Gene expression studies of hepatitis virus-induced woodchuck hepatocellular carcinoma in correlation with human results

FANGJING WANG[1], PAUL W. ANDERSON[1], NICOLAS SALEM[1], YU KUANG[1], BUD C. TENNANT[2] and ZHENGHONG LEE[1,3]

[1]Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44106; [2]Department of Clinical Sciences, Cornell University, Ithaca, NY 14853; [3]Department of Radiology, University Hospitals of Cleveland, Cleveland, OH 44106, USA

Received July 19, 2006; Accepted September 20, 2006

Abstract. The lack of good molecular markers for diagnosis as well as treatment assessment has rendered the hepatocellular carcinoma (HCC) a major challenge in health care. In this study, woodchucks were used as an animal model for hepatitis virus-induced HCC, and gene expression studies were performed using a human oligonucleotide microarray. An analysis approach combing supervised significant analysis of microarray (SAM), prediction analysis of microarray (PAM), and unsupervised hierarchical cluster methodologies statistically determined 211 upregulated and 78 downregulated genes between liver cancer and non-cancer liver tissues, and demonstrated ≥93% accuracy in classifying the tissue samples. RT-PCR results confirmed the differential expression of selected sequenced woodchuck genes (SAT, IDH3B, SCD) in the microarray. Our study showed that differentially expressed genes were involved in transcription, RNA splicing, translation, cell cycle, metabolism, protein folding and degradation, apoptosis, immune response, metal binding, etc. Interestingly, some genes were involved with signaling pathways such as Ras/MAPK (MAPKAP1), Src-dependent pathways (CSK), hedgehog signaling pathway (HHIP), while Wnt signaling pathway may not be dominant in woodchuck HCC as shown by the downregulation of ß-catenin (TNNB1) and the upregulation of CXXC4 and CSNK2B. Numerous genes found in this study were also differentially expressed in human HCC and many other human cancers including breast, prostate and lung cancers, etc., serving as tumor suppressors, promoters, prognostic markers or chemotherapy targets. In conclusion, this study has demonstrated the robustness of the data analysis and the potential of using human microarrays on woodchuck samples. In particular, some of the differentially expressed genes in the woodchuck HCC can be further explored for possible molecular imaging targets or biological markers in human HCC.

Correspondence to: Dr Zhenghong Lee, Department of Radiology, University Hospitals of Cleveland, 11100 Euclid Avenue, Cleveland, OH 44106, USA
E-mail: zxl11@case.edu

Key words: cross-species hybridization, microarray, woodchuck, hepatitis B virus, hepatocellular carcinoma, SAM

Introduction

Primary hepatocellular carcinoma (HCC) is the fifth most common malignancy in the world and the third most common cause of cancer-related death (1). It is most prevalent in some of the Asian and African countries, but the incidence of primary liver carcinoma in Western countries is on the rise in the past decades. HCC usually results from the chronic inflammation caused by hepatitis B virus (HBV), hepatitis C virus (HCV), or long-term exposure to alcohol or dietary aflatoxin B1. Despite the progress made in clinical studies, the molecular pathogenesis of HCC is still not well understood (2). Moreover, the overall 5-year survival rate is only about 3-25%, due to late diagnosis, high recurrence rates (84% within 3 years after hepatectomy), and metastasis (3). The lack of molecular markers to characterize tumorigenesis and progression poses a major obstacle for effective diagnosis, prognosis and treatment of HCC.

Patients chronically infected with HBV are among those with the highest risk for developing HCC. The studies of hepatitis B induced carcinoma have been mostly focused on human beings. The Eastern American woodchuck (*Marmota monax*) can develop hepatocellular carcinoma with a high prevalence within the first 2-4 years of life when infected with woodchuck hepatitis virus (WHV) at birth (4,5). WHV virus is similar to human HBV in both structure and replicative cycles. In addition, the WHV-induced liver carcinoma has similar pathology and natural history to human HCC (4,5). The woodchuck has proved to be a valid animal model for therapy of hepatitis B infection and human hepatoma, as well as for development of new imaging agents for enhanced detection of hepatic neoplasm by ultrasound and MRI. It also has great potential in the investigation of the viral and molecular mechanisms responsible for hepatocarcinogenesis, in the studies of molecular imaging as well as in the development of cancer treatments (4,5).

Microarrays have been extensively used on human HCC to examine the expression levels of thousands of genes simultaneously. This technique provides new insights into the molecular mechanisms of biological processes underlying

the development and progression of HCC, can efficiently and accurately detect multiple foci and screen for potential diagnosis and therapeutic markers. However, up to date, only a small fraction of woodchuck genes have been sequenced and no commercial microarray is available to study gene expressions in woodchucks. A cross-species hybridization analysis is one way to use on odd species. Some groups succeeded in using human Affymetrix GeneChips® on porcine samples (6,7). In particularly, another group demonstrated the feasibility of using commercially available human nylon membrane arrays on normal woodchuck liver samples (8).

Recently, our lab demonstrated the feasibility of the use of commonly available human oligonucleotide Affymetrix GeneChips® on woodchuck HCCs versus surrounding liver samples (9). In that study, gene expression profiles of the paired woodchuck HCCs and surrounding hepatic tissues were analyzed by GeneChip Operating Software from Affymetrix (GCOS). However, limited data sets, requirement for paired samples by GCOS, as well as the lack of a statistic analysis method for different pairs were some of the limitations. As a continuing effort, more samples were acquired and used for microarray experiments in the present study. Based on GCOS, supervised and unsupervised analysis methods, including significant analysis of microarray (SAM), prediction of analysis of microarray and hierarchical cluster, were further applied to statistically determine the differentially expressed genes and examine the accuracy of data analysis. By analyzing the differentially expressed genes in the woodchuck samples obtained from human microarray and RT-PCR, this work provided some insight into gene expression profiles and identified genes that could be potential targets for molecular imaging.

Materials and methods

RNA isolation. Woodchucks (acquired from Cornell University) were euthanized after PET imaging (reported separately). Liver tissues were immediately removed from the animals, snap frozen in liquid nitrogen and stored at -80˚C. RNA was extracted with RNeasy Midi Kit from Qiagen (Valencia, CA) according to the recommended protocol. Prior to microarray experiment, the RNA integrity number (RIN) was measured using a Bioanalyzer 2100 from Agilent (Palo Alto, CA).

Preparation of cRNA and microarray hybridization. Total RNA was given to the Gene Expression Array Core Facility at Case Western Reserve University. cRNA was prepared and hybridized to Affymetrix Human U133 plus 2.0 Gene Chips (Santa Clara, CA) according to the manufacturer's instructions.

Microarray data analysis. Two sample t-tests of quality control data were performed using Microsoft Excel. The hybridization results were first scaled to a target value of 15 using GCOS 1.2 from Affymetrix, and the differential gene expression of the HCCs versus the surrounding hepatic tissues was analyzed for each sample pair. A gene was considered upregulated in the HCCs if it was labeled as 'present' or 'marginal' in the HCC and if the fold change was 'increased' or 'marginally increased' according to GCOS. A gene was considered downregulated if it was labeled as 'present' or 'marginal' in the surrounding liver tissue, and if the change was 'decreased' or 'marginally decreased' in the HCCs. The call status of a gene in any of the two healthy livers could be used on the tumor-free tissue samples in order to favorably include more differentially expressed genes. Microsoft Excel and Access were used to extract the significant results. The genes that were differentially expressed in less than 2 sample pairs (out of 6 pairs) were not considered. The remaining genes were further analyzed by SAM (developed by the Stanford University). The data from SAM were used for PAM analysis and hierarchical cluster. The miscalculation errors as well as the cross-validated probability of the samples were computed using PAM by leave-out cross validation method.

RT-PCR. Available gene sequence information for woodchucks was found in the Nucleotide Database in the National Center for Biotechnology Information website (www.ncbi.nlm.nih.gov). The optimal TaqMan® primers and probe sequences for our genes of interest were designed and produced by Applied Biosystems (Foster City, CA) as listed in Table I. The primers, probes and the total RNA were given to the Gene Expression Array Core Facility. A 384-well plate set up was used on a PRISM® 7900HT Sequence Detection System from Applied Biosystems to perform RT-PCR.

Results

Quality control. As shown in Table II, the integrity number of all the extracted RNA samples was greater than 7.5, indicating that these RNA samples can be used for microarray experiments. It was found by GCOS that approximately 10% of human genes were present or marginally present in the woodchuck liver samples. The percentage of genes called 'present' or 'marginal' in the woodchuck tumor tissues was not significantly differently from that in non-tumor woodchuck liver tissues (P=0.06).

SAM analysis of differentially expressed genes. GCOS identified 615 gene probe sets that may distinguish HCCs from non-tumor liver tissue specimens. These genes were imported into SAM to statistically determine the significantly upregulated or downregulated genes. Table III shows the parameters used for SAM analysis. As the δ-value increased, fewer significant genes were found with a lower false discovery rate (FDR), indicating that fewer genes were likely to be falsely identified. In this study, we chose a δ-value of 1.0 and a fold change of 1.5 as a criterion, leading to 265 upregulated and 79 downregulated gene probes, which corresponds to 211 upregulated and 78 downregulated genes, respectively, in the woodchuck tumor tissues. As a result, false discovery analysis indicated a false-positive error of 0.18%, which was reasonably low.

Hierarchical cluster, tree view and PAM analysis. Based on the genes we obtained from SAM, hierarchical cluster method was used to classify the samples. As shown in Fig. 1, most of the tumor and non-tumor samples could be correctly identified with the only exception that W6380N (a surrounding hepatic sample) was falsely classified as a tumor tissue. We further

Table I. RT-PCR primers and probe sequences for selected genes.

Target gene	Forward primer sequence	Reverse primer sequence	Reporter probe sequence
Alpha-fetoprotein (AFP)	AGGCTGTCATTGCAGATTTCTCT	GGACCCTCTTCTGCAAAGCA	CTGGCAGCATGTCTCC
Hexokinase (glucokinase regulatory protein) (GCKR)	ATGCTGCAGCGGTTCTCT	GGATCGCTTGGAGGAGACTCT	AAGGCCCGATGCATTG
Isocitrate dehydrogenase (NAD+) beta (IDH3B)	TCTCAGCGGATTGCAAAGTTTG	CTTGTGGACAGCTGTGACCTT	CCCGCCCCTTCTTG
Fatty-acid-Coenzyme A ligase, very long-chain 1 (FACVL1)	GCGGGATGACACAGCAAAA	TCTTTCTGGAAATGTGAGTTATCCTTCTG	TCTGAAAATCTGAATATCCC
Spermidine/spermine N1-acetyltransferase (SAT)	TTTCATGCAACACTTGGTCTCTCT	CACTGGACTCCGGAAGGTAAC	CCCTCACCCAATCCAG
Carnitine palmitoyltransferase II (CPT2)	TGCCTATTCCCAAACTTGAAGACA	TCTTCCTGAACTGGCCATCATTC	TCAATGCACAGAAACCT
ELOVL family member 5, elongation of long chain fatty acids (ELOVL5)	TGCTGTCTCTCTACATGTTCTGTGA	GTGTGCCCTGACAAAAGAAGTTG	CCTTGCCACACTCCTG
Alcohol dehydrogenase 1 (class I, ADH1)	CCACCTGTGCTGTGTTTGG	CTGCTTTACAGCCCATGATAACAGA	CAGGCCAACTCCTCC
Stearoyl-CoA desaturase (delta-9-desaturase) (SCD)	TGCCCAGATGGCTTTAGAAAGG	GGAAGCTACAAGAGTGCCTAAGTTT	AACCTGGACAACCCC
3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (HMGCS2)	GGCCTTGGATCGATGTTACACA	CCAGCTTGCTTCCACTGTTTT	CATACCGCCAGAAAAT

Table II. Quality control data for RNA and microarray samples.

Sample[a]	RNA integrity no. (max = 10.0)	Average background	Noise	Present genes (%)	Marginally present genes (%)	Scaling factor
W5904N	7.5	33.8	1.00	9.0	1.3	2.96
W6361N	7.8	34.2	0.97	9.5	1.3	2.97
W6380N	9.7	32.1	0.93	9.9	1.2	2.96
W6503N	9.2	40.5	1.49	10.5	1.4	1.60
W5861N	8.8	42.9	1.53	10.2	1.4	1.65
W6337N	8.8	41.9	1.50	10.6	1.4	1.66
W6384N	9.2	39.5	1.40	10.0	1.3	1.85
W7034N	8.2	43.1	1.55	10.3	1.4	1.68
W5904T	9.8	36.9	1.07	10.3	1.3	2.21
W6361T	9.2	35.7	1.05	10.5	1.5	2.36
W6380T	9.9	31.6	0.94	10.8	1.3	2.46
W6337T	9.8	43.2	1.50	10.7	1.4	1.53
W6384T	9.7	42.3	1.52	10.0	1.2	1.73
W7034T	9.7	41.3	1.49	10.5	1.4	1.65

[a]N refers to non-tumor tissues; T refers to tumor tissues.

Table III. Significant gene analysis by SAM.

Parameter	Nos. called significant	Up	Down	FDR (%)
0.5	362	265	97	0.56
1.0	344	265	79	0.18
1.48	321	265	56	0.11
2.06	241	223	18	0.04
2.13	192	177	15	0

analyzed the gene expression data using PAM. As shown in Fig. 2, the miscalculation error was almost the same until the shrinkage parameter reached 3.5. At shrinkage values ≥3.5, less than 5 genes were positively identified and a drastically increased miscalculation error followed. When a shrinkage value of 2.0 was used, the resulted cross-out validation probability by PAM was 1, showing that W6380N is more like a tumor tissue. The use of 2.9 as a shrinkage value still led to a probability of 78%, but at this point, the healthy control samples (6503N and 5861N) as well as some tumor samples began to lose their clear identity. The fact that W6380N was not classified as a non-tumor tissue agrees with our histology and enzyme analysis. H&E staining demonstrated clear tumor

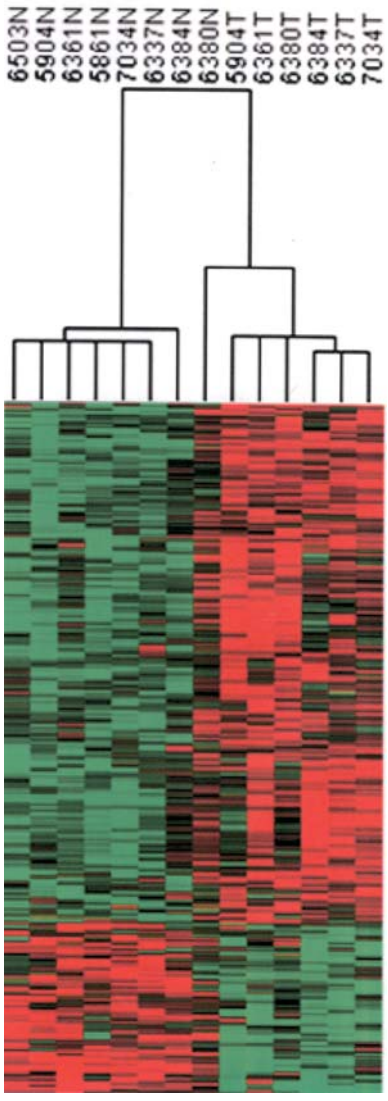

Figure 1. Hierarchical cluster of the 14 samples using all the genes obtained from the SAM analysis. Red, green, black and grey color represents up-regulation, downregulation, no change and no expression, respectively.

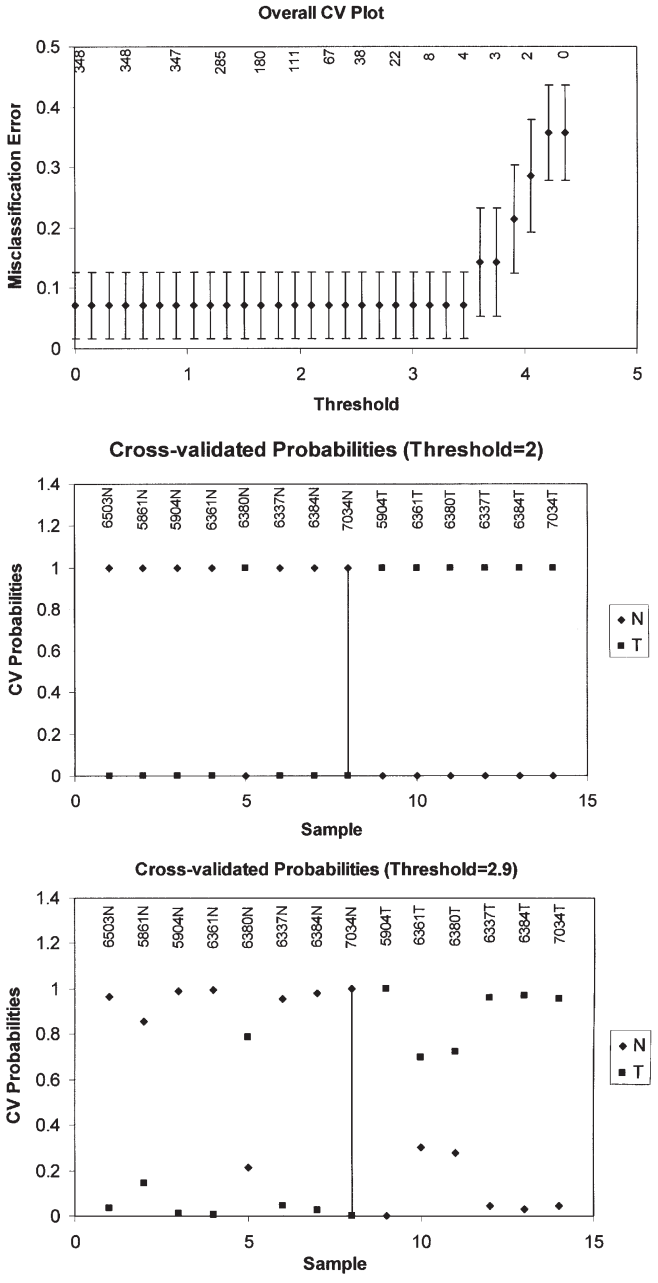

Figure 2. PAM prediction analysis of gene expression. Top, miscalculation errors; middle, cross-validated probabilities at threshold = 2.0; bottom, cross-validated probabilities at threshold = 2.9.

infiltration into the normal tissue, and the hexokinase activity for W6380N was significantly different from that of other non-tumor tissue (data not shown). These results indicate that an accuracy of more than 93% was achieved by using the genes in this study.

Upregulated and downregulated genes. Table IV shows some upregulated and downregulated genes. It is apparent that many classes of genes have been differentially expressed in woodchuck HCCs, including genes involved with transcription (e.g., ZNF6, EGR1, GTF2H4, CSDA), RNA processing (e.g., SYNCRIP, SIPI), DNA repair (e.g., H2AFZ, APEX1, H2AFY) and lipid metabolism (e.g., SCD, ELOVL6, DGAT2, CYP51A). Interestingly, a larger number of ribosomal proteins but RPL35, which constitute the ribosome organelles that catalyze protein synthesis, were upregulated (e.g., RPS18, RPL3). The 26 ribosomal proteins accounted for 59% in the translation-associated genes and 9% in the total number of genes that were differentially expressed in this study. Cell cycle regulators such as CDC2, CDK8, CDK4, CKS2, as well as CHES1, a checkpoint suppressor 1 gene, were found upregulated, while IGFBP2 (insulin-like growth factor binding protein 2) was significantly downregulated (4.8-fold). Apoptosis induction genes such as BCLAF1 (BCL2-associated transcription factor 1), PDCD4 (programmed death factor 4) were also found upregulated in woodchuck HCCs. Some genes were involved in the ubiquitin-proteasome pathway (e.g., PSMA7, PSMD10), and all these genes were upregulated in HCCs. It was also noted that the liver specific metallothionein genes (e.g., MT2A, MT1F, MT1G, MT1H, MT1X) that detoxify heavy metals were downregulated in the woodchuck HCCs, in correlation with human HCC (10,11) and prostate cancer (12,13). All these genes are located in 16q13 position in the human chromosomes as shown in Table IV.

Table IV. Selected genes that are differentially expressed.

Function	Gene symbol	Gene title	UniGene ID	Chromosomal location	Fold change[a]
Transcription factors	ZNF6	Zinc finger protein 6 (CMPX1)	Hs.326801	Xq21.1-q21.2	5.25
	HMGB3	High-mobility group box 3	Hs.19114	Xq28	4.13
	CSDA	Cold shock domain protein A	Hs.558424	12p13.1	2.76
	GTF2H4	General transcription factor IIH, polypeptide 4, 52 kDa	Hs.485070	6p21.3	2.04
	BTF3	Basic transcription factor 3	Hs.529798	5q13.2	1.96
	YY1	YY1 transcription factor	Hs.388927	14q	1.91
	SMARCA2	**SWI/SNF related, matrix associated, actin dependent regulator of chromatin a2**	**Hs.298990**	**9p22.3**	**0.54**
	BCL6	**B-cell CLL/lymphoma 6 (zinc finger protein 51)**	**Hs.478588**	**3q27**	**0.47**
	ETV6	**Ets variant gene 6 (TEL oncogene)**	**Hs.504765**	**12p13**	**0.46**
	ETS1	**v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)**	**Hs.369438**	**11q23.3**	**0.45**
	GATA3	**GATA binding protein 3**	**Hs.524134**	**10p15**	**0.37**
	EGR1	**Early growth response 1**	**Hs.326035**	**5q31.1**	**0.22**
RNA processing	SRrp35	Serine-arginine repressor protein (35 kDa)	Hs.254414	6q15	4.08
	SYNCRIP	Synaptotagmin binding, cytoplasmic RNA interacting protein	Hs.485877	6q14-q15	3.15
	SIP1	Survival of motor neuron protein interacting protein 1	Hs.533862	14q13	2.68
	SNRPB	Small nuclear ribonucleoprotein polypeptides B and B1	Hs.83753	20p13	2.33
	SFRS1	Splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor)	Hs.68714	17q21.3-q22	1.87
	HNRPH1	Heterogeneous nuclear ribonucleoprotein H1 (H)	Hs.202166	5q35.3	1.80
DNA replication repair	H2AFZ	H2A histone family, member Z	Hs.119192	4q24	3.73
	APEX1	APEX nuclease (multifunctional DNA repair enzyme) 1	Hs.73722	14q11.2-q12	3.23
	POLH	Polymerase (DNA directed), eta	Hs.439153	6p21.1	2.49
	H2AFY	H2A histone family, member Y	Hs.420272	5q31.3-q32	1.90
	SET	SET translocation (myeloid leukemia-associated)	Hs.436687	9q34	1.70
	PARP2	Poly(ADP-ribose) polymerase family, member 2	Hs.409412	14q11.2-q12	1.70
	RBMS1	**RNA binding motif, single stranded interacting protein 1**	**Hs.470412**	**2q24.2**	**0.36**
Translation	RPS18	Ribosomal protein S18	Hs.546290	6p21.3	3.85
	EIF5A	Eukaryotic translation initiation factor 5A	Hs.534314	17p13-p12	3.72
	EEF1G	Eukaryotic translation elongation factor 1 gamma	Hs.144835	11q12.3	2.63
	RPS21	Ribosomal protein S21	Hs.190968	20q13.3	2.59
	RPL3	Ribosomal protein L3	Hs.561637	22q13	2.25
	RPLP1	Ribosomal protein, large, P1	Hs.356502	15q22	2.23
	RPS10	Ribosomal protein S10	Hs.406620	6p21.31	2.21
	RPLP0	Ribosomal protein, large, P0 /// similar to ribosomal protein P0	Hs.448226	12q24.2	2.20
	RPS17	Ribosomal protein S17	Hs.512525	15q	2.06
	RPL10	Ribosomal protein L10	Hs.401929	Xq28	1.93
	RPL23	Ribosomal protein L23	Hs.406300	17q	1.88
	RPS6	Ribosomal protein S6	Hs.408073	9p21	1.86
	RPL7	Ribosomal protein L7	Hs.421257	8q21.11	1.76
	RPS7	Ribosomal protein S7	Hs.546287	2p25	1.71
	RPL36	Ribosomal protein L36	Hs.408018	19p13.3	1.70
	RPL6	Ribosomal protein L6	Hs.528668	12q24.1	1.52
	RPL35	**Similar to 60S ribosomal protein L35**	**Hs.449044**	**9q34.1**	**0.36**
	IMP-2	**IGF-II mRNA-binding protein 2**	**Hs.35354**	**3q27.2**	**0.13**
Cell cycle growth differentiation	CDC2	Cell division cycle 2, G1 to S and G2 to M	Hs.334562	10q21.1	2.91
	PTMA	Prothymosin, alpha (gene sequence 28)	Hs.459927	2q35-q36	2.66
	CDK8	Cyclin-dependent kinase 8	Hs.382306	13q12	2.49
	CKS2	CDC28 protein kinase regulatory subunit 2	Hs.83758	9q22	2.37
	CDK4	Cyclin-dependent kinase 4	Hs.95577	12q14	2.21
	MCC	Mutated in colorectal cancers	Hs.483104	5q21-q22	1.90
	CHES1	Checkpoint suppressor 1	Hs.434286	14q24.3-q32.11	1.61
	IGFBP2	**Insulin-like growth factor binding protein 2, 36 kDa**	**Hs.438102**	**2q33-q34**	**0.21**
Apoptosis	BCLAF1	BCL2-associated transcription factor 1	Hs.486542	6q22-q23	1.99
	PDCD4	Programmed cell death 4 (neoplastic transformation inhibitor)	Hs.232543	10q24	1.72
Immune response	MIF	Macrophage migration inhibitory factor (glycosylation-inhibiting factor)	Hs.407995	22q11.23	1.92
	ILF2	Interleukin enhancer binding factor 2, 45 kDa	Hs.75117	1q21.3	1.70
	IGHA1	**Immunoglobulin heavy constant alpha 1 etc.**	**Hs.510635**	**14q32.33**	**0.55**
	SERPING1	**Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1**	**Hs.384598**	**11q12-q13.1**	**0.51**
	HLA-DRB1	**Major histocompatibility complex, class II, DR beta 1**	**Hs.520049**	**6p21.3**	**0.50**
	ARMCX3	**Armadillo repeat containing, X-linked 3**	**Hs.172788**	**Xq21.33-q22.2**	**0.29**

Table IV. Continued.

Function	Gene symbol	Gene title	UniGene ID	Chromosomal location	Fold change[a]
Cell adhesion/ cytoskeletal organization	COL1A1	Collagen, type I, alpha 1	Hs.172928	17q21.3-q22.1	4.84
	CLDN10	Claudin 10	Hs.534377	13q31-q34	2.52
	COL1A2	Collagen, type I, alpha 2	Hs.489142	7q22.1	2.28
	ADAM10	ADAM metallopeptidase domain 10	Hs.172028	15q2|15q22	2.11
	ACTN1	Actinin, alpha 1	Hs.509765	14q24.1-q24.2|14q22-q24	1.90
	K-ALPHA-1	Tubulin, alpha, ubiquitous	Hs.524390	12q13.12	1.66
	SPTBN1	Spectrin, beta, non-erythrocytic 1	Hs.503178	2p21	1.63
	COL4A2	Collagen, type IV, alpha 2	Hs.508716	13q34	1.53
	PCDHA9	**Protocadherin alpha 9 etc.**	**Hs.199343**	**5q31**	**0.38**
Signal transduction	RAN	RAN, member RAS oncogene family	Hs.10842	12q24.3	3.62
	SKB1	SKB1 homolog (*S. pombe*)	Hs.367854	14q11.2-q21	3.48
	CXXC4	CXXC finger 4	Hs.12248	4q22-q24	3.07
	NPM1	Nucleophosmin (nucleolar phosphoprotein B23, numatrin)	Hs.557550	5q35	2.60
	FKBP1A	FK506 binding protein 1A, 12 kDa	Hs.471933	20p13	2.45
	GNB2L1	Guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1	Hs.5662	5q35.3	2.41
	PPP2R5E	Protein phosphatase 2, regulatory subunit B (B56), epsilon isoform	Hs.334868	14q23.1	2.17
	PTEN	Phosphatase and tensin homolog (mutated in multiple advanced cancers 1)	Hs.500466	10q23.3	1.86
	GCSH	Glycine cleavage system protein H (aminomethyl carrier)	Hs.435741	16q23.2	1.77
	CSK	c-src tyrosine kinase	Hs.77793	15q23-q25	1.74
	MAPKAP1	Mitogen-activated protein kinase associated protein 1	Hs.495138	9q33.3	1.72
	YWHAH	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	Hs.226755	22q12.3	1.71
	CSNK2B	Casein kinase 2, beta polypeptide	Hs.73527	6p21-p12|6p21.3	1.63
	YWHAQ	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide	Hs.74405	2p25.1	1.50
	FGG	**Fibrinogen gamma chain**	**Hs.546255**	**4q28**	**0.62**
	CTNNB1	**Catenin (cadherin-associated protein), beta 1, 88 kDa**	**Hs.476018**	**3p21**	**0.52**
	PDGFRA	**Platelet-derived growth factor receptor, alpha polypeptide**	**Hs.74615**	**4q11-q13**	**0.47**
	SPRY2	**Sprouty homolog 2 (*Drosophila*)**	**Hs.18676**	**13q31.1**	**0.39**
	CAMK2N1	**Calcium/calmodulin-dependent protein kinase II inhibitor 1**	**Hs.197922**	**1p36.12**	**0.30**
	HHIP	**Hedgehog interacting protein**	**Hs.507991**	**4q28-q32**	**0.29**
	JUNB	**Jun B proto-oncogene**	**Hs.25292**	**19p13.2**	**0.23**
Ubiquitin-proteasome pathway	RNF38	Ring finger protein 38	Hs.333503		2.11
	PSMD10	Proteasome (prosome, macropain) 26S subunit, non-ATPase, 10	Hs.522752	Xq22.3	1.86
	SUMO2	SMT3 suppressor of mif two 3 homolog 2 (yeast)	Hs.380973	17q25.1	1.73
	PSMA7	Proteasome (prosome, macropain) subunit, alpha type, 7	Hs.233952	20q13.33	1.61
Protein folding	HSPCB	Heat shock 90 kDa protein 1, beta	Hs.509736	6p12	2.56
	PPIA	Peptidylprolyl isomerase A (cyclophilin A)	Hs.356331	7p13-p11.2	2.00
	HSPA9B	Heat shock 70 kDa protein 9B (mortalin-2)	Hs.184233	5q31.1	1.81
	PPIB	Peptidylprolyl isomerase B (cyclophilin B)	Hs.434937	15q21-q22	1.74
	DNAJA1	**DnaJ (Hsp40) homolog, subfamily A, member 1**	**Hs.445203**	**9p13-p12**	**0.62**
	DNAJB9	**DnaJ (Hsp40) homolog, subfamily B, member 9**	**Hs.6790**	**7q31|14q24.2-q24.3**	**0.50**
Transport	VDAC3	Voltage-dependent anion channel 3	Hs.491597	8p11.2	3.18
	TOMM20	Translocase of outer mitochondrial membrane 20 homolog (yeast)	Hs.533192	1q42	2.60
	ARF6	ADP-ribosylation factor 6	Hs.525330	14q21.3	2.33
	KTN1	Kinectin 1 (kinesin receptor), protein disulfide isomerase family A, member 6	Hs.509414	14q22.1	2.17
	CSE1L	CSE1 chromosome segregation 1-like (yeast)	Hs.90073	20q13	1.93
	PDIA3	Protein disulfide isomerase family A, member 3	Hs.308709	15q15	1.93
	SLC30A1	**Solute carrier family 30 (zinc transporter), member 1**	**Hs.519469**	**1q32-q41**	**0.65**
	ATP1B1	**ATPase, Na^+/K^+ transporting, beta 1 polypeptide**	**Hs.291196**	**1q24**	**0.54**
Metabolism	ELOVL6	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like)	Hs.412939	4q25	5.77
	DGAT2	Diacylglycerol O-acyltransferase homolog 2 (mouse)	Hs.334305	11q13.5	2.78
	SCD	Stearoyl-CoA desaturase (delta-9-desaturase)	Hs.558396	10q23-q24	2.65
	UCK2	Uridine-cytidine kinase 2	Hs.458360	1q23	2.44
	ALDH6A1	Aldehyde dehydrogenase 6 family, member A1	Hs.293970	14q24.3	2.42
	TYMS	Thymidylate synthetase	Hs.369762	18p11.32	2.03
	CYP51A1	Cytochrome P450, family 51, subfamily A, polypeptide 1	Hs.417077	7q21.2-q21.3	1.86
	IDH3B	Isocitrate dehydrogenase 3 (NAD^+) beta	Hs.436405	20p13	1.72
	ALAD	**Aminolevulinate, delta-, dehydratase**	**Hs.1227**	**9q33.1**	**0.65**

Table IV. Continued.

Function	Gene symbol	Gene title	UniGene ID	Chromosomal location	Fold change[a]
	CP	**Ceruloplasmin (ferroxidase)**	**Hs.554736**	**3q23-q25**	**0.60**
	SLC27A2	**Solute carrier family 27 (fatty acid transporter), member 2**	**Hs.11729**	**15q21.2**	**0.57**
	ACAA2	**Acetyl-Coenzyme A acyltransferase 2**	**Hs.200136**	**18q21.1**	**0.45**
	CES1	**Carboxylesterase 1 (monocyte/macrophage serine esterase 1)**	**Hs.535486**	**16q13-q22.1**	**0.40**
	ASS	**Argininosuccinate synthetase**	**Hs.558301**	**9q34.1**	**0.19**
	SAT	**Spermidine/spermine N1-acetyltransferase**	**Hs.28491**	**Xp22.1**	**0.16**
Metal ion binding	**MT1G**	**Metallothionein 1G**	**Hs.433391**	**16q13**	**0.43**
	MT2A	**Metallothionein 2A**	**Hs.418241**	**16q13**	**0.34**
	MT1F	**Metallothionein 1F (functional)**	**Hs.513626**	**16q13**	**0.33**
	MT1X	**Metallothionein 1X**	**Hs.374950**	**16q13**	**0.32**
	MT1H	**Metallothionein 1H**	**Hs.438462**	**16q13**	**0.23**
Unknown	ProSAPiP1	ProSAPiP1 protein	Hs.90232	20p13	7.39
	LOC145853	Hypothetical LOC145853	Hs.438385	15q23	3.56
	MARCKSL1	MARCKS-like 1	Hs.75061	1p35.1	2.00
	MTDH	Metadherin	Hs.377155	8q22.1	1.92
	SERF2	Small EDRK-rich factor 2	Hs.424126	15q15.3	1.60
	C20orf24	Chromosome 20 open reading frame 24	Hs.184062	20q11.23	1.52
	LOC339562	**Similar to Ig kappa chain**	**Hs.449972**	**2p11.1**	**0.40**

[a]Fold change refers to the ratio of the tumors relative to the non-tumors, downregulated genes are highlighted in bold.

Notably, genes related to cell adhesion/cytoskeleton organization also formed a dominant gene group. Among these genes, collagen associated genes such as COL1A1, COL1A2, COL4A2, metallopeptidase related genes such as ADAM10, as well as tubulin related genes such as ubiquitous α-tubulin (K-ALPHA-1) were upregulated, while PCDHA9, a protocadherin α 9 gene, was downregulated. Upregulation of COL1A1, COL1A2 has been observed in human gastric carcinoma (14) and HCCs (15), respectively. A direct comparison of the surrounding hepatic portion of the HCC liver with the healthy control liver by SAM also showed the upregulation of COL1A1 in this study, probably indicating the roles of this gene in the early events of tumorigenesis.

In this study, the expression levels of 6 genes were altered by more than 5-fold. ProSAPIP1 (proline rich synapse associated protein interacting protein 1, undefined functions), had the highest upregulation of 7.3-fold followed by ELOVL6 (elongation of long chain fatty acid, 5.77-fold) and ZNF6, a zinc family protein associated with transcription. SAT (spermidine/spermine N1-acetyltransferase), which is the rate limiting enzyme in polyamine catabolism and decreases cellular spermidine/spermine pools, had a 6.2-fold decrease, only second to the last of IMP-2, an insulin-like growth factor 2 binding protein. Another downregulated gene, ASS (argininosuccinate synthetase) has a role in urea metabolisms.

More importantly, among the genes identified in this study, numerous genes (≥63) have been reported in human HCC. In particular, microarray studies performed on hepatitis B virus-positive hepatocellular carcinoma (10,11,15-17) have discovered many interesting genes, e.g., ALDH6A1 (aldehyde dehydrogenase 6 family, member A1), one of the superfamily of NAD(P)(+)-dependent enzymes implicating a role in the valine and pyrimidine catabolic pathways (15), and Ran, a member of RAS oncogene family (16). Other biological techniques have also identified some differentially expressed genes in human HCC, including CSDA (18), CSK (19), CYP51A1 (20), JUNB (21), TYMS (22), which will be analyzed further in the Discussion.

Quantitative RT-PCR results. To validate our microarray data, quantitative RT-PCR analysis was performed. Due to the limited number of sequenced woodchuck genes (around 600), we selected 10 sequenced genes for RT-PCR study, including genes involved in glycolysis and the tricarboxylic acid cycle (GCKR, IDH3B, and ADH1), fatty acid metabolism (FACVL1, CPT2, HMGCS2, SCD, ELOVL5) and hepatocellular carcinogenesis (AFP). As shown in Table V, 3 woodchuck genes (SAT, IDH3B, SCD) that were differentially expressed in the microarray match perfectly with RT-PCR. Three woodchuck genes that were not differentially expressed in microarray (FACVL1, GCKR, ELOVL5) correlate well with the RT-PCR results in ≥50% samples, assuming that like in microarray, RT-PCR has no significant change if the fold change is less than 1.5. Similarly, the gene expression levels of CPT2 and ADH1 had no consistent change in RT-PCR as shown in microarray. Although some genes showed upregulation (e.g., AFP) or downregulation (HMGCS2) only in RT-PCR, the correlation to the microarray result was still observed in some of the samples (especially for AFP). AFP is clinically used as a serum marker correlating with advanced stages of HCC (23), its upregulation in RT-PCR is consistent with expected results.

Discussion

The high overall incidence and death rate of hepatocarcinoma have prompted the development of reliable methods for early detection and effective treatment of this fatal disease.

Table V. RT-PCR results and comparison with the microarray data.

		W5904		W6361		W6380		W6337		W6384		W7034	
Gene	Process	RT-PCR	Gene chip	RT-PCR	Gene chip	RT-PCR	Gene chip	RT-PCR	Gene chip	RT-PCR	Gene chip	RT-PCR	Gene chip
Fatty-acid-Coenzyme A ligase, very long-chain 1 (FACVL1)	Lipid metabolism, fatty acid metabolism	0.19	down	undetermined	down	3.10	no change	0.605	down	0.33	down	3.65	no change
Alpha-fetoprotein (AFP)	Immune response, transport	320.63	up	3.64	up	4.97	up	42.80	no change	2.10	no change	17.30	no change
Hexokinase (GCKR)[a]	Carbohydrate metabolism	undetermined	no change	undetermined	up	3.58	no change	0.80	no change	1.21	no change	1.35	no change
Spermidine/spermine N1-acetyltransferase (SAT)	Polyamine metabolism	0.33	down	0.04	down	0.96	down	0.52	down	0.32	down	0.63	down
Isocitrate dehydrogenase 3 (NAD+) beta (IDH3B)	Tricarboxylic acid cycle	1.59	up	undeter mined	up	3.05	up	2.00	up	1.00	up	2.50	up
Carnitine palmitoyl-transferase II (CPT2)	Lipid metabolism, fatty acid metabolism, transport	2.81	no change	2.87	no change	1.12	no change	0.64	no change	0.51	no change	1.16	no change
ELOVL family member 5, elongation of long chain fatty acids (ELOVL5)	Protein biosynthesis	2.26	no change	1.13	no change	1.33	no change	1.19	no change	1.41	no change	1.22	no change
Alcohol dehydrogenase 1 (class I), ADH1	Alcohol metabolism/ oxidation	0.07	no change	0.12	no change	2.01	no change	0.414	no change	0.95	no change	0.71	no change
Stearoyl-CoA desaturase (delta-9-desaturase) (SCD)	Fatty acid biosynthesis, lipid biosynthesis	118.90	up	56.40	up	9.72	up	11.8	up	68.4	up	2.7	no change
3-hydroxy-3-methyl-glutaryl-Coenzyme A synthase 2 (HMGCS2)	Acetyl-CoA metabolism, cholesterol/steroid/ lipid/sterol biosynthesis	0.30	no change	0.32	no change	0.608	no change	0.39	no change	0.46	down	0.6	no change

[a]RT-PCR results are for glucokinase regulatory protein (GCKR).

In this study, we used human microarray chips to detect the differentially expressed genes between liver tumor and non-tumor liver tissues in the woodchuck model of HCC and to gain new insights into the molecular mechanisms of biological process underlying the development of HCC for identifying possible markers for early diagnosis as well as for effective treatments.

Gene chip data analysis. Due to absence of native gene chips for woodchucks, we used human microarrays to study the gene expression profiles in woodchuck HCC. It was found that over 10% of the total probe sets (>5500 probe sets) in the human array could be statistically detected on the woodchucks, suggesting that their corresponding gene sequences in humans and woodchucks may be conserved. Furthermore, alteration of the expressions of some genes in tumors could also be detected by using this microarray technique. We thus believe that the cross-hybridization method can be applied to the woodchuck model of HCC.

Analyses of gene expression profiles are not always straightforward because of the high-dimensional nature of the data set. An accurate and reliable data analysis is required to identify a robust set of genes to discriminate HCC from non-tumor liver tissues. Supervised and unsupervised methods are two commonly used strategies for gene expression analysis. Unsupervised method like hierarchical clustering can analyze all the microarray data to reveal the gene expression or array information, whereas supervised analysis is commonly used to identify groups of genes that are associated with biological or pathological features (24). In our previous study, GCOS was used solely for pair-wise data analysis (9). However, a straightforward statistic analysis method is required to differentiate the significantly changed genes from both paired and non-paired specimens. In the present study, SAM, a

supervised analysis method, was further used to find the significantly altered genes based on the GCOS analysis results. A further application of hierarchical cluster and PAM analysis examined the data analysis accuracy and showed only one misclassification. However, this tissue sample labeled as non-tumor (W6380N) was classified more like a tumor tissue, correlating with our histology data and hexokinase activity assay (data not shown). Hence, we were able to achieve an accuracy of ≥93% by using the 289 genes identified by SAM. The combined use of supervised and unsupervised studies made our data analysis more reliable and robust, and offered more insights into the sample information.

As shown in Table V, RT-PCR results matched well with microarray for the selected 3 differentially expressed genes (SAT, IDH3B, and SCD). Five other genes that were not differentially expressed in microarray (FACVL1, GCKR, ELOVL5, CPT2 and ADH1) also showed no consistent change in RT-PCR. Although some genes like AFP and HMGCS2 showed either upregulation or downregulation only in RT-PCR results, this discrepancy may be contributed to the difference in the woodchuck and human gene sequences, as the signal intensities of these disparate genes in the microarrays were always low. In this study, microarray data analysis first started with the detection of the 'presence' or 'absence' status of genes in the microarray statistically in GCOS, followed by the detection of the gene expression changes in each pair and subsequent determination of the robustly changed and present genes according to the rules described previously in Materials and methods. Based on this, supervised and unsupervised methods were further used to determine the significant genes. The use of such a stringent analysis strategy as well as the RT-PCR results especially of the differentially expressed genes can ensure the validity of our microarray results. Further investigation will be necessary to clarify the differences between the RT-PCR and microarray results on these two genes.

Differentially expressed genes in woodchuck HCC and the correlation with human data. In this study, many classes of genes were found differentially expressed, and involved in immune-response, detoxification, DNA replication/transcription/repair, RNA processing, signal transduction, cell-cell interaction/cell adhesion and metabolism. Hepatitis virus B encoded protein X (HBx) has been proved to be a transactivator, and plays essential roles in pathogenesis via stimulating cell proliferation, reducing DNA repair, inhibiting apoptosis as well as activating various signaling pathways (25). Considering the similarity of human HBV and woodchuck WHV in structure and replicative cycles (4), as well as the conservation of HBx protein among the mammalian hepadnaviruses (26), the observed upregulation of ribosomal, proteasomal and transcriptional genes may be caused by woodchuck hepatitis virus encoded protein, in correlation with the sage transcript analysis of normal primary human hepatocytes expressing oncogenic HBx (27). For example, HBx can interact with PSMA7, a proteasome subunit, and with the ERCC2 and ERCC3 subunits of the basic transcription factor GTF2H4 (28,29). Furthermore, several of the upregulated genes in the woodchuck HCCs with other biological functions may also be related to viral infection. SYNCRIP, a heterogeneous nuclear ribonucleoprotein (hnRNP) for RNA processing, was found to positively regulate mouse hepatitis virus (MHV) RNA synthesis (30). The binding of HBx to the mitochondrial VDAC3, a voltage-dependent anion channel for transport upregulated in this study, induced reactive oxygen species (ROS), which in turn contribute to pathogenesis (31). Heat shock protein 60 (HSP60) and heat shock protein 70 (HSP70) were shown to be the cellular target of HBx (29). CDK4, a cell cycle kinase, was elevated in response to HBx infection (25). These studies show that HBx can interact with multiple cellular genes, stimulating hepatocytes growth and thus, contributing to HCC development.

Among the genes found in this study, IGFBP2 and IGF2BP2 (IMP-2), two insulin-like growth factor related binding proteins, were highly downregulated (4.8- and 7.7-fold, respectively). The downregulation of IGFBP2 was also reported in human HCC (15). The reduced IGFBP expression may cause the overexpression of either IGFs and/or IGF-I receptor (IGFIR), which induce cell transformation and subsequent carcinogenesis by promoting cell cycle progression and inhibition of apoptosis. It is also interesting to note that the downregulation of the metallothionein genes in woodchuck HCCs may result in metal or ROS accumulation in hepatocytes, a condition associated with liver carcinoma, since these proteins serve to detoxify heavy metal ions such as Pt, Cu, Zn and Hg and scavenge ROS. Furthermore, their downregulation level was shown to correlate to the degree of histological differentiation and increased with increasing tumor stage in HCC (32). It is likely that the blockage of the detoxification system is a common pathway during carcinogenesis and/or progression of B-type HCC (11). Genes that are involved with clotting-fibronolytic pathway such as SERPING1 (serpin peptidase inhibitor, C1 inhibitor, member 1) and FGG (fibrinogen gamma) were abnormally expressed, suggesting their link to thromboembolism and cancer (33). Metal transport, blood coagulation and immune response may be associated with early-state liver carcinogenesis.

HBx has been found to activate many signal transduction cascades, including the Ras/MAPK, JNK, NF-κB, and Src-dependent pathways to initiate or promote tumor formation (31). MAPKAP1 (mitogen-activated protein kinase associated protein 1), upregulated in woodchuck HCCs, is a negative regulator of MEKK2 signaling and JNK-AP-1 pathway (34,35). PTEN (phosphatase and tensin homologous on chromosome 10), a tumor suppressor gene frequently mutated or deleted in a number of malignancies including human HCC, negatively regulate AKT/PKB signaling pathway. PTEN was also found to inactivate PKC, MAPK and MAPK kinase, which, in turn, eventually downregulates IGF-II expression caused by HBx during the formation of HCC (36). CSK (Src kinase), upregulated probably by HBx, is known to activate Ras and stimulate HBV replication (29). HHIP, a hedgehog interacting protein, was found downregulated in this study, in agreement with a recent report in human HCC (37), suggesting the possible involvement of hedgehog signaling pathway in the WHV induced woodchuck HCC. Furthermore, FKBP1A, NPM1 may be involved with the NF-κB signaling pathways. Wnt signal pathway was reported to contribute to early events in HCC formation (38). However, the downregulation of ß-catenin TNNB1 as well as the upregulation of CXXC4 and

CSNK2B, two negative regulators of the Wnt signaling pathway that enhance ß-catenin degradation, suggests that Wnt signaling pathway may not play an important role in the woodchuck HCC.

Among the downregulated genes identified in this study, some are known as tumor suppressors in human HCC or other tumors, implicating their crucial roles in woodchuck HCC formation. For example, EGR1 (early response 1), SPRY2 (sprouty homology 2 in *Drosophila*) and JUNB (jun-B proto-oncogene) were downregulated in human HCC (17,21,39), while gene loss (SMARCA2) (40), mutation (ETV6, MCC) (41,42), expression loss or reduction (ARMCX3, immune response gene) (43), were observed in human lung, prostate, colon, pancreas and ovarian cancers. Of interest, MCC (mutated in colorectal cancers) may be involved in early stages of colorectal neoplasia and blocks cell cycle progression from the G0/G1 to S-phase. Ribosomal protein L10 (RPL10), a putative suppressor gene, was upregulated in woodchuck HCCs, in coincidence with a report in prostate cancer cells (44). On the other hand, some upregulated genes in woodchuck HCC have been found to promote the formation of human HCC or other tumors. Among them is CSDA (cold shock protein), which can induce cell proliferation, transformation and genomic instability, contributing to hepatocarcinogenesis (18). H2AFZ (H2A histone family, member Z), upregulated in human HCC (17), is localized to the promoters of cyclin genes and induce their expression (45). The mutagenic or oncogenic APEX1 gene, encoding APEX nuclease 1 (multifunctional DNA repair enzyme) for base-excision repair and transcription, was elevated in 72% osteocarcinoma and associated with the patient survival time (46). Its chronic overexpression may facilitate peroxisome proliferator-induced HCC in rats (47). The metabolic ALDH genes such as ALDH6A1 may be related to the formation of hepatocarcinoma (48).

Some genes found in this study are potential markers in human HCC. For example, ETS1 has been linked to HCC progression, invasion and metastasis (49), whereas claudin-10 predicts disease recurrence after curative hepatectomy (50). The presence of PTEN may correlate with tumor grade and metastasis in HCC (51). CSE1L, a Ran-binding protein involved in nuclear transport (export), as well as in the mitotic spindle checkpoint for genomic stability during cell division, is strongly expressed in poorly differentiated hepatocellular carcinoma (49). Increased expression of CDK4 and CKS2 may be associated with poor human HCC prognosis (2). CDC2 overexpression seems to play a crucial role in modulating the cell cycle progression and cell proliferation of HCC, and significantly predicts recurrence (49). Higher expression of ubiquitination genes in human HCC such as PSMD10 (a subunit of the 26S proteasome), was highly correlated with cell proliferation and survival of patients (2). Enhanced activation of ubiquitin-dependent protein degradation may selectively degrade critical proteins such as cell cycle inhibitor RB and account for deregulation of cell cycle control and faster cell proliferation. Not surprisingly, anti-apoptotic genes such as PTMA, SET, YWHAH, and YWHAQ, were highly expressed in woodchuck HCC, in accordance with human HCC (2). In addition to inhibiting apoptosis, both PTMA and SET also work as part of the inhibitor of acetyltransferase complex that regulates histone modification and gene expression, suggesting their multiple roles in hepatocarcinogenesis (2). Other genes including YY1 (52), SIP1 (53), EIF5A (54), MTDH (55) were shown to be potential markers in a variety of other tumors. Genes contribute to tumor progression and/or metastasis by mediating the overexpression of MMP proteins for extracellular matrix degradation (ETS1), E-cadherin (YY1, SIP1), protein translation (EIF5A), cell-cell interaction (MTDH) as well as cell cycle progressions (CDK4, CDC2). NPM1 (nucleolar phosphoprotein B23, numatrin), which is upregulated in human HCC (17), is involved in centrosome duplication, and may be closely associated with tumor aneuploidy (56).

Important genes involved in HCC could be further identified by expanding these analyses to tumor treatment. MARCSKL1 (MARCKS-like 1) (57), RPL6 (58) as well as some metallothionein genes were found associated with drug resistance or tolerance (32). In particular, APEX1, POLH (polymerase eta) were involved in cellular tolerance to cis-platin- or H_2O_2-induced apoptosis (59,60). In contrast, some differentially expressed genes can be potential intracellular targets for cancer chemotherapy. It was found that vitamins K2 and K3 can induce potent antitumor effect on HCC, at least in part, by significantly reducing CD1 and CDK4 expression levels and inducing G1 arrest of the cell cycle (61). The antitumor activity of 3'-ethynyl nucleosides against human solid tumors is dependent on the uridine-cytidine kinase 2 (UCK2), which was upregulated in this study (62). Thymidylate synthase (TYMS), a gene upregulated in human and our woodchuck HCC, is considered to be the primary site of action for 5-fluorouracil, 5-fluoro-2-prime-deoxyuridine, and some folate analogs (22). Nonsteroidal anti-inflammatory drugs (NSAIDs) and synthetic alkylated polyamine analogs strongly induce SAT enzyme activity, and appear to function as anti-neoplastic agents (63). Recently, cyclopomine, a steroid alkaloid that inhibits Hedgehog signaling by binding to smo, was used to efficiently inhibit cell proliferation and increase apoptosis in hepatocytes (37).

Potential markers. In this study, genes related to lipid metabolism are particularly of our interest, as we aim to develop potential tracers for PET imaging. DGAT2, responsible for triglycerides, stearoyl-CoA desaturase (SCD), the rate-limiting enzyme in the synthesis of unsaturated fatty acids, and ELOVL6, which controls the rate of elongation of fatty acids, were all upregulated in woodchuck HCCs. Another upregulated gene, CYP51A1 (cytochrome P450, family 51, subfamily A, polypeptide 1), encoding monooxygenases which participate in cholesterol synthesis, were positively associated with HCC tumor size, and the expression of specific CYPs was altered in conjunction with progression of HCV-associated HCC (20). A 7.9-fold overexpression of acetyl-CoA synthetase 2 (Thiokinase) (ACSS2) was also observed in RT-PCR (data not shown), which agrees with the literature (64). This cytosolic enzyme catalyzes the activation of acetate for use in lipid synthesis and energy generation by producing acetyl-CoA from acetate. In contrast, HMGCS2 (3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2), ACAA2 (acetyl-coA acyltransferase 2), SLC27A2 (fatty acid transporter) were downregulated in this study. These gene chip results suggest that

fatty acid synthesis may play an important role in hepatocellular carcinoma. Some reports have shown that [^{11}C]-acetate can achieve high quality image, due to the enhanced lipid synthesis in the plasma membrane. The roles of fatty acid in [^{11}C]-acetate imaging are under investigation.

As an initial screening for differentially expressed genes, this microarray study found 211 upregulated genes in the woodchuck HCCs, some of which can be possible molecular imaging targets for hepatocellular carcinoma. For example, immunosuppressive (65) and anti-cancer (66) drugs that target FKBP1A (FK506 binding protein 1A, 12 kDa) are currently available. It was also reported that Eve-1c and Eve-1d, which are 397 and 420 amino acids respectively, can bind to ADAM10 (67), a disintegrin and metalloproteinase 10 (ADAM10), which is a surface protein and plays roles in cleavage of cadherins, cell-cell adhesion, migration, and ß-catenin signaling (68). These upregulated genes in the woodchuck HCC can thus be targeted for imaging purposes by radiolabeling their available antibodies or ligands.

Conclusions. In this study, microarray studies on 14 woodchuck samples were performed, and 289 genes involved in different biological functions were identified by using a robust analysis approach. Many genes in woodchucks were found differentially expressed in human HCC, and many were reported in other human tumors while their expressions in human HCC would be worthy of investigation. The gene expression profiles reported in this study show that hepatitis virus encoded protein plays an essential role in carcinogenesis by deregulating multiple classes of genes associated with ribosomal translation, proteasome degradation, transcription and other binding proteins. Multiple signaling pathways were found to be activated by the viral infection, contributing to tumor development. Tumor suppressor genes such as EGR1, SPRY2, PTEN, oncogenes such as CSDA, H2AFZ, APEX1, and molecular markers such as ETS1, claudin 10, CDK4, transcription factor YY1 may play important roles in tumor development, growth and metastasis. Genes like MT1F, ALDH and APEX1 have found to be related to anti-drug resistance mechanism, the modulation of their level will show significance in tumor treatment. For example, strong downregulation of APEX1 can stop cell proliferation and activate apoptosis (69). Furthermore, CDK4, TYMS, and SAT can serve as potential targets for effective HCC chemotherapy. These human microarray results provide the foundation for our imaging work to explore the woodchuck model of virus-induced hepatocellular carcinoma. Future work will include SAGE transcript analysis or woodchuck gene sequencing to further confirm our results. We will also investigate the regulation of pathways involving PET imaging tracers using enzyme assays and immunohistochemistry. Additional efforts will be focused on developing new radiolabeled antibodies and/or ligands for the imaging targets that were differently expressed in woodchuck HCC.

Acknowledgements

We thank Dr Omer Koc for use of his laboratory equipment, Dr Patrick Leahy for many insightful conversations, and Steve Schomisch and Joe Molter for assistance maintaining the woodchucks. Also, we thank pathologists Dr Gregory MacLennan and Dr Vinayak Mehta. This research was supported by the Gene Expression Array Core Facility of the Comprehensive Cancer Center of Case Western Reserve University and University Hospitals of Cleveland (P30 CA43703). This work was supported in part by an NIH grant CA095307 (Z.L., P.I.).

References

1. Llovet JM, Burroughs A and Bruix J: Hepatocellular carcinoma. Lancet 362: 1907-1917, 2003.
2. Lee JS, Chu IS, Heo J, *et al*: Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology 40: 667-676, 2004.
3. Wang Y, Wu MC, Sham JST, Zhang WG, Wu WQ and Guan XY: Prognostic significance of c-myc and AIB1 amplification in hepatocellular carcinoma - A broad survey using high-throughput tissue microarray. Cancer 95: 2346-2352, 2002.
4. Tennant BC, Toshkov IA, Peek SF, *et al*: Hepatocellular carcinoma in the woodchuck mode of hepatitis B virus infection. Gastroenterology 127: S283-S293, 2004.
5. Gouillat C, Manganas D, Zoulim F, *et al*: Woodchuck hepatitis virus-induced carcinoma as a relevant natural model for therapy of human hepatoma. J Hepatol 26: 1324-1330, 1997.
6. Ji W, Zhou WL, Gregg K, Yu N, Davis S and Davis S: A method for cross-species gene expression analysis with high-density oligonucleotide arrays. Nucleic Acids Res 32: e93, 2004.
7. Shah GR, Azizian M, Bruch D, Mehta R and Kittur D: Cross-species comparison of gene expression between human and porcine tissue, using single microarray platform - preliminary results. Clin Transplant 18: 76-80, 2004.
8. Rinaudo JAS and Gerin JL: Cross-species hybridization: characterization of gene expression in woodchuck liver using human membrane arrays. J Med Virol 74: 300-313, 2004.
9. Anderson PW, Tennant BC and Lee Z: Cross-species hybridization of woodchuck hepatitis virus-induced hepatocellular carcinoma using human oligonucleotide microarrays. World J Gastroenterol 12: 4646-4651, 2006.
10. Kim BY, Lee JG, Park S, *et al*: Feature genes of hepatitis B virus-positive hepatocellular carcinoma, established by its molecular discrimination approach using prediction analysis of microarray. Biochim Biophys Acta 1739: 50-61, 2004.
11. Iizuka N, Oka M, Yamada-Okabe H, *et al*: Comparison of gene expression profiles between hepatitis B virus- and hepatitis C virus-infected hepatocellular carcinoma by oligonucleotide microarray data on the basis of a supervised learning method. Cancer Res 62: 3939-3944, 2002.
12. Chun JH, Kim HK, Kim E, *et al*: Increased expression of metallothionein is associated with irinotecan resistance in gastric cancer. Cancer Res 64: 4703-4706, 2004.
13. Garrett SH, Sens MA, Shukla D, Flores L, Somji S, Todd JH and Sens DA: Metallothionein isoform 1 and 2 gene expression in the human prostate: downregulation of MT-1X in advanced prostate cancer. Prostate 43: 125-135, 2000.
14. Oue N, Hamai Y, Mitani Y, *et al*: Gene expression profile of gastric carcinoma: identification of genes and tags potentially involved in invasion, metastasis, and carcinogenesis by serial analysis of gene expression. Cancer Res 64: 2397-2405, 2004.
15. Iizuka N, Oka M, Yamada-Okabe H, *et al*: Molecular signature in three types of hepatocellular carcinoma with different viral origin by oligonucleotide microarray. Int J Oncol 24: 565-574, 2004.
16. Neo SY, Leow CK, Vega VB, *et al*: Identification of discriminators of hepatoma by gene expression profiling using a minimal dataset approach. Hepatology 39: 944-953, 2004.
17. Xu XR, Huang J, Xu ZG, *et al*: 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, 2001.
18. Hayashi J, Kajino K, Umeda T, Takano S, Arakawa Y, Kudo M and Hino O: Somatic mutation and SNP in the promoter of dbpA and human hepatocarcinogenesis. Int J Oncol 21: 847-850, 2002.
19. Shih WL, Kuo ML, Chuang SE, Cheng AL and Doong SL: Hepatitis B virus X protein activates a survival signaling by linking Src to phosphatidylinositol 3-kinase. J Biol Chem 278: 31807-31813, 2003.
20. Tsunedomi R, Iizuka N, Hamamoto Y, *et al*: Patterns of expression of cytochrome P450 genes in progression of hepatitis C virus-associated hepatocellular carcinoma. Int J Oncol 27: 661-667, 2005.

21. Chang YS, Yeh KT, Yang MY, Liu TC, Lin SF, Chan WL and Chang JG: Abnormal expression of JUNB gene in hepatocellular carcinoma. Oncol Rep 13: 433-438, 2005.
22. Yasui K, Okamoto H, Arii S and Inazawa J: Association of overexpressed TFDP1 with progression of hepatocellular carcinomas. J Hum Genet 48: 609-613, 2003.
23. Patil MA, Chua MS, Pan KH, *et al*: An integrated data analysis approach to characterize genes highly expressed in hepatocellular carcinoma. Oncogene 24: 3737-3747, 2005.
24. Lee JS and Thorgeirsson SS: Genetic profiling of human hepatocellular carcinoma. Semin Liver Dis 25: 125-132, 2005.
25. Wu CG, Salvay DM, Forgues M, Valerie K, Farnsworth J, Markin RS and Wang XW: Distinctive gene expression profiles associated with hepatitis B virus x protein. Oncogene 20: 3674-3682, 2001.
26. Anzola M: Hepatocellular carcinoma: role of hepatitis B and hepatitis C viruses proteins in hepatocarcinogenesis. J Viral Hepat 11: 383-393, 2004.
27. Wu CG, Forgues M, Siddique S, Farnsworth J, Velerie K and Wang XW: SAGE transcript profiles of normal primary human hepatocytes expressing oncogenic hepatitis B virus X protein. FASEB J 16: 1665-1667, 2002.
28. Zhang ZS, Torii N, Furusaka A, Malayaman N, Hu ZY and Liang TJ: Structural and functional characterization of interaction between hepatitis B virus X protein and the proteasome complex. J Biol Chem 275: 15157-15165, 2000.
29. Zhang XD, Zhang H and Ye LH: Effects of hepatitis B virus X protein on the development of liver cancer. J Lab Clin Med 147: 58-66, 2006.
30. Choi KC, Mizutani A and Lai MMC: SYNCRIP, a member of the heterogeneous nuclear ribonucleoprotein family, is involved in mouse hepatitis virus RNA synthesis. J Virol 78: 13153-13162, 2004.
31. Waris G, Huh KW and Siddiqui A: Mitochondrially associated hepatitis B virus X protein constitutively activates transcription factors STAT-3 and NF-kappa B via oxidative stress. Mol Cell Biol 21: 7721-7730, 2001.
32. Endo T, Yoshikawa M, Ebara M, *et al*: Immunohistochemical metallothionein expression in hepatocellular carcinoma: relation to tumor progression and chemoresistance to platinum agents. J Gastroenterol 39: 1196-1201, 2004.
33. Lip GYH, Chin BSP and Blann AD: Cancer and the prothrombotic state. Lancet Oncol 3: 27-34, 2002.
34. Cheng J, Zhang DY, Kim K, Zhao YX, Zhao YM and Su B: Mip1, an MEKK2-interacting protein, controls MEKK2 dimerization and activation. Mol Cell Biol 25: 5955-5964, 2005.
35. Schroder W, Bushell G and Sculley T: The human stress-activated protein kinase-interacting 1 gene encodes JNK-binding proteins. Cell Signal 17: 761-767, 2005.
36. Kang-Park S, Lee YI and Lee YI: PTEN modulates insulin-like growth factor II (IGF-II)-mediated signaling; the protein phosphatase activity of PTEN downregulates IGF-II expression in hepatoma cells. FEBS Lett 545: 203-208, 2003.
37. Patil Mohini A, Zhang J, Ho C, Cheung S-T, Fan S-T and Chen X: Hedgehog signaling in human hepatocellular carcinoma. Cancer Biol Ther 5: 111-117, 2006.
38. Feitelson MA, Sun B, Tufan NLS, Liu J, Pan JB and Lian ZR: Genetic mechanisms of hepatocarcinogenesis. Oncogene 21: 2593-2604, 2002.
39. Fong CW, Chua MS, McKie AB, *et al*: Sprouty 2, an inhibitor of mitogen-activated protein kinase signaling, is down-regulated in hepatocellular carcinoma. Cancer Res 66: 2048-2058, 2006.
40. Reisman DN, Sciarrotta J, Wang WD, Funkhouser WK and Weissman BE: Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis. Cancer Res 63: 560-566, 2003.
41. Kibel AS, Faith DA, Bova GS and Isaacs WB: Mutational analysis of ETV6 in prostate carcinoma. Prostate 52: 305-310, 2002.
42. Matsumine A, Senda T, Baeg GH, *et al*: MCC, a cytoplasmic protein that blocks cell cycle progression from the G(0)/G(1) to S phase. J Biol Chem 271: 10341-10346, 1996.
43. Kurochkin IV, Yonemitsu N, Funahashi S and Nomura H: ALEX1, a novel human armadillo repeat protein that is expressed differentially in normal tissues and carcinomas. Biochem Biophys Res Commun 280: 340-347, 2001.
44. Karan D, Kelly DL, Rizzino A, Lin MF and Batra SK: Expression profile of differentially-regulated genes during progression of androgen-independent growth in human prostate cancer cells. Carcinogenesis 23: 967-975, 2002.
45. Dhillon N, Oki M, Szyjka SJ, Aparicio OM and Kamakaka RT: H2A.Z functions to regulate progression through the cell cycle. Mol Cell Biol 26: 489-501, 2006.
46. Wang D, Luo MH and Kelley MR: Human apurinic endonuclease 1 (APE1) expression and prognostic significance in osteosarcoma: enhanced sensitivity of osteosarcoma to DNA damaging agents using silencing RNA APE1 expression inhibition. Mol Cancer Ther 3: 679-686, 2004.
47. Holmes EW, Bingham CM and Cunningham ML: Hepatic expression of polymerase beta, Ref-1, PCNA, and Bax in WY 14,643-exposed rats and hamsters. Exp Mol Pathol 73: 209-219, 2002.
48. Park KS, Cho SY, Kim H and Paik YK: Proteomic alterations of the variants of human aldehyde dehydrogenase isozymes correlate with hepatocellular carcinoma. Int J Cancer 97: 261-265, 2002.
49. Qin LX and Tang ZY: The prognostic molecular markers in hepatocellular carcinoma. World J Gastroenterol 8: 385-392, 2002.
50. Cheung ST, Leung KL, Ip YC, *et al*: Claudin-10 expression level is associated with recurrence of primary hepatocellular carcinoma. Clin Cancer Res 11: 551-556, 2005.
51. Lu DD, Zhang XR and Cao XR: Expression and significance of new tumor suppressor gene PTEN in primary liver cancer. J Cell Mol Med 7: 67-71, 2003.
52. Wang CC, Tsai MF, Hong TM, *et al*: The transcriptional factor YY1 upregulates the novel invasion suppressor HLJ1 expression and inhibits cancer cell invasion. Oncogene 24: 4081-4093, 2005.
53. Comijn J, Berx G, Vermassen P, *et al*: The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. Mol Cell 7: 1267-1278, 2001.
54. Chen G, Gharib TG, Thomas DG, *et al*: Proteomic analysis of eIF-5A in lung adenocarcinomas. Proteomics 3: 496-504, 2003.
55. Brown DM and Ruoslahti E: Metadherin, a cell surface protein in breast tumors that mediates lung metastasis. Cancer Cell 5: 365-374, 2004.
56. Okuda M, Horn HF, Tarapore P, *et al*: Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. Cell 103: 127-140, 2000.
57. Du JP, Shi YQ, Pan YL, *et al*: Regulation of multidrug resistance by ribosomal protein L6 in gastric cancer cells. Cancer Biol Ther 4: 242-247, 2005.
58. Hsia TC, Lin CC, Wang JJ, Ho ST and Kao A: Relationship between chemotherapy response of small cell lung cancer and P-glycoprotein or multidrug resistance-related protein expression. Lung 180: 173-179, 2002.
59. Yang S, Irani K, Heffron SE, Jurnak F and Meyskens FL: Alterations in the expression of the apurinic/apyrimidinic endonuclease-1/redox factor-1 (APE/Ref-1) in human melanoma and identification of the therapeutic potential of resveratrol as an APE/Ref-1 inhibitor. Mol Cancer Ther 4: 1923-1935, 2005.
60. Albertella MR, Green CM, Lehmann AR and O'Connor MJ: A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. Cancer Res 65: 9799-9806, 2005.
61. Hitomi M, Yokoyama F, Kita Y, *et al*: Antitumor effects of vitamins K1, K2 and K3 on hepatocellular carcinoma *in vitro* and *in vivo*. Int J Oncol 26: 713-720, 2005.
62. Murata D, Endo Y, Obata T, *et al*: A crucial role of uridine/cytidine kinase 2 in antitumor activity of 3'-ethynyl nucleosides. Drug Metab Dispos 32: 1178-1182, 2004.
63. Kee K, Foster BA, Merali S, *et al*: Activated polyamine catabolism depletes acetyl-CoA pools and suppresses prostate tumor growth in TRAMP mice. J Biol Chem 279: 40076-40083, 2004.
64. Kurowska EM, Manthey JA, Casaschi A and Theriault AG: Modulation of HepG2 cell net apolipoprotein B secretion by the citrus polymethoxyflavone, tangeretin. Lipids 39: 143-151, 2004.
65. Formica RN, Lorber KM, Friedman AL, Bia MJ, Lakkis F, Smith JD and Lorber MI: The evolving experience using everolimus in clinical transplantation. Transplant Proc 36: 495S-499S, 2004.
66. Sawyers CL: Will mTOR inhibitors make it as cancer drugs? Cancer Cell 4: 343-348, 2003.
67. Tanaka M, Nanba D, Mori S, *et al*: ADAM binding protein Eve-1 is required for ectodomain shedding of epidermal growth factor receptor ligands. J Biol Chem 279: 41950-41959, 2004.
68. Maretzky T, Reiss K, Ludwig A, *et al*: ADAM10 mediates E-cadherin shedding and regulates epithelial cell-cell adhesion, migration, and, beta-catenin translocation. Proc Natl Acad Sci USA 102: 9182-9187, 2005.
69. Fung H and Demple B: A vital role for Ape1/Ref1 protein in repairing spontaneous DNA damage in human cells. Mol Cell 17: 463-470, 2005.