Exploration of the regulatory mechanism for hepatocellular carcinoma based on ceRNA networks analysis and immune infiltration

Ning Huang
Yunnan Agricultural University https://orcid.org/0000-0002-0269-0012
Qiang Chen (chq@sjtu.edu.cn)
Yunnan Agricultural University
Xiaoyi Wang
Yunnan Agricultural University

Research

Keywords: Hepatocellular carcinoma, weighted gene co-expression network analysis, competing endogenous RNA, immune infiltration

Posted Date: September 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-885697/v1

Abstract

Background

Hepatocellular carcinoma (HCC) as malignant cancer has been deeply investigated for its widespread distribution and extremely high mortality rate worldwide. Despite efforts to understand the regulatory mechanism in HCC, it remains largely unknown.

Methods

The RNA (mRNAs, lncRNAs, and miRNAs) profiles were downloaded from The Cancer Genome Atlas (TCGA) database. Based on the Weighted Gene Co-expression Network Analysis (WGCNA), the hub differentially expressed RNAs (DERNAs) were screened out. The competing endogenous RNA (ceRNA) and Protein and Protein Interaction (PPI) network were constructed based on the hub DERNAs. The Cox and LASSO regression analysis were used to find the independent prognostic ceRNAs. We performed the “CIBERSORT” algorithm estimate the abundance of immune cells. The correlation analysis was applied to determine the relationship between HCC-related immune cells and prognostic ceRNAs. GEPIA and TIMER database were used to explore the association of critical genes with survival and immune cell infiltration, respectively.

Results

A total of 524 hub RNAs (507 DEmRNAs, 13 DElncRNAs and 4 DEmiRNAs) were identified in the turquoise module (cor = 0.78, P = 4.7e − 198) using WGCNA algorithm. PPI network analysis showed that NDC80, BUB1B and CCNB2 as the critical genes in HCC. Subsequently, survival analysis revealed that the low expression of NDC80 and BUB1B resulted in a longer overall survival (OS) time for HCC patients in GEPIA database. These critical genes and several immune cells were all significantly positive correlated in TIMER database. The ceRNA network were establish, and were incorporated to risk model. Subsequently, ROC curve showed that the area under the curve (AUC) of the 1-, 3-, and 5-year were 0.762, 0.705, and 0.688, respectively. Out of the 22 cell types, T cells CD4 memory resting were identified as the HCC-related immune cells by systematic analysis. The correlation analysis shown that T cells CD4 memory resting is negatively associated with both AL021453.1 (R = − 0.44, P = 0.00049) and CCDC137 (R = − 0.47, P = 2e-04).

Conclusion

The current study provide potential prognostic signatures and therapeutic targets for HCC.

Background

Liver cancer is the fourth leading cause of cancer-related deaths worldwide [1]. Hepatocellular carcinoma (HCC) accounts for more than 90% of primary liver cancer [1]. At present, surgical treatment is the main treatment for early HCC patients, but there is still a high recurrence rate for HCC patients [2]. Intermediate HCC patients is treated with arterial radiotherapy embolization (TACE) [3]. Even so, the overall survival of HCC patients is not ideal. The numerous patients are diagnosed in the advanced stage of the disease. It is reported that approximately 80% of patients with HCC may miss the opportunity for surgical treatment, liver transplantation and radiofrequency ablation in the early stages of cancer [4]. Molecular targeted drugs are currently the only effective treatment for advanced HCC, however the overall survival rate of patients is still not ideal [5].Therefore, it is necessary to find new prognostic biomarker and therapeutic targets for immunotherapy of HCC.

Accumulating evidences have showed that tumor cells and tumor-infiltrating immune cells (TICs) play an important role in the occurrence and development of tumors by changing the immune microenvironment [6]. Recent studies have shown that the crosstalk between tumor cells and TICs is usually regulated by a ceRNA network, which contains messenger RNAs (mRNAs), long noncoding RNAs (lncRNAs), and microRNAs (miRNAs) [7]. LncRNAs as a menber member of ceRNA network can adjust the related genes expression level by competitively combining with microRNA, thereby regulating various biological processes such as tumor occurrence and progression [8]. Thus, it is essential to explore the potential regulatory mechanism underlying the ceRNA network

and immune cells to combat HCC. However, a lack of studies consider the co-expression feature in ceRNA network, which may be difficult to target the key ceRNA network from the the numerous ceRNA networks. Weighted gene co-expression network analysis (WGCNA) is an efficient analysis approach which is widely used to identify the most significant module, and further to discover the key molecules in the network [9].

In our study, we used the WGCNA to identify the hub differential expressed RNAs, including DEmRNAs, DElncRNAs and DEmiRNAs and established a ceRNA network, based on HCC-related transcriptome data from the Cancer Genome Atlas (TCGA) database. The "CIBERSORT" algorithm to calculate the abundance of immune cells. Subsequently, the ceRNA-related prediction risk model was constructed to predict the prognosis in HCC patients. Moreover, correlation analysis was performed to explore the potential regulatory mechanism between ceRNA network and the differential immune cells (between two risk groups).To further investigate the molecular and immune cells, the critical gene were screened out from the hub DEmRNAs by the centrality analysis. Meanwhile, TIMER database was applied to assess the relationships between these critical genes and several immune cells.

Results

Identification of differentially expressed RNAs

The workflow of bioinformatics analysis in the present study was showed in Fig. 1. A total of 2293 DERNAs were screened out between the HCC samples and tumor-adjacent samples based on the cut-off threshold ($P < 0.05$ and the log2 (fold-change) > 1). Among them, 2,028 DEmRNAs (1222 up-regulated and 806 down-regulated), 136 DElncRNAs (104 up-regulated and 32 down-regulated) and 129 DEmiRNAs (107 up-regulated and 21 down-regulated) were identified (Fig. 2A-2F). All up- and down-regulated DERNAs were summarized in Table S1-S3.

Construction of weighted gene co-expression module network

The DERNAs were used to establish the co-expression networks. In the WGCNA package of R software, The "pick soft threshold" function automatically selected the most suitable β value of 4, which ensured the scale independence value (R^2) > 0.9 and lower mean connectivity (Fig. 3A). Based on the hierarchical clustering and dynamic cut tree algorithm, we got 7 co-expression modules (Fig. 3B). In addition, based on the correlation coefficient and the corresponding p value, we found that turquoise module had the most highly association with tumor (cor = − 0.70, P = 2e − 63, Fig. 3C). Meanwhile, Fig, 3D showed the closest correlation between MM and GS in turquoise module (cor = 0.78, p = 4.7e − 198). Thus, the turquoise module including 965 DERNAs was selected as the candidate module for further analysis.

Functional annotation analysis

GO and KEGG pathway enrichment analysis were used to explore the potential functions of 507 hub genes with the cut-off criteria ($|MM| \geq 0.6$ and $|GS| \geq 0.2$) in turquoise module. GO annotation showed that 499 BPs, 65 CCs and 63 MFs were significantly enriched and the most significant BP, CC, and MF were organelle fission, chromosomal region, catalytic activity, respectively (Fig. 4A). KEGG analysis revealed that the most significant pathway was the Cell cycle (Fig. 4B, 4C).

PPI network construction and hub gene validation

In the turquoise module, 507 hub genes were used to conduct PPI network analysis. Using the STRING database, we got a PPI network including 175 nodes and 705 edges (Fig. 5A). Based on MCODE, we selected the most significant subnetwork (Fig. 5B) with 18 nodes and 139 edges (score = 16.353) for the next centrality analysis in the Cytoscape. Centrality analysis indicated that BUB1B (logFC = 3.511, P = 4.52E-64), NDC80 (logFC = 3.308, P = 9.99E-75) and CCNB2 (logFC = 3.724, P = 1.75E-80) were screened as the critical genes in the subnetwork1 (Table 2). In the GEPIA database, the expression levels of three genes in cancer tissues were significantly higher than in normal tissues (Additional file 4: Figure S6A-S6C). Meanwhile, the three genes were confirmed to be significantly correlated with pathological stages (Additional file 4: Figure S1D-S1F). In addition, the low expression levels of BUB1B and NDC80 were significantly related to a longer patients' OS time in the Kaplan Meier analysis (Additional file 1: Figure S1G-S1I). The findings suggested that BUB1B and NDC80 might sever as the OS-associated critical genes in HCC patients.

Table 2
Centrality analysis results of top 3 genes in subnetwork.

Rank	Gene	Subgragh	Degree	Eigenvector	Information	LAC	Betweenness	Closeness	Network
1	BUB1B	394316.12	17	0.2539486	9.004098	14.3529415	2.438739	1	17
2	NDC80	394313.53	17	0.2539478	9.004098	14.3529415	2.438739	1	17
3	CCNB2	394313.53	17	0.2539478	9.004098	14.3529415	2.438739	1	17

Association between critical genes and tumor-infiltrating immune cells

Due to infiltrating immune cells play an important role in the tumor landscape, we analyzed the correlation of critical gene with several immune cells based on the TIMER database. The results showed that CCNB2, NDC80 and BUB1B were positively correlated with immune cells (including B cells, CD8 + T cells, CD4 + T cells, macrophages, neutrophils and dendritic cells) (Additional file 5: Fig. 2SA,2SB). The results indicated that CCNB2, NDC80 and BUB1B possessed the potential to serve as immunotherapy targets in HCC.

Construction of ceRNA network and survival analysis for HCC

To better understand the interactions among DEmRNAs, DElncRNAs and DEmiRNAs, we constructed a ceRNA network that the DElncRNA (AL021453.1) that regulates the expression of 3 DEmRNAs (ABCC5, RACGAP1 and CCDC137) by competitively binding the DEmiRNAs (hsa-miR-93-5p) (Fig. 6A; Table 1). The Cox regression and KM (LR test) analysis were used to identify ceRNA biomarkers associated with the OS and prognosis of HCC patients. With $P < 0.05$ cut-off threshold, ABCC 5 ($P = 0.001$), RACGAP1 ($P = 0.001$), AL021453.1 ($P = 0.041$), CCDC137 ($P = 2.87e-05$) were detected to have significantly associated with the HCC patients' OS in KM analysis (Fig. 6B–6E). Five ceRNAs all were determined as the key signature and incorporated into the multivariate Cox regression model (Fig. 7A). In addition, the results of LASSO regression suggested that the 5 ceRNA biomarkers were crucial for the modeling and non-overfitting for the model (Fig. 7B-7D). ROC and the calibration curves manifested that the risk prediction model had high accuracy and good discrimination for the prognosis of patients with HCC (Fig. 7E, 7F). The AUC of 1-, 3-, 5-year survival was 0.762, 0.705, and 0.688, respectively (Fig. 7E). Furthermore, compared with the low-risk group, the high-risk group showed the worse survival probability (Fig. 7G).

Table 1
Hypergeometric testing and correlation analysis results based on ceRNA network.

LncRNAs	Genes	MiRNA	Hypergeometric test P	Correlation P
AL021453.1	ABCC5	hsa-miR-93-5p	0.031241496	6.46E-19
AL021453.2	RACGAP1	hsa-miR-93-5p	0.003238409	3.68E-33
AL021453.3	CCDC137	hsa-miR-93-5p	0.001739558	3.85E-25

The composition of immune cells in HCC

We use the "CIBERSORT" algorithm to estimate the composition of 22 immune cells in each sample (Fig. 8A, 8B). Using Wilcoxon rank-sum test, the Macrophages M2 ($P = 0.019$) possessed significantly higher infiltration level in the high-risk group, while T cells CD4 memory resting ($P = 0.035$) had significantly higher infiltration level in the low-risk group (Fig. 8C). The abundance of T cells CD4 memory resting ($P = 0.037$) and T cells follicular helper ($P = 0.021$) was found to be significantly related to patient's OS rate (Additional file 3: Fig. 3SA, 3SB). Based on the Wilcoxon rank-sum test, we found that the fraction of T cells CD4 memory resting ($P = 0.031$) was significantly negatively associated with clinical grades (Supplementary Fig. 3C), while the T cells follicular helper ($P = 0.021$) was significantly positively associated with the clinical grades (Additional file 6: Fig. 3SD). Thus, T cells follicular helper was identified as the key HCC-related immune cell.

Pearson correlation analysis

To explore whether the five ceRNA signatures could reflect the state of immune cells in tumor microenvironment, we evaluated the association between the ceRNA signatures and the immune cells via the correlation analysis (Fig. 9A, 9B). The results indicated that

T cell CD4 memory resting is significantly negatively associated with both AL021453.1 (R = − 0.44, P = 4.9e-40.0005) and CCDC137 (R = − 0.47, P = 2e − 04), with P < 0.001 and |R|>0.3 cut-off threshold (Fig. 9C, 9D).

Discussion

The occurrence and development of HCC is a multi-factor and multi-stage complex pathogenesis, involving the aberrant expression of many genes and changes in the immune microenvironment [10]. HCC is a cancer with a high fatality rate in the world, and there is currently not effective diagnosis and treatment method [11, 12]. Increasing experimental evidences demonstrated that the ceRNAs can regulate the interaction between invading immune cells and tumor cells and influence tumor proliferation, metastasis, drug resistance and the prognostic outcome of patients [8, 13, 14]. Therefore, it is extremely urgent to find potential novel therapeutic targets and prognostic markers.

In our study, we identified two survival-related critical genes (NDC80 and BUB1B) by WGCNA, PPI network construction, centrality analysis and survival analysis. Generally, NDC80 (nuclear division cycle 80) is a core component of the NDC80 kinetochore complex, which consists of NDC80/HEC1, CDCA1, SPBC24 and SPBC25 [15]. A previous study showed that NDC80 complexes plays important roles in chromosome segregation, and spindle assembly checkpoint function [16]. In the tumorigenesis, overexpression of NDC80 leads to the hyper-activation of spindle assembly checkpoint and finally causes chromosome instability, which is generally considered to be landmark events of tumor growth [17]. Furthermore, higher expression level of NDC80 have been reported in HCC tissues, compared to adjacent tissues, which is insistent with our research result. Previous researches manifested that higher levels of NDC80 in tumor tissues promote HCC progression by anti-apoptotic effects and overcoming cell cycle arrest. BUB1B (BUB1 Mitotic Checkpoint Serine/Threonine Kinase B) is an essential component of the spindle checkpoint [18]. Increasing studies indicate overexpression of BUB1B promotes the tumor proliferation and invasion by BUB1B/mTORC1 signaling pathway in HCC [19]. The TICs, as a component of the tumor microenvironment, exert enormous influence on tumor growth, invasion, and metastasis [20]. The tumor microenvironment has different abilities to induce adverse and beneficial consequences of tumorigenesis [21]. In this study, we suggested that the high expression level of NDC80 and BUB1B was significantly related with the patients' short OS time and high immune infiltrates.

In the current study, we found that the HCC-related T cells CD4 memory resting is negatively associated with AL021453.1 (R = − 0.44, P = 4.9e-4) and CCDC137 (R = − 0.47, P = 2e − 04) by correlation analysis, which suggested that AL021453.1 and CCDC137 may play a role in immune landscapes of HCC.

AL021453.1 is an endogenous lncRNAs, has so far been reported to aberrant expression in the Pheochromocytoma and Paraganglioma [22]. But its function entirely unknown in cancer. This study is the first to show AL021453.1 as an independent biomarker high expression associated with poor overall survival and and high risk scores in HCC patients. CCDC family proteins can participate in intracellular signal transduction, genetic signal transcription, molecular recognition and cell cycle regulation [23, 24]. Similar to AL021453.1, the CCDC137 as an independent signature was highly expressed in HCC tissue, and HCC patients with high expression possessed poor overall survival rate and prognosis. Some previous studies showed that the CCDC137(coiled-coil domain containing 137) is the member of CCDC family proteins contributing to various diseases and may contribute to elevated infiltration of tumor microenvironment and was associated with tumor immunosuppressive status [25]. Besides, it had been reported to involve in the human immunodeficiency virus type 1 (HIV-1) regulating immune cell death [26]. Meanwhile, accumulating studies identified that metabolic changes in the tumor microenvironment may impact immune metabolism, thus promoting or damaging anti-cancer immunity [27]. For example, Van et al found thatnon-alcoholic fatty liver disease (NAFLD) cause apoptosis of intrahepatic CD4 + T cells and promote the tumorigenesis [28]. In adition, Brown et al further certified that down-regulated expression of Carnitine palmitoyltransferase gene induces CD4 + T cell apoptosis, thereby facilitating the HCC development through in vivo experiments [29]. In our study, T cells CD4 memory resting in high risk groups possessed lower fraction than in the lower risk groups. In clinical and survival analysis, high expression of T cells CD4 memory resting showed higher survival probability, and higher overall survival [30, 31]. Thus, T cells CD4 memory resting may function as HCC suppressors. After a systematic literature review, we found no direct reports on AL021453.1 association with dendritic cells and tumor immunity. However, the hypothesis that ceRNAs are involved in immune regulation and affecting HCC tumorigenesis and development have been validated by various experiments [7]. Meanwhile, the hypergeometric testing and correlation analysis results of the ceRNAs network revealed that CCDC137 (protein-coding RNA), hsa-miR-93-5p (miRNA), and AL021453.1 (LncRNA) were significantly

correlated. Thus, we infered that AL021453.1 might indirectly regulate CCDC137 by interacting with hsa-miR-93-5p, and further regulate T cells CD4 memory resting. We had such a conserved discussion because of the lack of literature. Further studies to explore the relationship between the ceRNAs and T cells CD4 memory resting are needed.

Conclusion

In conclusion, a novel ceRNA-related nomogram was established to predict HCC patients' the prognosis and survival, and certified the utility by the high AUC values. The proposed ceRNA prognostic signatures might promote better individual treatment and management of HCC patients. Moreover, the immune-related ceRNA network mechanism might further elucidate the regulatory mechanisms underlying HCC. However, an important limitation of our study must be noted. All of our results are found based on TCGA using a pure bioinformatics analysis. Although, these results have been proved by GEPIA database, further biological and molecular basic experiments to verify our results are needed.

Methods

Datasets from TCGA

RNA profiles (mRNAs, lncRNAs and miRNAs) of 424 samples including 50 para-carcinoma HCC and 374 HCC tissues samples were retrieved from The Cancer Genome Atlas (TCGA) database (https://tcga-data.nci.nih.gov/tcga/). Simultaneously, the corresponding clinical information of HCC patients were collected from the TCGA database. Among them, 5 HCC patients with incomplete follow-up information were excluded for survival analysis. The present study does not require the approval from the ethics committee.

RNA-seq data preprocessing and differential expression analysis

The TMM (trimmed mean of M-values) algorithm was used to normalize RNA-seq data. The cut-off criteria of |logFC| >2 and false discovery rate (FDR) P-value < 0.05 were setted to differential expression analysis for the RNA data between the HCC and normal tissue samples. In addition, the ggplot2 and pheatmap package of R software was used to draw differentially expressed RNAs (DERNAs, including DEmRNAs, DElncRNAs and DEmiRNAs) by volcano maps and heat maps, respectively.

Weighted gene co-expression network analysis

The DERNAs co-expression network was established based on the WGCNA package in R software [32]. According to the results of scale independence and mean connectivity analysis, we can choose power applicable soft threshold to construct an adjacency matrix. The soft threshold was selected based on the scale independence value (R^2) > 0.9 and mean connectivity. Subsequently, the adjacency matrix was transformed into a topological overlap matrix (TOM) that was applied to access the correlation of gene pairs. Based on the corresponding dissimilarity (1-TOM), the hierarchical clustering and dynamic tree cut analysis were performed to classify the genes with similar expression profiles into different gene co-expression modules. The clinical trait information was applied to calculate the module-trait associations. The most significant and correlated module was selected as the candidate module for subsequent analysis. Module membership (MM) and gene significance (GS) as the vital indicator identifying the central genes in the modules were calculated.

PPI network construction and analysis

The hub genes with |MM| ≥ 0.6 and |GS| ≥ 0.2 in the candidate module were used to construct a PPI (protein–protein interactions) network using STRING database [33]. In the STRING database, the genes with combined score ≥ 0.99 were selected to establish a network visualized in Cytoscape v.3.5.1 [34]. In the network, the most significant correlated subnetwork was selected for subsequent analysis by a plugin MCODE (threshold parameters: Degree Cutoff = 2, Node Score Cutoff = 0.4, K-Core = 4 and Max. Depth = 100 cut-off threshold) in Cytoscape. Subsequently, the centrality analysis containing eight Centrality methods (Subgraph Centrality, Degree Centrality, Eigenvector Centrality, Information Centrality, LAC Centrality, Betweenness Closeness Closeness Centrality and Network Centrality) was used to identify the critical genes (top 3 genes in the Centrality scores) by using the plugin CytoNCA of Cytoscape.

Gene function analysis and critical gene validation

To understand the possible functions of the hub genes, the clusterProfiler package of R software [35] was applied to conduct the GO and KEGG pathways enrichment analysis among hub genes, with a cut-off criterion of adjusted $p < 0.05$. GO annotation includes Biological Process (BP), Cellular Component (CC) and Molecular Function (MF). Using the online GEPIA (http://gepia.cancer-pku.cn) database, the expression of critical genes in PPI network between HCC and normal tissues was validated, and the association between overall survival and the critical genes was determined.

Exploration of associations between critical genes and immune cells

The TIMER (https://cistrome.shinyapps.io/timer/) database, as a comprehensive website, can systematically evaluate the immune infiltration levels of various cancer types. In the website, immune cell types mainly include B cells, CD8 + T cells, CD4 + T cells, macrophages, neutrophils and dendritic cells. The Correlation analysis between the critical genes and immune cell was based on "Gene" part of the website. Scatter plot was online generated and the Pearson's correlation coefficient and corresponding P value were calculated in the website.

Construction of lncRNA-miRNA-mRNA ceRNA network

In the candidate module (turquoise module), the hub DERNAs with $|MM| \geq 0.6$ and $|GS| \geq 0.2$ were used to construct the ceRNA network. The miRNA–mRNA interaction relationships verified by experimental methods was obtained from miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/index.html) [36]. The lncRNA–miRNA interaction relationships were predicted using lncbase v.2 Experimental Module (http://carolina.imis.athenainnovation.gr/diana_tools/web/index.php?r=lncbasev2%2Findex-experime ntal) [37]. Lastly, the intersecting mRNAs, lncRNAs and miRNAs, with statistical significance in hypergeometric testing and correlation analysis, were selected to establish the ceRNA network. The network was visualized using the Cytoscape software.

Identification of HCC-specific prognostic ceRNA signatures

The relationships between the ceRNAs and the overall survival (OS) of HCC patients were explored by using the Kaplan Meier (KM) analysis and Log-rank (LR) test in the survminer package of R software. Moreover, ceRNAs were selected to construct Cox proportional hazards model. Then, the ceRNAs in the Cox model was selected to perform the Least Absolute Shrinkage And Selection Operator (LASSO) regression analysis for reducing the impact of collinearity among the variables and preventing over-fitting of variables in the prognostic risk model. According to the risk models, the rms package of R software was used to construct a nomogram to assess the prognostic value for patients with HCC. These patients were classified into high- and low-risk groups based on the the median of patients' risk scores. The ROC and calibration curves were applied to evaluate the sensitivity and concordance of the risk model.

CIBERSORT estimation

To further explore the relationship of TICs and ceRNA signatures. The "CIBERSORT" algorithm [38] was used to predict the abundances of 22 immune cells in the HCC tissues samples and tumor-adjacent normal tissues samples. The Wilcoxon rank-sum test was performed to compare the differential proportion of immune cells between two risk groups (high risk group *vs* low-risk groups) and two clinical grades (G1 + G2 *vs* G3 + G4). In addition, KM analysis was implemented to assess the relationship between the proportion of immune cells and HCC patients' OS rates. Pearson correlation was used to examine the relationship between ceRNA signatures and the HCC-related immune cells that was simultaneously associated with the patients' risk levels, clinical grades and overall survival. $P < 0.001$ and $|R| > 0.3$ were set as the cut-off threshold of Pearson correlation.

Statistical analysis

All statistical analysis were enforced with R version 4.0.2 statistical software. The two-sided P value of less than 0.05 was considered to be statistical significance expect specified cut-off values.

Abbreviations

HCC	Hepatocellular carcinoma
TCGA	The Cancer Genome Atlas
WGCNA	Weighted Gene Co-expression Network Analysis
DERNAs	Differentially expressed RNAs
DEmRNAs	Differentially expressed mRNAs
DEmiRNAs	Differentially expressed miRNAs
DElncRNAs	Differentially expressed lncRNAs
CeRNA	Competing endogenous RNA
PPI	Protein and Protein Interaction
CIBERSORT	Cell type identification by estimating relative subsets of RNA transcripts Treated
TACE	Treated with arterial radiotherapy embolization
TICs	Tumor-infiltrating immune cells
MM	Module membership
GS	Gene significance
TOM	Topological overlap matrix
BP	Biological Process
CC	Cellular Component
MF	Molecular Function
OS	Overall survival
KM	Kaplan Meier
LR	Log-rank
LASSO	Least Absolute Shrinkage And Selection Operator
NAFLD	Non-alcoholic fatty liver disease
Human	Immunodeficiency virus type 1
CCDC137	Coiled-coil domain containing 137
BUB1B	BUB1 Mitotic Checkpoint Serine/Threonine Kinase B
NDC80	Nuclear division cycle 80

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The raw data of this study are acquired from the TCGA database (https://portal.gdc.cancer.gov/), which is publicly available database.

Competing interests

The authors declare no conflict of interest.

Funding

The work was supported by Yunnan Province Applied Basic Research Projects (no 2016FB146).

Authors' contributions

QC conceived and designed the project. NH and XW performed the experiments and analyzed the data. QC and NH wrote the paper. All authors reviewed the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Oshi M, Kim TH, Tokumaru Y, Li Y, Takabe K. Enhanced DNA Repair Pathway is Associated with Cell Proliferation and Worse Survival in Hepatocellular Carcinoma (HCC). Cancers. 2021;13(2):323.
2. Poon R. Optimal initial treatment for early hepatocellular carcinoma in patients with preserved liver function: transplantation or resection? Ann Surg Oncol. 2007;14(2):541–7.
3. Wang EA, Stein JP, Bellavia RJ, Broadwell SR. Treatment options for unresectable HCC with a focus on SIRT with Yttrium-90 resin microspheres. International Journal of Clinical Practice. 2017;e12972.
4. Greten TF, Wang XW, Korangy F. Current concepts of immune based treatments for patients with HCC: from basic science to novel treatment approaches. Gut. 2015;64(5):842–8.
5. Raoul J-L, Kudo M, Finn RS, Edeline J, Reig M, Galle PR. Systemic therapy for intermediate and advanced hepatocellular carcinoma: Sorafenib and beyond. Cancer treatment reviews. 2018;68:16–24.
6. Gajewski TF, Schreiber H, Fu Y-X. Innate and adaptive immune cells in the tumor microenvironment. Nature immunology. 2013;14(10):1014–22.
7. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. Nature. 2014;505(7483):344–52.
8. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: The rosetta stone of a hidden RNA language? Cell. 2011;146(3):353–8.
9. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. Bmc Bioinformatics. 2008;9(1):559.
10. Hu B, Lin JZ, Yang XB, Sang XT. Aberrant lipid metabolism in hepatocellular carcinoma cells as well as immune microenvironment: A review. Cell proliferation. 2020;53(3):e12772.
11. Ruf B, Heinrich B, Greten TF. Immunobiology and immunotherapy of HCC: spotlight on innate and innate-like immune cells. Cell Mol Immunol. 2021;18(1):112–27.
12. Liu G-M, Zeng H-D, Zhang C-Y, Xu J-W. Identification of a six-gene signature predicting overall survival for hepatocellular carcinoma. Cancer cell international. 2019;19(1):1–13.
13. Rohr-Udilova N, Klinglmüller F, Schulte-Hermann R, Stift J, Herac M, Salzmann M, Finotello F, Timelthaler G, Oberhuber G, Pinter M. Deviations of the immune cell landscape between healthy liver and hepatocellular carcinoma. Scientific reports. 2018;8(1):1–11.
14. Huang R, Meng T, Chen R, Yan P, Zhang J, Hu P, Zhu X, Yin H, Song D, Huang Z: The construction and analysis of tumor-infiltrating immune cell and ceRNA networks in recurrent soft tissue sarcoma. Aging (Albany NY). 2019; 11(22).
15. Ciferri C, Pasqualato S, Screpanti E, Varetti G, Santaguida S, Reis GD, Maiolica A, Polka J, Luca J, Wulf PD. Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex. Cell. 2008;133(3):427–39.
16. Lampert F, Mieck C, Alushin GM, Nogales E, Westermann S. Molecular requirements for the formation of a kinetochore–microtubule interface by Dam1 and Ndc80 complexes. J Cell Biol. 2013;200(1):21–30.

17. Linton A, Cheng YY, Griggs K, Kirschner MB, Gattani S, Srikaran S, Kao CH, Mccaughan BC, Klebe S, Zandwijk NV. An RNAi-based screen reveals PLK1, CDK1 and NDC80 as potential therapeutic targets in malignant pleural mesothelioma. Br J Cancer. 2014;110(2):510–9.
18. Reid S, Plaja A, Firth H, Fitzpatrick D, Kidd A, Méhes K, Nash R, Robin N, Hanks S, Coleman K. Constitutional aneuploidy and cancer predisposition caused by biallelic mutations in BUB1B. Nat Genet. 2004;36(11):1159–61.
19. BUB1B promotes. hepatocellular carcinoma progression via activation of the mTORC1 signaling pathway. Cancer Medicine. 2020; 9(21).
20. Ino Y, Yamazaki-Itoh R, Shimada K, Iwasaki M, Kosuge T, Kanai Y, Hiraoka N. Immune cell infiltration as an indicator of the immune microenvironment of pancreatic cancer. British Journal of Cancer.
21. Chen L, Zou W, Zhang L, Shi H, Li Z, Ni C. ceRNA network development and tumor-infiltrating immune cell analysis in Hepatocellular Carcinoma. Med Oncol. 2021;38(7):1–14.
22. Wang Z, Li Y, Zhong Y, Wang Y, Peng M. Comprehensive analysis of aberrantly expressed competitive endogenous rna network and identification of prognostic biomarkers in pheochromocytoma and paraganglioma. OncoTargets therapy. 2020;13:11377.
23. Zhang X, Qin Z, Chen W, Zhou H, Wang E. CCDC106 promotes non-small cell lung cancer cell proliferation. Oncotarget. 2017; 8(16).
24. Wei G, Wei L, Fan Y, Ye Z, Zhang L. Overexpression of CCDC34 in colorectal cancer and its involvement in tumor growth, apoptosis and invasion. Mol Med Rep. 2017;17(1):465–73.
25. Guo L, Li B, Lu Z, Liang H, Yang H, Chen Y, Zhu S, Zeng M, Wei Y, Liu T. CCDC137 is a prognostic biomarker and correlates with immunosuppressive tumor microenvironment based on pan-cancer analysis. Frontiers in molecular biosciences. 2021; 8.
26. Zhang F, Bieniasz PD. HIV-1 Vpr induces cell cycle arrest and enhances viral gene expression by depleting CCDC137. Elife. 2020;9:e55806.
27. Lyssiotis CA, Kimmelman AC. Metabolic interactions in the tumor microenvironment. Trends in cell biology. 2017;27(11):863–75.
28. Van Herck MA, Weyler J, Kwanten WJ, Dirinck EL, De Winter BY, Francque SM, Vonghia L. The differential roles of T cells in non-alcoholic fatty liver disease and obesity. Frontiers in immunology. 2019;10:82.
29. Brown ZJ, Fu Q, Ma C, Kruhlak M, Zhang H, Luo J, Heinrich B, Yu SJ, Zhang Q, Wilson A. Carnitine palmitoyltransferase gene upregulation by linoleic acid induces CD4 + T cell apoptosis promoting HCC development. Cell death disease. 2018;9(6):1–14.
30. Tran E, Turcotte S, Gros A, Robbins PF, Lu Y-C, Dudley ME, Wunderlich JR, Somerville RP, Hogan K, Hinrichs CS. Cancer immunotherapy based on mutation-specific CD4 + T cells in a patient with epithelial cancer. Science. 2014;344(6184):641–5.
31. Liu MX, Juan L, Ming X, Gao ZK, Wang XH, Zhang Y, Shang MH, Yin LH, Pu YP, Ran L. miR-93-5p transferred by exosomes promotes the proliferation of esophageal cancer cells via intercellular communication by targeting PTEN. Biomedical environmental sciences. 2018;31(3):171–85.
32. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinform. 2008;9(1):1–13.
33. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P: The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic acids research. 2016;gkw937.
34. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B. Integration of biological networks and gene expression data using Cytoscape. Nature protocols. 2007;2(10):2366–82.
35. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics: a journal of integrative biology. 2012;16(5):284–7.
36. Chou C-H, Shrestha S, Yang C-D, Chang N-W, Lin Y-L, Liao K-W, Huang W-C, Sun T-H, Tu S-J, Lee W-H. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic acids research. 2018;46(D1):D296–302.
37. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, Zagganas K, Tsanakas P, Floros E, Dalamagas T. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. Nucleic acids research. 2016;44(D1):D231–8.

38. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA: Profiling tumor infiltrating immune cells with CIBERSORT. Methods in molecular biology (Clifton, NJ). 2018; 1711:243.

Figures

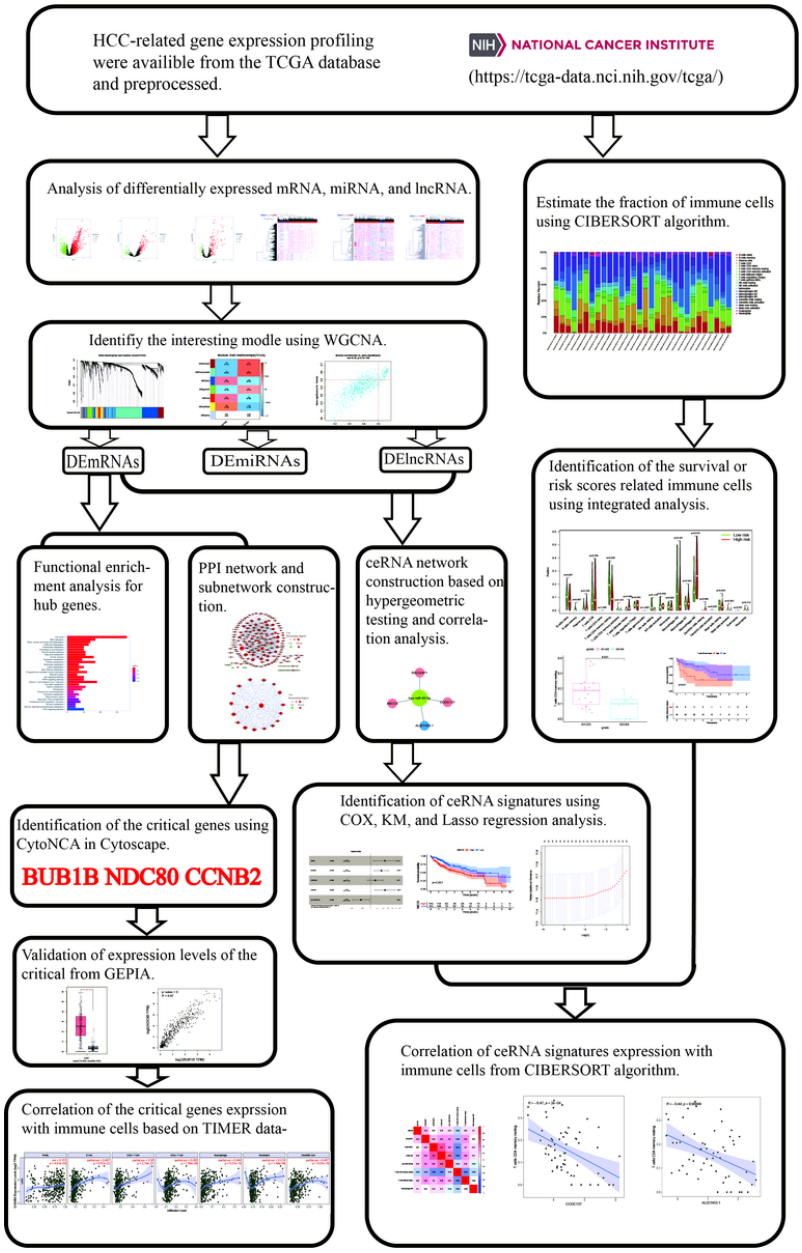

Figure 1

The main workflow of our study. Abbreviations: HCC, Hepatocellular carcinoma; WGCNA, Weighted Gene Co-expression Network Analysis; KM, Kaplan–Meier; PPI, Protein and Protein Interaction; ceRNA, competing endogenous RNA.

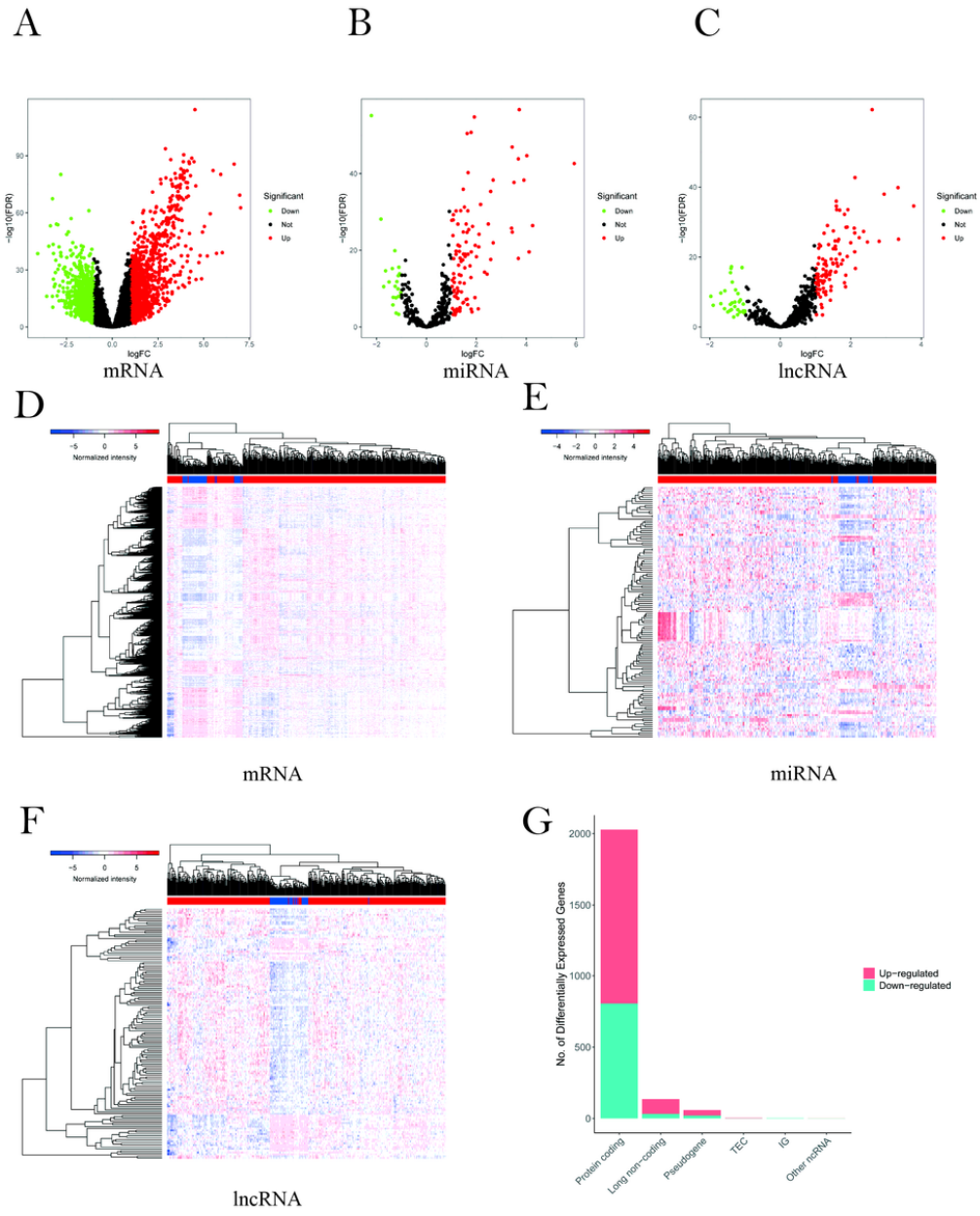

Figure 2

The differentially expressed RNAs in HCC. The volcano plot (A,B,C) and the heatmap (D,E,F) were depicted to show the 2028 DEmRNAs(A,D), 128 DEmiRNAs(B,E), 136 DElncRNAs (C,F) between HCC tissues samples and tumor-adjacent tissues samples; (G) The fraction of of differentially expressed genes. Abbreviations: HCC, Hepatocellular carcinoma; DERNAs, differentially expressed RNAs; DEmRNA: differentially expressed messenger RNAs; DEmiRNAs, differentially expressed microRNAs; DElncRNAs, differentially expressed long noncoding RNAs.

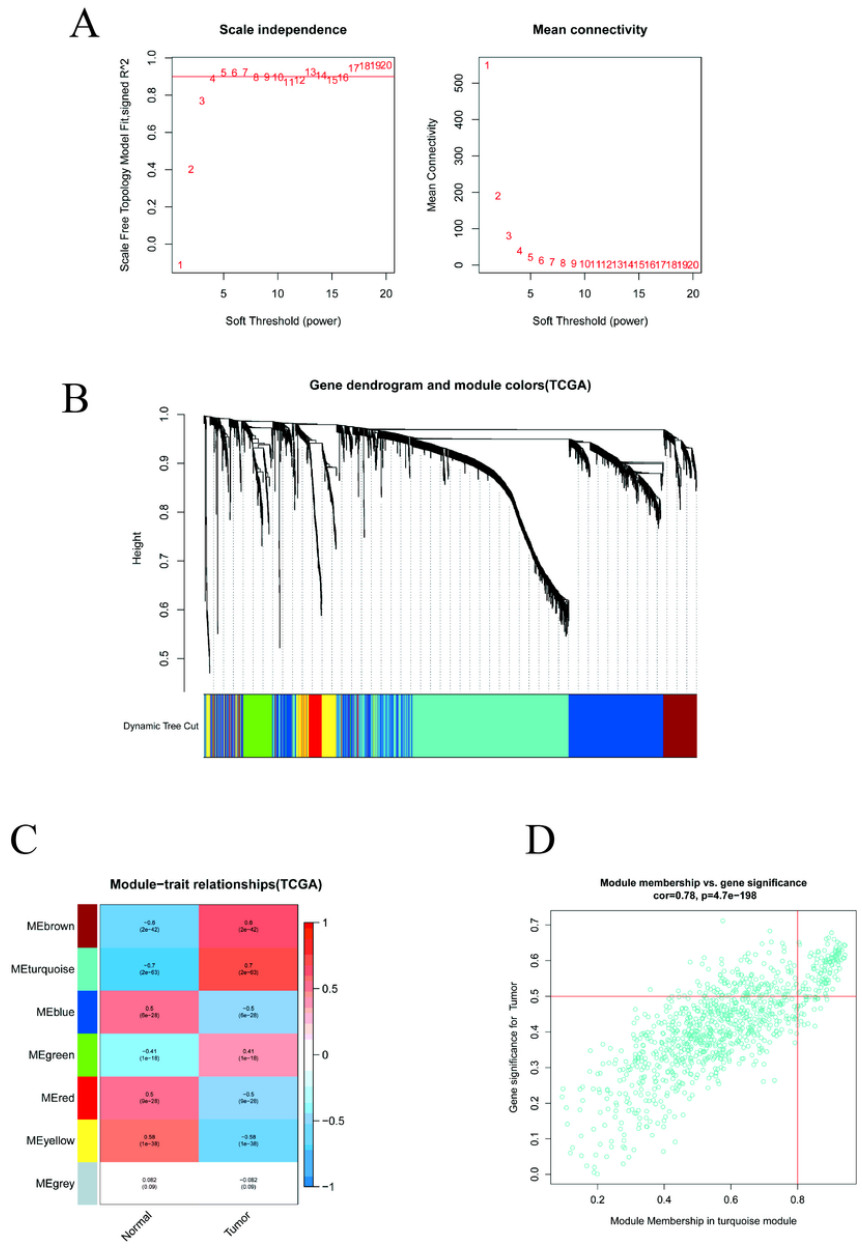

Figure 3

Identification of key module by WGCNA. (A) Scale independence and mean connectivity analysis. (B) Clustering dendrograms of DERNAs, by hierarchical clustering of adjacency-based dissimilarity. (C) The Module-trait relationships plot (D) The associations of turquoise module membership and gene significance for tumor. Abbreviations: WGCNA, weighted gene co-expression network analysis; TOM, topological overlap measure.

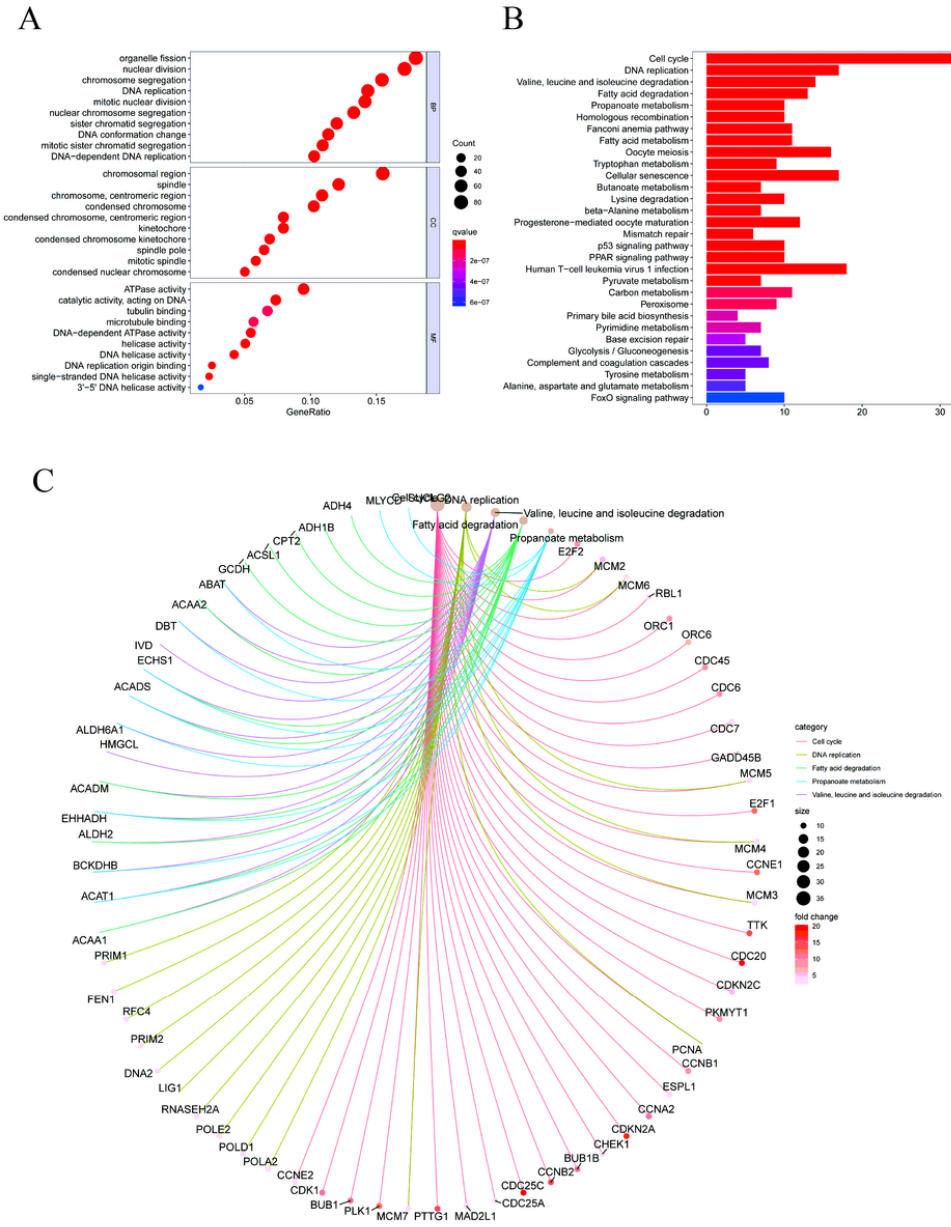

Figure 4

GO and KEGG pathway enrichment analysis. The top 10 enriched GO terms (A) and KEGG pathways (B) for the hub gene in the turquoise module. (C) The top 5 pathways and pathway-gene network. Abbreviations: GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, Biological process; CC, Cellular component; MF, Molecular function.

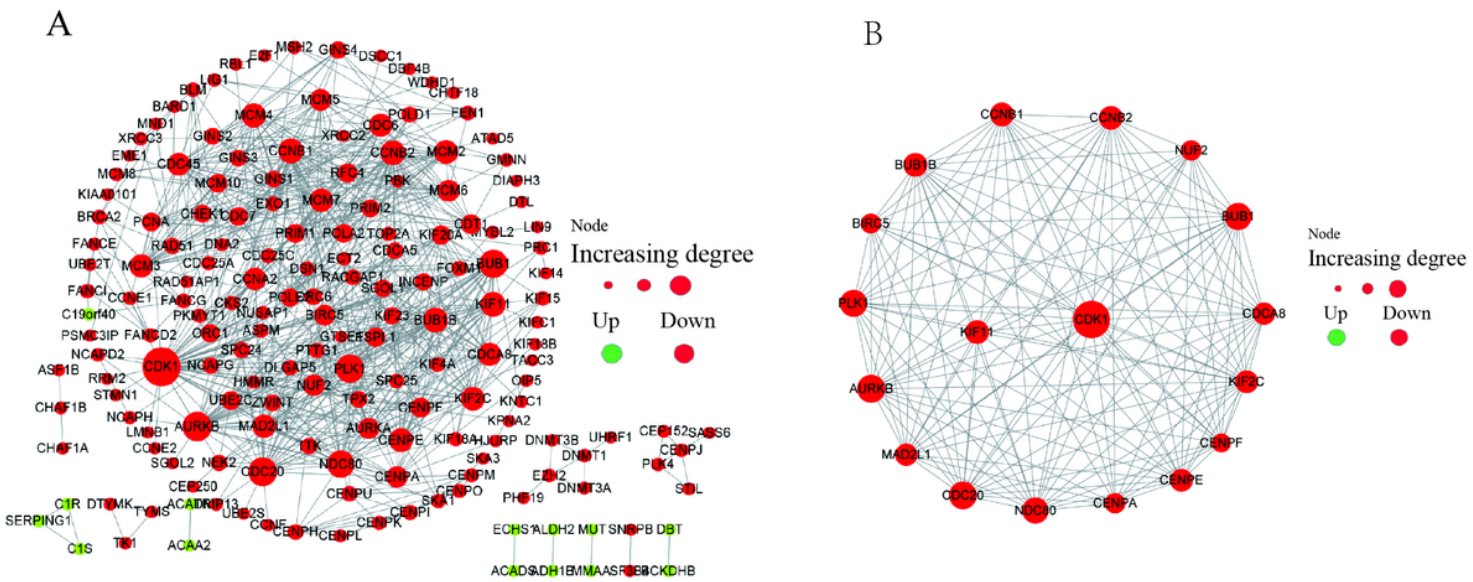

Figure 5

PPI network and corresponding subnetwork1 (A) The hub genes of turquoise module were constructed the PPI network including 175 DEmRNAs. (B) Identification of the subnetwork1 including 18 critical genes by a plugin MCODE (version 2.1.6) in Cytoscape. Abbreviations: DERNAs, differentially expressed RNAs; PPI, Protein and Protein Interaction;

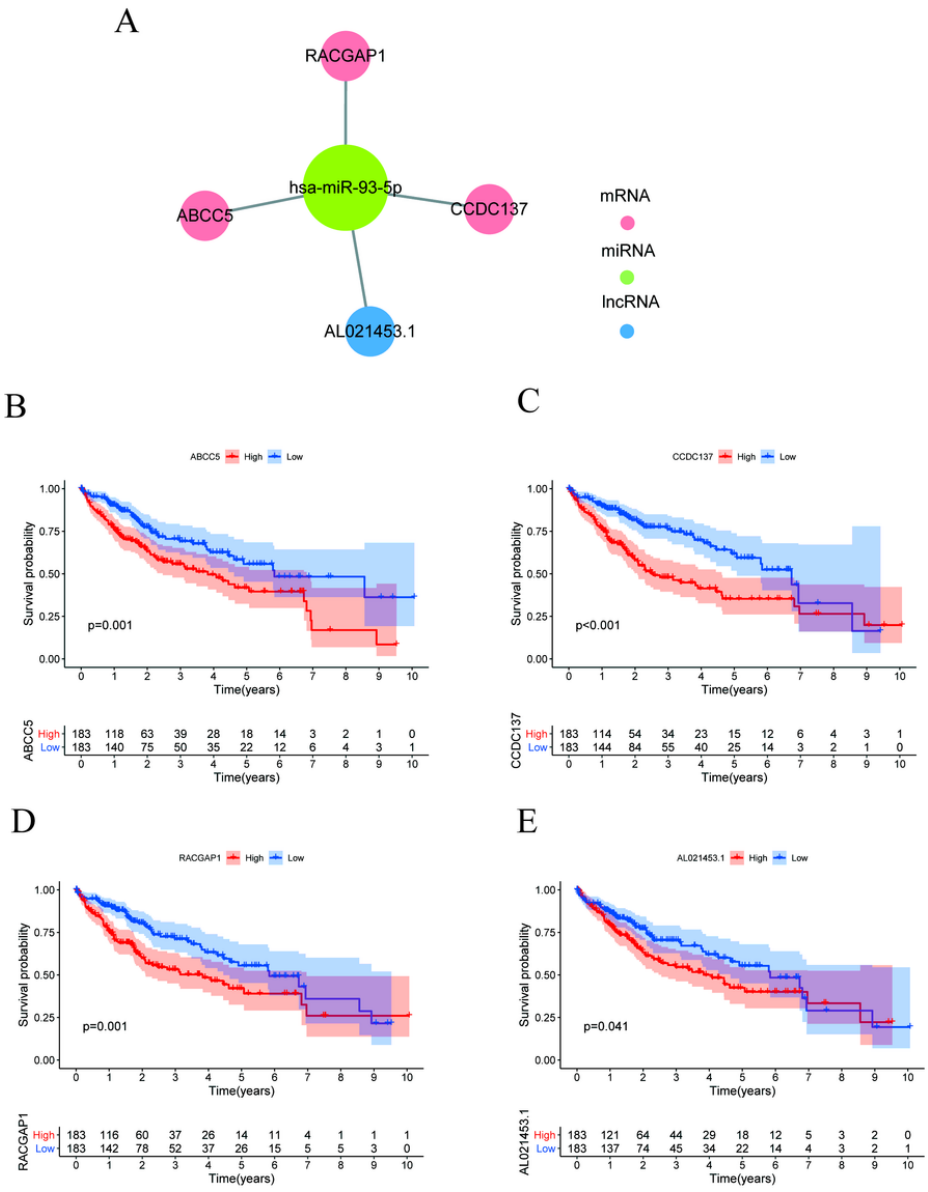

Figure 6

(A) Construction of the HCC-related ceRNA network. The survival curve showed that low ABCC5 (B), CCDC137(C), RACGAP1 (D), and AL021453.1 (E) expression associated with HCC patients' higher OS rate. Abbreviations: ceRNA, competing endogenous RNA network; DEmRNA, differentially expressed messenger RNAs; DEmiRNAs, differentially expressed microRNAs; DElncRNAs, differentially expressed long noncoding RNAs.

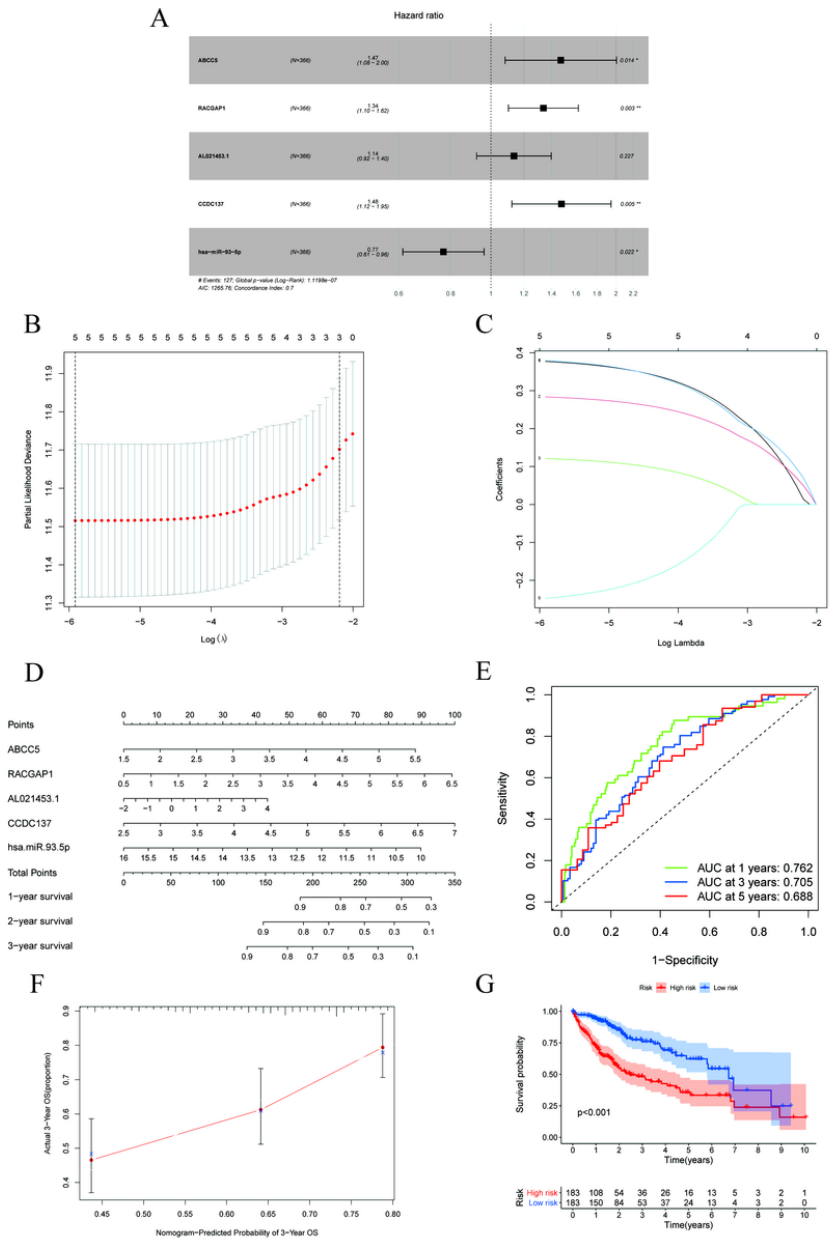

Figure 7

(A) Cox proportional hazards regression mode (forest plots), including 3 mRNAs, 1 miRNAs, and 1 lncRNA. (B, C) Lasso regression analysis demonstrated that all five ceRNA signatures were important for modeling. (D) Nomogram was established based on the the multivariable model. (E,F) The ROC curve and the calibration diagram manifested that the accuracy and discrimination of risk prediction model. (G) Kaplan-Meier analysis showed that the HCC patients with high risk score had a lower OS rate. Abbreviations: Lasso, least absolute shrinkage and selection operator; ROC: receiver operating characteristic; AUC, Area Under Curve.

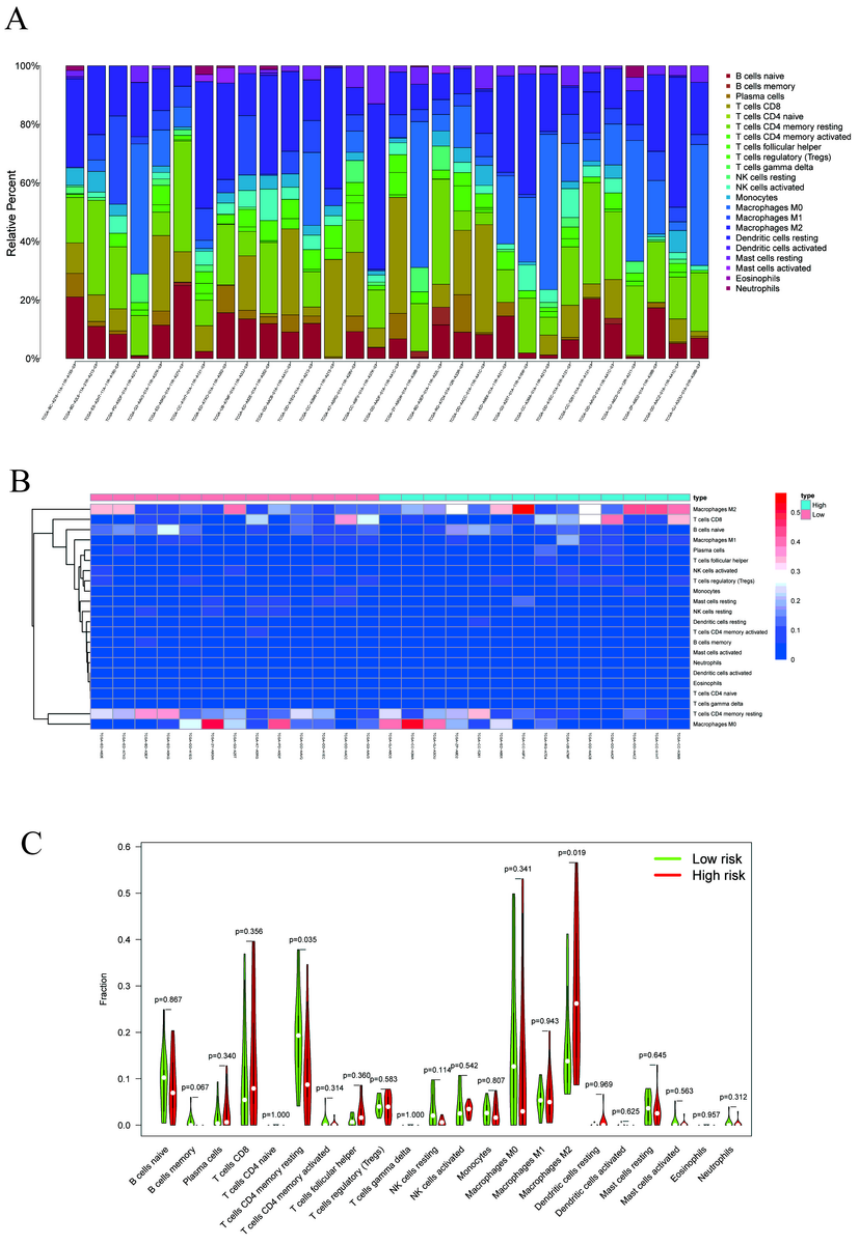

Figure 8

(A) The fractions of immune cells estimated by CIBERSORT algorithm in HCC patients, were annotated by different colors. The heatmap (B) and violin plot (C) exhibited the differential distributions of immune cells between high- and low- risk groups. (C) Macrophages M2 (P=0.019) and T cells CD4 memory resting (P=0.035) showed that the opposite immune infiltration level in in high- and low- risk groups. Abbreviations: CIBERSORT; Cell type identification by estimating relative subsets of RNA transcripts.

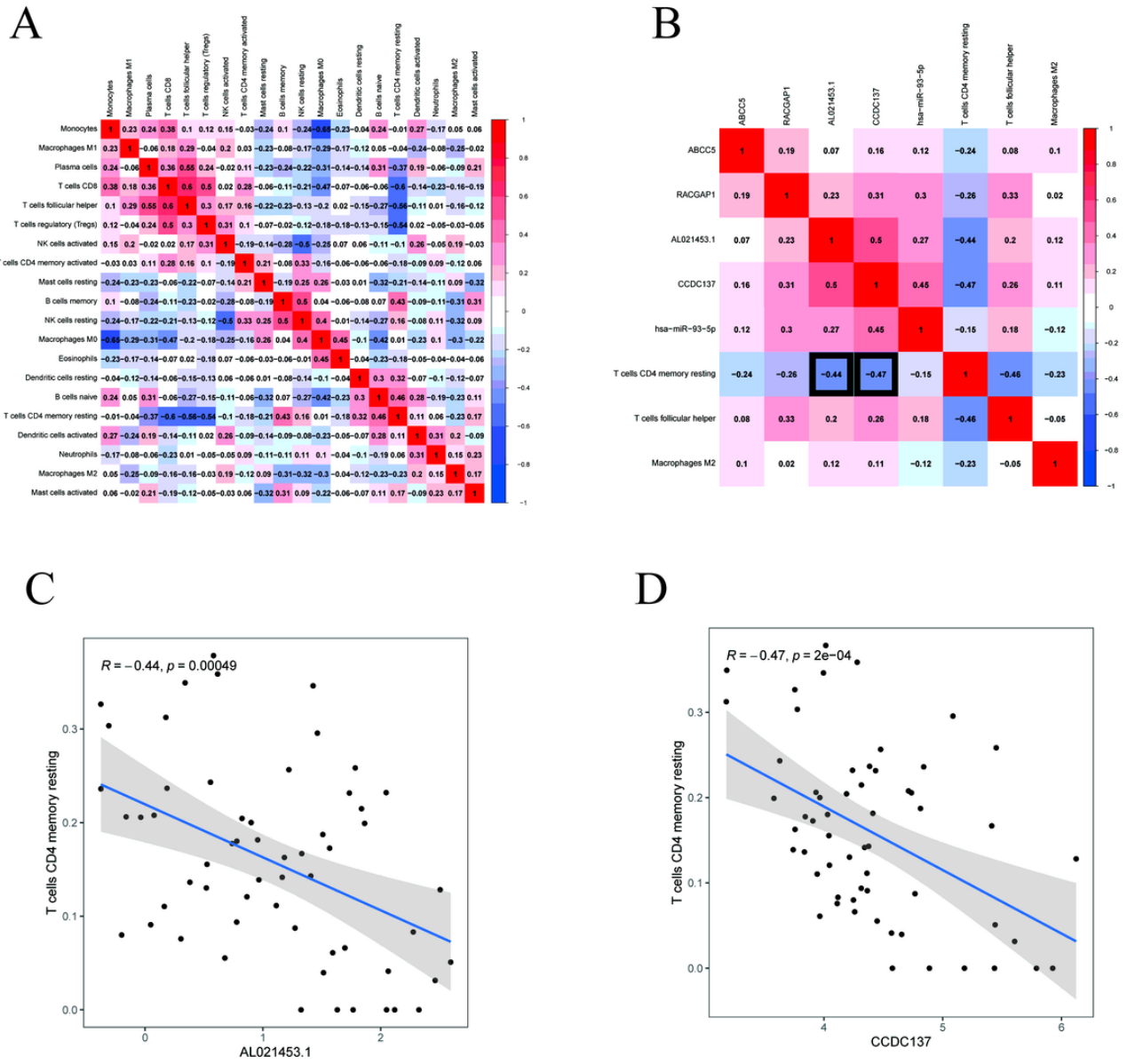

Figure 9

Correlation of ceRNA signatures with the fraction of T cells CD4 memory resting. (A) co-expression analysis among 22 types immune cells. (B) Correlation of ceRNA signatures with the fraction of survival or (and) risk score related immune cells. (C, D) AL021453.1 (R = -0.44, P = 0.00049) and CCDC137(R = -0.47, P = 2e-04) expression were significantly correlated with T cells CD4 memory resting. P<0.001 and |R|>0.3 cut-off threshold. Abbreviations: ceRNA, competing endogenous RNA.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS1.tif
- FigureS2.tif
- FigureS3.tif
- TableS1.ThedifferentiallyexpressedmRNAsofHCC.xlsx
- TableS2.ThedifferentiallyexpressedmiRNAsofHCC.xlsx

- TableS3.ThedifferentiallyexpressedLncRNAsofHCC.xlsx