Original paper

Bioinformatics analysis of the gene expression profile of hepatocellular carcinoma: preliminary results

Jia Li[1,2], Zhongxi Huang[3*], Lixin Wei[1*]

[1]Department of Pathology, Chinese PLA General Hospital, Beijing, China
[2]Medical College, Nankai University, Tianjing, China
[3]Institute of Oncology, Nanfang Medical University, Guangzhou, Guangdong, China

These authors contributed equally to this work.

Aim of the study: To analyse the expression profile of hepatocellular carcinoma compared with normal liver by using bioinformatics methods.
Material and methods: In this study, we analysed the microarray expression data of HCC and adjacent normal liver samples from the Gene Expression Omnibus (GEO) database to screen for differentially expressed genes. Then, functional analyses were performed using GenCLiP analysis, Gene Ontology categories, and aberrant pathway identification. In addition, we used the CMap database to identify small molecules that can induce HCC.
Results: Overall, 2721 differentially expressed genes (DEGs) were identified. We found 180 metastasis-related genes and constructed co-occurrence networks. Several significant pathways, including the transforming growth factor β (TGF-β) signalling pathway, were identified as closely related to these DEGs. Some candidate small molecules (such as betahistine) were identified that might provide a basis for developing HCC treatments in the future.
Conclusions: Although we functionally analysed the differences in the gene expression profiles of HCC and normal liver tissues, our study is essentially preliminary, and it may be premature to apply our results to clinical trials. Further research and experimental testing are required in future studies.

Key words: hepatocellular carcinoma, bioinformatics analysis, differentially expressed genes, functional analysis.

Contemp Oncol (Pozn) 2016; 20 (1): 20–27
DOI: 10.5114/wo.2016.58497

Introduction

Hepatocellular carcinoma (HCC) is the fifth most frequent cancer worldwide and is also the third leading cause of cancer-related mortality [1, 2]. The World Health Organisation (WHO) has estimated that there are nearly 56,400 new cases of HCC around the world per year [3], and the incidence is much higher in men than in women. The highest liver cancer rates are found in developing countries, especially in East Asia and Malaysia, South Africa, and Sub-Saharan Africa, whereas rates are lower in Europe, North and South America, Australia, and New Zealand [4]. HCC can be induced by several risk factors, such as chronic infection with hepatitis B virus (HBV) or hepatitis C virus (HCV), hepatic cirrhosis, alcoholic liver disease, and exposure to aflatoxins [5–7].

Hepatocellular carcinoma always results from accumulative, long-term interactions between environmental and genetic factors. The multifactorial progression of HCC involves the activation of oncogenes, the inactivation of tumour suppressor genes, gene mutations, and irreversible cell damage. Many studies have focused on the genetic mutations and the overexpression of abnormal genes that promote malignant progression, such as Cyclin D1 (CCND1), v-raf murine sarcoma viral oncogene homolog B (BRAF), epidermal growth factor receptor (EGFR), c-myc, Ras, AKT, Yap, and baculoviral IAP repeat containing 2 (BIRC2) [8, 9], as well as on the deletion or loss of heterozygosity (LOH) in the chromosomal regions of tumour suppressor genes, such as CDKN2A, RB1, TP53, and PTEN [9, 10].

Although many genes that can promote or suppress HCC have been identified, the molecular mechanisms underlying HCC initiation, progression, metastasis, or targeted therapy remain unclear. High-throughput microarray technology, which enables investigators to obtain massive expression data sets, has been demonstrated to be a useful approach for identifying new tumour marker genes for tumour diagnosis or targeted treatment [11–13].

The aim of this study was to analyse the expression profile of hepatocellular carcinoma compared with normal liver by using bioinformatics methods.

Material and methods

Affymetrix microarray data from hepatocellular carcinoma and adjacent normal liver tissues

To investigate the change in expression profile between hepatocellular carcinoma tissues and adjacent normal liver tissues and to explore the mechanisms that may be involved in hepatocarcinogenesis, we downloaded and analysed the gene expression profile of GSE33006 from the Gene Expression Omnibus – a public functional genomics data repository (http://www.ncbi.nlm.nih.gov/geo/). The dataset, which was submitted by Huang

et al. (2011), contains three HCC tissue chips and three adjacent normal liver tissue chips from patients who underwent surgery, and it is based on the Affymetrix GPL570 Platform (Affymetrix GeneChip Human Genome U133 Plus 2.0 Array). Total RNA was extracted from biopsied samples using TRIzol reagent for further individual on-chip analysis.

Screening of differentially expressed genes

The original CEL files were downloaded and analysed using the R package (3.0.2) (http://www.r-project.org/). The robust multi-array average (RMA) method and Affymetrix Microarray Suite version 5 (MAS5) were used for data normalisation and background correcting. We used a classical t-test to identify differentially expressed genes with a variation > 2-fold, and we defined $p < 0.05$ to be statistically significant. The probe set ID list of selected differentially expressed genes was then uploaded to the NetAffx™ Analysis Centre (http://Affymetrix.com/analysis/index/affx) to obtain the corresponding gene symbols and gene titles. The expression data were deleted if there was no corresponding gene symbol for the probe set or if more than one gene symbol corresponded to a probe set. Only probe sets that mapped to a unique gene were suitable for further analysis. If there were multiple probe sets corresponding to the same gene, the expression values of these probe sets were then averaged.

Functional analysis of differentially expressed genes

GenCLiP (a software program for clustering gene lists by literature profiling, and for constructing gene co-occurrence networks related to keywords of interest; http://ci.smu.edu.cn/GenCLip/) [14, 15] was used to analyse the differentially expressed genes, explore the pathogenesis, and construct gene networks related to important biological processes in tumours, such as metastasis, cell growth, and cell cycle progression.

Gene Ontology Analysis (GO) is a common useful approach for annotating genes and gene products and for predicting gene function for high-throughput genome or transcriptome data [16, 17]. To better investigate the function of these DEGs, a functional annotation tool, DAVID (Database for Annotation, Visualisation, and Integrated Discovery) v6.7 (http://david.abcc.ncifcrf.gov/), was used to cluster enriched function-related gene groups according to gene ontology (GO) terms, including molecular function, biological process, and cellular component [18]. We also used DAVID to visualise genes on KEGG (Kyoto Encyclopaedia of Genes and Genomes) pathway maps to investigate the dysregulated biological pathways in which the DEGs may participate. The cut-off criteria were that a pathway must contain at least two differentially expressed genes and have a p-value < 0.05.

Small molecule identification

The Connectivity Map (CMap, build version build 02, http://www.broadinstitute.org/cmap/) was used to compare the differentially expressed genes with those in the CMap database, to identify the small molecules associated with these DEGs. First, we divided the DEGs into two groups with an upper limit of 500 probe sets per group: the up-regulated group and the down-regulated group. Then, the probe sets from the two groups were preloaded into their sandbox with the GRP format for gene set enrichment analysis. Finally, the enrichment scores, which ranged from +1 to –1, were calculated.

The enrichment value represented the association between the preloaded query signature and the gene profile for a small molecule treatment. A high positive up score (close to +1) indicates that the corresponding small molecule induces the expression of the probe sets in the up tag list (hepatocellular carcinoma), whereas a high negative up score (close to –1) indicates greater similarity between the genes induced by the small molecule and the probe sets in the down tag list (adjacent normal liver).

Results

Identification of differentially expressed genes

Using bioinformatics analysis we found 4233 probe set IDs that differed between the HCC and adjacent normal liver tissues. After the gene symbols from Affymetrix database were matched and the substandard expression data were removed, altered expression was identified for 2721 probe set IDs (corresponding to 2721 genes); these genes were marked for further analysis.

GenCLiP analysis of the differentially expressed genes

Of the 2721 analysed genes, 2701 had related literature, with an average of 661 literature matches per gene; these genes were subsequently used for cluster analysis (Fig. 1).

We used the "literature mining gene networks" function of GenCLiP to search for related genes and to construct co-occurrence gene networks among the DEGs using the keyword "metastasis". A total of 180 known metastasis-related genes were identified, and 180 genes formed 268 related gene pairs (Fig. 2A). To determine whether the 180 known metastasis-related genes were identified randomly, PubMed was used to search for the occurrence of each gene set of the microarray with the keyword "metastasis", and then 300 random simulations were performed. The resulting distribution of the number of metastasis-related genes and gene pairs derived from random genes was similar to the normal distribution, and the probability that a set of 2701 randomly selected genes contained more than 180 metastasis-related genes or 268 gene pairs was $p = 0.00000$ for both distributions (Fig. 2B). Furthermore, gene networks related to the keywords "cell growth" and "cell cycle" were also constructed (Supplement Fig. S1).

Gene Ontology analysis of differentially expressed genes

These 2721 differentially expressed genes were functionally classified into three Gene Ontology (GO) categories using the online analysis tool DAVID. For the biological process category, a cut-off level of $p < 0.001$ was

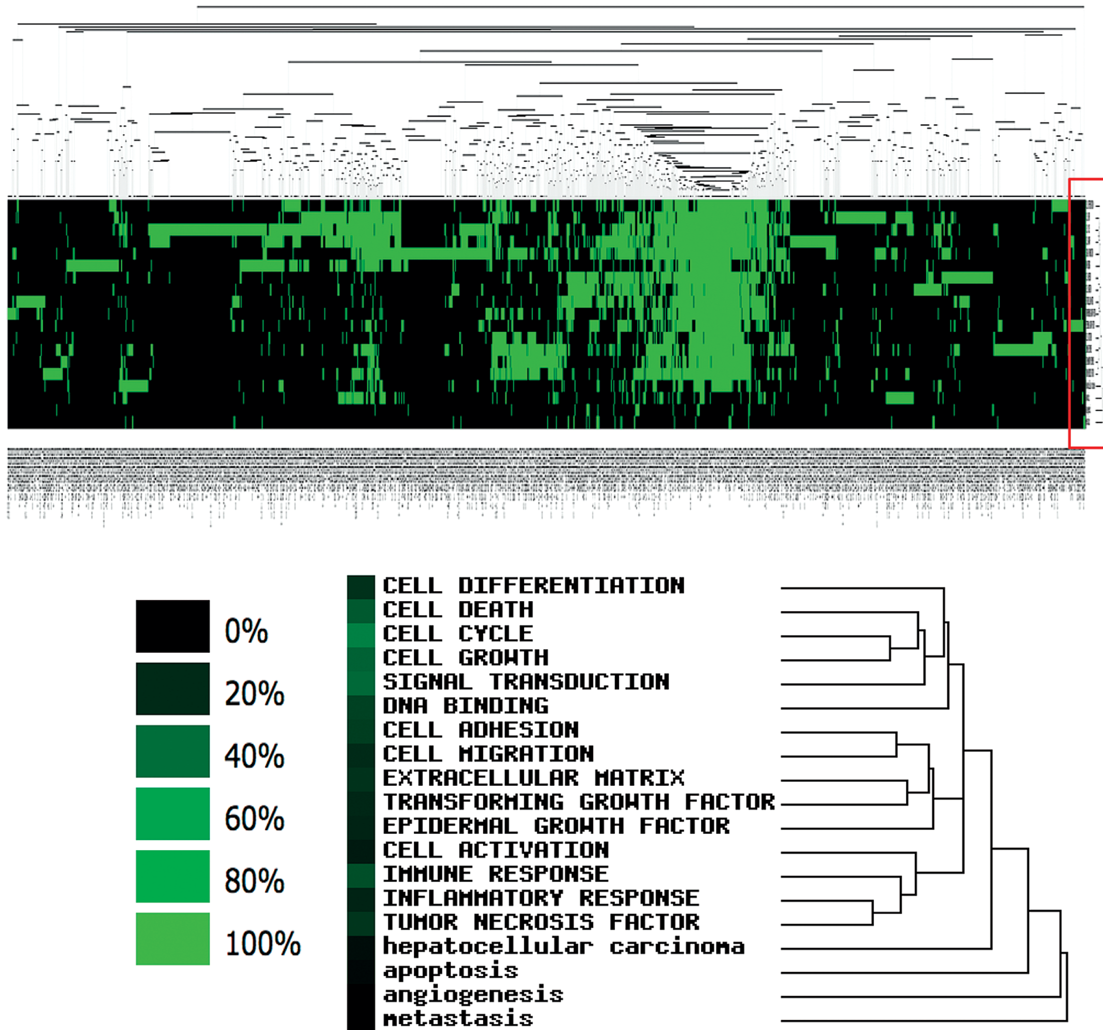

Fig. 1. Literature profiling of differentially expressed genes. Clustergram generated from analysing the keyword occurrence of the analysed genes. Green represents the corresponding gene-term associations that were reported to be positive

used to identify significant enrichment of genes with the corresponding GO terms. Analysis revealed changes in the biological processes of the immune system, such as the positive regulation of immune system processes, leukocyte activation, innate immune response, T cell activation, B cell-mediated immunity, and immunoglobulin-mediated immune response. The DEGs also showed significant enrichment in processes related to the regulation of cell growth, such as the regulation of cell proliferation, the cell cycle, cell differentiation, nuclear division, and M phase of the mitotic cell cycle. Changes in the expression profile also affected the biological processes of angiogenesis and signal transduction (Table 1).

In the category of cellular component, the most enriched GO term was cytoplasm (1089 genes). In addition, the DEGs were also enriched in cellular components related to the plasma membrane, extracellular region, chromosome, secretory granules, and the cytoplasmic membrane-bound vesicle lumen ($p < 0.01$) (Table 2). Table 3 shows the clustered GO terms in the molecular function category for the differentially expressed genes ($p < 0.01$). The HCC expression profile indicated that the activity of some enzymes changed, including the activities of oxidoreductase, endopeptidase inhibitor, transmembrane receptor protein tyrosine kinase, etc. The binding abilities of some materials (polysaccharides, glycosaminoglycans, etc.) were also changed.

KEGG pathway enrichment analysis

The changed gene expression profile of hepatocellular carcinoma may result in many dysregulated signal pathways. We used DAVID to cluster the DEGs for KEGG pathway enrichment analysis, and $p < 0.05$ was set as the cut-off criteria for statistical significance. As shown in Table 4, 20 dysregulated pathways were identified from the changes in HCC; of these pathways, the complement and coagulation cascades and cell adhesion molecules (CAMs) showed the most significant enrichment ($P = 2.79E-18$ and $P = 3.17E-04$, respectively). As previously reported, some altered pathways were highly related to the initiation or progression of malignant tumours; these pathways included the TGF-β signalling pathway, T cell/B cell receptor signalling pathway, and pathways related to DNA replication and cytokine-cytokine receptor interactions.

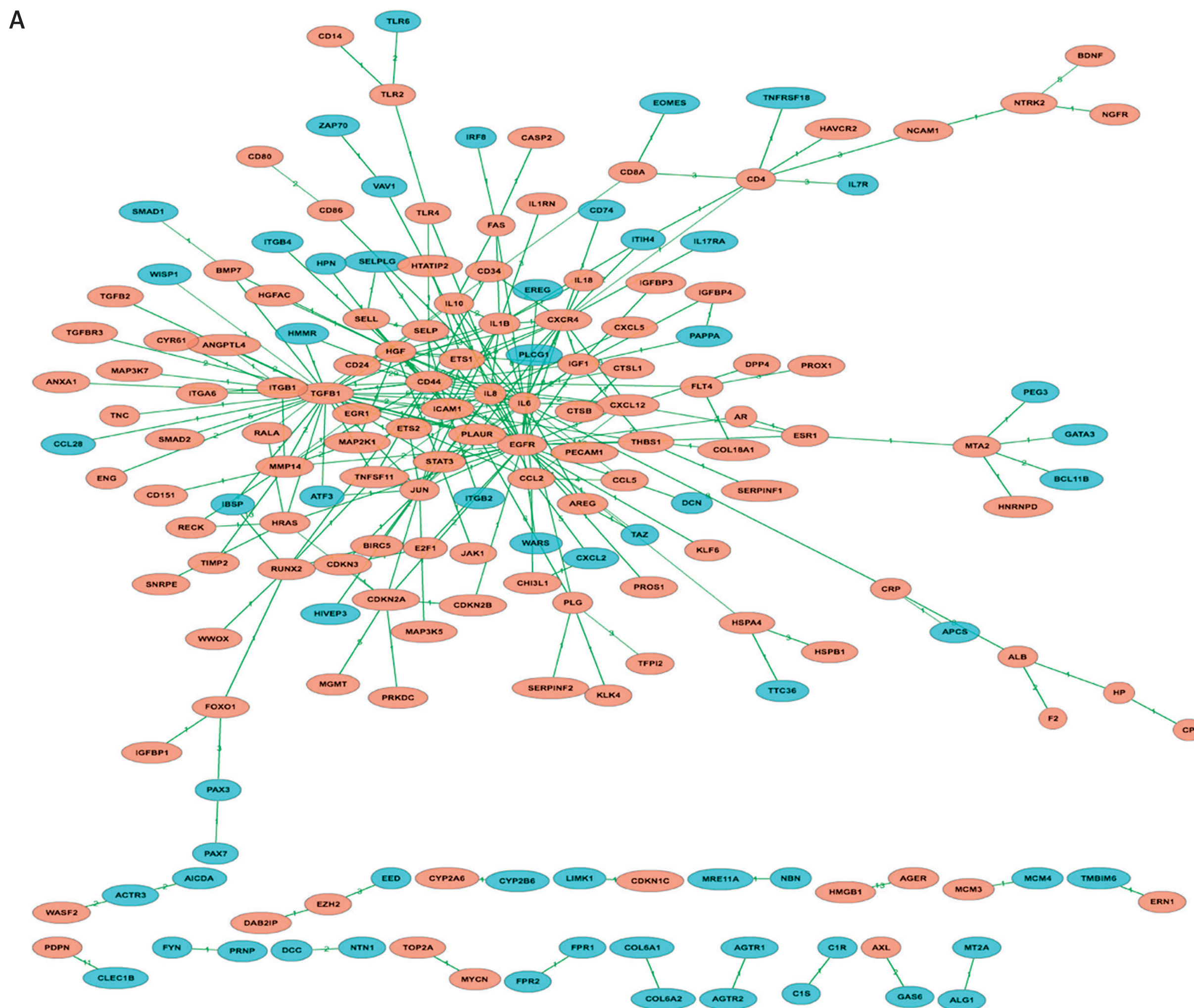

Fig. 2. Constructing a gene co-occurrence network related to the keyword "metastasis" from the differentially expressed genes. A) The network of analysed genes. The orange node represents known hepatocellular carcinoma-related genes. B) The distribution of the number of metastasis-related genes and gene pairs derived from the random genes was similar to the normal distribution

Identification of candidate small molecules

To identify candidate small molecules that could reverse the gene expression changes of hepatocellular carcinoma, the DEGs were divided into two groups: up-regulated and down-regulated, which were uploaded to the CMap database for Gene Set Enrichment Analysis and then matched to treatments with small molecules. The 20 most significant small molecules are listed in Table 5 with their enrichment scores and p-values. As shown in Table 5, the small molecules vorinostat (enrichment score = 0.973) and trichostatin A (enrichment score = 0.895) were associated with a highly significant positive score and could partially imitate the status of hepatocellular carcinoma. These small molecules may be strong induction factors for HCC. In contrast, cromoglicic acid (enrichment score = –0.927) and ranitidine (enrichment score = –0.837) were associated with highly significant negative scores and may imitate the normal liver status. These small molecules could reverse the tumoral status of HCC and therefore provide novel ideas and molecular mechanisms for developing new drugs for treating HCC in the future. However, these candidate small molecules still require further detailed research.

Table 1. Gene Ontology enrichment of differentially expressed genes in biological process

Term	Count	*p*-value
GO:0006955~immune response	182	9.01E-18
GO:0002376~immune system process	239	1.60E-17
GO:0002526~acute inflammatory response	49	2.01E-16
GO:0006952~defense response	157	7.08E-14
GO:0048518~positive regulation of biological process	402	1.34E-13
GO:0002460~adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin	39	2.32E-13
GO:0002682~regulation of immune system process	110	3.04E-13
GO:0002541~activation of plasma proteins involved in acute inflammatory response	27	3.09E-12
GO:0045087~innate immune response	53	6.50E-12
GO:0002449~lymphocyte-mediated immunity	35	7.52E-12
GO:0019724~B cell-mediated immunity	30	3.50E-11
GO:0016064~immunoglobulin-mediated immune response	29	7.34E-11
GO:0002455~humoral immune response-mediated by circulating immunoglobulin	20	1.96E-09
GO:0022610~biological adhesion	152	5.46E-08
GO:0007155~cell adhesion	151	8.58E-08
GO:0042110~T-cell activation	41	4.11E-07
GO:0008283~cell proliferation	98	3.28E-06
GO:0045595~regulation of cell differentiation	107	5.46E-06
GO:0001944~vasculature development	62	1.37E-05
GO:0050848~regulation of calcium-mediated signalling	12	2.32E-05
GO:0001568~blood vessel development	59	4.95E-05
GO:0007049~cell cycle	150	5.66E-05
GO:0007059~chromosome segregation	25	2.69E-04
GO:0009966~regulation of signal transduction	162	3.13E-04
GO:0007169~transmembrane receptor protein tyrosine kinase signalling pathway	52	3.76E-04
GO:0000087~M phase of mitotic cell cycle	52	3.76E-04
GO:0040011~locomotion	88	3.93E-04
GO:0007067~mitosis	51	4.46E-04
GO:0007167~enzyme linked receptor protein signalling pathway	72	5.78E-04
GO:0006928~cell motion	94	7.27E-04
GO:0001525~angiogenesis	37	7.33E-04
GO:0040007~growth	43	9.98E-04

Discussion

By analysing the differentially expressed genes using GenCLiP software, 180 metastasis related genes were identified and used to construct co-occurrence gene networks. Of the metastasis-related genes, *TGFB1* and *EGFR* had the largest number of co-occurring genes (34 and 33, respectively) and were located in the centre of the network. As recognised dysregulated growth factors, TGFB1/EGFR and their downstream signalling pathway components contribute to the proliferation and invasive behaviour of liver cancer cells [19–21]. As a transmembrane glycoprotein, CD44 was shown to interact with 28 genes. Reports have revealed that through the interaction of CD44 and its association molecules, CD44 can regulate cancer cell proliferation, adhesiveness, migration, and metastasis [22]. CD44 was also demonstrated to be closely associated with the extrahepatic metastasis of HCC [23]. Co-occurrence networks involving large numbers of related genes enable researchers to identify critical genes and their possible interactional networks, which may provide a new direction for the diagnosis and targeted therapy of HCC.

Gene Ontology (GO) analysis revealed that these DEGs were closely related to multiple biological processes involved in the mechanism of most malignant tumours, such as the regulation of the immune system, cell growth, the cell cycle, angiogenesis, and signal transduction. Sev-

Table 2. Gene Ontology enrichment of differentially expressed genes in cellular component

Term	Count	*p*-value
GO:0044421~extracellular region part	212	2.83E-13
GO:0005615~extracellular space	160	3.50E-12
GO:0031226~intrinsic to plasma membrane	248	1.60E-11
GO:0005887~integral to plasma membrane	242	3.67-11
GO:0005576~extracellular region	366	8.21E-11
GO:0009897~external side of plasma membrane	52	2.24E-08
GO:0009986~cell surface	86	3.78E-08
GO:0031093~platelet alpha granule lumen	21	7.13E-08
GO:0060205~cytoplasmic membrane-bounded vesicle lumen	21	3.09E-07
GO:0031983~vesicle lumen	21	7.49E-07
GO:0000793~condensed chromosome	37	1.54E-05
GO:0005737~cytoplasm	1089	2.58E-05
GO:0005886~plasma membrane	593	2.81E-05
GO:0005579~membrane attack complex	7	4.00E-05
GO:0043235~receptor complex	32	1.41E-04
GO:0005694~chromosome	91	2.32E-04
GO:0030141~secretory granule	43	2.92E-04
GO:0044433~cytoplasmic vesicle part	41	0.0024
GO:0005604~basement membrane	21	0.0034
GO:0000777~condensed chromosome kinetochore	17	0.0039
GO:0031012~extracellular matrix	66	0.0041
GO:0000775~chromosome, centromeric region	29	0.0046
GO:0031410~cytoplasmic vesicle	112	0.0048
GO:0034358~plasma lipoprotein particle	12	0.0054
GO:0000779~condensed chromosome, centromeric region	18	0.0064
GO:0030055~cell-substrate junction	26	0.0082

Table 3. Gene Ontology enrichment of differentially expressed genes in molecular function

Term	Count	*p*-value
GO:0005515~protein binding	1282	5.58E-10
GO:0030247~polysaccharide binding	49	4.16E-08
GO:0005539~glycosaminoglycan binding	45	1.15E-07
GO:0030246~carbohydrate binding	87	1.71E-07
GO:0003823~antigen binding	20	1.46E-04
GO:0008201~heparin binding	30	1.49E-04
GO:0019955~cytokine binding	31	1.80E-04
GO:0019838~growth factor binding	30	2.14E-04
GO:0046983~protein dimerization activity	105	5.49E-04
GO:0030234~enzyme regulator activity	151	0.001
GO:0004896~cytokine receptor activity	18	0.0011
GO:0019825~oxygen binding	15	0.0017
GO:0004866~endopeptidase inhibitor activity	35	0.0017
GO:0030414~peptidase inhibitor activity	36	0.0023
GO:0004497~monooxygenase activity	25	0.0037
GO:0003824~catalytic activity	787	0.0051
GO:0019899~enzyme binding	96	0.0052
GO:0019842~vitamin binding	30	0.0074
GO:0019901~protein kinase binding	33	0.0074
GO:0008172~S-methyltransferase activity	5	0.0097
GO:0016491~oxidoreductase activity	120	0.0099

Table 4. The altered KEGG pathways

Term	Description	Count	*p*-value
hsa04610	complement and coagulation cascades	44	2.79E-18
hsa04514	cell adhesion molecules (CAMs)	38	3.17E-04
hsa00071	fatty acid metabolism	15	0.0027
hsa04640	hematopoietic cell lineage	25	0.0037
hsa00380	tryptophan metabolism	14	0.0077
hsa04672	intestinal immune network for IgA production	16	0.0081
hsa04060	cytokine-cytokine receptor interaction	58	0.0084
hsa04660	T cell receptor signalling pathway	28	0.011
hsa04110	cell cycle	31	0.013
hsa00330	arginine and proline metabolism	16	0.017
hsa05214	glioma	18	0.019
hsa03030	DNA replication	12	0.022
hsa04512	ECM-receptor interaction	22	0.023
hsa00590	arachidonic acid metabolism	16	0.028
hsa04350	TGF-β signalling pathway	22	0.033
hsa04810	regulation of actin cytoskeleton	46	0.034
hsa05322	systemic lupus erythematosus	24	0.040
hsa00280	valine, leucine, and isoleucine degradation	13	0.041
hsa04662	B cell receptor signalling pathway	19	0.048
hsa04650	natural killer cell mediated cytotoxicity	30	0.049

Table 5. List of top 20 identified small molecules

Cmap name	Enrichment score	*p*-value
vorinostat	0.973	0
trichostatin A	0.895	0
geldanamycin	0.705	0
fluphenazine	0.629	0
trifluoperazine	0.625	0
thioridazine	0.599	0
sirolimus	0.491	0
valproic acid	0.359	0
pioglitazone	–0.646	0.00004
6-bromoindirubin-3’-oxime	–0.769	0.00008
withaferin A	0.896	0.0001
wortmannin	0.501	0.0001
ivermectin	0.858	0.00012
prochlorperazine	0.524	0.00014
suloctidil	0.888	0.00016
cephaeline	0.848	0.00018
PNU-0293363	–0.954	0.00022
ranitidine	–0.837	0.0003
meptazinol	–0.825	0.00177
betahistine	–0.822	0.00189

eral significant altered pathways were identified by KEGG pathway analysis. The TGF-β signal pathway has been reported to be functionally impaired in hepatocarcinogenesis [20]. Interactions between the extracellular matrix (ECM) receptor and cells play a vital role in cell adhesion and form a crucial step in tumour cell migration and invasion into the extracellular matrix [24].

Analysis using the CMap database identified a set of small molecules that may imitate the status of hepatocellular carcinoma or a normal liver. The candidate small molecules that were associated with highly significant negative enrichment scores may reverse the abnormal gene expression profile of HCC; this information will be beneficial to investigators who may develop new target therapeutic drugs against HCC. Histamine has been demonstrated to be involved in cell proliferation and tumour growth by the activation of histamine receptors [25]. As an agonist of histamine that interacts with the H1 and H3 receptors [26], betahistine may play a role in tumour biology through the regulation of histamine receptors.

In conclusion, we identified 2271 differentially expressed genes in hepatocellular carcinoma, and the co-occurrence networks related to “metastasis”, “cell growth”, and “cell cycle” were constructed. Furthermore, we identified significant biological processes and abnormally altered pathways that were related to the development of HCC. We also screened a set of candidate small molecules, some of which may induce the initiation of HCC, and some reversed the expression profile of HCC. These small molecules may be candidates for therapeutic drugs that are capable of targeting hepatocellular carcinoma. However, the number of samples involved in this study was limited, and the analysed results contained a massive amount of

information, which requires thorough research and must be experimentally validated in future studies.

The authors declare no conflict of interest.

References

1. Parkin DM. Global cancer statistics in the year 2000. Lancet Oncol 2001; 2: 533-43.
2. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology 2007; 132: 2557-76.
3. Bosch FX, Ribes J, Diaz M, Cleries R. Primary liver cancer: worldwide incidence and trends. Gastroenterology 2004; 127: S5-S16.
4. Turdean S, Gurzu S, Turcu M, Voidazan S, Sin A. Current data in clinicopathological characteristics of primary hepatic tumors. Rom J Morphol Embryol 2012; 53: 719-24.
5. Severi T, van Malenstein H, Verslype C, van Pelt JF. Tumor initiation and progression in hepatocellular carcinoma: risk factors, classification, and therapeutic targets. Acta Pharmacol Sin 2010; 31: 1409-20.
6. Michielsen P, Ho E. Viral hepatitis B and hepatocellular carcinoma. Acta Gastroenterol Belg 2011; 74: 4-8.
7. McGivern DR, Lemon SM. Virus-specific mechanisms of carcinogenesis in hepatitis C virus associated liver cancer. Oncogene 2011; 30: 1969-83.
8. Zender L, Spector MS, Xue W, et al. Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. Cell 2006; 125: 1253-67.
9. Wang XW, Hussain SP, Huo TI, Wu CG, Forgues M, Hofseth LJ, Brechot C, Harris CC. Molecular pathogenesis of human hepatocellular carcinoma. Toxicology 2002; 181-182: 43-7.
10. Buendia MA. Genetics of hepatocellular carcinoma. Seminars in Cancer Biology 2000; 10: 185-200.
11. Golub TR, Slonim DK, Tamayo P et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999; 286: 531-7.
12. Hoefnagel JJ, Dijkman R, Basso K, Jansen PM, Hallermann C, Willemze R, Tensen CP, Vermeer MH. Distinct types of primary cutaneous large B-cell lymphoma identified by gene expression profiling. Blood 2005; 105: 3671-8.
13. Okabe H, Satoh S, Kato T et al. Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. Cancer Res 2001; 61: 2129-37.
14. Huang ZX, Tian HY, Hu ZF, Zhou YB, Zhao J, Yao KT. GenCLiP: a software program for clustering gene lists by literature profiling and constructing gene co-occurrence networks related to custom keywords. BMC Bioinformatics 2008; 9: 308.
15. Li J, Fan Y, Chen J, Yao KT, Huang ZX. Microarray analysis of differentially expressed genes between nasopharyngeal carcinoma cell lines 5-8F and 6-10B. Cancer Genet Cytogenet 2010; 196: 23-30.
16. Gene Ontology C. The Gene Ontology (GO) project in 2006. Nucleic Acids Res 2006; 34: D322-326.
17. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25-9.
18. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4: 44-57.
19. Li GC, Ye QH, Dong QZ, Ren N, Jia HL, Qin LX. TGF beta1 and related-Smads contribute to pulmonary metastasis of hepatocellular carcinoma in mice model. J Exp Clin Cancer Res 2012; 31: 93.
20. Breuhahn K, Longerich T, Schirmacher P. Dysregulation of growth factor signaling in human hepatocellular carcinoma. Oncogene 2006; 25: 3787-800.
21. Han C, Michalopoulos GK, Wu T. Prostaglandin E2 receptor EP1 transactivates EGFR/MET receptor tyrosine kinases and enhances invasiveness in human hepatocellular carcinoma cells. J Cell Physiol 2006; 207: 261-70.
22. Marhaba R, Zoller M. CD44 in cancer progression: adhesion, migration and growth regulation. J Mol Histol 2004; 35: 211-31.
23. Hirohashi K, Yamamoto T, Uenishi T, et al. CD44 and VEGF expression in extrahepatic metastasis of human hepatocellular carcinoma. Hepatogastroenterology 2004; 51: 1121-23.
24. Lara-Pezzi E, Majano PL, Yáñez-Mó M, Gómez-Gonzalo M, Carretero M, Moreno-Otero R, Sánchez-Madrid F, López-Cabrera M. Effect of the hepatitis B virus HBx protein on integrin-mediated adhesion to and migration on extracellular matrix. J Hepatol 2001; 34: 409-15.
25. Blaya B, Nicolau-Galmes F, Jangi SM, et al. Histamine and histamine receptor antagonists in cancer biology. Inflamm Allergy Drug Targets 2010; 9: 146-57.
26. Arrang JM, Garbarg M, Quach TT, Dam Trung TuongM, Yeramian E, Schwartz JC. Actions of betahistine at histamine receptors in the brain. Eur J Pharmacol 1985; 111: 73-84.

Address for correspondence

Zhongxi Huang
Institute of Oncology
Nanfang Medical University
Guangzhou, Guangdong, 510515, China
tel. +86 20 61647129
e-mail: huangzhongxi@gmail.com

Lixin Wei
Department of Pathology
Chinese PLA General Hospital
Beijing, 100853, China
tel. +86 10 66939726
e-mail: weilx301@263.net;

Submitted: 27.11.2013
Accepted: 16.07.2014