The competitive endogenous RNA regulatory network reveals potential prognostic biomarkers for overall survival in hepatocellular carcinoma

Zhiqiao Zhang1 | Jing Li1 | Tingshan He1 | Yanling Ouyang1 | Yiyan Huang1 | Qingbo Liu2 | Peng Wang1 | Jianqiang Ding1

Abstract

The aim of the present study is to construct a competitive endogenous RNA (ceRNA) regulatory network by using differentially expressed long noncoding RNAs (lncRNAs), microRNAs (miRNAs), and mRNAs in patients with hepatocellular carcinoma (HCC), and to construct a prognostic model for predicting overall survival (OS) of HCC patients. Differentially expressed lncRNAs, miRNAs, and mRNAs were explored between HCC tissues and normal liver tissues. A prognostic model was built for predicting OS of HCC patients and receiver operating characteristic curves were used to evaluate the performance of the prognostic model. There were 455 differentially expressed lncRNAs, 181 differentially expressed miRNAs, and 5035 differentially expressed mRNAs. A ceRNA regulatory network was constructed based on 43 lncRNAs, 37 miRNAs, and 105 mRNAs. Eight mRNA biomarkers (H2AFX, SQSTM1, ITM2A, PFKP, TPD52L1, ACSL4, STRN3, and CPEB3) were identified as independent risk factors by multivariate Cox regression and were used to develop a prognostic model for OS. The C-indexes in the model group were 0.776 (95% confidence interval [CI], 0.730–0.822), 0.745 (95% CI, 0.699–0.791), and 0.789 (95% CI, 0.743–0.835) for 1-, 3-, and 5-year OS, respectively. The current study revealed potential molecular biological regulation pathways and prognostic biomarkers by the ceRNA regulatory network. A prognostic model based on prognostic mRNAs in the ceRNA network might be helpful to predict the individual mortality risk for HCC patients. The individual mortality risk calculator can be used by visiting the following URL: https://zhangzhiqiao.shinyapps.io/Smart_cancer_predictive_system_HCC/.

KEYWORDS

competitive endogenous RNA, hepatocellular carcinoma, mRNA, overall survival, prognostic model

Abbreviations: ACSL4, acyl-CoA synthetase long chain family member 4; ceRNA, competitive endogenous RNA; C-index, Harrell’s concordance index; CPEB3, cytoplasmic polyadenylation element binding protein 3; CI, confidence interval; DCA, decision curve analysis; GEO, Gene Expression Omnibus; GO, Gene Ontology; HCC, hepatocellular carcinoma; HR, hazard ratio; KEGG, Kyoto Encyclopedia of Genes and Genomes; lncRNA, long noncoding RNA; miRNA, microRNA; mRNA, messenger RNA; OS, overall survival; ROC, receiver operating characteristic; TARDBP, Tat activating regulatory DNA-binding protein; TCGA, The Cancer Genome Atlas.
1 | INTRODUCTION

Hepatocellular carcinoma is the seventh most prevalent malignant tumor and the third leading cause for tumor-associated mortality. It was estimated that there were 841,080 newly diagnosed HCC patients and 781,631 HCC patients died in 2018. Most HCC patients were diagnosed at an advanced stage because early stage HCC patients often have no obvious clinical symptoms. Although great progress has been made in the diagnosis and treatment of HCC, the 5-year OS rate in advanced stage HCC patients is still less than 20%. It has been reported that the average OS of advanced stage HCC patients who did not receive systematic treatments was merely 71 months. To date, the molecular biological mechanisms affecting the progression and prognosis of HCC have not been fully illuminated. From the point of view of clinical practice, exploration of molecular biological mechanisms affecting the prognosis of HCC and identification of reliable prognostic biomarkers are valuable for optimizing individualized treatments and improving clinical prognosis.

Emerging evidence has revealed potential molecular biological mechanisms and regulatory pathways in the progression and prognosis in HCC patients. The hypothesis of ceRNA, proposed by Salmena et al., depicted a specific molecular biological regulatory mechanism for posttranscriptional regulation. Long non-coding RNAs could indirectly regulate the expression of mRNAs through competitively binding response elements of miRNAs as endogenous molecular sponges. Several studies have explored the molecular biological regulatory mechanisms for prognosis of HCC patients through the ceRNA regulatory network. Lin et al. constructed 3 prognostic prediction models by using IncRNAs, miRNAs, and mRNAs, whereas Bai et al. constructed an IncRNA prognostic model for OS in HCC patients based on 13 IncRNAs identified by a ceRNA regulatory network. However, these prognostic models were limited by the following reasons: complex calculation formulas, results with unclear prognostic significance, and lack of external validation. Therefore, a prognostic model with a simple calculation formula, easily interpreted results, and reliable external repeatability is of great significance for optimizing individualized therapy.

The nomogram, as a graphical display tool, was convenient to calculate and interpret the predictive results. Therefore the aim of the present study is to develop and validate a prognostic nomogram for OS of HCC patients based on gene expression data identified by a ceRNA regulatory network.

2 | MATERIALS AND METHODS

2.1 | Protocol approval

All datasets in the present study were downloaded from public databases, including TCGA and the cBioPortal and GEO database. These public databases allowed researchers to download and analyze public datasets for scientific purposes and thus ethics approval was not required.

2.2 | Gene expression information of the model dataset

We downloaded the original RNA expression dataset from TCGA generated on the Illumina HiSeq 2000 RNA Sequencing platform. The RNA symbols were determined based on GENCODE version 29 (https://www.gencodegenes.org/human/). The RNA expression dataset contained 1093 IncRNAs and 13,356 mRNAs from 371 HCC samples and 50 normal samples. The RNA expression values were standardized by the trimmed mean of M method. The miRNA expression dataset was obtained from the University of California Santa Cruz Xena database (https://xena.ucsc.edu/). The miRNA expression dataset involved 2172 miRNAs from 371 HCC and 49 normal liver samples. The miRNA expression values were standardized with the scale method by the “limma” package.

2.3 | Differentially expressed analyses

The differentially expressed analyses were carried out between HCC tissues and adjacent normal liver tissues, with a defined P of .05 and a ratio of 1.5 times between tumor tissue expression and adjacent normal liver tissue expression.

2.4 | Clinical dataset of the model group

According to the medians of the original gene expression values, the original expression values were converted into “0” for lower expression and “1” for higher expression. The clinical dataset of 442 HCC patients was downloaded from the cBioPortal database (http://www.cbioportal.org/study?xml:id=lihc_tcga) as the model dataset. The HCC patients without OS information or OS time less than 1 month (for living patients) were excluded from the present study (n = 87). After intersection between the gene dataset and the clinical dataset, the final model dataset contained 348 HCC patients with adequate gene expression information and OS information. The maximum OS time was 120.7 months and the minimum OS time was 0.3 months in the model dataset. The mean ± SD age of HCC patients in the model dataset was 59.5 ± 13.4 years. One hundred and thirty (37.4%) patients out of 348 HCC patients died within the follow-up period (mean ± SD, 843 ± 714 days).

2.5 | Validation group dataset

The validation datasets were downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/). We collected potential datasets in the GEO database by the following criteria: (a) gene expression values were available; (b) OS time and OS status were available; and (c) patient number more than 80. Finally, we identified GSE14520 and GSE76427 as external validation datasets. There were 203 HCC patients in GSE14520 (https://www.ncbi.
2.6 | Construction and verification of the prognostic nomogram

The "rms" package was used to generate a prognostic nomogram for OS. Calibration plots were carried out to evaluate the predictive performance of the prognostic nomogram. Time-dependent ROC curves were used to verify the predictive performance of the prognostic nomogram. The score of the prognostic nomogram was calculated by 2 steps: (a) multiply each gene expression value by the Cox regression coefficient to obtain the corresponding score for 1 gene; and (b) calculate the sum of the scores of all genes.

2.7 | Decision curve analysis

The DCA has been recommended to evaluate and compare the performance of predictive models or prognostic models.20-22 The DCA analyses were carried out to assess the predictive performance of the prognostic nomogram.

| TABLE 1 | Clinical features of hepatocellular carcinoma patients in the model group and validation group |
|---------------------------------|---------------------------------|---------------------------------|
| | Model group (n = 348) | Validation group (n = 301) | P value |
| Death, n (%) | 130 (37.4) | 104 (34.6) | .458 |
| Total survival time, mean (min, max), mo | 20.5 (0.3, 120.7) | 36.4 (0.8, 93.1) | <.001 |
| Survival time for dead patients, mean (min, max), mo | 13.7 (0.3, 107.3) | 15.8 (0.8, 75.5) | <.001 |
| Survival time for living patients, mean (min, max), mo | 22.9 (3.1, 120.7) | 52.9 (1.2, 93.1) | <.001 |
| Age, mean ± SD, y | 59.5 ± 13.4 | 55.0 ± 12.3 | <.001 |
| Male sex, n (%) | 236 (67.8) | 189 (62.8) | .179 |
| AJCC Stage, IV/III/II/I/NA | 185/79/80/4/0 | 127/103/68/3/0 | .056 |
| AJCC PT, T4/T3/T2/T1/NA | 14/74/87/173/0 | NA | NA |
| AJCC PN, N2/N1/N0/NA | 100/3/245/0 | NA | NA |
| AJCC PM, M2/M1/M0/NA | 96/4/248/0 | NA | NA |
| Grade, IV/III/II/I/NA | 12/115/168/53/0 | NA | NA |
| Radiation, Yes/no/NA | 4/230/114 | NA | NA |
| Pharmaceutical adjuvant, Yes/No/NA | 15/215/118 | NA | NA |
| Ablation embolization, Yes/no/NA | 13/219/116 | NA | NA |
| Grade, IV/III/II/I/NA | 12/115/163/53/5 | NA | NA |
| Vascular invasion, Macro/micro/no/NA | 16/87/191/54 | NA | NA |
| Surgical margin resection status, RX/R2/R1/R0/NA | 19/1/15/306/7 | NA | NA |
| Child-Pugh, 3/2/1/NA | 1/20/210/117 | NA | NA |
| Ishak fibrosis, 6/5/3-4/1-2/0/NA | 67/8/26/31/71/145 | NA | NA |
| ECOG score, 4/3/2/1/0/NA | 2/12/25/81/159/69 | NA | NA |

Continuous variables were compared by t test or Mann-Whitney U test as appropriate. Categorical variables were compared by χ² test or Fisher’s exact test as appropriate. NA, missing data; RX, uncertain surgical margin resection status.
Gene ontology functional enrichment analyses were carried out to explore the biological significance of the prognostic mRNAs in the ceRNA network by using the "clusterProfiler" package. $P < .05$ was defined as the threshold of statistical significance for Fisher’s test.

### Table 2: Top 10 upregulated and downregulated differentially expressed RNAs in hepatocellular carcinoma samples

| Type   | Symbol      | logFC  | P value | Symbol      | logFC  | P value |
|--------|-------------|--------|---------|-------------|--------|---------|
| mRNA   | EBF2        | 5.112  | 1.05E-42| CRHBP       | -4.435 | 5.77E-48|
| mRNA   | C20orf204   | 5.128  | 1.06E-18| PZP         | -4.282 | 1.16E-35|
| mRNA   | MUC13       | 5.223  | 1.05E-13| FCN3        | -4.059 | 1.35E-44|
| mRNA   | ALDH3A1     | 5.277  | 5.36E-10| HAMP        | -3.568 | 1.42E-19|
| mRNA   | TGM3        | 5.397  | 4.13E-21| VIPR1       | -3.444 | 3.85E-48|
| mRNA   | NXPH4       | 5.709  | 3.28E-20| CFP         | -3.36  | 2.88E-58|
| mRNA   | THBS4       | 5.78   | 7.93E-27| OIT3        | -3.233 | 7.22E-69|
| mRNA   | S100P       | 5.966  | 1.74E-13| PTH1R       | -3.215 | 2.05E-49|
| mRNA   | GPC3        | 5.981  | 2.74E-25| FOSB        | -3.206 | 7.88E-30|
| mRNA   | AFP         | 7.07   | 4.90E-13| DBH         | -3.178 | 8.00E-44|
| miRNA  | hsa-miR-452-3p| 1.421  | 5.45E-11| hsa-miR-675-3p| -2.716 | 2.16E-14|
| miRNA  | hsa-miR-21-5p| 1.524  | 1.61E-34| hsa-miR-144-3p| -2.663 | 1.57E-30|
| miRNA  | hsa-miR-452-5p| 1.859  | 9.50E-18| hsa-miR-424-5p| -2.637 | 8.27E-51|
| miRNA  | hsa-miR-96-5p| 1.966  | 4.86E-14| hsa-miR-139-3p| -2.442 | 1.13E-39|
| miRNA  | hsa-miR-182-5p| 1.977  | 5.04E-09| hsa-miR-199b-3p| -2.324 | 1.71E-17|
| miRNA  | hsa-miR-183-5p| 2.338  | 1.96E-10| hsa-miR-199a-3p| -2.32  | 1.88E-17|
| miRNA  | hsa-miR-224-5p| 2.598  | 2.47E-18| hsa-miR-3607-3p| -2.161 | 1.66E-24|
| miRNA  | hsa-miR-10b-5p| 2.906  | 8.15E-23| hsa-miR-5589-5p| -2.133 | 2.01E-19|
| miRNA  | hsa-miR-1269a| 3.266  | 4.22E-08| hsa-miR-139-5p| -2.093 | 8.71E-28|
| lncRNA | PVT1        | 2.197  | 1.08E-15| LINC02428   | -2.138 | 7.08E-14|
| lncRNA | AC111000.4  | 2.361  | 4.24E-14| AL161668.4  | -2.121 | 7.78E-21|
| lncRNA | AC092171.2  | 2.492  | 6.54E-32| FAM99A      | -2.013 | 8.62E-09|
| lncRNA | AC010719.1  | 2.628  | 3.46E-27| AC020978.4  | -1.969 | 3.04E-18|
| lncRNA | LINC00665   | 2.681  | 1.18E-12| AC008549.1  | -1.908 | 4.06E-11|
| lncRNA | MAFG-DT     | 2.722  | 5.59E-21| LINC01767   | -1.826 | 1.19E-18|
| lncRNA | LINC01451   | 3.089  | 1.19E-13| AC112206.2  | -1.795 | 1.45E-08|
| lncRNA | CRNDE       | 3.147  | 6.76E-20| AL359715.3  | -1.704 | 4.11E-34|
| lncRNA | AL365181.3  | 3.52   | 1.36E-14| AP003716.1  | -1.641 | 4.70E-09|
| lncRNA | HAGL4       | 6.147  | 2.23E-21| LINC01702   | -1.629 | 7.04E-08|

**Type**, long noncoding RNA; **logFC**, log fold change; **miRNA**, microRNA.
prognostic indicators for OS. To compare the difference of survival curves between high-risk and low-risk groups, Kaplan-Meier survival analyses were carried out. Time-dependent ROC curves were used to assess the predictive performance of prognostic models. The statistical analyses were carried out through SPSS Statistics 19.0 (SPSS Inc.) and R software (version 3.5.2) with the following packages: “GOplot”, “timeROC “, “rms”, “pROC”, “survival”, “clusterProfiler” and “glmnet”. P value <.05 was the threshold value for statistical significance.

3 | RESULTS

3.1 | Study cohort

There were 348 HCC patients in the model group and 301 HCC patients in the validation group. In the model group, 130 patients (37.4%) died during the follow-up period; 104 patients (34.6%) died during the follow-up period in the verification group. The basic clinical features of HCC patients in the model and verification groups are shown in Table 1.

3.2 | Differentially expressed analyses

According to the given threshold, 5035 mRNAs (3329 upregulated and 1706 downregulated), 455 lncRNAs (316 upregulated and 139 downregulated), and 181 miRNAs (58 upregulated and 123 downregulated) were identified between HCC tissues and normal liver tissues. The top 10 upregulated and downregulated mRNAs, miRNAs, and lncRNAs are summarized in Table 2. The volcano plots of mRNAs (Figure 1A), miRNAs (Figure 1B), and lncRNAs (Figure 1C) were presented in Figure 1.

3.3 | Screening of prognostic mRNAs for OS

The univariate Cox analyses were used to screen the potential candidates for OS through TCGA clinical dataset based on previous differentially expressed mRNAs. There were 1918 differentially expressed mRNAs identified as potential prognostic mRNAs by using univariate Cox analyses.

3.4 | Construction of ceRNA network

The differentially expressed lncRNAs (n = 455), miRNAs (n = 181), and prognostic mRNAs (n = 1918) were used to construct a ceRNA network. To obtain the biological targets of lncRNAs, the miRcode database (http://www.mircode.org/) was searched according to the differentially expressed lncRNAs (n = 455). To obtain the biological targets of miRNAs, TargetScan (http://www.targetscan.org/), miRDB (http://www.mirdb.org/), and miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/php/index.php) were searched according to the previous lncRNA-targeted miRNAs (n = 208). The miRNA-predicted mRNAs that could be searched in the 3 databases were defined as the miRNA-targeted mRNAs (n = 1269). Then we undertook an interaction between miRNA-targeted mRNAs and prognostic mRNAs to gain the prognosis-associated miRNA-targeted mRNAs (n = 105). Finally, 43 lncRNAs, 37 miRNAs, and 105 mRNAs were used to build the ceRNA network for OS of HCC patients. The regulatory network and potential pathways among lncRNA-miRNA-mRNA are summarized in Table S1.
Cytoscape software was used to demonstrate the ceRNA network (Figure 2).

There were 536 connections between 43 lncRNAs and 37 miRNAs, whereas there were 169 connections between 37 miRNAs and 105 mRNAs.

### 3.5 Functional enrichment analyses

Gene Ontology and KEGG analyses were carried out to explore the functions associated with the prognostic mRNAs in the ceRNA network. Figure 3 shows the barplot and the dotplot of GO terms for OS.

![Figure 2: Competitive endogenous RNA network chart to identify prognostic biomarkers for overall survival in hepatocellular carcinoma. Blue triangles represent 37 microRNAs, green circles represent 105 mRNAs, and red triangles represent 43 long noncoding RNAs.](image)

![Figure 3: Barplot (A) and dotplot (B) of Gene Ontology terms for prognostic mRNAs for overall survival in hepatocellular carcinoma.](image)
The GO terms (Figure 4) and KEGG pathways (Figure 5) are displayed by the “GOplot” package. The top 5 GO terms of prognostic mRNAs in the ceRNA network were mainly enriched in DNA replication, extrinsic apoptotic signaling pathway, signal transduction in absence of ligand, extrinsic apoptotic signaling pathway in absence of ligand, and regulation of apoptotic signaling pathway. The KEGG pathways of these prognostic mRNAs were mainly enriched in terms of cell cycle, focal adhesion, regulation of actin cytoskeleton, and cellular senescence.

**FIGURE 4** Chord chart of Gene Ontology (GO) terms for prognostic mRNAs for overall survival in hepatocellular carcinoma. logFC, log fold change

**FIGURE 5** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways associated with prognostic mRNAs for overall survival in hepatocellular carcinoma. GO, Gene Ontology; logFC, log fold change
### 3.6 Development of a prognostic signature

Based on 105 prognostic mRNAs involved in the ceRNA network, multivariate Cox regression was carried out to build a prognostic model for OS. Eight prognostic mRNAs were determined as independent prognostic indicators for OS. Information for 8 prognostic mRNAs is shown in Table 3. The risk score of the prognostic model could be calculated by using the following formula:

\[
\text{risk score} = (H2AFX^*0.5202) + (SQSTM1^*0.5561) - (ITM2A^*0.7905) + (PFKP^*0.5082) + (TPD52L1^*0.4272) + (ACSL4^*0.3985) + (STRN3^*0.4861) - (CPEB3^*0.4412)
\]

The further prognostic nomogram chart is presented in Figure 6.

### 3.7 Distribution characteristics of prognostic model score

To display the distribution characteristics of prognostic model score, a scatter plot (Figure 7A) and the interaction distribution scatter plot (Figure 7B) are presented.

#### Table 3 Model information of prognostic mRNAs in univariate and multivariable Cox regression analyses

| Gene       | Univariate analysis | Multivariate analysis |
|------------|---------------------|-----------------------|
|            | HR                  | 95% CI                | P value   | Coefficient | HR                  | 95% CI                | P value   |
| H2AFX (high/low) | 1.996               | 1.400–2.846           | .001      | 0.520       | 1.682               | 1.160–2.440           | .006      |
| SQSTM1 (high/low) | 1.802               | 1.268–2.560           | .001      | 0.556       | 1.744               | 1.200–2.533           | .004      |
| ITM2A (high/low) | 0.601               | 0.424–0.851           | .004      | −0.791      | 0.454               | 0.312–0.660           | <.001     |
| PFKP (high/low) | 1.828               | 1.279–2.613           | <.001     | 0.508       | 1.662               | 1.143–2.417           | .008      |
| TPD52L1 (high/low) | 1.680               | 1.184–2.384           | .004      | 0.427       | 1.533               | 1.068–2.201           | .021      |
| ACSL4 (high/low) | 1.519               | 1.070–2.159           | .020      | 0.399       | 1.490               | 1.030–2.154           | .034      |
| STRN3 (high/low) | 1.524               | 1.076–2.157           | .018      | 0.486       | 1.63                | 1.117–2.368           | .011      |
| CPEB3 (high/low) | 0.633               | 0.446–0.898           | .010      | −0.441      | 0.643               | 0.439–0.942           | .023      |

Medians of long noncoding RNA (lncRNA) expression values were used as cut-off values to stratify lncRNA expression values into the high expression group (as value 1) and low expression group (as value 0). CI, confidence interval; HR, hazard ratio.

#### Figure 6 Prognostic nomogram for overall survival in patients with hepatocellular carcinoma
3.8 | Associations between prognostic mRNAs and OS

Survival curves were drawn to explore the associations between prognostic mRNAs and OS with the Kaplan-Meier method (Figure 8). The differences of survival curves between the high-risk group and low-risk group were evaluated by using the log-rank test. The results indicated that all 8 prognostic mRNAs were significantly related with OS (P < .05).

3.9 | Predictive performance in model dataset

According to the median value of the prognostic model score, HCC patients in the model cohort (n = 348) were stratified into the high-risk group (n = 174) and low-risk group (n = 174). The OS rate (Figure 9A) in the high-risk group was significantly poorer than that in the low-risk group (P < .001). The C-indexes of the prognostic signature for OS in the model group were 0.776 (95% CI, 0.730-0.822), 0.745 (95% CI, 0.699-0.791) and 0.789 (95% CI, 0.743-0.835) for 1-year, 3-year, and 5-year OS, respectively (Figure 9B). Calibration curves for 1-year OS (Figure 10A), 3-year OS (Figure 10B), and 5-year OS (Figure 10C) indicated good agreement between predictive mortality probability and actual mortality percentage.

3.10 | External validation of prognostic model

We externally validated the predictive performance of the prognostic model using the GSE14520 and GSE76427 datasets. Kaplan-Meier analysis (Figure 11A) indicated that there was significant difference for OS between low-risk and high-risk groups in the validation dataset (P < .001). The C-index of the prognostic model for OS in the validation dataset were 0.944 (95% CI, 0.898-0.990), 0.839 (95% CI, 0.793-0.885) and 0.823 (95% CI, 0.777-0.869) for 1-year OS, 3-year OS, and 5-year OS, respectively (Figure 11B). There was good agreement between predictive mortality probability and actual mortality percentage in calibration curves of 1-year OS (Figure 12A), 3-year OS (Figure 12B), and 5-year OS (Figure 12C).

3.11 | Independence assessment for prognostic model

We carried out multivariate Cox regression analyses to assess the independence of prognostic model for OS in HCC patients. Table 4 indicated that the prognostic model and AJCC stage were independent risk factors for OS in the model dataset by multivariate Cox regression analyses. In the validation dataset, multivariate Cox regression analyses indicated that the prognostic model and AJCC stage were independent risk factors for OS.
FIGURE 8  Survival curves of patients with hepatocellular carcinoma, stratified into high-risk and low-risk groups, carrying identified prognostic genes. HR, hazard ratio.
3.12 | Decision curve analyses

As shown in Figure 13, the prognostic model (red line) had the higher net benefit than pathological stage (green line). The DCA indicated that the prognostic model could gain more benefit than either the dead-all-patients scheme or the dead-none-patients scheme for prediction of 1-year OS (Figure 13A). The clinical impact curve (Figure 13B) depicted the prediction of risk stratification of 1000 patients by using the resample bootstrap method.

3.13 | Subgroup analyses

Pathological stage was an important influencing factor for prognosis of HCC patients. As shown in Figure 14, the OS rate in the high-risk group was significantly poorer than in the low-risk group for all subgroups, indicating that the prognostic model had a stable and reliable predictive performance in different pathological stage subgroups.

3.14 | Correlation analysis

We undertook a correlation analysis to explore the potential association of the prognostic signature and 8 prognostic genes with HCC etiological variables. The results indicated that the prognostic signature was significantly correlated with pathological stage, AJCC PT, grade, vital status, vascular invasion, Ishak score, α-fetoprotein, and Child-Pugh score. The results are shown in Figure 15 for correlation significance and Figure 16 for correlation coefficient.

3.15 | Exploration of survival curves in various subgroups

To further explore survival curves of previously identified prognostic genes in different gender and pathological stage subgroups, we developed a new online program, the Gene Survival Analysis Screening System. Gene Survival Analysis Screening System is available at the following URL: https://zhangzhiqiao3.shinyapps.io/Gene_Survival_Analysis_Screening_System/.

4 | DISCUSSION

At present, the molecular biological mechanisms and regulatory pathways affecting the progression and prognosis of HCC remain unclear. It is important and helpful to optimize the individualized treatment by
exploring the molecular biological regulatory network and the prognostic biomarkers for prognosis of HCC. The ceRNA regulatory network was presented based on 43 lncRNAs, 37 miRNAs, and 105 mRNAs.

According to the 105 prognostic mRNAs in the ceRNA network, we constructed and validated an 8-mRNA prognostic nomogram for OS. The 8-mRNA prognostic nomogram was helpful to identify the HCC patients with high mortality. The OS rate in the high-risk group was significantly poorer than that in the low-risk group in both the model and validation groups. We further developed a web calculator based on this 8-mRNA prognostic nomogram, providing a full-time
individualized risk predictive tool for HCC patients by using the following URL: https://zhangzhiquiao.shinyapps.io/Smart_cancer_predictive_system_HCC/.

Several studies have explored the prognostic biomarkers and possible molecular regulatory pathways of HCC based on competitive endogenous regulatory networks, but did not construct corresponding prognostic models for OS of HCC patients. One mRNA-based prognostic model was built for OS of HCC patients through a single dataset from TCGA, but did not carry out external validation. Additionally, the calculation formula of this previous prognostic model was too complex for clinical utility and the results were difficult to interpret. Therefore the present study screened the prognostic mRNA biomarkers and developed a prognostic model through the model dataset from TCGA, whereas the external validation was undertaken by using an independent external validation dataset from the GEO database.

The ceRNA regulation network is a method for describing post-transcriptional gene regulation, which is helpful for us to depict and understand the complex molecular biological regulation mechanisms. A lot of potential biomarkers in the current study were found to be closely related to OS in HCC. Through the ceRNA regulatory network, the potential IncRNA-miRNA-mRNA regulatory pathways were depicted in Figure 2 and Table S1, which might be helpful to elucidate the potential biological mechanisms and regulatory pathways for OS of HCC. The GO terms and KEGG pathways were further explored to determine the potential molecular mechanisms for OS of HCC.

It has been reported that SQSTM1 played an important role in mediation of selective phagophore recruitment, degradation, fusion with lysosome, and autophagosome formation. Silencing of SQSTM1 could inhibit degradation of Q6-induced hypoxia-inducible factor 1A. Decreased autophagy activity was related with tumorigenesis in different tumors. SQSTM1 could interact with cyclinD1 and regulate the process of cell cycle. PFKP might promote cell proliferation and inhibit apoptosis in renal carcinoma. TARDBP could upregulate PFKP expression through downregulating miR-520 expression. The TARDBP-miR-520-PFKP axis is important for proliferation and prognosis of HCC cells. ACSL4 was prominently correlated with grade and stage. Sung et al. reported that ACSL4 was overexpressed in HCC and increased ACSL4 could promote the proliferation of HCC cells. Triacsin C, an ACSL4 inhibitor, could suppress the proliferation of Hep3B cells through inducing apoptosis. Further survival analysis indicated that ACSL4 expression was an independent prognostic risk factor for OS in HCC.
There was a significant relationship between $CPEB3$ expression and poor prognosis in HCC patients. Increased miR-107 could promote the progression of HCC through inhibiting $CPEB3$ expression, which was identified as a tumor suppressor in the epidermal growth factor receptor molecular regulatory pathway.

The results in the present study were in good agreement with the results of previous studies. MicroRNA-23b-3p is significantly downexpressed in HCC and was helpful to screen high-risk HCC patients. MicroRNA-27a-3p can regulate tumor cell invasion and migration through dual
TABLE 4 Univariate and multivariable Cox regression analyses for independence assessment of prognostic model

|                      | Univariate analysis | Multivariate analysis |
|----------------------|---------------------|-----------------------|
|                      | HR                  | 95% CI                | P value | Coefficient | HR      | 95% CI                | P value |
| **Model group (n = 348)** |                     |                       |         |             |         |                       |         |
| Age, high/low        | 1.186               | 0.839-1.676           | .334    | 0.111       | 1.117   | 0.787-1.585           | .536    |
| Gender, male/female  | 0.817               | 0.573-1.164           | .264    | -0.063      | 0.939   | 0.656-1.342           | .728    |
| AJCC stage, IV,III/II,I | 2.225              | 1.556-3.182           | <.001   | 0.675       | 1.963   | 1.369-2.816           | <.001   |
| Prognostic model, high/low | 3.302           | 2.279-4.784           | <.001   | 1.127       | 3.085   | 2.117-4.496           | <.001   |
| **Validation group (n = 301)** |                     |                       |         |             |         |                       |         |
| Age, high/low        | 0.973               | 0.661-1.432           | .889    | -0.028      | 0.973   | 0.637-1.486           | .898    |
| Gender, male/female  | 1.383               | 0.896-2.136           | .144    | 0.279       | 1.321   | 0.825-2.117           | .247    |
| AJCC stage, IV,III/II,I | 3.084              | 2.060-4.618           | <.001   | 0.991       | 2.694   | 1.782-4.072           | <.001   |
| Prognostic model, high/low | 4.035           | 2.669-6.099           | <.001   | 1.339       | 3.817   | 2.509-5.806           | <.001   |

The median of prognostic model scores was used as the cut-off value to stratify gastric cancer patients into high-risk and low-risk groups. CI, confidence interval; HR, hazard ratio.

FIGURE 13 Decision curve analysis of the prognostic model of patients with hepatocellular carcinoma. A, Decision curve analysis for 1-y overall survival. Black line, net benefit of treating no patients; gray line, net benefit of treating all patients; red line, prognostic model. The green, red, and blue lines represent the decision curves of pathological stage, prognostic signature, and age, respectively. B, Clinical impact curve. The y-axis represents the net benefit.
specificity protein phosphatase. MicroRNA-17-5p can exert an antitumor immune response by inhibiting dendritic cells. MicroRNA-20b-5p-mimic can inhibit tumorigenicity of HCT-116 cells in Balb/c mice. The miR-217/MAPK1 axis is related with migration and proliferation of HCC cells. These studies indicated that miRNAs play important roles in the development and prognosis of cancer, which are of great clinical significance and are necessary to further elucidate the potential mechanisms and regulatory pathways by experimental studies.

High expression of KTN1-AS1 was significantly related with poor prognosis in HCC patients. By increasing the expression level of transforming growth factor-β1, MEG3 can regulate the migration,
proliferation, and invasion of HCC cells.\textsuperscript{40} The IncRNA MCM3AP-AS1 can upregulate the expression of FOXA1 by competitive combination with miR-194-5p and promotes HCC cell proliferation.\textsuperscript{41} The IncRNA CRNDE can promote the invasive ability of HCC cell through the Wnt/β-catenin pathway.\textsuperscript{42} The high expression of PVT1 can promote the occurrence of HCC through effecting expression level of miR-150 and HIG2.\textsuperscript{43} These studies indicated that IncRNA plays an important role in the occurrence and development of HCC. It is necessary to undertake experimental studies to elucidate the relevant pathogenesis and biological regulatory pathways.

The current research has the following advantages. First, the prognostic nomogram is convenient to provide the individual mortality percentages at different time points. Second, a free web calculator for OS in HCC patients was developed and maintained to facilitate users to assess the individual mortality percentages at different time points. Finally, as a noninvasive predictive tool, the prognostic nomogram provides an alternative option for HCC patients who cannot tolerate or are unwilling to undergo surgery.

The current research also has the following shortcomings. First, the current research was carried out using gene expression data generated on gene chip detection platforms, and lacks basic experimental research by using quantitative RT-PCR method and other detection methods. Second, the study data were generated on different detection platforms, affecting the repeatability of research results in different populations. Third, the predictive performance of the prognostic nomogram in the validation group was superior to that in the model group. The median survival times were 13.7 and 15.8 months ($P < .001$) for dead patients in the model and validation groups, respectively, whereas the median survival times were 22.9 and 52.9 months ($P < .001$) for living patients in the model and validation groups, respectively. The long follow-up period in the validation group might partly explain the good predictive performance of the prognostic model compared with the model group. Finally, the model and validation datasets did not contain detailed study information such as drug regimen and other postoperative treatments, which might influence the therapeutic effect and clinical prognosis.

The current study revealed potential molecular biological regulation pathways and prognostic biomarkers by using a ceRNA regulatory network. A prognostic model based on mRNAs in the ceRNA network might be of great assistance in improving the prediction of the individual mortality risk for OS in HCC patients.
ACKNOWLEDGMENTS

This study was funded by Guangdong Provincial Health Department and Guangdong Provincial Financial Department (A2016450 and B2018237). We thank TCGA and the GEO and cBioPortal databases. The idea of the web calculator in this article was inspired by QCancer tools designed by Mr Gary S Collins and his group, and we would like to express our sincere thanks to them. We sincerely thank Qingmei Liu, a computer specialist, for her support on program coding and software development.

DISCLOSURE

None.

AVAILABILITY OF DATA AND MATERIALS

The datasets analyzed in the current study are provided at the following URL: https://zhangzhiquiao3.shinyapps.io/Gene_Survival_Analysis_Screening_System/.

ORCID

Zhijiao Zhang https://orcid.org/0000-0003-4631-8818

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Kudo M. Systemic therapy for hepatocellular carcinoma: 2017 update. Oncology. 2017;93(Suppl 1):135-146.
3. Bosetti C, Taurati F, La Vecchia C. Hepatocellular carcinoma epidemiology. Best Pract Res Clin Gastroenterol. 2014;28(5):753-770.
4. Kulik L, El-Serag HB. Epidemiology and management of hepatocellular carcinoma. Gastroenterology. 2019;156(2):477-491.e471.
5. Yu H, Liu F, Sansas B, et al. Typing of killer-cell immunoglobulin-like receptors and their cognate human leukocyte antigen class I ligands predicts survival of Chinese Han patients with metastatic non-small-cell lung cancer. Mol Clin Oncol. 2017;6(2):279-285.
6. Zeng J, Cai X, Hao X, et al. LncRNA FUNDCC2P4 down-regulation promotes epithelial-mesenchymal transition by reducing E-cadherin expression in residual hepatocellular carcinoma patients with insufficient radiofrequency ablation. Int J Hyperthermia. 2018;34(6):802-811.
7. Zhong X, Long Z, Wu S, Xiao M, Hu W. LncRNA-SNHG7 regulates proliferation, apoptosis and invasion of bladder cancer cells assurance guideline ines. J buon. 2018;23(3):776-781.
8. Shi X, Zhao Y, He R, et al. Three-lncRNA signature is a potential prognostic biomarker for pancreatic adenocarcinoma. Oncotarget. 2018;9(36):24248-24259.
9. Huang Y, Xiang B, Liu Y, Wang Y, Kan H. LncRNA CDKN2B-AS1 promotes tumor growth and metastasis of human hepatocellular carcinoma by targeting let-7c-5p/NAP1L1 axis. Cancer Lett. 2018;437:56-66.
10. 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):353-358.
11. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. Nat Rev Genet. 2016;17(5):272-283.
12. Yue C, Ren Y, Ge H, et al. Comprehensive analysis of potential prognostic genes for the construction of a competing endogenous RNA regulatory network in hepatocellular carcinoma. Onco Targets Ther. 2019;12:561-576.
13. Lin F, Wen DY, Li Q, He Y, Yang H, Chen G. Genome-wide analysis of prognostic lncRNAs, miRNAs, and mRNAs forming a competing endogenous RNA network in hepatocellular carcinoma. Cell Physiol Biochem. 2018;48(5):1953-1967.
14. Tang Z, Ding Y, Shen Q, et al. KIAA1199 promotes invasion and migration in non-small-cell lung cancer (NSCLC) via PI3K-Akt mediated EMT. J Mol Med (Berl). 2019;97(1):127-140.
15. Bai Y, Long J, Liu Z, et al. Comprehensive analysis of a ceRNA network reveals potential prognostic cytoplasmic lncRNAs involved in HCC progression. J Cell Physiol. 2019;234(10):18837-18848.
16. Song W, Miao DL, Chen L. Nomogram for predicting survival in patients with pancreatic cancer. Onco Targets Ther. 2018;11:539-545.
17. Peng P. A prognostic 3-long noncoding RNA signature for patients with gastric cancer. J Cell Biochem. 2018;119(11):9261-9269.
18. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11(3):R25.
19. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
20. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inform Decis Mak. 2008;8:53.
21. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26(6):565-574.
22. Localio AR, Goodman S. Beyond the usual prediction accuracy metrics: reporting results for clinical decision making. Ann Intern Med. 2012;157(4):294-295.
23. Wu SY, Lan SH, Liu HS. Degradative autophagy selectively regulates CCND1 (cyclin D1) and MIR224, two oncogenic factors involved in hepatocellular carcinoma tumorigenesis. Autophagy. 2019;15(4):729-730.
24. Liu XW, Cai TY, Zhu H, et al. A novel hypoxia-targeted drug, regulates hypoxia-inducible factor signaling via an autophagy-dependent mechanism in hepatocellular carcinoma. Autophagy. 2014;10(1):111-122.
25. Belaid A, Ndiaye PD, Cerezo M, et al. Autophagy and SQSTM1 on the RHOA(d) again: emerging roles of autophagy in the degradation of signaling proteins. Autophagy. 2014;10(2):201-208.
26. Gao C, Cao W, Bao L, et al. Autophagy negatively regulates Wnt signalling by promoting Dishevelled degradation. Nat Cell Biol. 2010;12(8):781-790.
27. Wu SY, Lan SH, Wu SR, et al. Hepatocellular carcinoma-related cyclin D1 is selectively regulated by autophagy degradation system. Hepatology. 2018;68(1):141-154.
28. Kisiel JB, Dukek BA, Kanipakam VSRR, et al. Hepatocellular carcinoma detection by plasma methylated DNA: discovery, phase I pilot, and phase II C clinical validation. Hepatology. 2019;69(3):1180-1192.
29. Park YY, Kim SB, Han HD, et al. Tat-activating regulatory DNA-binding protein regulates glycolysis in hepatocellular carcinoma by regulating the platelet isoform of phosphofructokinase through microRNA 520. Hepatology. 2013;58(1):182-191.
30. Sun XJ, Xu GL. Overexpression of Acyl-CoA Ligase 4 (ACSL4) in patients with hepatocellular carcinoma and its prognostic significance. Med Sci Monit. 2017;23:4343-4350.
31. Sung YK, Park MK, Hong SH, et al. Regulation of cell growth by fatty acid-CoA ligase 4 in human hepatocellular carcinoma cells. Exp Mol Med. 2007;39(4):477-482.
32. Liu H, Wang Y, Chen B, Shen X, Li W. Effects of lidocaine-mediated CPEB3 upregulation in human hepatocellular carcinoma cell proliferation in vitro. Biomed Res Int. 2018;2018:8403157.
33. Zou CD, Zhao WM, Wang XN, et al. MicroRNA-107: a novel promoter of tumor progression that targets the CPEB3/EGFR axis in human hepatocellular carcinoma. Oncotarget. 2016;7(1):266-278.
34. Sun Q, Li J, Jin B, Wang T, Gu J. Evaluation of miR-331-3p and miR-23b-3p as serum biomarkers for hepatitis c virus-related hepatocellular carcinoma at early stage. Clin Res Hepatol Gastroenterol. 2019. https://doi.org/10.1016/j.clinre.2019.03.011
35. Li JM, Zhou J, Xu Z, Huang HJ, Chen MJ, Ji JS. MicroRNA-27a-3p inhibits cell viability and migration through down-regulating DUSP16 in hepatocellular carcinoma. J Cell Biochem. 2018;119(7):5143-5152.
36. Cui ZJ, Xie XL, Qi W, et al. Cell-free miR-17-5p as a diagnostic biomarker for gastric cancer inhibits dendritic cell maturation. Onco Targets Ther. 2019;12:2661-2675.
37. Tang D, Yang Z, Long F, et al. Long noncoding RNA MALAT1 mediates stem cell-like properties in human colorectal cancer cells by regulating miR-20b-5p/Oct4 axis. J Cell Physiol. 2019;234(11):20816-20828.
38. Wang H, Ke J, Guo Q, Barnabo Nampoukime KP, Yang P, Ma K. Long non-coding RNA CRNDE promotes the proliferation, migration and invasion of hepatocellular carcinoma cells through miR-217/MAPK1 axis. J Cell Mol Med. 2018;22(12):5862-5876.
39. Zhang L, Wang L, Wang Y, et al. LncRNA KTN1-AS1 promotes tumor growth of hepatocellular carcinoma by targeting miR-23c/ERBB2IP axis. Biomed Pharmacother. 2019;109:1140-1147.
40. Dong H, Zhang Y, Xu Y, et al. Downregulation of long non-coding RNA MEG3 promotes proliferation, migration, and invasion of human hepatocellular carcinoma cells by upregulating TGF-β1. Acta Biochim Biophys Sin (Shanghai). 2019;51(6):645-652.
41. Wang Y, Yang L, Chen T, et al. A novel lncRNA MCM3AP-AS1 promotes the growth of hepatocellular carcinoma by targeting miR-194-5p/FOX A1 axis. Mol Cancer. 2019;18(1):28.
42. Zhu L, Yang N, Du G, et al. LncRNA CRNDE promotes the epithelial-mesenchymal transition of hepatocellular carcinoma cells via enhancing the Wnt/β-catenin signaling pathway. J Cell Biochem. 2019;120:1156-1164.
43. Xu Y, Luo X, He W, et al. Long non-coding RNA PVT1/miR-150/HIG2 axis regulates the proliferation, invasion and the balance of iron metabolism of hepatocellular carcinoma. Cell Physiol Biochem. 2018;49(4):1403-1419.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Zhang Z, Li J, He T, et al. The competitive endogenous RNA regulatory network reveals potential prognostic biomarkers for overall survival in hepatocellular carcinoma. Cancer Sci. 2019;110:2905-2923. https://doi.org/10.1111/cas.14138