RESEARCH ARTICLE

Bicluster and Pathway Enrichment Analysis of HCV-induced Cirrhosis and Hepatocellular Carcinoma

Peng Cheng[1&], You Cheng[2&], Mei X Su[1], Dong Li[1], Guo Z Zhao[1], Hui Gao[1], Yan Li[1], Jie Y Zhu[1], Hua Li[1], Tao Zhang[1]*

Abstract

    <u>Background</u>: Hepatocellular carcinoma (HCC) is the sixth most common cancer worldwide and the most common form of liver cancer. However, while it is associated frequently with hepatitis C virus (HCV) there is only an elementary understanding of its molecular pathogenesis. <u>Methods</u>: To gain insight into the molecular mechanisms of HCV-induced hepatocarcinogenesis, we performed microarray analysis on 75 surgical liver samples from 48 HCV-infected patients. <u>Results</u>: There were 395 differentially expressed geness between cirrhotic samples and HCC samples. Of these, 125 genes were up-regulated and 270 genes were down-regulated. We performed pathway enrichment analysis and screened as described previously. <u>Conclusions</u>: The differentially expressed genes might be involved in hepatocarcinogenesis through upregulating the pathways of ECM-receptor interaction, focal adhesion, cell adhesion molecules and other cancer-related pathways, and downregulating the pathways of "complement and coagulation cascades". We hope our results could aid in seeking of therapeutic targets for HCV-induced hepatocellular carcinoma.

Keywords: Bicluster analysis - pathway enrichment - cirrhosis - hepatocellular carcinoma

Introduction

Hepatocellular carcinoma is the sixth most common cancer worldwide and the most common form of liver cancer, being responsible for 80% of the primary malignant liver tumors in adults (Avila et al., 2006; Ferlay et al., 2010). In 2008, an estimated 748,000 new cases of liver cancer occurred and approximately 696,000 people died of this cancer worldwide (Abedi-Ardekani et al., 2011). This disease is most prevalent in Eastern and Southeastern Asia, and Middle Africa, with more than half of the patients being reported from China (Avila et al., 2006). HCC is the third most common cancer in China with the age-standardized rate of 37.4 and 34.1 per 100,000 person-years in males and females, respectively (Ferlay et al., 2010).

Much is known about the development and causes of HCC. The risk factors for liver carcinogenesis include chronic infections with hepatitis B (HBV) and C (HCV) viruses, chronic alcohol consumption (Donato et al., 2002), consumption of aflatoxin B1 (AFB1) (Soini et al., 1996), contaminated food (Abedi-Ardekani et al., 2011) and type 2 Diabetes (Wideroff et al., 1997; El-Serag et al., 2006). Most HCC is thought to be associated with either chronic hepatitis C virus (HCV) or hepatitis virus (HBV) infection (Barazani et al., 2007). HCV causes a chronic

infection in 70-80% of infected individuals, 10–20% of which develop liver fibrosis and cirrhosis within two decades (Befeler and Di Bisceglie, 2002). Once cirrhosis is pronounced, HCC develops at a rate between 1.5 and 8% per year (Marcellin, 1999; Degos et al., 2000; Bruix et al., 2001).

Current evidence suggests that the early stages in HCC development are characterized by certain common traits governed by both genetic and epigenetic mechanisms. A variety of genomic studies have identified potential biomarkers for detection of early HCC (Marrero and Lok, 2004), such as GPC3 (Capurro et al., 2003), TERT, STK15, PLA2 (Smith et al., 2003) and HSP70 (Chuma et al., 2003). Most of the protein markers used for histopathological analysis are associated with cell proliferation and resistance to cell death (Hui et al., 1998; Thorgeirsson and Grisham, 2002). However, the process of HCV-induced carcinogenesis is still poorly understood. Therefore, it is important to better understand the roles of deregulated genes in the progression from healthy liver to cirrhotic and then hepatocellular carcinoma.

Factor Analysis for Bicluster Acquisition (FABIA) is based on a multiplicative model, which accounts for linear dependencies between gene expression and conditions, and also captures heavy-tailed distributions as observed in real-world transcriptomic data. The generative framework

[1]*Department of Oncology, PLA general hospital of Chengdu Commond, Clinical Medical College of The Third Medical Military University, Chengdu,* [2]*Department of Otolaryngology–Head and Neck Surgery, PLA General Hospital of Nanjing Commond, Clinical Medical College of Nanjing University, Nanjing, China* &Equal contributors *For correspondence: zhangt46@gmail.com*

allows to utilize well-founded model selection methods and to apply Bayesian techniques (Hochreiter et al., 2010). In this present study, we performed bicluster analysis using fabia to classify the genes in the progression and identified the differentially expressed genes among healthy liver samples, cirrhotic samples and HCC samples. Furthermore, we performed pathway enrichment analysis using BiNGO to identify the significant pathways related to the DEGs. We sought to explore the underlying molecular mechanism of the progression of HCV-induced hepatocarcinogenesis.

Materials and Methods

The transcription profile of GSE6764 was downloaded from GEO (http://www.ncbi.nlm.nih.gov/geo/), a public functional genomics data repository, which are based on the Affymetrix GPL570 platform data (Affymetrix Human Genome U133 plus 2.0 Array). Total 75 samples including 65 samples of HCV- induced HCC and 10 samples of normal liver were used to analysis. Sixty-five samples were obtained from 38 patients with HCV infection representing 13 samples from cirrhotic tissue, 17 dysplastic nodules, and 35 HCCs.

Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/) is a collection of online databases dealing with genomes, enzymatic pathways, and biological chemicals (Kanehisa, 2002). The 'pathway' database records networks of molecular interactions in the cells, and variants of them specific to particular organisms. FABIA is developed to find biclusters that have correlated rows and columns. More precisely, the rows in the bicluster need to be only correlated across the columns of the bicluster and vice versa. Bicluster analysis with the FABIA using the fabia package in Bioconductor (Gentleman et al., 2004) was performed to find the enriched biclusters. The package allows to extract biclusters from data sets based on a generative model according to the FABIA method (Hochreiter et al., 2010).

Gene Ontology (GO) analysis has become a commonly used approach for functional studies of large-scale genomic or transcriptomic data (Hulsegge et al., 2009).

BiNGO (Maere et al., 2005) is an open-source Java tool to determine which Gene Ontology (GO) terms are significantly overrepresented in a set of genes. We used the BiNGO to identify over-represented GO categories in biological process. Moreover, hypergeometric tests were used for statistical analysis and the BH method (Benjamini, 1995) False Discovery Rate (FDR) procedure was used for the multiple testing correction with the hypergeometric distribution. The FDR less than 0.05 was chosen as the threshold. We performed GO enrichment analysis for each bicluster respectively, and selected the most significant GO term of each bicluster as its GO term annotations.

For the GSE6764 dataset, the SAM method (Li and Tibshirani, 2011) (http://www-stat.stanford.edu/~tibs/SAM/) was used to identify DEGs. The DEGs only with fold change value larger than 2 were selected.

Results

For dataset GSE6764, we performed data preprocessing using R and Bioconductor. After deleting the unqualified probes, we got 13127 probes to perform bicluster analysis with fabia package. Total 19 biclusters were obtained (Table 1). GO enrichment analysis was performed for each bicluster respectively, and selected the most significant GO term of each bicluster as its GO term annotations. As shown in table 1, several biological processes were enriched, such as "cellular metabolic process" (bicluster7, 8, 10, 13, 14, 16), "immune system process" (15, 19), "response to chemical stimulus" (12, 18), "calcium-dependent cell-cell adhesion" (5) and so on.

In order to get differentially expressed genes, we obtained publicly available microarray dataset GSE6764 from GEO. For the GSE6764 dataset, the SAM method was used to identify DEGs. After microarray analysis, the differentially expressed genes with the fold change value larger than 2 were selected. We got 501 DEGs between healthy liver samples and cirrhotic samples, with 81% (408) up-regulation and 19% (93) down-regulation. Total

**Table 1. GO Enrichment Analysis in the Total 22 Biclusters**

bicluster ID	number of conditions	number of genes	GO term	p-value	FDR
1	14	62	developmental induction	2.84E-06	2.62E-03
2	17	37	--	--	--
3	19	150	--	--	--
4	9	42	calcium-dependent cell-cell adhesion	1.40E-07	1.20E-04
5	23	104	macromolecule localization	3.78E-05	2.24E-02
6	15	62	nuclear division	1.67E-06	5.45E-04
7	18	266	cellular metabolic process	6.69E-21	1.68E-17
8	13	498	cellular metabolic process	9.90E-09	3.21E-05
9	15	253	organelle organization	3.48E-11	9.15E-08
10	8	1504	cellular metabolic process	9.36E-52	5.37E-48
11	9	273	cellular process	2.45E-07	4.00E-04
12	10	362	response to chemical stimulus	3.15E-19	1.06E-15
13	18	162	nitrogen compound metabolic process	3.10E-08	6.24E-05
14	10	560	cellular catabolic process	9.01E-12	3.17E-08
15	21	221	antigen processing and presentation of peptide antigen	1.70E-07	3.45E-04
16	11	779	cellular metabolic process	4.07E-43	1.88E-39
17	23	49	cellular respiration	2.54E-05	2.85E-02
18	7	442	response to chemical stimulus	7.75E-20	2.68E-16
19	20	405	immune system process	9.96E-32	3.22E-28

Table 2. GO Enrichment Analysis of the Overlapping DEGs

GO-ID	p-value	FDR	Description
7159	4.69E-05	1.43E-02	leukocyte cell-cell adhesion
2000056	2.53E-04	2.56E-02	regulation of Wnt receptor signaling pathway involved in digestive tract morphogenesis
2000057	2.53E-04	2.56E-02	negative regulation of Wnt receptor signaling pathway involved in digestive tract morphogenesis
46794	5.06E-04	3.07E-02	virion transport
46795	5.06E-04	3.07E-02	intracellular virion transport
46968	7.58E-04	3.84E-02	peptide antigen transport

Table 3. Pathway Enrichment Analysis of the Up-regulated Genes Between Healthy Liver Samples and Cirrhotic Samples

Pathway Term	P-Value	Bonferroni	Benjamini	FDR
hsa04512:ECM-receptor interaction	2.82E-11	3.75E-09	3.75E-09	3.30E-08
hsa04510:Focal adhesion	5.85E-10	7.78E-08	3.89E-08	6.84E-07
hsa04514:Cell adhesion molecules (CAMs)	1.51E-08	2.01E-06	6.71E-07	1.77E-05
hsa04612:Antigen processing and presentation	3.92E-06	5.22E-04	1.30E-04	0.004589
hsa05340:Primary immunodeficiency	1.78E-05	0.002366	4.74E-04	0.020822
hsa05222:Small cell lung cancer	2.46E-05	0.00327	5.46E-04	0.02879
hsa04670:Leukocyte transendothelial migration	4.44E-05	0.005886	8.43E-04	0.051888
hsa04940:Type I diabetes mellitus	7.27E-05	0.009618	0.001207	0.084933
hsa05330:Allograft rejection	1.82E-04	0.023889	0.002683	0.212348
hsa05332:Graft-versus-host disease	3.07E-04	0.040011	0.004075	0.358344
hsa05200:Pathways in cancer	3.08E-04	0.040136	0.003717	0.359488
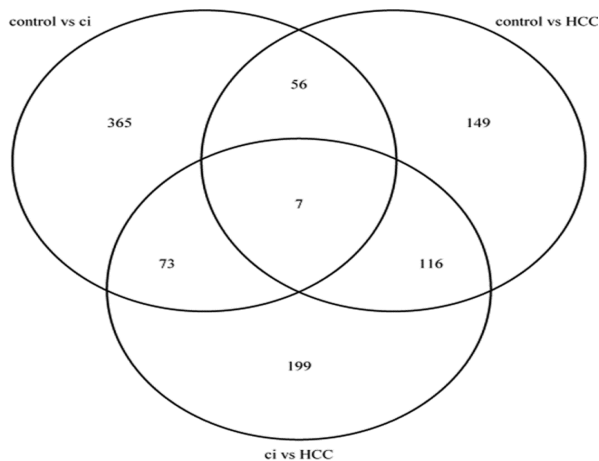
Figure 1. VENN Graph Display the Information of Our Data. Total 7 overlapping genes were selected as DEGs from GSE6764. The 'control' represents for the healthy liver samples, 'ci' represents for cirrhotic samples and 'HCC' represents for hepatocellular carcinoma samples

328 DEGs with 31% (102) up-regulation and 69% (226) down-regulation were identified between healthy liver samples and HCC samples and 395 DEGs with 32% (125) up-regulation and 68% (270) down-regulation were identified between cirrhotic samples and HCC samples (Figure 1). There were 7 overlapping genes which expressed differentially in all the three type samples. The 7 DEGs are CLEC4M, MAP2, APOF, NAT2, CHST4, PROM1, and SFRP5. We held the opinion that these 7 genes play an important role in the progression from healthy liver to cirrhotic and then HCC. To identify the functional annotation of the 7 DEGs, we performed GO enrichment analysis, and 6 biological processes were enriched, including "leukocyte cell-cell adhesion", "virion transport", "peptide antigen transport" and "regulation of Wnt receptor signaling pathway" (Table 2).

For the DEGs of GSE6764, we performed pathway enrichment analysis to find the most significant pathways related to the progression of HCV-induced hepatocarcinogenesis.

There were 501 DEGs between healthy liver samples and cirrhotic samples. Of these, 408 genes were up-regulated and 93 genes were down-regulated. We performed pathway enrichment analysis of the up-regulated and down-regulated genes, respectively. After p-value correction, the pathways whose Bonferroni, Benjamini, and FDR at least two values less than 0.05 were selected as the significant pathways.Finally, we got 11 enriched pathways in up-regulated genes and 0 enriched pathways in down-regulated genes (Table 3). The enriched pathways include "ECM-receptor interaction", "focal adhesion", "cell adhesion molecules (CAMs)" and so on. There were 395 DEGs between cirrhotic samples and HCC samples. Of these, 125 genes were up-regulated and 270 genes were down-regulated. We performed pathway enrichment analysis and screened as described previously. Finally, we got none enriched pathways in up-regulated genes and 1 enriched pathway in down-regulated genes. The enriched pathway was "complement and coagulation cascades"

Discussion

Hepatocellular carcinoma (HCC) is a common primary cancer associated frequently with hepatitis C virus (HCV) (Smith et al., 2003). To gain insight into the molecular mechanism of HCV-induced hepatocarcinogenesis, we performed bicluster analysis of the cDNA microarray obtained from GEO on 75 surgical liver samples from 48 HCV-infected patients, and identified the differentially expressed genes in the progressive stages of hepatocarcinogenesis. Furthermore, we performed pathway enrichment analysis of the DEGs.

In our result of bicluster analysis of the gene expression profile, we can find that total 19 clusters were enriched. The "cellular metabolic process" was the most significant biological process. Cellular metabolism is the set of chemical reactions that occur in living organisms in

order to maintain life. This process allows organisms to grow and reproduce, maintain their structures, and respond to environmental changes. Cells from some tumors use an altered metabolic pattern compared with that of normal differentiated cells in the body (Levine and Puzio-Kuter, 2010). Otto Warburg postulated that change in metabolism is the fundamental cause of cancer (Warburg, 1956).

The other significant biological process was "immune system process". This reflects the transition from normal to carcinoma, where the liver function becomes impaired and extracellular matrix deposition increases (Bieche et al., 2005). The immune system can specifically identify and eliminate tumor cells on the basis of their expression of tumor-specific antigens or molecules induced by cellular stress (Swann and Smyth, 2007). However, HCV has evolved a variety of mechanisms to establish persistent infections and evade the host immune response (Block et al., 2003). Previous studies have shown that expression of HCV core activate NF-κβ, a transcription factor that is involved in regulating the immune response (Zhu et al., 2001). Hepatocytes from patients chronically infected with HCV show elevated levels of NF-κβ protein and increased NF-κβ DNA binding activity (Tai et al., 2000). It is suggested by Timothy et al. that HCV core protein promotes persistent infection by downregulating the host immune system.

The "response to chemical stimulus" was also observed in our bicluster analysis (bicluster 12 and bicluster 18). We checked the condition12 and condition18 and found that these conditions only contain the samples from early, middle and advanced stage of HCC. Based upon these observations, we hypothesized that these biclusters were result from the chemotherapy to the tumor.

Total 7 genes were identified as DEGs of GSE6764, including CLEC4M, MAP2, APOF, NAT2, CHST4, PROM1, and SFRP5. This result suggesting that these genes play an important role in the progression of this disease, especially the CLEC4M and CHST4.

CLEC4M(C-type lectin domain family 4 member M), also been designated as CD209L, encodes a transmembrane receptor L-SIGN. Liver/lymph node-specific intercellular adhesion molecule-3-grabbing integrin (L-SIGN) is a calcium-dependent lectin expressed mainly on endothelial cells of liver and lymph nodes (Koppel et al., 2005). Jason et al. have demonstrated that L-SIGN is a liver-specific receptor for HCV, and may play important roles in HCV infection and immunity (Gardner et al., 2003; Cormier et al., 2004). L-SIGN may establish cellular interactions with T cells (Bashirova et al., 2001) and enable activated T cells to recirculate to the liver or to the lymph nodes through interactions with L-SIGN (Cocquerel et al., 2006). L-SIGN and DC-SIGN may rather contribute to the establishment or persistence of infection, both by the capture and delivery of HCV to the liver and by modulating dendritic-cell functions as suggested by Cormier et al. (2004) and Lozach et al. (2004).

CHST4 (Carbohydrate sulfotransferase) encodes an N-acetylglucosamine 6-O sulfotransferase. This protein is involved in the modification of glycan structures on ligands of the lymphocyte homing receptor L-selectin

and plays a central role in lymphocyte trafficking during chronic inflammation (Wu et al., 2011).

Of the 11 up-regulated genes related pathways, the "ECM- receptor interaction" was the most significant. The extracellular matrix (ECM) consists of a complex mixture of structural and functional macromolecules and serves an important role in tissue and organ morphogenesis and in the maintenance of cell and tissue structure and function. Cirrhotic liver contains approximately six times more ECM overall than normal liver (Kumar and Sarin, 2007). ECM can directly influence the function of surrounding cells through interaction with cell surface receptors, including integrins and nonintegrin matrix receptors. The interaction of ECM with integrins is thought to be an important factor in cell-cell and cell-substrate adhesions in tumors, tumor invasion and metastases (Jaskiewicz et al., 1993). ECM can also indirectly affect cell function via release of soluble cytokines, which in turn are controlled by local metalloproteinases (Seyer et al., 1977; Jaskiewicz et al., 1993). This pathway is not only related to the transition from normal to cirrhosis, and also related to the hepatocarcinogenesis of cirrhosis.

Only one pathway were enriched in the down-regulated genes between cirrhosis and HCC, that is "complement and coagulation cascades". This result suggesting that the activity of blood coagulation factor decreased significantly in the progression from cirrhosis to HCC, which may because of the damage of normal liver cells in the progression. Previous studies suggested that the level of coagulation factor can be used to diagnose cirrhosis and primary HCC, and is an indicator of prognosis of liver disease (Amitrano et al., 2002). Therefore, further analysis and investigation of this pathway will be of clinical significance to understand this carcinogenesis.

In conclusion, we have used bioinformatics methods to analyse the potential molecular mechanism of HCV-induced hepatocarcinogenesis. Our analysis indicated several differentially expressed genes might play crucial roles in hepatocarcinogenesis, including CLEC4M, MAP2, APOF, NAT2, CHST4, PROM1, and SFRP5. Furthermore, our results predicted the differentially expressed genes might be involved in hepatocarcinogenesis through upregulating the pathways of ECM-receptor interaction, focal adhesion, cell adhesion molecules and other cancer-related pathways, and downregulating the pathways of "complement and coagulation cascades". We expect numerous advanced researches in HCV-induced hepatocellular carcinoma in the coming years based on our Meta-analysis.

Acknowledgements

References

Abedi-Ardekani B, Gouas D, Villar S, et al (2011). TP53 Mutations and HBX Status Analysis in Hepatocellular Carcinomas from Iran: Evidence for Lack of Association between HBV Genotype D and TP53 R249S Mutations. *Hepat Res Treat*, **2011**, 475965.

Amitrano L, Guardascione MA, Brancaccio V, Balzano A (2002). Coagulation disorders in liver disease. *Semin Liver Dis*, **22**, 83-96.

Avila MA, Berasain C, Sangro B, Prieto J (2006). New therapies for hepatocellular carcinoma. *Oncogene*, **25**, 3866-84.

Barazani Y, Hiatt JR, Tong MJ, Busuttil RW (2007). Chronic viral hepatitis and hepatocellular carcinoma. *World J Surg*, **31**, 1243-8.

Bashirova AA, Geijtenbeek TB, van Duijnhoven GC, et al (2001). A dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin (DC-SIGN)-related protein is highly expressed on human liver sinusoidal endothelial cells and promotes HIV-1 infection. *J Exp Med*, **193**, 671-8.

Befeler AS, Di Bisceglie AM (2002). Hepatocellular carcinoma: diagnosis and treatment. *Gastroenterology*, **122**, 1609-19.

Benjamini YH, Y (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J Royal Stat Society Series B (Methodological)*, **57**, 289-300.

Bieche I, Asselah T, Laurendeau I, et al (2005). Molecular profiling of early stage liver fibrosis in patients with chronic hepatitis C virus infection. *Virology*, **332**, 130-44.

Block TM, Mehta AS, Fimmel CJ, Jordan R (2003). Molecular viral oncology of hepatocellular carcinoma. *Oncogene*, **22**, 5093-107.

Bruix J, Sherman M, Llovet JM, et al (2001). Clinical management of hepatocellular carcinoma. Conclusions of the Barcelona-2000 EASL conference. European Association for the Study of the Liver. *J Hepatol*, **35**, 421-30.

Capurro M, Wanless IR, Sherman M, et al (2003). Glypican-3: a novel serum and histochemical marker for hepatocellular carcinoma. *Gastroenterology*, **125**, 89-97.

Chuma M, Sakamoto M, Yamazaki K, et al (2003). Expression profiling in multistage hepatocarcinogenesis: identification of HSP70 as a molecular marker of early hepatocellular carcinoma. *Hepatology*, **37**, 198-207.

Cocquerel L, Voisset C, Dubuisson J (2006). Hepatitis C virus entry: potential receptors and their biological functions. *J Gen Virol*, **87**, 1075-84.

Cormier EG, Durso RJ, Tsamis F, et al (2004). L-SIGN (CD209L) and DC-SIGN (CD209) mediate transinfection of liver cells by hepatitis C virus. *Proc Natl Acad Sci U S A*, **101**, 14067-72.

Degos F, Christidis C, Ganne-Carrie N, et al (2000). Hepatitis C virus related cirrhosis: time to occurrence of hepatocellular carcinoma and death. *Gut*, **47**, 131-6.

Donato F, Tagger A, Gelatti U, et al (2002). Alcohol and hepatocellular carcinoma: the effect of lifetime intake and hepatitis virus infections in men and women. *Am J Epidemiol*, **155**, 323-31.

El-Serag HB, Hampel H, Javadi F (2006). The association between diabetes and hepatocellular carcinoma: a systematic review of epidemiologic evidence. *Clin Gastroenterol Hepatol*, **4**, 369-80.

Ferlay J, Shin HR, Bray F, et al (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*, **127**, 2893-917.

Gardner JP, Durso RJ, Arrigale RR, et al (2003). L-SIGN (CD 209L) is a liver-specific capture receptor for hepatitis C virus. *Proc Natl Acad Sci USA*, **100**, 4498-503.

Gentleman RC, Carey VJ, Bates DM, et al (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*, **5**, R80.

Hochreiter S, Bodenhofer U, Heusel M, et al (2010). FABIA: factor analysis for bicluster acquisition. *Bioinformatics*, **26**, 1520-7.

Hui AM, Makuuchi M, Li X (1998). Cell cycle regulators and human hepatocarcinogenesis. *Hepatogastroenterology*, **45**, 1635-42.

Hulsegge I, Kommadath A, Smits MA (2009). Globaltest and GOEAST: two different approaches for Gene Ontology analysis. *BMC Proc*, **3** Suppl 4, S10.

Jaskiewicz K, Chasen MR, Robson SC (1993). Differential expression of extracellular matrix proteins and integrins in hepatocellular carcinoma and chronic liver disease. *Anticancer Res*, **13**, 2229-37.

Kanehisa M (2002). The KEGG database. *Novartis Found Symp*, **247**, 91-101; discussion -3, 19-28, 244-52.

Koppel EA, van Gisbergen KP, Geijtenbeek TB, van Kooyk Y (2005). Distinct functions of DC-SIGN and its homologues L-SIGN (DC-SIGNR) and mSIGNR1 in pathogen recognition and immune regulation. *Cell Microbiol*, **7**, 157-65.

Kumar M, Sarin SK (2007). Is cirrhosis of the liver reversible? *Indian J Pediatr*, **74**, 393-9.

Levine AJ, Puzio-Kuter AM (2010). The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. *Science*, **330**, 1340-4.

Li J, Tibshirani R (2011). Finding consistent patterns: A nonparametric approach for identifying differential expression in RNA-Seq data. *Stat Methods Med Res*.

Lozach PY, Amara A, Bartosch B, et al (2004). C-type lectins L-SIGN and DC-SIGN capture and transmit infectious hepatitis C virus pseudotype particles. *J Biol Chem*, **279**, 32035-45.

Maere S, Heymans K, Kuiper M (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **21**, 3448-9.

Marcellin P (1999). Hepatitis C: the clinical spectrum of the disease. *J Hepatol*, **31** Suppl 1, 9-16.

Marrero JA, Lok AS (2004). Newer markers for hepatocellular carcinoma. *Gastroenterology*, **127**, S113-9.

Seyer JM, Hutcheson ET, Kang AH (1977). Collagen polymorphism in normal and cirrhotic human liver. *J Clin Invest*, **59**, 241-8.

Smith MW, Yue ZN, Geiss, et al (2003). Identification of novel tumor markers in hepatitis C virus-associated hepatocellular carcinoma. *Cancer Res*, **63**, 859-64.

Soini Y, Chia SC, Bennett WP, et al (1996). An aflatoxin-associated mutational hotspot at codon 249 in the p53 tumor suppressor gene occurs in hepatocellular carcinomas from Mexico. *Carcinogenesis*, **17**, 1007-12.

Swann JB, Smyth MJ (2007). Immune surveillance of tumors. *J Clin Invest*, **117**, 1137-46.

Tai DI, Tsai SL, Chen YM, et al (2000). Activation of nuclear factor kappaB in hepatitis C virus infection: implications for pathogenesis and hepatocarcinogenesis. *Hepatology*, **31**, 656-64.

Thorgeirsson SS, Grisham JW (2002). Molecular pathogenesis of human hepatocellular carcinoma. *Nat Genet*, **31**, 339-46.

Warburg O (1956). On the origin of cancer cells. *Science*, **123**, 309-14.

Wideroff L, Gridley G, Mellemkjaer L, et al (1997). Cancer incidence in a population-based cohort of patients hospitalized with diabetes mellitus in Denmark. *J Natl Cancer Inst*, **89**, 1360-5.

Wu ZL, Prather B, Ethen CM, et al (2011). Detection of specific glycosaminoglycans and glycan epitopes by in vitro sulfation using recombinant sulfotransferases. *Glycobiology*, **21**, 625-33.

Zhu N, Ware CF, Lai MM (2001). Hepatitis C virus core protein enhances FADD-mediated apoptosis and suppresses TRADD signaling of tumor necrosis factor receptor. *Virology*, **283**, 178-87.