Modern Pathology (2008) 21, 505–516
© 2008 USCAP, Inc All rights reserved 0893-3952/08 $30.00
www.modernpathology.org

Gene expression profiling in hepatocellular carcinoma: upregulation of genes in amplified chromosome regions

Britta Skawran[1], Doris Steinemann[1], Anja Weigmann[1], Peer Flemming[2], Thomas Becker[3], Jakobus Flik[4], Hans Kreipe[2], Brigitte Schlegelberger[1] and Ludwig Wilkens[1,2]

[1]*Institute of Cell and Molecular Pathology, Hannover Medical School, Hannover, Germany;* [2]*Institute of Pathology, Hannover Medical School, Hannover, Germany;* [3]*Department of Visceral and Transplantation Surgery, Hannover Medical School, Hannover, Germany and* [4]*Institute of Virology, Hannover Medical School, Hannover, Germany*

Cytogenetics of hepatocellular carcinoma and adenoma have revealed gains of chromosome 1q as a significant differentiating factor. However, no studies are available comparing these amplification events with gene expression. Therefore, gene expression profiling was performed on tumours cytogenetically well characterized by array-based comparative genomic hybridisation. For this approach analysis was carried out on 24 hepatocellular carcinoma and 8 hepatocellular adenoma cytogenetically characterised by array-based comparative genomic hybridisation. Expression profiles of mRNA were determined using a genome-wide microarray containing 43 000 spots. Hierarchical clustering analysis branched all hepatocellular adenoma from hepatocellular carcinoma. Significance analysis of microarray demonstrated 722 dysregulated genes in hepatocellular carcinoma. Gene set enrichment analysis detected groups of upregulated genes located in chromosome bands 1q22–42 seen also as the most frequently gained regions by comparative genomic hybridisation. Comparison of significance analysis of microarray and gene set enrichment analysis narrowed down the number of dysregulated genes to 18, with 7 genes localised on 1q22 (*SCAMP3*, *IQGAP3*, *PYGO2*, *GPATC4*, *ASH1L*, *APOA1BP*, and *CCT3*). In hepatocellular adenoma 26 genes in bands 11p15, 11q12, and 12p13 were upregulated. However, the respective chromosome bands were not gained in hepatocellular adenoma. Expression analysis and comparative genomic hybridisation identified an upregulation of genes in amplified regions of 1q. These results may serve to further narrow down the number of candidate driver genes in hepatocarcinogenesis.
Modern Pathology (2008) **21,** 505–516; doi:10.1038/modpathol.3800998; published online 15 February 2008

Keywords: hepatocellular carcinoma; hepatocellular adenoma; gene expression; chromosomal instability

During recent years, large-scale gene expression profiling has provided important insights into the biology of hepatocellular carcinomas. The results provided new information on classification, aetiology, survival prediction, and identification of signalling pathways that could serve as therapeutic targets.[1–5] Functional genomics integrating comparison with genetically modified mice as models for human hepatocellular carcinoma[6] as well as data from promoter regions,[7] expression of non-coding genes, that is, microRNA,[8] or array-based comparative genomic hybridisation (aCGH)[9] has further increased the reliability and significance of the biological and clinical conclusions drawn from gene expression profiles. This will be the basis for developing new targeted therapies, an urgent need to reduce the mortality from hepatocellular carcinoma, which represents the fourth most common malignant tumour with more than one million patients affected worldwide per year, and usually has a very poor prognosis.[10]

Recently, we have shown that hepatocellular carcinoma is characterised by chromosomal instability.[11] In contrast, hepatocellular adenoma, a rare benign tumour occurring mainly in premenopausal women after oral contraceptive use,[12] shows a low rate of chromosomal imbalance. This qualifies hepatocellular adenoma as a comparison group since the surrounding tissue in livers harbouring

Correspondence: Dr L Wilkens, MD, Institute of Pathology, University of Bern, Murtenstrasse 31, Postfach 62, Bern CH 3010, Switzerland.
E-mail: ludwig.wilkens@pathology.unibe.ch
Received 6 February 2007; revised 14 May 2007; accepted 17 May 2007; published online 15 February 2008

hepatocellular carcinoma often contains significant alterations in gene expression and on the genomic level similar to those found in hepatocellular carcinoma.[13–15]

During dedifferentiation of hepatocellular carcinoma, chromosomal instability increases from differentiated G1 to undifferentiated G3 tumours. Recurrent chromosome aberrations of hepatocellular carcinoma are gains of 1q, 8q, 6q, and 20q and losses of 4q, 8p, 13q, and 17p. Gains of 1q are by far the most frequent aberrations and occur early during tumorigenesis.[11,16]

Here we aimed to verify recent data suggesting that hepatocellular adenoma can be separated from hepatocellular carcinoma based on gene expression profiles, and to identify candidate driver genes in gained or amplified regions, in particular in 1q. For this purpose, 32 hepatocellular adenoma and hepatocellular carcinoma, cytogenetically characterised by aCGH, were analysed for gene expression using a whole genome microarray. Alterations in gene expression were then set in correlation to genomic imbalances using novel bioinformatic models that enabled us to dissect the overexpressed genes in amplified regions.

Materials and methods

Materials

Analysis was carried out on tumour specimens of 24 patients suffering from hepatocellular carcinoma and of 8 patients with hepatocellular adenoma. All patients were treated at Hannover Medical School. Clinical, serological, and pathomorphological data are given in Table 1. Tumours were obtained as permitted by the Ethics Committee of the Hannover Medical School from surgical specimens taken for diagnostic purposes. Unfixed tissue samples were snap-frozen within 30 min after resection and stored at −80°C.

A piece of tissue was cut from the tumour for RNA extraction and another piece of tissue was fixed in formalin and was paraffin-embedded for additional histological examination. This ensured the closest proximity and homogeneity of the tissues taken for the different experiments.

For all tumours included in this study, cytogenetic characterisation was already carried out using aCGH with detailed information given in an earlier study.[11] The karyotypes obtained by this approach are listed in Table 2. In brief, the most commonly gained region was 1q22.1–23.1 in 22 out of 24 hepatocellular carcinoma (92%). With a decreasing number of tumours affected, a gain of 8q was observed in 14 hepatocellular carcinoma (58%), losses for 8p in 12 hepatocellular carcinoma (50%) losses for 4q in nine cases (38%), losses of 13q in eight cases (33%), losses of 6q in eight cases (33%), gains for 6p in eight cases (33%), losses of 17p in eight cases (33%), and gains for 20q in eight cases (33%). In the group of hepatocellular adenoma, sample N103 (13%) showed a gain of 1q22–q23.1 as also seen in hepatocellular carcinoma. In this case, adenomatosis hepatis was diagnosed because of more than 20 separate tumour nodules. In tumours N103 and N48, a loss of 8p was found. A gain of 20q13.33–20qter was observed in six out of eight cases (75%). Five hepatocellular adenoma revealed a gain of the subtle region 22q12.3–q13.1. These two regions were rarely affected in hepatocellular carcinoma.

Methods

cDNA Microarray Analysis

Total RNA was isolated using Trizol (Invitrogen, Karlsruhe, Germany) followed by RNeasy Mini Kit (Qiagen, Hilden, Germany) and linearly amplified (MessageAmp aRNA kit, Ambion, Huntingdon, UK). A measure of 1.5 μg of each of amplified RNA from the tumours and the human reference (Universal Human Reference RNA, Stratagene, La Jolla, CA, USA) were reverse transcriptase labelled with Cy3 (reference) and Cy5 (tumours) (Amersham Biosciences, Little Chalfont, UK). Labelled RNA of samples and references were comparatively hybridised overnight onto a cDNA chip containing the 'Resgen clone set' with more than 43 000 spots covering the entire human genome (Stanford Functional Genomics Facility, Stanford, CA, USA) according to the posted protocol (http://brownlab.Stanford.edu/protocols.html). Arrays were imaged using a GenePix 4000B scanner and dependent software (GenePix Pro 4.1, Axon Instruments, Ismaning, Germany). For normalisation of each array, data were imported to the Stanford Microarray Database (available at http://smd.stanford.edu). Data were excluded if the regression correlation was less than 0.6 or the mean spot intensity/median background intensity was less than 1.5. The well-measured genes were subsequently 'mean-centred' (i.e., reported for each gene relative to the mean ratio across all samples, rather than relative to the control spots).

Normalised and log-transformed expression data served as input for significance analysis of microarray with standard parameters given in the literature.[17]

For detection of modest but coordinate changes on gene sets that share common chromosomal location, gene set enrichment analysis was used based on the C1 and C2 databases (version 1.0, March 2005) and the software package developed by Subramanian *et al*.[18]

RT-PCR

To confirm the array results, quantitative RT-PCR was performed on the selected genes *CCT3*, *CABC1*, and *CDKN1B*, found upregulated in hepatocellular carcinoma in comparison to hepatocellular adenoma.

Table 1 Clinical, morphological, and serological data of the 32 patients analysed

Laboratory no.	*Age (years)*	*Gender*	*pTNM classification*	*Diameter (cm)*	*Histological grading*	*Hepatitis serology*	*ISHAK score*	*Additional findings*	*Alive … months after surgery*	*Dead … months after surgery*
31	33	M		7.5	HCA	Neg	No cirrhosis		53	
73	23	F		12	HCA	B	No cirrhosis		69	
11	29	F		0.7	HCA	B	No cirrhosis		59	
30	39	F		7	HCA	Neg	No cirrhosis		73	
104	21	M		8	HCA	Neg	No cirrhosis		65	
48	30	F		7	HCA	Neg	No cirrhosis		69	
70	27	F		11	HCA	Neg	No cirrhosis		66	
103	41	F		4.5?	HCA	Neg	No cirrhosis	Adenomatosis	40	
81	43	M	pT2, NX, MX	7	1	Neg	No cirrhosis		57	
25	54	M	pT2, N0, MX	1.5	1	B	A1B0C1D2F6		Dead	
20	68	F	pT2, NX, MX	5	1	Neg	NA		94	
77	53	F	pT3, NX, MX	18	1	Neg	No cirrhosis		59	
15	68	M	pT3, NX, M1	6.5	2	Neg	A2B0C1D2F6	Steatosis hepatis	58	
62	75	M	pT3, N0, MX	2.5	2	Neg	A2B0C1D1F	Steatosis hepatis	70	
80	65	M	pT3, NX, MX	9	2	B	No cirrhosis		54	
87	47	M	pT2, NX, MX	3.3	2	C	A3B0C1D3F6		54	
9	81	M	pT2, NX, MX	9	2	B	No cirrhosis	Steatosis hepatis	Dead	30
13	65	F	pT4, NX, MX	14	2	Neg	No cirrhosis	Steatosis hepatis	Dead	21
69	73	M	pT2, NX, MX	2.5	2	Neg	A2B0C1D2F4		70	
1	10	M	pT2, N0, MX	2.5	2	Neg	A1B0C0D1F6	Tyrosinaemia	66	
37	71	M	pT2, NX, MX	6	2	Neg	No cirrhosis		Dead	46
74	35	M	pT2, NX, MX	6	2	Neg	No cirrhosis		NA	
107	66	M	pT4, NX, MX	2.8	2	B	A2B0C1D1F6	RE-Tx	60	
8	63	M	pT2, NX, MX	9	2–3	C	A4B0C2D4F2		54	
45	73	F	pT2, NX, MX	10	2–3	C	A2B0C1F2		Dead	15
86	57	M	pT3, NX, MX	8	2–3	Neg	No cirrhosis	Steatosis hepatis	65	
47	65	M	pT3, NX, MX	8	2–3	B	A1B0C1D1F2	Steatosis hepatis	0	
71	59	M	pT3, NX, MX	17	2–3	B	No cirrhosis		Dead	16
91	37	M	pT2, NX, MX	15	2–3	B	A1B0C1D2F0		54	
90	9	M	pT3, NX, M1	3	2–3	B	No cirrhosis		Dead	5
40	50	M	pT3, N0, MX	11	3	B	No cirrhosis	Steatosis hepatis	40	
82	19	M	pT3, N0, MX	4	3	B	A1B0C1D1F6		55	

In a total volume of 120 μl, 1 μg of total RNA was reverse transcribed by M-MuLV reverse transcriptase using random hexamer primers (RevertAid First Strand cDNA synthesis kit; Fermentas, St-Leon-Rot, Germany). Two microliters of *CCT3* cDNA samples were amplified on an iCycler iQ real-time detection system (Bio-rad Laboratories, München, Germany) using a QuantiTect SYBR Green PCR kit (Qiagen, Hilden, Germany). Five microliters of each cDNA sample was used for the analysis of CABC1 and CDKN1B expression. This expression was performed using an iCycler iQ real-time detection system (Bio-rad Laboratories, München, Germany). Primers and probes for *CABC1* and *CDKN1B* were synthesised by TibMolBiol (Berlin, Gerrmany). Detailed PCR conditions and primer sequences for all genes used are shown in Table 3. β-Actin was used as internal control gene. For calibration the human reference (Stratagene, Amsterdam, The Netherlands) was used. Relative quantification of gene expression was determined according to the $\Delta\Delta C_T$ method using an iCycler iQ real-time detection system software version 3.1 (Bio-rad Laboratories). All samples were analysed in triplicates.

Results

Gene Expression Analysis of Hepatocellular Adenoma Compared to Hepatocellular Carcinoma

Evaluating microarray data from 24 hepatocellular carcinoma and eight hepatocellular adenoma, 7733 cDNA clones with the most variable expression among tumour and reference DNA were selected. Hierarchical clustering algorithm was applied based on the Pearson's correlation (centred) coefficients

to group the genes, as well as the samples, on the basis of similarity in their expression pattern. This analysis revealed two major clusters of gene expression pattern (Figure 1), one representing hepatocellular carcinoma tumours and the other representing hepatocellular adenoma tumours, with only one exception (N37).

We then used significance analysis of microarray to identify those genes significantly deregulated. By this approach, 171 genes were expressed at a higher level and 551 genes were expressed at a lower level in hepatocellular carcinoma in comparison to hepatocellular adenoma, as listed in detail in Supplementary Table 1. Comparison to the list of genes most significantly dysregulated in hepatocellular carcinoma as reported by Chen *et al*[19] revealed nine genes detected in both studies.

Table 2 List of the chromosome bands containing set of genes significantly upregulated ($P<0.01$) in HCC or in HCA (complete data are given in Supplementary Table 2)

Chromosomal region	*No. of genes*	*NOM p-val*	*FDR q-val*
Upregulated in HCC			
1q22	37	0.009	0.223
1q32	88	0.004	0.226
1q41	14	0.010	0.174
1q42	46	0.004	0.220
Yp11	14	0.005	0.344
Upregulated in HCA			
11p15	121	0.002	0.037
11q12	62	0.000	0.031
12p13	89	0.008	0.199

Chromosome bands found aberrant in dedifferentiated HCC were only excluded to avoid a systematic bias. In addition, the chromosome bands for the gonosomes are given as an internal control. Since HCC occur more frequently in men than in women, as seen also in this study, an upregulation of genes located on Y has to be expected and is detected here. Conversely, an upregulation of genes located on the X chromosome is not assumed due to the inactivation of one of the two of these chromosomes normally found in female cells. Accordingly, no upregulation was seen by gene set enrichment analysis.

Genes Commonly Upregulated in Amplified Regions

As the second mathematical approach, gene set enrichment analysis was applied. Gene set enrichment analysis evaluates microarray data not at the level of single genes but at the level of gene sets, that is, groups of genes that share a common biological function, regulation, or chromosomal location. As the basic parameter, sets of genes were analysed with regard to their localisation in the genome, that is, their chromosome band annotation.

First, upregulated genes were determined in hepatocellular carcinoma in comparison to hepatocellular adenoma. Second, downregulated genes in hepatocellular carcinoma were identified. Six sets of genes significantly upregulated in hepatocellular carcinoma ($P<0.01$) were localised in chromosome bands 1q22 (37 genes), 1q32 (88 genes), 1q41 (14 genes), 1q42 (46 genes), 2p22 (21 genes), and Yp11 (14 genes), respectively (Figure 2a, Supplementary Table 2). However, since chromosome band 2p22 was found altered in dedifferentiated but not in well-differentiated hepatocellular carcinoma, this gene set was excluded from further evaluation to avoid a systematic bias with regard to the dedifferentiation of hepatocellular carcinoma. We also excluded Yp11, because 6 out of 8 patients suffering from hepatocellular adenoma were females, whereas 21 out of 24 patients with hepatocellular carcinoma

Table 3 Sequences of primer, probes, and cycler conditions used for verification of array results by RT-PCR

Gene	*Primer forward*	*Primer reverse*	*Cycler conditions*		
Real-time RT-PCR with SYBR green					
β-Actin	GATATCGCCGC GCTCGTCGTC	GGCTGGGGTGTTG AAGGTCTC	95°C	10 min	
CCT3	CGGAAGTAGAA CGCAACCTC	TCAGGGTACGAGG AATGACC	95°C	15 s	40 ×
			60°C	30 s	
			72°C	30 s	
Real-time RT-PCR with probe					
β-Actin	AgCCTCgCCTTTgCCgA	CTggTgCCTggggCg	95°C	10 min	
CABC1	gCTCCgAggACCCCTCAg	CTgCATTggCCTCggACA	95°C	15 s	40 ×
CDKN1B	gAgAgCCAggATgTCAgCg	TTgTTTTgAgTAgA AgAATCgTCggT	60°C	1 min	

Gene	*Probe*
β-Actin	FAM-CCgCCgCCCg TCCACACCCgCC XT-p
CABC1	CCgTgCTgggTTC CAgTCCTTC
CDKN1B	FAM-CCTTTAATTggg gCTCCggCTAACT-TMR

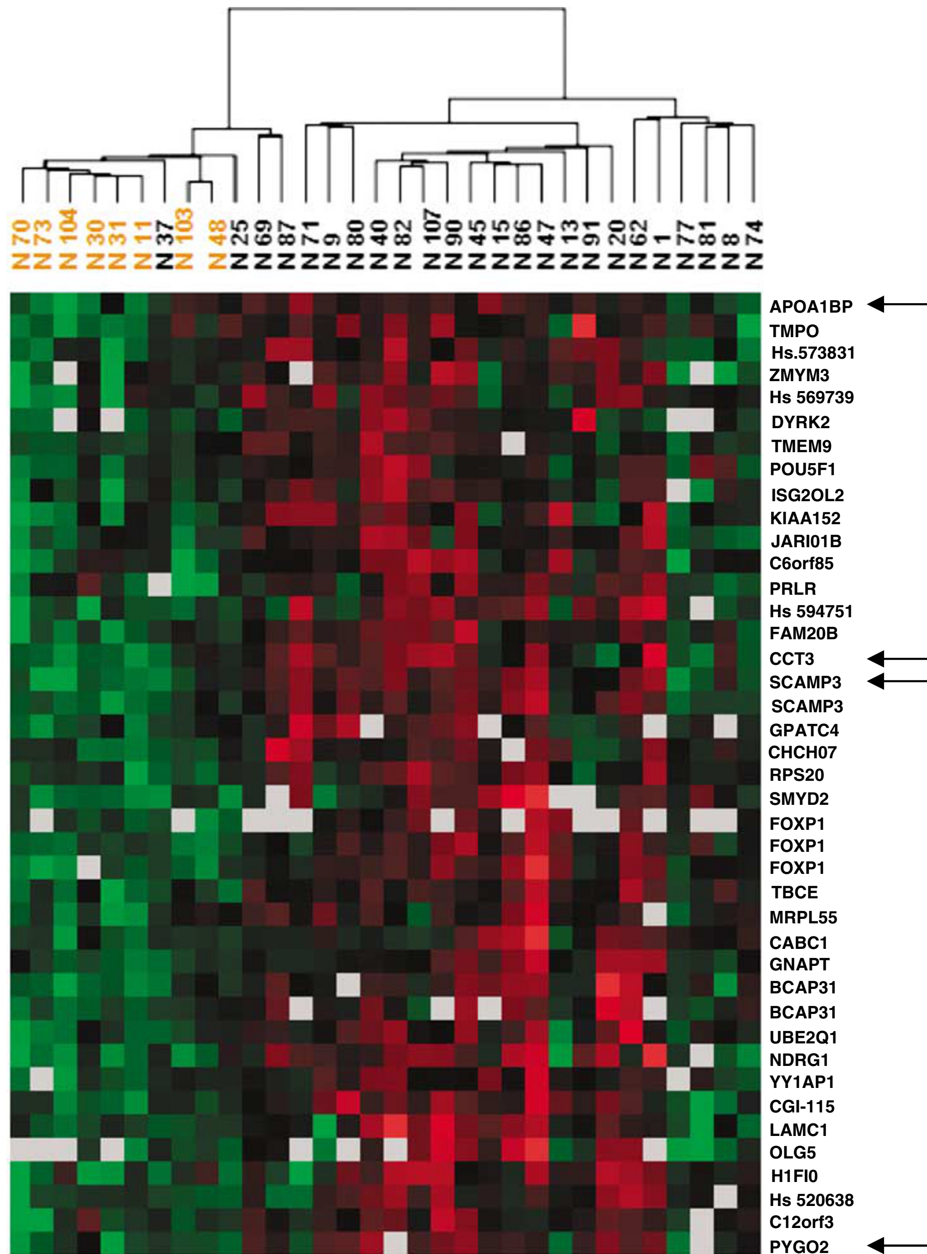


Figure 1 Subset of the heat map using hierarchical clustering based on significance analysis of microarray. Upregulated genes are marked in red, downregulated genes in green. As seen at the top, annotations of hepatocellular adenoma are printed in green, hepatocellular carcinoma in black. Hierarchical clustering including 7733 genes informative for this approach branched all hepatocellular adenoma from hepatocellular carcinoma with only one hepatocellular carcinoma sample (N37) occurring in the cluster of hepatocellular adenoma. Conversely, no hepatocellular adenoma was seen in the group of hepatocellular carcinoma. The detailed search for a clustered subset of genes includes four genes upregulated in hepatocellular carcinoma and localised in 1q22 (*CCT3*, *SCAMP3*, *ApoA1BP*, and *PYGO2*), indicating the importance of these genes in the definition of hepatocellular carcinoma by expression analysis.

were males, resulting in an expected bias. Nevertheless, this finding demonstrates the reliability of the technique. All genes dysregulated in the remaining chromosome bands are named in Table 4.

Moreover, 10 gene sets commonly downregulated in hepatocellular carcinoma were identified as significantly deregulated (complete set of data is given in Supplementary Table 2). Seven gene sets were again localised at chromosome bands affected only in dedifferentiated hepatocellular carcinoma and were therefore also excluded. The remaining three sets were localised in 11p15 (121 genes), 11q12 (62 genes), and 12p13 (83 genes), respectively (Figure 2b).

Comparison of the data obtained by gene set enrichment analysis and significance analysis of microarray analyses showed that seven genes (*SCAMP3*, *IQGAP3*, *PYGO2*, *GPATC4*, *ASH1L*, *APOA1BP*, and *CCT3*) located in 1q were found to be significantly upregulated in both mathematical

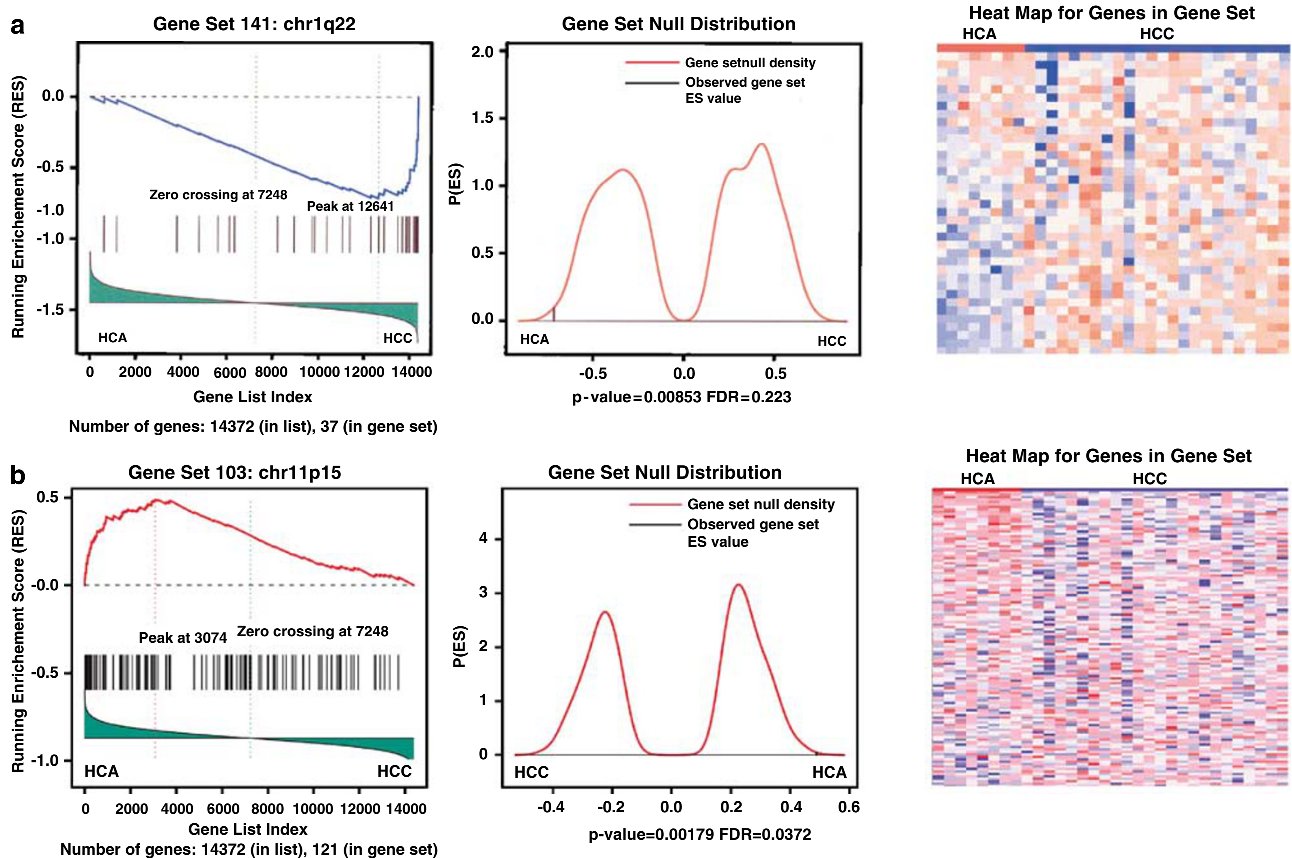


Figure 2 (**a**) Gene set enrichment analysis for comparison of hepatocellular carcinoma with hepatocellular adenoma. Four chromosome bands were found to contain highly significant ($P<0.01$) upregulated gene sets in hepatocellular carcinoma. All these four chromosome bands are localised in 1q (1q22 (shown here), 1q32, 1q41, and 1q42) known as most frequently gained in hepatocellular carcinoma with 1q22 aberrant in 92% of cases in aCGH. Visualisation of the results is carried out using a gene list index (left), null distribution graph (middle), and heat map (right) with downregulated genes marked in blue and upregulated genes marked in red. (**b**) Gene set enrichment analysis data obtained when comparing hepatocellular adenoma with hepatocellular carcinoma. Three bands (11p15 (shown here), 11q12, and 12p13) were found.

models (Table 5). For chromosome band 1q32, two out of three genes were found to be deregulated both in gene set enrichment analysis and significance analysis of microarray analyses. For chromosome band 1q41, 1 out of 1 gene, for 1q42 5 out of 6 genes, for 11p15 10 out of 10, for 11q15 8 out of 8 genes, and for 12p13 2 out of 8 genes were seen as deregulated in both mathematical models.

According to the literature (detailed information given in: http://smd.stanford.edu/cgi-bin/source/sourceResult), the majority of the genes, that is, 24 out of 45 genes, have a function in the induction of proliferation and cell cycle control (Table 6).

Comparison of mRNA Expression Analysis and aCGH

The upregulation of genes may be induced by copy number changes. Therefore, we investigated whether the gene sets significantly upregulated in hepatocellular carcinoma in comparison to hepatocellular adenoma are located in gained or amplified regions. This was true for chromosome bands 1q22, 1q32, 1q41, and 1q42 that were gained in 22 out of 24 hepatocellular carcinoma analysed by both gene expression and aCGH analyses. In chromosome band 11p15, significance analysis of microarray and gene set enrichment analysis revealed a set of upregulated genes. However, only 4 out of 22 hepatocellular carcinoma demonstrated a gain of this chromosomal region. The same is true for chromosome band 11q12, where no chromosomal imbalances were found. For 12p12–13, three hepatocellular carcinoma showed a loss, whereas three hepatocellular carcinoma showed a gain of at least part of this region.

RT-PCR

Results obtained for mRNA expression analysis were proven exemplarily for genes *CCT3*, *CABC1*, and *CDKN1*, respectively. *CCT3* is localised on 1q22, *CABC1* on 1q42, and *CDKN1* on 12p13. RT-PCR confirmed the results of array analysis revealing a systematic increase of mRNA expression of *CCT3*

Table 4 List of chromosome bands and their annoted gene sets found upregulated most significantly by gene set enrichment analysis in hepatocellular carcinoma and adenoma, respectively

Upregulated in hepatocellular carcinoma

1q22	*1q32*		*1q41*	*1q42*
RHBG	*PLXNA2*	*RABIF*	*TAF1A*	*GNPAT*
SHC1	*NEK2*	*PTPN7*	*GPATC2*	*TBCE*
PYGO2	*ZNF281*	*TRAF5*	*DISP1*	*ARID4B*
SCAMP3	*RIPK5*	*LRRN5*	*KCTD3*	*CABC1*
C1 or f2	*MCP*	*PPFIA4*	*SMYD2*	*GGPS1*
GPATC4	*JARID1B*	*IL19*	*LYPLAL1*	*PARP1*
HCN3	*TMEM9*	*MYBPH*	*EPRS*	*TAF5L*
ASH1L	*PPP2R5A*	*PPP1R12B*	*KCNK2*	*LBR*
KRTCAP2	*RNPEP*	*TIMM17A*	*HLX1*	*GNG4*
CCT3	*RBBP5*	*SERTAD4*	*PTPN14*	*WDR26*
RUSC1	*MDM4*	*CHI3L1*	*PROX1*	*TRIM11*
MTX1	*LPGAT1*	*KISS1*	*MARK1*	*ARV1*
APOA1BP	*ADIPOR1*	*ARL10B*	*USH2A*	*TARBP1*
IQGAP3	*SOX13*	*TNNT2*	*DUSP10*	*ABCB10*
PMVK	*C1 or f48*	*BTG2*		*NUP133*
GBA	*NR5A2*	*SLC30A1*		*TOMM20*
PKLR	*NQO3A2*	*ADORA1*		*COG2*
FDPS	*HHAT*	*IKBKE*		*RAB4A*
DAP3	*RAB7L1*	*SLC41A1*		*KCNK1*
EFNA1	*DAF*	*SRGAP2*		*GUK1*
SSR2	*MYOG*	*ATP2B4*		*CDC42BPA*
THBS3	*CHIT1*	*DYRK3*		*TTC13*
ARHGEF2	*PFKFB2*	*C4BPA*		*ACBD3*
CLK2	*C1 or f37*	*PIGR*		*SLC35F3*
RIT1	*CSRP1*	*PKP1*		*PSEN2*
BGLAP	*C4BPB*	*CD34*		*ENAH*
MEF2D	*CR1L*	*LGR6*		*PGBD5*
ADAM15	*PPP1R15B*	*MAPKAPK2*		*TP53BP2*
PMF1	*PHLDA3*	*SYT14*		*EPHX1*
LMNA	*IPO9*	*RASSF5*		*EGLN1*
EFNA4	*PCANAP6*	*KLHL12*		*CAPN9*
SYT11	*LGTN*	*TNNI1*		*NID*
PBXIP1	*RPS6KC1*	*LMOD1*		*DISC1*
RAB25	*ATF3*	*ETNK2*		*FBXO28*
KCNN3	*RCOR3*	*CACNA1S*		*C1 or f35*
SEMA4A	*KIF21B*	*IRF6*		*PCNXL2*
ZFP67	*KIF14*	*HSD11B1*		*NVL*
	LAD1	*IL24*		*KIAA0133*
	PIK3C2B	*CNTN2*		*AGT*
	LAMB3	*CR2*		*TSNAX*
	C1 or f40	*FMOD*		*SIPA1L2*
	PLEKHA6	*REN*		*CAPN2*
	ELK4	*PRELP*		*ITPKB*
	NAV1	*CTSE*		*OBSCN*
				TM7SF1
				GALNT2

Upregulated in hepatocellular adenoma

12p13		*11q12*		*11p15*		
CD163	*ACRBP*	*MS4A6E*	*C11 or f31*	*XLKD1*	*CARS*	*PDE3B*
C3AR1	*JARID1A*	*MS4A6A*	*B3GAT3*	*LDHA*	*RNF141*	*TRIM6*
CLECSF2	*SLC2A3*	*MS4A7*	*ASRGL1*	*H19*	*TSSC4*	*TNNT3*
C1RL	*KCNA5*	*UBE2L6*	*TAF6L*	*SAA4*	*PLEKHA7*	*MUC5B*
CD4	*PRR4*	*YPEL4*	*TCN1*	*SAA2*	*DCHS1*	*FXC1*
CDKN1B	*TAS2R14*	*HRASLS3*	*STX5A*	*SAA1*	*AP2A2*	*TALDO1*
PZP	*GNB3*	*RARRES3*	*SSRP1*	*IFITM1*	*MRPL17*	*STIM1*
KLRG1	*ETV6*	*MS4A4A*	*TNKS1BP1*	*ADM*	*HRAS*	*ARFIP2*
GABARAPL1	*LEPREL2*	*ROM1*	*SCGB2A2*	*ZDHHC13*	*CALCB*	*TOLLIP*
CACNA1C	*VWF*	*SYT7*	*DDB1*	*TRIM22*	*PTH*	*ILK*
KLRB1	*PHC1*	*SLC15A3*	*C11 or f9*	*ARNTL*	*KCNQ1DN*	*PIK3C2A*
RPS27	*SLC2A14*	*SLC22A9*	*SLC22A8*	*SMPD1*	*NUCB2*	*EPS8L2*
HEBP1	*GDF3*	*LPXN*	*ZDHHC5*	*IRF7*	*COPB*	*INS*
PTMS	*CD9*	*RAB3IL1*	*OSBP*	*HPS5*	*CALCA*	*USH1C*

Table 4 Continued

Upregulated in hepatocellular adenoma						
12p13		*11q12*		*11p15*		
SCNN1A	PEX5	CD6	FADS1	OSBPL5	TEAD1	CHRNA10
TNFRSF7	WNT5B	SLC43A3	NXF1	IFITM3	STK33	EIF4G2
PRH1	DYRK4	CTNND1	PRP19	TM4SF7	WEE1	ABCC8
TNFRSF1A	GRIN2B	SERPING1	SLC3A2	PSMA1	SLC25A22	RPL27A
PTPN6	DUSP16	AGTRL1	RBM21	IFITM2	MLSTD2	TPH1
M6PR	LRP6	GLYAT	FEN1	ZNF143	TNNI2	PHLDA2
TEAD4	LTBR	MS4A1	EEF1G	SOX6	C11 or f13	ZNF214
FBXL14	RAD51AP1	STX3A	BSCL2	PARVA	NRIP3	DKK3
C1S	FOXJ2	DTX4		UBQLN3	DUSP8	TAF10
APOBEC1	LOH12CR1	SCGB2A1		LMO1	NELL1	CCKBR
MFAP5	GSG1	MRPL16		PSMD13	USP47	C11 or f17
AICDA	NDUFA9	CD5		LDHC	NUP98	SLC6A5
CREBL2	ENO2	PGA5		TH	KCNC1	CSRP3
EMP1	GUCY2C	SLC22A6		ODF3	TRIM3	SH2BP1
KLRD1	ATF7IP	VMD2		HPX	SLC22A1LS	EIF3S5
A2M	SLC6A13	SLC43A1		IGF2	NAV2	GTF2H1
SLC6A12	C12orf4	INCENP		CTSD	PHEMX	MRPL23
KLRC1	GPR19	MS4A8B		LSP1	RNH	CDKN1C
CD69	RAD52	FADS2		TSG101	CD151	TRIM66
CLECSF6	BCL2L14	MED19		ZNF195	PNPLA2	HTATIP2
DCP1B	FOXM1	PRG2		TUB	APBB1	RRM1
RBP5	DRPLA	CHRM1		HRMT1L3	SLC22A18	PKP3
TMEM16B	OLR1	MS4A2		DEAF1	C11 or f15	LRDD
CCND2	RAB6IP2	MS4A5		SCUBE2	MYOD1	ST5
AKAP3	FKBP4	FABP5		PRKCDBP	AMPD3	MRVI1
MLF2	USP5	SCGB1A1		PPFIBP2	HBE1	MUC6
	CHD4					BET1L
	KCNA1					
	NINJ2					
	NOL1					
	CDCA3					
	TULP3					
	CLSTN3					
	VAMP1					
	ZNF384					

Table 5 Comparison of gene set enrichment analysis and significance analysis of microarray results

Chromosome band						
1q22	*1q32*	*1q41*	*1q42*	*11p15*	*11q12*	*12p13*
SCAMP3	HSPC150	SMYD2	LBR	XLKD1	MS4A7	C1S
IQGAP3	JARID1B		GNPAT	ARNTL	UBE2L6	CDKN1B
PYGO2	NEK2		MRPL55	TRIM22	MS4A4A	KLRK1
GPATC4			CABC1	SAA1	MS4A6A	CD163
ASH1L			ARID4B	SAA4	DTX4	C3AR1
CCT3			TBCE	LDHA	MS4A6E	C1RL
APOA1BP				ADM	HRASLS3	CLEC2
				SMPD1	SERPING1	CD4
				IRF7		
				IFITM1		

Listed are the genes found upregulated by significance analysis of microarray in regard to the chromosome bands seen upregulated by gene set enrichment analysis. Thirty-nine out of forty-three genes dysregulated by significance analysis of microarray were also aberrant in gene set enrichment analysis. Four genes were found by significance analysis of microarray only, and are set in grey. For 11q21 no genes were found by significance analysis of microarray.

and *CABC1*, and a decrease for *CDKN1* as also found by significance analysis of microarray as well as gene set enrichment analysis, as summarised graphically in Figure 3 (raw data given in Supplementary Table 3).

Discussion

Array-based gene expression analysis has been established as a highly informative technique in hepatocellular carcinoma. Recent studies have

Table 6 Genes found dysregulated in SAM and GSEA

Symbol	*Cytoband*	*Gene ontology annotations*	*q-value (%)*	*Log 2 ratio*
SCAMP3	1q22	Protein transport	1.57	1.39
IQGAP3	1q22	Ras GTPase activator activity, signal transduction	2.23	1.47
PYGO2	1q22	Wnt receptor signaling, pathway, regulation of transcription	1.57	1.39
GPATC4	1q22	Nucleic acid binding	3.30	1.01
ASH1L	1q22	Regulation of transcription, cell-cell signalling, DNA packaging	3.90	0.96
CCT3	1q22	ATP binding, protein folding	4.35	1.23
APOA1BP	1q22	Molecular function unknown	5.46	0.68
NEK2	1q32	Regulation of mitosis, may also play a role in meiosis, may have a role at the g2-m transition	3.12	1.59
JARID1B	1q32	Tumor-suppressive function partially mediated by pRb modulation, regulation of transcription	5.46	0.95
HSPC150	1q32	Ubiquitination	1.57	1.46
SMYD2	1q41	Molecular function unknown	5.46	1.61
LBR	1q42	Anchors the lamina and the heterochromatin to the inner nuclear membrane	2.59	1.03
GNPAT	1q42	Glycerone-phosphate O-acyl-transferase activity	5.46	0.99
MRPL55	1q42	Structural constituent of ribosome	3.12	0.93
CABC1	1q42	Mediating p53-inducible apoptosis, unfolded protein binding (Mitochondrion)	5.46	1.00
TBCE	1q42	Protein amino acid glycosylation	1.85	1.08
ARID4B	1q42	Chromatin binding	2.59	0.76
CD14	5q22	Apoptosis, inflammatory response	2.08	−0.77
CDO1	5q22	Oxidoreductase activity, L-cysteine metabolism	4.35	−0.71
XLKD1	11p15	Morphogenesis, cell motility	0.37	−1.43
ARNTL	11p15	Regulation of transcription	1.85	−0.88
TRIM22	11p15	Regulation of transcription, immune response	3.90	−0.58
SAA1	11p15	Tricarboxylic acid cycle intermediate metabolism, glycolysis	0.31	−1.67
ADM	11p15	Cell-cell signalling, signal transduction	1.01	−0.88
IRF7	11p15	Regulation of transcription, immune response, inflammatory response	0.00	−1.70
IFITM1	11p15	Negative regulation of cell proliferation, immune response	0.00	−1.37
SMPD1	11p15	Carbohydrate metabolism, signal transduction, neurogenesis	1.57	−0.93
LDHA	11p15	Tricarboxylic acid cycle intermediate metabolism, anaerobic glycolysis	0.78	−0.77
SAA4	11p15	Regulation of protein secretion, positive regulation of cell adhesion, negative regulation of inflammatory response	0.00	−1.62
MS4A7	11q12	Receptor complex involved in signal transduction	0.00	−1.62
UBE2L6	11q12	Ubiquitin cycle	0.00	−1.27
MS4A4A	11q12	Receptor activity, signal transduction	2.08	−1.09
MS4A6A	11q12	Receptor activity, signal transduction	0.00	−1.34
DTX4	11q12	Protein ubiquitination	1.85	−1.02
HRASLS3	11q12	tumor suppressor, negative regulation of cell cycle	3.30	−0.98
SERPING1	11q12	Immune response, complement activation, blood coagulation	0.78	−0.84
C1S	12p13	Proteolysis and peptidolysis, complement component, immune response	5.46	−0.68
CDKN1B	12p13	Involved in g1 arrest, regulation of cyclin dependent protein kinase activity	4.35	−0.52
CD163	12p13	Antimicrobial humoral response (sensu Vertebrata)	0.00	−1.26
C3AR1	12p13	Cell motility, signal transduction, receptor activity	0.00	−1.10
C1RL	12p13	Intracellular transport of retinol	0.78	−1.05
CD4	12p13	Gluconeogenesis	0.00	−1.24
CLEC2	12p13	Cell surface receptor linked signal transduction, sugar binding	3.90	−1.07

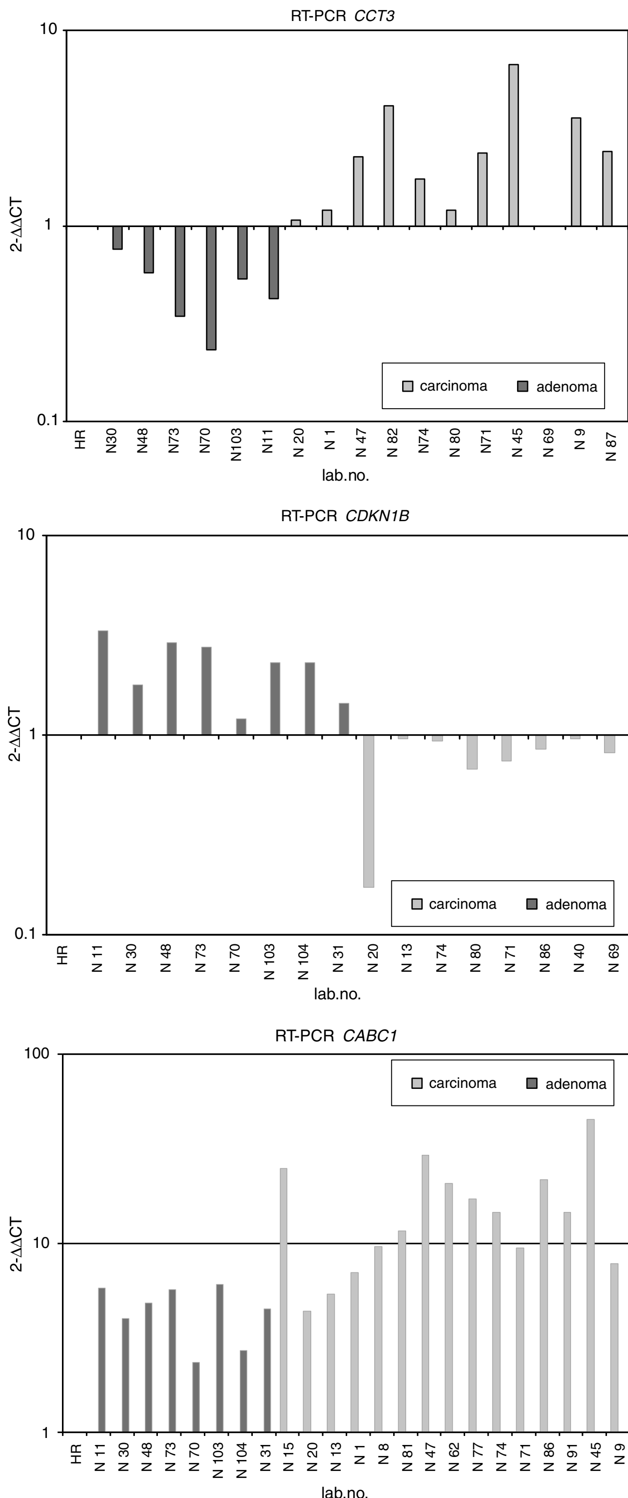
Gene symbol, chromosome band, short description of function, *q*-values, and log 2 ratio are given. Genes linked to signal transduction and/or cell cycle regulation are set in grey.

focused on the correlation of mRNA expression with aetiology, clinicopathological findings, grading, and survival.[20–23] However, mRNA expression analysis is hampered, since in livers bearing hepatocellular carcinoma the non-neoplastic tissues already reveal distinct dysregulated sets of genes as also found in hepatocellular carcinoma.[24–26] We therefore focused on the comparison of hepatocellular carcinoma and hepatocellular adenoma, a rare tumour, almost never reported as becoming malignant.

Until now, global gene expression analysis in hepatocellular adenoma has only been described in a study by Chen *et al*.[27] Their attempt was to define a set of diagnostically dysregulated genes separating hepatocellular adenoma from well-differentiated hepatocellular carcinoma. When comparing the results of Chen *et al* and to those of our study, in the former study, five genes were downregulated in hepatocellular carcinoma and 39 upregulated, whereas in our study, 171 genes were upregulated and 551 genes upregulated in hepatocellular adenoma. This difference in the detection of significantly altered genes is a phenomenon also reported in other studies and has to be seen together with variant mathematical models used, different settings of patients examined, and variant experimental settings.[28] Sample sets in particular are not directly comparable since our pool of hepatocellular carcinoma includes not only well differentiated but also poorly differentiated tu-

mours. Furthermore, a sample size of six hepatocellular adenoma *vs* eight well-differentiated hepatocellular carcinoma is borderline to achieve statistically significant results in view of single genes.

Nevertheless, at least nine genes were dysregulated concordantly in both studies, including genes for metabolism (*ALDH2*, *ADK*, *SLC7A2*, and *CYP4V2*), hormonal receptor (*ESR1*), complement activation (*C8A*), DNA repair (*PCNA*), and cell signalling (*PTPN3*), which is important in cell cycle regulation.



As a further mathematical attempt to better understand the overwhelming data obtained, gene set enrichment analysis was performed. This statistical model was chosen as a second approach to detect coordinate changes on gene sets defined as groups of genes that share common biological function, regulation, or chromosomal location.

In our study, we focused on the evaluation of gene sets defined by their location on distinct chromosome bands because of the known recurrent structural chromosome imbalances reported for hepatocellular carcinoma, but not hepatocellular adenoma, respectively. When performing gene set enrichment analysis, to look for gene sets upregulated in hepatocellular adenoma in comparison to hepatocellular carcinoma, 10 sets were found with significant *P*-values <0.01. However, evaluation was impaired, since seven of these chromosome bands were found altered in dedifferentiated hepatocellular carcinoma, but not in well-differentiated hepatocellular carcinoma. To avoid a systematic bias, these gene sets were excluded from further analysis.

Four sets of upregulated genes are located in 1q. The remaining three gene sets are located in chromosome bands found aberrant by aCGH in only a minority of cases. Therefore, it seems unlikely to assume that gene copy number changes have induced the overexpression of these genes as shown for 1q. The gene sets in 1q contained 37, 88, 14, and 46 genes within chromosome bands 1q22, 1q32, 1q41, and 1q42, respectively. Reliability of these results was underscored by comparing gene set enrichment analysis and significance analysis of microarray. Thirty-nine out of 44 (89%) upregulated genes identified by significance analysis of microarray were also detected by gene set enrichment analysis. Most interestingly, a gain of 1q has been reported as the most frequent imbalance (for review see progentix.com[29]) and is known as the chromosomal imbalances separating hepatocellular carcinoma from hepatocellular adenoma most significantly.[16]

In particular, chromosome band 1q22 was described as the commonly gained region in hepatocellular carcinoma. Within this small region eight

Figure 3 Results obtained by global expression analysis were proven by RT-PCR exemplarily for *CCT3* and *CABC1*, two of the genes found upregulated in hepatocellular carcinoma by significance analysis of microarray as well as gene set enrichment analysis. Experiments were performed for those cases for which exactly the same lot of mRNA was available for RT-PCR as also used in array-based experiments. This was carried out to avoid alterations of the results due to differences based in the laboratory work. The increased gene expression on the mRNA level for these two genes was also detectable by RT-PCR, as demonstrated by significantly higher $\Delta\Delta C_T$ values in hepatocellular carcinoma compared to hepatocellular adenoma. Conversely, *CDKN1B* was seen downregulated in hepatocellular carcinoma, as expected after array-based expression analysis.

genes were detected in both mathematical approaches, significance analysis of microarray and gene set enrichment analysis, respectively. This group of genes includes *SCAMP3*, *IQGAP3*, *PYGO2*, *GPATC4*, *ASH1L*, *APOA1BP*, and *CCT3*. Most genes play a role in proliferation and cell cycle control. Members of the IQGAP proteins are integral components of cytoskeletal regulation.[30] They are known to act on Ca(2+) calmodulin signalling, cytoskeletal architecture, CDC42 and Rac signalling. Furthermore, they are part of E-cadherin-mediated cell–cell adhesion and β-catenin-mediated transcription control.[31] Hepatocellular carcinoma A2 contains two Src homology 3 (SH3) domains indicating a function in intracellular signal transduction. *PYGO2* is a member of the canonical Wnt pathway and after recruitment to nuclear β-catenin permits transcriptional activation of WNT target genes.[32] The function of secretory carrier membrane proteins (*SCAMP3*) has yet not been characterised in detail. Singleton *et al*[33] found a ubiquitous distribution in nearly all tissues examined. They discuss the possible function of this gene in trafficking processes, probably of the post-Golgi apparatus. *CCT3* is a chaperone playing an important role in the folding of cytoskeletal components.[34] Elevated *CCT3* expression possibly impairs correct folding and assembly of complex proteins. For *CCT3*, significant overexpression in hepatocellular carcinoma has been reported by Wong *et al*.[35]

In conclusion, gene expression profiling suggests a coordinated upregulation of genes localised in amplified chromosome regions, in particular of 1q. This underlines the importance of this genomic region for the development of hepatocellular carcinoma regardless of the aetiological factors as already recognised on the genomic level. Therefore, further attempts should be made to narrow down this number of candidate driver genes in carcinogenesis of hepatocellular carcinoma.

Acknowledgement

This work was supported by the Deutsche Forschungsgemeinschaft (Grant KFO 119); Deutsche Krebshilfe (Grant 106200).

References

1 Wu CG, Salvay DM, Forgues M, *et al.* Distinctive gene expression profiles associated with Hepatitis B virus x protein. Oncogene 2001;20:3674–3682.
2 Smith MW, Yue ZN, Geiss GK, *et al.* Identification of novel tumor markers in hepatitis C virus-associated hepatocellular carcinoma. Cancer Res 2003;63: 859–864.
3 Okabe H, Satoh S, Kato T, *et al.* Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. Cancer Res 2001;61:2129–2137.
4 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 2004;24:565–574.
5 Neo SY, Leow CK, Vega VB, *et al.* Identification of discriminators of hepatoma by gene expression profiling using a minimal dataset approach. Hepatology 2004;39:944–953.
6 Farazi PA, Glickman J, Jiang S, *et al.* Differential impact of telomere dysfunction on initiation and progression of hepatocellular carcinoma. Cancer Res 2003;63:5021–5027.
7 Calvisi DF, Ladu S, Gorden A, *et al.* Ubiquitous activation of Ras and Jak/Stat pathways in human HCC. Gastroenterology 2006;130:1117–1128.
8 Murakami Y, Yasuda T, Saigo K, *et al.* Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. Oncogene 2006;25:2537–2545.
9 Patil MA, Gutgemann I, Zhang J, *et al.* Array-based comparative genomic hybridization reveals recurrent chromosomal aberrations and JAB1 as potential target for 8q gain in hepatocellular carcinoma. Carcinogenesis 2005;26:2050–2057.
10 Schirmacher P, Dienes HP. Hepatocellular carcinoma. In: Kurzrock R, Talpaz M (eds). Molecular Biology in Cancer Medicine. Dunitz: London, 1999, pp 355–366.
11 Steinemann D, Skawran B, Becker T, *et al.* Assessment of differentiation and progression of hepatic tumors using array-based comparative genomic hybridization. Clin Gastroenterol Hepatol 2006;4:1283–1291.
12 Ishak KG, Goodman ZD, Stocker JT. Tumors of the Liver and Intrahepatic Bile Ducts. Armed Forces Institute of Pathology: Washington, DC, USA, 2001.
13 Kim JW, Wang XW. Gene expression profiling of preneoplastic liver disease and liver cancer: a new era for improved early detection and treatment of these deadly diseases? Carcinogenesis 2003;24: 363–369.
14 Kim JW, Ye Q, Forgues M, *et al.* Cancer-associated molecular signature in the tissue samples of patients with cirrhosis. Hepatology 2004;39:518–527.
15 Anders RA, Yerian LM, Tretiakova M, *et al.* cDNA microarray analysis of macroregenerative and dysplastic nodules in end-stage hepatitis C virus-induced cirrhosis. Am J Pathol 2003;162:991–1000.
16 Wilkens L, Bredt M, Flemming P, *et al.* Differentiation of liver cell adenomas from well-differentiated hepatocellular carcinomas by comparative genomic hybridization. J Pathol 2001;193:476–482.
17 Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 2001;98: 5116–5121.
18 Subramanian A, Tamayo P, Mootha VK, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545–15550.
19 Chen ZM, Crone KG, Watson MA, *et al.* Identification of a unique gene expression signature that differentiates hepatocellular adenoma from well-differentiated hepatocellular carcinoma. Am J Surg Pathol 2005;29: 1600–1608.
20 Neo SY, Leow CK, Vega VB, *et al.* Identification of discriminators of hepatoma by gene expression profil-

ing using a minimal dataset approach. Hepatology 2004;39:944–953.
21 Wu CG, Salvay DM, Forgues M, *et al.* Distinctive gene expression profiles associated with Hepatitis B virus x protein. Oncogene 2001;20:3674–3682.
22 Smith MW, Yue ZN, Geiss GK, *et al.* Identification of novel tumor markers in hepatitis C virus-associated hepatocellular carcinoma. Cancer Res 2003;63:859–864.
23 Shirota Y, Kaneko S, Honda M, *et al.* Identification of differentially expressed genes in hepatocellular carcinoma with cDNA microarrays. Hepatology 2001;33:832–840.
24 Kim JW, Wang XW. Gene expression profiling of preneoplastic liver disease and liver cancer: a new era for improved early detection and treatment of these deadly diseases? Carcinogenesis 2003;24:363–369.
25 Anders RA, Yerian LM, Tretiakova M, *et al.* cDNA microarray analysis of macroregenerative and dysplastic nodules in end-stage hepatitis C virus-induced cirrhosis. Am J Pathol 2003;162:991–1000.
26 Kim JW, Ye Q, Forgues M, *et al.* Cancer-associated molecular signature in the tissue samples of patients with cirrhosis. Hepatology 2004;39:518–527.
27 Chen ZM, Crone KG, Watson MA, *et al.* Identification of a unique gene expression signature that differentiates hepatocellular adenoma from well-differentiated hepatocellular carcinoma. Am J Surg Pathol 2005;29:1600–1608.
28 Subramanian A, Tamayo P, Mootha VK, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545–15550.
29 Moinzadeh P, Breuhahn K, Stutzer H, *et al.* Chromosome alterations in human hepatocellular carcinomas correlate with aetiology and histological grade—results of an explorative CGH meta-analysis. Br J Cancer 2005; 92:935–941.
30 Briggs MW, Sacks DB. IQGAP proteins are integral components of cytoskeletal regulation. EMBO Rep 2003;4:571–574.
31 McCallum SJ, Wu WJ, Cerione RA. Identification of a putative effector for Cdc42Hs with high sequence similarity to the RasGAP-related protein IQGAP1 and a Cdc42Hs binding partner with similarity to IQGAP2. J Biol Chem 1996;271:21732–21737.
32 Kramps T, Peter O, Brunner E, *et al.* Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. Cell 2002;109:47–60.
33 Singleton DR, Wu TT, Castle JD. Three mammalian SCAMPs (secretory carrier membrane proteins) are highly related products of distinct genes having similar subcellular distributions. J Cell Sci 1997;110: 2099–2107.
34 Roobol A, Holmes FE, Hayes NV, *et al.* Cytoplasmic chaperonin complexes enter neurites developing *in vitro* and differ in subunit composition within single cells. J Cell Sci 1995;108:1477–1488.
35 Wong N, Chan A, Lee SW, *et al.* Positional mapping for amplified DNA sequences on 1q21-q22 in hepatocellular carcinoma indicates candidate genes over-expression. J Hepatol 2003;38:298–306.

Supplementary Information accompanies the paper on Modern Pathology website (http://www.nature.com/modpathol)