

COMMENTARY

Gene expression profiling of preneoplastic liver disease and liver cancer: a new era for improved early detection and treatment of these deadly diseases?

Jin Woo Kim and Xin Wei Wang¹

Building 37, Room 2C25, Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD 20892-4255, USA

¹To whom correspondence should be addressed
Email: xw3u@nih.gov

Hepatocellular carcinoma (HCC) is a multi-step process associated with changes in gene expression. Currently, several technologies enable global gene expression profiling. The number of studies to probe global gene expression profiles of HCC or preneoplastic chronic liver diseases has increased exponentially in recent years. These studies have quickly provided rich information and some additional clues to the genesis of liver cancer. The application of gene expression profiling to preneoplastic liver diseases and HCC is growing in importance and practicality. In this commentary, we review the recent advances in the utilization of global gene expression profiling to liver cancer, which have provided new insight into the molecular mechanisms underlying the development of HCC. We have also discussed the problems related to these new technologies, as well as their contributions and implications. By recognizing the shortcomings, we can reassess our current approaches, which allow us to better design and analyze global gene expression-based experiments. These new approaches will undoubtedly contribute to a better understanding of hepatocarcinogenesis.

Liver cancer is the third most deadly cancer worldwide (estimated 548 554 deaths in 2000), and the fifth in the number of cases (1). Hepatocellular carcinoma (HCC) is one of the most common malignant liver tumors in the world with a high prevalence in Asia and sub-Saharan Africa (2). Recent studies have shown that the incidence of HCC has substantially increased in the USA as well as in other areas including Japan and Europe (3–5). HCC is one of the few human cancers in which an underlying etiology can often be identified in most cases. It is frequently associated with chronic liver diseases including cirrhosis (6). Most of the etiological factors associated with HCC, such as chronic viral hepatitis, alcohol abuse, metabolic disorders or other environmental agents, also lead to cirrhosis (7). However, it is not clear whether these factors induce HCC directly or whether they act indirectly by producing chronic liver injury and cirrhosis.

The molecular mechanisms of hepatocarcinogenesis are not well understood, although aneuploidy, and multiple genetic alterations are often present. For example, mutations of p53, Rb and β -catenin have been reported in HCC, while c-myc

Abbreviations: CCP, compound covariate predictor; DD, differential display; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; MMP, matrix metalloproteinase; PCR, polymerase chain reaction; RDA, representational difference analysis; SAGE, serial analysis of gene expression; SSH, suppression subtractive hybridization.

and cyclin D1 are frequently overexpressed (8–12). Loss of heterozygosity (LOH) may be associated with the inactivation of tumor suppressor genes. LOH at 1p, 4q, 6q, 8p, 13q, 16q and 17p loci are frequently reported in human HCC (13–16). In addition, several growth factors have been implicated in the development of HCC, including transforming growth factor α and β (17,18). Unfortunately, these fragmented findings do not piece together a clear portrait of HCC. This disease has been referred to as very heterogeneous (19).

With the recent development of the gene expression analysis methodologies, gene expression on a global scale can be determined in tissue samples and cell lines. The use of gene expression profiling is particularly important in cancer. This is because the initiation and malignant progression of cancers result from the accumulation and combining effects of many changes in the sequence or expression level of cancer initiation genes.

In this paper, we review recent progress on gene expression profiling of HCC and discuss problems and contributions as well as their implications.

Technologies to analyze global gene expression

Several techniques are now available to monitor gene expression on a genomic scale. Differential display (DD) is the most common technique used in the last 10 years to examine differences in gene expression between two samples (20). This technique identifies differentially expressed mRNAs based on their ability to hybridize with specific primer sequences. In DD, multiple biological samples are amplified in parallel, resolved by electrophoresis and visualized. Several investigators have used the DD technique to analyze hepatocarcinogenesis (21–24).

Suppression subtractive hybridization (SSH) is a recently developed technology designed for identifying differentially expressed genes (25). One critical feature of this technique is the combination of suppression polymerase chain reaction (PCR) with subtraction. To optimize the subtraction, the cDNA is digested with a restriction enzyme to generate small DNA fragments (~500 nucleotides in length). Specially designed DNA adapters are ligated onto one of the cDNA pools, which are followed by two rounds of hybridization and PCR. Miyasaka *et al.* used this technique to determine changes in gene expression profiles in HCC (26).

Representational difference analysis (RDA) is a variant of differential hybridization that incorporates a PCR amplification step (27). The procedure is based on the generation of two complex cDNA populations that are cut, ligated to suitable adapters and subsequently amplified by PCR. The products represent the RNA populations at reduced complexity (therefore, called representations) and are the starting material for the comparison. Differences are accumulated by a combined process of subtraction and kinetic enrichment. After repeated cycles of RDA, the differentially expressed products are cloned and sequenced. This method was used to identify CRG-L1

(musculus cancer related gene-liver 1) in a mouse model of HCC (28).

The DD, SSH and RDA methods provide some sensitivity in identifying novel genes that may be associated with hepatocarcinogenesis. However, these methods are not quantitative and are labor-intensive. The developments of Serial Analysis of Gene Expression (SAGE) and cDNA microarray techniques, for the first time have allowed the quantitative measurement of gene expression on a global scale.

SAGE was first described by Velculescu *et al.* in 1995 (29). In SAGE, a short unique sequence tag is isolated from each gene by a PCR-based strategy. Concatemered tags are sequenced, and the abundance of these tags provides a measure of the level of gene expression in the starting material. These SAGE tags can be linked to a specific transcript designation in an appropriate database of unique transcripts, such as UniGene. SAGE is essentially an accelerated technique for cDNA library sequencing. Because it requires intensive sequencing, SAGE is not well suited for the analysis of large numbers of samples. However, because SAGE does not require a prior knowledge of the pattern of gene expression in a given mRNA source, the same biochemical and bioinformatics procedure can be applied to any sample, given the availability of the appropriate reference database of SAGE tags.

In 1995, Brown *et al.* at Stanford University published the first paper on cDNA microarrays, describing them as a high-capacity system developed to monitor RNA levels of numerous genes simultaneously, using two-color fluorescence (30). cDNA microarrays consist of thousands of different cDNA clones spotted in an array manner on a glass slide or a nylon membrane. These slides or membranes can be hybridized with two cDNA probes made from two different samples, each fluorescently labeled with spectrally distinct dyes. The ratio of fluorescent signal intensity at any location represents the ratio of corresponding mRNA molecules in the two samples hybridized to the array. In contrast to other techniques, microarray can only measure the expression of genes that correspond to sequences included in the array fabrication process. The advantage of cDNA microarrays is that it does not require a prior knowledge of a cDNA sequence because clones can be used and sequenced later if they prove of interest.

Instead of using one cDNA PCR product to capture targets representing a particular gene of interest, multiple oligonucleotide probes have been designed from various regions along the 3' end of the gene (31). One advantage of using oligonucleotides for microarray probes is their increased specificity when compared with cDNA clones. Because oligonucleotide arrays are designed and synthesized based on sequence information, physical intermediates such as cloning and PCR are not required. Specific sequences, which are non-overlapping, can be designed to increase the hybridization sensitivity, even with shorter sequences. Furthermore, oligonucleotides can distinguish between single nucleotide polymorphisms and splice variants. Current cDNA microarrays contain up to 32 000 transcripts, and oligonucleotide arrays can contain up to 49 000 unique sequences.

Examples of molecular profiling of HCC

Hepatocarcinogenesis is a long-term process (19). During the long preneoplastic stage leading to HCC, alterations in gene expression are almost entirely quantitative. Thus, gene expression profiling of preneoplastic stage liver will be indispensable

to understanding the molecular mechanism of HCC. Several groups have analyzed global gene expression in preneoplastic liver diseases (32,33). Honda *et al.* compared the gene expression profiles of chronic hepatitis B and C patients by cDNA microarray (32). SAGE has also been applied to liver disease and HCC (33,34). Because hepatitis B virus (HBV) is a major risk factor contributing to the development of HCC, several groups have compared gene expression profiles in cultured liver cells expressing HBV or the HBV-encoded oncogene, HBx (34–36). We have performed both microarray and SAGE analyses on freshly isolated human primary hepatocytes transiently expressing HBx (34,35). This system served as an ideal experimental control as the relatively low abundance of HBx transcripts (0.14%) in the HBx-infected hepatocyte library is very similar to the abundance of HBx (0.11%) from the HBV-positive HCC library (T.Yamashita, personal communication). Using SAGE, we have identified 31 novel transcripts that are potential targets for HBx. These studies also allow us to generate the hypothesis that HBx may function as a major regulator in common cellular pathways that, in turn, regulate protein synthesis, gene transcription and protein degradation.

Molecular profiling has been successfully used to identify candidate genes for HCC in human and animal model systems. We have listed selected references to these studies (Table I).

Several groups have used cDNA microarray to identify genes that regulate the composition of the extracellular matrix and the cytoskeleton. Matrix metalloproteinase 14 (MMP14) and osteonectin (SPARC) are up-regulated in HCC, which is consistent with previous findings (37–40). MMPs play a crucial role in tumor invasion by degrading basement membranes and the stromal extracellular matrix. Osteonectin is a glycoprotein involved in extracellular matrix remodeling. Large amounts of osteonectin mRNA and protein have been detected in the tumor capsule, in the fibrous bands and along capillaries within HCCs.

Rho B (ARHB) is down-regulated in HCC, but another Ras homolog gene family, member A (ARHA) shows an opposite trend (36–38,41). These results are consistent with another recent result showing that the Ras/Erk and the Ras/RhoA pathways negatively regulate cytokine-induced NOS2 in the normal human liver cell line, while RhoB enhances it (42). This is the first demonstration that genes promoting malignant transformation such as Ras and RhoA inhibit NOS2 induction, while genes with tumor suppressor activity such as RhoB enhance the process. Also Rho E, another Ras homolog gene family member E (ARHE), is down-regulated as several studies have reported (37,41,43).

Major histocompatibility complex class I C (HLA-C) is up-regulated in HCC. This result has been confirmed through two separate reports using different techniques (33,44). Class I genes belong to a large multigene family within the major histocompatibility complex (MHC) including HLA-A, HLA-B and HLA-C. There are several reports about MHC and HCC (45,46). Results of these studies have shown that HLA is up-regulated in both viral hepatitis and HCC.

Dyneins are molecular motors that translocate along microtubules. The expression of cytoplasmic dynein light chain 1 (PIN, *ddlc1*) gene is enhanced in HCC, and in chronic hepatitis B or C liver (32,44). An anti-apoptotic effect of this gene has been demonstrated recently in a study of *Drosophila* (47). Furthermore, its interaction with I κ B has been reported, and its involvement in nuclear factor- κ B-dependent gene regulation has been suggested (48). Therefore, this gene

Table 1. Molecular profiling of liver diseases.

Disease	Object	Method	Genes up-regulated	Genes down-regulated	Ref.
HCC	Human	cDNA	SPARC, MMP14, VIM	SOD1, VTN, MST1, ARHB, ARHE	37
HCC	Human	cDNA		IGFBP3, FOS, HGFAL, ARHE, A2M	43
HCC	Human	cDNA	165 genes (MMP11)	170 genes (IGFBP3, CCND1, TP53)	53
HCC	Human	cDNA	MMP14, SPARC, ARHA	DCN, ARHB, IGFBP3, 4, CDKN1A	38
HCC	Human	cDNA	MME		70
HCC	Human	cDNA	84 genes	72 genes	71
HCC	Human	cDNA	884 genes	1369 genes	72
HCC	Human	cDNA	10 genes (PIN, HLA-C, RPL6, ARHB)	9 genes	44
HCC	Human	cDNA	GTR2-2, RPL6, RPS16, CKYP, G3PDH,	147 genes	73
HCC	Human	cDNA	1640 genes		41
Cell lines	Human	cDNA	MYC, IRS-1, FGFR1, PKC	CDKN1A, EGFR	75
Cell lines	Human	cDNA	11 genes	7 genes	76
HepG2+HBx	Human	cDNA	IGFR2, ARHA, MLK3, HSP27, BAK, CDKN1A		36
Hepatocyte+HBx	Human	cDNA	22 genes	15 genes	35
Hepatitis B, C	Human	cDNA	IFNA, IRF7, ILF1, PIN, CDKN2A	IGFBP1, CDC27, VTN	32
HCC	Human	genome DNA	CCND1, SAS, MET, TERC, FGF3,MYC		74
HCC	Human	DD	ER-1, MVDP, AR		21
HCC	Human	DD	RP S8, L12, L23a, L27, L30		23
HCC	Human	DD	TM4SF3	CD9	24
HCC	Human	DD	29 genes	27 genes	22
HCC	Human	oligo	842 genes	393 genes	77
Huh7+HCV core	Human	oligo	MT, NNMT, JUNB, JUND, OPN, GPLP		78
HCC	Human	oligo			79
Hepatitis, HCC	Human	SAGE			33
Hepatocyte+HBx	Human	SAGE	57 genes	46 genes	34
HCC	Human	SSH	FAK, GS,PGC, M130, DCC, OAT	DCN	26
HCC, cell lines	Human	RDA		DCL-1	51
cell line	Rat	cDNA, SSH			80
tumor	Mouse	oligo, RDA			28
HCV infection	Chimpanzee	oligo	327 genes		81

HCV, hepatitis C virus.

may be involved, not only in the cell morphology of HCC, but also in carcinogenesis as a result of its anti-apoptotic effect.

SSH analysis of HCC has identified decorin as being down-regulated in HCC (26). Decorin (DCN) is a small proteoglycan, and is known to bind to and inhibit transforming growth factor- β (TGF- β). It may directly interfere with the cell cycle via the induction of cyclin-dependent protein kinase, p21 (49). It has been reported that DCN suppresses tumorigenicity when expressed in colon cancer cells (50). These results suggest that DCN might have an inhibitory effect on HCC tumorigenesis or progression.

RDA has identified the deleted liver cancer 1 (DLC1), a gene commonly deleted in HCC (51). DCL1 exerts an inhibitory effect on HCC cell growth (52).

Several insulin-like growth factor binding proteins are differentially expressed in HCC and non-cancer samples (38,41,43,53). IGFBP-3 was reported to be a growth suppressor in various pathways. In the IGF receptor-dependent pathway, IGFBP-3 mediates a wide variety of growth suppression signals such as TGF- β , retinoic acid, TNF- α and p53, among others. Reduced expression of IGFBP-3 has been reported in HCC. Also, other family members, i.e. IGFBP-1 and 4 have been shown to be

differentially expressed in HCC when compared with normal liver (32,38).

Potential pitfalls and ways to improve expression-profiling techniques

Problems with sample choice and processing

Gene expression profiling of tissues is probably dependent on several factors including individual variation and sample variation. The ethnicity, sex, age and genetic background of a given patient probably affect gene expression. Potential errors caused by those biological variations can be minimized only by increasing the number of subjects included in an analysis.

The selection of an appropriate control as a reference is an important element for microarray study. The reference serves to control variations in the size of corresponding spots on different arrays and variations in sample distribution over the slide (54). In most liver cancer studies, the basic design of gene expression profiling is to compare normal with cancer tissues. The debatable issue is whether adjacent non-HCC liver tissue samples from the same patient should be used when comparing it with HCC or whether a common reference

derived from a cell line pool or pooled normal liver tissues from healthy donors. Use of non-cancerous tissues from the same patient can provide a clue for individual variation due to the difference in genetic background. However, tissues adjacent to an area of HCC may not be normal, despite the absence of gross histopathological evidence indicative of cancerous changes in apparently normal tissues. Adjacent tissues near a cancer could be genetically altered or exhibit an altered gene expression profile. There may be a significant difference in the conclusions reached by two similar expression-profiling studies when different references are used.

Liver cancer may be a heterogeneous mixture of different cell types. This heterogeneity can complicate the interpretation of gene expression studies. Thus, sample selection is an important issue. The recently developed technique of laser capture microdissection can be used to isolate a defined cell population from specific areas of tissues under direct microscopic visualization (55). Although this can be a time-consuming process, several groups have been able to apply it to microarray analysis (56–58). For example, Okabe *et al.* has used this technique for cDNA microarray analysis of HCC (53). However, laser capture microdissection generates a very small amount of RNA, which then needs to be amplified to attain the sensitivity required by microarray. Two methods have been developed to accommodate this need. One approach uses an RT-PCR based method that can be performed on single cells, but suffers from the observation that cDNA abundance does not correlate with original mRNA levels. The other uses a linear amplification method based on cDNA synthesis by reverse transcription from primers that are tagged with a bacteriophage RNA polymerase promoter sequence (57). This approach has been reported to be reproducible and allows the use of as little as 1–50 ng of total RNA. However, whether amplified RNA samples produced by these techniques are a true representation of the original RNA population still remains to be determined.

We have performed microarray analysis using multiple samples from the same chronic hepatitis B patient (manuscript in preparation). These samples had different histological features under the microscope. Hierarchical clustering analysis of global gene expression was used to compare these samples with other samples from different patients. The samples from the same patient always clustered together and were well separated from other samples of different patients. These results indicate that sample variation is not as significant a problem as variation between individuals, and suggest that gross-dissected liver samples without laser capture microdissection are sufficient for microarray analysis.

Statistical methods

Data management is an essential element in global gene expression profiling. To date, various approaches have been developed for the analysis and exploration of gene expression data, particularly microarray data. However, investigators are confronted with the problem of deciding which expression ratios to regard as significant because there are no standard criteria for the selection of differentially expressed genes. Most investigators typically apply an arbitrary global threshold of a 2–4-fold change for differences in expression that might be considered biologically interesting or significant. While it seems reasonable to assume that the largest changes in mRNA levels are the ones most likely to be biologically significant, there are problems with this approach. Two-fold differences

in the ratio of some mRNAs could have more biological impact than 3-fold differences in the expression of other mRNAs. Fold threshold selection does not have any statistical validation (59). For example, in a 10 000 gene array, one would expect to have over 500 non-specific genes that appear differentially expressed by chance at the 95% significant level, regardless of the correlation patterns of the genes (54). Therefore, it is essential to select genes with some statistical confidence to avoid non-specificity.

Several statistical algorithms have been developed, which allow some confidence in assigning significance to candidate genes. Statistical analysis of gene expression data has centered on two approaches, unsupervised and supervised algorithms. Unsupervised methods require no prior knowledge on the sample classification. They are geared towards the discovery of patterns in the data unbiased by outside knowledge. One of the most frequently used unsupervised analysis methods is the hierarchical clustering developed by Eisen *et al.* (60). However, supervised algorithms require the conditions to be associated with labels that provide information about a pre-existing classification. This information comes from outside the gene expression experiment and might include knowledge of disease subtype or tissue origin of a cell type. Supervised algorithms can be used to identify small changes among the groups that cannot be distinguished by unsupervised methods.

One of the goals of supervised expression data analysis is to construct classifiers, such as linear discriminant analysis, support vector machines, artificial neural networks or K-nearest neighbors (61–64). To date, supervised analysis algorithms have not been used for HCC cases. Recently, we have compared gene expression profiles between metastatic HCC liver samples and non-metastatic HCC by supervised algorithms (65). We used cDNA microarray to investigate the gene expression profiles of primary and metastatic tumors from surgical specimens of 40 HCC patients. Gene expression profiles were obtained by comparing primary or metastatic HCC with its corresponding adjacent non-tumorous tissues. While unsupervised methods did not identify any difference, the supervised compound covariate predictor (CCP) analysis was used to classify primary and metastatic samples, and to identify genes that discriminate between these two groups (66). CCP analysis correctly grouped 19 of 20 primary HCC samples with metastatic lesions and nine of 13 primary HCC without metastasis. The cross-validated misclassification rate was significantly lower than expected by chance ($P = 0.002$). In contrast, CCP analysis did not lead to a statistically significant difference between primary and metastatic HCC sample pairs ($P = 0.328$). These results indicate that the alterations of metastasis-associated genes may occur in primary lesions with metastatic potential. This analysis may allow us to build a training module useful for the diagnosis and prognostic assessment of advanced HCC patients and to identify metastasis-associated candidate genes that may be used as potential targets for molecular therapy. These gene expression based classifiers offer a potential for adjuvant therapy.

New areas of gene expression profiling of HCC

Early detection markers and diagnosis

Most HCC patients currently identified in clinics are often at an advanced stage of the disease with a very poor prognostic outcome. Therefore, early detection, using a tumor marker, is essential to reduce the fatality of this disease.

An ideal tumor marker should be specific for that particular type of cancer, and produced only by it and not by any non-malignant condition. It should be absent or present in low amounts in normal individuals, but become elevated in the presence of a very small tumor. The biological half-life should be short, and the quantity of the marker should correlate with tumor load so that it can be used in staging the disease and monitoring an individual's response to treatment. It should also be easily and rapidly measurable by available laboratory methods. While alpha-fetoprotein is the best marker presently available and remains the most widely used, it fulfils only some of these criteria.

The use of gene expression profiling for cancer diagnosis has been demonstrated in acute leukemia samples by oligonucleotide microarray analysis technique (67). Using unsupervised learning, leukemia samples were neatly clustered into the two known subsets solely on the basis of gene expression. In addition, using supervised learning, gene sets differentially expressed in the two subsets were used to correctly classify a group of known samples into the correct categories. This study provided strong evidence that tumor expression profiles can be used for cancer classification. Thus far, this application has not been used for HCC. Recently, we have performed microarray analysis on several chronic liver diseases, and have applied supervised statistical algorithms, including leave-one-out cross-validated CCP and k-nearest neighbor classifications. These approaches have allowed us to successfully classify these preneoplastic liver diseases according to their risk for developing HCC (68). The development of such molecular signatures may provide a platform that would allow an early diagnosis of HCC patients, and the genes in this classifier may help us obtain a better understanding of the initiation of HCC.

Outcome prediction and drug development

Another promising application of global gene expression profiles in oncology is in the development of new anticancer treatments. One obvious application relies on the identification of new genes and pathways involved in cancer progression that could become therapeutic targets. Large-scale analysis of the changes in gene expression, induced by treatments or by candidate drugs, may reveal the participation and behavior of both known and previously unsuspected genes. Such studies will allow the identification of new molecular targets downstream of the primary one, and will also help to predict the potential toxicity of treatments. Using cDNA arrays, Huang *et al.* have monitored the differences in gene expression between untreated melanoma cells and the same cells treated with a combination of recombinant interferon and the anti-leukemic compound (69). They have identified the expression changes that correlate with and potentially control these processes.

Molecular characterization of cancers by the global gene expression profile approach may allow for the earlier detection of cancers at a stage when they may be responsive to treatment. It may also be used to identify high-risk individuals or screen relatives in families with inherited cancers. Gene expression-profiling techniques could distinguish neoplastic from benign tumors when histological methods are inconclusive. It may add prognostic information to traditional methods of patient staging to improve the prediction of clinical outcome. It may also be able to predict the needed response to treatment so that treatment approaches can be individualized. In addition, it may lead to the development of novel treatment strategies.

In general, HCC is considered to be a fatal disease because of its poor prognosis, with the exception of a few patients who have received liver transplantation. This is largely because of the lack of a method for early diagnosis, and the lack of information on the phenotypic changes associated with the development of HCC. Changes in gene expression profiles during the genesis of HCC are also largely unknown. The comparison of global gene expression patterns between normal liver and tumor samples may allow us to understand better the molecular mechanism of hepatocarcinogenesis.

Conclusions

Global gene expression profiling, for the first time, will assess molecular changes associated with liver cancer on a global scale. The newly developed state-of-the-art technologies such as cDNA microarray and SAGE, demonstrate a promising future for the study of liver cancer. However, at present, there is a lack of comparability among these studies, which still precludes a reasonable interpretation for a consensus change that allows us to depict a true portrait of HCC on a global scale. This is largely from, not only the lack of standardization, but also the lack of a good biostatistical approach and the use of different techniques. A clear objective for the design of global gene expression profiling, in combination with good statistical tools, can offer a solid foundation that will yield a meaningful outcome. The two emerging focuses should be considered in the field of liver cancer study. The first need is to develop a gene-expression-based model that allows for identifying individuals with a potential for developing HCC. One way is to compare gene expression profiles between the high-risk and rare-risk groups of chronic liver disease patients with cirrhosis, as well as HCC patients, and to identify a set of significant genes that can be used to classify high-risk individuals with a potential for developing HCC. This approach may offer an early diagnosis for HCC. The second need is to develop a molecular fingerprint that would help to distinguish HCC patients with a potential for developing metastasis or a recurrence after curative surgery. The high mortality of HCC patients is mainly because of the occurrence of metastasis and recurrence. Distinguishing HCC with metastasis potential may offer an opportunity for patients with a poor prognosis to adjuvant therapy. These studies also have the potential to identify novel diagnostic markers and candidate therapeutic targets for effective therapy. The ultimate goal of these efforts will be to decrease HCC aggressiveness and increase patient survival.

Acknowledgements

We thank Ana I. Robles for comments, and Dorothea Dudek and the CCR Fellows Editorial Board for editorial assistance.

References

1. Ferlay, J., Bray, F., Pisani, P. and Parkin, D.M. (2001) *GLOBOCAN 2000: Cancer Incidence, Mortality and Prevalence Worldwide*. IARC Scientific Publications, IARC Press, Lyon.
2. Parkin, D.M., Pisani, P. and Ferlay, J. (1999) Global cancer statistics. *CA Cancer J. Clin.*, **49**, 33–64.
3. El-Serag, H.B. and Mason, A.C. (1999) Rising incidence of hepatocellular carcinoma in the United States. *N. Engl. J. Med.*, **340**, 745–750.
4. Okuda, K., Fujimoto, I., Hanai, A. and Urano, Y. (1987) Changing incidence of hepatocellular carcinoma in Japan. *Cancer Res.*, **47**, 4967–4972.
5. Taylor-Robinson, S.D., Foster, G.R., Arora, S., Hargreaves, S. and Thomas, H.C. (1997) Increase in primary liver cancer in the UK, 1979–94. *Lancet*, **350**, 1142–1143.

6. del Olmo, J.A., Serra, M.A., Rodriguez, F., Escudero, A., Gilabert, S. and Rodrigo, J.M. (1998) Incidence and risk factors for hepatocellular carcinoma in 967 patients with cirrhosis. *J. Cancer Res. Clin. Oncol.*, **124**, 560–564.
7. Tiribelli, C., Melato, M., Croce, L.S., Giarelli, L., Okuda, K. and Ohnishi, K. (1989) Prevalence of hepatocellular carcinoma and relation to cirrhosis: comparison of two different cities of the world—Trieste, Italy, and Chiba, Japan. *Hepatology*, **10**, 998–1002.
8. Hsu, I.C., Tokiwa, T., Bennett, W., Metcalf, R.A., Welsh, J.A., Sun, T. and Harris, C.C. (1993) p53 gene mutation and integrated hepatitis B viral DNA sequences in human liver cancer cell lines. *Carcinogenesis*, **14**, 987–992.
9. Murakami, Y., Hayashi, K., Hirohashi, S. and Sekiya, T. (1991) Aberrations of the tumor suppressor p53 and retinoblastoma genes in human hepatocellular carcinomas. *Cancer Res.*, **51**, 5520–5525.
10. Nhieu, J.T., Renard, C.A., Wei, Y., Cherqui, D., Zafrani, E.S. and Buendia, M.A. (1999) Nuclear accumulation of mutated beta-catenin in hepatocellular carcinoma is associated with increased cell proliferation. *Am. J. Pathol.*, **155**, 703–710.
11. Peng, S.Y., Lai, P.L. and Hsu, H.C. (1993) Amplification of the c-myc gene in human hepatocellular carcinoma: biologic significance. *J. Formos. Med. Assoc.*, **92**, 866–870.
12. Nishida, N., Fukuda, Y., Komeda, T. *et al.* (1994) Amplification and overexpression of the cyclin D1 gene in aggressive human hepatocellular carcinoma. *Cancer Res.*, **54**, 3107–3110.
13. Kuroki, T., Fujiwara, Y., Tsuchiya, E., Nakamori, S., Imaoka, S., Kanematsu, T. and Nakamura, Y. (1995) Accumulation of genetic changes during development and progression of hepatocellular carcinoma: loss of heterozygosity of chromosome arm 1p occurs at an early stage of hepatocarcinogenesis. *Genes Chromosomes Cancer*, **13**, 163–167.
14. Konishi, M., Kikuchi-Yanoshita, R., Tanaka, K., Sato, C., Tsuruta, K., Maeda, Y., Koike, M., Tanaka, S., Nakamura, Y. and Hattori, N. (1993) Genetic changes and histopathological grades in human hepatocellular carcinomas. *Jpn. J. Cancer Res.*, **84**, 893–899.
15. De Souza, A.T., Hankins, G.R., Washington, M.K., Fine, R.L., Orton, T.C. and Jirtle, R.L. (1995) Frequent loss of heterozygosity on 6q at the mannose 6-phosphate/insulin-like growth factor II receptor locus in human hepatocellular tumors. *Oncogene*, **10**, 1725–1729.
16. Nishida, N., Fukuda, Y., Kokuryu, H., Sadamoto, T., Isowa, G., Honda, K., Yamaoka, Y., Ikenaga, M., Imura, H. and Ishizaki, K. (1992) Accumulation of allelic loss on arms of chromosomes 13q, 16q and 17p in the advanced stages of human hepatocellular carcinoma. *Int. J. Cancer*, **51**, 862–868.
17. Hsia, C.C., Axiotis, C.A., Di Bisceglie, A.M. and Tabor, E. (1992) Transforming growth factor-alpha in human hepatocellular carcinoma and coexpression with hepatitis B surface antigen in adjacent liver. *Cancer*, **70**, 1049–1056.
18. Ito, N., Kawata, S., Tamura, S., Takaishi, K., Shirai, Y., Kiso, S., Yabuuchi, I., Matsuda, Y., Nishioka, M. and Tarui, S. (1991) Elevated levels of transforming growth factor beta messenger RNA and its polypeptide in human hepatocellular carcinoma. *Cancer Res.*, **51**, 4080–4083.
19. Thorgeirsson, S.S. and Grisham, J.W. (2002) Molecular pathogenesis of human hepatocellular carcinoma. *Nature Genet.*, **31**, 339–346.
20. Liang, P. and Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, **257**, 967–971.
21. Scuric, Z., Stain, S.C., Anderson, W.F. and Hwang, J.J. (1998) New member of aldose reductase family proteins overexpressed in human hepatocellular carcinoma. *Hepatology*, **27**, 943–950.
22. Kim, M.Y., Park, E., Park, J.H., Park, D.H., Moon, W.S., Cho, B.H., Shin, H.S. and Kim, D.G. (2001) Expression profile of nine novel genes differentially expressed in hepatitis B virus-associated hepatocellular carcinomas. *Oncogene*, **20**, 4568–4575.
23. Kondoh, N., Shuda, M., Tanaka, K., Wakatsuki, T., Hada, A. and Yamamoto, M. (2001) Enhanced expression of S8, L12, L23a, L27 and L30 ribosomal protein mRNAs in human hepatocellular carcinoma. *Anticancer Res.*, **21**, 2429–2433.
24. Kanetaka, K., Sakamoto, M., Yamamoto, Y., Yamasaki, S., Lanza, F., Kanematsu, T. and Hirohashi, S. (2001) Overexpression of tetraspanin CO-029 in hepatocellular carcinoma. *J. Hepatol.*, **35**, 637–642.
25. Diatchenko, L., Lau, Y.F., Campbell, A.P. *et al.* (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl Acad. Sci. USA*, **93**, 6025–6030.
26. Miyasaka, Y., Enomoto, N., Nagayama, K., Izumi, N., Marumo, F., Watanabe, M. and Sato, C. (2001) Analysis of differentially expressed genes in human hepatocellular carcinoma using suppression subtractive hybridization. *Br. J. Cancer*, **85**, 228–234.
27. Lisitsyn, N., Lisitsyn, N. and Wigler, M. (1993) Cloning the differences between two complex genomes. *Science*, **259**, 946–951.
28. Graveel, C.R., Jatke, T., Madore, S.J., Holt, A.L. and Farnham, P.J. (2001) Expression profiling and identification of novel genes in hepatocellular carcinomas. *Oncogene*, **20**, 2704–2712.
29. Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K.W. (1995) Serial analysis of gene expression. *Science*, **270**, 484–487.
30. Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467–470.
31. Lipshutz, R.J., Fodor, S.P., Gingeras, T.R. and Lockhart, D.J. (1999) High density synthetic oligonucleotide arrays. *Nature Genet.*, **21**, 20–24.
32. Honda, M., Kaneko, S., Kawai, H., Shirota, Y. and Kobayashi, K. (2001) Differential gene expression between chronic hepatitis B and C hepatic lesion. *Gastroenterology*, **120**, 955–966.
33. Yamashita, T., Kaneko, S., Hashimoto, S., Sato, T., Nagai, S., Toyoda, N., Suzuki, T., Kobayashi, K. and Matsushima, K. (2001) Serial analysis of gene expression in chronic hepatitis C and hepatocellular carcinoma. *Biochem. Biophys. Res. Commun.*, **282**, 647–654.
34. Wu, C.G., Forgues, M., Siddique, S., Farnsworth, J., Valerie, K. and Wang, X.W. (2002) SAGE transcript profiles of normal primary human hepatocytes expressing oncogenic hepatitis B virus X protein. *FASEB J.*, **16**, 1665–1667.
35. Wu, C.G., Salvay, D.M., Forgues, M., Valerie, K., Farnsworth, J., Markin, R.S. and Wang, X.W. (2001) Distinctive gene expression profiles associated with hepatitis B virus x protein. *Oncogene*, **20**, 3674–3682.
36. Han, J., Yoo, H.Y., Choi, B.H. and Rho, H.M. (2000) Selective transcriptional regulations in the human liver cell by hepatitis B viral X protein. *Biochem. Biophys. Res. Commun.*, **272**, 525–530.
37. Goldenberg, D., Ayesh, S., Schneider, T. *et al.* (2002) Analysis of differentially expressed genes in hepatocellular carcinoma using cDNA arrays. *Mol. Carcinogen.*, **33**, 113–124.
38. Delpuech, O., Trabut, J.B., Carnot, F., Feuillard, J., Brechet, C. and Kremsdorf, D. (2002) Identification, using cDNA macroarray analysis, of distinct gene expression profiles associated with pathological and virological features of hepatocellular carcinoma. *Oncogene*, **21**, 2926–2937.
39. Harada, T., Arai, S., Mise, M. *et al.* (1998) Membrane-type matrix metalloproteinase-1 (MT1-MMP) gene is overexpressed in highly invasive hepatocellular carcinomas. *J. Hepatol.*, **28**, 231–239.
40. Le Bail, B., Faouzi, S., Boussarie, L., Guirouilh, J., Blanc, J.F., Carles, J., Bioulac-Sage, P., Balabaud, C. and Rosenbaum, J. (1999) Osteonectin/SPARC is overexpressed in human hepatocellular carcinoma. *J. Pathol.*, **189**, 46–52.
41. Chen, X., Cheung, S.T., So, S. *et al.* (2002) Gene expression patterns in human liver cancers. *Mol. Biol. Cell*, **13**, 1929–1939.
42. Delarue, F.L., Taylor, B.S. and Sebt, S.M. (2001) Ras and RhoA suppress whereas RhoB enhances cytokine-induced transcription of nitric oxide synthase-2 in human normal liver AKN-1 cells and lung cancer A-549 cells. *Oncogene*, **20**, 6531–6537.
43. Hanafusa, T., Yumoto, Y., Nouse, K. *et al.* (2002) Reduced expression of insulin-like growth factor binding protein-3 and its promoter hypermethylation in human hepatocellular carcinoma. *Cancer Lett.*, **176**, 149–158.
44. Shirota, Y., Kaneko, S., Honda, M., Kawai, H.F. and Kobayashi, K. (2001) Identification of differentially expressed genes in hepatocellular carcinoma with cDNA microarrays. *Hepatology*, **33**, 832–840.
45. Paterson, A.C., Sciot, R., Kew, M.C., Callea, F., Dusheiko, G.M. and Desmet, V.J. (1988) HLA expression in human hepatocellular carcinoma. *Br. J. Cancer*, **57**, 369–373.
46. Yang, P.M., Sung, J.L. and Chen, D.S. (1989) HLA-A, B, C and DR antigens in chronic hepatitis B viral infection. *Hepatogastroenterology*, **36**, 363–366.
47. Dick, T., Ray, K., Salz, H.K. and Chia, W. (1996) Cytoplasmic dynein (ddlc1) mutations cause morphogenetic defects and apoptotic cell death in *Drosophila melanogaster*. *Mol. Cell. Biol.*, **16**, 1966–1977.
48. Crepieux, P., Kwon, H., Leclerc, N., Spencer, W., Richard, S., Lin, R. and Hiscott, J. (1997) I kappaB alpha physically interacts with a cytoskeleton-associated protein through its signal response domain. *Mol. Cell. Biol.*, **17**, 7375–7385.
49. Stander, M., Naumann, U., Wick, W. and Weller, M. (1999) Transforming growth factor-beta and p-21: multiple molecular targets of decorin-mediated suppression of neoplastic growth. *Cell Tissue Res.*, **296**, 221–227.
50. Santra, M., Skorski, T., Calabretta, B., Lattime, E.C. and Iozzo, R.V. (1995) *De novo* decorin gene expression suppresses the malignant phenotype in human colon cancer cells. *Proc. Natl Acad. Sci. USA*, **92**, 7016–7020.
51. Yuan, B.Z., Miller, M.J., Keck, C.L., Zimonjic, D.B., Thorgeirsson, S.S. and Popescu, N.C. (1998) Cloning, characterization, and chromosomal

- localization of a gene frequently deleted in human liver cancer (DLC-1) homologous to rat RhoGAP. *Cancer Res.*, **58**, 2196–2199.
52. Ng,I.O., Liang,Z.D., Cao,L. and Lee,T.K. (2000) DLC-1 is deleted in primary hepatocellular carcinoma and exerts inhibitory effects on the proliferation of hepatoma cell lines with deleted DLC-1. *Cancer Res.*, **60**, 6581–6584.
 53. Okabe,H., Satoh,S., Kato,T., Kitahara,O., Yanagawa,R., Yamaoka,Y., Tsunoda,T., Furukawa,Y. and Nakamura,Y. (2001) 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.*, **61**, 2129–2137.
 54. Simon,R., Radmacher,M.D. and Dobbin,K. (2002) Design of studies using DNA microarrays. *Genet. Epidemiol.*, **23**, 21–36.
 55. Emmert-Buck,M.R., Bonner,R.F., Smith,P.D., Chuaqui,R.F., Zhuang,Z., Goldstein,S.R., Weiss,R.A. and Liotta,L.A. (1996) Laser capture microdissection. *Science*, **274**, 998–1001.
 56. Miura,K., Bowman,E.D., Simon,R. *et al.* (2002) Laser capture microdissection and microarray expression analysis of lung adenocarcinoma reveals tobacco smoking- and prognosis-related molecular profiles. *Cancer Res.*, **62**, 3244–3250.
 57. Luo,L., Salunga,R.C., Guo,H. *et al.* (1999) Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nature Med.*, **5**, 117–122.
 58. Sgroi,D.C., Teng,S., Robinson,G., LeVangie,R., Hudson,J.R. Jr and Elkahoun,A.G. (1999) *In vivo* gene expression profile analysis of human breast cancer progression. *Cancer Res.*, **59**, 5656–5661.
 59. Claverie,J.M. (1999) Computational methods for the identification of differential and coordinated gene expression. *Hum. Mol. Genet.*, **8**, 1821–1832.
 60. Eisen,M.B., Spellman,P.T., Brown,P.O. and Botstein,D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA*, **95**, 14863–14868.
 61. Hamadeh,H.K., Bushel,P.R., Jayadev,S. *et al.* (2002) Prediction of compound signature using high density gene expression profiling. *Toxicol. Sci.*, **67**, 232–240.
 62. Brown,M.P., Grundy,W.N., Lin,D., Cristianini,N., Sugnet,C.W., Furey,T.S., Ares,M. Jr and Haussler,D. (2000) Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc. Natl Acad. Sci. USA*, **97**, 262–267.
 63. Khan,J., Wei,J.S., Ringner,M. *et al.* (2001) Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nature Med.*, **7**, 673–679.
 64. Yeang,C.H., Ramaswamy,S., Tamayo,P., Mukherjee,S., Rifkin,R.M., Angelo,M., Reich,M., Lander,E., Mesirov,J. and Golub,T. (2001) Molecular classification of multiple tumor types. *Bioinformatics*, **17** (suppl. 1), S316–S322.
 65. Ye,Q., Qin,L., Forgues,M. *et al.* (2002) cDNA microarray approach to analyze metastasis-associated genes in human hepatocellular carcinoma. *Proc. Am. Assoc. Cancer Res.*, **43**, 902.
 66. Radmacher,M.D., McShane,L.M. and Simon,R. (2002) A paradigm for class prediction using gene expression profiles. *J. Comput. Biol.*, **9**, 505–511.
 67. Golub,T.R., Slonim,D.K., Tamayo,P. *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 531–537.
 68. Kim,J.W., Sime,J., Forgues,M., He,P., Ye,Q., Kaul,R. and Wang,X.W. (2002) Molecular characterization of preneoplastic liver diseases by cDNA microarray. *Proc. Am. Assoc. Cancer Res.*, **43**, 461–462.
 69. Huang,F., Adelman,J., Jiang,H., Goldstein,N.I. and Fisher,P.B. (1999) Identification and temporal expression pattern of genes modulated during irreversible growth arrest and terminal differentiation in human melanoma cells. *Oncogene*, **18**, 3546–3552.
 70. Xiao,S.Y., Wang,H.L., Hart,J., Fleming,D. and Beard,M.R. (2001) cDNA arrays and immunohistochemistry identification of CD10/CALLA expression in hepatocellular carcinoma. *Am. J. Pathol.*, **159**, 1415–1421.
 71. Xu,L., Hui,L., Wang,S., Gong,J., Jin,Y., Wang,Y., Ji,Y., Wu,X., Han,Z. and Hu,G. (2001) Expression profiling suggested a regulatory role of liver-enriched transcription factors in human hepatocellular carcinoma. *Cancer Res.*, **61**, 3176–3181.
 72. Xu,X.R., Huang,J., Xu,Z.G. *et al.* (2001) Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver. *Proc. Natl Acad. Sci. USA*, **98**, 15089–15094.
 73. Lau,W.Y., Lai,P.B., Leung,M.F., Leung,B.C., Wong,N., Chen,G., Leung,T.W. and Liew,C.T. (2000) Differential gene expression of hepatocellular carcinoma using cDNA microarray analysis. *Oncol. Res.*, **12**, 59–69.
 74. Takeo,S., Arai,H., Kusano,N. *et al.* (2001) Examination of oncogene amplification by genomic DNA microarray in hepatocellular carcinomas: comparison with comparative genomic hybridization analysis. *Cancer Genet. Cytogenet.*, **130**, 127–132.
 75. Lee,J.S. and Thorgeirsson,S.S. (2002) Functional and genomic implications of global gene expression profiles in cell lines from human hepatocellular cancer. *Hepatology*, **35**, 1134–1143.
 76. Kawai,H.F., Kaneko,S., Honda,M., Shirota,Y. and Kobayashi,K. (2001) Alpha-fetoprotein-producing hepatoma cell lines share common expression profiles of genes in various categories demonstrated by cDNA microarray analysis. *Hepatology*, **33**, 676–691.
 77. Tackels-Horne,D., Goodman,M.D., Williams,A.J., Wilson,D.J., Eskandari,T., Vogt,L.M., Boland,J.F., Scherf,U. and Vockley,J.G. (2001) Identification of differentially expressed genes in hepatocellular carcinoma and metastatic liver tumors by oligonucleotide expression profiling. *Cancer*, **92**, 395–405.
 78. Li,K., Prow,T., Lemon,S.M. and Beard,M.R. (2002) Cellular response to conditional expression of hepatitis C virus core protein in Huh7 cultured human hepatoma cells. *Hepatology*, **35**, 1237–1246.
 79. Midorikawa,Y., Tsutsumi,S., Taniguchi,H., Ishii,M., Kobune,Y., Kodama,T., Makuuchi,M. and Aburatani,H. (2002) Identification of genes associated with dedifferentiation of hepatocellular carcinoma with expression profiling analysis. *Jpn. J. Cancer Res.*, **93**, 636–643.
 80. Tanaka,F., Hori,N. and Sato,K. (2002) Identification of differentially expressed genes in rat hepatoma cell lines using subtraction and microarray. *J. Biochem. (Tokyo)*, **131**, 39–44.
 81. Bigger,C.B., Brasky,K.M. and Lanford,R.E. (2001) DNA microarray analysis of chimpanzee liver during acute resolving hepatitis C virus infection. *J. Virol.*, **75**, 7059–7066.

Received November 5, 2002; accepted December 5, 2002