Oncogene (2001) 20, 2704–2712

www.nature.com/onc

Expression profiling and identification of novel genes in hepatocellular carcinomas

Carrie R Graveel[1], Tim Jatkoe[2], Steven J Madore[2], Alison L Holt[1] and Peggy J Farnham*[,1]

[1]*McArdle Laboratory for Cancer Research, University of Wisconsin-Madison, Madison, Wisconsin, WI 53706, USA;* [2]*Genomics and Bioinformatics, Pfizer Inc, Ann Arbor, Michigan, MI 48105, USA*

Liver cancer is the fifth most common cancer worldwide and unlike certain other cancers, such as colon cancer, a mutational model has not yet been developed. We have performed gene expression profiling of normal and neoplastic livers in C3H/HeJ mice treated with diethylnitrosamine. Using oligonucleotide microarrays, we compared gene expression in liver tumors to three different states of the normal liver: quiescent adult, regenerating adult, and newborn. Although each comparison revealed hundreds of differentially expressed genes, only 22 genes were found to be deregulated in the tumors in all three comparisons. Three of these genes were examined in human hepatocellular carcinomas and were found to be upregulated. As a second method of analysis, we used Representational Difference Analysis (RDA) to clone mRNA fragments differentially expressed in liver tumors versus regenerating livers. We cloned several novel mRNAs that are differentially regulated in murine liver tumors. Here we report the sequence of a novel cDNA whose expression is upregulated in both murine and human hepatocellular carcinomas. Our results suggest that DEN-treated mice provide an excellent model for human hepatocellular carcinomas. *Oncogene* (2001) **20,** 2704–2712.

Keywords: hepatocellular carcinoma; representational difference analysis; expression profiling; oligonucleotide microarrays

Introduction

Liver cancer is the fifth most common cancer worldwide with 437 000 cases reported in 1990 (Parkin *et al*., 1999) with hepatocellular carcinomas (HCC) accounting for 85% of liver cancer cases. Several risk factors have been correlated with the development of liver cancer, such as hepatitis B and C infection and exposure to aflatoxin β1. Accordingly, the occurrence of liver cancer is highest in areas of Asia and Africa where the population shows a high prevalence of hepatitis B and C infection (Bosch *et al*., 1999). Unlike other cancers such as skin and colon, a clear mutational model has not been developed for liver cancer and there is currently a lack of successful treatment options. The identification of specific genes that are deregulated in liver cancer is a critical first step in developing a successful strategy for the treatment of liver cancer.

In an effort to address these questions, we attempted to identify genes that show differential expression in liver tumors as compared to normal liver tissue. We used C3H/HeJ male mice due to their high susceptibility to diethylnitrosamine (DEN)-induced liver tumors. In this study, the liver tumors were classified as hepatomas (mixed types A and B). We compared gene expression profiles in liver tumors to three different models of normal proliferation: the quiescent liver, regenerating liver and newborn liver. It is likely that many genes are differentially expressed in a quiescent liver when compared to a liver tumor; however, our goal was to identify genes which are critical to the development of the tumor and not just a consequence of increased proliferation. Because of the unique proliferative abilities of the liver in response to injury, such as partial hepatectomy, we have been able to compare gene expression in nonneoplastic proliferating livers to liver tumors. Finally, because HCC often presents gene expression profiles characteristic of less differentiated hepatocytes (Kojiro and Nakashima, 1999), we have also compared the tumors to newborn livers. We utilized two methods, each with distinct advantages and disadvantages, to assess gene expression changes in liver tumors. We began by using oligonucleotide microarrays due to their ability to detect modest changes in gene expression and their ability to easily compare multiple samples. Using this approach we were able to examine the mRNA expression of 6500 murine genes in three pairwise comparisons of liver tumors to quiescent, regenerating, and newborn livers. However, oligonucleotide arrays do not detect genes at extremely low levels or genes whose probes are absent from the microarray cannot be detected. Therefore, we also examined gene expression differences between liver tumors and regenerating livers using RDA. As reported below, the combination of these two techniques is an effective method for identifying gene expression alterations in liver cancer

*Correspondence: PJ Farnham
Received 6 November 2000; revised 7 February 2001; accepted 12 February 2001

and for isolating previously unknown tumor-specific genes.

Results

Identification of gene expression alterations in liver tumors with oligonucleotide microarrays

To analyse gene expression alterations during hepatocarcinogenesis, we used oligonucleotide microarrays to compare gene expression profiles of 6500 murine genes in liver tumors to three states of normal proliferation: quiescent liver, regenerating liver and newborn liver. A comparison of liver tumors to quiescent livers identified genes whose expression was altered after neoplastic transformation. The comparison of liver tumors to two normal states of proliferation, the regenerating and newborn livers allowed the identification of genes that are tumor-specific as opposed to proliferation-specific. As a control, we also compared gene expression in quiescent livers to regenerating livers; as expected, we observed several genes, such as *cdc2* and *cyclin B2*, whose expression was increased 65- and fivefold respectively, in the regenerating liver (data not shown). Information about all genes that showed a greater than 2.5-fold difference in each comparison can be found on our website: http://mcardle.oncology.wisc.edu/farnham.

Because of the immense number of genes differentially expressed in each pairwise comparison, we used the Absolute Call, a value determined by the expression algorithm, to prioritize our results. On the microarrays, each gene is represented by 20 pairs of 25 oligomers called probe pairs. Each probe pair contains a perfect match and a mismatch oligomer, which is identical to the perfect match except for a mismatched base at the center position. The Absolute Call of *present, absent* or *marginal* is based on the number of probe pairs for each gene or EST cluster that showed a positive hybridization signal and the ratio of hybridization to the perfect match and mismatch probes. The Absolute Call is useful for eliminating signals that may be due to nonspecific hybridization. We initially focused on those genes considered to be *present* in each pairwise comparison. Shown in Figure 1a is a Venn diagram representing the number of differentially expressed genes in each normal sample as compared to the liver tumors. For instance, 162 genes were changed in expression by at least 2.5-fold in the regenerating liver *vs* liver tumor comparison while 297 genes were differentially expressed in the newborn *vs* tumor comparison; 69 genes were differentially expressed in both the regenerating and newborn livers as compared to liver tumors. Although hundreds of genes were differentially expressed in a single pairwise comparison, only 22 genes were altered in all three comparisons. The names and fold difference levels for each of these 22 genes are shown in Table 1a. Although the reliability and reproducibility of data obtained using oligonucleotide microarrays has been previously documented (Coller *et al*., 2000), we confirmed the differential expression of several of the genes listed in Table 1a using RT–PCR. For these experiments, seven individual tumors were analysed since each hepatocellular carcinoma arises independently and may present different gene expression patterns (Farber, 1976). Shown in Figure 1b is the differential expression of seven genes identified as upregulated and four genes identified as downregulated in liver tumors. GAPDH expression was examined in the RNA samples to confirm equal quantitation (data not shown). The RT–PCR results were very similar to the oligonucleotide microarray results, therefore, we did not perform RT–PCR on all 22 genes listed in Table 1a.

In addition to identifying genes which are differentially regulated in liver tumors as compared to all three normal liver samples, we wanted to identify genes which were significantly differentially expressed in tumors as compared to the normal quiescent state. Such genes may serve as markers for the altered proliferation characteristic of HCC. Shown in Table 1b is a list of genes which were greater than 10-fold upregulated or downregulated in the liver tumor as compared to the quiescent liver. As noted above, the Absolute Call can be used to eliminate nonspecific hybridization signals. However, an *absent* call may also be made if the expression of an mRNA is extremely low in a sample. Since, in this case, we are interested in identifying mRNAs that are extremely low in quiescent livers, we have included both *absent* and *present* genes in Table 1b. We have examined the expression of three highly differentially expressed genes by RT–PCR (Figure 1c). We note that only two of the 16 genes in Table 1b are present in Table 1a, which is most likely due to extremely low expression of those genes in quiescent livers. The results of the RT–PCR confirm the differential expression in quiescent liver versus the liver tumor but also show very little differential expression when newborn and regenerating livers are compared to the liver tumors.

Finally, we were interested in knowing if liver tumors represent a less differentiated state. Therefore, we have identified mRNAs which are newborn specific and deregulated in the adult liver tumors. In the regenerating *vs* newborn comparison we identified genes which are newborn-specific and not just involved in increased proliferation. Of the 194 mRNAs which are deregulated in the quiescent *vs* tumor comparison, 88 of them are differentially expressed in the regenerating *vs* newborn comparison. Genes such as *H19* and *insulin-like growth factor binding protein-1* (*IGFBP-1*) are known to be expressed during fetal development. In this study, both *H19* and *IGFBP-1* were upregulated in liver tumors and in newborn livers. Other genes that were deregulated in the tumor and newborn liver are *CD63, testosterone 16α-hydroxylase* and *intestinal trefoil factor*. These results suggest that the liver tumors are less differentiated than adult hepatocytes and these genes may be markers of a less differentiated state.

We also wanted to examine the expression of a few of the most highly deregulated genes in human

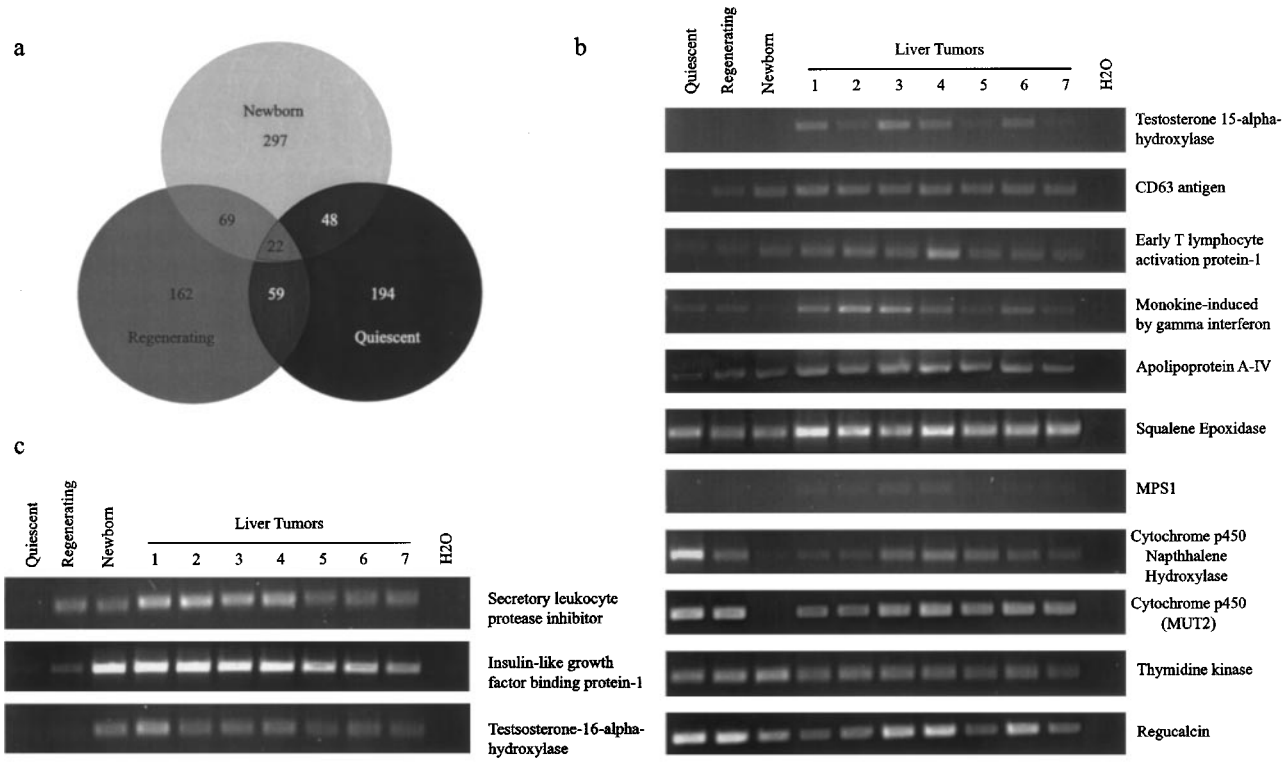

Figure 1 Analysis and confirmation of oligonucleotide microarray results. (**a**) Venn diagram of the genes found to be present and at least 2.5-fold differentially expressed when compared to tumors by oligonucleotide microarrays. (**b**) Confirmation of oligonucleotide microarray results by RT–PCR. Quiescent and regenerating samples are a combination of four livers and newborn samples are a combination of eight livers. Tumor samples represent individual tumors. Tumors 1–4 were used in oligonucleotide microarray experiments. For each reaction, 100 ng of cytoplasmic RNA was used. (**c**) Confirmation by RT–PCR of genes found to be differentially expressed in the quiescent *vs* tumor comparison. RNA samples were pooled as described in (**b**)

hepatocellular carcinomas. Using RT–PCR (Figure 2), we analysed the expression of *CD63, osteopontin* (the human homologue of *Eta-1*), and *monokine induced by γ interferon* (*MIG*) in HepG2 cells (a hepatoma derived cell line), human hepatocellular adenomas and carcinomas. Signal intensity was normalized to GAPDH in each reaction. The expression of these genes was measured in three individual HCC samples because of the heterogeneity of each tumor. Due to the rarity of adenomas, only one sample was examined. For *CD63* and *MIG* we observed an overall upregulation of expression in both the adenoma and HCCs, but osteopontin was variably upregulated in the three HCCs but not the adenoma. This pattern of upregulated expression in human HCC is similar to the results observed in the mouse liver tumors. These results indicate the relevancy of the DEN mouse model in studying the molecular profile of human HCC.

Identification of gene expression alterations in liver tumors by RDA

Although oligonucleotide microarrays provide a rapid means of screening many genes, they are not suitable for detection of rare transcripts and they limit analyses to previously cloned mRNAs. Therefore, we utilized a second technique, RDA, to identify gene expression differences in tumors. We performed two comparisons between the regenerating livers and liver tumors; in one comparison, the tester was the liver tumor (to identify genes upregulated in liver tumors) and in the other comparison, the regenerating liver was the tester (to identify genes upregulated in regenerating livers). Enrichment of difference products (DP) was observed after the second round of subtraction and amplification (DP2). We observed no difference products after the third subtraction and amplification (DP3), perhaps due to the high tester/driver ratio. Consequently, we cloned mRNAs from the DP2 samples. DP2 products that were cloned and sequenced are listed in Table 2. It has been observed previously that the relative difference in expression level of a transcript in the tester *vs* driver usually correlates with the frequency of isolation of that gene by RDA (Welford *et al*., 1998). We noted in parentheses the number of times a gene was cloned from our RDA difference products. We also note that several of the RDA products were also found to be differentially expressed in the oligonucleotide microarray comparisons (see Table 2, italics). The fact that several genes were found to differentially expressed by both RDA and oligonucleotide microarrays demonstrates that the two techniques are complementary. Since RDA has a false positive rate of approximately 10% (Hubank and Schatz, 1999), it is possible that

Table 1 Fold change in microarray comparisons

(a) Genes that are at least 2.5-fold differentially expressed in all three comparisons (quiescent *vs* tumor, regenerating *vs* tumor and newborn *vs* tumor) and considered to be *present* on the oligonucleotide microarrays

Database accession number	*Entrez definition*	*Regenerating* vs *tumor*	*Quiescent* vs *tumor*	*Newborn* vs *tumor*
J03549	Mouse testosterone 15 α-hydroxylase mRNA type 1, complete cds	24.4	4.3	52.7
X58196	Mus musculus H19 mRNA	21.8	22.8	−7.8
D16432	Mouse murine CD63 mRNA for murine homologue of CD63/ME491, complete cds	14.3	28.6	3.6
X16151	Mouse mRNA for early T-lymphocyte activation 1 protein (ETa-1)	14.2	8.8	6.4
M34815	Mouse monokine induced by γ interferon (MIG) mRNA, complete cds	8.5	5.1	14.6
M64250	Mouse apolipoprotein A-IV gene, complete cds, clone Apo4.5	6.5	9.3	11.5
D42048	Mouse mRNA for squalene epoxidase, complete cds	6.2	5.7	3.8
L20315	Mus musculus MPS1 gene and mRNA, 3end	4.9	25.1	6.9
M38337	Mouse milk fat globule membrane protein E8 mRNA, complete cds	4.1	4.9	2.8
W34349	Homologous to sp P00748: Coagulation factor XII precursor (EC 3.4.21.38)	3.8	2.8	2.8
M17440	Mouse sex-limited protein (SlpA) gene, exons 24-41, and cytochrome P-450 (Cyp21(w7A)) gene, exons 1-7	3.1	3.4	5.9
AA116604	Homologous to sp P05689: Cathepsin (EC 3.4.22.-) (fragment).	3	3	4.4
Z22216	M. musculus APOC2 gene, complete CDS, and exons 2 and 3	2.8	5.6	3.8
W21013	Homologous to sp P31361: Brain specific homeobox	2.8	4.3	2.7
AA145371	Homologous to sp P09912: Interferon-induced protein 6-16 precursor	2.6	6.9	4.7
D26137	Mouse cytochrome P450IIIA mRNA, complete cds	−2.6	−3.3	−19.6
M27796	Mouse carbonic anhydrase III (CAIII) mRNA, complete cds	−2.6	−6.6	6.1
M63244	Mus musculus amino levulinate synthase mRNA, complete cds	−2.7	−4.5	−15.6
M77497	Mus musculus cytochrome P-450 naphthalene hydroxylase mRNA, complete cds	−3.3	−8.2	−3.2
D17674	Mouse mRNA for cytochrome P-450, complete cds	−3.6	−5.4	15.5
M68489	Mouse cytosolic thymidine kinase mRNA clone pMtk4, complete cds	−6.8	−4.9	−3.9
U28937	House mouse; Musculus domesticus liver mRNA for regucalcin, complete cds	−10.7	−12.6	−4.4

(b) Fold change in quiescent *versus* tumor comparison. Genes that are greater than 10-fold differentially expressed in the quiescent *vs* tumor comparison on the oligonucleotide microarrays. Positive values represent higher mRNA expression in the tumors and negative values represent lower mRNA expression in the tumors

Database accession number	*Entrez definition*	*Quiescent* vs *tumor*
U73004	Mus musculus secretory leukocyte leukocyte protease inhibitor mRNA, complete cds	51.5
W13166	Mouse SV-40 induced 24p3 mRNA	34.1
D16432	Mouse murine CD63 mRNA for murine homologue of CD63/ME491, complete cds	28.6
X81579	M. musculus mRNA for insulin-like growth factor binding protein-1	26.3
L20315	Mus musculus MPS1 gene and mRNA, 3end	25.1
M60273	Mouse testosterone 16-α-hydroxylase mRNA, complete cds, clone pf26	24.2
X58196	Mus musculus H19 mRNA	22.8
X14194	M. musculus nid gene (exons 19 & 20)	21.9
M26270	Mouse stearoyl-CoA desaturase (SCD2) mRNA, complete cds	15.5
D38410	Mouse mRNA for intestinal trefoil factor, complete cds	10.2
M13522	Mus musculus serum amyloid A protein isoform 2 mRNA, complete cds	−11.3
M26005	Mouse endogenous retrovirus truncated gag protein, complete cds, clone del env-1 3.1	−11.8
U28937	House mouse; Musculus domesticus liver mRNA for regucalcin, complete cds	−12.6
U36993	Mus musculus cytochrome P450 Cyp7b1 mRNA, complete cds	−24
AA139907	M. musculus spot14 gene	−25.1

some of the genes cloned by RDA are not truly differentially expressed. We felt confident that the genes identified by both RDA and the oligonucleotide microarrays were true positives and thus concentrated on confirming those genes only identified by RDA, especially those genes which were cloned more than once. The expression of genes which were only identified by RDA but cloned multiple times was examined by RT–PCR (Figure 3). As shown, *aurora-related kinase 1* and the *ubiquitin-conjugating enzyme* (the mouse homologue of human *cyclin-selective ubiquitin carrier*) have slightly higher expression in the regenerating liver than the liver tumors, but have undetectable expression in the quiescent livers. This suggests these genes are proliferation specific. *Serine proteinase inhibitor 2* (*SPI-2*) was not differentially expressed and is considered a false positive.

Isolation and expression analysis of a novel cDNA

A unique advantage of RDA is the ability to identify previously uncharacterized genes. We isolated five novel genes using liver tumors as the tester and one gene using the liver tumors as the driver. These novel transcripts have no significant homology to any gene in the NCBI database. However, two of the genes

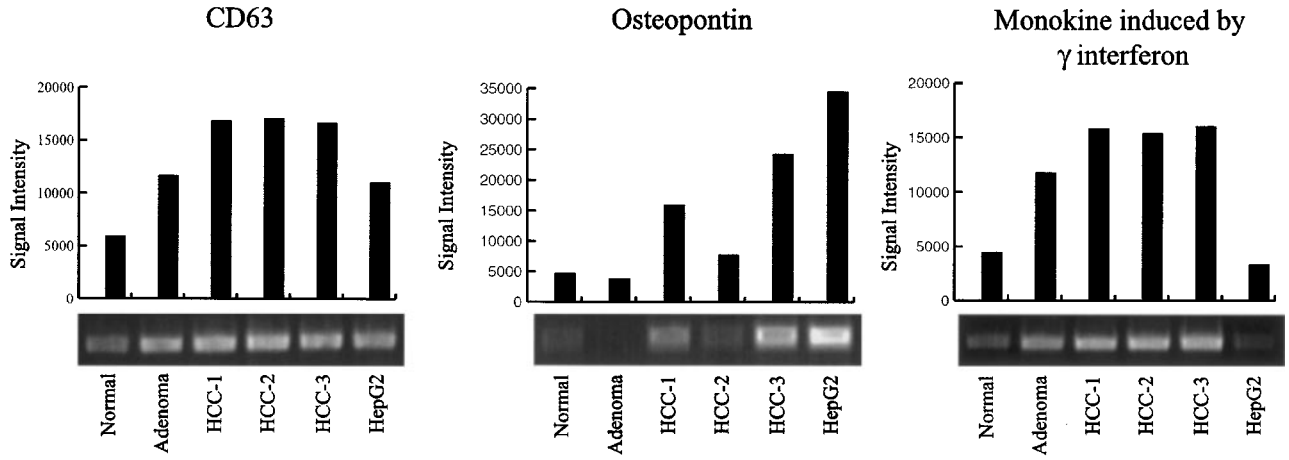

Figure 2 Expression of highly deregulated genes in human HCC. Expression of *CD63, osteopontin*, and *monokine induced by γ interferon* by RT–PCR in normal human liver, hepatocellular adenoma, three independent hepatocellular carcinomas, and HepG2 cells. For each reaction, 100 ng of cytoplasmic RNA was used. Signal intensity was normalized to GAPDH in each reaction

Table 2 Genes found to be differentially expressed by RDA

mRNAs higher in liver tumors than in regenerating livers	*mRNAs lower in liver tumors than in regenerating livers*
Testosterone 15α-hydroxylase (7)	*Viral envelope like protein (11)*
Testosterone 16α-hydroxylase (2)	*Cyclin B1 (6)*
Apolipoprotein A-IV (2)	*SPI-2 gene (6)*
H19 (2)	*Human Cdc20 (4)*
Lipoprotein lipase (1)	Human cyclin-selective ubiquitin carrier (3)
Carboxypeptidase E (1)	Aurora-related kinase 1 (2)
Carboxypeptidase H (1)	*Cyclin B2 (2)*
Rat 10α-hydroxysteroid dehydrogenase (1)	Fibrinogen A-α chain (1)
Unknown 1 (15)	Fibrinogen B-β chain (1)
Unknown 2 (1)	Rabkinesin-6 (1)
Unknown 3 (1)	Complement component C3 (1)
Unknown 4 (1)	Alpha-tubulin isotype (1)
Unknown 5 (1)	Adenovirus type 2 (1)
	RNA1 homolog Fug1 (1)
	LLRep3 protein (1)
	Class 1 alcohol dehydrogenase (1)
	Rat kinesin-related protein (1)
	Nonmuscle tropomyosin 5 (1)
	Serum albumin (1)
	Human ATP synthase (1)
	Haptoglobin (1)
	nimA-related kinase 2 (1)
	Unknown 6 (1)

Those genes in italics were also found to be differentially expressed by oligonucleotide microarray comparisons. The number in parentheses refers to the number of times a gene was isolated from the difference products

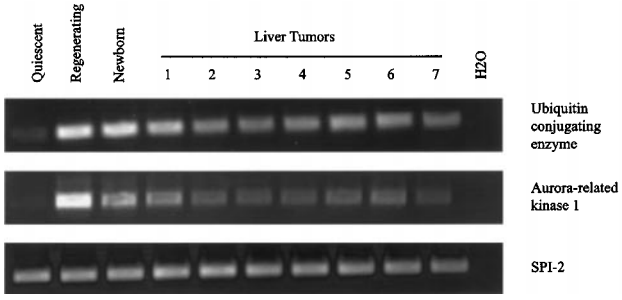

Figure 3 Confirmation by RT–PCR of gene products isolated by RDA. This includes those genes isolated more than one time which were not previously confirmed using oligonucleotide microarrays. RNA samples were pooled as described in Figure 2

upregulated in tumors and the one downregulated gene in tumors did have significant homology to uncharacterized mouse ESTs. To confirm the levels of differential expression of the novel genes, we examined their expression by RT–PCR in quiescent, regenerating and newborn livers as well as in seven liver tumors (Figure 4a). While we observed only modest differential expression for novel genes 1 and 2, novel genes 3 and 4 were highly differentially expressed. Novel gene 5 was not differentially expressed and was considered a false positive. We note that although novel gene 6 showed a slightly higher expression in the regenerating livers than liver tumors, the differential expression in the quiescent liver *vs* liver tumors is much more dramatic, suggesting that this gene may be a proliferation marker.

We focused on novel gene 4 because it was most highly differentially expressed and contained a significant ORF. We examined the expression of novel gene 4 in eight mouse tissues and four embryonic stages using a Multiple Tissue cDNA Panel. Gene 4 was most highly expressed in the heart, lung and testis (Figure 4b). Fairly low expression was seen in the liver sample, a finding in agreement with our previous data showing that this gene is normally expressed at low levels in a quiescent liver when compared to a liver tumor. We cloned the mouse cDNA of novel gene 4 by screening a Rapid-Screen Mouse Liver cDNA Library with primers designed from the isolated RDA fragment. A 4.175 kb cDNA was isolated and sequenced and is now called *CRG-L1* (cancer related gene–liver 1). Significant homology to two mouse ESTs (GenBank accession numbers, AW701866 and AW490555) was

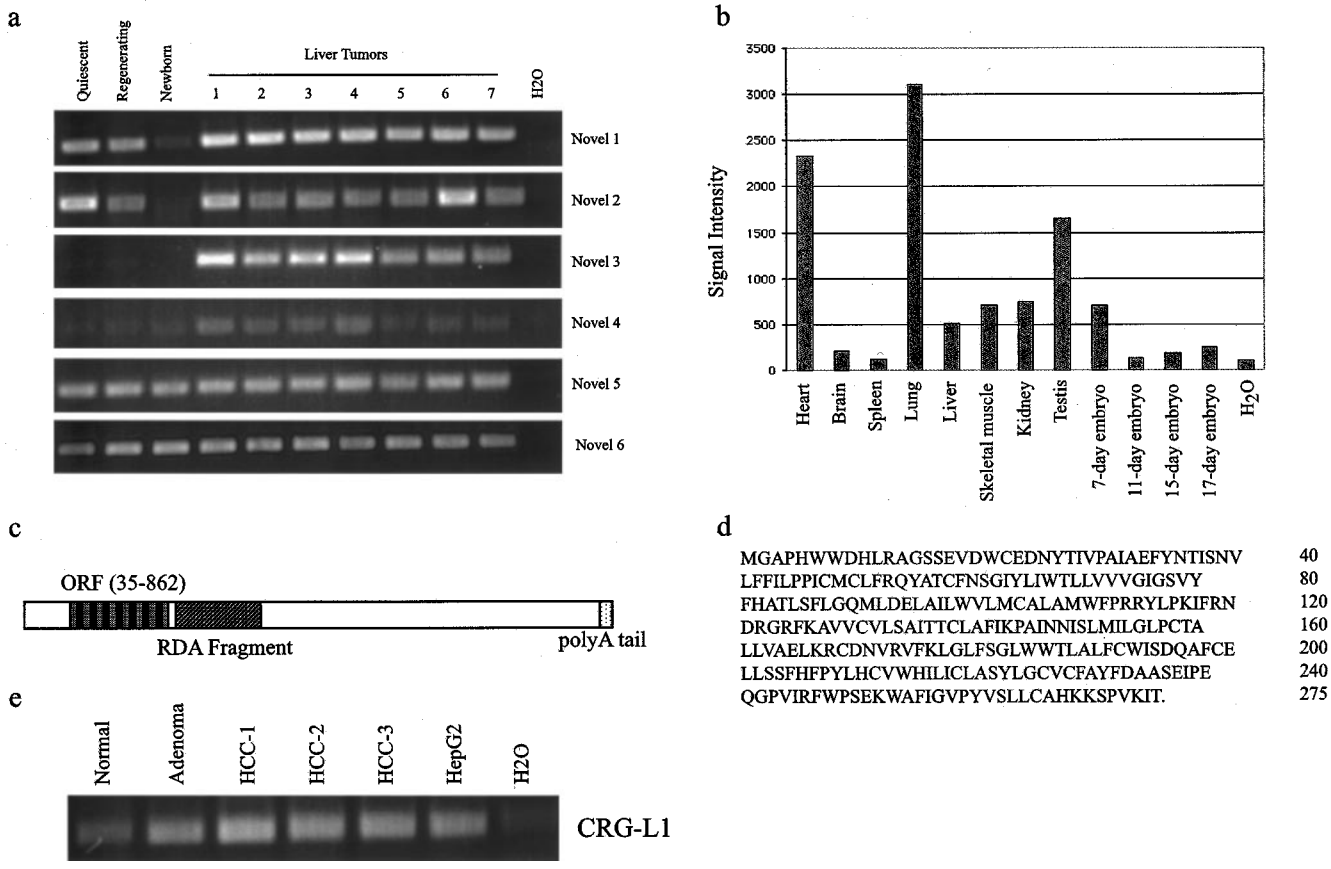

Figure 4 Analysis of novel genes. (**a**) Confirmation of differential expression by RT–PCR of the six novel genes isolated by RDA. RNA samples were pooled as described in Figure 2. (**b**) Relative expression levels of novel gene 4 (*CRG-L1*) in multiple murine tissues. (**c**) Schematic of *CRG-L1* cDNA (GenBank AF282864). Putative transmembrane domains are represented by black boxes. (**d**) Protein sequence of novel 4 ORF (bases 35–862). (**e**) Expression of CRG-L1 in HCC. Expression of *CRG-L1* by RT–PCR in normal human liver, hepatocellular adenoma, three independent hepatocellular carcinomas and HepG2 cells. For each reaction, 100 ng of cytoplasmic RNA was used

found at the 3′ end. *CRG-L1* had no significant homology to any gene in the NCBI database. The RDA fragment that was originally cloned from the DP2 products mapped to bases 996–1383 of *CRG-L1*. A potential ATG translation initiation site (GCGGCCATGG) was found at position 35 with an open reading frame extending to nucleotide 862. The predicted translation contains 275 amino acids with a molecular weight of 31.4 kD (Figure 4d). To search for any known protein motifs, the protein sequence was analysed by SMART (Simple Modular Architecture Research Tool) (Schultz *et al*., 1998, 2000). Seven possible transmembrane domains were found within the protein spanning amino acids 33–53, 62–82, 91–111, 123–143, 146–166, 174–194 and 212–232. No other known protein motifs have been identified.

We also examined the expression of *CRG-L1* in normal human liver, a hepatocellular adenoma, HepG2 cells, and three independent HCC samples. We observed a significant increase in expression in each HCC and HepG2 cells when compared to the normal liver (Figure 4e). The level of expression in the adenoma was slightly less than the HCC levels, suggesting that *CRG-L1* expression correlates with malignancy in the liver.

Discussion

The goal of this study is to identify genes that are important in liver tumorigenesis. One problem with the identification of tumor-specific genes is the lack of a good non-tumorigenic control. Therefore, we began with a mouse model, which allowed us to compare gene expression between liver tumors and three normal states of proliferation using both oligonucleotide microarrays and representational difference analysis. This approach has allowed us to identify both known and novel genes involved in dedifferentiation, proliferation and neoplasia.

By consideration of the mRNAs that are deregulated by at least 2.5-fold in all three normal samples as compared to tumors and of the mRNAs that are greater than 10-fold different in quiescent livers *versus* liver tumors, we have found several different classes of deregulated mRNAs. One category of genes includes

liver metabolic enzymes, such as *naphthalene hydroxylase*, which has been shown previously to be downregulated during mouse hepatocarcinogenesis (Ye *et al*., 1997; Yamada *et al*., 1999). Other liver metabolic genes that we found to be significantly deregulated, such as *apolipoprotein A-IV* and *squalene epoxidase*, have not previously been examined in liver cancer. Another category of genes which are deregulated in the mouse hepatocellular carcinomas are involved in cell adhesion and motility. A few of the genes identified in this study have been previously linked to HCC (i.e. *H19*; Ariel *et al*., 1998), and some genes have been previously shown to be deregulated in different human cancers. For example, *osteopontin* (*Eta-1* homologue) is upregulated in gastric carcinomas, lung adenocarcinomas, and breast cancer (Tuck *et al*., 1998; Ue *et al*., 1998; Shijubo *et al*., 1999). In breast cancer, the expression of this protein is correlated with increased invasiveness and decreased survival (Singhal *et al*., 1997; Tuck *et al*., 1998, 1999).

It is also of interest to consider mRNAs which were not identified as deregulated in the liver tumors. For example, several genes such as *β-catenin, p53* and *Rb* are known to be mutated within hepatocellular carcinomas (Buendia, 2000). *β-catenin* is frequently mutated at the GSK-phosphorylation site in HCC resulting in accumulation of *β-catenin* protein within the nucleus. Although altered *β-catenin* mRNA levels would not be expected to be detected by the microarray studies, it was possible that mRNAs from genes that are regulated by *β-catenin* would have been identified. However, we did not detect an increase in expression of any of the putative *β-catenin* target genes, such as c-*myc*. It is possible that some of the genes found to be deregulated in this study are unidentified targets of *β-catenin*. Alternatively, mutated *β-catenin* may not affect gene expression but inhibit apoptosis (Chen *et al*., 2001). Expression of p53 and Rb may not have been altered in the tumors derived from the DEN-treated mice. Alternatively, proteins highly related to p53 (e.g. p73) and Rb (e.g. p107) may have compensated for such mutations.

To determine if the DEN-treated mice are a good model for HCC, we tested three of the tumor-specific genes in human liver tissue samples. We observed upregulated expression of *osteopontin, CD63* and *MIG*. *CD63* and *MIG* were upregulated in both the adenoma and the three carcinomas, but *osteopontin* was only upregulated in the carcinomas. Therefore, *osteopontin* may not be a marker of proliferation, but may be an indicator of the metastatic potential of a tumor, similar to breast cancer (Tuck *et al*., 1999), which would explain the variable expression of *osteopontin* in each HCC. *MIG* is known to be overexpressed in the endothelium of HCC and may be involved in lymphocyte recruitment (Yoong *et al*., 1999).

Using RDA, we have also identified five novel genes that are differentially expressed in liver tumors as compared to at least one normal state of proliferation. We isolated the mouse cDNA of *CRG-L1* which contains an open reading frame of 827 bases. Seven putative transmembrane domains have been identified, but no other known protein motifs have been distinguished. *CRG-L1* is also upregulated in human hepatocellular adenomas and carcinomas. These results suggest this novel gene may be a critical factor in liver tumor development. Efforts are underway to determine its expression in other human cancers.

In summary, we have identified a number of known and novel genes which are deregulated in murine and human HCC. Importantly, our results indicate that DEN treated mice provide a good model for studying the molecular changes within human HCC. Since studies using human tumors cannot provide information concerning whether the deregulation is an early or late step in carcinogenesis, we plan to examine gene expression in preneoplastic foci in murine livers. These studies will aid in the classification of genes as markers of early or late stages of tumor development. Other future goals are to determine which genes are general tumor markers, which genes are specifically expressed in HCC, and which represent early changes in neoplastic transformation. We hope that our results will lead to a better understanding of the cellular pathways involved in tumor progression.

Materials and methods

Mouse husbandry

Inbred C3H/HeJ mice were bred and housed in the McArdle Laboratory Animal Facilities in plastic cages on corncob bedding from Bed-O'Cobs (Anderson Cob Division) and fed Mouse Chow 9F (Purina). Food and acidified water were available *ad libitum*. To obtain regenerating livers, partial hepatectomies were performed on male, 6-week-old mice as described previously (Lukas *et al*., 1999). Animals were sacrificed 36 h after the surgery, which corresponds to peak DNA synthesis (Bennett *et al*., 1995), and the liver remnants were harvested. Quiescent livers were harvested from 6-week-old, male mice. Newborn livers, from both male and female mice, were isolated within 24 h after birth. The liver tumors (hepatomas type A and B) were taken from male mice that were treated with diethylnitrosamine (DEN) (0.1 μM/g body weight) at 12 days of age and sacrificed at 32 weeks of age.

Human tissue

Human tissue was procured at the University of Wisconsin Surgical Pathology department and through the National Disease Research Interchange. All tissues analysed were primary and noncirrhotic. HCC-2 was hepatitis C positive. As required by our IRB protocol, the identity of the patients was unknown. The excess tissue was frozen after surgery and stored at -70°C.

Preparation of RNA

Total RNA was extracted from liver using guanidine thiocyanate/CsCl as described previously (Lukas *et al*., 1999). Poly(A)$^+$ mRNA was isolated from 250 μg of total RNA using Oligotex mRNA Kit (Qiagen). For the oligonucleotide microarrays, mRNA was purified twice using Oligotex mRNA Kit (Qiagen) and electrophoresed on a 1%

agarose/1 × MAE buffer (50% formamide, 2.2 M formaldehyde, 1 mM 4-morpholinopropanesulfonic acid (MOPS) (pH 7.0), 0.4 M NaOAc, 0.05 mM EDTA) gel to examine for degradation. For oligonucleotide microarray experiments, poly(A)$^+$ RNA samples were pooled from eight newborn livers while the quiescent, regenerating and tumor samples were pooled from four livers to avoid mouse-to-mouse variation. HepG2 RNA was made as described previously (Slansky *et al*., 1993).

Oligonucleotide microarrays

A complete protocol for converting RNA into 'target' for hybridization to microarrays is available at our website: http://mcardle.oncology.wisc.edu/farnham. In brief, twice-purified poly(A)$^+$ RNA from quiescent, regenerating and newborn livers, or liver tumors was used to create cDNA with a T7-polyT primer and reverse transcriptase Superscript II (GIBCO/BRL). Approximately 1 μg of cDNA was subjected to *in vitro* transcription (Ambion) in the presence of biotinylated UTP and CTP (Enzo Diagnostics). The cRNA was fragmented and combined with BSA (0.5 mg/ml) in a buffer containing 2 × MES, 1.7 M NaCl, 40 mM EDTA and 0.02% Tween 20. Target cRNA (10 μg) was hybridized for 16 h at 40°C to each oligonucleotide array (Mu6500 tetraset; Affymetrix) containing probes for more than 6500 murine genes and ESTs. Arrays were washed in the Affymetrix Fluidics Station at 50°C with 6 × SSPE-T (0.9 M NaCl, 60 mM Na_2HPO_4, 6 mM EDTA, 0.005% Triton X-100, pH 7.6) then at 40°C with 0.5 × SSPE-T. Arrays were then stained with streptavidin phycoerythrin (Molecular Probes) and washed with 6 × SSPE-T. Fluorescent intensities were measured with a laser confocal scanner (Hewlett-Packard) and analysed with GeneChip software (Affymetrix).

Expression analysis by representational difference analysis

The protocol developed by Hubank and Schatz (1994) was followed in detail using polyA RNA from regenerating livers and liver tumors. In the first subtractive round, the representations were hybridized to each other in a 1 : 100 tester/driver ratio. The second and third difference products used a tester/driver ratio of 1 : 800 and 1 : 400 000 respectively. There were no products visualized by agarose gel in the third difference products (DP3), so the difference products were subcloned from DP2 into pBSM13+. Cloned products were sequenced by Big Dye (ABI) in the McArdle Laboratory Sequencing Facility.

RT–PCR

Each reaction contained 100 ng of cytoplasmic RNA, 1 × EZ buffer (Perkin Elmer), 0.4 M Betaine (Sigma), 60 nM primers, 300 mM dNTPs, 2.5 mM $Mn(OAc)_2$ and 5U rTth polymerase (Perkin Elmer). After 27–40 cycles of amplification, products were electrophoresed on a 1% agarose gel and visualized by ethidium bromide staining. Details of the primers used for confirmation of differentially expressed mRNAs can be found on our website. All primers were synthesized at the UW Biotechnology Center.

Cloning CRG-L1

The mouse cDNA of novel gene 4 was identified from a Rapid-Screen cDNA Mouse Liver Library Panel (Origene) using primers designed to the cDNA fragment cloned by RDA. A Multiple Tissue cDNA Panel (Clontech) was used to analyse expression in multiple mouse tissues. A 4.175 kb cDNA was isolated and sequenced (Genbank Accession AF282864) in both directions by Big Dye (ABI) in the McArdle Laboratory Sequencing Facility.

Acknowledgments
The authors thank J Messamore, E Perguson, S Bartley, Z Diaz, E Lukas and E Armstrong for technical assistance; G Kennedy for the HepG2 cells; J Menetski for collaborative efforts; and the Farnham lab and M Ryan for helpful discussions and critical review of the manuscript. We also thank Dr T Warner and the National Disease Research Interchange for the human tissues. This work was supported in part by research grants from National Cancer Institute/National Institutes of Health CA22484 and National Institute of General Medical Sciences/National Institutes of Health GM08349.

References

Ariel I, Miao HQ, Ji XR, Schneider T, Roll D, de Groot N, Hochberg A and Ayesh S. (1998). *Mol. Path.,* **51,** 21–25.
Bennett LM, Farnham PJ and Drinkwater NR. (1995). *Mol. Carcin.,* **14,** 46–52.
Bosch FX, Ribes J and Borras J. (1999). *Semin. Liver Dis.,* **19,** 271–285.
Buendia MA. (2000). *Semin. Cancer Biol.,* **10,** 185–200.
Chen S, Guttridge DC, You Z, Zhang Z, Fribley A, Mayo MW, Kitajewski J and Wang C. (2001). *J. Cell Biol.,* **152,** 87–96.
Coller HA, Grandori C, Tamayo P, Colbert T, Lander ES, Eisenman RN and Golub TR. (2000). *Proc. Natl. Acad. Sci. USA,* **97,** 3260–3265.
Farber E. (1976). The pathology of experimental liver cell cancer. *Liver Cell Cancer*. Cameron HM, Linsell DA and Warwick GP. (eds). Elsevier Scientific Publishing Company: New York. pp. 243–277.
Hubank M and Schatz DG. (1994). *Nucleic Acids Res.,* **22,** 5640–5648.
Hubank M and Schatz DG. (1999). *Methods Enzymol.,* **303,** 325–349.
Kojiro M and Nakashima O. (1999). *Semin. Liver Dis.,* **19,** 287–296.
Lukas ER, Bartley SM, Graveel CR, Diaz ZM, Dyson N, Harlow E, Yamasaki L and Farnham PJ. (1999). *Mol. Carcin.,* **25,** 295–303.
Parkin DM, Pisani P and Ferlay J. (1999). *Int. J. Cancer,* **80,** 827–841.
Schultz J, Copley RR, Doerks T, Ponting CP and Bork P. (2000). *Nucleic Acids Res.,* **28,** 231–234.
Schultz J, Milpetz F, Bork P and Ponting CP. (1998). *Proc. Natl. Acad. Sci. USA,* **95,** 5857–5864.
Shijubo N, Uede T, Kon S, Maeda M, Segawa T, Imada A, Hirasawa M and Abe S. (1999). *Am. J. Respir. Crit. Care Med.,* **160,** 1269–1273.
Singhal H, Bautista DS, Tonkin KS, O'Malley FP, Tuck AB, Chambers AF and Harris JF. (1997). *Clin. Cancer Res.,* **3,** 605–611.

Slansky JE, Li Y, Kaelin WG and Farnham PJ. (1993). *Mol. Cell. Biol.,* [author's correction: **13,** 7201].

Tuck AB, Arsenault DM, O'Malley FP, Hota C, Ling MC, Wilson SM and Chambers AF. (1999). *Oncogene,* **18,** 4237 – 4246.

Tuck AB, O'Malley FP, Singhal H, Harris JF, Tonkin KS, Kerkvliet N, Saad Z, Doig GS and Chambers AF. (1998). *Int. J. Cancer,* **79,** 502 – 508.

Ue T, Yokozaki H, Kitadai Y, Yamamoto S, Yasui W, Ishikawa T and Tahara E. (1998). *Int. J. Cancer,* **79,** 127 – 132.

Welford SM, Gregg J, Chen E, Garrison D, Sorensen PH, Denny CT and Nelson SF. (1998). *Nucleic Acids Res.,* **26,** 3059 – 3065.

Yamada Y, Karasaki H, Matsushima K, Lee GH and Ogawa K. (1999). *Lab. Invest.,* **79,** 1059 – 1067.

Ye X, Lu L and Gill SS. (1997). *Biochem. Biophys. Res. Commun.,* **239,** 660 – 665.

Yoong KF, Afford SC, Jones R, Aujla P, Qin S, Price K, Hubscher SG and Adams DH. (1999). *Hepatology,* **30,** 100 – 111.