RESEARCH ARTICLE

Significantly different expression levels of microRNAs associated with vascular invasion in hepatocellular carcinoma and their prognostic significance after surgical resection

Sung Kyu Song1*, Woon Yong Jung2*, Seung-Keun Park3, Chul-Woon Chung1, Yongkeun Park1

1 Department of Surgery, Catholic Kwandong University International St. Mary’s Hospital, Incheon, Republic of Korea, 2 Department of Pathology, Hanyang University Guri Hospital, Hanyang University College of Medicine, Gyeonggi-do, Republic of Korea, 3 Department of Supercomputing M&S Technology Development, Korea Institute of Science and Technology Information, Daejeon, Republic of Korea

*These authors contributed equally to this work.
* ykpark@ish.ac.kr

Abstract

Background

Although gross vascular invasion (VI) has prognostic significance in patients with hepatocellular carcinoma (HCC) who have undergone hepatic resection, few studies have investigated the relationship between gross VI and aberrant expression of microribonucleic acids (miRNAs and miRs). Thus, the objective of this study was to identify miRNAs selectively expressed in HCC with gross VI and investigate their prognostic significance.

Materials and methods

Eligible two datasets (accession number: GSE20594 and GSE67140) were collected from the National Center for Biotechnology Information’s (NCBI) Gene Expression Omnibus (GEO) database to compare miRNAs expression between HCC with and without gross VI. Differentially expressed miRNAs were externally validated using expression data from The Cancer Genome Atlas (TCGA) database. Prognostic significance and predicted functions of selected miRNAs for HCC were also investigated.

Results

Thirty-five miRNAs were differentially expressed between HCC with and without gross VI in both datasets. Among them, three miRNAs were validated using TCGA database. miR-99a, miR-100, and miR-148a were downregulated to a greater extent in patients with HCC and gross VI than in those with HCC but no gross VI. Receiver operating characteristic (ROC) curve analysis showed discriminatory power of these miRNAs in predicting gross VI. Multivariate survival analysis revealed that types of surgery, advanced tumor node metastasis
TNM) stage, and low expression of miR-100-5p were independently associated with tumor recurrence. It also revealed that types of surgery, advanced TNM stage, low expression of miR-100-5p and miR-148a-3p were independent risk factors for overall survival (OS) after hepatic resection for HCC. A text mining analysis revealed that these miRNAs were linked to multifaceted hallmarks of cancer, including “invasion and metastasis.”

Conclusions
Low expressions of miR-100-5p and miR-148a-3p were associated with gross VI and poor survival of patients after hepatic resection for HCC.

Introduction
Hepatocellular carcinoma (HCC) has received increasing attention because of its frequent diagnosis worldwide with a dismal prognosis [1]. Although various curative or palliative therapeutic modalities have been administered, long-term outcomes of patients with HCC have remained poor [2]. A significant contributor to poor outcomes is the tendency of HCC toward vascular invasion (VI) which reflects the tumor’s aggressiveness [3,4]. Therefore, altered or disrupted regulatory mechanisms contributing to VI in tumor cells a barrier to overcoming this cancer and so are candidate targets for a new therapeutic trial. These targets should be distinct genetic or pathologic features of HCC with VI definitely different from those of HCC without VI.

It is well known that unique gene expressions are highly related to tumor progression [5]. Recent studies have shown that aberrant epigenetic gene regulations also play a critical role in tumor progression [6]. Epigenetic alterations include deoxyribonucleic acid (DNA) methylation, histone modification, and ribonucleic acid (RNA) interference. Of these alterations, RNA interference can cause silencing of gene expression after introduction of sense–antisense RNA pairs [7]. Cells have a large number of noncoding RNA (ncRNA) molecules, many of which are capable of RNA interference. Recently, critical roles of small ncRNAs including micro-RNAs (miRNAs, miRs) and piwi-interacting RNAs (piRNAs) in a variety of human diseases, particularly cancers, have been well elucidated [8,9]. These RNAs have been shown to play an important role in tumorigenesis and/or tumor progression. Some of these RNAs can regulate epithelial–mesenchymal transition (EMT) directly or indirectly. Thus, they are involved in tumor invasion or metastasis [10]. Recent studies have identified clinically significant abnormal patterns of miRNA expression in HCC, with some miRNAs showing marked association with aggressive tumor phenotypes or poor survival [11,12].

Several studies have demonstrated correlations of VI and aberrant expression with specific miRNAs (down-regulation of miR-30a-3p and miR-34c-3p) [13,14]. However, these studies did not analyze all miRNA profiling with RNA sequencing technology or microarray analysis. They only focused on selected miRNAs in order to elucidate whether these miRNAs were distinctly expressed between HCC with VI and HCC without VI. Therefore, the objective of this study was to identify miRNAs that could act as important contributors to VI in HCC.

Materials and methods
Study data and screening of differentially expressed miRNAs
High-throughput miRNA expression data were obtained from the National Center for Biotechnology Information’s (NCBI) Gene Expression Omnibus (GEO) dataset [15] (http://www.ncbi.nlm.nih.gov/geo/) under accession numbers GSE20594 and GSE67140, including
miRNA expression profiles of HCC tumor samples and information about VI status. Differentially expressed miRNAs between tumors on the basis of the absence or presence of VI were detected with both Wilcoxon ranked-sum test with multiple test corrections (Benjamini–Hochberg-adjusted \( p \)-value < 0.05) and RankProd method [16]. Differentially expressed miRNAs between groups were validated and their prognostic values were additionally assessed using expression data obtained from The Cancer Genome Atlas (TCGA) database. Fig 1A shows a schematic flowchart of this study.

Clinical data for 377 patients with HCC were obtained from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA LIHC) database (see https://tcga-data.nci.nih.gov and http://gdac.broadinstitute.org/) available to the public. Patients, their corresponding clinical data, and miRNA and messenger RNA (mRNA) expression profiles from the dataset were categorized into two groups on the basis of the absence or presence of gross VI. Patients with missing gross VI values were included only for survival analysis according to the expression value of each candidate miRNA. Descriptive statistics were used to compare patient characteristics including sex, age, etiology, laboratory data, types of surgery, and pathologic data. A major hepatic resection was defined as the removal of three or more segments of the liver whereas a minor resection was defined as the removal of two segments or less.
MiRNA and mRNA expression data were generated using Illumina HiSeq 2000 sequencing platforms (Illumina Inc., San Diego, CA, USA). Raw read counts were used to analyze differentially expressed miRNAs and mRNAs. Processed RNA sequencing (RNA-Seq) data that were normalized according to reads per million (RPM) of miRNAs and mRNAs were also used in survival and correlation analyses. This study complied with publication guidelines provided by the TCGA (http://cancergenome.nih.gov/publications/publicationguidelines). All TCGA data are now available without restrictions on their use in publications or presentations. Differentially expressed sequence (DESeq2) package of the statistics software R version 3.3.1 was used to assess differentially expressed miRNAs on the basis of raw read counts [17]. Significantly different miRNA expression between groups was determined by false discovery rate (FDR) correction for multiple hypothesis testing using Benjamini–Hochberg-adjusted p-values with a threshold of FDR < 0.05. Receiver operating characteristic (ROC) curve analysis was performed to define optimal cutoff values of selected miRNAs on the basis of the absence or presence of gross VI. RPM expression values of selected miRNAs were used to build a logistic regression model to predict the absence or presence of gross VI.

Survival analysis
Recurrence-free survival (RFS) and overall survival (OS) were estimated for all patients. The expression values (RPM) of selected miRNAs were used as continuous variables. We also calculated optimal cutoff values for RPM values of selected miRNAs and continuous variables of clinical parameters for use in Kaplan–Meier survival analysis, which was calculated by ROC curve for 1-year tumor recurrence. In addition, survival rates and curves were estimated by the Kaplan–Meier method and compared using the log-rank test. To calculate hazard ratios (HRs) and 95% confidence intervals (CIs) and identify independent factors associated with RFS and OS, multivariate analysis was performed using Cox regression proportional hazards model. All statistical analyses were performed using Epi, GGally, moonBook and Survival packages of R version 3.3.1 [18].

Prediction of miRNA targets, functional analysis and text mining analysis
To identify target genes for the above candidate miRNAs, we obtained related genome data from NCBI GEO database which revealed changes in gene expression profile after overexpression of candidate miRNA mimic. Differentially expressed genes between control and candidate miRNA mimic-treated samples were analyzed using GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/). |log₂FC| > 1 and adjusted p-value < 0.05 were used as cut-off criteria to define statistically significant difference. Selected genes were compared to a validated target module of MiRWalk2.0 database [19]. Only common genes were chosen for further analysis. We also performed Pearson’s correlation analysis to calculate correlations between the expression of candidate miRNAs and that of predicted target genes with all available samples in the TCGA LIHC database, filtering out genes with meaningful negative correlations (FDR < 0.05). To investigate the biological processes in which selected miRNAs may be involved, functional analyses were performed. Gene ontology analyses for those filtered genes were performed by DAVID (http://david.ncifcrf.gov). Cancer Hallmarks Analytics Tool (CHAT) was used to figure out the association between selective miRNAs and documented evidence of these in hallmarks of cancer [20].

Results
miRNA differential expression analysis
As described in the detailed flowchart shown in Fig 1A, in this study, we obtained information about expression levels of miRNAs from NCBI GEO database [15]. To screen miRNAs that
were significantly more or less abundant in HCC with gross VI compared to those in HCC without gross VI, we performed differential expression analysis by comparing the expression level of each miRNA from two microarray datasets (accession number: GSE20594 and GSE67140). These datasets included 31 samples of HCC without gross VI vs. 46 samples of HCC with gross VI (accession number: GSE20594) and 91 samples of HCC without gross VI vs. 81 samples of HCC with gross VI (accession number GSE67140). We then identified differentially expressed miRNAs based on the Wilcoxon ranked-sum test and the RankProd method [16]. A total of 49 miRNAs in GSE20594 were upregulated in HCC with VI compared to those in HCC without VI. In addition, 185 miRNAs in GSE67140 showed increased expression levels in HCC with VI compared to those in HCC without VI. Ten miRNAs were upregulated in these two microarray data sets. Forty-one miRNAs (GSE20594) and 184 miRNAs (GSE67140) were downregulated in HCC with VI compared to HCC without VI while 25 miRNAs were downregulated in both data sets. These 35 miRNAs were predicted to be candidate biomarkers for gross VI of HCC (Fig 1B). They were validated using expression data obtained from TCGA database.

Data collection and overall evaluation of TCGA LIHC

We found no miRNA expression data for 5 of 377 patients with HCC in the TCGA LIHC database. The remaining 372 patients were divided into three groups by gross VI status as follows: group A (n = 300), patients with HCC without gross VI; group B (n = 17), patients with HCC and gross VI; and group C (n = 55), patients for whom no gross VI status was available. Groups A and B were included in the study to find miRNAs expressed differentially between HCC with and without gross VI. Group C was included in survival analysis according to expression value of each candidate miRNA.

Table 1 summarizes clinical characteristics of groups A, B, and C. Table 1 summarizes the clinical characteristics of groups A, B, and C. The percentage of patients who had better grades of tumor differentiation and background liver inflammation was significantly lower in group C compared to A and B (P = 0.03 and 0.003, respectively). Aside from these differences, the three groups were comparable. The overall median follow-up period was 16 months (range, 0–122 months).

Differentially-expressed miRNAs on based on the absence or presence of gross VI in the TCGA database

We obtained information about the expression levels of miRNAs from the TCGA expression data. A total of 12 miRNAs were differentially expressed between groups A and B and three miRNAs were on the above list of candidate miRNAs in two GEO datasets (Fig 1B). miR-99a, miR-100, and miR-148a were downregulated to a greater extent in group B than in group A (log2 FC = −1.06 with FDR = 0.018, log2 FC = −1.33 with FDR = 0.004, and log2 FC = −0.87 with FDR = 0.018, respectively) (Fig 2A).

Next, we assessed the discriminatory power of the three miRNAs to predict gross VI. To evaluate the predictive value, we used ROC curves to analyze sensitivity and specificity. As shown in Fig 2B, the ROC curve of miR-99a-5p showed an area under the curve (AUC) of 71.3% (sensitivity, 64.7%; specificity, 76.0%), similar to the AUC of miR-100-5p at 73.5% (sensitivity, 94.1%; specificity, 49.0%), and the AUC of miR-148a-3p at 70.7% (sensitivity, 76.5%; specificity, 63.7%). Around 88% (15 out of 17) of HCC with gross VI showed two or more aberrant expression pattern in these three miRNAs, while only 33% of HCC without gross VI did (P < 0.001).
Table 1. Comparison of clinicopathological data of hepatocellular carcinoma patients categorized into three groups on the basis of the absence or presence of gross VI.

|                       | Group A (N = 300) | Group B (N = 17) | Group C (N = 55) | P value |
|-----------------------|-------------------|------------------|------------------|---------|
| Sex                   |                   |                  |                  |         |
| Male                  | 201 (67.0%)       | 9 (52.9%)        | 42 (76.4%)       | 0.161   |
| Female                | 99 (33.0%)        | 8 (47.1%)        | 13 (23.6%)       |         |
| Age                   | 59.3 ± 13.7       | 60.6 ± 14.7      | 59.5 ± 11.6      | 0.872   |
| Viral hepatitis       |                   |                  |                  |         |
| No                    | 127 (43.3%)       | 10 (62.5%)       | 19 (34.5%)       | 0.129   |
| Yes                   | 166 (56.7%)       | 6 (37.5%)        | 36 (65.5%)       |         |
| History of other malignancy |          |                  |                  | 0.065   |
| No                    | 268 (89.3%)       | 14 (82.4%)       | 54 (98.2%)       |         |
| Yes                   | 32 (10.7%)        | 3 (17.6%)        | 1 (1.8%)         |         |
| Platelet count (×1000/μL) |         |                  |                  | 0.171   |
| W.N.L.                | 213 ± 101         | 271 ± 85         | 227 ± 93         |         |
| B.N.L.                | 251 (95.1%)       | 14 (100.0%)      | 19 (95.0%)       | 0.696   |
| Serum creatinine      |                   |                  |                  |         |
| W.N.L.                | 13 (4.9%)         | 0 (0.0%)         | 1 (5.0%)         |         |
| B.N.L.                | 166 (56.7%)       | 6 (37.5%)        | 36 (65.5%)       |         |
| Serum albumin         |                   |                  |                  | 0.619   |
| W.N.L.                | 202 (76.2%)       | 13 (86.7%)       | 14 (73.7%)       |         |
| B.N.L.                | 44 (24.2%)        | 2 (13.3%)        | 5 (26.3%)        |         |
| Serum total bilirubin |                   |                  |                  | 0.465   |
| W.N.L.                | 156 (84.8%)       | 16 (94.1%)       | 15 (88.2%)       |         |
| B.N.L.                | 63 (23.8%)        | 1 (5.9%)         | 2 (11.8%)        |         |
| Child-Turcotte-Pugh classification |      |                  |                  | 0.489   |
| A                     | 202 (91.0%)       | 11 (84.6%)       | 8 (100.0%)       |         |
| B or C                | 20 (9.0%)         | 2 (15.4%)        | 0 (0.0%)         |         |
| Alpha-fetoprotein (ng/mL) | 14628.1 ± 132595.5 | 16856.2 ± 34681.1 | 2078.1 ± 6346.0 | 0.705   |
| Types of hepatic resection |              |                  |                  | 0.054   |
| Major                 | 169 (56.7%)       | 9 (52.9%)        | 21 (38.9%)       |         |
| Minor                 | 129 (43.3%)       | 8 (47.1%)        | 33 (61.1%)       |         |
| Histologic grading by Edmondson and Steiner's classification |      |                  |                  | 0.03    |
| 1–2                   | 182 (61.1%)       | 8 (47.1%)        | 41 (77.4%)       |         |
| 3–4                   | 116 (38.9%)       | 9 (52.9%)        | 12 (22.6%)       |         |
| Residual tumor        |                   |                  |                  | 0.421   |
| Negative              | 265 (95.3%)       | 15 (88.2%)       | 45 (93.8%)       |         |
| Positive              | 13 (4.7%)         | 2 (11.8%)        | 3 (6.2%)         |         |
| Ishak fibrosis staging system |            |                  |                  | 0.22    |
| 0                     | 68 (34.5%)        | 5 (55.6%)        | 1 (11.1%)        |         |
| 1–4                   | 54 (27.4%)        | 2 (22.2%)        | 5 (55.6%)        |         |
| 5–6                   | 75 (38.1%)        | 2 (22.2%)        | 3 (33.3%)        |         |
| Grade of liver inflammation |              |                  |                  | 0.003   |
| 0                     | 96 (45.3%)        | 9 (81.8%)        | 12 (92.3%)       |         |
| 1–2                   | 98 (46.2%)        | 2 (18.2%)        | 1 (7.7%)         |         |
| 3–4                   | 18 (8.5%)         | 0 (0.0%)         | 0 (0.0%)         |         |

Notes: Group A, no gross vascular invasion; Group B, gross vascular invasion; Group C, no available information regarding vascular invasion. Mismatches of patients numbers across the variables occurs due to some missing values. B.N.L.: Beyond normal range; W.N.L.: With normal range.

https://doi.org/10.1371/journal.pone.0216847.t001
The model for predicting gross VI

We performed binary logistic regression with gross VI status as the dependent variable and expression values of three miRNAs as the independent variables in the data cohort of GSE67140. The model for predicting gross VI (MVI) was a linear combination of miRNAs, which were weighted by regression coefficients. The model is as follows: $MVI = \frac{e^{VS}}{1+e^{VS}}$, where $e^{VS}$ is the natural exponential value of the gross VI score (VS) and $VS = -0.9415 \times (\log_2 \text{value of miR-100-5p}) + 5.6175$ (P = 0.001). When performing similar logistic regression analyses with the TCGA cohort, the expression value of miR-100-5p as a predictor was significant at P = 0.03.

Survival analysis

Univariate analysis using expression values of three miRNAs as continuous variables was performed to estimate the predictive powers of three selected miRNAs for tumor recurrence and survival. Kaplan-Meier analysis indicated that the expression of miR-148a-3p was associated with poor RFS (P = 0.036) and the expression of all three miRNAs was associated with OS. ROC-curve-determined optimal cutoff values of miRNA expression for 1-year HCC recurrence in study participants (n = 372) was used to classify patients into low and high expression groups for each miRNA. The cutoff values for miR-99a-5p, miR-100-5p, and miR-148a-3p expression were 430.1, 2821.4 and 100929.3 RPM, respectively. Kaplan–Meier analysis indicated that low expression of miR-100-5p and miR-148a-3p were significantly associated with poor RFS (log-rank test: P < 0.001 and P = 0.015, respectively) and worse OS (log-rank test: P < 0.001 and P = 0.009, respectively) (Table 2). However, low expression of miR-99a-5p showed a significant difference in OS (log-rank test: P = 0.003) but not in RFS.
Table 2. Univariate analysis of factors predictive of recurrence-free survival and overall survival.

| Factors                              | No. of patients | MDFST (95% CI)       | P value | Factors                              | No. of patients | MOST (95% CI)       | P value |
|--------------------------------------|-----------------|----------------------|---------|--------------------------------------|-----------------|---------------------|---------|
| **Sex**                              |                 |                      |         | **Hepatitis B or C infection status**|                 |                     |         |
| Male                                 | 222             | 21.2 (16.5–33.0)     | 0.617   | Negative                             | 126             | 19.2 (15.7–33.0)    | 0.510   |
| Female                               | 100             | 19.2 (14.8–35.8)     |         | Positive                             | 188             | 21.6 (16.5–35.8)    |         |
| **Age (years)**                      |                 |                      |         | **Prothrombin time (sec.)**          |                 |                     |         |
| <68                                  | 228             | 19.2 (15.4–35.6)     | 0.501   | With normal range                    | 220             | 23.0 (19.2–33.9)    | 0.474   |
| ≥68                                  | 94              | 23.0 (18.4–36.7)     |         | Beyond normal range                  | 40              | 29.3 (18.3-NA)      |         |
| **Hepatitis B or C infection status**|                 |                      |         | **Serum creatinine (mg/dL)**         |                 |                     |         |
| Negative                             | 126             | 19.2 (15.7–33.0)     | 0.510   | With normal range                    | 251             | 23.0 (19.2–33.9)    | 0.564   |
| Positive                             | 188             | 21.6 (16.5–35.8)     |         | Beyond normal range                  | 10              | 45.8 (11.8-NA)      |         |
| **Serum albumin (g/dL)**             |                 |                      |         | **Serum total bilirubin (mg/dL)**    |                 |                     |         |
| With normal range                    | 206             | 21.6 (18.3–35.6)     | 0.622   | With normal range                    | 230             | 24.8 (19.4–36.7)    | 0.145   |
| Beyond normal range                  | 58              | 25.5 (15.9–71.1)     |         | Beyond normal range                  | 37              | 25.5 (14.8-NA)      |         |
| **Child–Turcotte–Pugh classification**|                 |                      |         | **Alpha-fetoprotein (ng/mL)**        |                 |                     |         |
| A                                    | 196             | 23.6 (18.6–36.7)     | 0.153   | A                                    | 220             | 2542 (1694-NA)      | 0.311   |
| B or C                               | 19              | 23.9 (8.6–NA)        |         | B or C                               | 22              | 887 (601-NA)        |         |
| **Extent of resection**              |                 |                      |         | **Residual tumor**                   |                 |                     |         |
| Major                                | 169             | 35.6 (23.0–53.5)     | 0.116   | No                                   | 288             | 21.6 (18.2–33.0)    | 0.016   |
| Minor                                | 152             | 13.6 (9.89–19.2)     |         | Yes                                  | 15              | 15.6 (10.8–NA)      |         |
| **Histologic grading by Edmondson and Steiner’s classification**|                 |                      |         | **Histologic grading by Edmondson and Steiner’s classification**|                 |                     |         |
| I–II                                 | 200             | 21.6 (18.4–30.0)     | 0.421   | I–II                                 | 230             | 1791 (1386–2542)    | 0.654   |
| III–IV                               | 118             | 15.7 (11.7–42.0)     |         | III–IV                               | 136             | 1622 (1149-NA)      |         |
| **Ishak fibrosis staging system**    |                 |                      |         | **Grade of liver inflammation**      |                 |                     |         |
| 0                                    | 60              | 23.9 (15.7–NA)       | 0.472   | 0                                    | 74              | 74 (2131 (931-NA)    | 0.472   |
| 1–4                                  | 57              | 24.8 (16.5–36.7)     |         | 1–2                                  | 61              | 61 (1791 (1372-NA)   |         |
| 5–6                                  | 73              | 18.3 (12.9–47.0)     |         | 3–4                                  | 79              | 79 (NA (1685-NA)     |         |
| **American Joint Committee on Cancer TNM stage** |                 |                      |         | **American Joint Committee on Cancer TNM stage**|                 |                     |         |
| I–II                                 | 225             | 30.0 (23.6–42.0)     | <0.001  | I–II                                 | 257             | 2456 (1694-NA)      | <0.001  |
| III–IV                               | 76              | 9.6 (8.3–14.2)       |         | III–IV                               | 89              | 660 (445–1791)      |         |

(Continued)
Multivariate analysis

In this study, to identify possible predictors of tumor recurrence and poor survival of patients with HCC after hepatic resection, we analyzed other associated clinicopathological factors by univariate and multivariate methods (Table 2). Univariate analysis showed significant associations of tumor recurrence with types of surgery ($P < 0.001$) and advanced TNM stage ($P < 0.001$). Multivariate logistic regression was performed for each predictor of recurrence, including gross VI, low expressions of miR-100-5p and miR-148a-3p, and expression level of miR-148a as a continuous variable. We found that types of surgery (HR = 1.716; 95% CI: 1.234–2.387; $P = 0.001$), advanced TNM stage (HR = 1.800; 95% CI: 1.262–2.567; $P = 0.001$), and low expressions of miR-100-5p (HR = 1.487; 95% CI: 1.031–2.146; $P = 0.033$) remained significant independent risk factors for tumor recurrence (Fig 3). Types of surgery, concomitant viral hepatitis, and advanced TNM stage were associated with poorer OS in univariate Kaplan–Meier survival estimation with log-rank comparisons ($P = 0.024$, $P = 0.018$, and $P < 0.001$, respectively). The multivariate proportional hazards model identified that types of surgery (HR = 1.948; 95% CI: 1.168–3.248; $P = 0.011$), advanced TNM stage (HR = 2.477; 95% CI: 1.471–4.170; $P < 0.001$), low expressions of miR-100-5p (HR = 2.565; 95% CI: 1.478–4.453; $P < 0.001$), and miR-148a-3p (HR = 1.767; 95% CI: 1.042–2.997; $P = 0.034$) were independent risk factors for OS after hepatic resection in HCC (Fig 3).

Functional analyses

A total of 508 validated target genes were identified by searching the three selected miRNAs in the miRWalk2.0 database. Among them, 50 genes were common targets of two of the selected miRNAs but none were for all three. Pearson correlation tests were conducted on the three miRNAs and their target genes. Pairs of each selected miRNA and their target genes with significantly negative correlation (FDR < 0.05) were filtered, resulting in 191 meaningful pairs (miRNA and its target gene). To further investigate the regulatory effects of these significantly

### Table 2. (Continued)

| Factors          | No. of patients | MDFST$^1$ (95% CI) | $P$ value | Factors          | No. of patients | MOST$^2$ (95% CI) | $P$ value |
|------------------|-----------------|--------------------|-----------|------------------|-----------------|-------------------|-----------|
| miR-99a-5p (RPM) | 0.081           |                    |           | miR-99a-5p (RPM) | 0.036           |                    |           |
| $\geq$430.09     | 212             | 23.6 (19.7–33.9)   |           | $\geq$430.09     | 236             | 1791 (1490–NA)    |           |
| <430.09          | 110             | 12.9 (9.76–30.0)   |           | <430.09          | 134             | 887 (633–NA)      |           |
| $^3$miR-99a-5p (RPM) as a CV | 0.304 |                    |           | $^3$miR-99a-5p (RPM) as a CV | <0.001 |                    |           |
| miR-100-5p (RPM) | -0.001          |                    |           | miR-100-5p (RPM) | 0.009 |                    |           |
| $\geq$2821.42    | 245             | 27.2 (20.9–36.7)   |           | $\geq$2821.42    | 277             | 2116 (1622–3258)  |           |
| <2821.42         | 77              | 10.1 (7.85–15.7)   |           | <2821.42         | 93              | 600 (431–1490)    |           |
| $^3$miR-100-5p (RPM) as a CV | 0.220 |                    |           | $^3$miR-100-5p (RPM) as a CV | 0.005 |                    |           |
| miR-148a-3p (RPM) | 0.015           |                    |           | miR-148a-3p (RPM) | 0.002 |                    |           |
| $\geq$100929.3   | 144             | 25.5 (21.02–47.7)  |           | $\geq$100929.3   | 162             | 2542 (1622–NA)    |           |
| <100929.3        | 178             | 16.1 (13.1–25.3)   |           | <100929.3        | 208             | 1386 (1147–1791)  |           |

Notes:

1 Median disease-free survival time (month)

2 Median overall survival time (day). Mismatches of patients numbers across the variables occurs due to some missing values.

3 miRNA as a continuous variable (CV)

Abbreviations: CI (confidence interval), RPM (reads per million), TNM (Tumor Node Metastasis), NA (not available).

https://doi.org/10.1371/journal.pone.0216847.t002
negative relationships, 191 target genes were mapped into the DAVID database and statistically enriched in biological process (BP), cellular components (CC), and molecular functions (MF), in which 133, 55, and 37 terms were identified, respectively. The FDRs were < 0.05 in four terms of BP, 12 of CC, and nine of MF (Fig 4). The top terms of BP, CC, and MF were the translational elongation (GO: 0006414; FDR = 2.04x10^-15), the cytosol (GO: 0005829; FDR = 4.92x10^-12) and structural constituent of ribosome (GO: 0003735; FDR = 2.30x10^-11), respectively. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were also identified and the ribosomal pathway was associated with target genes (hsa03010; FDR = 4.92x10^-12).

Fig 3. Summary of statistically significant factors on tumor recurrence and overall survival using Cox regression proportional hazards model. TNM, American Joint Committee on Cancer Tumor Node Metastasis; HR, hazard ratio; CI, confidence interval.

https://doi.org/10.1371/journal.pone.0216847.g003

Fig 4. GO functional annotation for significantly related targets of selected miRNAs (miR-99a-5p, miR-100-5p and miR-148a-3p). BP, biological process; CC, cellular component; MF, molecular function.

https://doi.org/10.1371/journal.pone.0216847.g004
Prediction of miRNA targets and text mining analysis

To identify target genes for miR-100-5p, we obtained two published microarray data from the NCBI website which revealed changes of gene expression profile after overexpression of miR-100-5p mimic (GEO accession numbers: GSE16571 and GSE20668). Each dataset included three biological replicates for both control and miRNA mimics treated cells (a total of six samples: GSE16571) and two replicates (a total of 4 samples: GSE20668). A total of 1181 genes were significantly down-regulated in miR-100-5p mimic treated samples ($\log_2 FC < -1$ and adjusted $p$-value $< 0.05$) in GSE16571 dataset. However, there was no gene expression change fulfilling the cut-off criteria in GSE20668 dataset. We also found a list of 243 predicted target genes of miR-100-5p using miRNA–mRNA pairs experimentally validated in the MiRWalk2.0 database [19]. The overlap of both lists comprised 21 genes (Fig 5A). Next, we performed Pearson’s correlation analysis to calculate correlations between the expression of miR-100-5p and of the above 21 genes with all available 424 samples (both mRNA and miRNA) from the TCGA LIHC database. Among 21 genes, expression values of ACSL3 and CTDSPL had significant negative correlations with miR-100-5p (Pearson’s $r = –0.310$ and $–0.330$, respectively, FDR $< 0.001$, Fig 5B). There was no available dataset for finding targets of miR-148a from NCBI GEO database.

A text mining analysis was conducted to determine which hallmarks of cancer were affected when miR-99a/-100/-148a were dysregulated. All miRNAs were associated with at least eight hallmarks of cancer. Especially, invasion and metastasis, immune destruction, inducing angiogenesis, resisting cell death, sustaining proliferative signaling, and tumor promoting inflammation were affected by all three miRNAs (Fig 5C).

Discussion

In this study, we used publicly available data from TCGA repositories to find specific miRNAs associated with gross VI in HCC. We found that down-regulation of miR-99a/-100/-148a were commonly associated with gross VI in three publicly open datasets (TCGA and GEO, accession numbers: GSE20594 and GSE67140). Specifically, aberrant expressions of two miRNAs
(miR-100-5p and miR-148a-3p) were independent risk factors for HCC recurrence and/or survival after hepatic resection for HCC.

Tumor biology research has focused on miRNAs. Many studies have demonstrated the involvement of miRNAs in cancer metastasis [21]. Compared with parenchyma tumors of HCC, one study has shown increased expression of miR-135a in a portal vein tumor thrombus [22]. However, it is difficult to find studies comparing expression values of all miRNAs on the basis of the absence or presence of gross VI. To the best of our knowledge, no previous study has investigated the association between aberrant expression of these three miRNAs (miR-99a, miR-100 and miR-148a, and miR-582) and gross VI in HCC tissue samples.

Gross VI is a well-known factor that contributes to tumor recurrence and poor prognosis. A recent study has reported that its prevalence in newly detected HCC is 35% [23]. Notably, gross VI is found to be a major prognostic factor following hepatic resection for HCC [24]. Unfortunately, the optimal treatment strategy for HCC with gross VI remains inconclusive. The current Barcelona clinic liver cancer staging system recommends systemic therapy with sorafenib at this stage. However, outcome following this treatment is not promising [25].

Under careful selection based on good hepatic reserve and favorable preoperative predictors, hepatic resection has been justified in patients with HCC and gross VI [26]. Although hepatic resection has clinical significance, there is limited information about the underlying mechanism for gross VI. One possible inference is that gross VI merely occurs coincidentally, with high arterial pressure in the tumor acting as a driving force of cancer cells to neighboring portal or venous vessels. However, recent genomic data analyses have demonstrated that unique genes and ncRNAs play an important role in metastasis [27].

miRNA is one of endogenously expressed small ncRNAs (~22 nucleotides in length) with capacity as posttranscriptional regulators through degradation of their target mRNA or by inhibiting translation [28]. miRNA has been known to have a regulatory role in tumor progression such as migration, angiogenesis, and invasion as a powerful regulator of critical biological processes, including cell cycle, metabolism, development, and cell differentiation [13, 29]. miR-100 has been shown to be a prognostic factor for different types of cancer [30]. One study has concluded that miR-100 down-regulation is correlated with progressive pathological features and poor prognosis in patients with HCC [31]. However, the exact role of miR-100 in preventing cancer progression has not been fully elucidated yet, although its importance has been recognized. Recently, miR-100 has been shown to be involved in suppressing migration and invasion of nasopharyngeal carcinoma cells by targeting insulin-like growth factor 1 receptor [32]. In our study, mRNA expressions of ACSL3 and CTDSPL showed significant negative correlations with miR-100, consistent with results of an earlier study showing that the expression of CTDSPL (also known as RBSP3) protein had the most inverse correlation with miR-100 in clinical samples of acute myeloid leukemia [33]. Altered fatty acid metabolism in cancer has been increasingly recognized. Long-chain acyl-CoA synthetases (ACSLs) with responsibility for activation of long-chain fatty acids are commonly deregulated in cancer. Among ACSL family, ACSL3 was found to be overexpressed in different types of cancer [34]. However, the prognostic role of ACSL3 remains unclear.

Recently, a similarly designed study using the TCGA database for investigating gross VI–related miRNAs expression profiles has been published [35]. Results of that study concluded that 16 miRNA-based classifiers could effectively identify gross VI. They also play a role as prognostic factors. However, miRNAs selected in that study were completely different from those selected in our study. In the TCGA LIHC database, we can recognize two types of VI from the pathologic data of every patient, McVI and gross VI. There is no mention about such information in previous studies. This is why different miRNAs are sorted. Because there is little evidence that McVI can be considered the first step of metastatic dissemination via the
vascular route and McVI will advance to gross VI, we thought it was not justified that McVI and gross VI seemed to be the same pathobiological feature. Also, the prognostic impact of McVI remains inconclusive [36,37]. Therefore, patients with HCC with McVI were excluded from analysis in our study.

Over the past two decades, there has been a dramatic increase in genomic databases for diseases, including those for cancer. The TCGA provides a large volume of publicly available cancer genomic data. Researchers interested in molecular biology of cancer can access this valuable source of data. In the present study, we used epigenetic profiles of each HCC tumor sample with clinical parameters, enabling the establishment of new knowledge about the role of miRNAs in HCC with gross VI. In the near future, it is likely that newly discovered molecular targets based on the TCGA database will be applied to clinical cancer practice, including for early detection, treatment, and even prevention of cancers. However, effective translation of cancer genomics or proteomics to clinical practice requires progress in analytics, which in turn, requires close cooperation between bioinformaticians, mathematicians, and oncologists.

A limitation of our study was that it was based on secondary data analysis. Therefore, there was a lack of information about important perioperative data such as liver enzyme profiles, antiviral drug use, and postoperative progression of underlying liver disease. These are believed to influence tumor recurrence and de novo malignancy. Additionally, data on tumor size and the number of tumors in each patient are omitted from the TCGA database, although these two parameters are considered important prognostic factors. Unlike miR-100, no target candidate gene of miR-148a could be identified despite its prognostic significance. It is known that miRNAs can inhibit target genes via degradation of their target mRNAs or by inhibition of translation. If a specific miRNA functions through suppressing translation but not through degrading mRNA, it is difficult to find a target just from miRNA and mRNA expression profiles and their correlation analysis. In the TCGA database, protein expression data are limited compared with miRNA and mRNA data. Therefore, experimental verification is required in the future.

**Conclusions**

This study showed that miR-100-5p and miR-148a-3p were underexpressed to a great extent in HCC tissues with gross VI than that in HCC tissues without gross VI. In addition, their aberrant expressions were significantly associated with poor survival of patients after hepatic resection for HCC.

**Author Contributions**

**Data curation:** Sung Kyu Song, Seung-Keun Park, Chul-Woon Chung, Yongkeun Park.

**Formal analysis:** Sung Kyu Song, Woon Yong Jung, Seung-Keun Park, Chul-Woon Chung, Yongkeun Park.

**Funding acquisition:** Sung Kyu Song, Yongkeun Park.

**Investigation:** Sung Kyu Song, Woon Yong Jung, Yongkeun Park.

**Supervision:** Chul-Woon Chung, Yongkeun Park.

**Validation:** Yongkeun Park.

**Visualization:** Yongkeun Park.

**Writing – original draft:** Sung Kyu Song, Woon Yong Jung, Yongkeun Park.

**Writing – review & editing:** Seung-Keun Park, Chul-Woon Chung, Yongkeun Park.
References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA: a cancer journal for clinicians. 2015; 65:87–108.

2. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. Hepatology 2011; 53:1020–1022. https://doi.org/10.1002/hep.24199 PMID: 21374666

3. Zhang G, Li R, Zhao X, Meng S, Ye J, Zhao L. Validation of the American Joint Committee on Cancer eighth edition staging system in patients undergoing hepatectomy for hepatocellular carcinoma: a US population-based study. J Surg Res. 2017; 222:55–68. https://doi.org/10.1016/j.jss.2017.09.044 PMID: 29273376

4. Park YK, Song SK, Kim BW, Park SK, Chung CW, Wang HJ. Prognostic significance of microvascular invasion in tumor stage for hepatocellular carcinoma. World J Surg Oncol. 2017; 15:225. https://doi.org/10.1186/s12957-017-1292-3 PMID: 29258907

5. Chen H, Miao J, Li H, Wang C, Li J, Zhu Y, et al. Expression and prognostic significance of p21-activated kinase 6 in hepatocellular carcinoma. J Surg Res. 2014; 189:81–88. https://doi.org/10.1016/j.jss.2014.01.049 PMID: 24576777

6. Sarkar S, Horn G, Moulton K, Oza A, Byler S, Kokolus S, et al. Cancer development, progression, and therapy: an epigenetic overview. Int J Mol Sci. 2013; 14:21087–21113. https://doi.org/10.3390/ijms141021087 PMID: 24152442

7. Wall NR, Shi Y. Small RNA: can RNA interference be exploited for therapy? Lancet. 2003; 362:1401–1403. https://doi.org/10.1016/S0140-6736(03)14637-5 PMID: 14585643

8. Esteller M. Non-coding RNAs in human disease. Nature reviews Genetics. 2011; 12:861–874. https://doi.org/10.1038/nrg3074 PMID: 22094949

9. Liu Y. MicroRNAs and PIWI-interacting RNAs in oncology. Oncol Lett. 2016; 12:2289–2292. https://doi.org/10.3892/ol.2016.4996 PMID: 27696791

10. Li PF, Chen SC, Xia T, Jiang XM, Shao YF, Xiao BX, et al. Non-coding RNAs and gastric cancer. World J Gastroenterol. 2014; 20:5411–5419. https://doi.org/10.3748/wjg.v20.i18.5411 PMID: 24833871

11. Kojima K, Takata A, Vadnais C, Otsuka M, Yoshikawa T, Akanuma M, et al. MicroRNA122 is a key regulator of alpha-fetoprotein expression and influences the aggressiveness of hepatocellular carcinoma. Nature communications. 2011; 2:338. https://doi.org/10.1038/ncomms1345 PMID: 21654638

12. Zheng F, Liao YJ, Cai MY, Liu YH, Liu TH, Chen SP, et al. The putative tumour suppressor microRNA-124 modulates hepatocellular carcinoma cell aggressiveness by repressing ROCK2 and EZH2. Gut. 2012; 61:278–289. https://doi.org/10.1136/gut.2011.239145 PMID: 21672940

13. Wang W, Lin H, Zhou L, Zhu Q, Gao S, Xie H, et al. microRNA-30a-3p inhibits tumor proliferation, invasiveness and metastasis and is downregulated in hepatocellular carcinoma. Eur J Surg Oncol. 2014; 40:1586–1594. https://doi.org/10.1016/j.ejso.2013.11.008 PMID: 24203072

14. Song J, Wang Q, Luo Y, Yuan P, Tang C, Hui Y, et al. miR-34c-3p inhibits cell proliferation, migration and invasion of hepatocellular carcinoma by targeting MARK5. Int J Clin Exp Pathol. 2015; 8:12728–12737. PMID: 26722462

15. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnevd E, Evangelista C, et al. NCBI GEO: mining tens of millions of expression profiles—data base and tools update. Nucleic Acids Res. 2007; 35:D760–765. https://doi.org/10.1093/nar/gkl887 PMID: 17099226

16. Hong F, Breitling R, McEntee CW, Wittner BS, Nemhauser JL, Chory J. RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. Bioinformatics. 2006; 22:2825–2827. https://doi.org/10.1093/bioinformatics/btl476 PMID: 16982708

17. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. https://doi.org/10.1186/s13059-014-0550-8 PMID: 25516281

18. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016. at http://www.R-project.org/.

19. Dweep H, Gretz N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nature methods. 2015; 12:697. https://doi.org/10.1038/nmeth.3485 PMID: 26226356

20. Baker S, Ali I, Silins I, Pyyysalo S, Guo Y, Hogberg J, et al. Cancer Hallmarks Analytics Tool (CHAT): A text mining approach to organize and evaluate scientific literature on cancer. Bioinformatics. 2017; 33:3973–3981 https://doi.org/10.1093/bioinformatics/btx454 PMID: 29036271

21. Chan SH, Wang LH. Regulation of cancer metastasis by microRNAs. J Biomed Sci 2015; 22:9. https://doi.org/10.1186/s12929-015-0113-7 PMID: 25614041
22. Liu S, Guo W, Shi J, Li N, Yu X, Xue J, et al. MicroRNA-135a contributes to the development of portal vein tumor thrombus by promoting metastasis in hepatocellular carcinoma. J Hepatol. 2012; 56:389–396. https://doi.org/10.1016/j.jhep.2011.08.008 PMID: 21888875

23. Lee YH, Hsu CY, Huang YH, Hsia CY, Chou YY, Su CW, et al. Vascular invasion in hepatocellular carcinoma: prevalence, determinants and prognostic impact. J Clin Gastroenterol. 2014; 48:734–741. https://doi.org/10.1097/MCG.0b013e3182a8a254 PMID: 24100755

24. Ruzzene A, Capra F, Pachera S, Iacono C, Piccirillo G, Lunardi M, et al. Is liver resection justified in advanced hepatocellular carcinoma? Results of an observational study in 464 patients. J Gastrointest Surg. 2009; 13:1313–1320. https://doi.org/10.1007/s11605-009-0903-x PMID: 19418103

25. Bruix J, Cheng AL, Meinhardt G, Nakajima K, De Sanctis Y, Llovet J. Prognostic factors and predictors of sorafenib benefit in patients with hepatocellular carcinoma: Analysis of two phase III studies. J Hepato.2017; 67:999–1008.

26. Glantzounis GK, Paliouras A, Stylianidi MC, Milionis H, Tzimas P, Roukos D, et al. The role of liver resection in the management of intermediate and advanced stage hepatocellular carcinoma. A systematic review. Eur J Surg Oncol. 2018; 44:195–208. https://doi.org/10.1016/j.ejso.2017.11.022 PMID: 29258719

27. Filippowicz W, Jaskiewicz L, Kolb FA, Pillai RS. Post-transcriptional gene silencing by siRNAs and miRNAs. Curr Opin Struct Biol. 2005; 15:331–341. https://doi.org/10.1016/j.sbi.2005.05.006 PMID: 15925505

28. Wang J, Yu M, Guan S, Zhang G, Wang J, Cheng Y. Prognostic significance of microRNA-100 in solid tumors: an updated meta-analysis. Oncotargets Ther. 2017; 10:493–502. https://doi.org/10.2147/OTT.S122774 PMID: 28176958

29. Sun X, Liu X, Wang Y, Yang S, Chen Y, Yuan T. miR-100 inhibits the migration and invasion of nasopharyngeal carcinoma by targeting IGF1R. Oncol Lett. 2018; 15:8333–8338. https://doi.org/10.3892/ol.2018.8420 PMID: 29805566

30. Zheng YS, Zhang H, Zeng CW, et al. MiR-100 regulates cell differentiation and survival by targeting RBSP3, a phosphatase-like tumor suppressor in acute myeloid leukemia. Oncogene. 2012; 31:80–92. https://doi.org/10.1038/onc.2011.208 PMID: 21643017

31. Tang Y, Zhou J, Hooi SC, Jiang YM, Lu GD. Fatty acid activation in carcinogenesis and cancer development: Essential roles of long-chain acyl-CoA synthetases. Oncol Lett. 2018; 16:1390–1396. https://doi.org/10.3892/ol.2018.8843 PMID: 30008815

32. Lin Z, Cai YJ, Chen RC, Chen BC, Zhao L, Xu SH, et al. A microRNA expression profile for vascular invasion can predict overall survival in hepatocellular carcinoma. Clin Chim Acta. 2017; 469:171–179. https://doi.org/10.1016/j.cca.2017.03.026 PMID: 28365450

33. Huang C, Zhu XD, Ji Y, Ding GY, Shi GM, et al. Microvascular invasion has limited clinical values in hepatocellular carcinoma patients at Barcelona Clinic Liver Cancer (BCLC) stages 0 or B. BMC Cancer. 2017; 17:58. https://doi.org/10.1186/s12885-017-3050-x PMID: 28095820

34. Shen J, Wen J, Li C, Wen T, Yan L, Li B, et al. The prognostic value of microvascular invasion in early-intermediate stage hepatocellular carcinoma: a propensity score matching analysis. BMC Cancer. 2018; 18:278. https://doi.org/10.1186/s12885-018-4196-x PMID: 29530006