

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

journal homepage: [www.JournalofSurgicalResearch.com](http://www.JournalofSurgicalResearch.com)

## Secondary use of existing public microarray data to predict outcome for hepatocellular carcinoma

Curtis J. Wray, MD,<sup>a,\*</sup> Tien C. Ko, MD,<sup>a</sup> and Filemon K. Tan, MD, PhD<sup>b</sup>

<sup>a</sup>Department of Surgery, University of Texas Medical School at Houston, Houston, Texas

<sup>b</sup>Division of Rheumatology, Department of Medicine, University of Texas Medical School, Houston, Texas

### ARTICLE INFO

#### Article history:

Received 15 July 2013

Received in revised form

5 December 2013

Accepted 13 December 2013

Available online 6 January 2014

#### Keywords:

Hepatocellular carcinoma

Gene regulation

Chronic hepatitis

Cancer genetics

### ABSTRACT

**Background:** Since 1990, numerous public repositories of microarray data have been created to store vast genomic data sets. Our hypothesis is that a secondary analysis of an available hepatocellular carcinoma (HCC) public data set could generate new findings and additional hypotheses.

**Methods:** The Gene Expression Omnibus at the National Center for Biotechnology Information was queried for available data sets specific for ‘HCC’ and ‘clinical data.’ Genes that passed filtering and normalization criteria were analyzed using the class comparison and prediction functions in BRB-ArrayTools. Ingenuity pathway analysis software was used to identify potential gene networks up- or down-regulated.

**Results:** The file GDS274, which measured gene expression in primary HCC lesions with or without hepatic metastases from a cohort of Chinese patients, was identified as an appropriate data set and was imported into BRB-ArrayTools. 9984 genes passed filtering criteria. Clinical data demonstrated alpha fetoprotein (AFP) >100 ng/mL predictive of worse survival (HR 5.87, 95% confidence interval: 1.11–31.0). A class comparison between patients with an AFP >100 and those with AFP <100 demonstrated 92 genes to be differentially expressed. Ingenuity pathway analyses demonstrated the top networks associated with the observed gene expression.

**Conclusions:** Using available HCC microarray data, we identified genes differentially expressed based on AFP >100. Canonical pathway analysis demonstrated functional gene pathways and associated upstream regulators. This study maximizes the use of publicly available data by generating new findings. Secondary analyses of these data sets should be considered by investigators before embarking on new genomic experiments.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

High throughput genomic technologies are increasingly being used to identify therapeutic targets and risk factors for specific diseases in this era of personalized medicine [1–3]. Gene expression microarrays have been used to differentiate types of leukemia, B-cell lymphoma, breast cancer, and lung cancer

[4–6]. The use of high throughput technologies has generated vast amounts of genomic data. Since 1990, numerous public repositories of microarray data have been created. At the present time, a prerequisite to the publication of microarray data is that the results must be publicly available to the research community [7]. The data should be in a form that permits conclusions to be evaluated independently [8].

\* Corresponding author. Department of Surgery, University of Texas Medical School at Houston, Houston, TX. Tel.: +1 713 566 5095; fax: +1 713 566 4583.

E-mail address: [curtis.j.wray@uth.tmc.edu](mailto:curtis.j.wray@uth.tmc.edu) (C.J. Wray).  
0022-4804/\$ – see front matter © 2014 Elsevier Inc. All rights reserved.  
<http://dx.doi.org/10.1016/j.jss.2013.12.013>

Authors describing a newly sequenced genome, gene, or protein must deposit the primary data in a permanent, public data repository, such as the DNA Data Bank of Japan, European Bioinformatics Institute, and the National Center for Biotechnology Information (NCBI) or ArrayExpress [9].

The established databases allow researchers, at their discretion, to submit some or all of the clinical data associated with a microarray experiment. The standardization of data formatting facilitates further data analyses. This common format makes it easier for researchers to access, query, and share data [10]. Our research aims were [1] to search for a publicly available hepatocellular carcinoma (HCC) gene expression data set that also included clinical patient data and [2] to generate hypotheses using this data set.

---

## 2. Methods

### 2.1. Online search

The online Gene Expression Omnibus, a public functional genomics data repository at the NCBI (<http://www.ncbi.nlm.nih.gov/gds>) was queried for available data sets. The specific search included (“carcinoma, hepatocellular” [MeSH Terms] OR HCC [All Fields]) AND (“patients” [MeSH Terms] OR patient [All Fields]) AND (“mortality” [Subheading] OR “survival” [MeSH Terms] OR survival [All Fields]). The genomic data (GDS274) file met the search criteria and was imported into BRB-ArrayTools version 4.2 (National Cancer Institute), available at <http://linus.nci.nih.gov/BRB-ArrayTools.html> [11]. This data set represents primary lesions with or without hepatic metastases in patients with hepatitis B-induced HCC.

### 2.2. Clinical data

Deidentified available patient data were analyzed for overall survival using Kaplan–Meier survival analysis. Data included age, primary tumor size, type of surgical resection, portal vein involvement, presence of multiple tumors, cirrhosis, serum alpha fetoprotein (AFP), vital status, and survival time. A Cox proportional hazards model was used to determine the effects of multiple independent predictor variables on overall survival. The final multivariate model was created using the backward, stepwise method of covariate elimination to consider a wide range of possible best models [12]. Covariates that were significant below a *P* value <0.20 were included in the final multivariate model analysis [13]. STATA 12 (Stata-Corp, College Station, TX) statistical software was used for all analyses.

### 2.3. Microarray class comparison

Using BRB-ArrayTools, genes that passed filtering and normalization criteria were analyzed using the class comparison, which compares gene expression among predefined classes and presumes the data consists of experiments of different samples representative of the classes. We identified genes that were differentially expressed among classes using a multivariate permutation test [14–17]. The test statistics used were

random variance *t*-statistics for each gene. Although *t*-statistics were used, the multivariate permutation test is nonparametric and does not require the assumption of Gaussian distributions. In the class comparison analysis, technical replicates of the same sample were averaged.

### 2.4. Canonical pathway analysis

Interactive pathway analysis (IPA) of complex genomics data software (Ingenuity Systems, [www.ingenuity.com](http://www.ingenuity.com), Redwood City, CA) was used to examine differentially expressed genes [18,19]. The analysis settings reference set was the Ingenuity Knowledge Bases (genes + endogenous chemicals). IPA was used to assess for network-associated functions and well-characterized molecular signaling (canonical) pathways. This computational approach investigates the network behavior as a system. The Ingenuity software scans the list of input genes to identify networks (i.e., relationships between genes) using data in the Ingenuity Pathways Knowledge Base, a manually curated database of functional interactions extracted from peer-reviewed publications [20]. A Fisher exact test is performed to determine the likelihood of obtaining at least the equivalent numbers of genes by chance (i.e., from a random input gene set) as actually overlap between the input gene set and the genes present in each identified network. IPA predicts which upstream regulators are activated or inhibited, based on known relationships, to explain the up- and down-regulated genes. The IPA software describes an “upstream regulator” as any molecule that can affect the expression of another molecule.

---

## 3. Results

### 3.1. Online search

The data set GDS274 “HCC metastasis” was identified at Gene Expression Omnibus and imported into BRB-ArrayTools. The data from this microarray experiment were obtained from hepatitis B virus (HBV) positive HCC patients (*n* = 40) in China. GDS274 included primary HCC tumors and matched intrahepatic metastases (i.e., a primary tumor and an intrahepatic metastasis from the same patient). As originally published by Ye *et al.* [11], the mean patient age was 50 y (range: 36–74). The median diameter of the primary HCC was 7.2 cm (range 1.3–17.5). Thirty-two cases (80%) had underlying cirrhosis and 98% of the patients were HBV-positive. Serum AFP was >20 ng/mL in 68% of patients.

### 3.2. Clinical data

Deidentified, individual patient data were included in the GDS274 data set. A Cox proportional hazards model was created to determine predictors of survival. Age, tumor size, portal vein involvement, stage, and AFP >100 were found to have a *P* value <0.20 on univariate analysis. In the final multivariate analysis, only AFP >100 (HR 5.87) was predictive of worse survival (Table 1).

**Table 1 – Cox proportional hazards model.**

	Hazard ratio	P value	95% CI
Age	0.98	0.69	0.88–1.08
Tumor size	1.15	0.22	0.92–1.44
PV involvement	1.42	0.64	0.32–6.33
Stage	2.02	0.57	0.18–22.1
AFP >100	5.87	0.03	1.11–31.0

CI = confidence interval; PV = portal vein.

**3.3. Class comparison and IPA**

After filtering and normalization, 9984 genes passed inclusion criteria. A class comparison (modified t-test) demonstrated 92 genes were differentially expressed based on classifying patients as either (a) AFP >100 or (b) AFP ≤100. The full list of genes is available online at [http://www.uth.tmc.edu/scleroderma/Supplemental\\_data.html](http://www.uth.tmc.edu/scleroderma/Supplemental_data.html). The list of 92 genes was imported into Ingenuity software. The top nine upregulated and nine downregulated genes are listed in Table 2. The top network-associated functions were [1] neurologic disease/cardiovascular disease/heart failure (Figure) and carbohydrate metabolism/lipid metabolism/small molecule biochemistry. The top biologic functions are listed in Table 3. The top canonical pathways involved, which is defined by the ratio of observed up- or down-regulated genes that are belong to a defined pathway are tyrosine degradation I and lipopolysaccharide/interleukin 1 mediated inhibition of retinoid X receptor function (Table 4). In addition, IPA analyzed upstream regulators of the observed gene set and found peroxisome proliferator-activated receptor (PPARα) and hepatocyte nuclear factor 4 alpha (HNF4α) are upstream regulators (Table 4).

**Table 2 – Top up- and down-regulated genes.**

Upregulated genes	Fold change
HP	1.490
HSD17B6	1.450
CAMP	1.430
OTC	1.430
SLC10A1	1.430
CPS1	1.420
PLA2G1B	1.420
TAT	1.380
HPD	1.370
PAEP	1.370
Downregulated genes	Fold change
Ethylene glycol	-1.818
chloroquine	-1.429
S100P	-1.408
IGF2BP3	-1.250
SNRPN	-1.250
DUSP6	-1.205
STK17 A	-1.136
STEAP4	-1.136

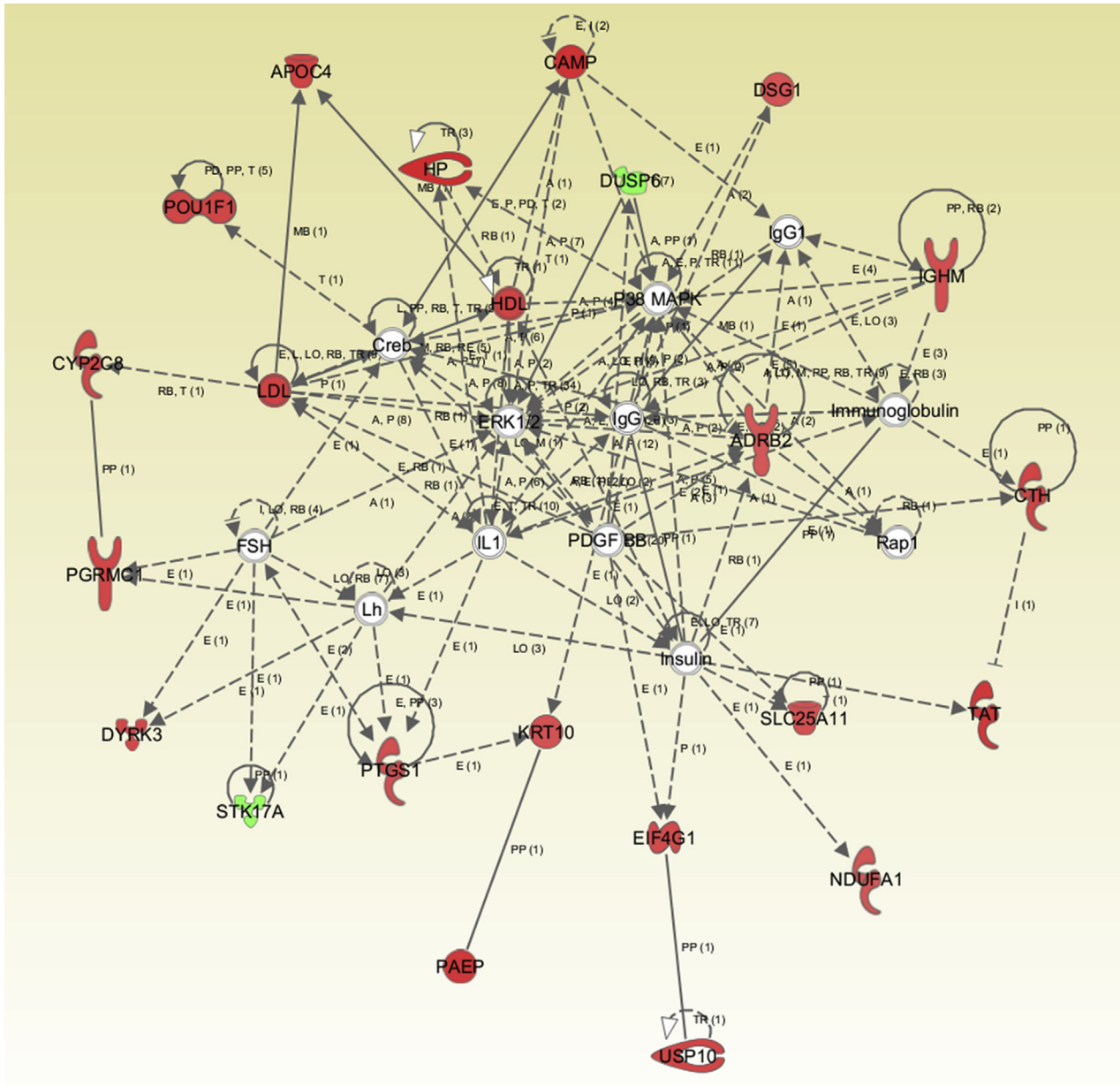
**4. Discussion**

Using an online public repository of microarray data, we performed a secondary analysis of existing HCC gene expression data. The initial microarray experiment investigated differential gene expression between primary HBV-positive HCC tumors and paired intrahepatic metastases [11]. The study authors found Osteopontin was upregulated in metastatic HCC. For our present study, we analyzed the clinical and gene expression data as well as known canonical pathways and networks. Elevated AFP >100 was [1] predictive of worse survival and [2] associated with differential gene expression. AFP has been established as a tumor marker for HCC, and serum levels predict survival; however, the regulation of AFP gene expression is less clear [21–25].

Other studies have investigated gene expression profiles predictive of survival in HCC. Hoshida et al. [26] examined primary HCC tumors and did not identify a gene profile associated with survival. However, analysis of adjacent noncancerous liver parenchyma identified a gene profile related to normal liver function highly correlated with survival. They also identified a gene signature associated with inflammation that predicted poor survival. Sun et al. [27] examined expression between primary HCC and surrounding liver tissues, which resulted in a nine-gene profile associated with cell cycle and immune response that predicted survival in HCC samples. Although there is discrepancy between many of these studies, it is important to note the clinical specimens, patient populations, experimental design, and means of measurement all differ. The samples used to produce our results are from primary tumors and matched intrahepatic metastases, a distinctly different experimental model.

Although important for prognosis, gene expression profiles predictive of survival may have limited therapeutic utility as the interaction and significant relationships are not well characterized in a simple list of up- or down-regulated genes. Since this GDS274 was created, various pathway analysis tools became available to identify biological pathways and to unravel the intricate complexity of gene expression. Using software capable of modeling and understanding genomic networks, we were able to build on the initial data gained from the significant analysis of microarrays experiment. Analyses using this approach generate lists of differential gene expression; however, the biological relevance of the list of up- or down-regulated genes is not readily apparent [28]. Using IPA to better understand the output of the list of genes, we applied our results of differential gene expression to a known gene ontology database to examine potential gene pathways and networks. This subsequent analysis provides a broad understanding of functional gene expression as it goes beyond simple gene clustering [29].

The genes we identified involve signaling pathways implicated in hepatocarcinogenesis and other candidate genes not well characterized. We found S100p to be downregulated, and this gene has been studied in various gastrointestinal cancers [30,31]. Further investigation of the IPA network data, illustrated in Figure, depicts interactions with transcription factors. In turn, a transcription factor, such as



**Figure – Neurologic disease, cardiovascular disease, and heart failure network. Caption: solid line = direct interaction, dashed line = indirect interaction, A = activation, B = binding, E = expression, I = inhibition, LO = localization, M = biochemical modification, P = phosphorylation, PP = protein–protein binding, RB = regulation of binding, T = transcription, and TR = translocation.**

nuclear factor-kappa B, interacts with mitogen-activated protein kinase transcription factors, which have been shown to increase the proliferation and invasion of HCC cells *in vitro* [32]. The IPA demonstrated upstream regulators (PPAR $\alpha$  and HNF4 $\alpha$ ) of the AFP >100 class may also interact with these transcription factors. Sustained activation of PPAR $\alpha$  by agonists has been linked to HCC due to sustained oxidative stress, endoplasmic reticulum stress, and liver cell proliferation [33]. HNF4 $\alpha$  is a nuclear transcription factor essential for liver development and hepatocyte function [34]. These preliminary observations may generate hypotheses for future studies by providing investigators with candidate molecular targets for novel therapeutic agents.

There are several important limitations of this study. Foremost, the researcher must rely on secondhand data and has no ability to double check the internal validity of the data. However, each microarray chip has internal controls as a quality control measure. Second, the generalizability of the results to the U.S. patients is unknown because >70% of HCC patients in the U.S. present with advanced cirrhosis. This data set included Chinese HCC patients that were primarily HBV-positive, which is in contrast to HCC patients in the U.S. that are typically hepatitis C virus positive. Third, in this data set, surgical resection was the only therapy provided, thus additional locoregional or systemic therapies could not be included as covariates. Since all patients included in this

**Table 3 – Top biologic functions.**

	P value	Number of molecules
<b>Diseases and disorders</b>		
Developmental disorder	$4.55 \times 10^{-9}$ to $2.11 \times 10^{-2}$	16
Hereditary disorder	$4.55 \times 10^{-9}$ to $1.93 \times 10^{-2}$	22
Metabolic disease	$4.55 \times 10^{-9}$ to $8.48 \times 10^{-3}$	10
Immunologic disease	$6.22 \times 10^{-5}$ to $2.11 \times 10^{-2}$	17
Inflammatory response	$1.77 \times 10^{-4}$ to $2.11 \times 10^{-2}$	14
<b>Molecular and cellular function</b>		
Amino acid metabolism	$6.86 \times 10^{-9}$ to $2.11 \times 10^{-2}$	14
Small molecule biochemistry	$6.86 \times 10^{-9}$ to $2.11 \times 10^{-2}$	41
Energy production	$8.63 \times 10^{-6}$ to $2.11 \times 10^{-2}$	13
Lipid metabolism	$8.63 \times 10^{-6}$ to $2.11 \times 10^{-2}$	25
Carbohydrate metabolism	$3.02 \times 10^{-5}$ to $2.11 \times 10^{-2}$	15
<b>Development and function</b>		
Endocrine system	$8.63 \times 10^{-6}$ to $8.48 \times 10^{-3}$	5
Connective tissue	$3.34 \times 10^{-4}$ to $1.27 \times 10^{-2}$	4
Organismal survival	$4.90 \times 10^{-4}$ to $5.28 \times 10^{-4}$	12
Cardiovascular system	$5.57 \times 10^{-4}$ to $2.11 \times 10^{-2}$	9
Tissue morphology	$5.57 \times 10^{-4}$ to $2.11 \times 10^{-2}$	12

study underwent surgical resection, most patients likely had preserved hepatic function with mild cirrhosis. Lastly, long-term follow-up of patients in this data set is not available.

In the initial publication by Ye et al. [11], Osteopontin was found to be upregulated in metastatic HCC. In our current results, there was no correlation between Osteopontin and the

**Table 4 – Top canonical pathways and upstream regulators.**

Gene name	P value	Ratio
Tyrosine degradation I	$3.22 \times 10^{-5}$	3/15 (0.20)
LPS/IL-1 mediated inhibition of RXR function	$4.17 \times 10^{-5}$	7/239 (0.03)
Xenobiotic metabolism signaling	$1.39 \times 10^{-3}$	6/299 (0.02)
Noradrenaline and adrenaline degradation	$1.4 \times 10^{-3}$	3/52 (0.06)
Oleate biosynthesis II	$2.61 \times 10^{-3}$	2/18 (0.11)
<b>Upstream regulator</b>	<b>P value of overlap</b>	<b>Predicted activation state</b>
Methylprednisolone	$7.62 \times 10^{-10}$	Activated
Dexamethasone	$2.98 \times 10^{-7}$	
Ciprofibrate	$1.13 \times 10^{-6}$	
PPAR $\alpha$	$1.97 \times 10^{-6}$	
HNF4 $\alpha$	$1.83 \times 10^{-5}$	

IL-6 = interleukin 6; LPS = lipopolysaccharide; Ratio = percentage of genes involved in the canonical pathway; RXR = retinoid X receptor.

list of upregulated genes in the AFP >100 class. There may be several reasons for this observation. In the original study, Osteopontin was found to be upregulated when global gene expression was compared between 10 primary HCC lesions and 10 HCC lesions with portal vein tumor thrombus. Our study used the same data set, but included all samples from 40 patients and technical replicates from the same patient were averaged.

Currently, the Functional Genomics Data Society requires all authors using microarray data to submit a complete data set to the NCBI [35]. This is to be compliant with the Minimum Information About a Microarray Experiment standard [8]. Despite the availability of public access to large data sets, secondary analyses are rarely performed [36]. This study maximizes the utility of available data by generating additional findings and hypotheses. Subsequent studies are needed to validate these results. Nonetheless, secondary analyses of existing microarray and clinical data sets should be considered by investigators before embarking on new, costly genomic experiments. Resulting gene profiles may be useful in elucidating mechanisms of carcinogenesis, identifying novel therapeutic targets, and stratifying patients in clinical trials. Considering the enormous amount of genomic data stored in public repositories, further analyses of this data with newer software may prove useful.

### Acknowledgment

Author contributions: W.J.C. was responsible for conception and study design and data collection, manuscript composition. K.C.T. and T.K.F. were responsible for scientific review and manuscript revision. W.J.C. and T.K.F. were responsible for data analysis and interpretation.

Financial support: Funded in part by American Cancer Society (MSRG-12-178-01-PCSM).

### REFERENCES

- [1] Liu X, Niu T, Liu X, et al. Microarray profiling of HepG2 cells ectopically expressing NDRG2. *Gene*; 2012.
- [2] Joyce T, Oikonomou E, Kosmidou V, et al. A molecular signature for oncogenic BRAF in human colon cancer cells is revealed by microarray analysis. *Curr cancer drug targets* 2012;12:873.
- [3] Ong PS, Chan SY, Ho PC. Microarray analysis revealed dysregulation of multiple genes associated with chemoresistance to As(2)O(3) and increased tumor aggressiveness in a newly established arsenic-resistant ovarian cancer cell line, OVCAR-3/AsR. *Eur j pharm sci* official j Eur Fed Pharm Sci 2012;45:367.
- [4] Kao KJ, Chang KM, Hsu HC, Huang AT. Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization. *BMC cancer* 2011;11:143.
- [5] Szczepanek J, Styczynski J, Haus O, Tretyn A, Wysocki M. Relapse of acute lymphoblastic leukemia in children in the context of microarray analyses. *Archivum immunologiae et therapiae experimentalis* 2011;59:61.

- [6] Neumann J, Feuerhake F, Kayser G, et al. Gene expression profiles of lung adenocarcinoma linked to histopathological grading and survival but not to EGF-R status: a microarray study. *BMC cancer* 2010;10:77.
- [7] Oliver S. On the MIAME standards and central repositories of microarray data. *Comp funct genomics* 2003;4:1.
- [8] Brazma A. Minimum information about a microarray experiment (MIAME)—successes, failures, challenges. *TheScientificWorldJournal* 2009;9:420.
- [9] Barrett T, Troup DB, Wilhite SE, et al. NCBI GEO: archive for high-throughput functional genomic data. *Nucleic acids res* 2009;37:D885.
- [10] Stoeckert CJ Jr, Causton HC, Ball CA. Microarray databases: standards and ontologies. *Nat Genet* 2002;32(Suppl):469.
- [11] Ye QH, Qin LX, Forgues M, et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nat med* 2003;9:416.
- [12] Harrell WA, Spaulding LM. Social psychological models of choice behavior and drivers' left turns. *J Soc Psychol* 2001; 141:714.
- [13] Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. *Bmj* 2009;338:b604.
- [14] Simon RM. Design and analysis of DNA microarray investigations. New York: Springer; 2003.
- [15] Yang YH, Dudoit S, Luu P, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic acids res* 2002;30:e15.
- [16] Wright GW, Simon RM. A random variance model for detection of differential gene expression in small microarray experiments. *Bioinformatics* 2003;19:2448.
- [17] Korn EL, Li MC, McShane LM, Simon R. An investigation of two multivariate permutation methods for controlling the false discovery proportion. *Stat med* 2007;26:4428.
- [18] Siripurapu V, Meth J, Kobayashi N, Hamaguchi M. DBC2 significantly influences cell-cycle, apoptosis, cytoskeleton and membrane-trafficking pathways. *J Mol Biol* 2005;346:83.
- [19] Raponi M, Belly RT, Karp JE, et al. Microarray analysis reveals genetic pathways modulated by tipifarnib in acute myeloid leukemia. *BMC cancer* 2004;4:56.
- [20] Calvano SE, Xiao W, Richards DR, et al. A network-based analysis of systemic inflammation in humans. *Nature* 2005; 437:1032.
- [21] Marrero JA, Feng Z, Wang Y, et al. Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. *Gastroenterology* 2009;137:110.
- [22] Kohles N, Nagel D, Jungst D, et al. Prognostic relevance of oncological serum biomarkers in liver cancer patients undergoing transarterial chemoembolization therapy. *Tumour biol j Int Soc Oncodevelopmental Biol Med* 2012;33: 33.
- [23] Beale G, Chattopadhyay D, Gray J, et al. AFP, PIVKAI, GP3, SCCA-1 and follistatin as surveillance biomarkers for hepatocellular cancer in non-alcoholic and alcoholic fatty liver disease. *BMC cancer* 2008;8:200.
- [24] Yuen MF, Lai CL. Serological markers of liver cancer. *Best practice & research. Clin gastroenterol* 2005;19:91.
- [25] Zhang L, He T, Cui H, et al. Effects of AFP gene silencing on apoptosis and proliferation of a hepatocellular carcinoma cell line. *Discov med* 2012;14:115.
- [26] Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *The New Engl j med* 2008;359:1995.
- [27] Sun M, Wu G, Li Y, et al. Expression profile reveals novel prognostic biomarkers in hepatocellular carcinoma. *Front biosci* 2010;2:829.
- [28] Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98:5116.
- [29] Altman RB, Raychaudhuri S. Whole-genome expression analysis: challenges beyond clustering. *Curr opin struct biol* 2001;11:340.
- [30] Hamada S, Satoh K, Hirota M, et al. Calcium-binding protein S100P is a novel diagnostic marker of cholangiocarcinoma. *Cancer Sci* 2011;102:150.
- [31] Kim JK, Jung KH, Noh JH, et al. Targeted disruption of S100P suppresses tumor cell growth by down-regulation of cyclin D1 and CDK2 in human hepatocellular carcinoma. *Int J Oncol* 2009;35:1257.
- [32] Wu R, Duan L, Ye L, et al. S100A9 promotes the proliferation and invasion of HepG2 hepatocellular carcinoma cells via the activation of the MAPK signaling pathway. *Int J Oncol* 2013; 42:1001.
- [33] Pyper SR, Viswakarma N, Yu S, Reddy JK. PPARalpha: energy combustion, hypolipidemia, inflammation and cancer. *Nucl receptor signaling* 2010;8:e002.
- [34] Hatziaepostolou M, Polytaichou C, Aggelidou E, et al. An HNF4alpha-miRNA inflammatory feedback circuit regulates hepatocellular oncogenesis. *Cell* 2011;147:1233.
- [35] Quackenbush J. Data reporting standards: making the things we use better. *Genome med* 2009;1:111.
- [36] Ventura B. Mandatory submission of microarray data to public repositories: how is it working? *Physiol genomics* 2005;20:153.