doi:10.4149/neo_2018_161215N642

Integrated network analysis to identify the key genes, transcription factors, and microRNAs involved in hepatocellular carcinoma

S.Q. SHI[1], J.-J. KE[1], Q.S. XU[1], W.Q. WU[1], Y.Y. WAN[2,*]

[1]Department of Gastroenterology, Zhejiang Provincial People's Hospital, Hangzhou, 310014, China; [2]Departments of Gastrointestinal Surgery, Zhejiang Provincial People's Hospital, Hangzhou, 310014, China

**Correspondence: yuanyuwany06@aliyun.com*

Received December 15, 2016 / Accepted May 29, 2017

HCC (hepatocellular carcinoma), which can be induced by cirrhosis and viral hepatitis infection, is the most frequent form of liver cancer. This study is performed to investigate the mechanisms of HCC. GSE57957 was obtained from Gene Expression Omnibus database, including 39 HCC samples and 39 adjacent non-tumorous samples. The DEGs (differentially expressed genes) were screened using the limma package in R, and then were conducted with enrichment analysis using "BioCloud" platform. Using STRING database, WebGestalt tool, as well as ITFP and TRANSFAC databases, PPI (protein-protein interaction) pairs, miRNA (microRNA)-target pairs, and TF (transcription factor)-target pairs separately were predicted. Followed by integrated network was constructed by Cytoscape software and module analysis was performed using the MCODE plugin of Cytoscape software. There were 518 DEGs identified from the HCC samples, among which 17 up-regulated genes (including *MCM2*, *MCM6*, and *CDC20*) and 5 down-regulated genes could also function as TFs. In the integrated network for the down-regulated genes, *FOS* and *ESR1* had higher degrees, and both of them were targeted by *miR-221* and *miR-222*. Additionally, *MCM2* had interaction with *MCM6* in the up-regulated module with the highest score. *MCM2*, *MCM6*, *CDC20*, *FOS*, *ESR1*, *miR-221* and *miR-222* might affect the pathogenesis of HCC.

Key words: carcinoma, hepatocellular, gene expression, transcription factors, microRNAs, gene regulatory networks

HCC (hepatocellular carcinoma, also named malignant hepatoma), is the most frequent form of liver cancer [1]. Most cases of HCC are induced by cirrhosis and viral hepatitis infection (hepatitis B or hepatitis C) [2–4]. Only 10–20% cases of HCC can be cured by complete resection, thus the usual outcome of HCC is poor [5]. As one of the most common cancers around the world, HCC usually occurs in males from 30 to 50 years old [6]. According to statistics, HCC results in 662,000 deaths per year globally and approximately half of them in China [7]. Therefore, revealing the mechanisms of HCC is important for developing novel therapies and improving its prognosis.

Previous study reports that the overexpression of *UHRF1* (Ubiquitin-like with PHD and RING finger domains 1) affects DNA hypomethylation in HCC cells and that senescence is a main means of inhibiting tumorigenesis induced by epigenetic disruption [8, 9]. *c-Myc* silences tumor-suppressive miRNAs (microRNAs) in the process of hepatocarcinogenesis through collaborating with EZH2 (enhancer of zeste homolog 2)-containing PRC2 (polycomb repressive complex 2) complex and can serve as potential candidate for the treatment of human HCC [10]. *SALL4* (spalt-like transcription factor 4) is found to be useful for the diagnosis and therapy of HCC with the characteristics of stem cells [11, 12]. Li et al. demonstrate that the miR-224/HOXD10 (homeobox D10)/p-PAK4 (phosphorylated p21 protein (Cdc42/Rac)-activated kinase 4/MMP-9 (matrix metallopeptidase 9) signaling pathway promotes the regulation of cell invasion and migration and provides promising therapeutic targets for HCC [13]. Zhang et al. assume that *miR-7* plays tumor-suppressive role during hepatocarcinogenesis via inhibiting the expression of oncogene *CCNE1* (cyclin E1) and can be used for HCC treatment [14]. However, the pathogenesis of HCC has not been completely reported yet.

In 2014, Mah et al. [15] detected the gene expression profiles and methylation profiles of HCC patients, finding that inflammation via the NF-κB (nuclear factor-kappa B) pathway functions in regulating gene expression of HCC patients by methylation. Nevertheless, they have not fully analyzed the gene expression patterns of HCC patients using

in-depth bioinformatics analysis. Using the gene expression profiles deposited by Mah et al. [15], differential expression analysis, enrichment analysis, and integrated regulatory network analysis were successively carried out to identify the key genes associated with the mechanisms of HCC.

Materials and methods

Microarray data. Microarray data of GSE57957, which was based on the platform of GPL10558 Illumina HumanHT-12 V4.0 expression beadchip, was obtained from GEO (Gene Expression Omnibus, http://www.ncbi.nlm.nih.gov/geo/) database. GSE19701 contained 39 HCC samples and 39 adjacent non-tumorous samples. HCC tissues and adjacent non-tumorous liver tissues were from the NCCS (National Cancer Centre of Singapore)/SingHealth Tissue Repository, and all patients have given their informed consent. Tissue samples were isolated, frozen and then stored at −80 °C. The Qiagen RNeasy mini kit (Qiagen, Germany) was used to extract RNA from the tissue samples. Mah et al. [15] deposited GSE57957, and their research got the approval of the SingHealth Centralized Institutional Review Board.

Data preprocessing and DEGs (differentially expressed genes) screening. The raw data was normalized by the RMA (Robust Multiarray Average) method [16] of the Affy package in R. To identify the DEGs, we grouped the 39 HCC samples together and compared the gene expression to that in 39 non-tumorous samples based on the Bayesian method in the R package limma (Linear Models for Microarray Analysis, http://www.bioconductor.org/packages/release/bioc/html/limma.html) [17]. The genes with |logFC (fold change)| >1.5 and adjusted p-value <0.05 were selected as DEGs.

Functional and pathway enrichment analysis. GO (Gene Ontology, http://www.geneontology.org) database has a series of controlled, structured vocabularies for annotating genes, gene sequences and products [18]. The KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.ad.jp/kegg) database is developed for exploring gene functions, linking genomic information with corresponding functional information [19]. Reactome (http://www. reactome. org) is a knowledgebase of pathways, reactions and biological processes [20]. "BioCloud" online platform can be used to solve computational problems of high-throughput data. Based on "BioCloud" platform (http://www.biocloudservice.com), GO functional, KEGG pathway and Reactome pathway enrichment analyses were conducted for the DEGs.

PPI (protein-protein interaction) network analysis. STRING (Search Tool for the Retrieval of Interacting Genes, http://string-db.org/) database provides direct and indirect PPIs associated with over 1100 organisms [21]. Using STRING database [21], PPIs were predicted for the DEGs, with the threshold of combined score >0.4. Then, PPI network was constructed by Cytoscape software (http://www.cytoscape.org) [22].

Integrated network analysis. Using WebGestalt (WEB-based gene set analysis toolkit, http://www.webgestalt.org) tool [23], the miRNAs targeting the nodes of the PPI network were predicted, with the number of target genes ≥4 as the cut-off criterion. According to the adjusted p-values, the top 10 predicted results for miRNAs were obtained for further analysis. Based on the ITFP (integrated transcription factor platform, http://itfp.biosino.org/itfp) [24] and TRANSFAC databases (http://www.gene-regulation.com/pub/databases.html) [25], the TFs (transcription factors) among the DEGs and the DEGs targeted by them were predicted. Using Cytoscape software [22], an integrated network involving the PPI pairs, miRNA-target pairs, and TF-target pairs were constructed. Furthermore, module analysis was performed for the integrated network using the MCODE (Molecular Complex Detection) plugin [26] of Cytoscape software, with the default parameters.

Results

DEGs analysis. With the thresholds of |logFC| >1.5 and adjusted p-value <0.05, a total of 518 DEGs (194 up-regulated and 324 down-regulated) were identified in the HCC samples compared with the adjacent non-tumorous samples. There were more down-regulated genes relative to up-regulated genes.

Functional and pathway enrichment analysis. The up-regulated genes and the down-regulated genes separately were conducted with enrichment analysis. The up-regulated genes were enriched in 390 GO terms, 11 KEGG pathways, and 87 Reactome pathways. The top 5 terms in each category are listed in Table 1, mainly including mitotic cell cycle (GO, p-value = 1.26E-12), ECM-receptor interaction (KEGG pathway, p-value = 1.47E-06), and Cell Cycle, Mitotic (Reactome, p-value = 1.76E-11). Besides, the down-regulated genes were involved in 29 KEGG pathways, 828 GO terms, and 67 Reactome pathways. For the down-regulated genes, the enriched terms mainly include organic acid metabolic process (GO, p-value = 0), metabolic pathways (KEGG pathway, p-value = 4.22E-15), and metabolism (Reactome, p-value = 0) (Table 2).

Integrated network analysis. There were separately 314 and 547 interactions in the PPI network for the up-regulated genes and the PPI network for the down-regulated genes. Among the DEGs, 17 up-regulated genes (including *MCM2*, minichromosome maintenance complex component 2; *MCM6*, minichromosome maintenance complex component 6; and *CDC20*, cell division cycle 20) and 5 down-regulated genes could also be regarded as TFs (Table 3). Moreover, the 17 up-regulated TFs and the 5 down-regulated TFs targeted 52 up-regulated genes and 7 down-regulated genes, respectively. In addition, the top 10 miRNA predicted results for the up-regulated genes and the down-regulated genes (including *miR-22*1, and *miR-222*) and are listed in Table 4. The integrated network for the up-regulated genes is shown

Table 1. The top 5 GO (Gene Ontology) terms, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways, and Reactome pathways enriched for the up-regulated genes.

Category	Description	Count	p-value	Gene symbol
GO	GO:0000278~mitotic cell cycle	39	1.26E-12	*TOP2A, CDC20, PRC1, ASPM, AURKA, CDKN3, PTTG1, CDCA5, CCNB2, PTTG3P, NSMCE2, NUSAP1, NCAPG, RFC4, RRS1, MCM2, FAM83D, GMNN, PSME3, MCM4, MELK, KIF20A, MCM6, CDC45, TYMS, CDC25B, NABP2, KIFC1, NCAPD2, UBE2C, CENPW, NUP37, PSMC4, CENPN, CDC123, GINS2, CENPF, TPX2, NUP205*
GO	GO:1903047~mitotic cell cycle process	35	2.71E-12	*TOP2A, CDC20, PRC1, ASPM, AURKA, CDKN3, PTTG1, CDCA5, CCNB2, PTTG3P, NSMCE2, NUSAP1, NCAPG, RRS1, MCM2, FAM83D, PSME3, MCM4, MELK, KIF20A, MCM6, CDC45, TYMS, CDC25B, NABP2, KIFC1, NCAPD2, UBE2C, CENPW, NUP37, PSMC4, GINS2, CENPF, TPX2, NUP205*
GO	GO:0022402~cell cycle process	42	3.19E-11	*TOP2A, CDC20, PRC1, ASPM, AURKA, CDKN3, PTTG1, CDCA5, CCNB2, PTTG3P, NSMCE2, NUSAP1, BRSK1, NCAPG, CKS2, RFC4, RRS1, MCM2, FAM83D, PSME3, MCM4, MELK, KIF20A, MCM6, CDC45, TYMS, CDC25B, NABP2, KIFC1, NCAPD2, UBE2C, CENPW, NUP37, PSMC4, KIAA0196, CDC123, PEA15, GINS2, CENPF, TPX2, NUP205, RRAGD*
GO	GO:0000819~sister chromatid segregation	13	5.21E-10	*TOP2A, CDC20, PTTG1, CDCA5, PTTG3P, NSMCE2, NUSAP1, NCAPG, RRS1, KIFC1, NCAPD2, UBE2C, CENPF*
GO	GO:0007059~chromosome segregation	17	1.45E-09	*TOP2A, CDC20, PTTG1, CDCA5, PTTG3P, NSMCE2, NUSAP1, NCAPG, RRS1, FAM83D, KIFC1, NCAPD2, UBE2C, CENPW, NUP37, CENPN, CENPF*
KEGG	4512~ECM-receptor interaction	9	1.47E-06	*COL4A1, COL1A1, COL4A2, SPP1, COL5A2, THBS4, VWF, HMMR, COL1A2*
KEGG	4974~protein digestion and absorption	7	8.81E-05	*COL4A1, COL1A1, COL4A2, COL5A2, PRSS3, COL1A2, COL15A1*
KEGG	4110~cell cycle	8	0.000215	*CDC20, PTTG1, CCNB2, MCM2, MCM4, MCM6, CDC45, CDC25B*
KEGG	3030~DNA replication	4	0.001232	*RFC4, MCM2, MCM4, MCM6*
KEGG	4510~focal adhesion	8	0.004792	*COL4A1, COL1A1, COL4A2, SPP1, COL5A2, THBS4, VWF, COL1A2*
Reactome	69278~cell cycle, mitotic	25	1.76E-11	*TOP2A, CDC20, AURKA, PTTG1, CDCA5, CCNB2, NCAPG, RFC4, MCM2, GMNN, PSME3, MCM4, KIF20A, MCM6, CDC45, TYMS, CDC25B, NCAPD2, UBE2C, NUP37, PSMC4, CENPN, GINS2, CENPF, NUP205*
Reactome	1640170~cell cycle	26	2.15E-10	*TOP2A, CDC20, AURKA, PTTG1, CDCA5, CCNB2, NCAPG, RFC4, MCM2, GMNN, PSME3, MCM4, KIF20A, MCM6, CDC45, TYMS, CDC25B, NCAPD2, UBE2C, CENPW, NUP37, PSMC4, CENPN, GINS2, CENPF, NUP205*
Reactome	176974~unwinding of DNA	5	1.30E-07	*MCM2, MCM4, MCM6, CDC45, GINS2*
Reactome	1442490~collagen degradation	8	9.65E-07	*COL4A1, COL1A1, COL4A2, COL5A2, MMP9, MMP11, COL1A2, COL15A1*
Reactome	68886~M Phase	14	1.27E-06	*CDC20, PTTG1, CDCA5, CCNB2, NCAPG, PSME3, KIF20A, NCAPD2, UBE2C, NUP37, PSMC4, CENPN, CENPF, NUP205*

in Figure 1, and the nodes with degrees higher than 20 (including *CDC20*) are listed in Table 5A. Meanwhile, the integrated network for the down-regulated genes is shown in Figure 2, and the nodes with degrees higher than 20 (including *FOS*, FBJ murine osteosarcoma viral oncogene homolog; and *ESR1*, estrogen receptor 1) are listed in Table 5B. Especially, *miR-221* and *miR-222* could target both *FOS* and *ESR1* in the integrated network for the down-regulated genes.

A total of 7 modules were obtained from the integrated network for the up-regulated genes, among which, the module with the highest score (Mcode score = 9.333) had 13 nodes and 61 interactions (Figure 3). Importantly, *MCM2* interacted with *MCM6* in the module. The top 5 GO terms and KEGG pathways enriched for the 13 nodes are listed in Table 6, mainly including mitotic cell cycle process (GO, p-value = 0), and protein digestion and absorption (KEGG pathway, p-value = 8.93E-05). What's more, there were 13 modules identified from the integrated network for the down-regulated genes, and the module with the highest score (Mcode score = 7.4) had 11 nodes and 37 interactions (Figure 4). The top 5 terms enriched for the 11 nodes were listed in Table 7, mainly including epoxygenase P450 pathway (GO, p-value = 5.20E-13) and drug metabolism-cytochrome P450 (KEGG pathway, p-value = 2.45E-10).

Discussion

In this study, a total of 518 DEGs (194 up-regulated and 324 down-regulated) were identified from the HCC samples. Among the DEGs, 17 up-regulated genes (including *MCM2*, *MCM6*, and *CDC20*) and 5 down-regulated genes could also function as TFs. Overexpression of *CDC20* is reported to be related to the pathogenesis of HCC, and may be used as a potential therapeutic target for the disease [27]. The low expression of *CDC20* and *HPSE* (heparanase) help cell apoptosis and autophagy, and targeting inhibition of the expression of both *HPSE* and *CDC20* is a promising therapeutic strategy for HCC [28]. The mRNA and protein levels of *MCM6* in plasma can serve as independent biomarkers

Table 2. The top 5 GO (Gene Ontology) terms, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways, and Reactome pathways enriched for the down-regulated genes.

Category	Description	Count	p-value	Gene symbol
GO	GO:0006082~organic acid metabolic process	85	0	*CYP4A11, IVD, AASS, PFKFB1, APOA4, CSAD, TDO2, ETFDH, SLC25A15, GSTZ1, IGF1, ACSM2A, LIPC, FBP1, AKR1D1, MSRA, HPGD, HAAO, BHMT, OAT, PHGDH, ERLIN1, CYP4F2, HAL, HPD, CYP2C18, MTHFD1, ALDH8A1, ALDH6A1, G6PC, ACACB, PTGS2, CYP8B1, PDK4, MAT1A, ACADL, CYP2C19, CYP2C8, CYP4F12, AGXT2, SARDH, UGT2B10, CTH, VNN1, GCDH, GCKR, ACSM3, CYGB, SLC27A5, PLA2G16, GPT2, CYP1A1, GPT, ST3GAL6, FTCD, ACADS, SLC10A1, NCOR1, CYP2C9, OGDHL, GLYAT, BBOX1, PPARGC1A, HAO2, GHR, RBP1, HOGA1, APOA5, GLS2, CYP39A1, CYP2E1, DCN, CYP2A7, ENO3, IDO2, CYP2A6, KMO, TAT, GNMT, PCK1, SDS, SRD5A2, SLCO1B3, ASPG, CYP1A2*
GO	GO:0006805~xenobiotic metabolic process	28	0	*CYP4A11, GSTZ1, AKR7A3, CYP4F2, CYP2C18, CYP8B1, CYP3A7, MAT1A, ADH6, CYP2C19, CYP2C8, GSTA2, CYP4F12, ADH1C, CYP1A1, CYP2C9, GLYAT, CYP39A1, CYP2E1, CYP3A43, CYP2A7, NNMT, CYP2A6, ADH1B, ADH4, NAT2, CYP3A4, CYP1A2*
GO	GO:0009063~cellular amino acid catabolic process	22	0	*IVD, AASS, CSAD, TDO2, GSTZ1, HAAO, OAT, HAL, HPD, ALDH6A1, AGXT2, CTH, GCDH, GPT2, GPT, FTCD, HOGA1, GLS2, IDO2, KMO, TAT, SDS*
GO	GO:0009410~response to xenobiotic stimulus	28	0	*CYP4A11, GSTZ1, AKR7A3, CYP4F2, CYP2C18, CYP8B1, CYP3A7, MAT1A, ADH6, CYP2C19, CYP2C8, GSTA2, CYP4F12, ADH1C, CYP1A1, CYP2C9, GLYAT, CYP39A1, CYP2E1, CYP3A43, CYP2A7, NNMT, CYP2A6, ADH1B, ADH4, NAT2, CYP3A4, CYP1A2*
GO	GO:0016054~organic acid catabolic process	30	0	*CYP4A11, IVD, AASS, CSAD, TDO2, ETFDH, GSTZ1, AKR1D1, HAAO, OAT, CYP4F2, HAL, HPD, ALDH6A1, ACADL, CYP4F12, AGXT2, CTH, GCDH, GPT2, GPT, FTCD, ACADS, HOGA1, GLS2, CYP39A1, IDO2, KMO, TAT, SDS*
KEGG	1100~Metabolic pathways	71	4.22E-15	*CYP4A11, IVD, AASS, CSAD, TDO2, GSTZ1, TKFC, ACSM2A, LIPC, HSD11B1, FBP1, AKR1D1, HAAO, BHMT, OAT, PHGDH, CYP4F2, HAL, HPD, CYP2C18, MTHFD1, ALDH6A1, G6PC, ACACB, PTGS2, CYP8B1, HSD17B2, CYP3A7, MAT1A, ADH6, ACADL, CYP2C19, CYP2C8, AGXT2, SARDH, CDA, UGT2B10, CTH, GCDH, ACSM3, ALPL, ADH1C, SLC27A5, GPT2, CYP1A1, GPT, ST3GAL6, FTCD, ACADS, CYP2C9, OGDHL, HAO2, GLS2, CYP2E1, CYP3A43, CYP2A7, ENO3, NNMT, IDO2, CNDP1, CYP2A6, KMO, ADH1B, TAT, ADH4, PCK1, SDS, NAT2, DBH, CYP3A4, CYP1A2*
KEGG	982~Drug metabolism - cytochrome P450	19	9.77E-15	*GSTA5, GSTZ1, CYP2C18, CYP3A7, ADH6, CYP2C19, CYP2C8, GSTA2, UGT2B10, ADH1C, CYP2C9, CYP2E1, CYP3A43, CYP2A7, CYP2A6, ADH1B, ADH4, CYP3A4, CYP1A2*
KEGG	980~Metabolism of xenobiotics by cytochrome P450	18	8.53E-14	*GSTA5, GSTZ1, CYP2C18, CYP3A7, ADH6, CYP2C19, CYP2C8, GSTA2, UGT2B10, ADH1C, CYP1A1, CYP2C9, CYP2E1, CYP3A43, ADH1B, ADH4, CYP3A4, CYP1A2*
KEGG	830~Retinol metabolism	17	1.88E-13	*CYP4A11, CYP2C18, CYP3A7, ADH6, CYP2C19, CYP2C8, UGT2B10, ADH1C, CYP1A1, CYP2C9, CYP3A43, CYP2A7, CYP2A6, ADH1B, ADH4, CYP3A4, CYP1A2*
KEGG	591~Linoleic acid metabolism	9	3.51E-08	*CYP2C18, CYP3A7, CYP2C19, CYP2C8, CYP2C9, CYP2E1, CYP3A43, CYP3A4, CYP1A2*
Reactome	1430728~Metabolism	95	0	*CYP4A11, IVD, AASS, SLC25A37, PFKFB1, APOA4, CSAD, TDO2, ETFDH, SLC25A15, GSTZ1, TKFC, AKR7A3, HSD11B1, FBP1, AKR1D1, HPGD, HAAO, BHMT, OAT, PHGDH, CYP4F2, HAL, HPD, CYP2C18, MTHFD1, LYVE1, ALDH6A1, G6PC, ACACB, PTGS2, CYP8B1, CYP3A7, CA5A, PDK4, HBB, MAT1A, ADH6, ACADL, CYP2C19, CYP2C8, GSTA2, SLC19A3, CYP4F12, AGXT2, CDA, NPC1L1, CTH, GCDH, GCKR, CYGB, ADH1C, IYD, SLC27A5, PLA2G16, GPT2, CYP1A1, GPT, ST3GAL6, HBA2, NEU4, FTCD, ACADS, SLC10A1, NCOR1, CYP2C9, GLYAT, BBOX1, CETP, APOA5, GLS2, CYP39A1, GBA3, CYP2E1, CYP3A43, DCN, GCGR, ENO3, NNMT, IDO2, CYP2A6, KMO, ADH1B, TAT, ADH4, PCK1, SLC22A1, SRD5A2, NAT2, STAB2, LCAT, SLCO1B3, DBH, CYP3A4, CYP1A2*
Reactome	211859~Biological oxidations	27	2.22E-16	*CYP4A11, GSTZ1, AKR7A3, CYP4F2, CYP2C18, CYP8B1, CYP3A7, MAT1A, ADH6, CYP2C19, CYP2C8, GSTA2, CYP4F12, ADH1C, CYP1A1, CYP2C9, GLYAT, CYP39A1, CYP2E1, CYP3A43, NNMT, CYP2A6, ADH1B, ADH4, NAT2, CYP3A4, CYP1A2*
Reactome	211945~Phase 1 - Functionalization of compounds	20	3.33E-16	*CYP4A11, CYP4F2, CYP2C18, CYP8B1, CYP3A7, ADH6, CYP2C19, CYP2C8, CYP4F12, ADH1C, CYP1A1, CYP2C9, CYP39A1, CYP2E1, CYP3A43, CYP2A6, ADH1B, ADH4, CYP3A4, CYP1A2*
Reactome	71291~Metabolism of amino acids and derivatives	27	5.77E-14	*IVD, AASS, CSAD, TDO2, SLC25A15, GSTZ1, HAAO, BHMT, OAT, PHGDH, HAL, HPD, ALDH6A1, MAT1A, AGXT2, CTH, GCDH, IYD, GPT2, GPT, FTCD, BBOX1, GLS2, IDO2, KMO, TAT, DBH*
Reactome	211897~Cytochrome P450 - arranged by substrate type	16	1.24E-13	*CYP4A11, CYP4F2, CYP2C18, CYP8B1, CYP3A7, CYP2C19, CYP2C8, CYP4F12, CYP1A1, CYP2C9, CYP39A1, CYP2E1, CYP3A43, CYP2A6, CYP3A4, CYP1A2*

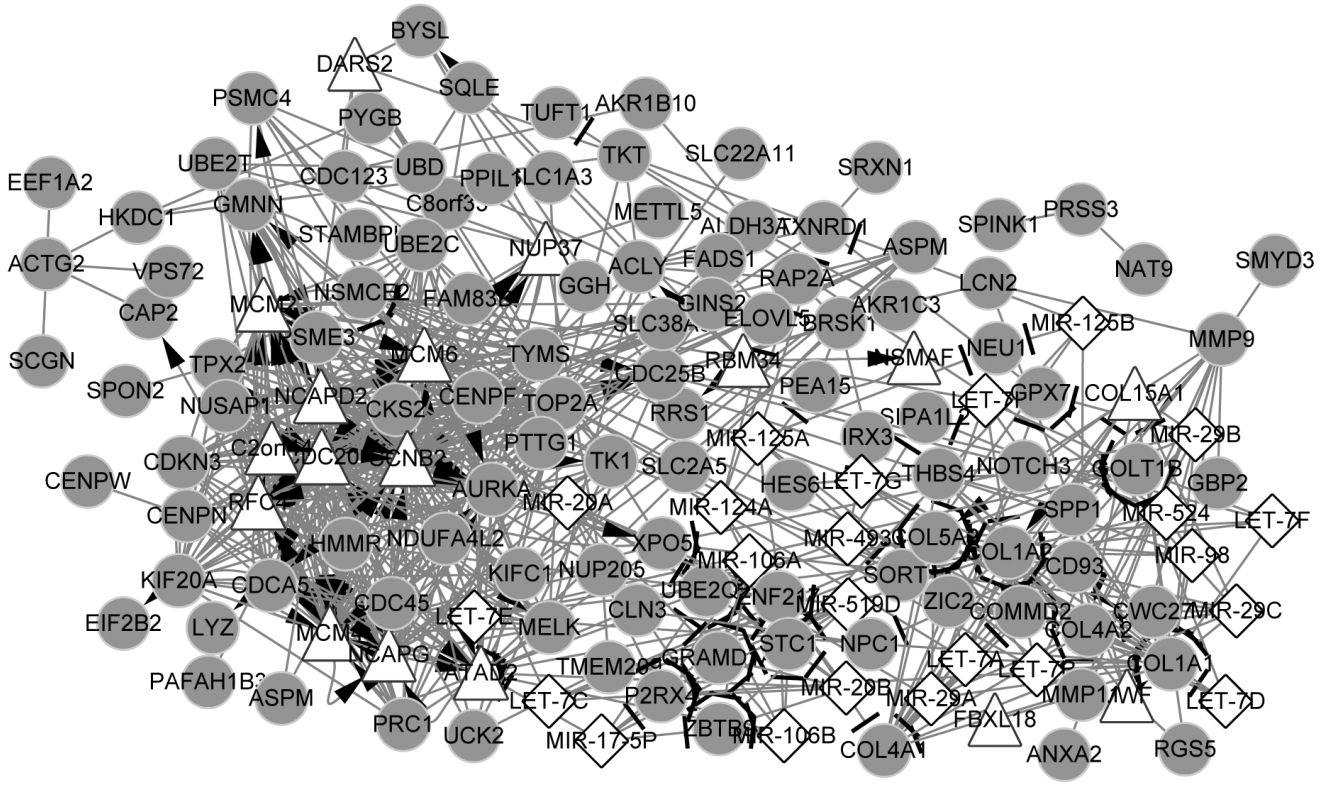

Figure 1. The integrated network for the up-regulated genes. The circles, triangles, and diamonds represent up-regulated genes, transcription factors, and microRNAs, respectively.

Table 3 The up-regulated genes and the down-regulated genes that could also be taken as TFs (transcription factors).

UP	DOWN
ATAD2,C2orf44,CCNB2,CDC20,COL15A1,DARS2,FBXL18,MCM2, MCM4,MCM6,NCAPD2,NCAPG,NSMAF,NUP37,RBM34,RFC4,VWF	*EGR1,FOSB,SARDH,SHBG,TACSTD2*

Table 4. The top 10 miRNA (microRNA) predicted results for the up-regulated genes and the down-regulated genes.

Category	microRNA	Count	Adjusted p-value
UP	*hsa_ATGTACA,MIR-493*	5	0.4988
	hsa_CTCAGGG,MIR-125B,MIR-125A	5	0.4988
	hsa_GCACTTT,MIR-17-5P,MIR-20A,MIR-106A,MIR-106B,MIR-20B,MIR-519D	7	0.4988
	hsa_CTACCTC,LET-7A,LET-7B,LET-7C,LET-7D,LET-7E,LET-7F,MIR-98,LET-7G,LET-7I	5	0.4988
	hsa_TGGTGCT,MIR-29A,MIR-29B,MIR-29C	8	0.4988
	hsa_CTTTGTA,MIR-524	5	0.4988
	hsa_TGCCTTA,MIR-124A	5	1.0000
DOWN	*hsa_ATGTAGC,MIR-221,MIR-222*	6	0.3202
	hsa_AACATTC,MIR-409-3P	6	0.3202
	hsa_TGCACTG,MIR-148A,MIR-152,MIR-148B	8	0.7214
	hsa_GTGCCAA,MIR-96	8	0.7214
	hsa_ACTGTGA,MIR-27A,MIR-27B	10	0.7386
	hsa_TATTATA,MIR-374	7	0.7386
	hsa_ATACTGT,MIR-144	5	0.7386
	hsa_TAGCTTT,MIR-9	5	0.7453
	hsa_TACTTGA,MIR-26A,MIR-26B	6	0.7453
	hsa_AAAGGGA,MIR-204,MIR-211	5	0.7453

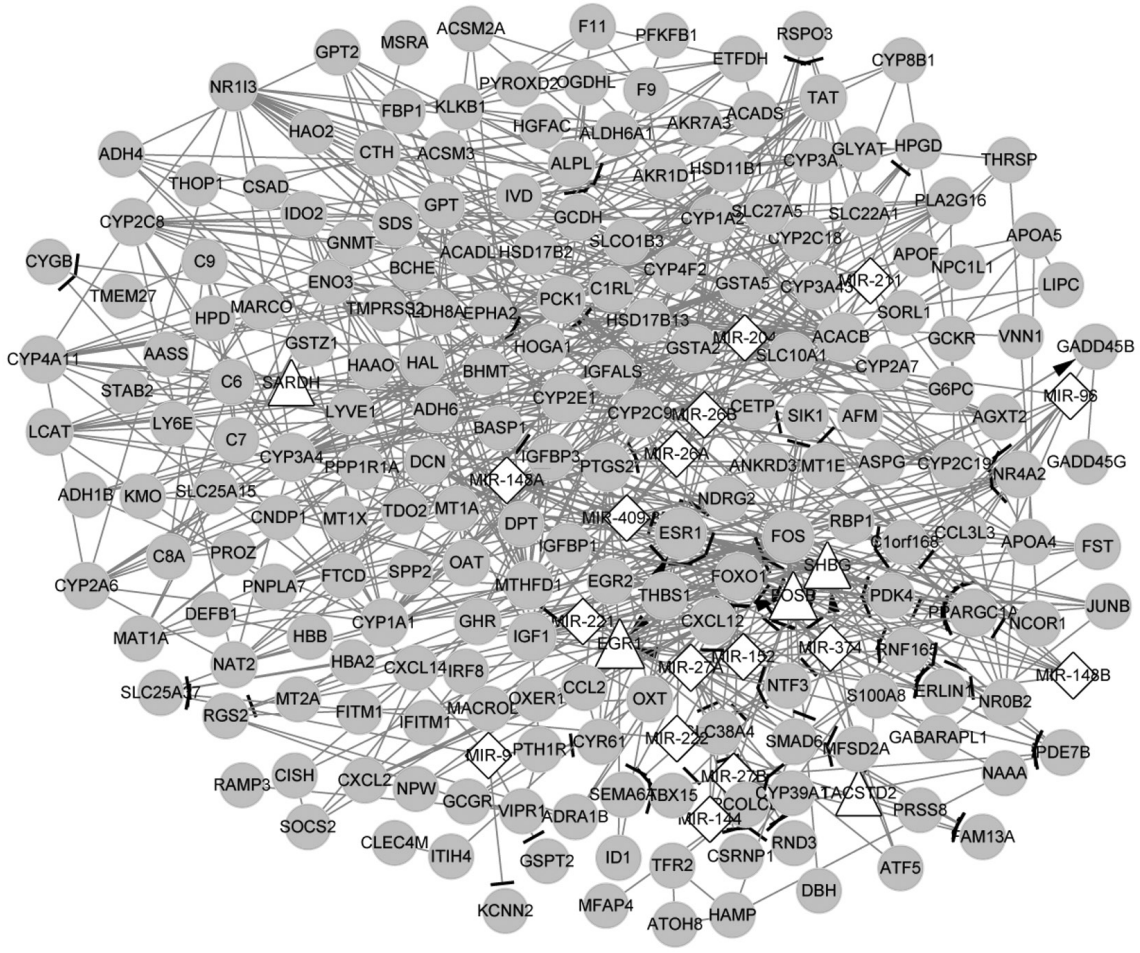

Figure 2. The integrated network for the down-regulated genes. The circles, triangles, and diamonds represent down-regulated genes, transcription factors, and microRNAs, respectively.

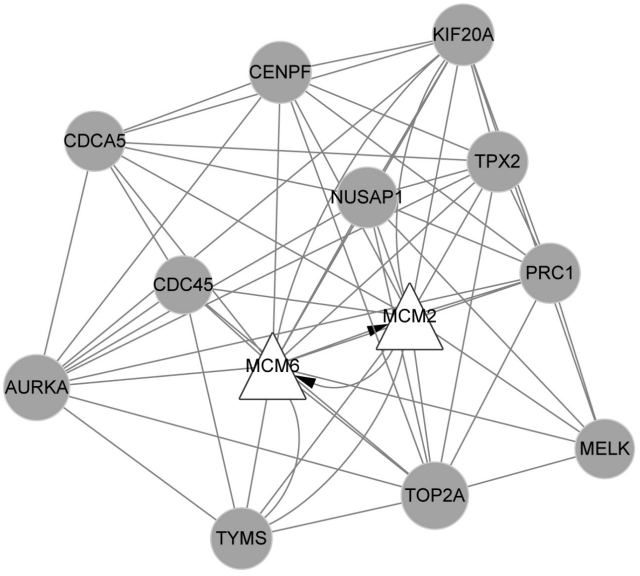

Figure 3. The most significant module obtained from the integrated network for the up-regulated genes. The circles and triangles represent up-regulated genes and transcription factors, respectively.

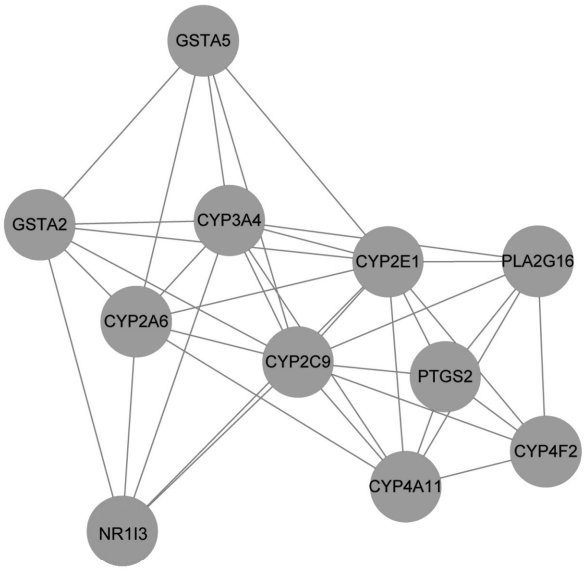

Figure 4. The most significant module obtained from the integrated network for the down-regulated genes. The circles represent down-regulated genes.

Table 5. The nodes with degrees higher than 20 in the integrated network for the up-regulated gene, and those in the integrated network for the down-regulated genes.

Gene	Degree	Gene	Degree
Up-regulated			
CDC20	69	NCAPD2	32
CCNB2	68	ATAD2	30
MCM4	54	NCAPG	30
TOP2A	39	COL1A1	24
RFC4	39	CENPF	23
MCM6	37	KIF20A	22
MCM2	37	TYMS	22
AURKA	35		
Down-regulated			
FOS	34	CYP2E1	22
ESR1	34	EGR1	22
ACACB	24	IGF1	21
PTGS2	23		

for HCC, especially in patients with AFP (α-fetoprotein) negative and small HCC [29]. Marshall et al. find that the early evaluation of *MCM*2 expression in hepatocyte contributes to predict the risk of progressive fibrosis for post-transplant HCV hepatitis [30, 31]. MCMs possess high sensitivity and specificity, thus MCMs are taken as prognostic and diagnostic markers in the clinical treatment of some types of human malignant tumors [32]. *MCM2* had interaction with *MCM*6 in the up-regulated module with the highest score. Thus, *MCM2*, *MCM6*, and *CDC20* might function in the pathogenesis of HCC, and *MCM2* and *MCM6* might affect HCC through interacting with each other.

Li et al. demonstrate that the expression of oncogene *FOS* is inhibited by *miR-101* which is dysregulated in HCC [33]. *RB1*, (retinoblastoma 1), *p53*, and *FOS* play important roles in HCC in Iran, and their simultaneous overexpression is remarkably correlated with their expression deletion during the development of HCC [34]. Fan et al. report that the derepression of *FOS* induced by the down-regulation

Table 6. The top 5 GO (Gene Ontology) terms and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways enriched for the nodes in the most significant module obtained from the integrated network for the up-regulated genes.

Category	Description	Count	p-value	Gene symbol
GO	GO:1903047~mitotic cell cycle process	13	0	*CENPF,AURKA,TPX2,NUSAP1,TYMS,MCM6,CDCA5,MCM2,PRC1,CDC45,KIF20A,MELK,TOP2A*
GO	GO:0000278~mitotic cell cycle	13	2.22E-16	*CENPF,AURKA,TPX2,NUSAP1,TYMS,MCM6,CDCA5,MCM2,PRC1,CDC45,KIF20A,MELK,TOP2A*
GO	GO:0022402~cell cycle process	13	3.55E-15	*CENPF,AURKA,TPX2,NUSAP1,TYMS,MCM6,CDCA5,MCM2,PRC1,CDC45,KIF20A,MELK,TOP2A*
GO	GO:0007049~cell cycle	13	1.35E-13	*CENPF,AURKA,TPX2,NUSAP1,TYMS,MCM6,CDCA5,MCM2,PRC1,CDC45,KIF20A,MELK,TOP2A*
GO	GO:0044772~mitotic cell cycle phase transition	8	4.19E-10	*CENPF,AURKA,TYMS,MCM6,CDCA5,MCM2,CDC45,MELK*
KEGG	4110~Protein digestion and absorption	3	8.93E-05	*MCM6,MCM2,CDC45*
KEGG	3030~ECM-receptor interaction	2	0.000362	*MCM6,MCM2*

Table 7. The top 5 GO (Gene Ontology) terms and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways enriched for the nodes in the most significant module obtained from the integrated network for the down-regulated genes.

Category	Description	Count	P-value	Gene symbol
GO	GO:0019373~epoxygenase P450 pathway	5	5.20E-13	*CYP4F2,CYP2C9,CYP2E1,CYP2A6,CYP4A11*
GO	GO:0019369~arachidonic acid metabolic process	6	8.91E-13	*CYP4F2,CYP2C9,CYP2E1,PTGS2,CYP2A6,CYP4A11*
GO	GO:0006805~xenobiotic metabolic process	7	3.56E-12	*CYP4F2,GSTA2,CYP3A4,CYP2C9,CYP2E1,CYP2A6,CYP4A11*
GO	GO:0071466~cellular response to xenobiotic stimulus	7	4.20E-12	*CYP4F2,GSTA2,CYP3A4,CYP2C9,CYP2E1,CYP2A6,CYP4A11*
GO	GO:0009410~response to xenobiotic stimulus	7	5.33E-12	*CYP4F2,GSTA2,CYP3A4,CYP2C9,CYP2E1,CYP2A6,CYP4A11*
KEGG	982~Drug metabolism - cytochrome P450	6	2.45E-10	*GSTA5,GSTA2,CYP3A4,CYP2C9,CYP2E1,CYP2A6*
KEGG	590~Arachidonic acid metabolism	5	1.06E-08	*CYP4F2,CYP2C9,CYP2E1,PTGS2,CYP4A11*
KEGG	980~~Metabolism of xenobiotics by cytochrome P450	5	2.73E-08	*GSTA5,GSTA2,CYP3A4,CYP2C9,CYP2E1*
KEGG	830~Retinol metabolism	4	1.56E-06	*CYP3A4,CYP2C9,CYP2A6,CYP4A11*
KEGG	591~Linoleic acid metabolism	3	9.92E-06	*CYP3A4,CYP2C9,CYP2E1*

of *miR-139* promotes the metastasis of HCC [35]. *ESR1* is a potential tumor suppressor gene in HCC and its promoter hypermethylation suppresses its expression, thus its hypermethylation level may be used to predict HCC status and progression [36, 37]. In the integrated network for the down-regulated genes, *FOS* and *ESR1* had degrees higher than 20, indicating that *FOS* and *ESR1* might be associated with the development and progression of HCC.

*miR-22*1, and *miR-222* were among the top 10 miRNA and predicted results for the down-regulated genes. *miR-221* inhibition suppresses cell proliferation, prevents cell cycle progression, and de-represses *p27* in HCC cells, thus *miR-221* may be a promising therapeutic target for HCC [38, 39]. *miR-221* can cause functional suppression or loss of the tumor suppressor *HDAC6* (histone deacetylase 6) by mediating NF-κB- and JNK (c-Met-mediated c-Jun NH2-terminal kinase)/c-Jun-signaling pathways in the process of liver tumorigenesis [40]. Yang et al. report that increased *miR-222* expression contributes to the proliferation of HCC HepG2 cells through reducing *p27* [41]. Ogawa et al. believe that *miR-221/222* may serve as novel markers for liver fibrosis progression and stellate cell activation [42]. In the integrated network for the down-regulated genes, *miR-221* and *miR-222* could target both *FOS* and *ESR1*, suggesting that *miR-221* and *miR-222* might also act in the mechanisms of HCC through targeting both *FOS* and *ESR1*.

In conclusion, a total of 518 DEGs were identified in the HCC samples using bioinformatics analysis. Besides, *MCM2*, *MCM6*, *CDC20*, *FOS*, *ESR1*, *miR-221* and *miR-222* might act in the pathogenesis of HCC. However, these genes still need to be confirmed by experimental research.

Supplementary information is available in the online version of the paper.

References

[1] MARENGO A, ROSSO C, BUGIANESI E. Liver cancer: connections with obesity, fatty liver, and cirrhosis. Annu Rev Med 2016; 67: 103–117. doi: 10.1146/annurev-med-090514-013832

[2] WHITE DL, KANWAL F, EL-SERAG HB. Association between nonalcoholic fatty liver disease and risk for hepatocellular cancer, based on systematic review. Clinical Gastroenterol Hepatol 2012; 10: 1342–1359. doi: 10.1016/j.cgh.2012.10.001

[3] TANAKA M, KATAYAMA F, KATO H, TANAKA H, WANG J et al. Hepatitis B and C virus infection and hepatocellular carcinoma in China: a review of epidemiology and control measures. J Epidemiol 2011; 21: 401–416.

[4] FAN H, CUI Z, ZHANG H, MANI S, DIAB A et al. DNA demethylation induces SALL4 gene re-expression in subgroups of hepatocellular carcinoma associated with Hepatitis B or C virus infection. Oncogene 2017; 36: 2435–2445. doi: 10.1038/onc.2016.399

[5] BRUIX J, SHERMAN M. Management of hepatocellular carcinoma: an update. Hepatology. 2011; 53(3): 1020–1022. doi: 10.1002/hep.24199

[6] THEISE ND. Liver and Gallbladder, chapter 18, pp. 812–882. In: V. Kumar, AK. Abbas, JC. Aster (Eds.). Robbins and Cotran Pathological Basis of Disease Professional Edition, 9th Edition, p. 1408. ISBN 978-0-323-26616-1

[7] BOSETTI C, TURATI F, LA VECCHIA C. Hepatocellular carcinoma epidemiology. Best Pract Res Clin Gastroenterol 2014; 28: 753–770. doi: 10.1016/j.bpg.2014.08.007

[8] MUDBHARY R, HOSHIDA Y, CHERNYAVSKAYA Y, JACOB V, VILLANUEVA A et al. UHRF1 Overexpression Drives DNA Hypomethylation and Hepatocellular Carcinoma. Cancer Cell 2014; 25: 196–209. doi: 10.1016/j.ccr.2014.01.003

[9] ZHUO H, TANG J, LIN Z, JIANG R, ZHANG X et al. The aberrant expression of MEG3 regulated by UHRF1 predicts the prognosis of hepatocellular carcinoma. Mol Carcinogenesis 2016; 55: 209–219. doi: 10.1002/mc.22270

[10] WANG L, ZHANG X, JIA LT, HU SJ, ZHAO J et al. c-Myc-mediated epigenetic silencing of MicroRNA-101 contributes to dysregulation of multiple pathways in hepatocellular carcinoma. Hepatology 2014; 59: 1850–1863. doi: 10.1002/hep.26720

[11] ZENG SS, YAMASHITA T, KONDO M, NIO K, HAYASHI T et al. The transcription factor SALL4 regulates stemness of EpCAM-positive hepatocellular carcinoma. J Hepatol 2014; 60: 127–134. doi: 10.1016/j.jhep.2013.08.024

[12] HAN SX, WANG JL, GUO XJ, HE CC, YING X et al. Serum SALL4 is a novel prognosis biomarker with tumor recurrence and poor survival of patients in hepatocellular carcinoma. J Immunol Res 2014; 2014: 262385. doi: 10.1155/2014/262385

[13] LI Q, DING C, CHEN C, ZHANG Z, XIAO H et al. miR-224 promotion of cell migration and invasion by targeting Homeobox D 10 gene in human hepatocellular carcinoma. J Gastroenterol Hepatol 2014; 29: 835–842. doi: 10.1111/jgh.12429

[14] ZHANG X, HU S, ZHANG X, WANG L, ZHANG X et al. MicroRNA-7 arrests cell cycle in G1 phase by directly targeting CCNE1 in human hepatocellular carcinoma cells. Biochem Biophys Res Commun 2014; 443: 1078–1084. doi: 10.1016/j.bbrc.2013.12.095

[15] MAH WC, THURNHERR T, CHOW PK, CHUNG AY, OOI LL et al. Methylation Profiles Reveal Distinct Subgroup of Hepatocellular Carcinoma Patients with Poor Prognosis. PLoS One 2014; 9: e104158. doi: 10.1371/journal.pone.0104158

[16] IRIZARRY RA, HOBBS B, COLLIN F, BEAZER-BARCLAY YD, ANTONELLIS KJ et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003; 4: 249–264.

[17] RITCHIE ME, PHIPSON B, WU D, HU Y, LAW CW et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47. doi: 10.1093/nar/gkv007

[18] GENE ONTOLOGY CONSORTIUM. Gene Ontology Consortium: going forward. Nucleic Acids Res 2015; 43: 1049–1056. doi: 10.1093/nar/gku1179

[19] KANEHISA M, SATO Y, KAWASHIMA M, FURUMICHI M, TANABE M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 2015; 44: D457–D462. doi: 10.1093/nar/gkv1070

[20] CROFT D, O'KELLY G, WU G, HAW R, GILLESPIE M et al. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res 2011; 39: 691–697. doi: 10.1093/nar/gkq1018

[21] FRANCESCHINI A, SZKLARCZYK D, FRANKILD S, KUHN M, SIMONOVIC M et al. STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 2013; 41: D808–D815. doi: 10.1093/nar/gks1094

[22] SAITO R, SMOOT ME, ONO K, RUSCHEINSKI J, WANG P-L et al. A travel guide to Cytoscape plugins. Nat Methods 2012; 9: 1069–1076. doi: 10.1038/nmeth.2212

[23] WANG J, DUNCAN D, SHI Z, ZHANG B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res 2013; 41: 77–83. doi: 10.1093/nar/gkt439

[24] ZHENG G, TU K, YANG Q, XIONG Y, WEI C et al. ITFP: an integrated platform of mammalian transcription factors. Bioinformatics 2008; 24: 2416–2417. doi: 10.1093/bioinformatics/btn439

[25] MATYS V, KEL-MARGOULIS OV, FRICKE E, LIEBICH I, LAND S. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 2006; 34: D108–D110.

[26] BADER GD, HOGUE CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 2003; 4: 2.

[27] LI J, GAO JZ, DU JL, HUANG ZX, WEI LX. Increased CDC20 expression is associated with development and progression of hepatocellular carcinoma. Int J Oncol 2014; 45: 1547–1555. doi: 10.3892/ijo.2014.2559

[28] LIU M, ZHANG Y, LIAO Y, CHEN Y, PAN Y et al. Evaluation of the Antitumor Efficacy of RNAi-Mediated Inhibition of CDC20 and Heparanase in an Orthotopic Liver Tumor Model. Cancer Biother Radiopharm 2015; 30: 233–239. doi: 10.1089/cbr.2014.1799

[29] ZHENG T, CHEN M, HAN S, ZHANG L, BAI Y et al. Plasma minichromosome maintenance complex component 6 is a novel biomarker for hepatocellular carcinoma patients. Hepatol Res 2014; 44: 1347–1356. doi: 10.1111/hepr.12303

[30] MARSHALL A, RUSHBROOK S, MORRIS LS, SCOTT IS, VOWLER SL et al. Hepatocyte expression of minichromosome maintenance protein-2 predicts fibrosis progression after transplantation for chronic hepatitis C virus: a pilot study. Liver Transpl 2005; 11: 427–433.

[31] UNITT E, GELSON W, DAVIES SE, COLEMAN N, ALEXANDER GJ. Minichromosome maintenance protein-2-positive portal tract lymphocytes distinguish acute cellular rejection from hepatitis C virus recurrence after liver transplantation. Liver Transpl 2009; 15: 306–312. doi: 10.1002/lt.21680

[32] GIAGINIS C, VGENOPOULOU S, VIELH P, THEOCHARIS S. MCM proteins as diagnostic and prognostic tumor markers in the clinical setting. Histol Histopathol 2010; 25: 351–370. doi: 10.14670/HH-25.351

[33] LI S, FU H, WANG Y, YI T, XING R et al. MicroRNA-101 regulates expression of the v-fos FBJ murine osteosarcoma viral oncogene homolog (FOS) oncogene in human hepatocellular carcinoma. Hepatology 2009; 49: 1194–1202. doi: 10.1002/hep.22757

[34] MOGHADDAM SJ, HAGHIGHI EN, SAMIEE S, SHAHID N, KERAMATI AR et al. Immunohistochemical analysis of p53, cyclinD1, RB1, c-fos and N-ras gene expression in hepatocellular carcinoma in Iran. World J Gastroenterol 2007; 13: 588–593.

[35] FAN Q, HE M, DENG X, WU WKK, ZHAO L et al. Derepression of c-Fos caused by MicroRNA-139 down-regulation contributes to the metastasis of human hepatocellular carcinoma. Cell Biochem Funct 2013; 31: 319–324. doi: 10.1002/cbf.2902

[36] HISHIDA M, NOMOTO S, INOKAWA Y, HAYASHI M, KANDA M et al. Estrogen receptor 1 gene as a tumor suppressor gene in hepatocellular carcinoma detected by triple-combination array analysis. International Journal of Oncology 2013; 43: 88–94. doi: 10.3892/ijo.2013.1951

[37] DAI B, GENG L, YU Y, SUI C, XIE F et al. Methylation patterns of estrogen receptor promoter correlate with estrogen receptor expression and clinicopathological factors in hepatocellular carcinoma. Exp Biol Med (Maywood) 2014; 239: 883–890.

[38] LEE EC, XU J, VINCENT T, CHEN J, NELSON M et al. Abstract 1459: miR-221: A potential therapeutic target for hepatocellular carcinoma. Cancer Res 2014; 74: 1459–1459. doi: 10.1158/1538-7445.AM2014-1459

[39] HE XX, GUO AY, XU CR, CHANG Y, XIANG GY et al. Bioinformatics analysis identifies miR-221 as a core regulator in hepatocellular carcinoma and its silencing suppresses tumor properties. Oncol Rep 2014; 32: 1200–1210. doi: 10.3892/or.2014.3306

[40] KIM HS, SHEN Q, EUN JW, SHIN WC, YANG HD et al. Abstract 3111: JNK/c-Jun- and NF-κB-mediated microRNA-221 governs tumor suppressor HDAC6 to potentiate malignant progression of liver cancer. Cancer Res 2015; 75: 3111–3111. doi: 10.1158/1538-7445.AM2015-3111

[41] YANG YF, WANG F, XIAO JJ, SONG Y, ZHAO YY et al. MiR-222 overexpression promotes proliferation of human hepatocellular carcinoma HepG2 cells by downregulating p27. Int J Clin Exp Med 2014; 7: 893–902.

[42] OGAWA T, ENOMOTO M, FUJII H, SEKIYA Y, YOSHIZATO K et al. MicroRNA-221/222 upregulation indicates the activation of stellate cells and the progression of liver fibrosis. Gut 2012; 61: 1600–1609.