Prospective evidence of a circulating microRNA signature as a non-invasive marker of hepatocellular carcinoma in HBV patients

Chun Wang[1,2,*], Hie-Won Hann[3,*], Zhong Ye[1,*], Richard S. Hann[3], Shaogui Wan[1,4], Xishan Ye[1], Peter D. Block[1], Bingshan Li[5], Ronald E. Myers[1], Xiaowei Wang[6], Hee-Soon Juon[1], Jesse Civan[3], Mimi Chang[7], Ho S. Bae[7], Jinliang Xing[8], Hushan Yang[1]

[1]Division of Population Science, Department of Medical Oncology, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA 19107, USA

[2]Department of Environmental Health, School of Public Health, Nantong University, Nantong, Jiangsu 226000, China

[3]Liver Disease Prevention Center, Division of Gastroenterology and Hepatology, Department of Medicine, Thomas Jefferson University, Philadelphia, PA 19107, USA

[4]Institute of Pharmacy, Pharmaceutical College, Henan University, Kaifeng, Henan 475004, China

[5]Center for Human Genetics Research, Department of Molecular Physiology & Biophysics, Vanderbilt University, Nashville, TN 37232, USA

[6]Department of Radiation Oncology, Washington University School of Medicine, Saint Louis, MO 63108, USA

[7]Asian Pacific Liver Center, Saint Vincent Medical Center, Los Angeles, CA 90057, USA

[8]Experimental Teaching Center, School of Basic Medicine, Fourth Military Medical University, Xi'an 710032, China

[*]These authors have contributed equally to this work

Correspondence to: Hushan Yang, *e-mail:* hushan.yang@jefferson.edu

ABSTRACT

The predictive value of circulating microRNAs (miRNAs) in hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC) has been demonstrated in retrospective studies, but it has rarely been tested in prospective studies. In a cohort of 373 cancer-free HBV patients with a median follow-up of 4.5 years, we measured the expression of 24 retrospectively identified HCC-related miRNAs in baseline serum samples. When we analyzed the prospective associations of miRNA expression with HCC risk using the Cox proportional hazards model, we found that 15 of the 24 miRNAs exhibited a significant association with HCC risk. In particular, 7 miRNAs (miR-122, miR-99a, miR-331, miR-125b, miR-23b, miR-92a, and miR-26a) were associated with an increased risk, and 8 miRNAs (miR-652, miR-23a, miR-27a, miR-34a, miR-145, miR-10a, miR-150, and let-7f) were associated with a decreased risk. Compared to HBV patients with a low miRNA-based risk score, those with a high miRNA-based risk score exhibited a significantly elevated HCC risk in both univariate (hazard ratio [HR] 6.56, 95% confidence interval [CI] 2.74-15.70) and multivariate (HR 3.57, 95% CI 1.34-9.48) analyses. The risk score significantly increased the HCC prediction performance of alpha-fetoprotein (concordance index increased from 0.68 to 0.82, *P* < 0.0001). *In silico* analyses indicated that the genes targeted by the 15 miRNAs are mainly enriched in the transforming growth factor-beta signaling pathway. Collectively, these results provide prospective evidence that circulating miRNAs serve as non-invasive markers for risk prediction of HCC in HBV patients.

INTRODUCTION

Hepatocellular carcinoma (HCC), the major histological subtype of primary liver cancers, is the fifth most prevalent malignancy and the third leading cause of cancer death [1, 2]. Chronic hepatitis B virus (HBV) infection is the most prevalent risk factor for HCC, accounting for more than 50% of HCC worldwide [3, 4].

HBV-related HCC has grown significantly in the United States during the past several decades, largely due to the influx of immigrants from HBV-endemic regions [5]. A recent report indicated that among the more than 40 million Americans born outside of the United States, 1.5 million had chronic HBV infection [5]. More than 60% of these HBV patients are relatively young and their risks of developing HCC increase significantly as they age [6]. The overall 5-year survival rate of HCC is extremely low due to limited therapeutic options [7], but survival is markedly improved in patients diagnosed and treated early [2]. These facts highlight the importance of identifying HBV patients at high-risk of HCC for targeted prevention and management.

MicroRNAs (miRNAs) play important roles in a wide range of biological processes [8–10]. MiRNAs could act as oncogenes or tumor suppressors by targeting and silencing mRNAs involved in carcinogenesis, and have been significantly implicated in the initiation and progression of various human cancers [11, 12]. HCC tumors exhibited different miRNA expression patterns compared to normal healthy liver or non-neoplastic peritumoral tissue, suggesting the potential of using miRNAs as tissue-based HCC biomarkers [13–16]. Recently, circulating miRNAs in serum or plasma have shown promise as risk prediction and early detection markers for HCC because of their stability in circulation, non-invasive procurement, and ease of detection [13, 17, 18]. Li *et al*. [19] identified 13 serum miRNAs that could distinguish HBV-positive HCC cases from healthy controls and cancer-free HBV cases. Zhou *et al*. [20] reported a plasma-based 7-miRNA panel with high diagnostic accuracy for HCC. The performance of this panel was independent of disease status and exhibited promising clinical value in diagnosing early-stage HBV-HCC patients [20]. Nevertheless, the vast majority of published studies measured circulating miRNAs from patients already diagnosed with HCC, especially late-stage HCC. Thus, whether the alterations of these identified miRNAs are the cause or consequence of hepatocarcinogenesis remains undetermined.

Mechanistic studies using animal model or liver tissue have suggested that aberrant expressions of certain miRNAs precede HCC development [21, 22]. Two recent multi-stage studies validated some retrospectively-identified circulating miRNAs using prospectively collected samples [23, 24]; however despite the promising findings, neither study reported an identical miRNA signature, even though both studies were conducted in Chinese HBV-HCC patients. To learn more about miRNA expression in HCC development, we tested the hypothesis that the expression profile of serum miRNAs alone, or combined with alpha-fetoproteins (AFP, a traditional HCC marker commonly used in current clinical settings), can prospectively predict HCC development in HBV patients. Our study took advantage of serum samples that were prospectively collected many years before HCC diagnosis in a unique population of Asian Americans with chronic HBV infection.

RESULTS

Characteristics of the study population

There were 379 cancer-free HBV patients who met the 1-year exclusion criteria of being followed for at least 1 year, during which time HCC did not develop. These exclusion criteria were used to minimize the confounding effects of undiagnosed HCC patients at study entry. After excluding 6 patients with a more than 50% failure rate in the detection of 24 miRNAs expression, 373 patients with a median age of 43.2 years (interquartile range [IQR] 36.6-50.8 years) were included in the primary prospective cohort. During a median follow-up of 4.5 years (IQR 2.4-7.8 years), 40 (10.7%) of the 373 patients developed HCC. The median time to HCC diagnosis was 4.0 years (IQR 2.0-9.6 years). Compared to the HBV patients who remained cancer-free during follow up, those who developed HCC were older ($P < 0.0001$), more likely to be ever-smokers ($P = 0.0301$), cirrhotic ($P < 0.0001$), and had high AFP levels ($P = 0.0005$). The demographics variables of the study subjects appear in Table 1.

Prospective association of the expression levels of baseline serum miRNAs with HCC risk

Among the 24 tested miRNAs, miR-98 and miR-338-3p were excluded from further analyses due to $\geq 50\%$ missing values in all patients. Significance Analysis of Microarrays (SAM) analysis identified 15 differentially expressed miRNAs that were significantly associated with HCC risk, including 7 miRNAs (miR-122, miR-99a, miR-331, miR-125b, miR-23b, miR-92a, and miR-26a) for which an elevated expression was associated with an increased risk (risk miRNAs), and 8 miRNAs (miR-652, miR-23a, miR-27a, miR-34a, miR-145, miR-10a, miR-150, and let-7f) for which an elevated expression was associated with a decreased risk (protective miRNAs) (Table 2).

A miRNA-based risk score

To assess these significant miRNAs as a panel, a risk score based on the linear combination of the expression levels of the 15 miRNAs, weighted by the SAM d-value (Table 2), was calculated for each patient. Figure 1 delineates the distributions of the risk scores (Figure 1A), follow-up time (Figure 1B), and serum miRNA expression levels (Figure 1C) of the 373 HBV patients. Each patient was then assigned to either a high- or a low-risk group according to the median cutoff of the risk scores. Compared to the patients with low-risk scores, those with

Table 1: Characteristics of the study population

Variables	Total subjects (N=373, 100%)	HBV patients who did not develop HCC (N=333, 89.28%)	HBV patients who developed HCC (N=40, 10.72%)	*P* value*
Age				
≤ **Median**	172 (46.11)	166 (49.85)	6 (15.00)	**<0.0001**
> **Median**	201 (53.89)	167 (50.15)	34 (85.00)	
Gender				
Female	115 (30.83)	108 (32.43)	7 (17.50)	0.0533
Male	258 (69.17)	225 (67.57)	33 (82.50)	
Smoking status				
Never	244 (65.42)	224 (67.27)	20 (50.00)	**0.0301**
Ever	129 (34.58)	109 (32.73)	20 (50.00)	
Drinking status				
Never	220 (58.98)	198 (59.46)	22 (55.00)	0.5880
Ever	153 (41.02)	135 (40.54)	18 (45.00)	
Family history of cancer				
No	240 (64.34)	217 (65.17)	23 (57.50)	0.3389
Yes	133 (35.66)	116 (34.83)	17 (42.50)	
Cirrhosis				
No	244 (65.42)	240 (72.07)	4 (10.00)	**<0.0001**
Yes	129 (34.58)	93 (27.93)	36 (90.00)	
AFP				
≤ **Median**	181 (48.53)	172 (51.65)	9 (22.50)	**0.0005**
> **Median**	192 (51.47)	161 (48.35)	31 (77.50)	

HBV: hepatitis B virus; HCC: hepatocellular carcinoma; AFP: alpha-fetoprotein.
*The distributions of host characteristics in HBV patients who developed HCC and who did not were compared by chi-square test.

high-risk scores tended to have more risk miRNAs and less protective miRNAs, and were more likely to develop HCC during follow-up (Figure 1).

Compared to the HBV patients with low-risk scores, those with high-risk scores had a 6.56-fold increased risk for HCC (95% confidence interval [CI] 2.74-15.70) and exhibited a significantly higher cumulative incidence of HCC during follow-up ($P_{logrank}$ = 1.23×10⁻⁶, Figure 2A). After adjustment for age, gender, smoking status, drinking status, family history of cancer, cirrhosis status, and AFP values, the association of the risk score with HCC risk remained significant (hazard ratio [HR] 3.57, 95% CI 1.34-9.48).

When we further restricted the prospective analyses to a sub-cohort of patients with a 2-year exclusion window (297 HBV patients, including 30 patients who developed HCC during follow-up), the result of the sub-cohort analysis ($P_{logrank}$ = 5.00×10⁻⁵, Figure 2B) was consistent with that of the primary cohort (Figure 2A). The association between risk scores and HCC risk in the sub-cohort subjects was statistically significant in both univariate (HR 5.88, 95% CI 2.24-15.45) and multivariate analyses (HR 2.96, 95% CI 1.05-8.32).

Prediction performance of the 15-miRNA risk scores and AFP

We then assessed the prediction performance of the 15-miRNA risk scores and AFP using concordance index (C-index) [25]. In the primary prospective cohort with the

Table 2: Differentially expressed miRNAs in HBV-HCC patients

ID	d-value	Numerator(r)	Denominator(s+s0)	q-value(%)
miR-122	2.8584	7.0516	2.4669	0
miR-99a	1.8061	2.2889	1.2673	0
miR-331	1.6990	1.7507	1.0305	0
miR-125b	1.6342	1.8521	1.1333	0
miR-23b	1.5654	1.4624	0.9342	0
miR-92a	1.2883	1.3291	1.0317	0
miR-26a	1.0043	1.1242	1.1193	3.63
miR-652	-2.4714	-4.2271	1.7104	0
miR-23a	-2.3286	-3.5700	1.5331	0
miR-27a	-2.2951	-3.2686	1.4242	0
miR-34a	-1.6421	-2.2733	1.3844	0
miR-145	-1.5791	-2.9245	1.8520	0
miR-10a	-1.5007	-2.2113	1.4735	0
miR-150	-1.1032	-1.9654	1.7815	1.95
let-7f	-1.0657	-1.0167	0.9540	1.95

HBV: hepatitis B virus; HCC: hepatocellular carcinoma.

1-year exclusion window, the C-index for the 15-miRNA risk scores was 0.74 (95% CI 0.66-0.82), based on a 5-year prediction (Table 3). A statistically significant increase in C-index was observed when 15-miRNA risk scores were incorporated into the model including AFP values only (0.68 for AFP only model *vs.* 0.82 for the model including miRNA risk scores plus AFP values, $P < 0.0001$). The sub-cohort analysis of patients with the 2-year exclusion window showed very similar results (Table 3).

In silico analyses for potential functions of the genes regulated by the 15 significant miRNAs

A total of 540 potential target genes were identified in the four algorithms of TargetScan, miRanda, miRDB, and PicTar. The Ingenuity Pathway Analysis (IPA) analysis suggested that cancer ($P = 9.88×10^{-16}$-$1.99×10^{-4}$) and gastrointestinal disease ($P = 4.97×10^{-12}$-$1.27×10^{-4}$) ranked as the top two diseases that were associated with these target genes. The top five enriched canonical pathways are shown in Figure 3A. The transforming growth factor-beta (TGF-β) signaling pathway was the most significant pathway regulated by these genes ($P = 5.14×10^{-8}$). Molecular and cellular functions related to these genes include cell growth and proliferation, cellular development, gene expression, cell morphology, and cell death and survival. Moreover, the top network identified with this method was defined as "Cell-To-Cell Signaling and Interaction, Tissue Development, Nervous System Development and Function" (Figure 3B). Functional descriptors assigned by the Ingenuity Knowledge Database to the molecules in the top network include organismal death, cancer, cell death, and apoptosis.

DISCUSSION

Aberrant expression of circulating miRNAs has been observed in HCC patients in retrospective studies, but it has rarely been tested in a prospective setting. In this study we demonstrated that a signature of 15 circulating miRNAs could prospectively predict HCC development in HBV patients, suggesting that these circulating miRNAs may be candidates for next-step validations as promising HCC risk predictors.

Recent studies identified differentially expressed circulating miRNAs in HCC patients that could separate HCC patients from cancer-free HBV patients [19, 20, 23, 24, 26, 27]. However, most studies adopted a retrospective design that used blood samples collected at or after HCC diagnosis. Therefore, their conclusions were significantly constrained by the reverse-causation limitation that is inherent in most retrospective studies. Our prospective analyses with a 1-year exclusion window circumvented the reverse-causation limitation. Using blood samples collected before HCC diagnosis, we demonstrated the prospective predictive value of a panel of circulating miRNAs, including miR-92a and miR-145, the two miRNAs which were recently validated in two Chinese cohorts [23, 24]. To

minimize the confounding effect of patients who actually had HCC but were not diagnosed at study entry, we further restricted the analyses to a sub-cohort of patients with a 2-year exclusion window. This more stringent analysis yielded very similar results, further substantiating the prospective value of our circulating miRNA signature.

The differential miRNAs expressions between HBV and HCC patients in our patient cohort are consistent with previous functional studies. MiR-92a is commonly considered an onco-miRNAs because it up-regulates angiogenesis and increases cell proliferation [28]. Several studies have detected significantly high levels of circulating miR-92a in HBV-related HCC patients [19, 24, 29], consistent with our findings. Let-7a, the most important member of the let-7 tumor suppressor family, inhibits HCC cell proliferation via down-regulating signal transducer and activator of transcription 3 and c-myc, and up-regulating p16[INK4A] [30, 31]. In line with our findings, a
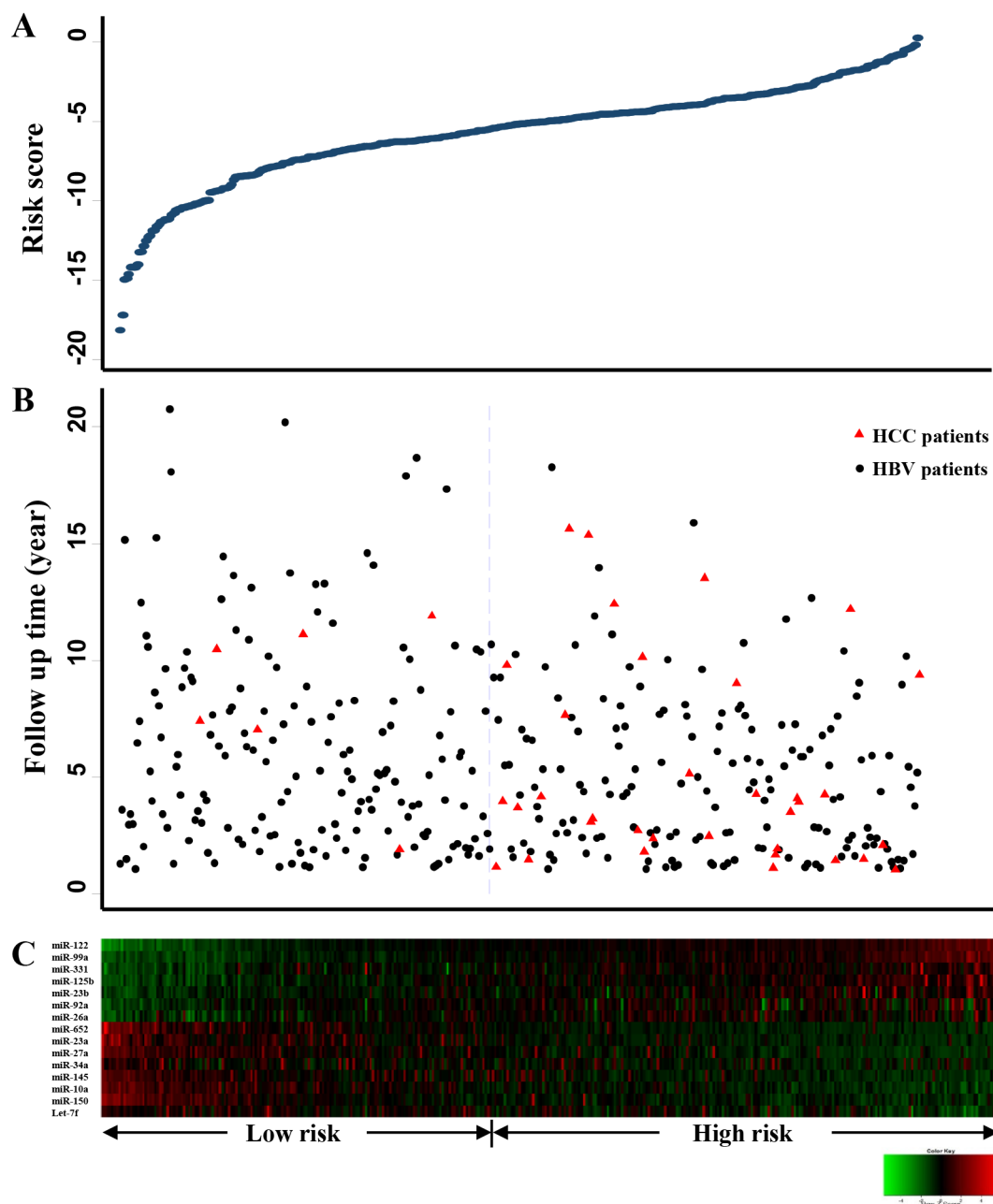

Figure 1: Risk score analyses of 373 HBV patients. A. Risk score distribution, X-axis represents patients, and each dot in the blue line represents the risk score for each patient; **B.** The status of HBV patients with/without hepatocellular carcinoma (HCC) and follow-up time, red triangles indicate HCC patients, and black dots indicate cancer-free HBV patients; **C.** Color-gram of miRNA expression profiles of HBV patients, rows represent miRNAs, and columns represent patients.
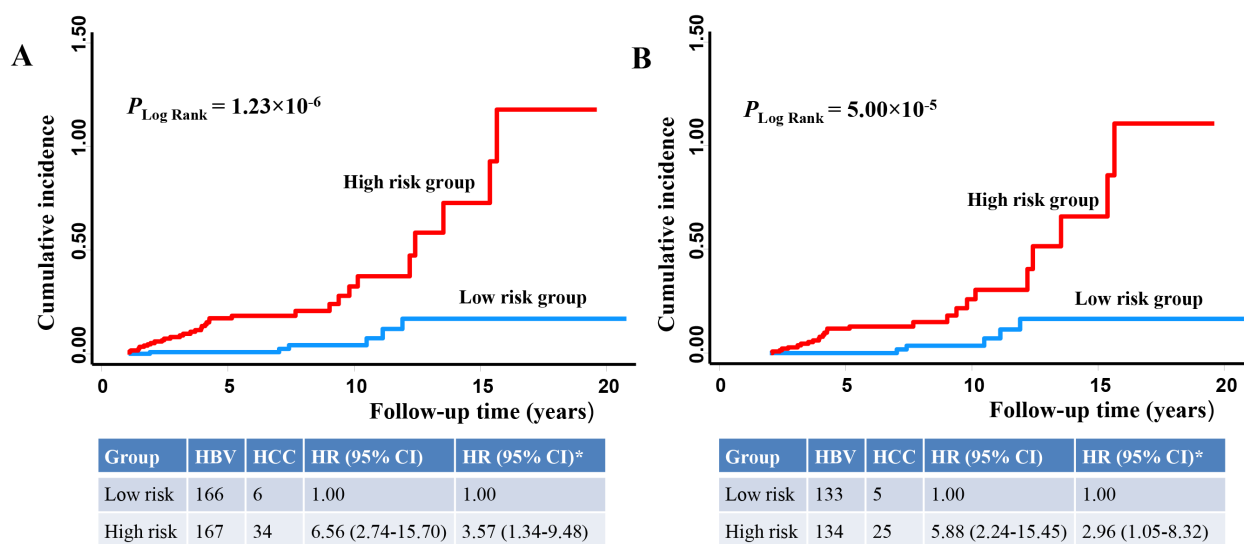
recent study [19] showed that 3 serum miRNAs, including let-7f, could separate HBV-HCC cases from controls. MiR-122, the most abundant liver-specific miRNA, is considered a suppressor of cell proliferation and malignant transformation of hepatocytes with remarkable tumor inhibition activity [32–34]. Interestingly, although down-regulation of miR-122 was frequently reported in the HCC tumor tissues [2, 34, 35], Qi *et al.* [36] showed that miR-122 in serum was elevated in HCC patients and markedly reduced in post-operative serum samples, which is consistent with the observation in our study. The paradox of miR-122 expression in blood and tissues may be attributed to the increased release of miRNAs from tumor cells into blood [13], and needs further investigation.

AFP is currently the only non-invasive circulating marker for HCC screening, but it has limited sensitivity and specificity [13, 37]. Identification of novel non-invasive markers that complement AFP is important to the early detection of HCC. Our results showed that serum miRNAs significantly improved the prediction performance of AFP (Table 3). During a 5-year follow-up period, miRNA risk scores identified 15 of 16 patients (94%) who were misclassified as cancer-free patients by AFP, significantly improving sensitivity (data not shown). In contrast, among the 57 patients who were misclassified as having HCC due to high AFP values, miRNA risk scores correctly re-classified 33 (58%), reducing false positive findings. Nonetheless, although miRNA risk scores significantly increased the sensitivity of HCC prediction, they did not significantly improve specificity, as evidenced by the large number of cancer-free HBV patients with high risk scores (Figure 1B). This was consistent with findings from recent prospective studies that reported miRNA signatures with high sensitivity and low specificity [23, 24]. Efforts are needed to identify additional miRNAs in order to further reduce false positives and increase specificity in HCC prediction.

Whether miRNAs behave as oncogenes or tumor suppressor genes depends on the cellular functions of their target genes [13, 38]. We conducted *in silico* analyses and found that the TGF-β signaling pathway is at the top of the list of canonical pathways regulated by the significant miRNAs identified in our study. TGF-β protein modulates a broad spectrum of cellular processes [39], has dual regulatory functions in the immune system, and is activated in response to viral infection, such as HBV [40]. Indeed, an initial step triggering HCC in HBV-infected patients is epithelial-to-mesenchymal transformation, which is modulated by TGF-β [41]. The inhibition of TGF-β is directed against some of the most fundamental pathological events which lead to chronic liver disease and subsequently to HCC [42]. Echoing these reports, our data highlighted the importance of the TGF-β signaling pathway in HCC initiation, suggesting a potential link between aberrant miRNA expression, TGF-β, and HCC development in HBV patients. Functional studies are warranted to further elucidate the biological mechanisms underlying these observations.

Our study has several strengths. It is among the very few studies to analyze circulating miRNAs in HCC development using a prospective design, which allows it to bypass the reverse-causation limitation that

A $P_{\text{Log Rank}} = 1.23 \times 10^{-6}$

Group	HBV	HCC	HR (95% CI)	HR (95% CI)*
Low risk	166	6	1.00	1.00
High risk	167	34	6.56 (2.74-15.70)	3.57 (1.34-9.48)

B $P_{\text{Log Rank}} = 5.00 \times 10^{-5}$

Group	HBV	HCC	HR (95% CI)	HR (95% CI)*
Low risk	133	5	1.00	1.00
High risk	134	25	5.88 (2.24-15.45)	2.96 (1.05-8.32)

HBV: hepatitis B virus; HCC: hepatocellular carcinoma; HR: hazard ratio; CI: confidence interval
*Adjusted for age, gender, smoking status, drinking status, family history of cancer, cirrhosis status, and alpha-fetoprotein values.

Figure 2: 15-miRNA risk scores associated with the risk of HBV-related hepatocellular carcinoma (HCC). The cumulative incidence of HCC in HBV patients at a high (red line) or a low (blue line) risk group based on the median cutoff of risk scores were plotted using the Nelson-Aalen method in patients with a **A.** 1-year or **B.** 2-year exclusion window.

Table 3: Discriminatory accuracy for predicting the development of HCC within 5 years

5-year risk prediction	C-index (95% CI)
1-year exclusion window	
miRNA risk scores only	0.74 (0.66-0.82)
AFP values only	0.68 (0.58-0.78)
miRNA risk scores plus AFP values	0.82 (0.74-0.89)
2-year exclusion window	
miRNA risk scores only	0.75 (0.66-0.83)
AFP values only	0.66 (0.53-0.79)
miRNA risk scores plus AFP values	0.83 (0.76-0.90)

HCC: hepatocellular carcinoma; AFP: alpha-fetoprotein; CI: confidence interval; C-index: concordance index.

is inherent in retrospective studies. The vast majority of our patients were infected with HBV at birth or childhood, making our population an ideal cohort to study the long-term risk of HCC development associated with HBV infection. The study was restricted to patients of Korean ancestry to eliminate the potential confounding effects of patient ethnicity. There are also limitations in our study. Firstly, rather than conducting a comprehensive genome-wide profiling using miRNA microarray or small RNA sequencing, we used a candidate approach to select promising miRNAs based on the findings of previous studies. Although cost-effective, this approach might miss novel important miRNAs not identified in previous studies. This limitation may account for the low specificity

of the miRNA risk score observed in our study, and could be remedied in the future by adding more significant miRNAs to the score. Secondly, given the potential interindividual, intraindividual, analytic, and preanalytic variations that may affect circulating miRNA levles [43], detecting miRNA expression using blood samples collected at only one time point, although routinely used in population studies [24, 44, 45], may introduce sample selection bias. Thirdly, we used cancer-free HBV patients, instead of HBV-free healthy subjects, as controls in this study, which may limit the generalization of our findings. Fourthly, malignant transformation induced by chronic HBV infection is a multistage pathogenic process and involves multiple risk predictors [37]. Several variables
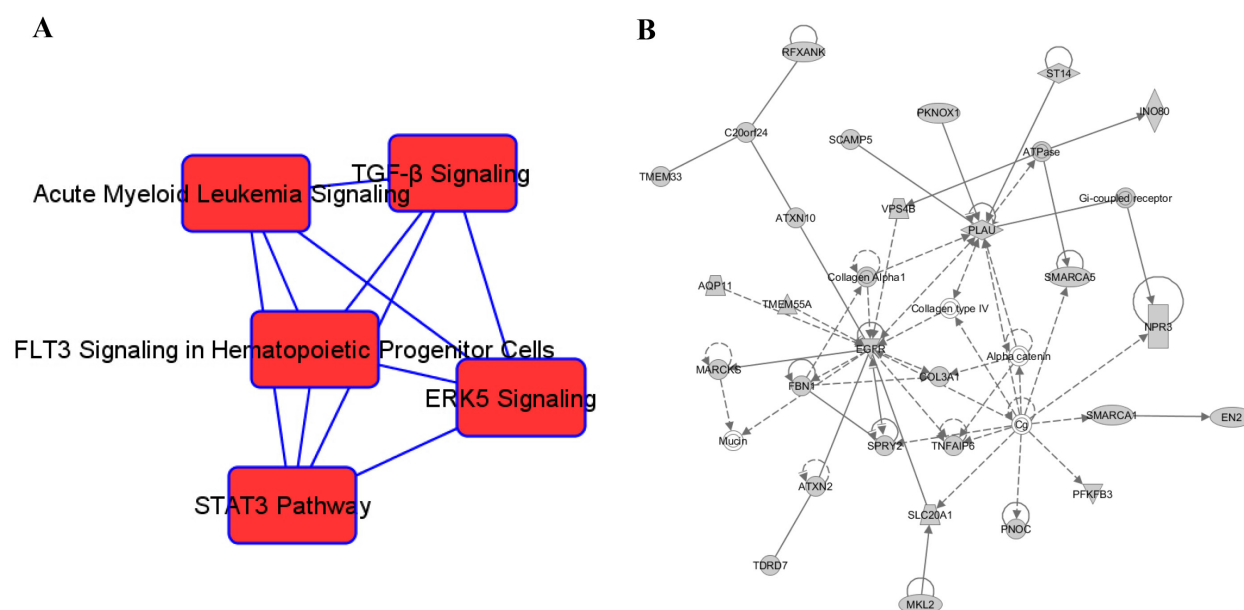

Figure 3: *In silico* **analyses for potential functions of the genes regulated by the 15 significant miRNAs. A.** The top five enriched canonical pathways implicated in miRNA-mediated HCC development and **B.** the top-ranked network formed by predicted miRNA target genes and functionally related molecules were identified.

usually collected in population studies were not recorded in our medical charts, such as dietary intake, occupational exposure, physical activity, and socioeconomic status. However, the majority of host factors that are consistently associated with HBV progression and HCC risk were available from our charts. Finally, despite the many strengths and unique prospective nature, our study has a relatively small number of patients who developed HCC and thus our findings need to be validated using larger and independent prospective patients cohorts.

In conclusion, using a prospective approach to analyze circulating miRNAs in baseline serum samples collected at least 1 year before HCC diagnosis, we identified aberrant serum miRNA expressions that were associated with HBV-related HCC risk. Our findings suggested that circulating miRNAs may serve as non-invasive biomarkers for the risk prediction of HCC in HBV patients.

MATERIALS AND METHODS

Study population

This prospective study was based on an existing and ongoing clinic-based cohort established in 1988 at the Liver Disease Prevention Center of the Thomas Jefferson University Hospital. Patients who visited the Center for treatment of liver-related diseases, including HCC, were consecutively enrolled [6]. There were no restrictions on age, gender, and disease etiology in patient recruitment. Approximately 90% of the patients had HBV infection, and over 90% were of Korean ancestry. Our study only included cancer-free HBV patients and HBV-related HCC patients of Korean ancestry in order to eliminate the confounding effects of population stratification and disease etiology. The patients included in the primary prospective cohort also met the following additional criteria: 1) had a 1-year exclusion window; 2) had baseline blood samples collected at least 1 year before the diagnosis of HCC; 3) had available major demographic and clinical data. This study was approved by the Institutional Review Board at Thomas Jefferson University.

Epidemiological and clinical data collection

Demographic and clinical data were collected for each patient from medical chart review and consultation with the treating physicians. Demographic variables included age, gender, ethnicity, smoking status, drinking status, cirrhosis, and family history of cancer. Ever smokers and drinkers were defined as described previously [6]. Liver cirrhosis and HCC were diagnosed by the combination of clinical diagnosis and imaging techniques (ultrasound, computed tomography, or magnetic resonance imaging), complemented by blood markers such as AFP.

Blood samples were drawn from each patient when taking clinical laboratory testing.

Total RNA isolation

Total RNA was isolated from 400 μL serum samples using the miRNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. To allow for normalization of sample-to-sample variation in RNA isolation, synthetic *C. elegans* miRNAs cel-miR-39 and cel-miR-54 (RNA oligonucleotides synthesized by Qiagen) were added to each sample [17, 34, 46]. The concentration of total small RNA molecules was quantified by a Qubit Fluorometer (Life Technologies). The integrity and size distribution of the total small RNAs were analyzed using the small RNA chip on a Bioanalyzer 2100 (Agilent).

Selection of miRNAs to be tested in this study

We conducted a comprehensive literature review to select candidate miRNAs to be included in this study. Literatures were identified by searching PubMed using keywords "microRNA", "serum", "plasma, "circulating", and "HCC" to identify studies that reported miRNAs related to the risk prediction and early detection of HCC. Additional studies were identified and supplied by checking the references of these original studies and related review articles. We excluded the publications that were 1) unrelated to miRNA expression, 2) unrelated to HCC risk, 3) related to HCV-HCC only, 4) related to miRNA expression in liver tissues only. Finally, we identified eight references (by 2014) on serum or plasma miRNA profiling in HBV/HCC patients [13, 19, 20, 34, 36, 37, 47, 48]. We then selected candidate miRNAs following the criteria: 1) miRNAs were expressed in human serum or plasma samples; and 2) the expression levels of miRNAs were significantly different between cancer-free HBV patients and HBV-related HCC patients. We selected a total of 24 miRNAs (let-7a, let-7c, let-7f, miR-10a, miR-21, miR-23a, miR-23b, miR-26a, miR-27a, miR-34a, miR-92a, miR-98, miR-99a, miR-122, miR-125b, miR-145, miR-150, miR-223, miR-331, miR-342-3p, miR-375, miR-338-3p, miR-423-3p, and miR-652) to determine their expression levels and explore their prospective association with HCC risk.

Determination of miRNA expression levels

The expression levels of selected miRNAs were quantitated using quantitative real-time PCR (qRT-PCR) assay as described previously [17, 49]. Briefly, purified serum RNA with a fixed volume was reverse transcribed using the TaqMan miRNA Reverse Transcription Kit and miRNA-specific stem-loop primers (Applied Biosystems, CA, USA) using an Eppendorf Thermal Cycler at 16°C for 30 min, 42°C for 30 min, and 85°C for 5 min. The 6 μL of PCR reaction included 2 μL of diluted (1:10) reverse

transcription product, 3 μL of TaqMan 2X Universal PCR Master Mix, 0.3 μL of 20X TaqMan miRNA Probe Assay, and 0.7 μL of nuclease-free water. Real-time PCR was carried out on a ViiA 7 real-time PCR platform (Life Technologies) at 95 °C for 10 min, followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min in a 384-well plate. All samples were run in duplicate. To ensure the reliability of experimental data, several negative controls, including no-template, no reverse-transcriptase, and no-reaction buffer negative controls were added in the reverse transcription phase, and no-template, water-template, no polymerase, and no-reaction buffer negative controls were added in the PCR phase. The *C. elegans* miRNAs were used as positive controls [17]. The raw cycle threshold (Ct) value for each miRNA was normalized to the raw Ct value for spike-in *C. elegans* miRNAs obtained from each individual sample to eliminate potential variations introduced from the processes of RNA isolation and quantification [49].

Identification of differentially expressed miRNAs

Using Relative Quantification Software version 2.3 (Life Technologies), the Ct values of miRNAs were determined using default threshold settings and the mean Ct was calculated. Data points that generated duplicated Ct values with more than one cycle variance were repeated. The delta Ct values were defined as the mean Ct value for a specific miRNA in an individual sample minus the averaged Ct value of spike-in synthetic *C. elegans* miRNAs for the same sample. Those miRNAs with ≥ 50% missing values (miR-98 and miR-338-3p) were excluded from further analysis. The delta Ct values were then log2 transformed and median centered by arrays using the adjusting data function of Cluster 3.0 software for cluster analysis (Stanford University). The significant differentially expressed miRNAs between HBV patients with and without HCC development were identified using SAM (version 4.0) [50]. SAM gives estimates of the false discovery rate (FDR), which is the proportion of genes likely to have been false positively identified by chance as being significant. The effect of multiple comparisons was controlled through FDR [51]. In our study, FDR was estimated with 5000 permutations (the maximum permissible number), with a threshold of ≤ 5% as being significant. Unsupervised hierarchical clustering using Euclidean distance and Centroid linkage in Cluster 3.0 was carried out to cluster variables into groups based on their similarity, and the results were visualized in a dendrogram by Java TreeView version1.1.6r4 [52].

miRNAs target prediction, pathway analysis, and network analysis

To explore the molecular functions of the significant miRNAs identified in our study, the potential target genes of these miRNAs were *in silico* predicted using four miRNA target prediction algorithms, including the TargetScan (http://www.targetscan.org), miRanda (http://www.microrna.org/microrna/home.do), miRDB (http://mirdb.org/miRDB), and PicTar (http://pictar.mdc-berlin.de). We compiled the target genes that are predicted by all of the four databases, and explored their functional interrelatedness using the IPA software (Ingenuity Systems, Redwood City, CA). Top-ranked pathways that are implicated in miRNA-mediated HCC development were determined based on the incidence of predicted miRNA targets in a list of canonical pathways provided by IPA. This software also generated the top-ranked networks comprising of the predicted miRNA targets. The biological functions associated with these networks were curated through Gene Ontology [53]. Fisher's exact test was used to calculate a *P* value determining the significance of the probability of the association between the genes in the dataset and the canonical pathways.

Statistical analyses

The clinical outcome analyzed in this study was the diagnosis of HCC. Time to the event was defined as the time from blood sample taken to the date of HCC diagnosis (for HCC patients), or last follow-up (for cancer-free HBV patients). All patients included in this study had been followed for at least 1 year and had not developed HCC within this 1-year follow-up (1-year exclusion window). To minimize the confounding effect of any patient who had undiagnosed HCC at baseline sample collection, we further restricted the analyses to a sub-cohort of patients with a 2-year exclusion window (followed for at least 2 years and had not developed HCC within these 2 years). Chi-square testing was used to compare the distributions of host characteristics in HBV patients who developed HCC and those who did not. To analyze the significant miRNAs as a panel, we assigned each patient a risk score that was developed based on a linear combination of the expression levels of the miRNAs multiplied by the SAM d-values, and then used the median risk score in cancer-free HBV patients as the cutoff point [44, 49, 54]. The cumulative incidence of HCC by follow-up years was plotted using the Nelson-Aalen method, and the difference between the patients with high- and low-risk scores was compared using log-rank test. The association between miRNAs risk scores and HCC risk was estimated using HR with 95% CI calculated by a univariate and multivariate Cox proportional hazards model, adjusting for age, gender, smoking status, drinking status, family history of cancer, cirrhosis status, and AFP values. Discrimination accuracy for predicting the development of HCC within 5 years after the baseline blood sample collection using miRNAs risk scores (alone or combined with baseline AFP) was evaluated by the C-index. The differences in discrimination accuracy between different

models (miRNA scores with and without AFP) were assessed by 10,000 bootstrap resampling. SAS (version 9.3, SAS Institute, Cary, NC) and STATA (version 11.0, STATA Corp., College station, TX) software packages were used for these analyses. All *P* values were 2-sided, with a *P* value of less than 0.05 considered the threshold for statistical significance.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

The authors declare that have no conflicts of interest.

GRANT SUPPORT

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin. 2011; 61:69-90.

2. Wang L, Zhang Z, Wang FS. The efficacy of miRNA122, a novel therapeutic target, for predicting the progression of hepatocellular carcinoma (HCC). Cell Mol Immunol. 2012; 9:103-104.

3. Nordenstedt H, White DL, El-Serag HB. The changing pattern of epidemiology in hepatocellular carcinoma. Dig Liver Dis. 2010; 42:S206-214.

4. McClune AC, Tong MJ. Chronic hepatitis B and hepatocellular carcinoma. Clin Liver Dis. 2010; 14:461-476.

5. Kim WR. Epidemiology of hepatitis B in the United States. Hepatology. 2009; 49:S28-34.

6. Fu X, Wan S, Hann HW, Myers RE, Hann RS, Au J, Chen B, Xing J, Yang H. Relative telomere length: a novel non-invasive biomarker for the risk of non-cirrhotic hepatocellular carcinoma in patients with chronic hepatitis B infection. Eur J Cancer. 2012; 48:1014-1022.

7. Blum HE. Hepatocellular carcinoma: therapy and prevention. World J Gastroenterol. 2005; 11:7391-7400.

8. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet. 2004; 5:522-531.

9. Chen K, Rajewsky N. The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet. 2007; 8:93-103.

10. Huntzinger E, Izaurralde E. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet. 2011; 12:99-110.

11. Imbeaud S, Ladeiro Y, Zucman-Rossi J. Identification of novel oncogenes and tumor suppressors in hepatocellular carcinoma. Semin Liver Dis. 2010; 30:75-86.

12. Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer. 2006; 6:857-866.

13. Giordano S, Columbano A. MicroRNAs: new tools for diagnosis, prognosis, and therapy in hepatocellular carcinoma? Hepatology. 2013; 57:840-847.

14. Gramantieri L, Fornari F, Callegari E, Sabbioni S, Lanza G, Croce CM, Bolondi L, Negrini M. MicroRNA involvement in hepatocellular carcinoma. J Cell Mol Med. 2008; 12:2189-2204.

15. Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, Shimotohno K. Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. Oncogene. 2006; 25:2537-2545.

16. Ji J, Wang XW. New kids on the block: diagnostic and prognostic microRNAs in hepatocellular carcinoma. Cancer Biol Ther. 2009; 8:1686-1693.

17. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A. 2008; 105:10513-10518.

18. Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, Guo J, Zhang Y, Chen J, Guo X, Li Q, Li X, Wang W, et al. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res. 2008; 18:997-1006.

19. Li LM, Hu ZB, Zhou ZX, Chen X, Liu FY, Zhang JF, Shen HB, Zhang CY, Zen K. Serum microRNA profiles serve as novel biomarkers for HBV infection and diagnosis of HBV-positive hepatocarcinoma. Cancer Res. 2010; 70:9798-9807.

20. Zhou J, Yu L, Gao X, Hu J, Wang J, Dai Z, Wang JF, Zhang Z, Lu S, Huang X, Wang Z, Qiu S, Wang X, et al. Plasma microRNA panel to diagnose hepatitis B virus-related hepatocellular carcinoma. J Clin Oncol. 2011; 29:4781-4788.

21. Sukata T, Sumida K, Kushida M, Ogata K, Miyata K, Yabushita S, Uwagawa S. Circulating microRNAs, possible indicators of progress of rat hepatocarcinogenesis from early stages. Toxicol Lett. 2011; 200:46-52.

22. Gao P, Wong CC, Tung EK, Lee JM, Wong CM, Ng IO. Deregulation of microRNA expression occurs early and accumulates in early stages of HBV-associated multistep hepatocarcinogenesis. J Hepatol. 2011; 54:1177-1184.

23. Lin XJ, Chong Y, Guo ZW, Xie C, Yang XJ, Zhang Q, Li SP, Xiong Y, Yuan Y, Min J, Jia WH, Jie Y, Chen MS,

et al. A serum microRNA classifier for early detection of hepatocellular carcinoma: a multicentre, retrospective, longitudinal biomarker identification study with a nested case-control study. Lancet Oncol. 2015; 16:804-815.

24. Wen Y, Han J, Chen J, Dong J, Xia Y, Liu J, Jiang Y, Dai J, Lu J, Jin G, Han J, Wei Q, Shen H, et al. Plasma miRNAs as early biomarkers for detecting hepatocellular carcinoma. Int J Cancer. 2015; 137:1679-1690.

25. Wen CP, Lin J, Yang YC, Tsai MK, Tsao CK, Etzel C, Huang M, Hsu CY, Ye Y, Mishra L, Hawk E, Wu X. Hepatocellular carcinoma risk prediction model for the general population: the predictive power of transaminases. J Natl Cancer Inst. 2012; 104:1599-1611.

26. Qu KZ, Zhang K, Li H, Afdhal NH, Albitar M. Circulating microRNAs as biomarkers for hepatocellular carcinoma. J Clin Gastroenterol. 2011; 45:355-360.

27. Gui J, Tian Y, Wen X, Zhang W, Zhang P, Gao J, Run W, Tian L, Jia X, Gao Y. Serum microRNA characterization identifies miR-885-5p as a potential marker for detecting liver pathologies. Clin Sci. 2011; 120:183-193.

28. Shigoka M, Tsuchida A, Matsudo T, Nagakawa Y, Saito H, Suzuki Y, Aoki T, Murakami Y, Toyoda H, Kumada T, Bartenschlager R, Kato N, Ikeda M, et al. Deregulation of miR-92a expression is implicated in hepatocellular carcinoma development. Pathol Int. 2010; 60:351-357.

29. Giray BG, Emekdas G, Tezcan S, Ulger M, Serin MS, Sezgin O, Altintas E, Tiftik EN. Profiles of serum microRNAs; miR-125b-5p and miR223-3p serve as novel biomarkers for HBV-positive hepatocellular carcinoma. Mol Biol Rep. 2014; 41:4513-4519.

30. Lan FF, Wang H, Chen YC, Chan CY, Ng SS, Li K, Xie D, He ML, Lin MC, Kung HF. Hsa-let-7g inhibits proliferation of hepatocellular carcinoma cells by downregulation of c-Myc and upregulation of p16(INK4A). Int J Cancer. 2011; 128:319-331.

31. Wang Y, Lu Y, Toh ST, Sung WK, Tan P, Chow P, Chung AY, Jooi LL, Lee CG. Lethal-7 is down-regulated by the hepatitis B virus x protein and targets signal transducer and activator of transcription 3. J Hepatol. 2010; 53:57-66.

32. Szabo G, Bala S. MicroRNAs in liver disease. Nat Rev Gastroenterol Hepatol. 2013; 10:542-552.

33. Hu J, Xu Y, Hao J, Wang S, Li C, Meng S. MiR-122 in hepatic function and liver diseases. Protein Cell. 2012; 3:364-371.

34. Mo MH, Chen L, Fu Y, Wang W, Fu SW. Cell-free circulating miRNA biomarkers in cancer. J Cancer. 2012; 3:432-448.

35. Negrini M, Gramantieri L, Sabbioni S, Croce CM. microRNA involvement in hepatocellular carcinoma. Anticancer Agents Med Chem. 2011; 11:500-521.

36. Qi P, Cheng SQ, Wang H, Li N, Chen YF, Gao CF. Serum microRNAs as biomarkers for hepatocellular carcinoma in Chinese patients with chronic hepatitis B virus infection. PLoS One. 2011; 6:e28486.

37. Chen CJ, Lee MH. Early diagnosis of hepatocellular carcinoma by multiple microRNAs: validity, efficacy, and cost-effectiveness. J Clin Oncol. 2011; 29:4745-4747.

38. Lujambio A, Lowe SW. The microcosmos of cancer. Nature. 2012; 482:347-355.

39. Majumdar A, Curley SA, Wu X, Brown P, Hwang JP, Shetty K, Yao ZX, He AR, Li S, Katz L, Farci P, Mishra L. Hepatic stem cells and transforming growth factor beta in hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2012; 9:530-538.

40. Karimi-Googheri M, Daneshvar H, Nosratabadi R, Zare-Bidaki M, Hassanshahi G, Ebrahim M, Arababadi MK, Kennedy D. Important roles played by TGF-beta in hepatitis B infection. J Med Virol. 2014; 86:102-108.

41. Cougot D, Neuveut C, Buendia MA. HBV induced carcinogenesis. J Clin Virol. 2005; 34:S75-78.

42. Giannelli G, Mazzocca A, Fransvea E, Lahn M, Antonaci S. Inhibiting TGF-beta signaling in hepatocellular carcinoma. Biochim Biophys Acta. 2011; 1815:214-223.

43. Zhao H, Shen J, Hu Q, Davis W, Medico L, Wang D, Yan L, Guo Y, Liu B, Qin M, Nesline M, Zhu Q, Yao S, et al. Effects of preanalytic variables on circulating microRNAs in whole blood. Cancer Epidemiol Biomarkers Prev. 2014; 23:2643-2648.

44. Hu Z, Chen X, Zhao Y, Tian T, Jin G, Shu Y, Chen Y, Xu L, Zen K, Zhang C, Shen H. Serum microRNA signatures identified in a genome-wide serum microRNA expression profiling predict survival of non-small-cell lung cancer. J Clin Oncol. 2010; 28:1721-1726.

45. Schultz NA, Dehlendorff C, Jensen BV, Bjerregaard JK, Nielsen KR, Bojesen SE, Calatayud D, Nielsen SE, Yilmaz M, Hollander NH, Andersen KK, Johansen JS. MicroRNA biomarkers in whole blood for detection of pancreatic cancer. JAMA. 2014; 311:392-404.

46. Wang X. A PCR-based platform for microRNA expression profiling studies. RNA. 2009; 15:716-723.

47. Xu J, Wu C, Che X, Wang L, Yu D, Zhang T, Huang L, Li H, Tan W, Wang C, Lin D. Circulating microRNAs, miR-21, miR-122, and miR-223, in patients with hepatocellular carcinoma or chronic hepatitis. Mol Carcinog. 2011; 50:136-142.

48. Zhang Q, Pu R, Du Y, Han Y, Su T, Wang H, Cao G. Non-coding RNAs in hepatitis B or C-associated hepatocellular carcinoma: potential diagnostic and prognostic markers and therapeutic targets. Cancer Lett. 2012; 321:1-12.

49. Wang Y, Gu J, Roth JA, Hildebrandt MA, Lippman SM, Ye Y, Minna JD, Wu X. Pathway-based serum microRNA profiling and survival in patients with advanced stage non-small cell lung cancer. Cancer Res. 2013; 73:4801-4809.

50. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001; 98:5116-5121.

51. Duale N, Lindeman B, Komada M, Olsen AK, Andreassen A, Soderlund EJ, Brunborg G. Molecular portrait of cisplatin induced response in human testis cancer cell lines based on gene expression profiles. Mol Cancer. 2007; 6:53.

52. Saldanha AJ. Java Treeview—extensible visualization of microarray data. Bioinformatics. 2004; 20:3246-3248.

53. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT. Gene Ontology: tool for the unification of biology. Nat Genet. 2000; 25:25-29.

54. Wang Z, Bao Z, Yan W, You G, Wang Y, Li X, Zhang W. Isocitrate dehydrogenase 1 (IDH1) mutation-specific microRNA signature predicts favorable prognosis in glioblastoma patients with IDH1 wild type. J Exp Clin Cancer Res. 2013; 32:59.