European Review for Medical and Pharmacological Sciences 2014; 18: 2102-2108

Genome-wide analysis of genetic variations assisted by Ingenuity Pathway Analysis to comprehensively investigate potential genetic targets associated with the progression of hepatocellular carcinoma

F. YU[1], X.-Y. SHEN[2], L. FAN[2], Z.-C. YU[2]

[1]Department of Medical Oncology, The People's Hospital of Shaanxi Province, Xi'an, P.R. China
[2]Department of Medical Oncology, Xijing Hospital, The Fourth Military Medical University, Xi'an, P.R. China

Abstract. – OBJECTIVES: Hepatocellular carcinoma (HCC) is one of the most common malignant tumors and the third cause of cancer death worldwide. The development of HCC is a multi-pathway regulated process including a variety of genetic mutations and variations.

MATERIALS AND METHODS: To get an insight into the overall genetic aberration, we analyzed the gene expression profile of 9 HCC samples and 9 normal liver tissues as control. Genes located in these two regions of 4p16.3 and 14q32.1 with gain or loss proceeded to in-depth analysis based on the Ingenuity Pathway Analysis (IPA). Through network analysis of these genes and pathways, molecules highly correlated with HCC and might be targeted for treatments were also hunted out.

RESULTS: Through this method, we found two MicroRNAs (miR-943 and miR-571) that never been reported in HCC samples before.

CONCLUSIONS: Our study has introduced an effective method for cancer research and our work will greatly promote the genetic knowledge of HCC.

Key words:
Hepatocellular carcinoma, Copy number variation, Ingenuity pathway analysis.

Introduction

Hepatocellular carcinoma (HCC), accounts for most of the primary hepatic carcinoma, is one of the most common malignant tumors that place it near the very top of causes for cancer death worldwide[1-3]. HCC is much more common in men than in women[4]. The incidence of it varies greatly by areas that are much higher in Eastern and Southern eastern Asia, Middle and Western Africa, Melanesia, and Micronesia/Polynesia than in developed regions such as most areas of American and Europe[2]. However, the incidence of HCC and mortality rates has increased rapidly in these areas for recent few years[5]. The induction factors are various and complicated including intemperate alcohol intake, aflatoxin poisoning and for the most part virus infection[6-9]. Since most of patients are diagnosed at later stage with other advanced diseases and liver dysfunction, the prognosis is generally very poor that the median survival times less than one year.

To some extent, cancer is the result of uncontrolled cell abnormal proliferation due to the mutations or other variations such as copy number variation (CNV) or DNA rearrangement of some gene regions that are critical to some cellular signal pathways in genetic regulating network[10]. Getting known the information about gene mutations and variations that are related with cancer development will assist us to make an earlier diagnosis and effective personalized treatment. For the past decades, with the progress in molecular mechanism of cancer research, several genes have been found to play an important role in carcinogenesis. The genes such as *TP53* and *PTEN* have been reported to suppress cancer growth while the gene of β-catenin, ErbB family, COX-2 and HGF can promote the development of cancer[11-15]. However, these are far from enough. The molecular mechanism is much more complicated than we have thought that different patients with the same cancer may have different genetic alterations and the variations usually appear in diverse combination. Therefore, much effort should be put to figure out the relationships between genes related with HCC and the function of them in the regulating network of cellular life circle.

In our study, we firstly identified the regions with gain or loss in HCC samples by analyzing the whole genome expression data of HCC samples

Corresponding Author: Zhaocai Yu, MD; e-mail: yzhaocai@fmmu.edu.cn

and non-cancer liver samples. For the differentially expressed regions, we then figured out the genes in these regions and the related pathways and regulating networks they participated in based on the method of Ingenuity Pathway Analysis (IPA). The interactions of these genes and the top pathways including related molecules that might play an important role in HCC pathogenesis were clearly showed up. By this way, potential gene targets might be searched out much more easily.

Material and Methods

Microarray data

Gene expression datasets (GSE40873, GSE6465 and GSE33006) were downloaded from the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) database. Sample information for 9 HCC samples and 9 noncancerous liver tissue samples is listed in Table I. These datasets were based on the GPL570 platform: Affymetrix GeneChip Human Genome U133 Plus 2.0Array.

Detection of chromosomal aberrations

Datasets (GSE40873, GSE6465 and GSE33006) of the 9 HCC samples and 9 noncancerous liver tissue samples were downloaded. Raw data from the CEL files was divided into test and control group respectively. Test data was normalized using the control dataset. The relative expression level was calculated by the mapping of probe sets to each chromosome, chromosomal arms, cytobands and chromosomal bands. Then, data were summarized and an iterative median polishing procedure was used to generate a single expression value for each probe set. The threshold of resulting log2-transformed robust multi-array analysis expression value was acquired to detect chromosomal gain or loss, respectively. A Fisher's exact test was used to compute the enrichment of under- and over-expressed probes in chromosomes or chromosomal regions and the P-value were Bonferroni-corrected by default[16].

IPA functional analysis

Gene list in the regions of chromosomal gain and loss were functionally analyzed based on the Ingenuity Pathway Analysis (IPA) softwarev14400082 (2012 Ingenuity Systems, Inc. http://www.ingenuity.com). The IPA software is based on computational algorithms that analyze the functional connectivity of the genes from information obtained within the IPA database. Canonical pathways were scored for degree by analyzing a ratio of the number of genes that map to the pathway. The created genetic networks describe functional relationships among genes or proteins based on known associations in the databases. Networks related were ranked according to their biological relevance to the gene list provided. The *p* value calculated by Fisher's exact test was to determine the probability of the association was not due to chance alone.

Table I. Sample information.

Number	GEO ID	Sample type
1	GSM1003891_LC_07N-051217.CEL	Noncancerous liver tissue
2	GSM1003892_L08_N-3_H_02.CEL	Noncancerous liver tissue
3	GSM1003893_chip_array_L10N.CEL	Noncancerous liver tissue
4	GSM1003894_chip_array_L18N.CEL	Noncancerous liver tissue
5	GSM1003895_chip_array_L31N.CEL	Noncancerous liver tissue
6	GSM1003896_chip_array_L42N.CEL	Noncancerous liver tissue
7	GSM148605.CEL	Primary hepatocellular carcinoma
8	GSM148606.CEL	Primary hepatocellular carcinoma
9	GSM148607.CEL	Primary hepatocellular carcinoma
10	GSM148608.CEL	Primary hepatocellular carcinoma
11	GSM148609.CEL	Primary hepatocellular carcinoma
12	GSM148610.CEL	Primary hepatocellular carcinoma
13	GSM818276_T_01.CEL	Adjacent_normal_liver
14	GSM818277_NT_01.CEL	HCC_tissue
15	GSM818278_T_02.CEL	Adjacent_normal_liver
16	GSM818279_NT_02.CEL	HCC_tissue
17	GSM818280_T_03.CEL	Adjacent_normal_liver
18	GSM818281_NT_03.CEL	HCC_tissue

Results

Through the analysis of chromosomal abnormalities from gene expression datasets, we detected that there was a gain and a loss in the region of 4p16.3 and 14q32.1 respectively for HCC samples controlled by the noncancerous liver tissue samples with *p* value of 0.0000585 and 0.00425. IPA analysis result showed that the pathways dominated in the region of 4p16.3 were functionally related with cell proliferation, signaling and metabolism, such as calcium signaling pathway, STAT3 pathway, cAMP-mediated signaling, bladder cancer signaling and dermatan sulfate degradation (Metazoa). These pathways played an important role in caner progression and other diseases like dermatological diseases and conditions. The network of these pathways that biologically related with HCC was showed in Figure 1. According to the IPA database, the molecules in these pathways have the hepatotoxicity that could cause hepatocellular carcinoma,

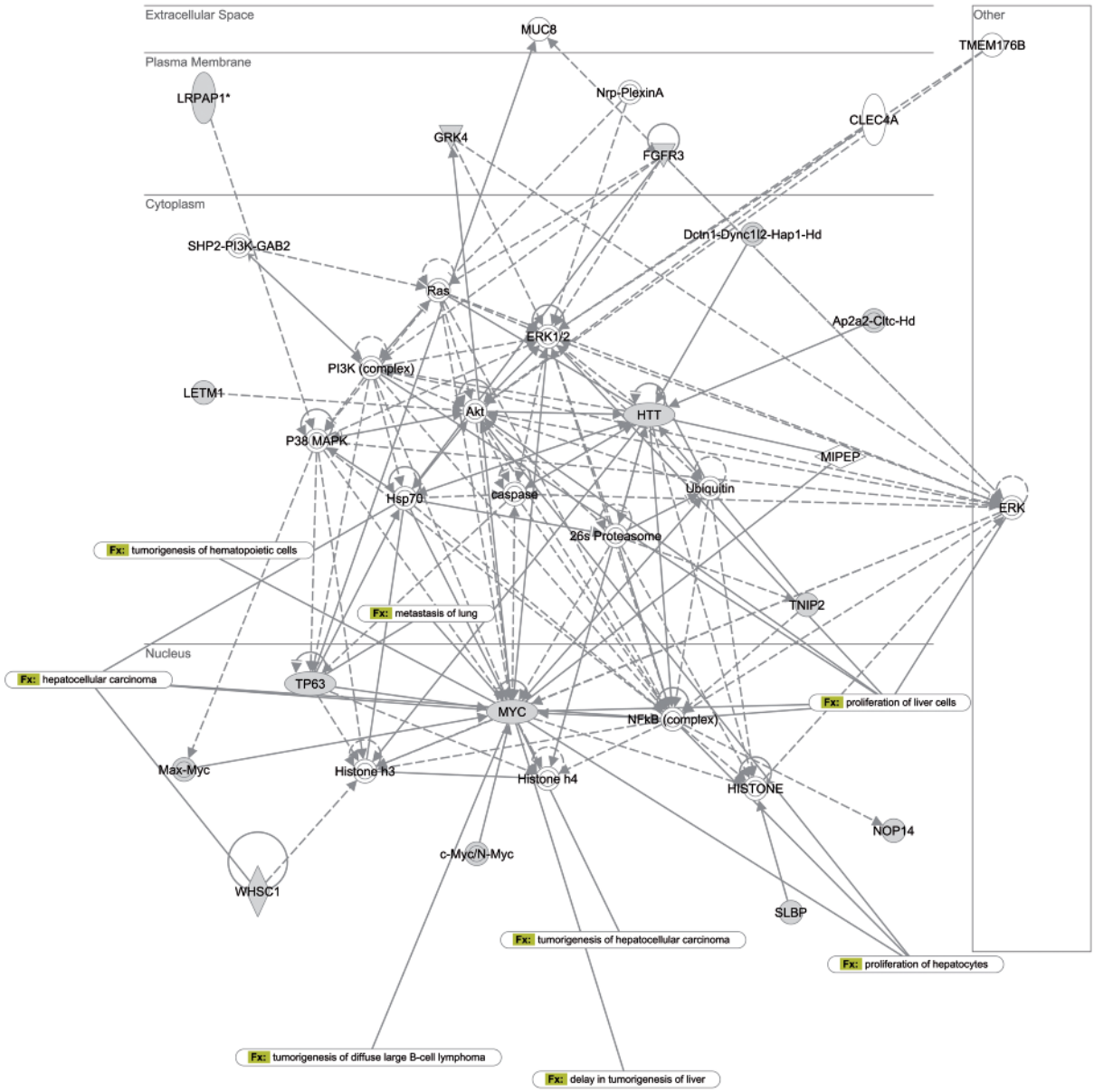

Figure 1. IPA based network for HCC in region 4p16.3. This figure illustrates the dominant pathways in the region of 4p16.3 for HCC samples, the critical transcription regulator, important enzymes and other molecules. The functional relationships of these molecules are illustrated using straight lines with the *arrow*.

liver hyperplasia, liver proliferation and liver necrosis. The top molecules from the IPA analysis that functioned with HCC were listed in Table II. Among these molecules, there were two non-coding RNA (miR-943 and miR-571) that had never been reported to play a role in HCC samples. Although not as significant as 4p16.3, we also performed IPA analysis to the region of 14q32.1. For this region, the top canonical pathways were coagulation system, acute phase response signaling, neuroprotective role of THOP1 in Alzheimer's disease, TGF-β signaling and role of NANOG in mammalian embryonic stem cell pluripotency, mostly related with metabolic diseases and neurological diseases. There were also some molecules having a hepatotoxicity to liver that could cause liver damage, liver hyperplasia, liver steatosis and liver cirrhosis. The function network of this region was showed in Figure 2.

Discussion

HCC is one of the significant cancers worldwide with poor prognosis. New and efficacious methods for early diagnosis and therapy are necessary. Accompanied with genetic malformations, some small molecules could be effective diagnosis marker for HCC. In our study, we have used the expression datasets of cancerous-non-

Table II. Molecules might contributed to the progression of HCC.

ID	Molecules in network	Score	Focus molecules	Top diseases and functions
1	ABCF2, AGPAT2, ANKRD26, ANKRD28, FAM175B, FAM193A, GID8, HAUS1, HAUS2, HAUS3, HAUS4, HAUS7, HAUS8, ILVBL, KIAA1009, KRI1, MAEA, MFSD10, RNF212, SDCCAG3, SHKBP1, SLC26A1, TACC3, TGS1, TMEM129, TTF2, TUBG2, UBC, UVSSA, ZFYVE28, ZNF141, ZNF595, ZNF718, ZNF721, ZNF732	32	15	Cellular assembly and organization, cellular function and maintenance, cell cycle
2	ADD1, APP, ATP5I, BUD31, DNAJC4, ECHDC2, FAM53A, heparan sulfate, HSP90AA1, IDUA, KCNG1, keratan sulfate, miR-4645-5p (and other miRNAsw/seed CCAGGCA), MSANTD1, MSMB, MXD4, MYL5, MYL10, MYLPF, myosin ATPase, AT6, NELFA, NKX3-2, OARD1, PIGG, Ppp1r12, Rock, SH3BP2, TGFB1, THAP4, TMEM175, TMEM126B, TSSK2, VCAM1, ZBTB49	24	13	Cell-to-cell signaling and interaction, hematological system development and function, immune cell trafficking
3	26sProteasome, Akt, Ap2a2-Cltc-Hd, c-Myc/N-Myc, caspase, CLEC4A, Dctn1-Dync1I2-Hap1-Hd, ERK, ERK1/2, FGFR3, GRK4, HISTONE, Histone h3, Histone h4, Hsp70, HTT, LETM1, LRPAP1, Max-Myc, MIPEP, MUC8, MYC, NFkB (complex), NOP14, Nrp-PlexinA, P38 MAPK, PI3K (complex), Ras, SHP2-PI3K-GAB2, SLBP, TMEM176B, TNIP2, TP63, Ubiquitin, WHSC1	20	11	Cancer, hematological disease, immunological disease
4	AGBL2, ATP, ATP6V0C, CARS,CBL, CPLX1, CTBP1-AS2, DCLRE1A, DGKQ, DOK7, HELQ, HIF1A, ITGB1, KIFC2, KRT3, MAGEA2/MAGEA2B, mir-202, miR-1237-5p (and other miRNAs w/seed GGGGGCG), miR-202-3p (and other miRNAs w/seed GAGGUAU), MutL alpha, MutS alpha, NAA30, NAE1, NAT8L, P2RX4, PCGF3, PCNA, PDE6B, POLN, RGS12, SPON2, TCN2, TMEM128, TP53, Ube3	20	11	Post-translational modification, protein folding, amino acid metabolism
5	CHTOP, CRIPAK, DDX21, DDX27, DDX50, DDX54, DHX30, DKC1, EGLN2, ESR1, FGF12, FTSJ3, GNL2, GNL3, GTPBP4, LUC7L3, LYAR, NAT10, NOP2,NSG1, PRKRA, PRPF4B, RBM23, RNF4, RPL21, RPL24, RPL18A, RPL36A/RPL36A-HNRNPH2, RRP1B, RSL1D1, SRPK2, SRSF5, TCOF1, TDG, YPEL3	6	4	RNA post-transcriptional modification, auditory and vestibular system development and function, organ morphology
6	miR-943 (miRNAs w/seed UGACUGU)	2	1	/
7	miR-571 (miRNAs w/seed GAGUUGG)	2	1	/

cancerous paired samples to detect the related genetic variations in HCC samples. For the detected gene regions with gain and loss compared with normal control samples, we analyzed the related genes based on the IPA software. The top related pathways and molecules in these regions were analyzed for the possible association with HCC. Functional network analysis in our study assisted for searching new biomarkers.

In our study, we detected a significantly gained region in 4p16.3. IPA analysis of this region showed that the top related pathways played an important role in cellular proliferation, signaling and metabolism, which would greatly promote the progression of HCC. The core regulating molecules were MYC and TP63 by our network analysis result showed in Figure 1. MYC and TP63 have been reported to play a key role in the regulation of gene expression with cancerous stages[17,18]. Our result was consistent with former study. However, we also found two miRNAs that were never been reported in HCC samples. One of these was miR943, and the other was miR571. It has been reported that non-coding RNAs play

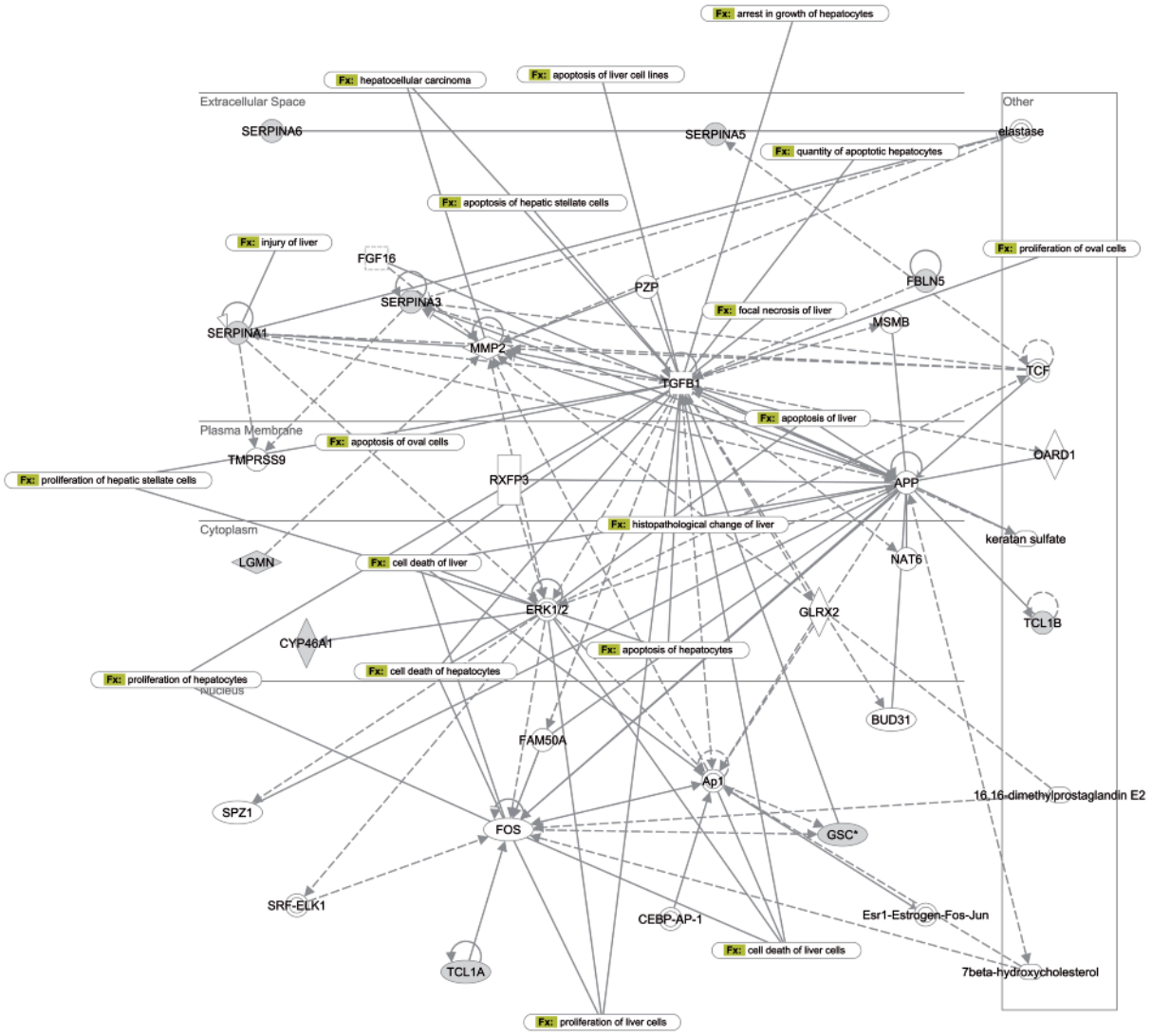

Figure 2. IPA based network for HCC in region of 14q32.1. This figure illustrates the dominant pathways in the region of 14q32.1 for HCC samples. Critical enzymes and functional molecules are illustrated and the functional relationships are illustrated as Figure 1.

an important role in cancer development[19,20]. MiR943 has been reported to expressed in human cervical cancer samples[21], which suggests great value of this molecule as a cancer biomarker. And miR571 has also been reported to play a role in pathogenesis of liver cirrhosis[22]. As researches founded that part of liver cirrhosis was an important factor for HCC[23], miR571 might be another critical molecule for the prevention of HCC. Although the gene loss in 14q32.1 was not so significant, the pathway and network analysis showed great functional potential of the variation in this region for HCC progression.

Conclusions

Our study has introduced an effective method for genome-wide detection of HCC related genetic variations using gene expression datasets. With the assistant of IPA software and the powerful database, the molecules related with HCC progression and potential biomarkers for diagnosis and therapy could be easily found. The mechanisms of actions could be clearly observed. Although the datasets needs to be enlarged for fully understand the cancerous progression of HCC and select the most effective biomarkers, our work could be a good model for further research and greatly promote the research of HCC and other cancers.

Acknowledgements

This project is supported by the National Natural Science Foundation of China (Grant No. 81372458).

Conflict of interest

The Authors declare that they have no conflict of interests.

Reference

1) Fernandez MP, Redvanly RD. Primary hepatic malignant neoplasms. Radiol Clin North Am 1998; 36: 333-348.
2) Bosch FX, Ribes J, Diaz M, Cleries R. Primary liver cancer: worldwide incidence and trends. Gastroenterology 2004; 127: S5-S16.
3) Organization WH. Mortality Database, WHO Statistical Information System. Available at http://www.who.int/whosis/en/ 2008.
4) Ries LAG HD, Krapcho M, Mariotto A, Miller BA, Feuer EJ, Clegg L, Eisner MP, Horner MJ, Howlader N, Hayat M, Hankey BF, Edwards BK. SEER Cancer Statistics Review, 1975-2003. Bethesda, MD: National Cancer Institute 2006; Available at http://seer.cancer.gov/csr/1975_2003/.
5) El-Serag HB. Hepatocellular carcinoma: recent trends in the United States. Gastroenterology 2004; 127: S27-34.
6) Simonetti RG, Camma C, Fiorello F, Politi F, D'Amico G, Pagliaro L. Hepatocellular carcinoma. A worldwide problem and the major risk factors. Dig Dis Sci 1991; 36: 962-972.
7) Bralet MP, Regimbeau JM, Pineau P, Dubois S, Loas G, Degos F, Valla D, Belghiti J, Degott C, Terris B. Hepatocellular carcinoma occurring in nonfibrotic liver: epidemiologic and histopathologic analysis of 80 French cases. Hepatology 2000; 32: 200-204.
8) Sangiovanni A, Prati GM, Fasani P, Ronchi G, Romeo R, Manini M, Del Ninno E, Morabito A, Colombo M. The natural history of compensated cirrhosis due to hepatitis C virus: A 17-year cohort study of 214 patients. Hepatology 2006; 43: 1303-1310.
9) Snowberger N, Chinnakotla S, Lepe RM, Peattie J, Goldstein R, Klintmalm GB, Davis GL. Alpha fetoprotein, ultrasound, computerized tomography and magnetic resonance imaging for detection of hepatocellular carcinoma in patients with advanced cirrhosis. Aliment Pharmacol Ther 2007; 26: 1187-1194.
10) Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de Rijn M, Waltham M, Pergamenschikov A, Lee JC, Lashkari D, Shalon D, Myers TG, Weinstein JN, Botstein D, Brown PO. Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 2000; 24: 227-235.
11) Karim S, Mirza Z, Naseer MI, Al-Qahtani MH, Ali A. Clinicopathological characteristics and chronology of p53 expression in the development of gastric cancer. Hepatogastroenterology 2013; 60: 2113-2118.
12) Chen Z, Trotman LC, Shaffer D, Lin HK, Dotan ZA, Niki M, Koutcher JA, Scher HI, Ludwig T, Gerald W, Cordon-Cardo C, Pandolfi PP. Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. Nature 2005; 436: 725-730.
13) Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, Teague JW, Campbell PJ, Stratton MR, Futreal PA. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res 2011; 39: D945-950.
14) Legan M. Cyclooxygenase-2, p53 and glucose transporter-1 as predictors of malignancy in the development of gallbladder carcinomas. Bosn J Basic Med Sci 2010; 10: 192-196.
15) Yu G, Jing Y, Kou X, Ye F, Gao L, Fan Q, Yang Y, Zhao Q, Li R, Wu M, Wei L. Hepatic stellate cells secreted hepatocyte growth factor contributes to the chemoresistance of hepatocellular carcinoma. PLoS One 2013; 8: e73312.

16) Bollen S, Leddin M, Andrade-Navarro MA, Mah N. CAFE: an R package for the detection of gross chromosomal abnormalities from gene expression microarray data. Bioinformatics 2014; 30: 1484-1485.

17) Denis N, Kitzis A, Kruh J, Dautry F, Corcos D. Stimulation of methotrexate resistance and dihydrofolate reductase gene amplification by c-myc. Oncogene 1991; 6: 1453-1457.

18) Wu G, Nomoto S, Hooque MO, Dracheva T, Osada M, Lee CC, Dong SM, Guo Z, Benoit N, Cohen Y, Rechthand P, Califano J, Moon CS, Ratovitski E, Jen J, Sidransky D, Trink B. DeltaNp63alpha and TAp63alpha regulate transcription of genes with distinct biological functions in cancer and development. Cancer Res 2003; 63: 2351-2357.

19) Ke TW, Hsu HL, Wu YH, Chen WT, Cheng YW, Cheng CW. MicroRNA-224 suppresses colorectal cancer cell migration by targeting Cdc42. Dis Markers 2014; 2014: 617150.

20) Chen Z, Li J, Tian L, Zhou C, Gao Y, Zhou F, Shi S, Feng X, Sun N, Yao R, Shao K, Li N, Qiu B, Tan F, He J. MiRNA expression profile reveals a prognostic signature for esophageal squamous cell carcinoma. Cancer Lett 2014; 350: 34-42.

21) Lui WO, Pourmand N, Patterson BK, Fire A. Patterns of known and novel small RNAs in human cervical cancer. Cancer Res 2007; 67: 6031-6043.

22) Roderburg C, Mollnow T, Bongaerts B, Elfimova N, Vargas Cardenas D, Berger K, Zimmermann H, Koch A, Vucur M, Luedde M, Hellerbrand C, Odenthal M, Trautwein C, Tacke F, Luedde T. Micro-RNA profiling in human serum reveals compartment-specific roles of miR-571 and miR-652 in liver cirrhosis. PLoS One 2012; 7: e32999.

23) Bartolomeo N, Trerotoli P, Serio G. Progression of liver cirrhosis to HCC: an application of hidden Markov model. BMC Med Res Methodol 2011; 11: 38.