Quality Improvement Study

Medicine®

OPEN

Systems analysis of key genes and pathways in the progression of hepatocellular carcinoma

Yu-Kui Shang, MS[a,b], Fanni Li, PhD[b], Yi Zhang, MD, PhD[b], Ze-Kun Liu, MD[b], Zi-Ling Wang, MD, PhD[a], Huijie Bian, MD, PhD[b,*], Zhi-Nan Chen, MD, PhD[a,b,*]

Abstract

The carcinogenesis of hepatocellular carcinoma (HCC) is a complex process, starting from a chronically altered hepatic microenvironment due to liver cirrhosis and ultimately progressing to HCC. However, the sequential molecular alterations driving the malignant transformation in liver cirrhosis are not clearly defined.

In this study, we obtained gene expression profiles of HCC, including 268 tumor tissues, 243 adjacent tumor tissues, and 40 cirrhotic tissues (GSE25097) from Gene Expression Omnibus (GEO), to comprehensively define changes in the transcriptome of HCC during the sequential evolution of liver cirrhosis into HCC.

We showed that changes in the molecular profiles of cirrhotic and adjacent tumor samples were small and quite uniform, whereas there was a striking increase in the heterogeneity of tumors in HCC tissues at the mRNA level. A massive deregulation of key oncogenic molecules and pathways was observed from cirrhosis to HCC tumors. In addition, we focused on *FOXO1* and *DCN*, 2 critical tumor suppressor genes that play an important role in liver cirrhosis and HCC development. *FOXO1* and *DCN* expression levels were significantly reduced in tumor tissues compared with adjacent tumor tissues in HCC. Kaplan–Meier analysis revealed that *FOXO1* and *DCN* expression was positively correlated with overall survival, defining *FOXO1* and *DCN* as adverse prognostic biomarkers for HCC.

This system-level research provided new insights into the molecular mechanisms of HCC carcinogenesis. *FOXO1* and *DCN* may be applied as potential targets for HCC treatment in the future.

Abbreviations: DEGs = differentially expressed genes, FDR = false discovery rate, GEO = Gene Expression Omnibus, HCC = hepatocellular carcinoma, KEGG = Kyoto Encyclopedia of Genes and Genomes, *SPIA* = Signalling Pathway Impact Analysis, TCGA = The Cancer Genome Atlas.

Keywords: DCN, FOXO1, gene expression omnibus, hepatocellular carcinoma, TCGA

1. Introduction

Hepatocellular carcinoma (HCC) is the fifth most common malignancy and the third leading cause of cancer-related death worldwide.[1] HCC accounts for 80% to 90% of primary liver cancers, and the incidence of HCC is growing globally by 3% to 9% annually.[2] HCC neoplasms detected at an early stage can be cured by mainly surgical resection. Treatment options for HCC at an advanced stage are often limited.[3,4] The survival duration of patients with advanced liver cancer is less than 12 months.[3] Early detection of HCC may help improve long-term survival rates.[4] Therefore, there is an urgent need for a deeper understanding of the molecular mechanisms underlying the initiation and progression of HCC, and this information might be helpful for designing novel therapeutic strategies in the future.

Because the liver is especially susceptible to chronic and acute viral injury, alcoholic insults, and nonalcoholic fatty liver disease, it is extremely prone to fibrotic remodeling.[5] Liver fibrosis usually progresses to cirrhosis, which can result in damage to the normal architecture of the liver, followed by an increased probability of the development of HCC.[5] HCC occurs at a rate of 1% to 4% per year once liver cirrhosis is established, and liver cirrhosis underlies HCC in approximately 80% to 90% of cases worldwide.[6] Increasing evidence has demonstrated that the carcinogenesis of HCC is a multistep process triggered by the accumulation of genetic alterations through the activation of different signaling pathways, which drives the transformation of normal cells into malignant cells.[5,7] However, the mechanism behind the progression from liver cirrhosis to HCC remains largely unknown. To the best of our knowledge, no systematic study has been performed to investigate the molecular events leading from liver cirrhosis to HCC. A definition of the sequential molecular events leading from cirrhosis to HCC is urgently needed, and it represents a major challenge in the clinical management of at-risk patients.

Editor: Shan Wang.

YKS, FL, and YZ have contributed equally to this work.

Funding/support: This work was funded by the National Science Foundation of China (31571434:H.Bian), National Basic Research Program of China (2015CB553701: Z.N. Chen) and China Postdoctoral Science Foundation (2017M613350:FN.Li).

The authors report no conflicts of interest.

Supplemental Digital Content is available for this article.

[a] College of Life Sciences and Bioengineering, Beijing Jiaotong University, Beijing, [b] State Key Laboratory of Cancer Biology, Cell Engineering Research Center & Department of Cell Biology, Fourth Military Medical University, Xi'an, China.

[*] Correspondence: Zhi-Nan Chen, College of Life Sciences and Bioengineering, Beijing Jiaotong University, Beijing, 100044, China (e-mail: znchen@fmmu.edu.cn); Huijie Bian, Fourth Military Medical University, Xi'an, Shanxi, 710032, China (e-mail: hjbian@fmmu.edu.cn).

Medicine (2018) 97:23(e10892)

Received: 8 October 2017 / Accepted: 8 May 2018

http://dx.doi.org/10.1097/MD.0000000000010892

In this study, we obtained a genome-wide expression profile of HCC, including 268 tumor tissues, 243 adjacent nontumor tissues, and 40 cirrhotic tissues (GSE25097),[8] to comprehensively define changes in the transcriptome of HCC during the sequential evolution of cirrhosis into HCC, and we validated some of these results with 3 other HCC datasets (GSE22508,[9] Oncopression database,[10] and TCGA_LICH), which only included adjacent nontumor and tumor tissues. We showed that changes in the molecular profiles of cirrhotic and adjacent nontumor tissues were small and quite uniform in contrast to the striking increase in heterogeneity of HCC tissues at the mRNA level. A massive deregulation of key oncogenic molecules and pathways was observed from cirrhosis to HCC. In addition, we focused on *FOXO1* and *DCN*, 2 critical tumor suppressor genes that play an important role in liver cirrhosis and HCC development. We detected the expression of *FOXO1* and *DCN* in HCC and analyzed their correlation with clinical pathological features. Our data indicated that low *FOXO1* or *DCN* expression was associated with poor prognosis of HCC.

2. Materials and methods

2.1. HCC datasets

The discovery dataset GSE25097 was obtained from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/)[8]. The validation datasets were extracted from the following 3 datasets: GSE22058, which includes 96 paired adjacent nontumor and tumor samples of HCC from the GEO database[9]; Oncopression datasets (http://www.oncopression.com),[10] including 524 tumor samples and 322 adjacent nontumor samples of HCC that integrate several gene expression datasets based on microarrays from different platforms into 1 large dataset; and the TCGA _LIHC dataset (http://tcga-data.nci. nih.gov, as of January 28, 2016), including 371 tumor samples and 50 adjacent nontumor samples of HCC with both mRNA expression data based on RNA-Seq and clinical feature information, which was used to perform the correlation analysis and survival analysis. All of the data in this study were based on previous published studies, and thus, no ethical approval and patient consent are required.

2.2. Functional enrichment analysis

Pathway analysis of different patterns of gene expression was performed using the *Sigora R* package version 2.0.1, which identified pathway enrichment based on statistically over-represented Pathway Gene-Pair Signatures.[11] Signalling Pathway Impact Analysis (*SPIA*) was used to assess the importance of enriched pathways in terms of their impact and ability to activate or inhibit a pathway.[12]*SPIA* analysis was accomplished using the *R* Bioconductor package *SPIA* (version 2.18.0). Entrez IDs, log2-fold changes, and Q-values of all genes were compiled. *SPIA* produces a *P* value, which represents the significance level at which a pathway is found to be perturbed, and a false discovery rate (FDR). We ran *SPIA* using the recommended value of 2000 bootstrap iterations, and all parameters were set to their default values. A pathway was significant if the FDR was less than 0.1.

2.3. Statistical analysis

A gene was considered differentially expressed when it was significant at 5% FDR (q-value method) and showed an absolute log2 mean difference higher than 1 (double expression). Single comparisons between the 2 groups were determined by a Student *t* test. Survival analysis was performed with the Kaplan–Meier method, and the log-rank test was used to evaluate the statistical significance of the differences. Differences were considered to be statistically significant when $P < .05$.

3. Results

3.1. Group comparison of different stages of hepatocarcinogenesis

We computed the Pearson correlation coefficient of each sample at the corresponding stage based on the mRNA expression values from GSE25097, GSE22058, Oncopression database, and TCGA_LIHC datasets. The gene expression profiles of adjacent nontumor samples were quite homogeneous (the mean coefficient of the Pearson correlation was 0.94 in GSE25097, 0.97 in GSE22058, 0.89 in Oncopression, and 0.89 in TCGA_LIHC). As anticipated, high homogeneity was also observed for cirrhotic samples (the mean coefficient of the Pearson correlation was 0.93 in GSE25097). However, the homogeneity dramatically decreased upon progression to HCC (the mean coefficient of the Pearson correlation was 0.84 in GSE25097, 0.87 in GSE22058, 0.80 in Oncopression, and 0.66 in TCGA_LIHC, $P < .0001$), reflecting the well-recognized prototypical heterogeneity of HCC (Fig. 1).

3.2. Identification of differentially expressed genes (DEGs) among cirrhotic, adjacent nontumor, and tumor tissues of HCC

To investigate the gene expression alterations associated with HCC progression, GSE25097 was used as the discovery dataset for the identification of DEGs among cirrhotic, adjacent nontumor, and HCC tumor samples. This discovery dataset included the gene expression profiles of 268 tumor, 243 adjacent nontumor, and 40 cirrhotic samples. A total of 1920 genes (961 upregulated genes and 959 downregulated genes) were differentially expressed with a FDR $< 5\%$ and a log2 mean difference > 1 between cirrhotic and adjacent nontumor samples, and 2007 genes (966 upregulated genes and 1041 downregulated genes) were differentially expressed between adjacent nontumor and tumor samples (Fig. 2A, B). As shown in Fig. 2C, 1047 genes were significantly differentially expressed among cirrhotic, adjacent nontumor, and tumor samples. The DEGs between cirrhotic and adjacent non-tumor samples were classified as "Tumor-like," "Trend," or "Adjacent-specific" patterns based on their level of expression in tumor tissues, as defined by Sanz-Pamplona et al.[13] There were 873 Tumor-like genes, 275 Trend genes, and 772 Adjacent-specific genes (Fig. 2D). Pathway analysis of different patterns of gene expression was conducted using *Sigora*, which limits the repetitive assignment of the same genes to multiple overlapping pathways. The enriched KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways for these gene patterns were Tumor-like pathways of Lysosome (hsa04142), Ribosome (hsa03010), Oxidative phosphorylation (hsa00190), ECM-receptor interaction (hsa04512), and NOD-like receptor signaling pathway (hsa04621). Trend pathways were enriched in Complement and coagulation cascades (hsa04610). Adjacent-specific pathways were most enriched in metabolism related pathways, such as glycine, serine, and threonine metabolism (hsa00260). These results suggest different functions for each gene expression pattern (see complete list in supplemental Tables S1, S2, and S3, http://links.lww.com/MD/C258).

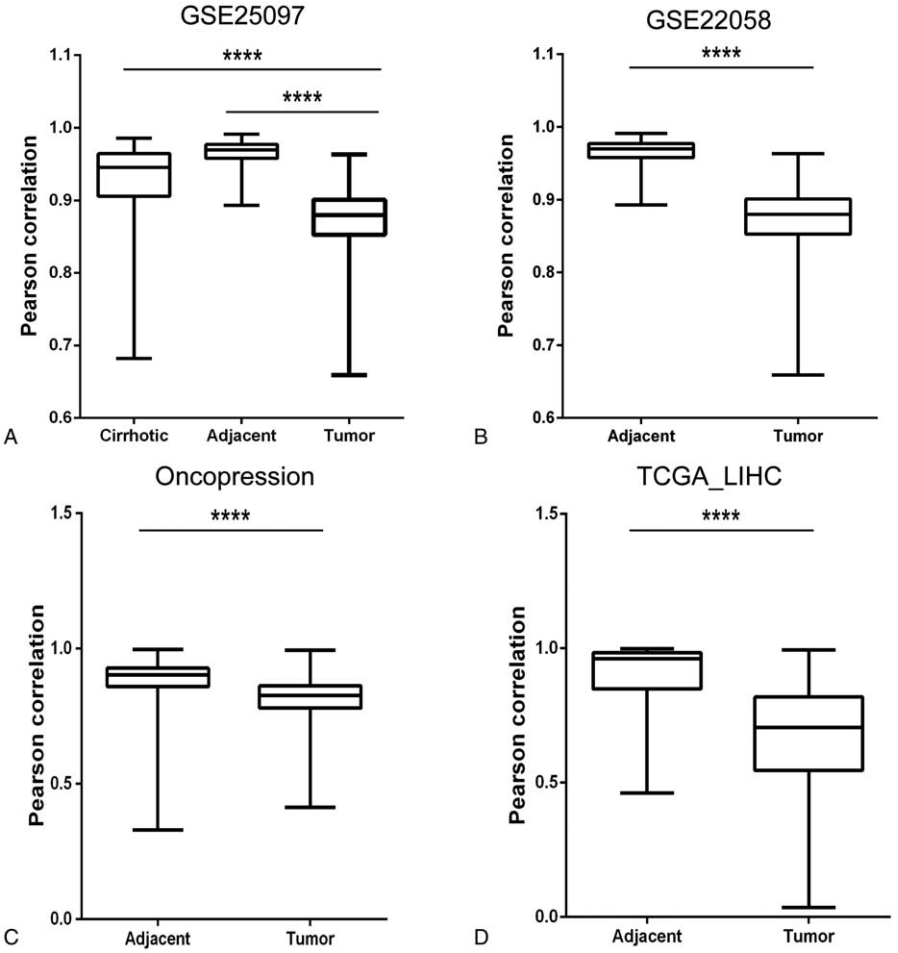

Figure 1. The Pearson correlation coefficients of each sample at the corresponding stage based on the mRNA expression values from GSE25097 (A), GSE22058 (B), Oncopression database (C), and TCGA_LIHC datasets (D). $^{*}P < .0001$.

3.3. Pathway enrichment analysis of DEGs among cirrhotic, adjacent nontumor, and tumor samples of HCC

Many of the existing pathway analysis methods are focused on either the number of DEGs in a pathway or on the correlation of genes in the pathway. Thus, information about complex gene interactions is disregarded. However, *SPIA* considers whether the DEGs found in a pathway have a meaningful impact within that pathway; thus, it addresses the topology of DEGs in pathways. Thus, in this study, we used *SPIA* to analyze the differences between aberrant pathways among cirrhotic, adjacent nontumor, and tumor tissues of HCC using the DEGs described above. A total of 59 KEGG pathways were identified as significantly perturbed in the progression from cirrhotic to adjacent nontumor (Table 1), and 40 KEGG pathways were significantly changed in the progression from adjacent nontumor to tumor (Table 2). Interestingly, most of the significantly perturbed pathways (50/59) during the transition from cirrhotic to adjacent nontumor were inhibited, whereas during the transition from adjacent nontumor to tumor, most of the pathways (29/39) were activated ($P < .0001$). Complement and coagulation cascades (hsa04610) and Antigen processing and presentation (hsa4612) were inhibited in the transitions from both cirrhotic to adjacent nontumor and adjacent nontumor to tumor; PPAR signaling pathway (hsa3320) was activated from both cirrhotic to adjacent nontumor and adjacent nontumor to tumor. ECM-receptor interaction (hsa4512), pathways in cancer (hsa5200), and insulin signaling pathway (hsa4910) were inhibited from cirrhotic to adjacent nontumor, but they were activated from adjacent nontumor to tumor. Focal adhesion (hsa4510) was activated from cirrhotic to adjacent nontumor, but it was inhibited from adjacent nontumor to tumor. The analysis also revealed that pathways involved in immune response were deregulated in the progression from cirrhotic to adjacent nontumor, including chemokine signaling pathway (hsa4062), natural killer cell mediated cytotoxicity (hsa4650), cytokine-cytokine receptor interaction (hsa4060), and Toll-like receptor signaling pathway (hsa4620). The tumor evolvement process was accompanied by an increase in the number of key oncogenic pathways associated with malignancy and metastatic spread, including p53 signaling pathway (hsa4155, activated), cell cycle (hsa4110, activated), transforming growth factor (TGF)-beta signaling pathway (hsa4350, activated), ErbB signaling pathway (hsa4012, activated), and NF-kappa B signaling pathway (hsa4064, inhibited). This functional analysis also suggested an active reaction of the adjacent nontumor related to the presence of the tumor or a more passive reaction induced by factors released from the tumor.

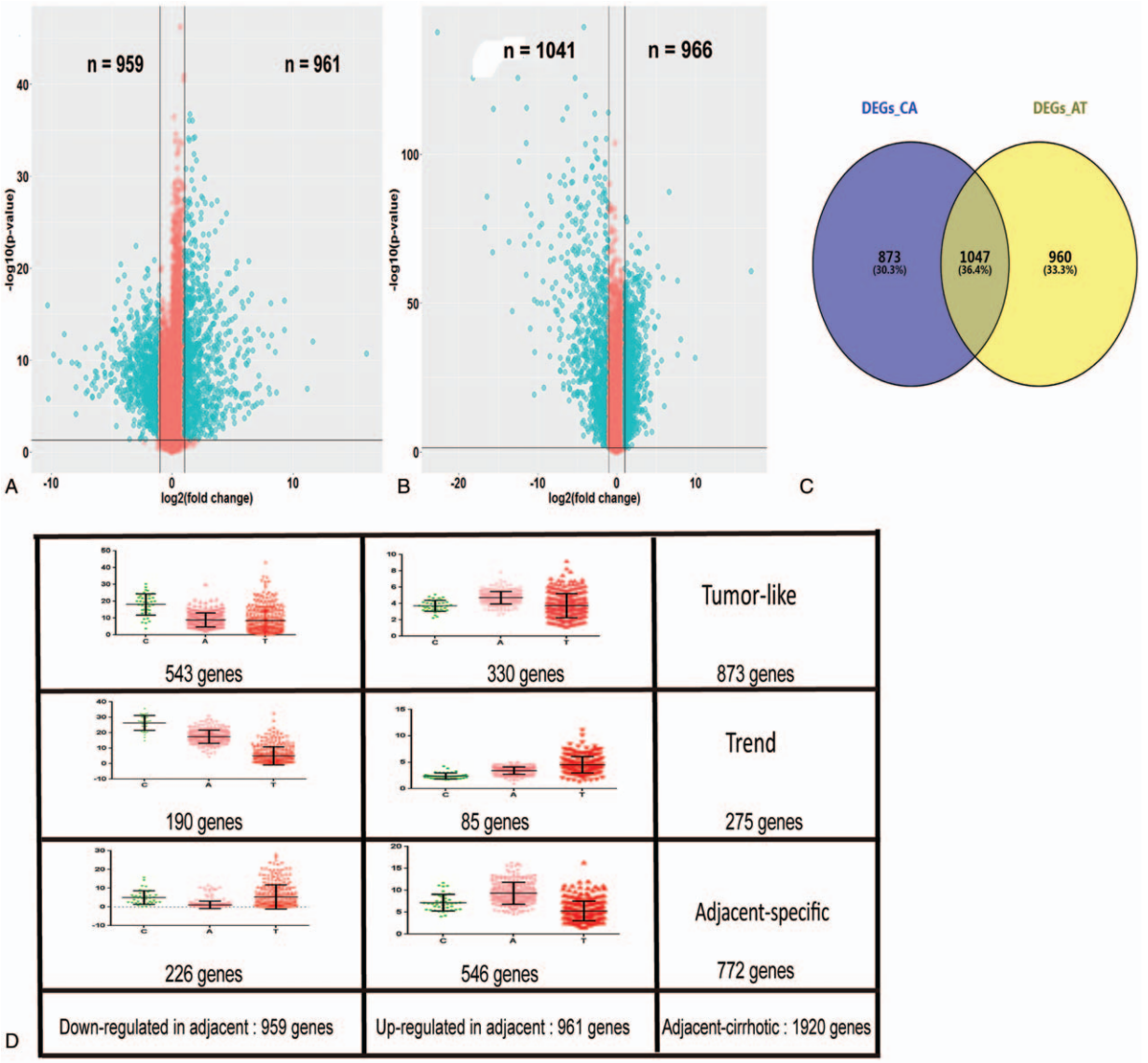

Figure 2. Differentially expressed genes (DEGs) among cirrhotic, adjacent nontumor, and tumor tissues of HCC. (A, B) Volcano plot of differential mRNA expression analysis. X-axis: log2-fold change; Y-axis: - log10 (FDR *P* value) of each probe; Vertical dotted lines: fold change ≥ 2 or ≤ 2; Horizontal dotted line: the significance cut-off (FDR = 5%). (A) There were 1920 genes identified as differentially expressed between cirrhotic and adjacent nontumor in GSE25097, including 961 upregulated genes and 959 downregulated genes. (B) Two thousand seven genes (1041 upregulated genes and 961 downregulated genes) were differentially expressed between adjacent nontumor and tumor in GSE25097. (C) Venn diagram showing the overlap of DEGs between cirrhotic and adjacent nontumor and tumor tissues of HCC. DEGs_CA were differentially expressed genes between cirrhotic and adjacent nontumor, and DEGs_AT were differentially expressed genes between adjacent nontumor and tumor of HCC. (D) Representative DEG patterns are displayed. DEG between adjacent and cirrhotic samples were classified as tumor-like, trend, and adjacent-specific genes.

3.4. FOXO1 and DCN (decorin) were underexpressed in HCC

Given that *FOXO1* was one of the top-ranked dysregulation transcription factors (TFs), playing a crucial role in the dynamic regulation of gene expression programs in tumors, and that *DCN* was the most downregulated gene in HCC progression from cirrhosis to adjacent nontumor and tumor, we focused on these 2 genes for further analysis in this study. *FOXO1* belonged to the adjacent-specific pattern. This classification highlighted its idiosyncratic role in the progression of HCC (Fig. 3A). There was a stepwise decrease of *DCN* mRNA expression levels in hepatocarcinogenesis from cirrhosis to adjacent nontumor and tumor tissues of HCC in microarray dataset GSE25097 (Fig. 3B). Furthermore, we validated their underexpression between adjacent nontumor and tumor using 3 other datasets (GSE22058, Oncopression, and TCGA_LIHC) from different platforms. Consistent with these findings, lower expression levels of *FOXO1* and *DCN* were seen in malignant samples of the validation cohort (Fig. 4A–F). All these results suggested that the underexpression of *FOXO1* and *DCN* was a common feature in HCC and that their dysregulation may be associated with tumorigenesis in HCC.

3.5. FOXO1 and DCN expression were prognostic indicators for patients with HCC

We also wanted to determine whether downregulated *FOXO1* and *DCN* levels could be used as prognostic indicators for HCC. Gene expression and clinical information from The Cancer Genome Atlas (TCGA) was collected for further investigation. A total of

Table 1

Significantly impacted pathways between cirrhotic and adjacent nontumor as determined by SPIA.

Status	Name	ID	pSize	pNDE	pGFdr
Activated	PPAR signaling pathway	3320	66	1.25E-15	1.42E-12
	Protein processing in endoplasmic reticulum	4141	151	3.06E-07	4.32E-05
	Osteoclast differentiation	4380	120	2.01E-06	0.000168608
	Adipocytokine signaling pathway	4920	66	3.73E-05	0.000477446
	Transcriptional misregulation in cancer	5202	166	0.001293694	0.008542133
	Type II diabetes mellitus	4930	46	0.000746657	0.009844659
	Carbohydrate digestion and absorption	4973	39	0.006384986	0.082097714
	Aldosterone-regulated sodium reabsorption	4960	37	0.091887967	0.087412368
	Neuroactive ligand-receptor interaction	4080	264	0.977787165	0.095337139
Inhibited	Complement and coagulation cascades	4610	66	4.19E-25	2.22E-21
	Staphylococcus aureus infection	5150	44	6.43E-15	3.03E-12
	Focal adhesion	4510	188	5.71E-10	3.32E-12
	ECM–receptor interaction	4512	80	1.05E-07	4.15E-10
	Pertussis	5133	69	1.57E-08	7.51E-08
	Pathogenic *Escherichia coli* infection	5130	43	1.24E-07	4.67E-07
	Bile secretion	4976	67	7.81E-09	9.35E-07
	Antigen processing and presentation	4612	55	3.97E-07	9.79E-07
	Leishmaniasis	5140	60	8.84E-08	3.32E-06
	Pathways in cancer	5200	309	4.65E-05	1.02E-05
	Regulation of actin cytoskeleton	4810	199	7.14E-07	1.19E-05
	Small cell lung cancer	5222	81	0.045596717	3.86E-05
	Tuberculosis	5152	158	0.000109774	0.000118125
	Leukocyte transendothelial migration	4670	103	6.71E-06	0.000162102
	Rheumatoid arthritis	5323	80	6.03E-06	0.000368915
	Toxoplasmosis	5145	123	1.02E-05	0.000368915
	Influenza A	5164	159	8.65E-06	0.000400371
	Prion diseases	5020	35	6.01E-06	0.000438155
	Chemokine signaling pathway	4062	174	0.00069752	0.000448442
	Amoebiasis	5146	99	9.52E-06	0.000477446
	Viral myocarditis	5416	57	1.48E-05	0.000760765
	Salmonella infection	5132	75	7.81E-05	0.000797777
	Bacterial invasion of epithelial cells	5100	62	5.24E-05	0.001060437
	Natural killer cell mediated cytotoxicity	4650	116	0.010323385	0.001699637
	Legionellosis	5134	51	4.78E-05	0.002183915
	Parkinson disease	5012	96	0.000384367	0.002434707
	HTLV-I infection	5166	240	0.000818551	0.003643844
	Insulin signaling pathway	4910	133	0.0001273	0.004903292
	Tight junction	4530	117	0.000259336	0.004903292
	Malaria	5144	47	0.000201796	0.004903292
	Chagas disease (American trypanosomiasis)	5142	98	0.000520637	0.005172963
	Prostate cancer	5215	81	0.001824529	0.009244886
	Alzheimer disease	5010	137	0.000508548	0.013631225
	Lysosome	4142	112	0.000611508	0.017335391
	Epithelial cell signaling in Helicobacter pylori infection	5120	65	0.001025904	0.018277346
	Shigellosis	5131	56	0.00176709	0.023407201
	Cytokine–cytokine receptor interaction	4060	243	0.084774545	0.027024436
	Gap junction	4540	83	0.054312893	0.030566225
	Hepatitis C	5160	122	0.002125553	0.037997457
	Epstein–Barr virus infection	5169	180	0.008571582	0.038669129
	Herpes simplex infection	5168	161	0.006091015	0.03977
	Toll-like receptor signaling pathway	4620	95	0.010629223	0.045695556
	Huntington disease	5016	151	0.002403594	0.046558925
	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	5412	69	0.058746986	0.049504699
	Systemic lupus erythematosus	5322	99	0.007692824	0.056782881
	Renal cell carcinoma	5211	63	0.013973267	0.069269248
	MAPK signaling pathway	4010	254	0.177507604	0.083010984
	Dilated cardiomyopathy	5414	86	0.066835773	0.093339512
	Glioma	5214	63	0.01820343	0.093339512
	Measles	5162	125	0.022908118	0.099741385

pGFdr = false discovery rate-adjusted global probability, pNDE = overrepresentation probability, pSize = pathway size.

Table 2

Significantly impacted pathways between adjacent nontumor and tumor as determined by SPIA.

Status	Name	ID	pSize	pNDE	pGFdr
Activated	PPAR signaling pathway	3320	66	2.74E-13	2.99E-10
	Bile secretion	4976	67	1.92E-08	7.09E-06
	Pathogenic Escherichia coli infection	5130	43	2.45E-07	1.30E-05
	Pertussis	5133	69	1.74E-07	4.50E-05
	ECM–receptor interaction	4512	80	4.12E-05	0.000249271
	Bacterial invasion of epithelial cells	5100	62	6.85E-06	0.000317131
	Insulin signaling pathway	4910	133	0.000273462	0.000470978
	p53 signaling pathway	4115	65	1.52E-05	0.000601282
	Protein processing in endoplasmic reticulum	4141	151	1.74E-05	0.001097691
	Pathways in cancer	5200	309	8.21E-05	0.001170439
	Cell cycle	4110	114	3.89E-05	0.002882713
	Focal adhesion	4510	188	0.000383626	0.005081632
	Adipocytokine signaling pathway	4920	66	0.000225152	0.006718334
	Thyroid cancer	5216	28	0.000558877	0.007882227
	Parkinson's disease	5012	96	0.000273576	0.008435032
	Transcriptional misregulation in cancer	5202	166	0.00496184	0.014245709
	TGF-beta signaling pathway	4350	80	0.05678399	0.027596232
	Chronic myeloid leukemia	5220	68	0.002705577	0.027596232
	Small cell lung cancer	5222	81	0.032155011	0.028052191
	Rheumatoid arthritis	5323	80	0.001007751	0.028052191
	Prostate cancer	5215	81	0.001181754	0.032750431
	HTLV-I infection	5166	240	0.001070874	0.034762572
	ErbB signaling pathway	4012	85	0.085597389	0.043444469
	Glioma	5214	63	0.008267628	0.043444469
	Chagas disease (American trypanosomiasis)	5142	98	0.002261424	0.043444469
	Leishmaniasis	5140	60	0.005260512	0.06165865
	Tight junction	4530	117	0.009314066	0.084253214
	Shigellosis	5131	56	0.0073537	0.086532876
	Epstein–Barr virus infection	5169	180	0.015384925	0.086532876
Inhibited	Complement and coagulation cascades	4610	66	8.76E-32	6.51E-30
	Staphylococcus aureus infection	5150	44	1.99E-11	1.40E-09
	Systemic lupus erythematosus	5322	99	0.000167758	5.45E-05
	Prion diseases	5020	35	1.72E-06	0.000150899
	Antigen processing and presentation	4612	55	3.70E-06	0.000256793
	Huntington disease	5016	151	0.000522774	0.027596232
	Viral myocarditis	5416	57	0.001097843	0.034009636
	Salmonella infection	5132	75	0.003086079	0.043444469
	NF-kappa B signaling pathway	4064	86	0.026467655	0.049012939
	Mineral absorption	4978	46	0.003629218	0.086532876
	Malaria	5144	47	0.004389869	0.096653767

pGFdr = false discovery rate-adjusted global probability, pNDE = overrepresentation probability, pSize = pathway size, TGF = transforming growth factor.

371 HCC patients were included, and their clinical characteristics are summarized in Table 3. *FOXO1* and *DCN* expression was remarkably negatively associated with pathologic T stage and tumor grade ($P < .05$). Both the low *FOXO1* expression group and the low *DCN* expression group had significantly poorer overall survival [*FOXO1*: $P = .0072$, hazard ratio: 0.6007, 95% confidence interval (95% CI): 0.4328–0.8514, Fig. 5A; *DCN*: $P = .0326$, hazard ratio: 0.8247, 95% CI: 0.5842–1.164, Fig. 5B]. The median survival period was 2116 days for the *FOXO1* high expression group, whereas it dropped to 1271 days in the *FOXO1* low expression group. The *DCN* low expression group had a reduced median survival period of 1397 days compared with a median survival period of 1694 days in the high expression group. These results indicated that *FOXO1* and *DCN* were beneficial factors for survival in HCC patients.

4. Discussion

The risk of HCC development is significantly increased among patients with advanced liver fibrosis caused by viral and nonviral etiologies, which then progressively evolves to cirrhosis. Although it is well-known that the prognosis of HCC patients is closely linked to levels of liver cirrhosis, the molecular mechanisms underlying the progression of liver cirrhosis to HCC remain unclear.

In this study, we found that the gene expression of cirrhotic samples and adjacent nontumor samples were surprisingly homogeneous; however, upon progression to HCC, the correlation between patients dramatically decreased, reflecting the well-recognized phenotypic heterogeneity of HCC. The high molecular heterogeneity of HCC may underscore the poor response to standard therapies in current clinical trials and the need for individualized treatment at progressed stages of HCC.

Previous studies have typically compared paired tumor and adjacent nontumor tissues, which can result in misleading interpretations. In this study, the inclusion of samples from cirrhotic tissues has allowed us to assess whether adjacent nontumor tissue from HCC patients differs from cirrhotic tissues due to tumor presence.

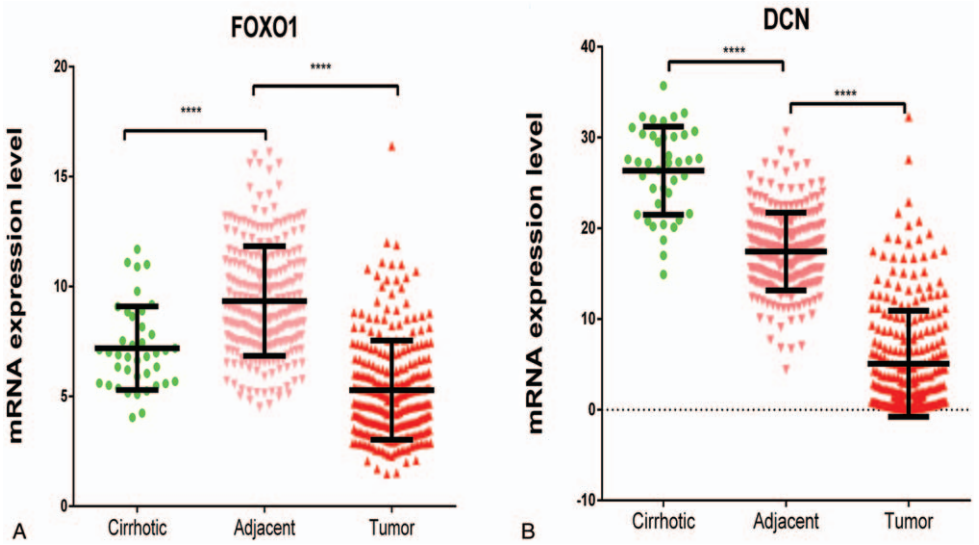

Figure 3. Gene expression levels of *FOXO1* (A) and *DCN* (B) in cirrhotic (green), adjacent nontumor (pink), and tumor tissue (red) of HCC from GSE25097. $^{*}P < .0001$.

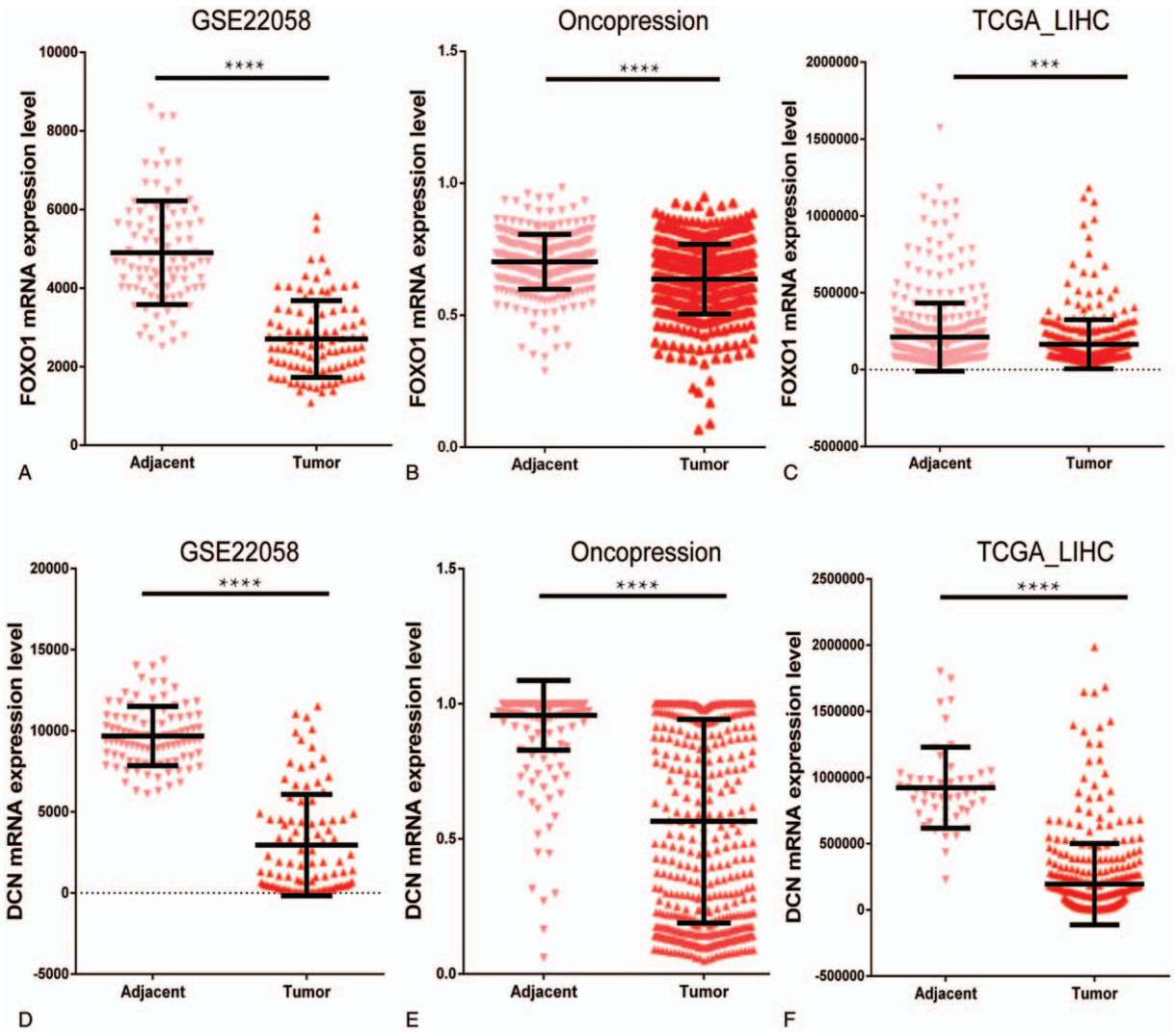

Figure 4. Validation of differential expression of *FOXO1* and *DCN* in validation datasets. (A–C) Reduced expression of FOXO1 in HCC. (D–F). Reduced expression of DCN in HCC. $^{*}P < .0001$, $^{\dagger}P = .001$.

Table 3

Clinical characteristics and correlations with the expression of *FOXO1* and *DCN*.

Characteristic	n = 371	*FOXO1* Low (185)	*FOXO1* High (186)	*P*	*DCN* Low (185)	*DCN* High (186)	*P*
Age, y							
≥ 65	149	73	76	.832	78	71	.525
< 65	221	112	109		107	114	
Gender							
Male	250	119	131	.253	124	126	.912
Female	121	66	55		61	60	
Race							
Asian	158	86	72	.159	88	70	.065
White	184	85	99		83	101	
Clinical stage							
I and II	257	125	132	.235	129	128	.626
III and IV	90	51	39		48	42	
Pathologic T stage							
T1	181	80	101	**.033**	80	101	**.037**
T2 + T3 + T4	188	105	83		104	84	
Grade							
G1 + G2	232	104	128	**.008**	104	128	**.017**
G3 + G4	134	80	54		78	56	

The significance of bold values $P < .05$.

Changes in gene expression and cellular signaling pathways cause significant changes in the transition from liver cirrhosis to HCC. A total of 1920 and 2007 genes were specifically associated with progression from cirrhotic to adjacent nontumor and from adjacent nontumor to tumor, respectively. DEGs between cirrhotic and adjacent nontumor tissues can be grouped into 3 altered patterns based on their level of expression in tumor tissues: "Tumor-like," "Trend," and "Adjacent-specific." Our results showed that each gene expression pattern has different functions. Moreover, *SPIA* was used for functional enrichment analysis of DEGs among cirrhotic, adjacent nontumor, and tumor samples of HCC. Most of the significantly perturbed pathways between cirrhotic to adjacent nontumor were inhibited, but between adjacent nontumor to tumor, most of these pathways were activated. Pathway analysis of cirrhotic to adjacent nontumor also revealed the deregulation of signaling pathways involved the immune response. Consistently, the majority of pathways activated during the conversion from adjacent nontumor to tumor conferred malignant and invasive properties, including p53 signaling pathway, cell cycle, TGF-beta signaling pathway, ErbB signaling pathway, and NF-kappa B signaling pathway. This analysis emphasized that the acquisition of malignant traits is a relatively late event. Our results also suggested that adjacent tumor tissue is abnormal. In fact, studies that only compare tumor and adjacent nontumor may miss good cancer biomarker candidates because many genes are deregulated in adjacent tumor tissue, mimicking tumor expression.

Another important finding of this study is identifying *FOXO1* and *DCN* preferentially upregulation in adjacent nontumor tissue and downregulation in tumor tissue of HCC. *FOXO1* belonged to the adjacent-specific gene pattern, and *DCN* belonged to the Trend gene pattern. The TF *FOXO1* is characterized by the forkhead DNA-binding domain, and its aberration influences multiple cellular functions, including apoptosis, cell cycle control, DNA damage repair, glucose metabolism, carcinogenesis, and tumor immunity.[14] In addition, as a critical modulator of many important stress pathways, the *FOXO1* TF, may regulate adaptation of the liver to stress.[15] Although *FOXO1* dysregulation has been observed in several human cancers,[16–18] the study of its expression in liver cirrhosis has been limited. It is well-known that hepatic stellate cells (HSCs) play a crucial role in the liver fibrotic response, as their activation, transdifferentiation, and proliferation are key steps in liver fibrosis. Intriguingly, *FOXO1* was reported to participate in the proliferation and transdifferentiation process of HSCs, which were enhanced by transcriptionally inactive *FOXO1*. In contrast, active *FOXO1* inhibited HSC proliferation by inducing cell cycle arrest to accumulate cells in the G0/G1 phase.[19] It is consistent with our results that *FOXO1* gene expression was downregulated in patients with cirrhosis. Our finding shows that the functional role of *FOXO1* may be linked to hepatocarcinogenesis at the stage in which cirrhosis progresses into tumor in HCC.

Another potential prognostic factor for HCC is *DCN*. The *DCN* gene encodes a member of the small leucine-rich proteoglycan

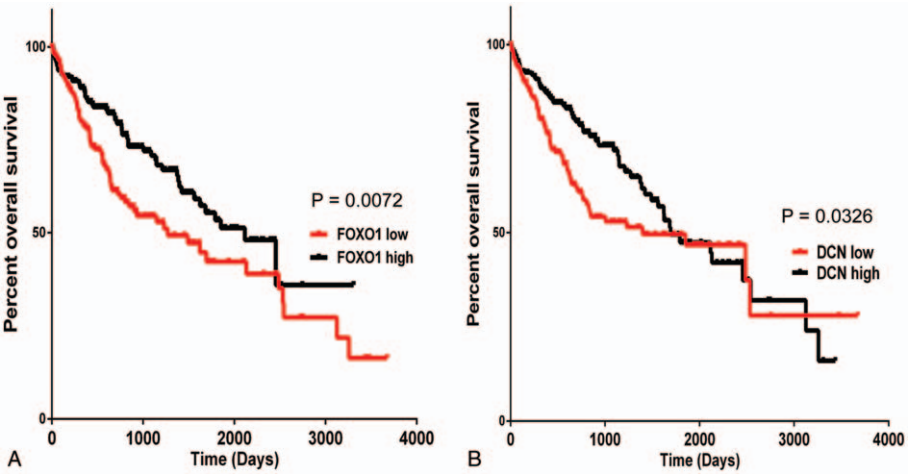

Figure 5. Overall survival curves based on *FOXO1* and *DCN* expression. (A) The survival curves of the *FOXO1* high expression group and the *FOXO*1 low expression group. *FOXO1* low expression was associated with poor overall survival ($P = .0072$). (B) The survival curves of the *DCN* high expression group and the *DCN* low expression group. *DCN* low expression was associated with poor overall survival ($P = .0326$). Patients were divided into low and high expression groups according to the gene's median mRNA expression value.

family of proteins regulating collagen fibrillogenesis during liver disease development. Previous studies showed that a low quantity of decorin is present in normal healthy liver, whereas the amount of decorin significantly increases during fibrogenesis.[20] Consistent with these findings, our results indicated that higher expression levels of *DCN* were seen in the progression of liver cirrhosis. Furthermore, the evidence suggests that decorin could act as a tumor repressor in a variety of cancers. An early study has shown that extopic expression of decorin could suppress the generalized growth of various neoplastic cells.[21] Regarding liver tumors, decorin inhibits the proliferation of HepG2 and Huh-7 hepatoma cell lines.[22] In addition, it has been shown that although decorin-deficient mice are not associated with the development of spontaneous tumors, their tissues are permissive for tumorigenesis.[23] In line with these reports, this study indicated that *DCN* mRNA expression is downregulated in HCC, as its lack is accompanied with significantly higher HCC prevalence.

5. Conclusion

We have provided the first evidence of the molecular mechanism in the transition from liver cirrhosis to HCC using bioinformatics techniques. In addition, we also found that *FOXO1* and *DCN* are underexpressed in HCC tissues and that their downregulation may be indicative of poor survival rates; furthermore, they could have a potential role as prognostic markers in HCC patients. Functional studies are needed to reveal the molecular mechanisms of *FOXO1* and *DCN* in HCC and their role in prognosis and therapeutic targets.

Author contributions

Data curation: Yu-Kui Shang, Fanni Li.
Formal analysis: Zi-Ling Wang.
Funding acquisition: Fanni Li, Huijie Bian, Zhi-Nan Chen.
Investigation: Zi-Ling Wang, Huijie Bian.
Methodology: Yu-Kui Shang, Fanni Li, Zekun Liu, Zi-Ling Wang.
Project administration: Zhi-Nan Chen.
Resources: Yi Zhang, Zekun Liu.
Software: Yi Zhang.
Validation: Yi Zhang, Huijie Bian, Zhi-Nan Chen.
Visualization: Yi Zhang, Huijie Bian.
Writing – original draft: Yu-Kui Shang, Fanni Li, Huijie Bian.
Writing – review & editing: Huijie Bian, Zhi-Nan Chen.

References

[1] Mortality GBD, Causes of Death CGlobal, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;385:117–71.
[2] Velazquez RF, Rodriguez M, Navascues CA, et al. Prospective analysis of risk factors for hepatocellular carcinoma in patients with liver cirrhosis. Hepatology 2003;37:520–7.
[3] Bruix J, Sherman M. American Association for the Study of Liver D. Management of hepatocellular carcinoma: an update. Hepatology 2011;53:1020–2.
[4] Dhir M, Melin AA, Douaiher J, et al. A review and update of treatment options and controversies in the management of hepatocellular carcinoma. Ann Surg 2016;263:1112–25.
[5] Buendia MA, Neuveut C. Hepatocellular carcinoma. Cold Spring Harb Perspect Med 2015;5:a021444.
[6] El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology 2007;132:2557–76.
[7] Moeini A, Cornella H, Villanueva A. Emerging signaling pathways in hepatocellular carcinoma. Liver Cancer 2012;1:83–93.
[8] Tung EK, Mak CK, Fatima S, et al. Clinicopathological and prognostic significance of serum and tissue Dickkopf-1 levels in human hepatocellular carcinoma. Liver Int 2011;31:1494–504.
[9] Burchard J, Zhang C, Liu AM, et al. microRNA-122 as a regulator of mitochondrial metabolic gene network in hepatocellular carcinoma. Mol Syst Biol 2010;6:402.
[10] Lee J, Choi C. Oncopression: gene expression compendium for cancer with matched normal tissues. Bioinformatics 2017;33:2068–70.
[11] Foroushani AB, Brinkman FS, Lynn DJ. Pathway-GPS and SIGORA: identifying relevant pathways based on the over-representation of their gene-pair signatures. PeerJ 2013;1:e229.
[12] Tarca AL, Draghici S, Khatri P, et al. A novel signaling pathway impact analysis. Bioinformatics 2009;25:75–82.
[13] Sanz-Pamplona R, Berenguer A, Cordero D, et al. Aberrant gene expression in mucosa adjacent to tumor reveals a molecular crosstalk in colon cancer. Mol Cancer 2014;13:46.
[14] Kim CG, Lee H, Gupta N, et al. Role of Forkhead Box Class O proteins in cancer progression and metastasis. Semin Cancer Biol 2017;[Epub ahead of print].
[15] Milkiewicz M, Kopycinska J, Kempinska-Podhorodecka A, et al. Ursodeoxycholic acid influences the expression of p27kip1 but not FoxO1 in patients with non-cirrhotic primary biliary cirrhosis. J Immunol Res 2014;2014:921285.
[16] Park J, Choi Y, Ko YS, et al. FOXO1 suppression is a determinant of acquired lapatinib-resistance in HER2-positive gastric cancer cells through MET upregulation. Cancer Res Treat 2018;50:239–54.
[17] Ushmorov A, Wirth T. FOXO in B-cell lymphopoiesis and B cell neoplasia. Semin Cancer Biol 2017;[Epub ahead of print].
[18] Zang Y, Wang T, Pan J, et al. miR-215 promotes cell migration and invasion of gastric cancer cell lines by targeting FOXO1. Neoplasma 2017;64:579–87.
[19] Adachi M, Osawa Y, Uchinami H, et al. The forkhead transcription factor FoxO1 regulates proliferation and transdifferentiation of hepatic stellate cells. Gastroenterology 2007;132:1434–46.
[20] Baghy K, Dezso K, Laszlo V, et al. Ablation of the decorin gene enhances experimental hepatic fibrosis and impairs hepatic healing in mice. Lab Invest 2011;91:439–51.
[21] Santra M, Mann DM, Mercer EW, et al. Ectopic expression of decorin protein core causes a generalized growth suppression in neoplastic cells of various histogenetic origin and requires endogenous p21, an inhibitor of cyclin-dependent kinases. J Clin Invest 1997;100:149–57.
[22] Zhang Y, Wang Y, Du Z, et al. Recombinant human decorin suppresses liver HepG2 carcinoma cells by p21 upregulation. Onco Targets Ther 2012;5:143–52.
[23] Baghy K, Horvath Z, Regos E, et al. Decorin interferes with platelet-derived growth factor receptor signaling in experimental hepatocarcinogenesis. FEBS J 2013;280:2150–64.