miR-9-5p Expression is Associated With Vascular Invasion and Prognosis in Hepatocellular Carcinoma

Yuan Chen
Yangzhou University Medical college

Xu Hao
Yangzhou University Medical college

Zhang Chi
Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

Shengjie Jin
Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

Dousheng Bai (✉ drbaidousheng@163.com)
Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

https://orcid.org/0000-0001-8032-7959

Research Article

Keywords: Hepatocellular carcinoma, Vascular invasion, Prognosis, TCGA

DOI: https://doi.org/10.21203/rs.3.rs-527753/v1

License: ☺ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Hepatocellular carcinoma (HCC) is a common liver malignancy. Early vascular invasion (VI) has been associated with poor prognosis in HCC patients. MicroRNAs (miRNAs) play a significant role in the emergence and development of many tumor types. In this study, we identified miR-9-5p that could predict VI and prognosis in HCC patients based on TCGA database. Further, we explored the possible mechanism of miR-9-5p by target gene prediction, functional enrichment analysis, and protein-protein interaction analysis. Univariate and multivariate analysis revealed that miR-9-5p was an independent risk factor for HCC. Finally, the nomogram based on miR-9-5p showed a good predictive value of HCC survival. In summary, miR-9-5p is associated with VI in HCC, and higher expression of miR-9-5p indicates poor prognosis in HCC.

1. Introduction

Hepatocellular carcinoma (HCC) is the fifth most common cancer, and the third leading cause of cancer-related deaths, worldwide[1]. The 5-year survival rate of HCC patients in Asia is less than 20%[2]. HCC is characterized by a high degree of malignancy, no obvious early symptoms, and intrahepatic and extrahepatic metastasis[3]. Currently, surgery remains the most efficient treatment in HCC patients, however, a great deal of patients have lost the opportunity of surgical treatment at the time of diagnosis and can only receive palliative care[4]. In the past few years, the global burden of liver cancer has been increasing[1, 5]. Therefore, early diagnosis of HCC is particularly important.

MicroRNA (miRNA) is a non-coding RNA composed of 18–25 nucleotides. Besides, they are highly expressed in different types of body fluids and are remarkably stable. They also target messenger RNAs (mRNAs) and inhibit their translation or induce degradation, thus providing a rapid and sensitive mechanism for regulating gene expression[6]. MiRNAs can also be used as a link to transmit information between cells[7]. Over the past few years, numerous studies have shown that miRNAs play a significant role in the occurrence and development of tumors. For example, up-regulation of miR-342-3p can inhibit HCC progression[8], down-regulation of miR-4521 inhibits gastric cancer progression[9], and miR-136-5p promotes the proliferation and metastasis of renal cell carcinoma[10]. With the recent development of liquid biopsy technology, some miRNAs have been used as biomarkers in different tumor types due to their high sensitivity and diagnostic efficiency[11]. Numerous studies have also demonstrated that various miRNAs are promising diagnostic and prognostic biomarkers for HCC. [12–15].

Vascular invasion (VI) is an important step in tumor invasion and metastasis. Studies have shown that VI is associated with poor prognosis in many tumor types[16–19]. But, there are few studies on the prediction of VI in HCC by miRNAs.

In this study, we investigated the expression and significance of miR-9-5p in predicting VI and prognosis in HCC patients using data from the TCGA database. Besides, we also determined the relationship
between miR-9-5p expression and survival of HCC patients, as well as the possible biological function of miR-9-5p.

2. Materials And Methods

2.1. Subjects and clinical characteristics

The clinical data, read counts information of miRNA-Seq, and RNA-Seq of HCC patients were derived from the TCGA database (https://portal.gdc.cancer.gov/). The information on gene length was obtained from GENCODE[20]. MiRNAs annotation information was obtained from miRbase[21]. To standardize the sequencing data and reduce errors, the read counts were standardized using the following formula[22].

\[
TPM (\text{transcripts per million}) = 10^6 \times \frac{\text{reads mapped to transcript}}{\text{transcript length}} \cdot \frac{\text{Sum(reads mapped to transcript / transcript length)}}{\text{transcript length}}
\]

The TCGA-LIHC data set contained 377 cases, and after sequentially excluding cases without complete clinical information (n = 56), histological type of non-hepatocellular carcinoma (n = 10), with a previous history of liver cancer (n = 21), with overall survival (OS) < 30 days (n = 41) and lacking both mRNAs and miRNAs data (n = 8), the final sample size was 241.

2.2. Screening of differentially expressed miRNAs (DEMs) related to VI in HCC

Patients were split into two groups according to whether or not they had VI. The DESeq2 package was used to analyze for any differences between groups, while the ggplot2 package was used to present the results in a volcano plot in R (Version 4.0.0). The cut-off value of | log₂FC | >1 and \( p \)-value < 0.05 were determined.

2.3. Identification of DEMs

To verify the relationship between DEMs and survival, GraphPad Prism (Version 8.0.2) was used to draw the Kaplan-Meier associated survival curves. MiRNAs related to survival were expressed and selected for further analysis.

2.4. Target gene prediction

We used three network databases to predict the potential target genes from differentially expressed miRNAs: TargetScan (targetscan.org)[23], microT-CDS (www.microrna.gr/microT-CDS)[24], and miRmap (mirmap.ezlab.org)[25]. Genes with a prediction score greater than 80 were considered as target genes. We identified the common prediction results in the three databases and considered them as potential target genes of miRNAs. A Venn diagram was used to visually represent the results(bioinformatics.psb.ugent.be/webtools/Venn/).

2.5. GO and KEGG Pathway Enrichment Analysis
Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were used to determine the possible enrichment functions and important pathways of identified miRNAs. R package was used to analyze and visualize the results. \( P < 0.05 \) was considered statistically significant.

### 2.6. Protein-protein interactions (PPI)

Relational data on PPI was downloaded from the STRING database (string-db.org). The genes with a minimum interaction score of 0.7 (high confidence) were selected and the PPI networks were established in Cytoscape (Version 3.7.2).

### 2.7. Nomogram Model Construction

To evaluate the prognosis in HCC patients, the miRNAs expression data were combined with clinical features (gender, age, grade, stage, and T), and used to construct a novel nomogram for the prediction of 1-, 2-, and 3-year OS in HCC patients. The "rms" package and the "survival" package were used to build and validate the model in R. The C-index and calibration plots were used to evaluate the predictive accuracy of the nomogram.

### 2.8. Statistical Analysis

R software (version 4.0.0) was used to perform all statistical analyses in this study. An unpaired t-test was used to determine the differences in expression levels of miRNAs between groups. The t-test or the chi-square test was used to analyze the clinical characteristics of patients in each group. Univariate/multivariate analysis was performed using the "survival" package and the results were visualized using the "survminer" package. The receiver operator characteristic (ROC) curve was used to assess the diagnostic accuracy of miRNAs in patient survival.

### 3. Results

#### 3.1. Identification of DEMs associated with VI in HCC

A total of 241 cases were obtained from the TCGA-LIHC data set, and the clinicopathological characteristics of the patients included in the analysis are shown in Table 1. The patients were divided into VI (+) group (n = 81) and VI (-) group (n = 160) based on the presence or absence of VI. The survival curves indicated that HCC patients with VI (+) had significantly worse survival compared with VI (-) (\( P = 0.0052 \), Fig. 1A) HCC patients. Further, 7 DEMs, including 5 up-regulated miRNAs (miR-9-5p, miR-9-3p, miR-1270, miR-196a-5p and miR-2114-5p) and 2 down-regulated miRNAs (miR-375-3p and miR-483-3p) were identified (Fig. 1B). Next, Kaplan–Meier survival analyses of the DEMs were performed to determine their prognostic significance. As shown in Fig. 2, the high expression levels of miR-9-5p and miR-9-3p were significantly correlated with poorer outcomes in HCC patients. Besides, we found that the expression level of miR-9-3p was relatively lower in HCC patients compared with the expression level miR-9-5p (Fig. 3). Consequently, miR-9-5p was selected for further studies.
Table 1
Clinicopathological characteristics of the TCGA cohort

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>VI (-)</th>
<th>VI (+)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>241</td>
<td>160</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>age (mean (SD))</td>
<td>59.5 (12.6)</td>
<td>59.6 (12.4)</td>
<td>59.5 (13.2)</td>
<td>0.968</td>
</tr>
<tr>
<td>gender = MALE (%)</td>
<td>165 (68.5)</td>
<td>106 (66.2)</td>
<td>59 (72.8)</td>
<td>0.372</td>
</tr>
<tr>
<td>VI (%)</td>
<td>&gt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macro</td>
<td>15 (6.2 )</td>
<td>0 (0.0 )</td>
<td>15 (18.5)</td>
<td></td>
</tr>
<tr>
<td>Micro</td>
<td>66 (27.4)</td>
<td>0 (0.0 )</td>
<td>66 (81.5)</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>160 (66.4)</td>
<td>160 (100.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>pathology grade (%)</td>
<td>0.206</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>31 (12.9)</td>
<td>26 (16.2)</td>
<td>5 (6.2)</td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td>110 (45.6)</td>
<td>73 (45.6)</td>
<td>37 (45.7)</td>
<td></td>
</tr>
<tr>
<td>G3</td>
<td>87 (36.1)</td>
<td>54 (33.8)</td>
<td>33 (40.7)</td>
<td></td>
</tr>
<tr>
<td>G4</td>
<td>11 (4.6 )</td>
<td>6 (3.8 )</td>
<td>5 (6.2)</td>
<td></td>
</tr>
<tr>
<td>unknow</td>
<td>2 (0.8 )</td>
<td>1 (0.6 )</td>
<td>1 (1.2)</td>
<td></td>
</tr>
<tr>
<td>clinical stage (%)</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage I</td>
<td>133 (55.2)</td>
<td>113 (70.6)</td>
<td>20 (24.7)</td>
<td></td>
</tr>
<tr>
<td>Stage II</td>
<td>50 (20.7)</td>
<td>18 (11.2)</td>
<td>32 (39.5)</td>
<td></td>
</tr>
<tr>
<td>Stage III</td>
<td>3 (1.2 )</td>
<td>2 (1.2 )</td>
<td>1 (1.2)</td>
<td></td>
</tr>
<tr>
<td>Stage II A</td>
<td>29 (12.0)</td>
<td>15 (9.4 )</td>
<td>14 (17.3)</td>
<td></td>
</tr>
<tr>
<td>Stage II B</td>
<td>5 (2.1 )</td>
<td>1 (0.6 )</td>
<td>4 (4.9)</td>
<td></td>
</tr>
<tr>
<td>Stage II C</td>
<td>4 (1.7 )</td>
<td>1 (0.6 )</td>
<td>3 (3.7)</td>
<td></td>
</tr>
<tr>
<td>Stage III</td>
<td>1 (0.4 )</td>
<td>0 (0.0 )</td>
<td>1 (1.2)</td>
<td></td>
</tr>
<tr>
<td>Stage III B</td>
<td>2 (0.8 )</td>
<td>2 (1.2 )</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>unknow</td>
<td>14 (5.8)</td>
<td>8 (5.0 )</td>
<td>6 (7.4)</td>
<td></td>
</tr>
<tr>
<td>T classification(%)</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>139 (57.7)</td>
<td>119 (74.4)</td>
<td>20 (24.7)</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>53 (22.0)</td>
<td>19 (11.9)</td>
<td>34 (42.0)</td>
<td></td>
</tr>
</tbody>
</table>
3.2. The relationship between miR-9-5p and clinical parameters.

The included samples were divided into the low expression group (n = 120) and high expression group (n = 121) according to the median expression level of miR-9-5p. Heatmap was used to visualize the expression patterns of miR-9-5p in HCC (Fig. 4A). Figure 4B shows the distribution of mir-9-5p expression, and the survival time and survival status of HCC patients were presented as a scatterplot (Fig. 4C).

Univariate Cox analysis and multivariate Cox analysis were performed to examine the effect of miR-9-5p on the clinical outcome of HCC patients. Univariate Cox analysis showed that clinical stage (P = 0.019, HR = 1.429, 95%CI = 1.061–1.925), pathological T stage (P = 0.015, HR = 1.44, 95%CI = 1.072–1.934), and miR-9-5p expression (P < 0.001, HR = 1.252, 95%CI = 1.126–1.3910) were closely related to the HCC outcome (Fig. 5A). Besides, multivariate Cox analysis confirmed that miR-9-5p expression was an independent prognostic factor for HCC patients (Fig. 5B). Moreover, the ROC curve showed that miR-9-5p expression had a better predictive ability compared with gender, age, grade, stage, T, and AFP (AUC = 0.693) (Fig. 5C).

3.3. Functional enrichment analysis of the target genes of miR-9-5p
A total of 245 mRNAs were identified from TargetScan, microT-CDS, and miRmap network databases, as the target genes of miR-9-5p (Fig. 6A). To further explore the function of miR-9-5p in HCC, GO and KEGG pathway analysis was performed. KEGG pathway analysis revealed the pathways in which the target genes are implicated. The results showed that the prolactin signaling pathway, transcriptional misregulation in cancer, and hepatocellular carcinoma were the main pathways regulated by the target genes (Fig. 6B). GO analysis revealed the function of the target genes based on the three GO terms. In the biological process (BP), the top three enriched terms were urogenital system development, hormone-mediated signaling pathway, and cellular response to steroid hormone stimulus. At the level of cellular component (CC), the target genes were mainly enriched in the basement membrane, messenger ribonucleoprotein complex, and SWI/SNF complex. Lastly, in terms of molecular function (MF), DNA-binding transcription factor binding, DNA-binding transcription repressor activity, RNA polymerase-specific, and DNA-binding transcription repressor activity were the top 3 most enriched terms (Fig. 6C).

3.4. PPI network analysis

The STRING online tool was used to construct the interaction network between miR-9-5p and the target genes (Fig. 7A). The top hub genes ranked by degree were selected for further analysis. These hub genes, including SIRT1, FOXO1, AR, SMARCD2, BCL6, SMARCE1, NCOR2, VCAN, CNTN4, and CNTN3, might play a critical role in cancer (Fig. 7B).

3.5. Construction of a nomogram model

We established and validate a nomogram model in the TCGA-LIHC cohort to predict the survival of HCC patients. This model predicted the 1-, 2- and 3-years overall survival, and age, grade, stage, T, and miR-9-5p expression variables were included in constructing the nomogram (Fig. 8A). The C-index of the nomogram was 0.697 (95%CI = 0.609–0.785). The calibration plot shows good consistency between the nomogram prediction probabilities and observed probability in predicting the 1-, 2- and 3-year survival rates in the TCGA-LIHC cohort (Fig. 8B-D).

4. Discussion

In 2020, there were over 906,000 new cases and 830,000 deaths of primary liver cancer, and HCC accounts for 75–85% of all primary liver cancers[1]. The global burden of HCC is high, and the prognosis of HCC is poor, which is related to early invasion and metastasis[26]. MiRNAs are members of the RNA family and play a key role in proliferation, invasion, metastasis, and immune escape of HCC[27–30]. Therefore, it is necessary to establish a prognostic model that includes the expression of miRNAs molecules related to HCC prognosis.

In previous studies, miR-9-5p has been reported to play an important role in a variety of diseases. In cervical cancer, miR-9-5p promotes tumor angiogenesis by targeting SOCS[31]. In prostate cancer, mir-9-5p regulates the expression of QKI-5, StarD, and other genes, thus affecting tumor invasion and
metastasis[32, 33]. However, Wang et al revealed that miR-9-5p inhibits tumor cell proliferation and promotes apoptosis by down-regulating PAK4 in colorectal cancer[34]. In HCC, several experimental studies have shown that miR-9-5p can promote the proliferation, invasion, and metastasis of tumor cells by down-regulating the expression of ESR1[35], Klf4[36], CNNM1[37], PPARA[38], and so on. In a previous study, high expression of miR-9-5p was associated with a poor prognosis of osteosarcoma[39]. However, patients with high expression of miR-9-5p are reported to have a better prognosis in ovarian cancer[40].

The present study focused on identifying miRNAs related to VI in HCC. Unlike other database-based studies[41–43], all the study subjects were HCC patients, to more accurately screen for differentially expressed miRNAs related to VI. Patients who met the inclusion criteria were divided into two groups based on whether or not they had VI, and the differences in miRNA expression levels analyzed. Seven differentially expressed miRNAs were identified, however, only miR-9-5p was highly expressed in HCC patients and correlated with OS. The potential targets of miR-9-5p were identified from three target gene prediction databases. GO and KEGG enrichment was used to predict the potential mechanism of miR-9-5p in HCC. PPI analysis was used to select the core target genes, and miR-9-5p was used in constructing a nomogram that had a significant prognostic value for predicting survival in HCC patients.

However, this study had some limitations. On the one hand, validation of the nomogram was not performed, due to lack of VI, survival, and miRNAs expression data using data from other databases. On the other hand, although many experimental studies have shown that miR-9-5p is related to the proliferation, invasion, and metastasis of HCC, the relationship between miR-9-5p expression and VI of HCC needs further experimental studies.

In conclusion, miR-9-5p is associated with VI in HCC and can be used as a potential diagnostic and prognostic signature for predicting the overall survival of HCC patients.

**Abbreviations**

HCC: hepatocellular carcinoma

VI: vascular invasion

miRNAs: microRNAs

mRNAs: messenger RNAs

TCGA: The Cancer Genome Atlas

DEMs: differentially expressed miRNAs

GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes
PPI: Protein-protein interaction
OS: overall survival
ROC: receiver operator characteristic
TPM: transcripts per million
AUC: area under curve

Declarations

Author Contribution

Yuan Chen conceived and designed the experiment, collected experimental data, wrote the manuscript draft, and finally approved the manuscript. Hao Xu collected and analyzed the experimental data and reviewed the manuscript. Chi Zhang and Shengjie Jin reviewed the manuscript. Dousheng Bai reviewed the manuscript and provided financial support.

Acknowledgments

This work was supported by The National Natural Science Foundation of China under Grant number 81871909; "13th five-year Plan" Science and Education strong Health Project leading personnel of Yangzhou under Grant number YZCXTD201801; and Provincial-level discipline leader of the NJPH under Grant number DTRC201809.

Declarations

Conflicts of Interest

All authors declare no conflicts of interest.

Consent for publication

Not applicable.

Ethics approval

Not applicable.
Consent for publication

All the authors consented to the publication of this research.

References


**Figures**

**A**

TCGA-VI(+) VS VI(-)

---

**B**

VI(-) VS VI(+)

Identification of differentially expressed miRNAs. A. Kaplan–Meier survival analysis VI (+) group and VI (-) group. B. Volcano plot showing miRNAs expression.
Figure 2

Kaplan–Meier survival analysis of the identified miRNAs. A. miR-9-5p, p= 0.0164. B. miR-9-3p, p= 0.0242. C. miR-196a-5p, p= 0.1845. D. miR-375-3p, p= 0.793. E. miR-483-3p, p= 0.9513. F. miR-1270, p= 0.8331. G. miR-2144-5p, p= 0.1119.
Figure 3

The expression of the selected miRNAs. A. Expression of miR-9-5p in VI (+) group, VI (-) group and normal group. B. Expression of miR-9-3p in VI (+) group, VI (-) group and normal group. *: p < 0.05, **: p < 0.01, ***: p < 0.001, ****: p < 0.0001
Figure 4
Risk score analysis of miR-9-5p in the TCGA cohort. A. The expression heatmap of miR-9-5p in HCC. B. The miR-9-5p distribution in the TCGA-LIHC cohort. C. The survival time and survival status with increasing miR-9-5p expression.
Figure 5

Figure 6

Figure 7

Protein-protein interaction analysis. A. PPI network of the target genes. B. Identified hub target genes of the PPI network.
Figure 8

A nomogram to predict the survival rate of patients with HCC. A. A nomogram containing age, grade, stage, T, and miR-9-5p expression. B-D. Calibration curves of the nomogram for predicting the 1-, 2-, and 3-year survival.