Original Article

Detection of Prognostic Biomarkers for Hepatocellular Carcinoma through CircRNA-associated CeRNA Analysis

Li Han[1], Maolong Wang[2], Yuling Yang[3], Hanlin Xu[2], Lili Wei[1] and Xia Huang[1*]

[1]Department of Nursing, Affiliated Hospital of Qingdao University, Qingdao, Shandong, China; [2]Department of Thoracic Surgery, The Affiliated Hospital of Qingdao University, Qingdao, Shandong, China; [3]Department of Infectious diseases, The Affiliated Hospital of Qingdao University, Qingdao, Shandong, China

Abstract

Background and Aims: The prognosis of hepatocellular carcinoma (HCC) is extremely poor; therefore, there is an urgent need for novel prognostic molecular biomarkers of HCC. The current investigation utilized circular (circ)RNA-associated competing endogenous (ce)RNAs analysis in order to identify significant prognostic biomarkers of HCC. ***Methods:*** CircRNAs and mRNAs that were differentially expressed between normal and HCC tissues were identified. Their respective functions were predicted with Gene Ontology enrichment and Kyoto Encyclopedia of Genes and Genomes enrichment analyses. A nomogram was used for model verification. ***Results:*** A ceRNA network composed of differentially expressed circRNAs and mRNAs was constructed. Significant hub nodes in the ceRNA network were hsa_circ_0004662, hsa_circ_0005735, hsa_circ_0006990, hsa_circ_0018403 and hsa_circ_0100609. By using this information, a prognostic risk assessment tool was developed based on the expressions of seven genes (PLOD2, TARS, RNF19B, CCT2, RAN, C5orf30 and MCM10). Furthermore, multivariate Cox regression analysis revealed risk and T-stage parameters as independent prognostic factors. The nomograms that were constructed from risk and T-stage groups were used to further assess the prediction of HCC patient survival rates. The nomogram, which consisted of risk and T-stage scores assessment models, was found to be an independent factor for predicting prognosis of HCC. ***Conclusions:*** Five circRNAs, including hsa_circ_0004662, hsa_circ_0005735, hsa_circ_0006990, hsa_circ_0018403 and hsa_circ_0100609, that may play key roles in the progression of HCC were identified. Seven gene signatures were identified, which were associated with the aforementioned circRNAs, including PLOD2, TARS, RNF19B, CCT2, RAN, C5orf30 and MCM10, all of which were significant genes involved in the pathophysiology of HCC. These genes may be used as a prognosticating tool in HCC patients.

Introduction

Hepatocellular carcinoma (HCC) ranks sixth among frequently encountered malignant tumors and is the third most common cause of cancer-associated mortality.[1] The global incidence of HCC has increased in recent years.[2] The most common causes of this cancer include alcohol abuse, aflatoxin exposure and hepatitis B and C viral infections.[3] These risk factors trigger a cascade of genetic and epigenetic changes that accumulate and result in the activation of oncogenes and in the inhibition of the function of tumor suppressor genes, leading to the development of HCC over time.[4] A total of 42,030 subjects were diagnosed with HCC in 2019 in the United States, of which 31,780 died due to this disease.[5] The current mortality to morbidity ratio of HCC is estimated to be 0.95, implying an extremely poor prognosis.[6] Therefore, there is an urgent need for novel prognostic molecular biomarkers of HCC.

The serum levels of carcinoembryonic antigen, alpha-fetoprotein, CA125, CA153, CA199, HSP90α and TK1 were significantly increased in patients with HCC.[7] The application of a single tumor marker in detecting HCC in combination with alpha-fetoprotein levels is an interesting approach. The combined detection of alpha-fetoprotein and HSP90 can improve the diagnostic value of alpha-fetoprotein for HCC and the cost is considerably lower. The combined detection of alpha-fetoprotein and HSP90α and CA125, CA153 and carcinoembryonic antigen can significantly improve the sensitivity and specificity of diagnosis.[8] This approach demonstrates optimal clinical value for early screening and early diagnosis of malignant tumors.[9] Although it has been shown that alpha-fetoprotein and other biomarkers are associated with the prognosis of HCC, their cut-off values have not been adequately defined.

Long intergenic non-coding RNA transcripts are regulated by competitive endogenous RNAs (ceRNAs) via competition for shared microRNAs (miRNAs/miRs).[10] Previous studies have documented the involvement of miRNA-mediated ceRNA modulator mechanisms in cancer development, in which circular RNA (circRNA) acts as a sponge for miRNA.

circRNA is an RNA transcript, which resembles mRNA. A limited number of circRNAs have the ability to encode proteins.[11] Ma *et al*.[12] reported that the dysregulation of hsa_circRNA_100290 reversed miR-378a-triggered inhibition of GLUT1 receptors by functioning as a ceRNA in oral squamous cell carcinoma. Jia *et al*.[13] further explored the potential of a three-circRNA signature as a group of therapeutic biomarkers for gastrointestinal stromal tumors. The roles of circRNA were also examined in HCC by several studies. One report highlighted that hsa_circ_0008287 and hsa_circ_0005027 significantly affected both Hippo and ErbB signaling pathways via miR-548c-3p.[14] However, the prognostic value of circRNA-associated ceRNA regulatory networks in HCC remains unclear.

The present study aimed to identify HCC-associated prognostic biomarkers from The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO) database and RNA sequencing through circRNA-associated ceRNA analysis. The data were processed and circRNAs were reannotated prior to differential expression analysis. Subsequently, a ceRNA network was constructed. The functional enrichment and pathway analyses of the target gene mRNAs were carried out. Potentially prognostic genes were screened prior to establishing a prognostic model. These genes were used to perform analysis of independent prognostic factors. Lastly, a line diagram model composed of independent prognostic factors was constructed for survival rate prediction in patients with HCC.

Methods

High-throughput data

Three datasets were extracted from the following sites: the GEO database at http://www.ncbi.nlm.nih.gov/geo; TCGA (https://portal.gdc.cancer.gov/) database; and RNA sequencing library. The dataset of the GEO database (accession number: GSE125469) was uploaded by Luo *et al*.,[15] which included circRNA data derived from three HCC and three paracancerous tissues. The data from TCGA database comprised RNA sequencing data from 369 HCC and 50 paracancerous tissues. The information was analyzed by the limma package and is presented as $\log_2$ (count+1). The extracted mRNA was rich in oligomeric (dT) beads. A RNA sequencing library was constructed and verified using Agilent 2100 before undergoing qRT-PCR quantification. Sample sequencing was performed with an Illumina Hiseq 4000 and subsequently analyzed by Aksomics (Shanghai, China). In addition, clinical demographics and patient survival data were extracted (download time: December 2019).

Patient samples

From June to December 2013, 82 samples of HCC tissues with their pairs of adjacent non-cancerous tissues were supplied by the Department of Hepatobiliary Surgery of Qingdao University Hospital (Supplementary Table 1). All samples were provided with complete clinical information. The patients provided informed consent regarding the use of their resected tissues for the study. As shown in Supplementary Table 2, the prognostic information is counted in detail. The protocols were approved by the Ethics Committee of Qingdao University Hospital (approval number QY-FYWZLL-25568). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Liquid nitrogen was used to freeze and preserve all samples until RNA extraction was performed. The 5-year prognosis of
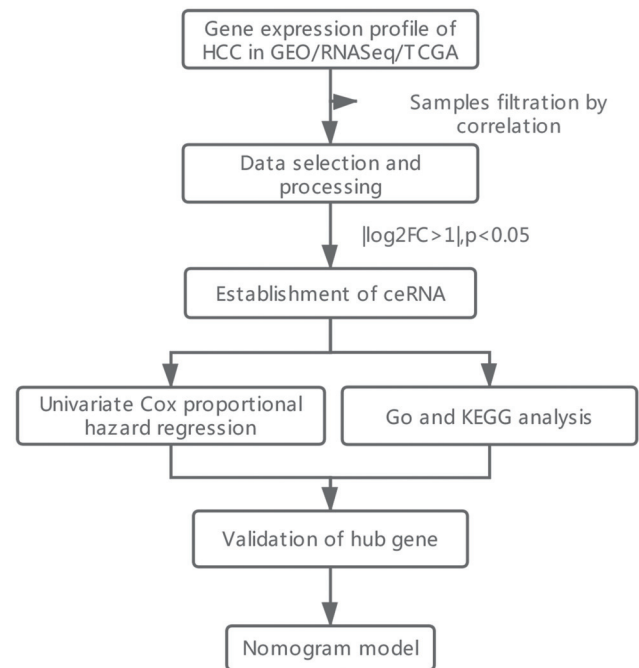
Fig. 1. Flow diagram of the analysis procedure indicating data collection, preprocessing, analysis and validation.

the patients was followed up by telephone after the operation.

Identification of differentially expressed circRNA (DEcircs)

The Limma software package[16] in R was used to determine DEcircs between HCC and adjacent normal tissues. False discovery rate (FDR) <0.05 and fold change >1 were regarded as the threshold of differential RNA expression. Figure 1 depicts the process of bioinformatics analysis. Heatmap and volcanic maps were illustrated with the ggplot2[17] software package in R.

ceRNA network construction and prediction of the target miRNA and circRNA

Bioinformatic analysis was used to predict the miRNA-binding sites of circRNAs, including circbank (http://www.circbank.cn/) and CircInteractome (https://circinteractome.nia.nih.gov/). Subsequently, a total of 10 potential miRNAs were selected. The targets of miRNAs were predicted and a possible association between miRNA and the targeted gene was identified using TargetScan (http://www.targetscan.org/vert_72/) and miRanda (http://www.microrna.org/). CircRNAs exert a 'sponge effect' on miRNAs and regulate their expression and activity, indicating that both circRNAs and mRNAs could be regulated by miRNAs in a competitive manner. Based on this hypothesis, a circRNA-miRNA-mRNA ceRNA network was constructed.

However, the ceRNA network construction and pairwise analysis are associated with several indirect interactions. These associations were eliminated with the use of HERMES (http://califano.c2b2.columbia.edu/hermes/).[18] Lastly, the circRNA-miRNA-mRNA pairs were combined with circRNA and mRNA coexpression (correlation coefficient t>0.6) to

produce a final ceRNA network, which was illustrated using the Cytoscape version 3.7.1.

Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses of DEcircs

In order to obtain a more comprehensive representation of gene and protein functions, a cluster profiler package was used to annotate, visualize and integrate the discovery of the genes.[19] KEGG is a resource dedicated to the detailed representation of functional and biological processes derived from large-scale molecular datasets generated using high-throughput experimental techniques.[20] GO is a verified genetic annotation tool that can also analyze target gene biological functions.[21] A *p*-value <0.05 was considered to indicate significant differences.

Establishment and validation of the prognostic gene signature

The intersection of genes differentially expressed between the three cohorts was used to build a survival prediction signature. The screening of differentially expressed mRNAs with prognostic value was performed in the ceRNA network. First, patients were grouped into low and high expression groups based on the median expression of each mRNA. The survival curves were constructed based on Kaplan-Meier survival analysis. The logarithmic rank test was used to calculate the *p*-values. The mRNAs were screened according to *p*<0.05. Univariate Cox regression analysis was used for screening mRNAs associated with prognosis.

SurvExpress is a tool used to verify biomarkers and to perform online survival analysis and was used to fit these prognostic genes from TCGA cohort into a multivariate Cox regression model. The risk score was calculated as follows: $=risk\ score=expr\ gene_1 \times \beta_1 gene_1 + expr\ gene_2 \times \beta_2 gene_2 + ... expr\ gene_n \times \beta_n gene_n$, where expr was the expression value of each gene and β indicated the prognostic correlation coefficient of each gene in each multiple Cox regression analysis.

Based on the median risk score, the samples were classified into low-and high-risk cohorts prior to Kaplan-Meier survival analysis and log-rank tests. The most significant prognostic model was the risk score model with the lowest *p*-value.

Independent prognostic factors

In order to determine prognosis-associated independent factors, univariate Cox regression analysis was performed on age, sex, TNM stage and tumor histological grade. The factors that were determined by *p*<0.05 were also included in the multivariate Cox regression analysis. The factor is regarded as an independent prognostic factor, if *p*<0.05 in multiple regression analysis.

Construction and verification of nomogram model for predicting the survival rate of HCC patients

A nomogram model was built that included description of independent prognostic factors. In order to verify the predictive ability of the line diagram, the consistency index (C-index), which was composed of independent prognostic factors and the nomogram model (fitted by T-stage and risk group and Coxph model), was calculated. Significant *p*-values were calculated using the resampling technique. The

fitting degree of independent prognostic factors and compound factors was assessed by the Coxph model and the calibration curve was derived from the component with the lowest *p*-value. The model with the best predictive ability was the model with an approximate calibration curve of 45°.

Real-time quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

Total RNA was extracted using the TRIzol® reagent (Invitrogen, Thermo Fisher Scientific, Inc., Waltham, MA, USA) from HCC tissues. qRT-PCR was performed using the SYBR Green PCR Kit (Takara Biotechnology Co., Ltd., Dalian, China). Specific primers (Tsingke Biological Technology, Qingdao, China) were used, which are listed in Supplementary Table 1. The steps were as follows: 95°C for 5 m, 95°C for 30 s, 60°C for 20 s and 72°C for 30 s. A total of 30 cycles were used. A final extension was performed at 72°C for 3 m. GAPDH was used as a positive control and all procedures were repeated three times. The relative expression levels of circRNAs were compared to those of internal controls and were analyzed by the $2^{-\Delta\Delta Ct}$ method.

Results

DEcircs screening in HCC

A total of 86 DEcircs were identified from GSE125469 with a fold-change of >1, of which 33 were down-regulated and 53 were up-regulated. A total of 48 up-regulated and 37 down-regulated DEcircs were identified between HCC and normal samples by RNA sequencing. Five circRNAs were common between the two datasets and were differentially expressed between HCC and normal samples as follows: hsa_circ_0004662, hsa_circ_0005735, hsa_circ_0006990, hsa_circ_0018403 and hsa_circ_0100609. The sequencing analysis was correct according to the database (Supplementary Fig. 1). The volcano plot and heatmap of the gene expressions are illustrated in Figure 2A, B and D. The intersection of the two datasets is shown in Figure 2C. The position of DEcircs is comprehensively displayed in the circos plot (Fig. 2E).

### Prediction of miRNA and ceRNA network construction

The circRNAs, miRNAs and mRNAs that were differentially expressed between HCC samples and adjacent normal tissues were extracted from the GEO and TCGA databases (absolute fold change >1; *p*<0.05). The current analysis resulted in a total of five circRNAs, ten miRNAs and seven hundred and seventy-two mRNAs that were coexpressed in the ceRNA network of HCC (Fig. 3).

Functional enrichment and pathway analyses for differentially expressed mRNAs

The dysfunctional mRNAs were further enriched and analyzed by clusterProfiler. The up-regulated mRNAs that were significantly enriched in the top 20 GO-biological process (BP) terms included regulation of RNA splicing, regulation of cell cycle arrest, maintenance of cell number and homeostasis of number of cells as well as KEGG pathways, including metabolic pathways, the cAMP signaling pathway, the MAPK signaling pathway and cancer-associated pathways (Fig. 4).
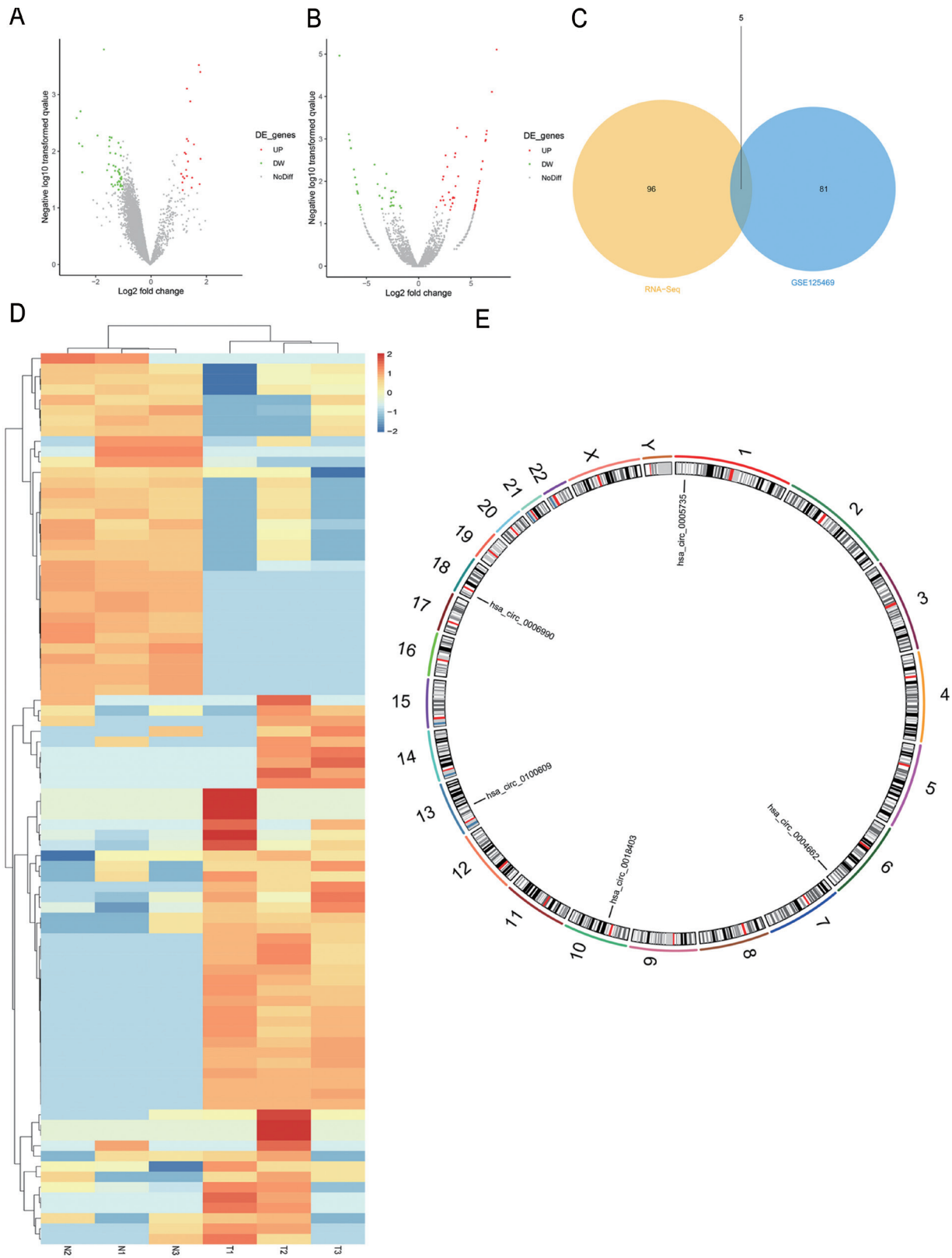
Fig. 2. Identification of DEcircs. (A, B) Volcano plot constructed using fold-change values and *p*-values for GSE125469 and RNA-Sequencing. (C) Intersection of the two datasets. (D) Cluster analysis of differentially expressed genes between T and N samples of HCC. Red represents up-regulated genes, green represents down-regulated genes, and white represents undifferentiated genes. (E) Circos plots indicating the specific chromosomal positions of DEcircs. N, normal; T, tumor.
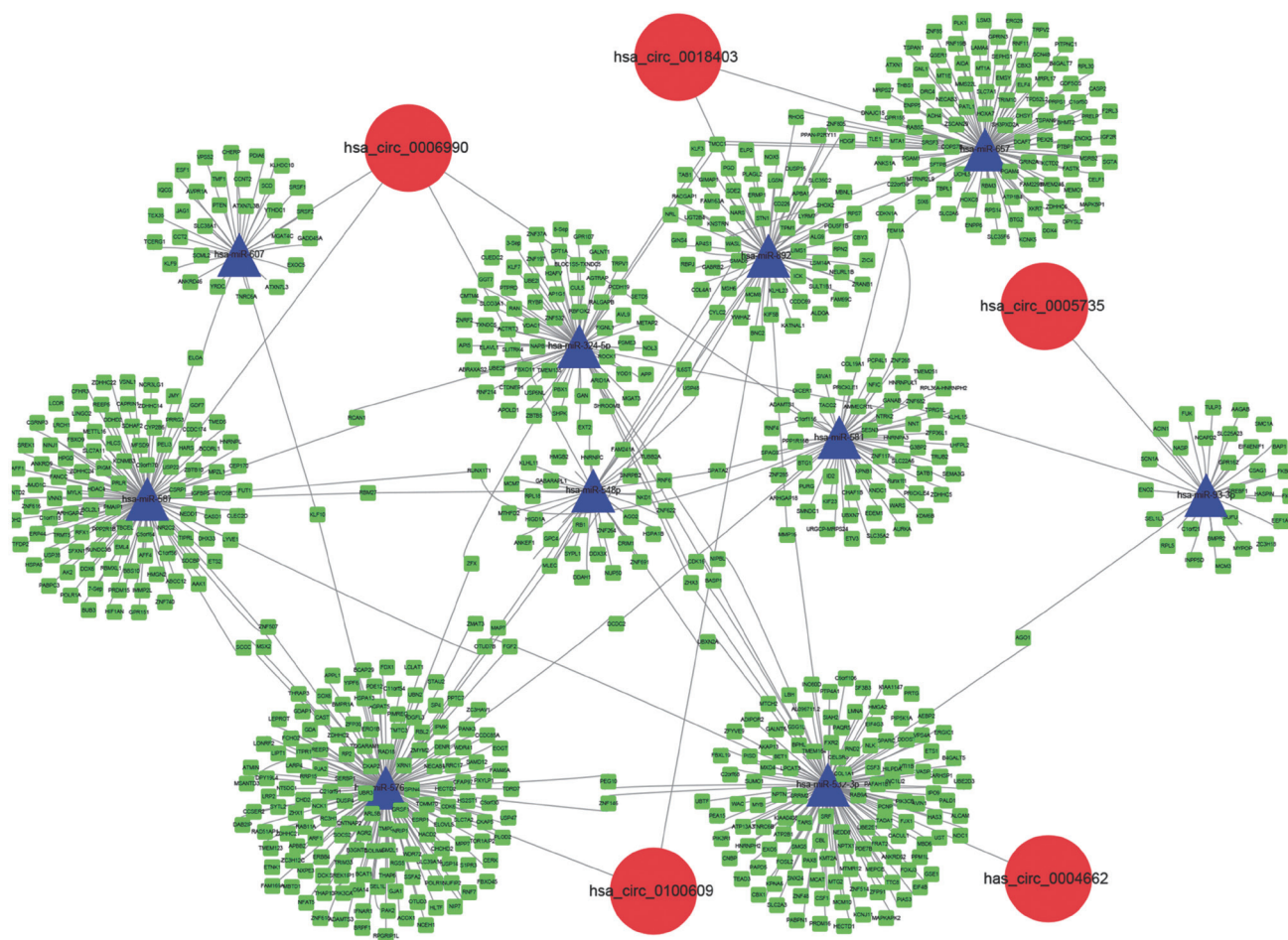
Fig. 3. ceRNA network. Red circle denotes deregulated circRNA. Blue triangle denotes miRNA. Green square denotes targeted mRNA.

Prognostic gene screening

The mRNA expression levels of 19 genes were compared with patient survival using the Kaplan-Meier plotter. The data indicated that the higher the mRNA expression levels of the genes, the shorter the survival time and the worse the prognosis of the patients. In other words, these 19 genes may be potential prognostic factors.

Univariate Cox regression analysis was performed for these 19 mRNAs and revealed that 7 mRNAs, including RAN, a member of the RAS oncogene family, procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (i.e. PLOD2), chaperonin containing TCP1 subunit 2 (i.e. CCT2), threonyl-tRNA synthetase (i.e. TARS), ring finger protein 19B (i.e. RNF19B), chromosome 5 open reading frame 30 (i.e. C5orf30) and minichromosome maintenance 10 replication initiation factor (i.e. MCM10), exhibited hazard ratios that were higher than 1, indicating their positive role in HCC (Supplementary Table 2).

Prognostic model construction

The patients were grouped into low- and high-risk groups based on their median risk scores. The survival time of the patients in the high risk group was shorter than that in the low risk group (Supplementary Fig. 2 and Fig. 5A). The risk scoring distribution of the high risk and low risk groups is

shown in Figure 5B. The patients of the former group demonstrated lower levels of the seven aforementioned genes in contrast to those of the latter (Fig. 5C). These results were also shown in the analysis of the 5-year survival curve performed in our hospital. The data indicated that the patients with lower risk exhibited a longer survival time (p=0.0003) (Fig. 5D). Based on these findings, it can be concluded that the seven predicted genes were protective against HCC progression.

Independent prognostic factors

Independent prognostic factors were determined using multivariate Cox regression analysis. Univariate and multivariate Cox regression analyses revealed that only the risk group and the T-stage groups exhibited $p<0.05$ (Supplementary Table 3), highlighting the utility of both these groups as independent prognostic factors.

Construction and verification of nomogram model for the survival rate prediction of HCC patients

The annual survival rates at 3 and 5 years were predicted upon construction of a nomogram based on the derived independent prognostic factors. Figure 6A indicates
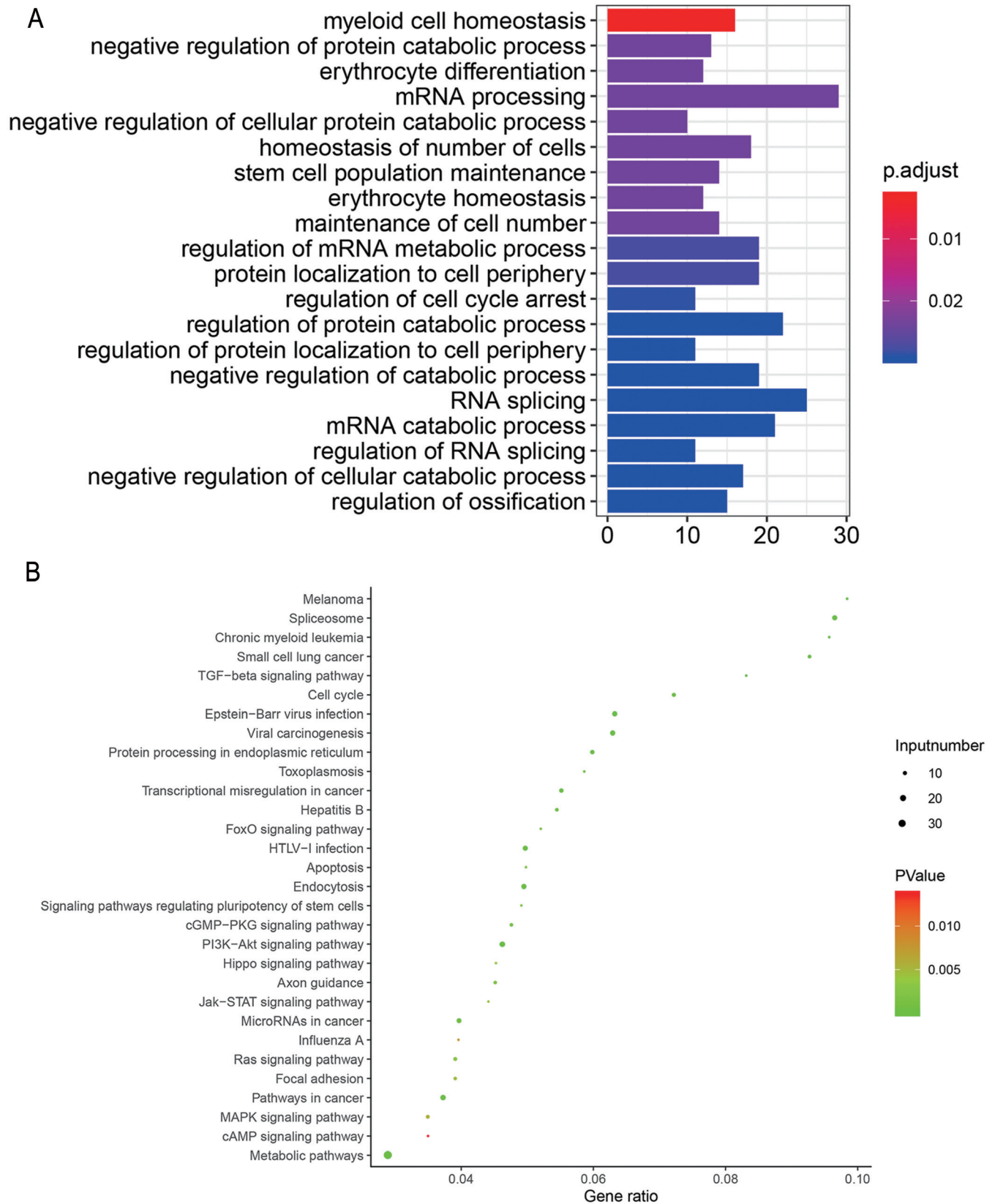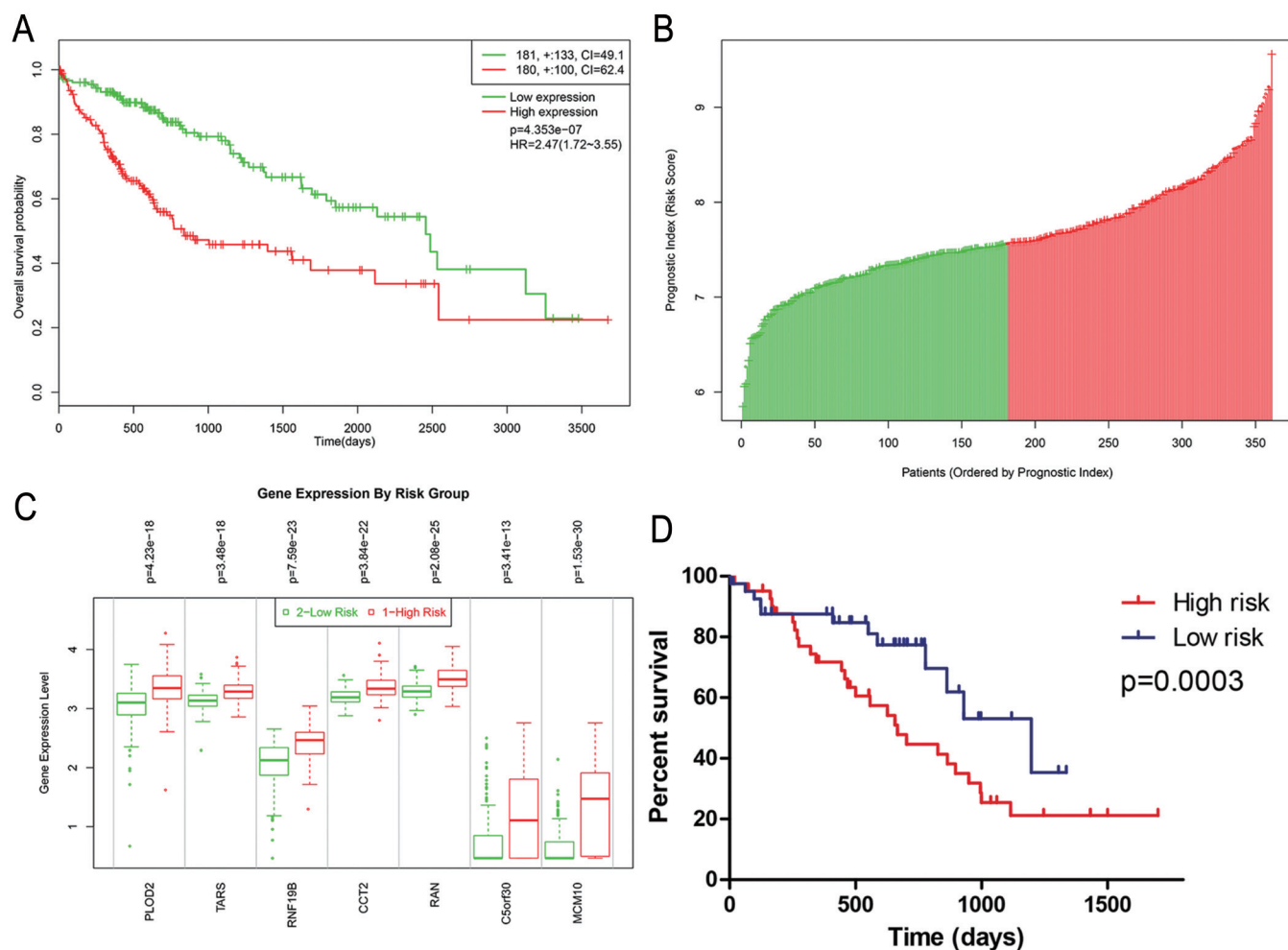
Fig. 4. GO of target genes (A) top 20 terms of GO-BP and (A) top 30 KEGG pathway terms.

Fig. 5. Construction of the seven-gene risk signature of HCC. (A) Kaplan-Meier survival curve for low- and high-risk groups in TCGA. (B) Risk scoring distribution of low- and high-risk groups. (C) Box plots of expression differences of the seven-gene signature in low (green) and high (red) risk groups of TCGA-HCC patients. (D) Kaplan-Meier survival plots demonstrating the association between raised expressions of the seven-gene signature and poor HCC survival outcomes.

a line-type model, comprised of pathological_T and risk groups. According to the nomogram, the 3- and 5-year annual survival rate of HCC patients was accurately predicted based on the T-stage and the risk group. The C-index of the T-stage, the risk group and the nomogram model (T-stage+risk group) were contrasted and the results indicated that the C-index of the risk group and the nomogram model (>0.7) were higher than those of the T-stage. Line diagram models exhibited the lowest *p*-value. The line chart model curve calibration indicated optimal consistency between the predicted annual overall and actual survival rates in 3 and 5 years. This indicated that the line chart model had optimal predictive ability (Fig. 6B).

### Validation of crucial circRNAs and mRNAs

qRT-PCR was used to predict the gene expression levels of hsa_circ_0004662, hsa_circ_0005735, hsa_circ_0006990, hsa_circ_0018403 and hsa_circ_0100609 in HCC and par-acancerous tissues. All genes were differentially expressed in tumor tissues compared to the corresponding expression noted in paracancerous tissues, suggesting their role in HCC development (Fig. 7).

Discussion

The present study analyzed circRNA-associated ceRNAs and revealed important biomarkers, which were associated with HCC prognosis. Hsa_circ_0004662, hsa_circ_0005735, hsa_circ_0006990, hsa_circ_0018403 and hsa_circ_0100609 were the key nodes in the ceRNA network.

A seven-gene prognostic model (PLOD2, TARS, RNF19B, CCT2, RAN, C5orf30 and MCM10) was built and found to correlate significantly with HCC prognosis. Multiple Cox regression analyses demonstrated an association between the prognostic model and the risk group. The T-stage and risk groups were regarded as independent prognostic factors. In addition, the line chart model derived from T-stage and risk groups exhibited optimal predictive ability with regard to the prognosis of HCC.

Previous evidence has shown that the serum levels of carcinoembryonic antigen, alpha-fetoprotein, CA125, CA153, CA199, HSP90α and TK1 in patients with HCC were higher than those in the normal groups. The increase in serum carcinoembryonic antigen levels may be due to the occurrence of HCC, which causes a large number of carcinoembryonic antigen to be stored in the digestive tract and enter the blood circulation. The increase in alpha-fetoprotein levels may be
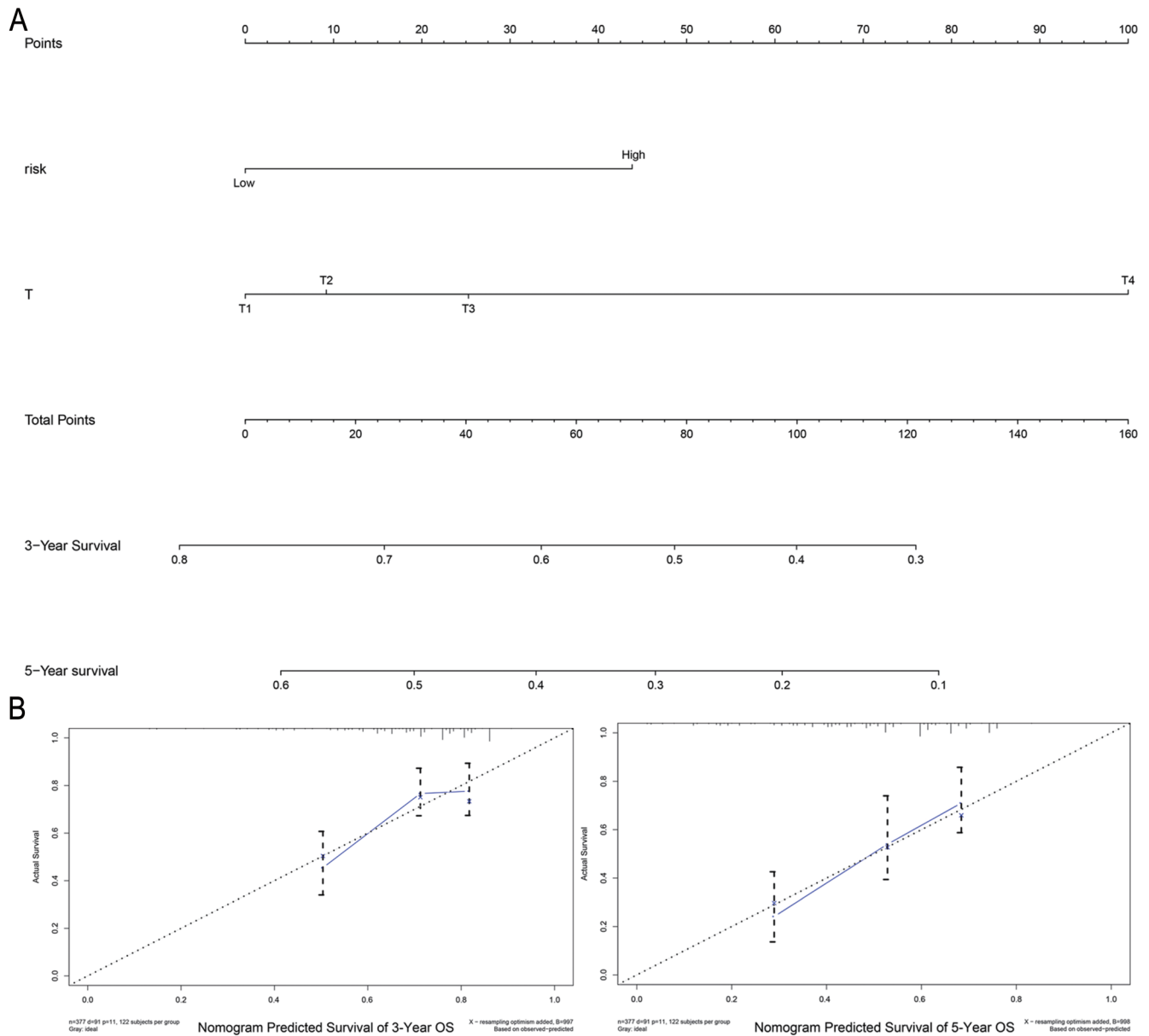
Fig. 6. Construction and verification of nomogram for evaluating OS of HCC patients. (A) Nomogram model, including T-stage and risk group prediction according to 3- and 5-year OS. Two components were present in this nomogram model: the T-stage and the risk group. Depending on the points and lines depicted, different parameters corresponded to different points. (B) The 3- and 5-year OS probability of certain patients can be derived based on the two-component total score of the aforementioned patients. OS, overall survival.

due to hepatocyte carcinogenesis, which results in excess production of alpha-fetoprotein. The increase in CA125 levels may be caused by liver metastasis and ascites. The increase in CA153 levels may be associated with the down-regulation of E-cadherin and of the mitogen-activated protein kinase signaling pathway as well as the invasion and metastasis of HCC. The increase in CA199 levels may be associated with the degree of liver injury, whereas the increase in HSP90α levels may be due to its extracellular secretion and the formation of a complex with HSP70, which is associated with microvascular infiltration, intrahepatic metastasis and HCC progression. The expression of TK1 increases gradually with the proliferation of cancer cells during the development of primary liver cancer. The diagnostic value of alpha-fetoprotein is considered the best, whereas that of CA199 is relatively poor, which may be due to the fact that CA199 has high specificity in the serum of liver cancer patients. Moreover, it exhibits a certain reference value for early diagnosis of primary liver cancer, whereas its lack of sensitivity and its application in the early diagnosis of HCC has some limitations. Although alpha-fetoprotein is the most widely used tumor biomarker for early detection of HCC, its sensitivity is limited. In addition, small liver tumors will cause the expression of alpha-fetoprotein to be lower than the detection limit; meanwhile, when the tumor is larger, the expression of alpha-fetoprotein will be delayed or higher than the detection limit, resulting in alpha-fetoprotein-negative primary liver cancer. Therefore, a single tumor marker is usually insufficient to diagnose HCC,
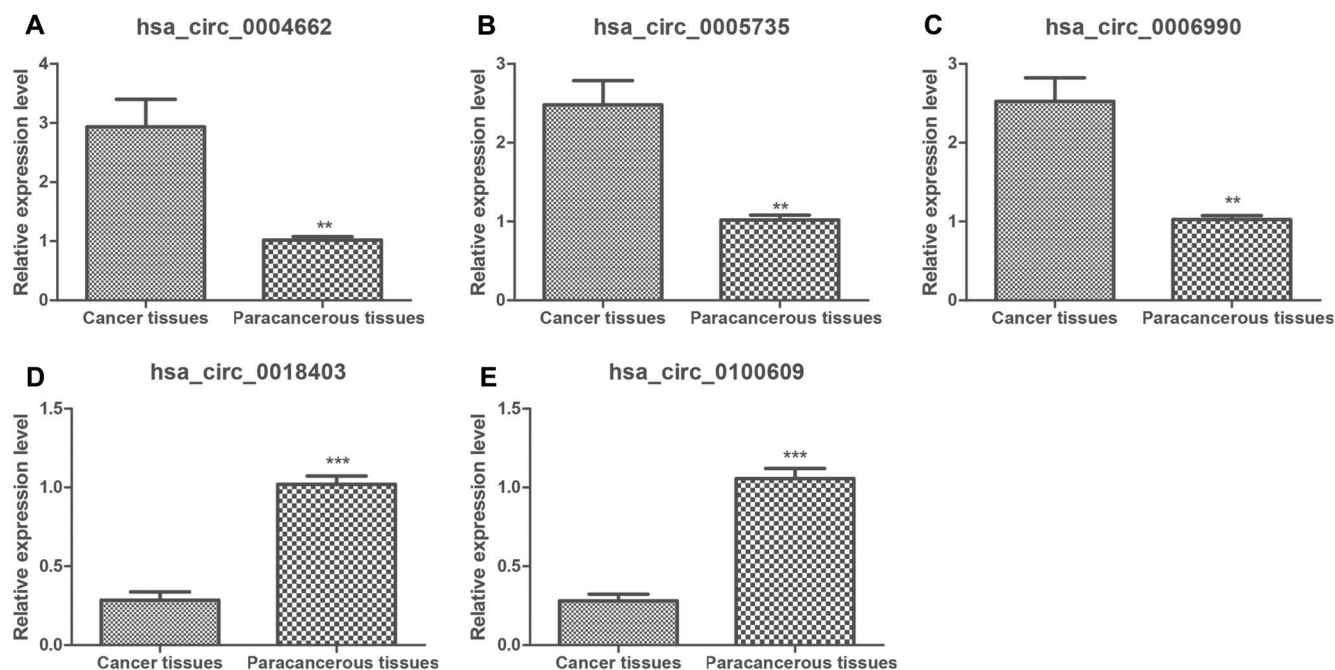
Fig. 7. Relative expression levels of crucial circRNAs detected by qRT-PCR. **p<0.01 and ***p<0.001. Detection of the expression levels of (A) hsa_circ_0004662, (B) hsa_circ_0005735, (C) hsa_circ_0006990, (D) hsa_circ_0018403 and (E) hsa_circ_0100609.

whereas the combined detection of multiple tumor markers can significantly improve the diagnostic value.

In order to obtain more accurate results, differentially expressed mRNAs and circRNAs between paracancerous and cancerous tissues were first identified on three separate platforms, namely GSE125469, RNA sequencing and TCGA. Functional enrichment analyses demonstrated that GO-BP terms, such as regulation of RNA splicing, regulation of cell cycle arrest, maintenance of cell number and cellular homeostasis, and specific KEGG-associated pathways, such as metabolic pathways, the cAMP signaling pathway, the MAPK signaling pathway and cancer-associated pathways, were significantly enriched by disordered genes. It is currently believed that the imbalance between cell numbers and regulatory pathways is crucial for the development of cancer.[22,23] Furthermore, previous studies have shown that the cell cycle and immune response play important roles in cancer development.[24,25] These findings have led to the conclusion that the dysfunctional nature of the identified genes is involved in the predicted GO-BP terms or pathways.

circRNAs have recently been highlighted as a key element in several malignancies. circRNAs typically work by modulating the function of their target genes.[26] In order to assess the role of differentially expressed circRNAs in HCC, a coexpression network was established with the functions of the top seven circRNAs predicted by the KEGG pathway enrichment analysis. Their GO-BP terms were cell adhesion, positive/negative regulation of gene expression and Golgi organization.[26] Fu *et al*.[27] identified that hsa_circ_0005075, hsa_circ_0003570 and hsa_circ_0004018 may be potential diagnostic indicators of HCC. Li *et al*.[28] reported the properties of hsa_circ_0006990 in sponging miR-101 in colorectal cancer.

Although previous studies have demonstrated the significance of molecular biomarkers in cancer, the single use of specific biomarkers is often unable to predict the survival of cancer patients.[29–31] Therefore, the seven-gene signature (PLOD2, TARS, RNF19B, CCT2, RAN, C5orf30 and MCM10) predicted in the present study is a potentially powerful prog-

nosticating model for HCC. The full name of PLOD2 is procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 and has been proven to be associated with tumor size (p=0.022) and macroscopic intrahepatic metastasis.[32] Hsu *et al*.[33] demonstrated that TARS is a candidate oncogene. Park *et al*.[34] suggested that CCT2 inhibited tumor induction via Gli-1. The mutation of RAN is associated with the survival rate of the patients with HCC.[35] C5orf30 was found to suppress CCNH, which is found in multiple cancer types.[36] Yang *et al*.[37] demonstrated that MCM10 promoted the invasion/migration potential of breast cancer cells via the Wnt/β-catenin signaling pathway, which was associated with poor breast cancer prognosis. However, no reports have been previously published on RNF19B. Based on this evidence, a risk model containing all this information may improve the prognosis of HCC.

Subsequent scrutinization of data revealed both risk and T-stage groups to be independent prognostic factors. A nomogram model (T-stage+risk group) based on both these groups was constructed and the data demonstrated satisfactory ability to prognosticate HCC accurately. This nomogram may be used to guide clinical management of HCC patients.

Conclusions

The present study identified five circRNAs, namely hsa_circ_0004662, hsa_circ_0005735, hsa_circ_0006990, hsa_circ_0018403 and hsa_circ_0100609 that may play key roles in the progression of HCC. In addition, seven gene signatures were identified, which were associated with the aforementioned circRNAs, including PLOD2, TARS, RNF19B, CCT2, RAN, C5orf30 and MCM10, all of which were significantly involved in the pathophysiology of HCC and may be used as a prognostic indices in HCC patients. Nevertheless, the clinical implementation of the nomogram combination model requires further refinement, along with the investigation of the exact molecular mechanisms of the five types of circRNAs identified.

Funding

Conflict of interest

The authors have no conflict of interests related to this publication.

Author contributions

Study design (LH), performance of experiments (XH, LH), analysis and interpretation of data (YY, MW), manuscript writing (LLW, LH), critical revision of the manuscript (YY, HX), statistical analysis (XH, LH), critical funding (XH, LH), administration (XH, LH), technical or material support (LW, LH).

Data sharing statement

The data that support the findings of this study are openly available in GSE125469, at https://www.ncbi.nlm.nih.gov/gds; the reference number is 31494761.

References

[1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021;71(1):7–33. doi:10.3322/caac.21654.
[2] EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol 2018;69(1):182–236. doi:10.1016/j.jhep.2018.03.019.
[3] Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;385(9963):117–171. doi:10.1016/S0140-6736(14)61682-2.
[4] Dhanasekaran R, Bandoh S, Roberts LR. Molecular pathogenesis of hepatocellular carcinoma and impact of therapeutic advances. F1000Res 2016;5(F1000 Faculty Rev):879. doi:10.12688/f1000research.6946.1.
[5] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69(1):7–34. doi:10.3322/caac.21551.
[6] Petrick JL, Braunlin M, Laversanne M, Valery PC, Bray F, McGlynn KA. International trends in liver cancer incidence, overall and by histologic subtype, 1978-2007. Int J Cancer 2016;139(7):1534–1545. doi:10.1002/ijc.30211.
[7] Wang T, Zhang KH. New blood biomarkers for the diagnosis of AFP-negative hepatocellular carcinoma. Front Oncol 2020;10:1316. doi:10.3389/fonc.2020.01316.
[8] Li Y, Li DJ, Chen J, Liu W, Li JW, Jiang P, et al. Application of joint detection of AFP, CA19-9, CA125 and CEA in identification and diagnosis of cholangiocarcinoma. Asian Pac J Cancer Prev 2015;16(8):3451–3455. doi:10.7314/apjcp.2015.16.8.3451.
[9] Edoo MIA, Chutturghoon VK, Wusu-Ansah GK, Zhu H, Zhen TY, Xie HY, et al. Serum biomarkers AFP, CEA and CA19-9 combined detection for early diagnosis of hepatocellular carcinoma. Iran J Public Health 2019;48(2):314–322.
[10] Chen K, Xie S, Jin W. Crucial lncRNAs associated with adipocyte differentiation from human adipose-derived stem cells based on co-expression and ceRNA network analyses. PeerJ 2019;7:e7544. doi:10.7717/peerj.7544.
[11] Li M, Ding W, Tariq MA, Chang W, Zhang X, Xu W, et al. A circular transcript of ncx1 gene mediates ischemic myocardial injury by targeting miR-133a-3p. Theranostics 2018;8(21):5855–5869. doi:10.7150/thno.27285.
[12] Ma X, Liu C, Gao C, Li J, Zhuang J, Liu L, et al. circRNA-associated ceRNA network construction reveals the circRNAs involved in the progression and prognosis of breast cancer. J Cell Physiol 2020;235(4):3973–3983. doi:10.1002/jcp.29291.
[13] Jia N, Tong H, Zhang Y, Katayama H, Wang Y, Lu W, et al. CeRNA expression profiling identifies KIT-related circRNA-miRNA-mRNA networks in gastrointestinal stromal tumour. Front Genet 2019;10:825. doi:10.3389/fgene.2019.00825.
[14] Feng C, Li Y, Lin Y, Cao X, Li D, Zhang H, et al. CircRNA-associated ceRNA network reveals ErbB and Hippo signaling pathways in hypopharyngeal cancer. Int J Mol Med 2019;43(1):127–142. doi:10.3892/ijmm.2018.3942.
[15] Luo Z, Mao X, Cui W. Circular RNA expression and circPTPRM promotes proliferation and migration in hepatocellular carcinoma. Med Oncol 2019;36(10):86. doi:10.1007/s12032-019-1311-z.
[16] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47. doi:10.1093/nar/gkv007.
[17] Ito K, Murphy D. Application of ggplot2 to pharmacometric graphics. CPT Pharmacometrics Syst Pharmacol 2013;2(10):e79. doi:10.1038/psp.2013.56.
[18] Talos F, Mitrofanova A, Bergren SK, Califano A, Shen MM. A computational systems approach identifies synergistic specification genes that facilitate lineage conversion to prostate tissue. Nat Commun 2017;8:14662. doi:10.1038/ncomms14662.
[19] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16(5):284–287. doi:10.1089/omi.2011.0118.
[20] Tanabe M, Kanehisa M. Using the KEGG database resource. Curr Protoc Bioinformatics 2012;38:1.12.1–1.12.43. doi:10.1002/0471250953.bi0112s38.
[21] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25(1):25–29. doi:10.1038/75556.
[22] Albanes D, Winick M. Are cell number and cell proliferation risk factors for cancer? J Natl Cancer Inst 1988;80(10):772–774. doi:10.1093/jnci/80.10.772.
[23] Fresno Vara JA, Casado E, de Castro J, Cejas P, Belda-Iniesta C, González-Barón M. PI3K/Akt signalling pathway and cancer. Cancer Treat Rev 2004;30(2):193–204. doi:10.1016/j.ctrv.2003.07.007.
[24] Collins I, Garrett MD. Targeting the cell division cycle in cancer: CDK and cell cycle checkpoint kinase inhibitors. Curr Opin Pharmacol 2005;5(4):366–373. doi:10.1016/j.coph.2005.04.009.
[25] Zhao F, Wang C, Yang Q, Han S, Hu Q, Fu Z. Titanium dioxide nanoparticle stimulating pro-inflammatory responses in vitro and in vivo for inhibited cancer metastasis. Life Sci 2018;202:44–51. doi:10.1016/j.lfs.2018.03.058.
[26] Qiu L, Wang T, Ge Q, Xu H, Wu Y, Tang Q, et al. Circular RNA signature in hepatocellular carcinoma. J Cancer 2019;10(15):3361–3372. doi:10.7150/jca.31243.
[27] Fu L, Jiang Z, Li T, Hu Y, Guo J. Circular RNAs in hepatocellular carcinoma: Functions and implications. Cancer Med 2018;7(7):3101–3109. doi:10.1002/cam4.1574.
[28] Li XN, Wang ZJ, Ye CX, Zhao BC, Huang XX, Yang L. Circular RNA circVAPA is up-regulated and exerts oncogenic properties by sponging miR-101 in colorectal cancer. Biomed Pharmacother 2019;112:108611. doi:10.1016/j.biopha.2019.108611.
[29] Zhu H, Bhaijee F, Ishaq N, Pepper DJ, Backus K, Brown AS, et al. Correlation of Notch1, pAKT and nuclear NF-κB expression in triple negative breast cancer. Am J Cancer Res 2013;3(2):230–239.
[30] Wang Y, Lu T, Wo Y, Sun X, Li S, Miao S, et al. Identification of a putative competitive endogenous RNA network for lung adenocarcinoma using TCGA datasets. PeerJ 2019;7:e6809. doi:10.7717/peerj.6809.
[31] Lu T, Chen D, Wang Y, Sun X, Li S, Miao S, et al. Identification of DNA methylation-driven genes in esophageal squamous cell carcinoma: a study based on The Cancer Genome Atlas. Cancer Cell Int 2019;19:52. doi:10.1186/s12935-019-0770-9.
[32] Noda T, Yamamoto H, Takemasa I, Yamada D, Uemura M, Wada H, et al. PLOD2 induced under hypoxia is a novel prognostic factor for hepatocellular carcinoma after curative resection. Liver Int 2012;32(1):110–118. doi:10.1111/j.1478-3231.2011.02619.x.
[33] Hsu CH, Hsu CW, Hsueh C, Wang CL, Wu YC, Wu CC, et al. Identification and characterization of potential biomarkers by quantitative tissue proteomics of primary lung adenocarcinoma. Mol Cell Proteomics 2016;15(7):2396–2410. doi:10.1074/mcp.M115.057026.
[34] Park SH, Jeong S, Kim BR, Jeong YA, Kim JL, Na YJ, et al. Activating CCT2 triggers Gli-1 activation during hypoxic condition in colorectal cancer. Oncogene 2020;39(1):136–150. doi:10.1038/s41388-019-0972-6.
[35] Kim MN, Kim JO, Lee SM, Park H, Lee JH, Rim KS, et al. Variation in the dicer and RAN genes are associated with survival in patients with hepatocellular carcinoma. PLoS One 2016;11(9):e0162279. doi:10.1371/journal.pone.0162279.
[36] Yu YP, Liu P, Nelson J, Hamilton RL, Bhargava R, Michalopoulos G, et al. Identification of recurrent fusion genes across multiple cancer types. Sci Rep 2019;9(1):1074. doi:10.1038/s41598-019-38550-6.
[37] Yang WD, Wang L. MCM10 facilitates the invaded/migrated potentials of breast cancer cells via Wnt/β-catenin signaling and is positively interlinked with poor prognosis in breast carcinoma. J Biochem Mol Toxicol 2019;33(7):e22330. doi:10.1002/jbt.22330.