T., Cui, L., Zhou, Y., Zhu, C., Fan, D., Gong, H., et al. (2015). Transcriptome-wide Investigation of Circular RNAs in rice. *RNA* 21 (12), 2076–2087. doi:10.1261/rna.052282.115

Lv, J., Zhao, H.-P., Dai, K., Cheng, Y., Zhang, J., and Guo, L. (2020). Circulating Exosomal miRNAs as Potential Biomarkers for Barrett's Esophagus and Esophageal Adenocarcinoma. *World. J. Gastroenterol.* 26 (22), 2889–2901. doi:10.3748/wjg.v26.i22.2889

Magny, E. G., Pueyo, J. I., Pearl, F. M. G., Cespedes, M. A., Niven, J. E., Bishop, S. A., et al. (2013). Conserved Regulation of Cardiac Calcium Uptake by Peptides Encoded in Small Open Reading Frames. *Science* 341 (6150), 1116–1120. doi:10.1126/science.1238802

Melo, S. A., Luecke, L. B., Kahlert, C., Fernandez, A. F., Gammon, S. T., Kaye, J., et al. (2015). Glypican-1 Identifies Cancer Exosomes and Detects Early Pancreatic Cancer. *Nature* 523 (7559), 177–182. doi:10.1038/nature14581

Morgan, J. R., Prasad, K., Jin, S., Augustine, G. J., and Lafer, E. M. (2003). Eps15 Homology Domain-NPF Motif Interactions Regulate Clathrin Coat Assembly during Synaptic Vesicle Recycling. *J. Biol. Chem.* 278 (35), 33583–33592. doi:10.1074/jbc.m304346200

Morishita, A., Fujita, K., Iwama, H., Chiyo, T., Fujihara, S., Oura, K., et al. (2020). Role of microRNA-210-3p in Hepatitis B Virus-Related Hepatocellular Carcinoma. *Am. J. Physiol. Gastrointest. Liver Physiol.* 318 (3), G401–G409. doi:10.1152/ajpgi.00269.2019

Nelson, B. R., Makarewich, C. A., Anderson, D. M., Winders, B. R., Troupes, C. D., Wu, F., et al. (2016). A Peptide Encoded by a Transcript Annotated as Long Noncoding RNA Enhances SERCA Activity in Muscle. *Science* 351 (6270), 271–275. doi:10.1126/science.aad4076

Niehof, M., and Borlak, J. (2008). EPS15R, TASP1, and PRPF3 Are Novel Disease Candidate Genes Targeted by HNF4α Splice Variants in Hepatocellular Carcinomas. *Gastroenterology* 134 (4), 1191–1202. doi:10.1053/j.gastro.2008.01.027

Olofson, A. M., Gonzalo, D. H., Chang, M., and Liu, X. (2018). Steatohepatitic Variant of Hepatocellular Carcinoma: A Focused Review. *Gastroenterol. Res.* 11 (6), 391–396. doi:10.14740/gr1110

Phelps, M., Coss, C., Wang, H., and Cook, M. (2016). Registered Report: Coding-independent Regulation of the Tumor Suppressor PTEN by Competing Endogenous mRNAs. *Elife* 5, e12470. doi:10.7554/eLife.12470

Rajappa, A., Banerjee, S., Sharma, V., and Khandelia, P. (2020). Circular RNAs: Emerging Role in Cancer Diagnostics and Therapeutics. *Front. Mol. Biosci.* 7, 577938. doi:10.3389/fmolb.2020.577938

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA Hypothesis: the Rosetta Stone of a Hidden RNA Language? *Cell* 146 (3), 353–358. doi:10.1016/j.cell.2011.07.014

Salzman, J., Gawad, C., Wang, P. L., Lacayo, N., and Brown, P. O. (2012). Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS One* 7 (2), e30733. doi:10.1371/journal.pone.0030733

Shi, Y., Liu, X., Sun, Y., Wu, D., Qiu, A., Cheng, H., et al. (2015). Decreased Expression and Prognostic Role of EHD2 in Human Breast Carcinoma:

Correlation with E-Cadherin. *J. Mol. Hist.* 46 (2), 221–231. doi:10.1007/s10735-015-9614-7

Smith, J. S., Tachibana, I., Pohl, U., Lee, H. K., Thanarajasingam, U., Portier, B. P., et al. (2000). A Transcript Map of the Chromosome 19q-Arm Glioma Tumor Suppressor Region. *Genomics* 64 (1), 44–50. doi:10.1006/geno.1999.6101

Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L.-H., et al. (2015). Exon Circularization Requires Canonical Splice Signals. *Cel Rep.* 10 (1), 103–111. doi:10.1016/j.celrep.2014.12.002

Tam, W. L., and Weinberg, R. A. (2013). The Epigenetics of Epithelial-Mesenchymal Plasticity in Cancer. *Nat. Med.* 19 (11), 1438–1449. doi:10.1038/nm.3336

Weng, W., Wei, Q., Toden, S., Yoshida, K., Nagasaka, T., Fujiwara, T., et al. (2017). Circular RNA ciRS-7-A Promising Prognostic Biomarker and a Potential Therapeutic Target in Colorectal Cancer. *Clin. Cancer Res.* 23 (14), 3918–3928. doi:10.1158/1078-0432.ccr-16-2541

Xie, G., and Diehl, A. M. (2013). Evidence for and against Epithelial-To-Mesenchymal Transition in the Liver. *Am. J. Physiol. Gastrointest. Liver Physiol.* 305 (12), G881–G890. doi:10.1152/ajpgi.00289.2013

Yang, Y., Gao, X., Zhang, M., Yan, S., Sun, C., Xiao, F., et al. (2018). Novel Role of FBXW7 Circular RNA in Repressing Glioma Tumorigenesis. *J. Natl. Cancer Inst.* 110 (3), 304–315. doi:10.1093/jnci/djx166

Zeng, Y., Du, W. W., Wu, Y., Yang, Z., Awan, F. M., Li, X., et al. (2017). A Circular RNA Binds to and Activates AKT Phosphorylation and Nuclear Localization Reducing Apoptosis and Enhancing Cardiac Repair. *Theranostics* 7 (16), 3842–3855. doi:10.7150/thno.19764

Zhao, J., Wu, J., Xu, T., Yang, Q., He, J., and Song, X. (2018). IRESfinder: Identifying RNA Internal Ribosome Entry Site in Eukaryotic Cell Using Framed K-Mer Features. *J. Genet. Genomics* 45 (7), 403–406. doi:10.1016/j.jgg.2018.07.006

Zhao, Y., Alexandrov, P. N., Jaber, V., and Lukiw, W. J. (2016). Deficiency in the Ubiquitin Conjugating Enzyme UBE2A in Alzheimer's Disease (AD) Is Linked to Deficits in a Natural Circular miRNA-7 Sponge (circRNA; ciRS-7). *Genes (Basel)* 7 (12), 116. doi:10.3390/genes7120116

Zheng, Q., Bao, C., Guo, W., Li, S., Chen, J., Chen, B., et al. (2016). Circular RNA Profiling Reveals an Abundant circHIPK3 that Regulates Cell Growth by Sponging Multiple miRNAs. *Nat. Commun.* 7, 11215. doi:10.1038/ncomms11215

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 10 March 2022
doi: 10.3389/fgene.2022.787734

Dysregulation of the HOTAIR-miR-152-CAMKIIα Axis in Craniosynostosis Results in Impaired Osteoclast Differentiation

*Chenbin Dong, Xiangqi Liu, Jun Li, Dongyi Lan and Shan Zheng**

Department of Plastic Surgery, Children's Hospital of Fudan University, Shanghai, China

Craniosynostosis is one of the most common craniofacial deformities demanding surgical treatment in infancy. LncRNA HOTAIR has verified its important role in osteogenesis and osteoarthritis. However, whether HOTAIR plays an essential role in the development of craniosynostosis is still unclear. In this study, we aimed to investigate the molecular role of HOTAIR in the osteoclast function and development of craniosynostosis.For osteoclast differentiation, RAW264.7 cells were induced by 50 ng/ml of RANKL and 10 ng/mL M-CSF, followed by TRAP staining. Cell proliferation and apoptosis were assayed by the CCK-8 kit and Annexin V-FITC apoptosis detection kit, respectively. The expression of HOTAIR was determined in PBMCs by qRT-PCR. Protein levels of all those involved genes were measured by Western blot assay. A luciferase reporter assay was used to determine the miRNA target validation. The HOTAIR expression in PBMCs from children with craniosynostosis was significantly downregulated. The results of cell proliferation and apoptosis assays indicated that silencing of HOTAIR could inhibit osteoclast differentiation and increase cell apoptosis. Moreover, the luciferase reporter assay revealed that the regulatory axis and HOTAIR-miR-152-CAMKIIα were the regulatory mechanisms of HOTAIR in the osteoclast function and development of craniosynostosis.In this study, our data showed that HOTAIR could promote osteoclast differentiation by binding miR-152. Furthermore, the HOTAIR/HOTAIR-miR-152-CAMKIIα axis was found to regulate osteoclast differentiation. These results indicate that the HOTAIR plays a crucial role in the development of osteoclasts.

Keywords: HOTAIR, miRNA-152, osteoclast, craniosynostosis impact statement, long nocoding RNA

OPEN ACCESS

Edited by:
Ramkrishna Mitra,
Thomas Jefferson University,
United States

Reviewed by:
Antonio Casado Díaz,
Centro de Investigación Biomédica en Red sobre Fragilidad y Envejecimiento Saludable (CIBERFES), Spain
Somnath Tagore,
Columbia University, United States

****Correspondence:***
Shan Zheng
szheng@shmu.edu.cn

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *01 October 2021*
Accepted: *21 February 2022*
Published: *10 March 2022*

Citation:
Dong C, Liu X, Li J, Lan D and Zheng S (2022) Dysregulation of the HOTAIR-miR-152-CAMKIIα Axis in Craniosynostosis Results in Impaired Osteoclast Differentiation. Front. Genet. 13:787734. doi: 10.3389/fgene.2022.787734

INTRODUCTION

Craniosynostosis is a heterogeneous disease defined as the premature fusion of cranial sutures. It is estimated that one in 2000–2,500 live births are affected (French et al., 1990; Singer et al., 1999; Hukki et al., 2008). This developmental abnormality leads to head strabismus and craniofacial asymmetry in children. It causes permanent neurological, eye, and respiratory dysfunction due to the inability of the cranial crest to adapt to physiological brain growth (Renier et al., 1982; Maliepaard et al., 2014). In recent decades, great progress has been made in technology and safety. Cranial vault reconstruction is still the main method for treating craniosynostosis, with a high incidence of complications and high mortality (Sloan et al., 1997; Clayman et al., 2007; Czerwinski et al., 2010). A comprehensive understanding of the basis of skull sutures is beneficial for the prevention, diagnosis, and treatment of the disease.

A variety of genetic and environmental factors have been reported to be associated with craniosynostosis. Specific single-gene mutations or chromosomal abnormalities account for at least 20% of all cases (Lajeunie et al., 1995; Wilkie, 1997; Singer et al., 1999; Boulet et al., 2008). The identified gene mutations mainly involve upregulation of osteogenesis, downregulation of osteoclast formation, cell proliferation and apoptosis, cell patterning, extracellular matrix regulation, or vascular function (Ma et al., 1996; Muenke et al., 1997; Wilkie, 1997; Fitzpatrick, 2013). Osteoclasts are cells that degrade bone to initiate normal bone remodeling. They mediate bone loss in pathologic conditions by increasing their resorptive activity (Boyce et al., 2009). Besides, skull sutures attain their complex shape at the same age when osteoclast number is highest along concave suture margins (Byron, 2006). Osteogenic dysfunction has been described as a major cause of premature ossification and fusion of the skull (Cohen, 2009). Bone homeostasis depends on a delicate balance between osteoclast and osteoblast-mediated bone resorption and bone deposition, respectively. Osteoclasts are desorbing bone cells derived from macrophage progenitor cells of bone marrow. This plays an essential role in maintaining the stability of the bone's internal environment.

Osteoclast activation mediated by the receptor activator of nuclear factor-kappa B (RANK)/RANK-ligand (RANKL) signaling plays a critical role in maintaining human cranial suture patency. Dysregulation of osteoclast differentiation and death contribute to premature suture fusion, leading to craniosynostosis (Lyon et al., 2018). Like osteoblast inhibition, osteoclast activation is considered a potential anti-craniosynostosis strategy (Beederman et al., 2014). However, the regulation of osteoclast differentiation in cranial suture biology remains largely unknown.

Long non-coding RNAs (lncRNAs) are a class of non-coding RNA molecules discovered in recent years. They are widely involved in many physiological and pathological processes, such as osteoblast and osteoclast-mediated bone remodeling. HOTAIR is the first trans-acting lncRNA abnormally overexpressed in various tumor tissues and cell lines. Microarray analysis of Xing et al. showed that HOTAIR was expressed in OA cartilage, and its level was higher than that of normal samples (Xing et al., 2014). In addition, it is revealed that the crosstalk between lncRNAs and mRNAs occurs through the competitive binding of microRNAs response elements (Cesana et al., 2011). A recent study showed that HOTAIR could act as a competitive endogenous RNA (ceRNA) for miR-331-3p and miR-124 and regulate their cellular level (Susiarjo et al., 2007). In addition, it is involved in the regulation of the expression of miR-17-5p in osteogenesis. In addition to physical association, HOTAIR could regulate microRNAs levels through direct recognition and target degradation (Huang et al., 2014). Moreover, bone homeostasis depends on the resorption of bones by osteoclasts and the formation of bones by the osteoblasts. Misawa et al., in their study, revealed that HOTAIR inhibits mineralization in osteoblasts. Wei et al. revealed that silencing of HOTAIR promotes osteoblast differentiation by upregulating miR-17-5p expression (Wei et al., 2017; Misawa and Orimo, 2018). However, it remains unclear whether HOTAIR plays a role in osteoclast differentiation.

MicroRNAs (miRNAs) are a class of non-coding RNAs that play essential roles in regulating gene expression. Based on the current state of knowledge, researches on miRNAs in craniosynostosis are scanty. However, Misra et al. stated that transcription factors and microRNA are associated with craniosynostosis (Misra et al., 2019). Through Bioinformatics prediction and RIP analysis, we found that HOTAIR could directly combine with miR-152. Feng et al. found that miR-152 can regulate osteoblast differentiation and influence osteoporosis (Feng et al., 2019). Ma et al. reported that miR-152-3p promotes osteoclastogenesis by targeting osteoclastogenic regulator MAFB (Ma et al., 2021). Moreover, calmodulin is an important regulator of osteoclast differentiation, function, and survival (Seales et al., 2006). It activates multifunctional CAMKs, a serine/threonine-protein kinase family, including CAMKI, CAMKII, and CAMKIV (Soderling and Stull, 2001). It has been reported that the CAMK-CREB pathway regulates osteoclast differentiation and function (Sato et al., 2006). In addition, through the Bioinformatics prediction (TargetScan, PicTar, and miRanda), we determined that $Ca2^{+}$/calmodulin dependent kinases 2-α (CAMKIIα) were the direct target of miR-152. CaMKIIα is an essential mediator of activity-dependent synaptic plasticity that possesses multiple protein functions. It is a critical gate controlling structural, functional, and behavioral expression of synaptic memory. CaMKIIα may result in brain hypoplasia, and craniofacial anomalies included craniosynostosis (Yamagata et al., 2009). Therefore, these data suggest that the HOTAIR-miR-152-CAMKIIα regulatory axis may be involved in osteoclastogenesis during the occurrence and development of cranial suture injury. This study aimed to investigate whether HOTAIR plays a regulatory role in osteoclast differentiation in craniosynostosis development through the miR-152/CAMKII α axis.

MATERIALS AND METHODS

Experiments on Human Study Subjects

The peripheral blood samples of children with craniosynostosis and healthy controls were collected from Children's Hospital of Fudan University. Peripheral blood mononuclear cells (PBMCs) were obtained from each subject. All patients have signed informed consent. This study has been approved by the research ethics committee of Children's Hospital of Fudan University (No. 2020-143). The baseline characteristics of the patients are shown in **Table 1**.

Cell Culture and Transfection

RAW264.7 cells as osteoclast precursors were obtained from the Cell Bank, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco, United States) supplemented with 10% fetal bovine serum

TABLE 1 | Baseline data of the patients.

No	Gender	Age at operation (month)	Weight at operation (kg)	Craniosynostosis type	Closed cranial suture	Open cranial suture
1	Female	8	8.5	simplex	sagittal suture	Coronal suture on the right
2	Male	16	10	Crouzon syndrome	Bilateral coronal sutures	sagittal suture
3	Male	24	14.5	Apert syndrome	Left temporal squamous suture, left coronal suture	sagittal suture
4	Female	22	11	Apert syndrome	Bilateral coronal suture, right herringbone suture	sagittal suture
5	Female	36	15	simplex	frontal seam	Coronal suture on the right
6	Female	8	9.5	simplex	Coronal suture on the right	sagittal suture
7	Female	9	9	simplex	sagittal suture	Coronal suture on the right
8	Female	21	13	Crouzon syndrome	Whole cranial seam	—
9	Male	47	16.5	Apert syndrome	Whole cranial seam	—
10	Male	6	8	Saethre-Chotzen syndrome	Coronal suture on the right	sagittal suture
11	Male	13	11	simplex	sagittal suture	Coronal suture on the right
12	Female	8	7.5	simplex	Coronal suture on the left	sagittal suture
13	Male	26	12	Crouzon syndrome	Whole cranial seam	—

	Craniosynostosis (n = 13)	Ctrl (n = 26)
male / female, (n)	6/7	13/13
Age, (y)	1.56 ± 0.99	8.38 ± 3.13
weight, (kg)	11.19 ± 2.74	N/A
Syndrome type/simple type, (n)	7/6	N/A

(FBS, Gibco, United States). Cells were then incubated at 37°C in a humidified chamber supplemented with 5% CO2. Small interfering RNAs (siRNAs), miR-152 mimics, inhibitors, and negative controls (NC) were purchased from GenePharma (Shanghai, China). They were transfected into cells using Lipofectamine RNAiMAX (Invitrogen), according to the manufacturer's instructions. The sequences of siRNAs, mimics, and controls are provided as follows: HOTAIR siRNA1, 5′-UAA CAA GAC CAG AGA GCU GTT-3' (sense); HOTAIRsiRNA2, 5′-GCA CAG AGC AAC UCU AUA ATT' (sense); CAMKIIα, 5′-TTGTGGCCCGGGAGT ATTACAGT′-3' (sense); miR-152 mimics, 5′-UCAGUG CAUGACAGAACUUGG-3' (sense).

Osteoclast Differentiation and Tartrate-Resistant Acid Phosphatase (TRAP) Staining

For osteoclast differentiation, RAW264.7 cells (5×10^3) cells were cultured in the presence of 50 ng/ml of RANKL (#462-TEC; R&D Systems, United States) and 10 ng/mL M-CSF (#416-ML; R&D Systems). Furthermore, for TRAP staining, cells were first washed twice with phosphate-buffered saline (PBS). Then, they were fixed in 4% paraformaldehyde (pH 7.4) at room temperature for 15 min and stained using leukocyte Acid phosphatase Kit 387-A (Sigma-Aldrich, St. Louis, MO, United States) according to the manufacturer's protocol.

Cell Apoptosis Assay

According to the manufacturer's instructions, Annexin V-staining was performed using an Annexin5-FITC apoptosis detection kit (BD Biosciences, United States). Briefly, cells were stained with Annexin5-FITC and propidium iodide in binding buffer for 15 min at 37°C in the dark. The samples were analyzed using a FACScan flow cytometer (BD Biosciences, United States).

Cell Proliferation Assay

Cells (1×10^3) were plated in 96-well plates until they reached 30–40% confluence, and 10 μL CCK-8 solutions (Dojindo, Japan) was added to each well. Cells were incubated at 37°C for 2 h. Then, a microplate spectrophotometer (BioTek Synergy HTX Multi-mode Reader, United States) was used to read the OD at 450 nm. Each sample was assessed at 24, 48, 72 h.

Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

According to the manufacturer's instructions, RNA was extracted from cells using TRIzol Reagent (Invitrogen, United States). cDNA was synthesized using the PrimeScript RT Reagent Kit (Takara, Japan). qRT-PCR was performed using the SYBR Premix Ex Taq RT-PCR kit (Takara, Japan) on LightCycler 480 II (Roche, Switzerland). The expression of miR-152 was quantified using a TaqMan miRNA assay kit (Life Technologies, United States). GAPDH and U6 expression levels were determined as internal controls. Primer sequences for all genes are shown in **Table 2**.

Western Blot Assay

The membranes were blocked using 5% BSA in TBST (#ST023; Beyotime, Shanghai, China) at room temperature for 1 h. Western blot assays were performed using antibodies directed against phosphorylated ERK 1/2, ERK 1/2, phosphorylated p65, p65, CAMKIIα (1: 1000; Cell Signaling Technology, United States), and β-Tubulin (1: 5,000; Abcam, UK). Correspondingly, β-Tubulin was used as a loading

TABLE 2 | Primer sequences for all genes.

Primer name	Forward	Reverse
Human GAPDH	5′-GTCTCCTCTGACTTCAACAGCG-3′	5′-ACCACCCTGTTGCTGTAGCCAA-3′
Mouse GAPDH	5′-CAGAACATCATCCCTGCATC-3′	5′-GCAGAGCCCTTTTTGATAATGT-3′
Human HOTAIR	5′-CAGTGGGGAACTCTGACTCG-3′	5′-GTGCCTGGTGCT CTCTTACC-3′
Mouse HOTAIR	5′-CAGTGGGGAACTCTGACTCG-3′	5′-GTGCCTGGTGCTCTCTTACC-3′
Mouse NFATc1	5′-GGTGCCTTTTGCGAGCAGTATC-3′	5′-CGTATGGACCAGAATGTGACGG-3′
Mouse Cathepsin K	5′-AGCAGAACGGAGGCATTGACTC-3′	5′-CCCTCTGCATTTAGCTGCCTTTG-3′
Mouse TRAP	5′-GCGACCATTGTTAGCCACATACG-3′	5′-CGTTGATGTCGCACAGAGGGAT-3′
Mouse RANK	5′-GGACAACGGAATCAGATGTGGTC-3′	5′-CCACAGAGATGAAGAGGAGCAG-3′
Mouse CAMKIIα	5′-AGCCATCCTCACCACTATGCTG-3′	5′-GTGTCTTCGTCCTCAATGGTGG-3′

control. The membrane was incubated with primary antibodies at 4 °C overnight, followed by incubation with HRP-linked anti-rabbit IgG (#7074; Cell Signaling Technology) or anti-mouse IgG (#7076; Cell Signaling Technology) at room temperature for 1 h. The protein bands were visualized using a Pierce ECL Western blotting substrate (#32106; Thermo Fisher Scientific, United States). Images were acquired using a ChemiDoc XRS$^+$ System (Bio-Rad, United States).

RNA Immunoprecipitation Assays

RNA immunoprecipitation (RIP) was performed using the EZ-Magna RIP kit (Millipore, United States). Cell extracts were co-immunoprecipitated with AGO2 (Millipore, United States) or normal rabbit IgG. The retrieved RNA was subjected to RT-qPCR analysis. On the day before transfection, RAW264.7 cells were inoculated in 10 cm Petri dishes with 10^6-10^7 cells in each dish and transfected with miR-152 mimics or an equivalent dose of unrelated control NC. After 48 h, removed culture medium, washed twice with PBS, and scraped cells off with a cell spatula. For each IP test using magnetic separation, 50 μL magnetic beads to 1.5 MLEP tubes, 5 μg IP antibody, RIP wash buffer, proteinase K Buffer, RIP rinse buffer, phenol-chloroform-isoamyl alcohol, chloroform, Salt Solution I 50 μL, Salt Solution II 15 μL, Precipitate Enhancer 5 μL, and Anhydrous alcohol 850 ml were used accordingly. Experiments were carried out following the manufacturer's instructions. Centrifugation at 4°C, rotation and incubation at 4°C overnight, water bath at 55°C for 30 min, and refrigeration at -80°C were performed. After completing the experiment, the tubes were then air-dried in an ultra-clean table, and the RNA was dissolved in DEPC water. The EP tubes were placed at 60°C for 3 min, and then, real-time PCR was used to detect gene expression.

Luciferase Reporter Assay

The psiCHECK-2 vector (Promega, Madison, WI) was used to clone the 3′-UTR of CAMKIIα. Primers used were as follows: CAMKIIα 3′-UTR, 5′-CCGCTCGAGTGCTTCCCTCGCAAA CT-3' (forward) and 5′- TTAGCGGCCGCTGGCTCTTCC TCCCCTAA-3' (reverse). RAW 264.7 cells were cultured in 96-well, and the cells were rinsed and then harvested 24 h after transfection. The ratio of Renilla luciferase activity to firefly luciferase activity was calculated.

Statistical Analysis

All statistical analyses were performed using GraphPad Prism 7.0 software. Differences between groups were analyzed using two-tailed Student's t-test and one-way ANOVA with post-hoc Bonferroni test. $p < 0.05$ was considered statistically significant.

Animal Experiments

Twelve two-day-old newborn mice were purchased and raised in the SPF animal room for the HOTAIR overexpressed mice model construction and detection. Two c57 female mice were breastfed together. At the age of 5 days, newborn rats were randomly divided into the experimental group (adenovirus loaded with Lnc injected locally) and control group (empty adenovirus/ADV6-NC injected locally into the skull). The animal experiments were approved by the research ethics committee of Children's Hospital of Fudan University.

The c57 Newborn Mouse Virus Injection and Sampling

The injection was administered once daily for 4 days (D5 to D8), with a virus titer of 10^9, 20 μL. After disinfecting the operation area with complexed iodine, injected horizontally from the center of the neck, back under the skin using a 10 μL micro-injector. At D9, three mice in each group were sacrificed by cervical dislocation. Their skulls were frozen and stored in liquid nitrogen to extract RNA. The remaining mice in each group continued to be kept for observation. By D20, all mice were sacrificed by cervical dislocation, photographed, and their skulls were fixed in 4% paraformaldehyde for 24 h, and then skull CT was taken. The skull specimen was embedded in paraffin and observed by HE staining. The mice were weighed every 3 days throughout the experiment.

Total RNA Extraction From C57 Neonatal Rat Skull

The skull was cut into strips, put in a 2 ml flat-bottomed EP tube, added TRIZOL 1ml, and put 2 3.0 mm magnetic beads in each tube. The grinding was carried out in a Mini bead-Beater-16 pearl-Magic Tissue Grinder with 60 Hz power and 2 min duration until no visible bone. After homogenate was obtained, subsequent RNA extraction, reverse transcription, and qRT-PCR were used to detect HOTAIR expression.

Cranial CT Scan and Data Analysis

After soaking in 4% paraformaldehyde for 24 h, the skull was placed along the long axis in a MICROCT imager (Quantum FX MicroCT). After MicroCT scanning, imported the scanned data into Analyze 12.0 (Mayo Clinical, United States) for data analysis. All the parietal bones were delimited as the target area to obtain bone density and bone mass. According to the standard curve provided by the software, the actual bone mineral density (mg/mm3) = (measured value + 2,769.7)/3.4445.

Pathological Sections and Staining of the Skull

After completing the CT scan, the skull tissue was removed from paraformaldehyde and immersed in a 10% EDTA decalcification solution. Wax block embedding was completed by the Shanghai Google Biological Co., Ltd. The wax block was taken and cooled in a refrigerator at -20°C for 30 min. Using a slicer, slice the slices to the thickness of 3 μm and bake them in the oven for 30 min. Routine dewaxing, HE staining, dehydration, and seal were performed. Microscopic examination was performed after preparation.

Statistical Analysis of Data

SPSS 18.0 software was used for statistical analysis, and GraphPad Prism 6.0 software was used for plotting. Continuous variables were compared between the two groups using unpaired Student's test. The 0.05 difference was considered statistically significant. (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

RESULTS

HOTAIR Promoted Osteoclast Differentiation

Knockdown of HOTAIR decreases the RANKL-induced osteoclast differentiation in RAW 264.7 cells

qRT-PCR results showed that HOTAIR expression was significantly lower in the craniosynostosis children's group than the control group ($p < 0.05$, **Figure 1A**). Furthermore, HOTAIR expression was significantly reduced by two specific siRNAs (**Figure 1B**). RANKL-induced RAW264.7 cells treated with siRNA negative control were able to differentiate into numerous TRAP-positive multinucleated osteoclasts. In contrast, HOTAIR siRNAs treated cells exhibited less TRAP-positive osteoclast formation (**Figure 1C,D**). Furthermore, the mRNA expression levels of osteoclast differentiation-related genes NFATc1, Cathepsin K, TRAP, and RANK were significantly downregulated in the absence of HOTAIR (**Figure 1E**).

HOTAIR Inhibits Apoptosis of Osteoclast Lineage Cells

The cell proliferation assay results indicated that knockdown of HOTAIR did not affect precursor proliferation during the early process of osteoclast differentiation (**Figure 2A**). In addition, we evaluated the influence of HOTAIR on the apoptosis of mature osteoclasts. It revealed that silencing of HOTAIR could markedly increase the frequency of both Annexin V single positive, Annexin V, and PI double-positive cell populations (**Figure 2B,C**). This indicated that more cells undergo apoptosis when lacking HOTAIR expression.

The HOTAIR-Overexpressing Mouse Model and Phenotype Detection

Overexpression of HOTAIR was verified by RT-qPCR on D9 post-injection. The relative expression of HOTAIR in the experimental group was 0.0040 ± 0.0012, whereas in the control group was 0 ($p = 0.041$). Also, local injection of HOTAIR loaded with adenovirus was performed, and skull samples were collected 1 day after injection for HOTAIR expression determination. The results showed a significant increase in HOTAIR expression in the skulls of mice in the experimental group. This suggests that local injection of adenovirus-loaded HOTAIR upregulates HOTAIR expression in the skull ($p < 0.01$). In addition, the mice in the experimental group showed a gradual slowdown in growth and development 4 days after injection with short body length, sparse hair, and dark yellow (**Figure 3A**). Bodyweight has been changed significantly between groups before administration (D5; control group (2.80 ± 0.27 g) vs experimental group (2.90 ± 0.29 g), $p = 0.68$) and at the end of observation (D20; control group (7.41 ± 0.48 g) vs. experimental group (4.84 ± 0.51 g), $p = 0.0032$) (**Figure 3B**).

Correspondingly, *in vitro* cranial CT scans (**Figure 3C**) showed that cranial sutures of the experimental group were open and wide, thin skull bones, and bone defects on the surface. On the other hand, skull bone was intact in the control group. The comparison of the coronal section showed that craniosynostosis in the control group was thicker, but narrower in the experimental group (**Figure 3C**, a4, b4). This suggests that local overexpression of HOTAIR can maintain the cranial suture opening. The cranial CT images were used to delimit the whole cranial range on the Analyze 12.0 software. A significant difference was observed in the average cranial bone mass between the control group (26.67 ± 2.51 mm (Singer et al., 1999)) and the experimental group (9.59 ± 2.08 mm (Singer et al., 1999)) ($p = 0.0008$; n = 3). The mean bone mineral density in the control group (1977 ± 27.06 mg/mm^3) was significantly higher than that in the experimental group (1841 ± 20.52 mg/mm^3) ($p = 0.0022$; n = 3). These results demonstrate that skull overexpression of HOTAIR could reduce bone formation, mass, and density (**Figure 3D**). Moreover, the HE pathological section showed that cranial bones of the posterior frontal suture were overlapped entirely and fused in the control group (**Figure 3E**, a2), with apparent gaps (**Figure 3E**). A large number of inflammatory cells could be seen infiltrating between b2 and b3, and the experimental group was thinner that lack obvious plate structure. These indicate that overexpression of HOTAIR could promote osteoclast differentiation, function and maintain cranial suture opening.

HOTAIR positively regulates osteoclast differentiation by downregulating the miR-152 expression.

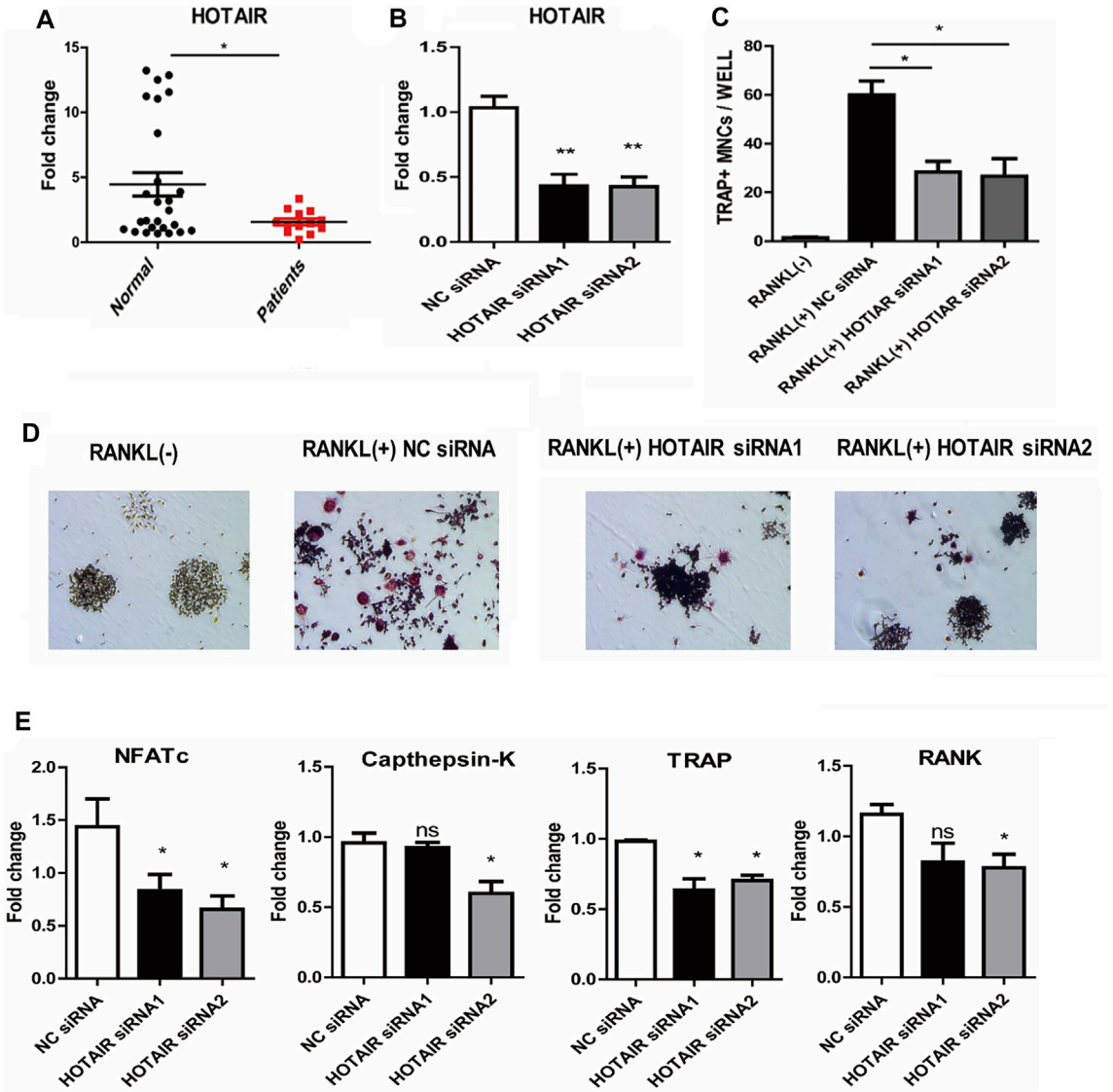


FIGURE 1 | The phenotype of HOTAIR in PBMCs of patients, and knockdown of HOTAIR decreases the RANKL-induced osteoclast differentiation in RAW 264.7 cells **(A)** The levels of HOTAIR were significantly lower in PBMCs of patients with craniosynostosis (N = 13) compared to normal people (N = 26) **(B)** Detection of HOTAIR expression in RAW 264.7 cells with HOTAIR knockdown by two distinct siRNA (200 nM), NC: negative control **(C and D)** RAW 264.7 cells were treated with RANKL and M-CSF for 5 days. Cells were fixed and stained for TRAP. TRAP$^+$ cells with more than three nuclei were counted as osteoclasts. Magnification: ×100. TRAP$^+$ cells were counted **(E)** The effects of HOTAIR knockdown on mRNA levels of osteoclast differentiation-related genes. Bars represent the mean ± SEM. $^*p < 0.05$, $^{**}p < 0.01$. The data are representative of three independent experiments.

We used starbase (http://starbase.sysu.edu.cn/) to identify the potential binding sites between HOTAIR transcript and miRNAs, such as miR-6807-3p, miR-148a/b-3p, miR-1227-5p, and miR-152 (**Figure 4A**). Among them, miR-152 expression in the craniosynostosis children's group was significantly upregulated (**Figure 4B**) compared with that in PBMCs of healthy controls. Moreover, the levels of miR-152 were notably increased in RAW264.7 cells treated with HOTAIR siRNA compared with those in cells treated with negative control siRNA (**Figure 4C**). The RIP results showed that HOTAIR was observed in miR-152-overexpressing RAW264.7 cells (**Figure 4D**), verifying the interaction between HOTAIR and miR-152. The TRAP staining results indicated that miR-152 overexpression could remarkably reduce the number of TRAP-positive multinucleated osteoclasts (**Figure 5A**). The mRNA levels of osteoclast differentiation-related genes, including *NFATc1, Cathepsin K, TRAP*, and *RANK* were significantly downregulated in miR-152 mimic-treated group (**Figure 5B**). Meanwhile, miR-152 mimic administration enhanced osteoclast apoptosis (**Figure 5C,D**). At the same time, it had little effect on precursor proliferation (**Figure 5E**), similar to the knockdown of HOTAIR data.

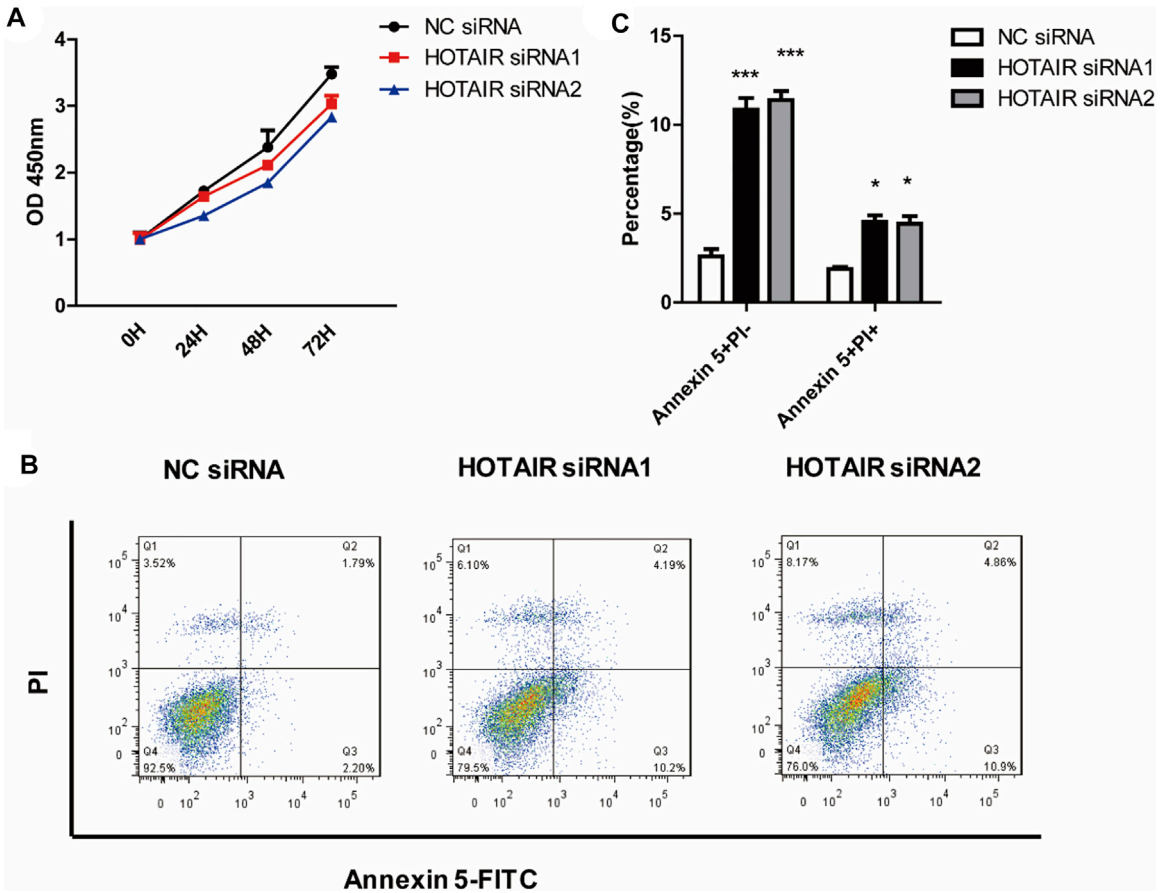


FIGURE 2 | HOTAIR suppresses the apoptosis of mature osteoclasts **(A)** RAW264.7 cells were cultured with M-CSF in the absence or presence of HOTAIR siRNA for 3 days. Then, the proliferation ability of early osteoclast precursors was analyzed by CCK-8 assay **(B)** After transfection with HOTAIR siRNA or the corresponding NCs, RAW 264.7 cells were treated with RANKL and M-CSF for 5 days and stained with Annexin 5-FITC and propidium iodide (PI). Cells were then analyzed using a flow cytometer. Early apoptotic cells were Annexin5$^+$ and PI$^-$ and late apoptotic cells were Annexin5$^+$ and PI$^+$ **(C)** Quantification of apoptosis cells was analyzed. Bars represent the mean ± SEM. $^*p < 0.05$, $^{***}p < 0.001$. The data are representative of three independent experiments.

HOTAIR-miR-152-CAMKIIαaxis Regulates Osteoclast Differentiation

qRT-PCR results showed that miR-152 overexpression or HOTAIR silencing could significantly reduce the mRNA levels of CAMKIIα in RAW264.7cells (**Figure 6A**). The luciferase assay revealed that miR-152 regulated the luciferase activity in a CAMKIIα 3′UTR-dependent manner (**Figure 6B,C**). Moreover, the results showed that HOTAIR knockdown or miR-152 overexpression could strongly attenuate p65 and MAPK/ERK 1/2 phosphorylation and protein level of CAMKIIα (**Figure 6D**). Furthermore, the mRNA expression levels of osteoclast differentiation-related genes *NFATc1, Cathepsin K, TRAP,* and *RANK* were significantly downregulated in the absence of CAMKIIα (**Figure 6E**).

DISCUSSION

In this study, our data showed that HOTAIR expression was significantly downregulated in PBMCs from children with craniosynostosis. However, approximately 50% of normal individuals had PBMC HOTAIR expression equal to or lower than patients with craniosynostosis, suggesting that low HOTAIR expression in PBMCs is related to craniosynostosis but might not be a suitable biomarker for craniosynostosis. The results of cell proliferation and apoptosis assays indicated that silencing of HOTAIR could inhibit osteoclast differentiation and increase cell apoptosis. Moreover, the luciferase reporter assay results showed that the regulatory axis, HOTAIR-miR-152-CAMKIIα, was the regulatory mechanism of HOTAIR in the osteoclast function and development of craniosynostosis. All those results implied that HOTAIR might regulate osteoclast formation to participate in the development of craniosynostosis.

There is increasing evidence that miRNAs are involved in the regulation of osteoclast differentiation and bone resorption (Tang et al., 2014). Ma et al. have shown that inhibition of miR-152-3p impairs osteoclastogenesis *in vitro* and reduces the osteolytic lesions while preserving trabecular architecture *in vivo*, suggesting that miR-152-3p promotes osteoclastogenesis (Ma et al., 2021). On the contrary, Xu et al. have found that

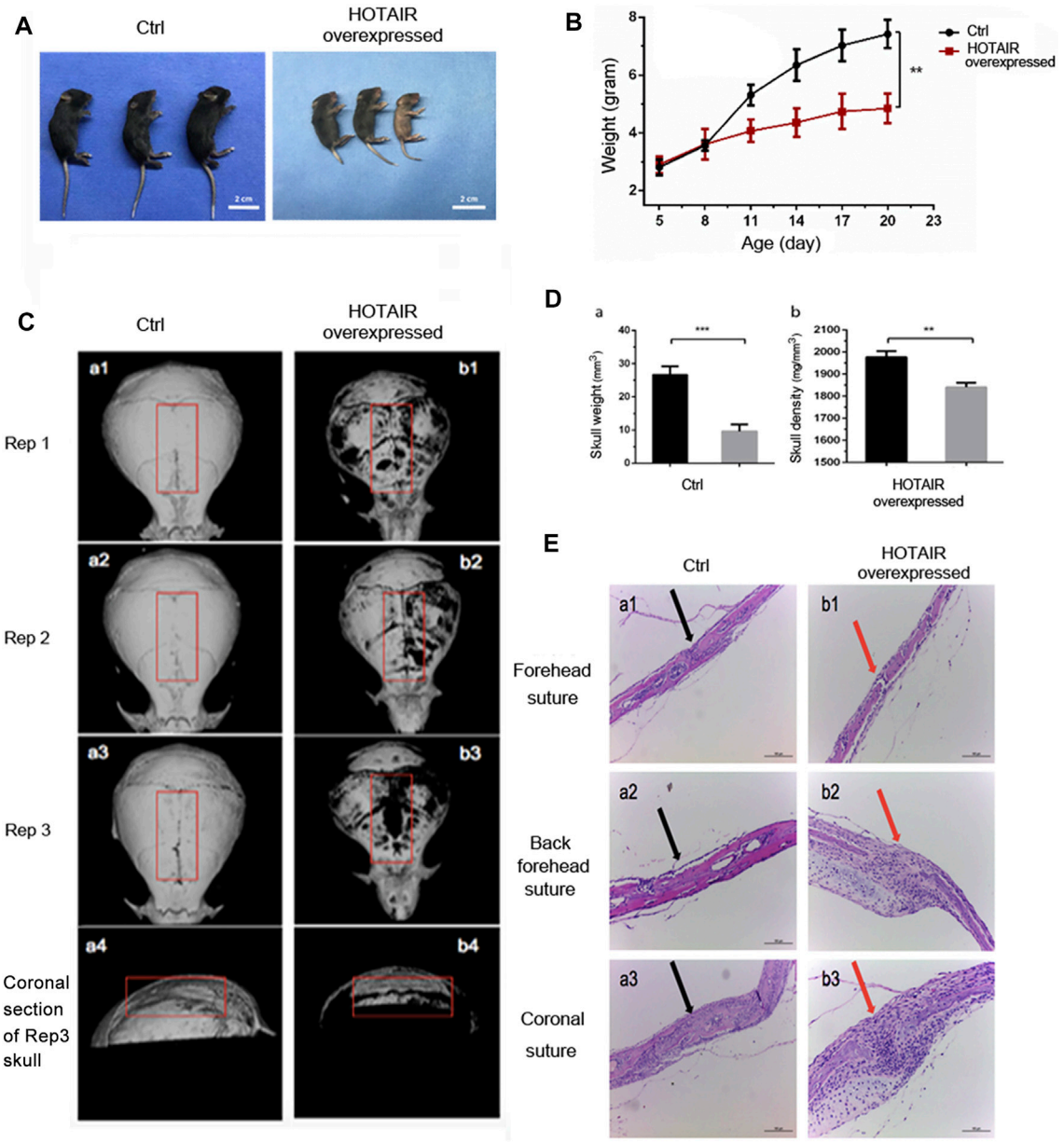


FIGURE 3 | The phenotype of HOTAIR in mouse model **(A)** The effect of HOTAIR on the physical development of C57 mice after 20 days' birth **(B)** Quantitative assay of the mice weighted along the days after birth **(C)** Images of CT scan of the mouse skull (a1-a3 show the CT scan images of the isolated skulls of three replicates of 20-day-old WT mice, a4 shows the coronal section of a3 skull. b1-b3 shows the CT scan images of the isolated skulls of three replicates of 20-day-old HOTAIR overexpressed mice, b4 show the coronal section of the b3 skull) **(D)** Quantitative assay of the weight and bone density of mouse skull with CT scan (**$p < 0.01$, ***$p < 0.001$) **(E)** HE assays of the effect of overexpression of HOTAIR on the skull sutures in mice. Arrows point to the cranial sutures of mice. The scale of the images is 150 um.

overexpression of miR-152 alleviates bone disruption in an intrabone multiple myeloma mouse model by targeting Dickkopf-1 (Xu et al., 2015). Since inhibition of Dickkopf-1 has been shown to increase osteoblastic differentiation while reducing osteoclast activity in a myelomatous mouse model (Yaccoby et al., 2007), the findings of Xu et al. suggest that miR-152 may suppress osteoclastogenesis. Consistently, our data indicate that miR-152 inhibits osteoclast formation and suggest that HOTAIR may regulate osteoclast formation by interfering with miR-152 expression and function.

The past decade has elucidated many important mechanisms in the biology of skull suture, including osteoblast dysfunction and regional dura involvement in regulating suture fusion (Bradley et al., 1997; Levine et al., 1998). These studies have become the basis for the development of molecular therapy. However, bone biology depends on the interaction between osteoblasts and osteoclasts. Therefore,

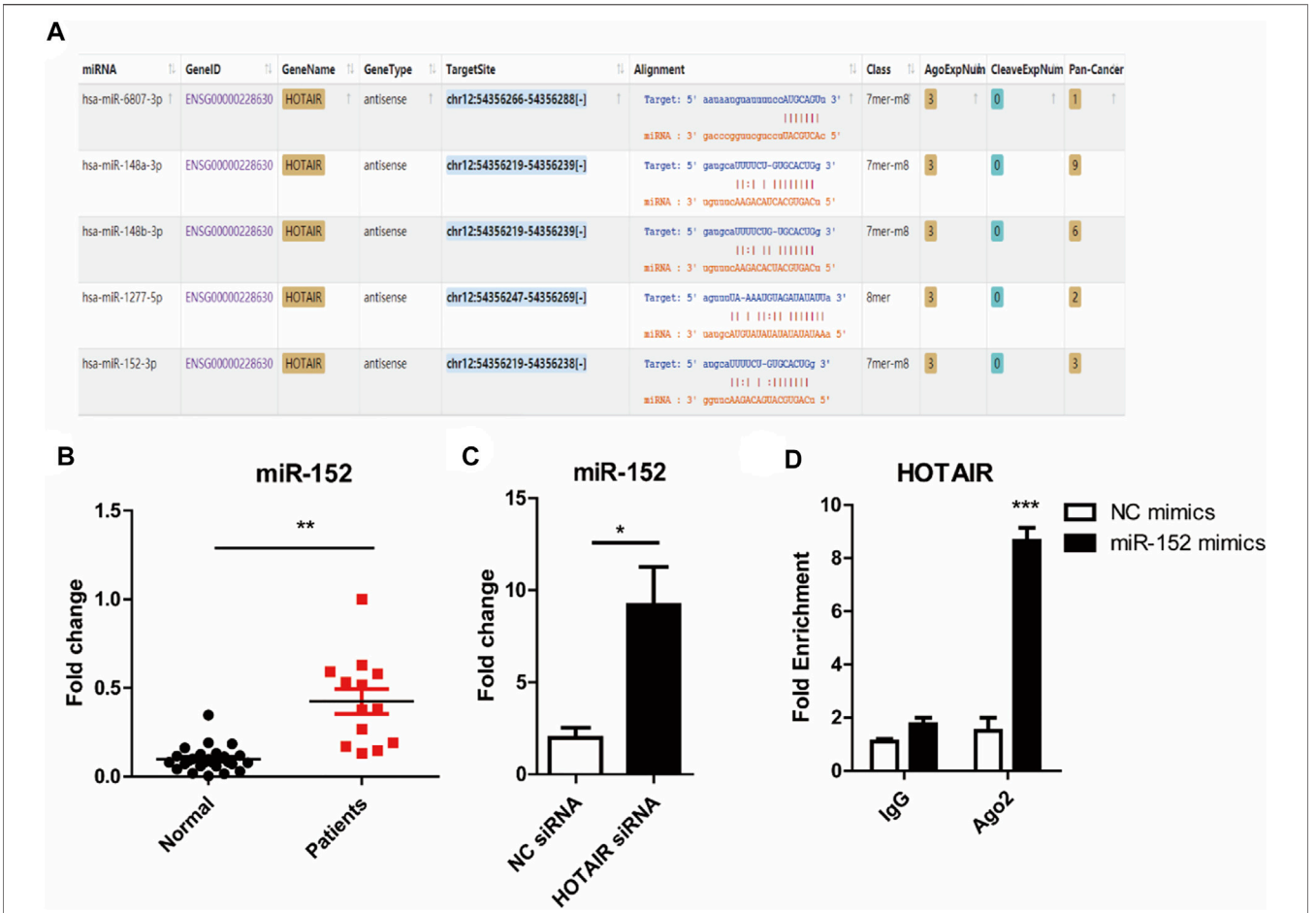


FIGURE 4 | HOTAIR acts as a sponge to inhibit miR-152 expression **(A)** Potential binding domains between HOTAIR transcript and miRNAs through Bioinformatics analysis using starbase Tools. The sequence alignments were visualized following the instruction on https://starbase.sysu.edu.cn/starbase2/index.php **(B)** Detection of miR-152 expression in PBMCs of patients with craniosynostosis, compared to control. $p < 0.01$ (Student's t-test) were considered statistical significant **(C)** RT-qPCR analysis of miR-152 expression levels in RANKL-induced osteoclasts with HOTAIR knockdown by two distinct shRNAs **(D)** RIP analyses of HOTAIR and miR-152 interaction in RANKL-induced osteoclast with miR-152 mimics or NC transfection (100 nM); RT-qPCR detected the mRNA levels HOTAIR; Bars represent the mean ± SD. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. The data are representative of three independent experiments.

bone pathology may also involve 2 cell types. Indeed, HOTAIR also affects osteoblastogenesis. Yuan et al. have shown that overexpression of HOTAIR inhibits osteogenic differentiation while promoting adipogenic differentiation of bone marrow stromal cells (Yuan et al., 2020). Wei et al. have found that HOTAIR inhibits osteogenic differentiation by suppressing miR-17-5p expression (Wei et al., 2017). HOTAIR also inhibits osteoblast differentiation of rat bone marrow stromal cells (Shen et al., 2019). Therefore, the results observed in our animal model and the possible role of HOTAIR in craniosynostosis seem to be related to the fact that HOTAIR favors bone resorption while inhibiting bone formation.

CaMKIIα is a vital gate that could control the structural, functional, and behavioral expression of synaptic memory. This may result in brain hypoplasia and craniofacial anomalies, including craniosynostosis and other anomalies (Yamagata et al., 2009). However, the rationality of the HOTAIR-miR-152-CAMKIIα pathway in the progression of craniosynostosis has not been explored in previous studies cordially. Nevertheless, the data of this study suggest that the HOTAIR-miR-152-CAMKIIα regulatory axis may be involved in forming osteoclasts during the occurrence and development of cranial suture injury. CAMKIIα is a serine/threonine protein kinase, which could play an important role in mediating NF-κB and MAPK/JNK signal transduction (Chang et al., 2008). NF-κB and MAPK signaling pathways play an essential role in regulating RANKL induced osteoclast formation. In the present study, CAMKIIα was identified as a direct target of miR-152. We found that HOTAIR knockout or miR-152 overexpression could inhibit CAMKIIα expression in RAW264.7 cells and weaken the activation of NF-κB. These data suggest that the HOTAIR-miR-152-CAMKIIα regulatory axis may be involved in forming osteoclasts during the occurrence and development of skull suture injury. However, the miR-152/CAMKIIα axis in

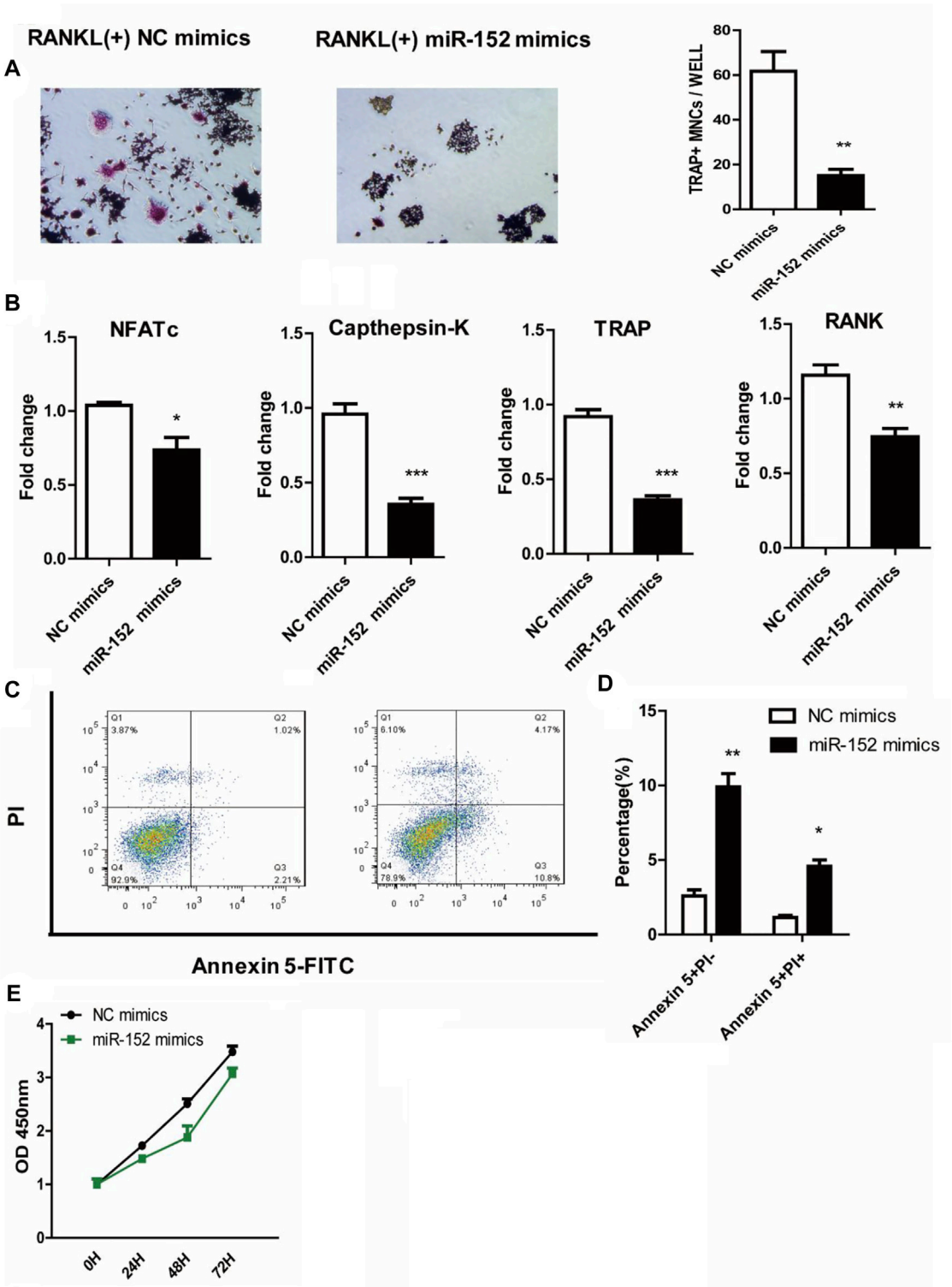


FIGURE 5 | miR-152 inhibits the RANKL-induced osteoclast differentiation in RAW 264.7 cells **(A)** Representative images and quantification of RANKL-induced osteoclasts transfected with NC mimics or miR-152 mimics were analyzed for TRAP + cells, Magnification: 100* **(B)** RT-qPCR analysis of osteoclast differentiation-related genes (NFATc1, Cathepsin K, TRAP, and RANK) expression in mature osteoclasts transfected with NC mimics or miR-152 mimics **(C)** After staining with Annexin 5-FITC and propidium iodide (PI), followed by analysis using a flow cytometer. Early apoptotic cells were Annexin5^{+} and PI^{-} and late apoptotic cells were Annexin5^{+} and PI^{+} **(D)** Analysis of mature osteoclasts apoptosis transfected with NC mimics or miR-152 mimics **(E)** RAW264.7 cells were cultured with M-CSF in the absence or presence of miR-152 mimics siRNA for 3 days, the proliferation ability of early osteoclast precursors was analyzed by CCK-8 assay. Bars represent the mean ± SEM.*p < 0.05, **p < 0.01, ***p < 0.001. The data are representative of three independent experiments.

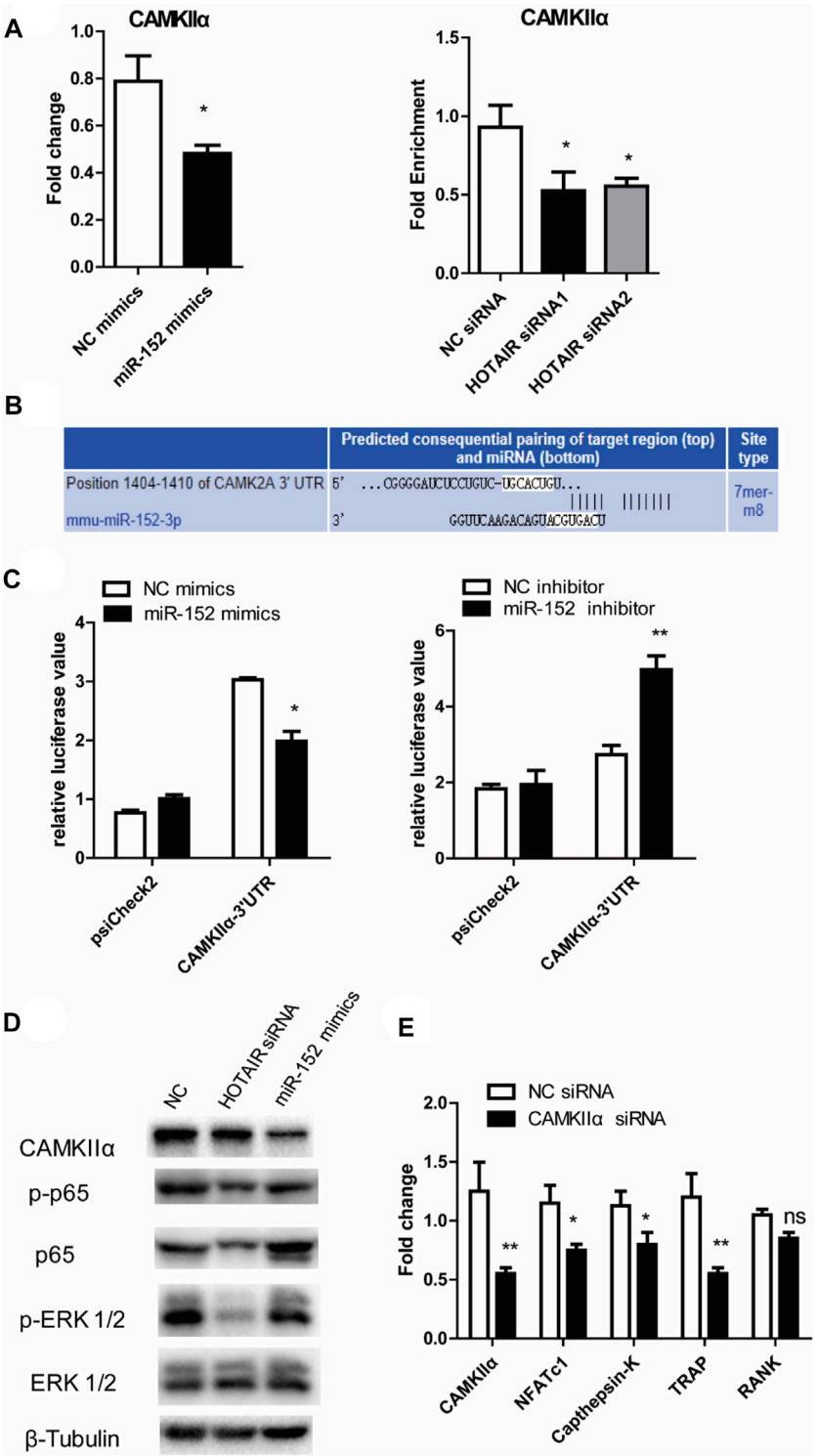


FIGURE 6 | HOTAIR-miR-152-CAMKIIα axis regulated osteoclast differentiation and apoptosis through NF-κB and MAPK/ERK 1/2 pathway **(A)** Expression of CAMKIIα mRNA in both miR-152 overexpressed (left) and HOTAIR-knockdown osteoclasts (right) was downregulated **(B)** Predicted binding sites for miR-152 in the 3′-UTR of CAMKIIα **(C)** RAW264.7 cells were transfected with the control construct (psiCHECK-2), or a construct encoding the wild-type CAMKIIα 3′-UTR, in addition to the miR-152 mimics (left) or miR-152 inhibitor (right). After 24 hs, luciferase activity in RAW264.7 lysates was detected **(D)** After transfection with HOTAIR siRNA or miR-152 mimics, the protein levels of CAMKIIα, Phosphorylated p65, p65, Phosphorylated ERK 1/2, and ERK 1/2 in mature osteoclasts were examined by western blotting. β-Tubulin was used to confirm equal protein loading **(E)** Effect of CAMKIIα on osteoclast differentiation-related gene expression in CAMKIIα siRNA-treated osteoclasts. Bars represent the mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$. The data are representative of three independent experiments.

osteoclast function and the occurrence and development of cranial suture disease are still unclear. It has been reported that miR-152/CAMKIIα axis is involved in immune homeostasis and immune regulation. Their data showed that miR-152 inhibited cytokine production, including IL-12, IL-6, TNF-α, and IFN-β. This has been upregulated MHC class II expression and inhibited DC-initiated Ag-specific T cell proliferation by targeting CAMKIIα(Liu et al., 2010). Therefore, it is suggested that miR-152/CaMKIIα may play an important role in osteoclast function. Through immunomodulation, it can cause the occurrence and development of craniosynostosis. This needs to be further explored.

In conclusion, craniosynostosis is the second most common cranial facial anomaly. The premature fusion of cranial sutures leads to deforming the skull shape and restricts brain growth. In this study, our findings demonstrated the role of HOTAIR in craniosynostosis through modulating miR-152 and its target gene CAMKIIα. This study provided a novel insight for understanding the potential molecular mechanism of osteogenic differentiation in craniosynostosis.

IMPACT STATEMENT

Craniosynostosis is a common congenital craniomaxillofacial malformation, and its pathogenesis is still unclear at present. Most studies have focused on the effect of osteoblasts on cranial suture. This study explores the opening and closing of cranial suture from the perspective of osteoclasts, and the role of lncRNA HOTAIR on the biological activity of osteoclasts, so that to explore a new idea on the mechanism of craniosynostosis.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Research ethics committee of Children's Hospital of Fudan Universityï (No. 2020-143). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. The animal study was reviewed and approved by Research ethics committee of Children's Hospital of Fudan University'No. 2020-143).

AUTHORS' CONTRIBUTIONS

All authors participated in the design, interpretation of the studies and analysis of the data and review of the manuscript, CD, XL, JL and DL conducted the experiment, CD wrote the manuscript.

FUNDING

This work was supported by the Shanghai Municipal Planning Commission of Science and Reserarch Fund (rant number 201740061); and Natural Science Foundation of Shanghai (grant number 22ZR1408300).

REFERENCES

Beederman, M., Farina, E. M., and Reid, R. R. (2014). Molecular Basis of Cranial Suture Biology and Disease: Osteoblastic and Osteoclastic Perspectives. *Genes Dis.* 1, 120–125. doi:10.1016/j.gendis.2014.07.004

Boulet, S. L., Rasmussen, S. A., and Honein, M. A. (2008). A Population-Based Study of Craniosynostosis in Metropolitan Atlanta, 1989-2003. *Am. J. Med. Genet.* 146A, 984–991. doi:10.1002/ajmg.a.32208

Boyce, B., Yao, Z., and Xing, L. (2009). Osteoclasts Have Multiple Roles in Bone in Addition to Bone Resorption. *Crit. Rev. Eukar Gene Expr.* 19, 171–180. doi:10.1615/critreveukargeneexpr.v19.i3.10

Bradley, J. P., Levine, J. P., McCarthy, J. G., and Longaker, M. T. (1997). Studies in Cranial Suture Biology: Regional Dura Mater Determines *In Vitro* Cranial Suture Fusion. *Plast. Reconstr. Surg.* 100, 1091–1099. doi:10.1097/00006534-199710000-00001

Byron, C. D. (2006). Role of the Osteoclast in Cranial Suture Waveform Patterning. *Anat. Rec.* 288A, 552–563. doi:10.1002/ar.a.20322

Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., et al. (2011). A Long Noncoding RNA Controls Muscle Differentiation by Functioning as a Competing Endogenous RNA. *Cell* 147, 358–369. doi:10.1016/j.cell.2011.09.028

Chang, E.-J., Ha, J., Huang, H., Kim, H. J., Woo, J. H., Lee, Y., et al. (2008). The JNK-dependent CaMK Pathway Restrains the Reversion of Committed Cells during Osteoclast Differentiation. *J. Cel Sci* 121, 2555–2564. doi:10.1242/jcs.028217

Clayman, M. A., Murad, G. J., Steele, M. H., Seagle, M. B., and Pincus, D. W. (2007). History of Craniosynostosis Surgery and the Evolution of Minimally Invasive Endoscopic Techniques. *Ann. Plast. Surg.* 58, 285–287. doi:10.1097/01.sap.0000250846.12958.05

Cohen, M. M., Jr (2009). Perspectives on Craniosynostosis20. *Suppl* 20, 646–651. doi:10.1097/scs.0b013e318193d48d

Czerwinski, M., Hopper, R. A., Gruss, J., and Fearon, J. A. (2010). Major Morbidity and Mortality Rates in Craniofacial Surgery: an Analysis of 8101 Major Procedures. *Plast. Reconstr. Surg.* 126, 181–186. doi:10.1097/prs.0b013e3181da87df

Feng, L., Xia, B., Tian, B.-F., and Lu, G.-B. (2019). MiR-152 Influences Osteoporosis through Regulation of Osteoblast Differentiation by Targeting RICTOR. *Pharm. Biol.* 57, 586–594. doi:10.1080/13880209.2019.1657153

Fitzpatrick, D. R. (2013). Filling in the Gaps in Cranial Suture Biology. *Nat. Genet.* 45, 231–232. doi:10.1038/ng.2557

French, L. R., Jackson, I. T., and Melton, L. J., 3rd (1990). A Population-Based Study of Craniosynostosis. *J. Clin. Epidemiol.* 43, 69–73. doi:10.1016/0895-4356(90)90058-w

Huang, J., Ke, P., Guo, L., Wang, W., Tan, H., Liang, Y., et al. (2014). Lentivirus-mediated RNA Interference Targeting the Long Noncoding RNA HOTAIR Inhibits Proliferation and Invasion of Endometrial Carcinoma Cells *In Vitro* and *In Vivo*. *Int. J. Gynecol. Cancer* 24, 635–642. doi:10.1097/igc.0000000000000121

Hukki, J., Saarinen, P., and Kangasniemi, M. (2008). Single Suture Craniosynostosis: Diagnosis and Imaging. *Front. Oral Biol.* 12, 79–90. doi:10.1159/000115033

Lajeunie, E., Merrer, M. L., Bonaïti-Pellie, C., Marchac, D., and Renier, D. (1995). Genetic Study of Nonsyndromic Coronal Craniosynostosis. *Am. J. Med. Genet.* 55, 500–504. doi:10.1002/ajmg.1320550422

Levine, J. P., Bradley, J. P., Roth, D. A., McCarthy, J. G., and Longaker, M. T. (1998). Studies in Cranial Suture Biology: Regional Dura Mater Determines Overlying Suture

Biology. *Plast. Reconstr. Surg.* 101, 1441–1447. doi:10.1097/00006534-199805000-00002

Liu, X., Zhan, Z., Xu, L., Ma, F., Li, D., Guo, Z., et al. (2010). MicroRNA-148/152 Impair Innate Response and Antigen Presentation of TLR-Triggered Dendritic Cells by Targeting CaMKIIα. *J.I.* 185, 7244–7251. doi:10.4049/jimmunol.1001573

Lyon, S. M., Mayampurath, A., Song, D., Ye, J., Januszyk, M., Rogers, M. R., et al. (2018). Whole-Proteome Analysis of Human Craniosynostotic Tissue Suggests a Link between Inflammatory Signaling and Osteoclast Activation in Human Cranial Suture Patency. *Plast. Reconstr. Surg.* 141, 250e–260e. doi:10.1097/PRS.0000000000004025

Ma, L., Golden, S., Wu, L., and Maxson, R. (1996). The Molecular Basis of Boston-type Craniosynostosis: the Pro148-->His Mutation in the N-Terminal Arm of the MSX2 Homeodomain Stabilizes DNA Binding without Altering Nucleotide Sequence Preferences. *Hum. Mol. Genet.* 5, 1915–1920. doi:10.1093/hmg/5.12.1915

Ma, Q., Liang, M., Wu, Y., Dou, C., Xu, J., Dong, S., et al. (2021). Small Extracellular Vesicles Deliver Osteolytic Effectors and Mediate Cancer-Induced Osteolysis in Bone Metastatic Niche. *J. Extracell Vesicles* 10, e12068. doi:10.1002/jev2.12068

Maliepaard, M., Mathijssen, I. M. J., Oosterlaan, J., and Okkerse, J. M. E. (2014). Intellectual, Behavioral, and Emotional Functioning in Children with Syndromic Craniosynostosis. *Pediatrics* 133, e1608–e1615. doi:10.1542/peds.2013-3077

Misawa, A., and Orimo, H. (2018). lncRNA HOTAIR Inhibits Mineralization in Osteoblastic Osteosarcoma Cells by Epigenetically Repressing ALPL. *Calcif Tissue Int.* 103, 422–430. doi:10.1007/s00223-018-0434-0

Misra, S., Shih, A., Yan, X., and Li, W. (2019). A Higher Proportion of Craniosynostosis Genes Are Cancer Driver Genes. *bioRxiv* 11, 872093. doi:10.1101/872093

Muenke, M., Gripp, K. W., McDonald-McGinn, D. M., Gaudenz, K., Whitaker, L. A., Bartlett, S. P., et al. (1997). A Unique point Mutation in the Fibroblast Growth Factor Receptor 3 Gene (FGFR3) Defines a New Craniosynostosis Syndrome. *Am. J. Hum. Genet.* 60, 555–564.

Renier, D., Sainte-Rose, C., Marchac, D., and Hirsch, J.-F. (1982). Intracranial Pressure in Craniostenosis. *J. Neurosurg.* 57, 370–377. doi:10.3171/jns.1982.57.3.0370

Sato, K., Suematsu, A., Nakashima, T., Takemoto-Kimura, S., Aoki, K., Morishita, Y., et al. (2006). Regulation of Osteoclast Differentiation and Function by the CaMK-CREB Pathway. *Nat. Med.* 12, 1410–1416. doi:10.1038/nm1515

Seales, E. C., Micoli, K. J., and McDonald, J. M. (2006). Calmodulin Is a Critical Regulator of Osteoclastic Differentiation, Function, and Survival. *J. Cel. Biochem.* 97, 45–55. doi:10.1002/jcb.20659

Shen, J. J., Zhang, C. H., Chen, Z. W., Wang, Z. X., Yang, D. C., Zhang, F. L., et al. (2019). LncRNA HOTAIR Inhibited Osteogenic Differentiation of BMSCs by Regulating Wnt/β-Catenin Pathway. *Eur. Rev. Med. Pharmacol. Sci.* 23, 7232–7246. doi:10.26355/eurrev_201909_18826

Singer, S., Bower, C., Southall, P., and Goldblatt, J. (1999). Craniosynostosis in Western Australia, 1980-1994: a Population-Based Study. *Am. J. Med. Genet.* 83, 382–387. doi:10.1002/(sici)1096-8628(19990423)83:5<382::aid-ajmg8>3.0.co;2-a

Sloan, G. M., Wells, K. C., Raffel, C., and McComb, J. G. (1997). Surgical Treatment of Craniosynostosis: Outcome Analysis of 250 Consecutive patientsE2. *Pediatrics* 100, E2. doi:10.1542/peds.100.1.e2

Soderling, T. R., and Stull, J. T. (2001). Structure and Regulation of Calcium/calmodulin-dependent Protein Kinases. *Chem. Rev.* 101, 2341–2352. doi:10.1021/cr0002386

Susiarjo, M., Hassold, T. J., Freeman, E., and Hunt, P. A. (2007). Bisphenol A Exposure In Utero Disrupts Early Oogenesis in the Mouse. *Plos Genet.* 3, e5. doi:10.1371/journal.pgen.0030005

Tang, P., Xiong, Q., Ge, W., and Zhang, L. (2014). The Role of microRNAs in Osteoclasts and Osteoporosis. *RNA Biol.* 11, 1355–1363. doi:10.1080/15476286.2014.996462

Wei, B., Wei, W., Zhao, B., Guo, X., and Liu, S. (2017). Long Non-coding RNA HOTAIR Inhibits miR-17-5p to Regulate Osteogenic Differentiation and Proliferation in Non-traumatic Osteonecrosis of Femoral Head. *PLoS One* 12, e0169097. doi:10.1371/journal.pone.0169097

Wilkie, A. (1997). Craniosynostosis: Genes and Mechanisms. *Hum. Mol. Genet.* 6, 1647–1656. doi:10.1093/hmg/6.10.1647

Xing, D., Liang, J.-q., Li, Y., Lu, J., Jia, H.-b., Xu, L.-y., et al. (2014). Identification of Long Noncoding RNA Associated with Osteoarthritis in Humans. *Orthopaedic Surg.* 6, 288–293. doi:10.1111/os.12147

Xu, Y., Chen, B., George, S. K., and Liu, B. (2015). Downregulation of MicroRNA-152 Contributes to High Expression of DKK1 in Multiple Myeloma. *RNA Biol.* 12, 1314–1322. doi:10.1080/15476286.2015.1094600

Yaccoby, S., Ling, W., Zhan, F., Walker, R., Barlogie, B., and Shaughnessy, J. D., Jr (2007). Antibody-based Inhibition of DKK1 Suppresses Tumor-Induced Bone Resorption and Multiple Myeloma Growth *In Vivo*. *Blood* 109, 2106–2111. doi:10.1182/blood-2006-09-047712

Yamagata, Y., Kobayashi, S., Umeda, T., Inoue, A., Sakagami, H., Fukaya, M., et al. (2009). Kinase-Dead Knock-In Mouse Reveals an Essential Role of Kinase Activity of Ca2+/Calmodulin-dependent Protein Kinase II in Dendritic Spine Enlargement, Long-Term Potentiation, and Learning. *J. Neurosci.* 29, 7607–7618. doi:10.1523/jneurosci.0707-09.2009

Yuan, S., Zhang, C., Zhu, Y., and Wang, B. (2020). Neohesperidin Ameliorates Steroid-Induced Osteonecrosis of the Femoral Head by Inhibiting the Histone Modification of lncRNA HOTAIR. *Dddt* Vol. 14, 5419–5430. doi:10.2147/dddt.s255276

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership