Integrated analysis of miRNA, gene, and pathway regulatory networks in hepatic cancer stem cells

Min Ding\textsuperscript{1†}, Jiang Li\textsuperscript{1†}, Yong Yu\textsuperscript{2†}, Hui Liu\textsuperscript{1}, Zi Yan\textsuperscript{1}, Jinghan Wang\textsuperscript{2} and Qijun Qian\textsuperscript{1*}

Abstract

Background: Hepatocellular carcinoma (HCC) is one of the most common malignancies worldwide. HCC has a poor prognosis associated with tumor recurrence and drug resistance, which has been attributed to the existence of hepatic cancer stem cells (HCSCs). However, the characteristics and regulatory mechanisms of HCSCs remain unclear. We therefore established a novel system to enrich HCSCs and we demonstrate that these HCSCs exhibit cancer stem cell properties.

Methods: We used miRNA and mRNA high-throughput sequencing data sets to determine molecular signatures and regulatory mechanisms in HCSCs. Paired miRNA and gene deep sequencing data in HCSCs versus HCC cells were used to identify candidate biomarkers of HCSCs. Using network analysis, we studied the relationship between miRNA and gene biomarkers, and KEGG pathway enrichment analysis was performed to study the function of candidate biomarkers.

Results: We identified 9 up- and 9 down-regulated miRNAs and 115 up- and 402 down-regulated genes in HCSCs compared with HCC cells. A miRNA-gene network was constructed using 651 miRNA–gene interactions (between 7 up-regulated miRNAs and 274 down-regulated genes), and 103 miRNA–gene interactions (between 9 down-regulated miRNAs and 62 up-regulated genes). Pathway enrichment analysis identified five tumor invasion- and metastasis-related pathways and MAPK signaling associated with HCSCs. We further discovered two novel pathways that likely play a role in the regulation of HCSCs.

Conclusions: We identified a molecular expression signature and pathway regulatory mechanisms in HCSCs with potential diagnostic and therapeutic value.

Keywords: Hepatic cancer stem cells, Hepatic cancer cells, miRNA, mRNA, Pathway analysis, Regulatory network

Background

Hepatocellular carcinoma (HCC) is one of the most common malignancies that accounts for 70–85% of liver cancers worldwide [1]. Although significant progress has been made in recent years regarding the treatment options for HCC, poor prognosis remains a problem because of late diagnosis, recurrence, and drug resistance [2]. While surgical intervention—the main treatment option for HCC—is effective in patients diagnosed at an early stage [3], the treatment of advanced liver cancer is more difficult and prognosis remains poor because of drug resistance [4], making recurrence almost inevitable [5]. Cancer stem cells (CSCs), which are critical for the initial growth and maintenance of the tumor, have been identified in liver cancers [6–8]. Recently, CSCs have been associated with tumor recurrence and drug resistance in HCC [9, 10]. CSCs are potential targets for HCC diagnosis and treatment—it is therefore crucial that we study the regulatory mechanisms of CSCs.
Several biological markers of hepatic cancer stem cells (HCSCs), including CD133 [10], CD90 [11], and EpCAM [12] have been identified. However, the characteristics and regulatory mechanisms of HCSCs remain unclear. We therefore established a novel system to enrich HCSCs and previously reported that these cells have CSC characteristics [13].

MicroRNAs (miRNAs), a class of small non-coding RNAs, have been shown to play an important role in a variety of biological processes. Abnormal expression of miRNAs may impact the expression of hundreds of genes. Recently, with the development of microarray and high-throughput sequencing technology, miRNA–mRNA interactions in cancers and other biological processes have been extensively studied [14–16]. However, genome-wide miRNA–mRNA interactions in HCSCs remain largely unknown.

In this study, we performed high-throughput sequencing of HCSC small RNA and mRNA, and integrated miRNA and mRNA data to identify biomarkers of HCSCs and to unravel HCSC regulatory networks. Our network analysis approach is summarized in Fig. 1.

**Methods**

**Data**

We used two human hepatoma cell lines (Hep3B and Huh7) from the American Type Culture Collection (ATCC) to culture stem-like cancer cells. We previously demonstrated that these cells have enhanced stem cell properties, drug resistance, properties of epithelial mesenchymal transition, and enhanced tumor-initiating capabilities [13]. Here, we analyzed the regulatory mechanism of these two types of HCSC using miRNA and RNA sequencing data. MiRNA and gene expression data of hepatic cancer stem cells and cancer cells were sequenced using the Illumina Genome Analyzer, to a depth of 30-fold coverage. Details of the data sets are shown in Table 1.

Small RNA and mRNA libraries were sequenced on an Illumina Genome Analyzer II (Illumina, San Diego, CA, USA) according to the manufacturer’s instructions. Raw RNAseq reads were filtered for adapters and/or low-quality reads, followed by alignment to the genome. Details of the data sets are shown in Table 1.

| Data type | Cell line | Total reads | Normalized |
|-----------|-----------|-------------|------------|
| miRNA     | Hep3B-C   | 15440438    | TPM        |
|           | Huh7-C    | 10714837    |            |
|           | Hep3B     | 16782602    |            |
|           | Huh7      | 10622276    |            |
| RNA       | Hep3B-C   | 5762058     | RPKM       |
|           | Huh7-C    | 5886142     |            |
|           | Hep3B     | 5998659     |            |
|           | Huh7      | 6116344     |            |

Paired miRNA and gene deep sequencing datasets from the same sample of two hepatic cancer stem cells and two hepatic cancer cells were used. Detailed information about these datasets is provided in this table.
human genome (NCBI Build 36.1) using NextGENe® software (Softgenetics, State College, PA, USA). Reads mapping to individual transcripts were counted digitally, and the expression levels for each gene were normalized using reads per kilobase of exon model per million mapped reads (RPKM) [17]. For small RNA analysis, after filtering contaminant and redundant reads, and reads smaller than 18 nt, clean reads were mapped to the human genome (NCBI Build 36.1) using SOAP [18]. Reads were then mapped to Ribosomal RNA (rRNA), Small cytoplasmic RNA (scRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), and transfer RNA (tRNA) from GenBank. The remaining reads were searched against miRBase (release 17) [19]. Read counts of the annotated miRNAs were normalized using transcripts per million (TPM) [20].

**Differential expression analysis**

Genes and miRNAs differentially expressed between hepatic cancer stem cells and hepatic cancer cells were identified by calculating fold change values and using an established statistical method based on the Poisson distribution [21–23] to calculate p-values. The Benjamini–Hochberg FDR method [24] was used to adjust p-values for multiple testing. MiRNAs and genes with an absolute log2 fold-change (expression of cancer stem cells/expression of cancer cells) ≥1 and an FDR ≤0.01 were considered statistically significant.

**miRNA target prediction**

Using the results from differential gene and miRNA expression (cancer stem cells vs. cancer cells), gene–miRNAs interactions were predicted using seven miRNA target computational prediction methods: MicroCosm (Version 5) [25], microT (version 3) [26], miRanda (August 2010 available at [http://www.microrna.org/microrna/home.do] [27], miRDB (version 4.0) [28], PicTar (four-way) [29], PITA (version 6) [30], and TargetScan (version 5.2) (with total context score ≥−0.3) [31]. Except for TargetScan, we used default cut-off values. Interactions that occurred in at least two of these sources were considered for downstream analyses.

**Pathway enrichment analysis**

Pathway enrichment analysis was employed to investigate the regulatory mechanisms of significantly differentially expressed miRNAs. KEGG and DAVID Bioinformatics Resources 6.7 databases were used for pathway enrichment analyses. The enriched pathways were defined by their enrichment of significantly differentially expressed miRNA target genes. For DAVID functional annotation, the Fisher’s exact test was used to calculate statistical significance (p values) of enriched annotation terms, where a smaller p value implies enrichment, and a p-value ≤0.05 was deemed significant.

**Results**

**Identification of candidate biomarkers in hepatic cancer stem cells**

Differential expression analysis was used to identify candidate biomarkers in hepatic cancer stem cells. Using deep sequencing, miRNAs and genes differentially expressed between stem cells (Hep3B-C, Huh7-C) and the paired cancer cells from which they were derived (Hep3B, Huh7) were identified. Data normalization and differential expression analysis are described in the “Methods” section.

For miRNAs, 250 (59.7%) up- and 18 (4.3%) down-regulated miRNAs were identified in Hep3B-C cells, while 23 (5.4%) up- and 128 (30.2%) down-regulated miRNAs were identified in Huh7-C cells. Finally, we selected 9 up- and 9 down-regulated miRNAs, that were consistently altered in both stem cell lines (compared with their corresponding cancer cells), as candidate miRNA biomarkers of HCSCs (Fig. 2a). Using the same two paired samples, differential gene expression analysis was performed, which produced 1928 (13.0%) up- and 4264 (28.8%) down-regulated genes in Hep3B-C (vs. Hep3B) cells and 1933 (13.1%) up- and 1935 (13.1%) down-regulated miRNAs in Huh7-C (vs. Huh-7) cells. As shown in Fig. 2b, 115 and 402 genes were consistently up- and down-regulated, respectively in the two stem cell lines; these miRNAs and genes were used for downstream analyses (Additional file 1: Table S1, Additional file 2: Table S2).

**miRNA-gene regulatory network analysis**

To investigate the function of differentially expressed miRNAs, miRNA target genes were identified and miRNA target networks were constructed. First, seven miRNA target computational prediction methods were used to predict potential miRNA target genes, while considering all human genes as potential targets. This yielded 54,933 miRNA–target interactions between 18 differentially expressed miRNAs and 15,058 genes. To verify the miRNA–gene regulatory relationship in HCSCs, differential gene expression was considered in the context of miRNA–mRNA interaction whereby 1,460 miRNA–gene interactions between 16 and 398 differentially expressed miRNAs and differentially expressed genes were selected (Table 2). Considering the mechanism of miRNA-mediated mRNA down-regulation, 651 miRNA–gene interactions between 7 up-regulated miRNAs and 274 down-regulated genes, and 103 miRNA–gene interactions between 9 down-regulated miRNAs and 62 up-regulated genes were selected (this workflow is summarized...
Fig. 2 Differentially expressed miRNAs and genes. a, b illustrate the selection of differentially expressed miRNAs and genes, respectively between stem and cancer cells. Two paired stem- and cancer-cell samples (Hep3B-C, Hep3B) and (Huh7-C, Huh7) were used here; miRNAs and genes consistently up- or down-regulated in both sample pairs were selected.

Table 2 Details of target genes regulated by differentially expressed miRNAs

| miRNA        | Up/down-regulation of miRNA (HCSC/HCC) | Number of target genes | Total predicted targets | Differentially expressed targets | Up-regulated targets (HCSC/HCC) | Down-regulated targets (HCSC/HCC) | Candidate targets |
|--------------|----------------------------------------|------------------------|-------------------------|----------------------------------|---------------------------------|------------------------------------|-------------------|
| hsa-miR-100  | Down                                   | 574                    | 22                      | 4                                | 18                              | 4                                  | 4                 |
| hsa-miR-210  | Down                                   | 1180                   | 33                      | 5                                | 28                              | 5                                  | 5                 |
| hsa-miR-29c  | Down                                   | 3533                   | 99                      | 26                               | 73                              | 26                                 | 26                |
| hsa-miR-181c | Down                                   | 5492                   | 141                     | 15                               | 126                             | 15                                 | 15                |
| hsa-miR-22*  | Down                                   | 2343                   | 55                      | 4                                | 51                              | 4                                  | 4                 |
| hsa-miR-15b* | Down                                   | 1641                   | 45                      | 4                                | 41                              | 4                                  | 4                 |
| hsa-miR-199a-3p | Down                              | 3707                   | 99                      | 11                               | 88                              | 11                                 | 11                |
| hsa-miR-199b-3p | Down                              | 3456                   | 96                      | 10                               | 86                              | 10                                 | 10                |
| hsa-miR-149  | Down                                   | 5100                   | 116                     | 24                               | 92                              | 24                                 | 24                |
| hsa-miR-378d | Up                                     | 7                      | 0                       | 0                                | 0                               | 0                                  | 0                 |
| hsa-miR-450b-5p | Up                                 | 5865                   | 146                     | 8                                | 138                             | 138                                | 138               |
| hsa-miR-338-5p | Up                                 | 5240                   | 156                     | 15                               | 141                             | 141                                | 141               |
| hsa-miR-760  | Up                                     | 4249                   | 113                     | 32                               | 81                              | 81                                 | 81                |
| hsa-miR-378  | Up                                     | 3486                   | 92                      | 20                               | 72                              | 72                                 | 72                |
| hsa-miR-215  | Up                                     | 2878                   | 81                      | 11                               | 70                              | 70                                 | 70                |
| hsa-miR-375  | Up                                     | 3312                   | 95                      | 10                               | 85                              | 85                                 | 85                |
| hsa-miR-1269b | Up                                   | 12                     | 0                       | 0                                | 0                               | 0                                  | 0                 |
| hsa-miR-1269 | Up                                     | 2877                   | 71                      | 7                                | 64                              | 64                                 | 64                |

Seven miRNA target computational prediction sources were used to identify the interactions of miRNAs and genes. The miRNA-target interactions which occurred in at least two of these sources were considered. For each miRNA, the expression level of the target genes in HCSC and HCC was considered. The detail result is listed in this table.
in Fig. 3). These interactions were then used as candidate miRNA–gene interactions to construct a miRNA–gene regulatory network (Fig. 4).

The network is an objective representation of all regulatory relationships between miRNAs and genes in HCSCs. Fig. 4a, b represent the regulatory networks of miRNAs up- and down-regulated in HCSCs compared with cancer cells, respectively. From Fig. 4, it is clear that up-regulated miRNAs interact with relatively more target genes compared with down-regulated miRNAs. Hsa-miR-338-5p and hsa-miR-450b-5p, which have a large number of target genes, were the most important components in the up-regulated miRNA regulatory network.

In the down-regulated miRNA regulatory network, four target genes were exclusively regulated by has-miR-15b* and the down-regulated miRNA regulatory network could be divided into two sub networks. To identify the significance of this feature, the Fisher’s exact test was used here [32]. However there was no significant difference (p = 0.1149), suggesting that this feature may be the result of random chance. The importance of this feature requires further validation.

Using a ‘novel out degree’ (NOD), defined by Zhang et al. [16], we measured the independent regulatory power of individual miRNAs. The NOD value represents the number of genes uniquely regulated by one
specific miRNA. The distribution of miRNA NOD values is shown in Fig. 5. Consistent with the above result, there were more up-regulated miRNAs with large NOD values. The up-regulated miRNAs had larger independent regulatory power.

**Analysis of stem-cell-associated miRNA pathways**

Investigating miRNA-regulated pathways should help uncover the underlying mechanisms of stem cells. Using KEGG pathways, we performed enrichment analysis to identify stem-cell-related pathways. Significantly enriched KEGG pathways (p value <0.05) are shown in Table 3. To investigate the relevance of these pathways to cancer, we searched PubMed for published papers describing the roles of these pathways; published cancer-associated functions are listed in Table 4. Five pathways, enriched with up-regulated miRNA targets, reportedly participate in tumor invasion and metastasis; these pathways are *Cytokine-cytokine receptor interaction, Regulation of actin cytoskeleton, Focal adhesion, Axon guidance,* and *Adipocytokine*...
signaling pathway. Another important cancer-related pathway, MAPK signaling pathway, was also enriched with up-regulated miRNA targets. To study the function of up-regulated miRNAs in the context of these cancer-related pathways, we constructed miRNA-gene-pathway regulatory networks. Fig. 6a shows the network for the five pathways related to tumor invasion and metastasis. Fig. 6b shows the network for the MAPK signaling pathway, which is involved in HCC growth and survival. In these six cancer-related pathways, four genes (IL8, PRLR, EFNA1, and CHP2) were uniquely regulated by one specific miRNA each (hsa-miR-338-5p_IL8, hsa-miR-338-5p_PRLR, hsa-miR-760_EFNA1, hsa-miR-450b-5p_CHP2).

However, two pathways, Terpenoid backbone biosynthesis (hsa00900) and Synthesis and degradation of ketone bodies (hsa00072), have not previously been related to cancer. These two pathways were enriched with the target genes of down-regulated miRNAs (hsa-miR-338-5p, and hsa-miR-450b-5p) that down-regulate hundreds of genes. The number of targets for hsa-miR-338-5p and hsa-miR-450b-5p were 141 and 138, respectively. The association of these two miRNAs with HCC and HCSCs has not previously been reported. Our results were obtained by computational methods only and further experimental validation is therefore required.

**Discussion**

To better understand the characteristics of HCSCs established in our lab, we identified differentially expressed miRNAs and genes in HCSCs compared with hepatic cancer cells based on high-throughput sequencing datasets (9 up-regulated miRNAs; 9 down-regulated miRNAs; 115 up-regulated genes; 402 down-regulated genes). The relationship between these miRNAs and genes, and their pathway-level involvement were analyzed. With the aim of investigating regulatory mechanisms in HCSCs, we constructed regulatory networks based on candidate biomarkers and enriched pathways in HCSCs. We found a complex relationship between differentially expressed miRNAs and genes in HCSCs. MiRNA–gene regulatory networks were constructed, and up-regulated miRNAs were found to regulate more target genes and have larger NOD values compared with down-regulated miRNAs. This implies that up-regulated miRNAs might be more important in the regulation of HCSCs. We identified two up-regulated miRNAs (hsa-miR-338-5p, and hsa-miR-450b-5p) that down-regulate hundreds of genes. The number of targets for hsa-miR-338-5p and hsa-miR-450b-5p were 141 and 138, respectively. The association of these two miRNAs with HCC and HCSCs has not previously been reported. Our results were obtained by computational methods only and further experimental validation is therefore required.

### Table 3  Significantly enriched KEGG pathways (p-value <0.05)

| miRNA   | Term_ID | Term_name                        | Gene_count | Genes                                | P-value  |
|---------|---------|----------------------------------|------------|--------------------------------------|----------|
| hsa-miR-29c | hsa00900 | Terpenoid backbone biosynthesis 3 | HMGR, HMGCS1, HMGCS2 | 0.00023 |
| hsa-miR-29c | hsa00072 | Synthesis and degradation of ketone bodies 2 | HMGCS1, HMGCS2 | 0.014 |
| hsa-miR-181c | hsa00900 | Terpenoid backbone biosynthesis 2 | HMGR, HMGCS1 | 0.012 |
| hsa-miR-338-5p | hsa05200 | Pathways in cancer 10 | RARB, FZD4, ITGA2, IL8, FGF12, DAPK1, EGLN3, PDGFR, LAMC2, FGF11 | 0.0036 |
| hsa-miR-338-5p | hsa04510 | Focal adhesion 7 | CAV1, CAV2, ITGA2, LAMC2, MYLK3, PDGFR, AVAV3 | 0.012 |
| hsa-miR-338-5p | hsa04810 | Regulation of actin cytoskeleton 7 | VAV3, MYLK3, PDGFR, ITGA2, FGF11, TIA2, FGF12 | 0.016 |
| hsa-miR-338-5p | hsa04920 | Adipocytokine signaling pathway 4 | JAK2, AC5L4, PPARGC1A, PRKAB2 | 0.026 |
| hsa-miR-338-5p | hsa04060 | Cytokine-cytokine receptor interaction 7 | CXCL12, GHR, INHHB, IL8, PDGFR, PRLR, TNFSF4 | 0.038 |
| hsa-miR-378 | hsa04510 | Focal adhesion 5 | COL6A2, ITGA2, MYLK3, PDGFR, AVAV3 | 0.0058 |
| hsa-miR-378 | hsa04810 | Regulation of actin cytoskeleton 5 | VAV3, MYLK3, PDGFR, ITGA2, FGF12 | 0.0073 |
| hsa-miR-378 | hsa04060 | Cytokine-cytokine receptor interaction 5 | CXCL12, EPO, INHHB, PDGFR, TGF2B | 0.014 |
| hsa-miR-378 | hsa05200 | Pathways in cancer 5 | FGF12, FZD4, ITGA2, PDGFR, TGF2B | 0.03 |
| hsa-miR-378 | hsa05210 | Colorectal cancer 3 | FZD4, PDGFR, TGF2B | 0.038 |
| hsa-miR-378 | hsa04600 | Cytokine-cytokine receptor interaction 6 | CXCL12, EPO, INHHB, PDGFR, TGF2B, TNFSF4 | 0.011 |
| hsa-miR-378 | hsa05200 | Pathways in cancer 6 | EMLN3, FGF11, FZD4, PDGFR, TGF2B, RALB | 0.027 |
| hsa-miR-378 | hsa04360 | Axon guidance 4 | CXCL12, EFNA1, NTX4, SEMA7A | 0.03 |
| hsa-miR-378 | hsa04510 | Pathways in cancer 5 | FGF11, FGF12, ITGA2, LAMC2, RALB | 0.014 |
| hsa-miR-450b-5p | hsa04060 | Cytokine-cytokine receptor interaction 7 | IL1R1, EPO, CXCL12, TNFSF4, TGF2B, PDGFR, GHR | 0.029 |
| hsa-miR-450b-5p | hsa04010 | MAPK signaling pathway 7 | IL1R1, DUSP110, TGF2B, PDGFR, CHOP1, MAP3K8, FGF12 | 0.052 |
| hsa-miR-450b-5p | hsa04810 | Regulation of actin cytoskeleton 6 | TIM2, FGF12, IGF2, ITGA2, PDGFR, AVAV3 | 0.043 |

KEGG database and DAVAD Bioinformatics Resources 6.7 were used for pathway enrichment analyses of genes regulated by identified HCSC miRNA biomarkers. Significant enriched pathways (P-value <0.05) are listed in this table.
Table 4: Published cancer-associated functions of enriched pathways

| Term_ID | miRNAs | Term_name                                      | Relevance to cancer                                                                                   |
|---------|--------|-----------------------------------------------|-------------------------------------------------------------------------------------------------------|
| hsa05200| hsa-miR-215, hsa-miR-338-5p, hsa-miR-378, hsa-miR-760 | Pathways in cancer                                                                                   |                                                                                                        |
| hsa04060| hsa-miR-338-5p, hsa-miR-378, hsa-miR-450b-5p, hsa-miR-760 | Cytokine–cytokine receptor interaction                                                               | Cytokines can control invasion and metastasis, and also function to inhibit tumor progression [40]   |
| hsa04810| hsa-miR-338-5p, hsa-miR-378, hsa-miR-450b-5p | Regulation of actin cytoskeleton                                                                   | Several studies revealed that molecules that link migratory signals to the actin cytoskeleton are upregulated in invasive and metastatic cancer cells [41] |
| hsa04510| hsa-miR-338-5p, hsa-miR-378 | Focal adhesion                                                                                       | Focal adhesion kinase, which plays an important role in tumor progression and metastasis, is overexpressed and activated in a variety of human cancers [42, 43] |
| hsa04010| hsa-miR-450b-5p | MAPK signaling pathway                           | The MAPK pathway plays an important role in HCC in that its activation is reportedly involved in HCC growth and survival [35] |
| hsa04360| hsa-miR-760 | Axon guidance                                     | The ligand/receptor pairs of axon guidance regulate tumor angiogenesis [44]                           |
| hsa04920| hsa-miR-338-5p | Adipocytokine signaling pathway                  | Adipocytokine signaling pathway has been demonstrated participate in breast cancer progression [45] |
| hsa05210| hsa-miR-378 | Colorectal cancer                                 |                                                                                                        |
| hsa09000| hsa-miR-181c, hsa-miR-29c | Terpenoid backbone biosynthesis                  |                                                                                                        |
| hsa00072| hsa-miR-29c | Synthesis and degradation of ketone bodies        |                                                                                                        |

We searched PubMed for published papers to explain the relevance of significantly enriched pathways to cancer. The published cancer-associated-functions of these pathways are listed in this table.

Fig. 6 miRNA–gene–pathway regulatory networks. **a** shows the network for five pathways related to tumor invasion and metastasis. **b** shows the network for the MAPK signalling pathway. **c** shows the network for two new pathways, which have not been previously associated with cancer. MiRNAs, genes, and pathways are represented by nodes (pink miRNAs; green genes; and blue pathways). Edges of dark color represent the relationship between genes and pathways; and edges of light color represent the relationship between miRNAs and genes. Nodes marked with a red asterisk refer to genes uniquely regulated by one specific miRNA.
Through functional analyses we uncovered important biological processes involved in the regulation of HCSCs. Ten KEGG pathways were enriched in HCSCs. Pathway-level text mining was used to evaluate the relevance of these enriched pathways in cancer. Interestingly, five of the pathways have previously been reported to be involved in tumor invasion and metastasis. Invasion and metastasis are the main causes of cancer deaths, and are complex multi-step processes [33, 34]. Pathway analysis results are supported by the fact that the HCSCs established in our lab have much stronger invasive capability than hepatic cancer cells. Four up-regulated miRNAs (hsa-miR-338-5p, hsa-miR-450b-5p, hsa-miR-378, and hsa-miR-760) were identified to take part in HCSC invasion-related pathways. Two genes, PDGFRA and CXCL12, were targeted by four invasional miRNAs and those that are involved in several invasion-related pathways, might play important roles in the regulation of this process (Fig. 6a). Another cancer-related pathway, MAPK signaling pathway, was also enriched with hsa-miR-450b-5p target genes. The MAPK signaling is a complex pathway involved in the regulation of a variety of growth and differentiation pathways and is reportedly involved in HCC growth and survival [35]. Four genes (IL8, PRLR, EFNA1, and CHP2) that were uniquely regulated by specific miRNAs, have been associated with HCC [36–39]. However, the regulatory mechanism of these genes in HCSCs requires further investigation. In addition to known cancer-related pathways, we also identified two novel HCSC-related pathways, Terpenoid backbone biosynthesis and Synthesis and degradation of ketone bodies. These two pathways were enriched with down-regulated miRNA target genes in HCSCs. Terpenoids are a large class of natural products consisting of isoprene units, while ketone bodies are produced by the liver from fatty acids and used peripherally as an energy source when glucose is not readily available. Their relevance to HCC and HCSCs has not been reported, and requires further validation.

We globally analyzed the molecular expression signature and regulatory mechanisms in HCSCs established in our lab. Although we have identified candidate molecular markers and important pathways in HCSCs, our results are preliminary and further experimental validation is required.

Conclusions

Using high-throughput sequencing data sets and bioinformatics analyses, we identified miRNA and mRNA signatures of HCSCs. Additionally, we combined miRNA, mRNA, and pathway analyses to study the regulatory mechanisms in HCSCs by constructing miRNA–mRNA and miRNA–mRNA–pathway regulatory networks. The molecular markers and pathways identified herein might be used as candidate biomarkers and drug targets for the diagnosis and treatment of hepatic cancer.

Additional files

Additional file 1: Table S1. Consistently differentially expressed miRNAs.

Additional file 2: Table S2. Consistently differentially expressed mRNAs.

Abbreviations

HCC: hepatocellular carcinoma; HCSCs: hepatic cancer stem cells; CSCs: cancer stem cells; KEGG: Kyoto Encyclopedia of Genes and Genomes; MAPK: mitogen-activated protein kinase; rRNA: ribosomal RNA; scRNA: small cytoplasmic RNA; snRNA: small nuclear RNA; snoRNA: small nucleolar RNA; tRNA: transfer RNA.

Authors’ contributions

MD designed the analysis pipeline, performed the statistical analysis, and drafted the manuscript. JL, JW, and ZY cultured HCSCs and HCCs for sequencing. HL and YY revised the manuscript. QQ conceived and coordinated the overall study and revised the manuscript. All authors read and approved the final manuscript.

Author details

Department of Viral and Gene Therapy, Eastern Hepatobiliary Surgery Hospital, The Second Military Medical University, Shanghai 200438, People’s Republic of China. The First Department of Biliary Surgery, Eastern Hepatobiliary Surgical Hospital, The Second Military Medical University, Shanghai 200438, People’s Republic of China.

Acknowledgements

This work was supported by the National Significant Science and Technology Special Projects of New Drugs Creation (No. 2013ZX10002-010-007, No. 2012ZX10002-014-005).

Accession number

miRNA-seq and miRNA-seq raw data are available at the Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) under accession numbers GSE70470 and GSE70537.

Compliance with ethical guidelines

Competing interests

The authors declare that they have no competing interest.

Received: 11 November 2014 Accepted: 17 July 2015

Published online: 11 August 2015

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D (2011) Global cancer statistics. CA Cancer J Clin 61(2):69–90. doi:10.3322/ caac.201077

2. Villanueva A, Minguet B, Forner A, Reig M, Llovet JM (2010) Hepatocellular carcinoma: novel molecular approaches for diagnosis, prognosis, and therapy. Annu Rev Med 61:317–328. doi:10.1146/annurev.med.080608.100623

3. Yang JD, Roberts LR (2010) Hepatocellular carcinoma: a global view. Nat Rev Gastroenterol Hepatol 7(8):448–458. doi:10.1038/nrgastro.2010.100

4. Yau T, Chan P, Epstein R, Poon RT (2008) Evolution of systemic therapy of advanced hepatocellular carcinoma. World J Gastroenterol WJG 14(42):6437–6441
5. Chun JM, Kwon HJ, Sohn J, Kim SG, Park JY, Bae HI et al (2011) Prognostic factors after early recurrence in patients who underwent curative resection for hepatocellular carcinoma. J Surg Oncol 103(2):148–151. doi:10.1002/jso.21786

6. Ma S, Chan KW, Hu L, Lee TK, WO JY, Ng IO et al (2007) Identification and characterization of tumorigenic liver cancer stem/progenitor cells. Gastroenterology 132(2):2542–2556. doi:10.1053/j.gastro.2007.04.025

7. Yang ZF, Ngai P, Ho DW, Yu WC, Ng MN, Lau CK et al (2008) Identification of local and circulating cancer stem cells in human liver cancer. Hepatology 47(3):919–928. doi:10.1002/hep.20282

8. Yin S, Li J, Hu C, Chen X, Yao M, Yan M et al (2007) CD133 positive hepatocellular carcinoma cells possess high capacity for tumorigenicity. Int J Cancer 120(7):1444–1450. doi:10.1002/ijc.22476

9. Guo Z, Li LQ, Jiang JH, Ou C, Zeng LX, Xiang BD (2014) Cancer stem cell markers correlate with early recurrence and survival in hepatocellular carcinoma. World J Gastroenterol 20(8):2098–2106. doi:10.3780/j.wjg.v20.i8.2098

10. Ma S, Lee TK, Zheng BJ, Chan KW, Guan XY (2008) CD133+ HCC cancer stem cells confer chemoresistance by preferential expression of the Akt/PKB survival pathway. Oncogene 27(12):1749–1758. doi:10.1038/sj.onc.1210811

11. Yang ZF, Ho DW, Ng MN, Lau CK, Yu WC, Ngai P et al (2008) Significance of CD90+ cancer stem cells in human liver cancer. Cancer Cell 12(3):153–166. doi:10.1016/j.ccc.2008.01.013

12. Yamashita T, Forquer M, Wang W, Kim YW, Ye Q, Jia H et al (2008) EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. Cancer Res 68(5):1451–1461. doi:10.1158/0008-5472.CAN-07-6013

13. Li J, Yu Y, Wang J, Yan Z, Liu H, Wang Y et al (2015) Establishment of a novel system for the culture and expansion of hepatic stem-like cancer cells. Cancer Lett 360(2):177–186. doi:10.1016/j.canlet.2015.02.006

14. Grigoriey VA, Kurian SM, Hart T, Nakorchevsky AA, Chen C, Campbell D et al (2011) MicroRNA regulation of molecular networks mapped by global microRNA, mRNA, and protein expression in activated T lymphocytes. J Immunol 187(5):2233–2243. doi:10.4049/jimmunol.1101233

15. Szeto CY, Lin CH, Cho SC, Yip TT, Ngan RK, Tsao GS et al (2014) Integrated analysis of genome-wide microRNA and mRNA transcriptome sequencing characterizes sequence variants and mRNA-microRNA regulatory networks in nasopharyngeal carcinoma model systems. FEBS Open Bio 4(12):140–140. doi:10.1002/fob.2014.01.004

16. Zhang W, Zang J, Jing X, Sun Z, Yan W, Yang D et al (2014) Identification of candidate miRNA biomarkers from miRNA regulatory network with application to prostate cancer. J Transl Med 12:66. doi:10.1186/1479-5876-12-66

17. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5(7):621–628. doi:10.1038/nmeth.1226

18. Li R, Li Y, Kristiansen K, Wang J (2008) SOAP: short oligonucleotide alignment program. Bioinformatics 24(5):713–714. doi:10.1093/bioinformatics/btn025

19. Kozomara A, Griffiths-Jones S (2011) miBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39(Database issue):D152–D157. doi:10.1093/nar/gkr1027

20. Hou J, Lin L, Zhou W, Wang Z, Ding G, Dong Q et al (2011) Identification of miRNomes in human liver and hepatocellular carcinoma reveals miR-199a/b-3p as therapeutic target for hepatocellular carcinoma. Cancer Cell 20(2):232–243. doi:10.1016/j.ccc.2011.01.001

21. Barrera-Figueroa BE, Gao L, Diop NN, Wu Z, Ehlers JD, Roberts PA et al (2011) Identification and comparative analysis of drought-associated microRNAs in two cowpea genotypes. BMC Plant Biol 11:127. doi:10.1186/1471-2229-11-127

22. Gonzalez-Ballester D, Casero D, Cokus S, Pellegrini M, Merchant SS, Grossman AR (2010) RNA-seq analysis of sulfur-deprived Chlamydomonas cells reveals aspects of acclimation critical for cell survival. Plant Cell 22(6):2058–2084. doi:10.1105/tpc.109.071167

23. Audic S, Claverie JM (1997) The significance of digital gene expression profiles. Genome Res 7(9):896–995