A Seven-Gene Set Associated with Chronic Hypoxia of Prognostic Importance in Hepatocellular Carcinoma

Hannah van Malenstein[1], Olivier Gevaert[3,6], Louis Libbrecht[1,5], Anneleen Daemen[3], Joke Allemeersch[4], Frederik Nevens[1], Eric Van Cutsem[2], David Cassiman[1], Bart De Moor[3], Chris Verslype[1,2], and Jos van Pelt[1]

Abstract

Purpose: Hepatocellular carcinomas (HCC) have an unpredictable clinical course, and molecular classification could provide better insights into prognosis and patient-directed therapy. We hypothesized that in HCC, certain microenvironmental regions exist with a characteristic gene expression related to chronic hypoxia which would induce aggressive behavior.

Experimental Design: We determined the gene expression pattern for human HepG2 liver cells under chronic hypoxia by microarray analysis. Differentially expressed genes were selected and their clinical values were assessed. In our hypothesis-driven analysis, we included available independent microarray studies of patients with HCC in one single analysis. Three microarray studies encompassing 272 patients were used as training sets to determine a minimal prognostic gene set, and one recent study of 91 patients was used for validation.

Results: Using computational methods, we identified seven genes (out of 3,592 differentially expressed under chronic hypoxia) that showed correlation with poor prognostic indicators in all three training sets (65/139/73 patients) and this was validated in a fourth data set (91 patients). Retrospectively, the seven-gene set was associated with poor survival (hazard ratio, 1.39; $P = 0.007$) and early recurrence (hazard ratio, 2.92; $P = 0.007$) in 135 patients. Moreover, using a hypoxia score based on this seven-gene set, we found that patients with a score of >0.35 ($n = 42$) had a median survival of 307 days, whereas patients with a score of ≤0.35 ($n = 93$) had a median survival of 1,602 days ($P = 0.005$).

Conclusions: We identified a unique, liver-specific, seven-gene signature associated with chronic hypoxia that correlates with poor prognosis in HCCs. *Clin Cancer Res; 16(16); 4278–88. ©2010 AACR.*

Hepatocellular carcinoma (HCC) is the sixth most common malignancy in the world and the third most common cause of cancer-related deaths (1). Every year, 600,000 new cases are diagnosed and almost just as many patients die annually of this disease. The most important risk factor for the development of HCC is cirrhosis, which is present in 80% of patients.

HCC generally develops over many years and unless the patients at risk are frequently screened, the disease is often diagnosed at a later stage (2). Only a small proportion of patients are eligible for surgical resection or transplantation; the majority have to rely on systemic chemotherapy, embolization, or supportive regimens (3). Patients with HCC have a heterogeneous course of disease. Therefore, therapeutic decisions for a specific patient would greatly benefit from a reliable prognostic indicator considering the treatment burden, duration, expected success rate, and costs.

Several studies have tried to identify gene sets with prognostic or diagnostic relevance by using microarray analysis. Each study resulted in its own classification with a specific separation into groups (4). All these microarray studies show remarkably little overlap and it is difficult to find a clear correlation between the molecular classes and prognosis. The results of these studies seem to be center-dependent because of the different microarray techniques used, the small heterogeneous cohorts that were studied, and the different clinical variables used for the evaluations (5).

Although it has been recognized that the microenvironment plays a role in tumor biology, it has never been used as a rational concept to study HCC. One of the factors that seems to affect cancer cell behavior and patient prognosis is hypoxia (6). Hypoxia is associated with poor prognosis in solid tumors and with the development of resistance to chemotherapeutic agents and radiation (7, 8). Although HCC is a hypervascular malignancy, there are regions with hypoxia (9). Hypoxic regions are already present in the early stage, when the vasculature is not sufficiently extended, and in more advanced stages, when rapid cell proliferation induces hypoxia (10). Furthermore, the neovascularization

American Association for Cancer Research

Translational Relevance

Hepatocellular carcinoma (HCC) has an increasing incidence and high mortality. Prognosis is difficult to predict for individual patients. We started with an *in vitro* model of chronic hypoxia and determined differentially expressed genes. The clinical value of these chronic hypoxia–related genes was assessed using bioinformatics in four data sets with HCC patients. We identified a very small gene set with important prognostic value in HCC patients. We developed a score based on these seven genes which correlates with survival and early recurrence. The score should make it feasible to better predict prognosis for an individual patient. This method requires the determination of only seven genes in contrast with most of the previous microarray studies that needed large cohorts for clustering and determining relative gene expression. Furthermore, we suggest that gene expression under chronic hypoxia might lead to new therapeutic targets, and the hypothesis-driven experimental methods we used could be applied to other malignancies as well.

is unorganized and the structural and functional defects lead to both acute hypoxia and chronic hypoxia (11, 12). Moreover, liver cancer usually develops in a cirrhotic environment in which the blood flow is already impaired.

We hypothesized that in HCC, there are regions with sustained hypoxia which induces a characteristic gene expression pattern. In addition, the extent of hypoxic gene expression determines the aggressiveness, or in general, the prognosis.

Materials and Methods

Chronic hypoxic gene expression

We used the human hepatoblastoma cell line HepG2 as an *in vitro* model. Cell culture experiments were done within 6 months after the cell line was obtained from American Type Culture Collection.

HepG2 cells were grown in a humidified incubator (Sanyo MCO-18M O2/CO$_2$ incubator; 20% O$_2$, 5% CO$_2$ at 37°C) in Williams medium E (WEM; InVitrogen) supplemented with 10% FCS, 2 mmol/L of L-glutamine, 20 mU/mL of insulin, 50 nmol/L of dexamethasone, 100 units/mL of penicillin, 100 μg/mL of streptomycin, 2.5 μg of fungizone, 50 μg/mL of gentamicin, and 100 μg/mL of vancomycin (WEM-C).

Dynamic behavior of hypoxia-related gene expression was studied over a period of 72 hours at different oxygen concentrations using real-time PCR and immunohistochemistry. From these results (Supplementary Figs. S1 and S2; Table 4), we concluded that 72 hours and 2% of O$_2$ represents a chronic hypoxia–like condition distinct from the acute situation during the first 24 hours.

Differentially expressed genes between HepG2 cells cultured for 72 hours at 2% O$_2$ (hypoxia) and cells cultured in parallel in 20% O$_2$ (normoxic) were determined with microarray analysis. For the microarray analyses, two experiments were executed in parallel. Cells were seeded at 3×10^6 in 75 cm^2 tissue culture flasks ($n = 4$) at 20% O$_2$ and were grown until 70% confluence. After reaching near-confluence, two flasks were placed in a humidified incubator under hypoxic conditions (2% O$_2$, 5% CO$_2$ at 37°C), whereas the other flasks ($n = 2$) remained in normoxic conditions (20% O$_2$). Cells were cultured for 72 hours in these different oxygen conditions, and after 3 days, cells were harvested. RNA samples were labeled and hybridized on dual-color Agilent Human Whole Genome Oligo Microarray (Agilent). The Agilent technology uses one glass array for the simultaneous hybridization of two populations of labeled, antisense cRNAs obtained from two samples (hypoxic and normoxic control) leading to a ratio for expression. Probes with corrected P values of <0.01 and a fold change of >2 were defined as highly significant and designated as the *in vitro* hypoxia gene set (Supplementary Table S2).

Clinical data sets and patient characteristics

To determine the clinical relevance of *in vitro* gene expression, we compared our findings in the largest microarray data sets with corresponding clinical information that is available in public databases. We used four publicly available HCC studies, which are summarized in Table 1 (13–17). By testing only the differentially expressed genes in the four data sets, the reliability of genome-wide research can be increased because the number of genes tested is in proportion with the number of patients studied (18).

Development of a prognostic gene set

We used the first three data sets as training sets to optimize our *in vitro* hypoxia gene set and to investigate the prognostic correlation. The fourth data set (Chiang et al.), was subsequently used to independently validate the signature. In the first step to defining a robust score from the different data sets, we used a global test (19) to investigate whether the hypoxia genes were associated with prognosis under a Q2-null hypothesis (20). This approach should be advantageous in that it is less dependent on the array platform used in different laboratories (Affymetrix, Agilent, Stanford, etc.). Moreover, by starting from a small subset of *in vitro*–determined hypoxia genes, this method provides more insight into the degree of relationship between the different genes found to be upregulated or downregulated. This enrichment-based method was then used to investigate whether our hypoxia set separates the good and poor prognostic characteristics in the three data sets individually. Because in all four data sets, another prognostic factor was reported, we also had to use a different prognostic factor in every data set. From Boyault et al. (13), the FAL index was used (21), which is a measure for chromosomal instability and a high score (>0.128) was associated with poor prognosis. In Wurmbach et al. (16),
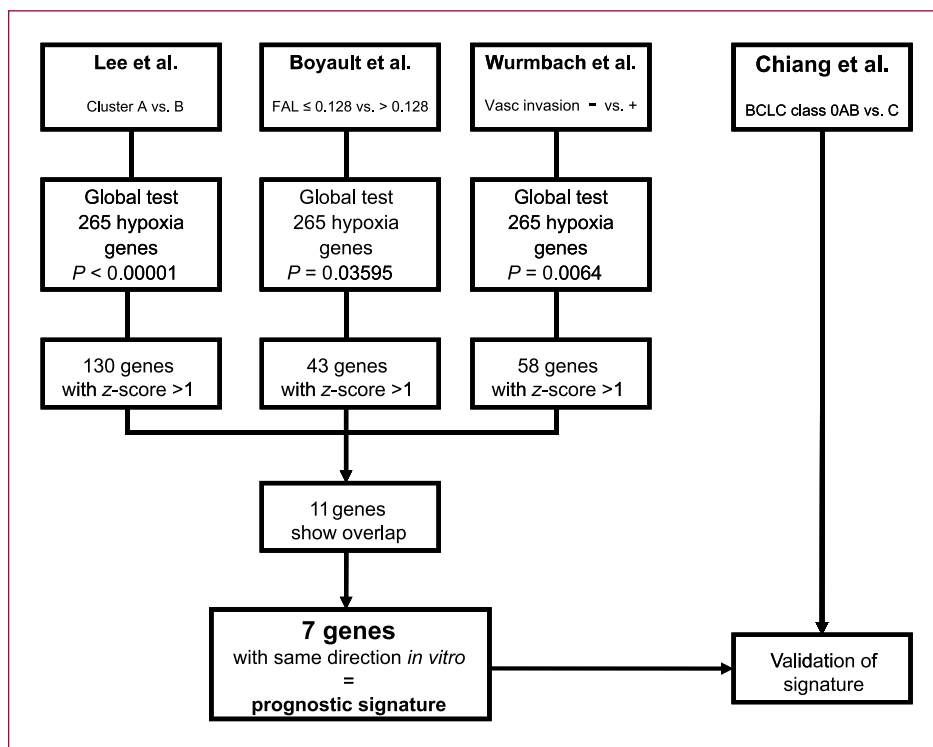
Lee et al.

Cluster A vs. B

Boyault et al.

FAL ≤ 0.128 vs. > 0.128

Wurmbach et al.

Vasc invasion − vs. +

Chiang et al.

BCLC class 0AB vs. C

Global test
265 hypoxia
genes
$P < 0.00001$

Global test
265 hypoxia
genes
$P = 0.03595$

Global test
265 hypoxia
genes
$P = 0.0064$

130 genes
with z-score >1

43 genes
with z-score >1

58 genes
with z-score >1

11 genes
show overlap

7 genes
with same direction *in vitro*
=
prognostic signature

Validation of
signature

Fig. 1. Selection of the seven-gene prognostic hypoxia gene set. We started with 265 genes that were identified from the microarray experiments with HepG2 cells. After several selection steps, we could identify a seven-gene set with prognostic value in our three training sets. The prognostic value was subsequently validated in the independent series of Chiang.

vascular invasion was used (22, 23). Lee et al. (14, 15) used the different prognostic clusters that correlate with survival (cluster A with poor prognosis and cluster B with good prognosis), and Chiang et al. (17) used the Barcelona Staging Classification (BCLC; ref. 24). Finally, based on the *z* scores, we determined which genes in our hypoxia set were most influential in differentiating between good and poor prognosis (19).

Results

Chronic hypoxia gene signature

We started with a cell culture model and determined the differentially expressed genes in HepG2 cells that were cultured for 72 hours at either 20% oxygen or in hypoxic conditions at 2% oxygen. We used Agilent technology with color flip on two independent experiments in duplicate, resulting in eight ratio values. A total of 37,707 spots showed a representative signal, of which 3,592 (8%; representing 3,259 genes) had a fold change of >2 and a *P* value of <0.05 (1,879 upregulated and 1,713 downregulated). To control the false discovery rate, multiple testing corrections were done and probes with a corrected *P* value of <0.01 and a fold change of >2 were selected (25). This resulted in a list of 265 highly significant genes (207 upregulated and 58 downregulated; Supplementary Table S2), designated as the *in vitro* hypoxic gene set (the microarray data has been deposited at the National Center for Biotechnology Information under no. GSE 15366).

Minimal prognostic gene set

We used presently available published data sets to investigate the prognostic correlation with our *in vitro*–derived hypoxia gene set and to develop an optimized, reduced prognostic gene list (Fig. 1). To test whether the overall expression pattern of these 265 hypoxia genes was significantly related to the prognostic factor considered for each of the three training data sets, the global test of Goeman et al. was used (19). This resulted in a significant enrichment of the hypoxia gene set for all three training sets ($P = 0.03595$ for Boyault, $P < 0.00001$ for Lee, and $P = 0.0064$ for Wurmbach).

Next, when keeping only the significant genes with a *z* score of >1, 130 genes remained for the data set of Lee et al., 43 genes for Boyault et al., and 58 genes for Wurmbach et al. Finally, genes for which the direction of altered expression did not correspond to the direction observed *in vivo* were removed. With this approach, we were able to downsize our hypoxia gene set to seven genes, the hypoxia signature of which was found to overlap between the three training data sets (Fig. 1).

Of the genes in our hypoxia signature, four genes were upregulated and three downregulated under hypoxia (Table 2). When we compared the expression of these seven genes in different normal tissues, we found that they have a unique expression pattern in liver not found in other tissues (Supplementary Fig. S3). Furthermore, we tested our seven-gene set related to HCC in other malignancies, such as colon carcinoma and colorectal liver metastasis (GSE3294, GSE 2630, and GSE6988). The seven-gene

set had a poor prognostic performance in these data sets (data not shown).

Chronic hypoxia as a prognostic marker

The seven genes were used to define the hypoxia score:

Hypoxia score = mean (expression ratio UP regulated genes in log base 2) − mean (expression ratio DOWN regulated genes in log base 2).

UP, the *in vivo* upregulated genes ($n = 4$), and DOWN, the *in vivo* downregulated genes ($n = 3$; Table 2).

This hypoxia score was used to classify patients in the three training sets and to calculate the area under the ROC curve (AUC), to assess the predictive performance in all data sets. The hypoxia score based on these seven genes could significantly divide patients into those with or without vascular invasion (Wurmbach; AUC, 88.9%), with a FAL index of >0.128 and ≤0.128 (Boyault; AUC,

Fig. 2. Seven-gene hypoxia score applied to patients with HCC. A, ROC-curves for the three training sets. The AUC for Wurmbach (vascular invasion) = 88.9%, the AUC for Boyault (FAL index) = 72.8% and the AUC for Lee (Clusters) = 84.9%. B, ROC curves for the validation set (Chiang) after application of the seven-gene prognostic signature. A division was made between BCLC stage O/A/B versus C (solid line; AUC, 91.0%), division between BCLC stage O/A versus B + C (dotted line; AUC, 71.5%).
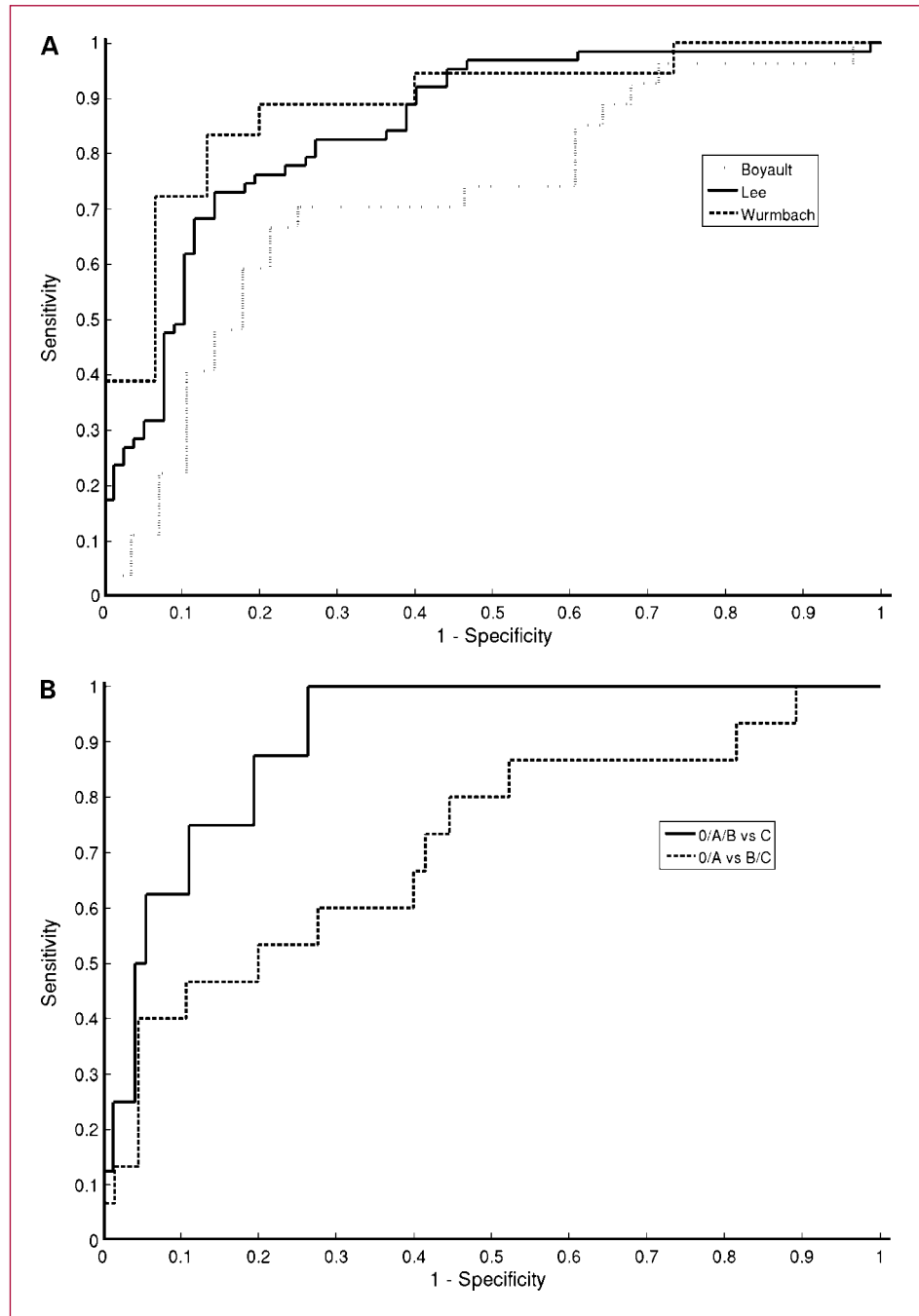
Table 1. Overview of published data sets that were used in this study

Available clinical data sets of HCC patients

	Boyault et al. (13)	Lee et al. (14, 15)	Wurmbach et al. (16)	Chiang et al. (17)
Data set ID	E-TABM-36	GSE1898 GSE4024	GSE6764	GSE9843
Array type	Affymetrix HG-U133A	Human Array-Ready Oligo Set, Qiagen	Affymetrix HG-U133A plus version 2.0	Affymetrix HG-U133A plus version 2.0
N array	65	139	73	91
N patients	60	139	48	91
N HCC	57	140	33	91
N control	5 pools of samples	19 pools of samples	10	8
N other (cirrhosis, adenoma, dysplasia)	3 adenoma = 3	none	30 cirrhosis = 13, dysplasia = 17	none
Sex	+	+	NA	+
M/F	47/13	102/37		54/27 (NA = 10)
Age	+	+	NA	+
Mean age (y)	61	56		65 (NA = 10)
Underlying liver disease	+/− HBV status + = 15	+ 14 crypto, 16 (N)ASH, 56 HBV, 14 HCV, 5 metabolic, 2 AIH, 1 PBC, 9 combi, 22 NA	+ All HCV	+ All HCV
Cirrhosis	NA	+ 50% positive, NA = 1	+ All cirrhosis	NA
AFP	NA	+ >300 = 55, NA = 11	NA	+ >300 = 15, NA = 22
Tumor size (5 cm)	NA	+ >5 = 77 NA = 1	+	NA (BCLC)
Differentiation	NA	+ 1 = 2, 2 = 57, 3 = 74, 4 = 6	+ 1 = 12, 2 = 9, 3-4 = 12	NA
Vascular invasion	NA	+ − = 21, + = 27, NA = 91	+ no = 15, micro = 11, macro = 7	NA (BCLC)
Prognostic clusters	NA	+ A = 60, B = 80	NA	NA
Satellite nodules*	+ 22/57 (39% +)	NA	+ 15/33 (45% +)	NA
BCLC score	NA	NA	NA	+ 0 = 9, A = 56, B = 7, C = 8, NA = 11
FAL index	+ − = 29, + = 26, NA = 5	NA	NA	NA
p53 mutation	+ − = 45, + = 14, NA = 1	NA	NA	+ − = 74, + = 11, NA = 6
β-Catenin mutation	+ − = 41, + = 18, NA = 1	NA	NA	+ − = 60, + = 27, NA = 4

Abbreviation: NA, not available.
*Satellite nodules were defined differently in Boyault and Wurmbach.

72.8%), and with cluster A and cluster B gene expression (Lee; AUC, 84.9%; Fig. 2A). To independently test the performance of our hypoxia score, we used the Chiang data set as validation set, with the BCLC classification as prognostic characteristic. The seven genes significantly separate the BCLC group 0/A/B from group C (AUC, 91%) as well as the group 0/A from B/C (AUC, 71.5%; Fig. 2B).
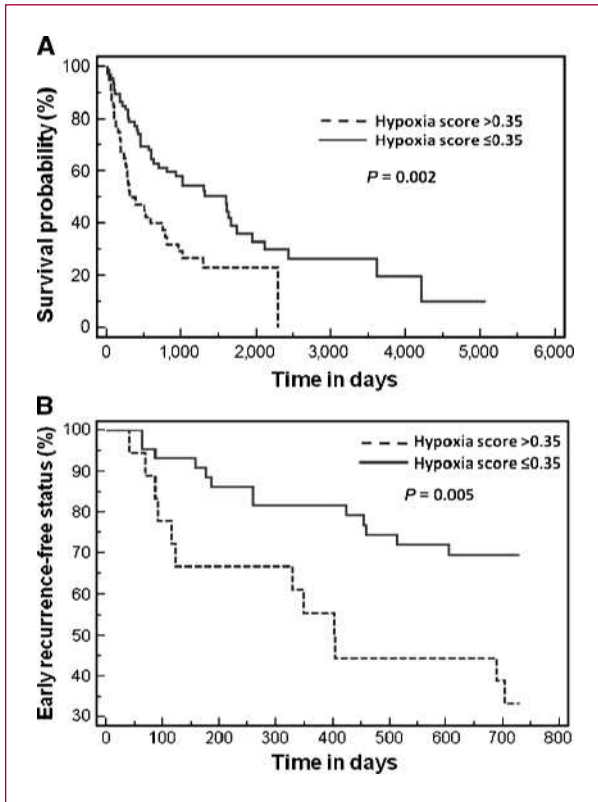
Fig. 3. Survival and early recurrence. A, Kaplan-Meier survival curve for 135 patients (14). The hypoxia score was used in a dichotomous manner (>0.35 versus ≤0.35). Patients with a score of >0.35 had a significantly poorer survival. B, Kaplan-Meier curve for early recurrence (<2 y) in 135 patients. Again, the hypoxia score was used in a dichotomous manner (>0.35 versus ≤0.35). Patients with a score of >0.35 had a significantly increased risk of early recurrence.

Survival and early recurrence

With the development of the hypoxia score, we were able to test whether the score correlates with survival and recurrence. We conducted a retrospective survival analysis on 135 patients of the study by Lee et al. (Med-Calc Software, version 11.0.1). The patient data is summarized in Table 3. We first determined the Cox proportional hazards ratio for survival because our hypoxia score is a continuous variable. Indeed, the hypoxia score significantly increased the risk of death (hazard ratio, 1.39; 95% confidence interval, 1.09-1.76; $P = 0.007$). If we used a cutoff value of 0.35 for the hypoxia score (log rank test; $P = 0.0018$), we were able to show significant differences in survival in 135 patients with a Kaplan-Meier survival curve (Fig. 3A). The median survival for patients with a hypoxia score of >0.35 ($n = 42$) was 307 days, whereas the median survival for patients with a hypoxia score of ≤0.35 ($n = 93$) was 1,602 days ($P = 0.002$). For recurrence in HCC patients, it has been suggested to make a differentiation between early recurrence (<2 years) and late recurrence (>2 years; refs. 26, 27). Early recurrence is the result of dissemination of the primary tumor and tumor characteristics determine the risk of recurrence. On the other hand, recurrence after 2 years is usually a second primary tumor that arises in a cirrhotic liver and has no relation with the first tumor. Because our hypoxia score is determined on the tumor tissue itself, we tested if it could predict early recurrence. We calculated a significant Cox proportional hazard ratio of 2.92 (95% confidence interval, 1.33-6.39; $P = 0.007$), which means that with an elevation of the hypoxia score with 0.1 point, the risk of developing a recurrence is 29.2% higher. Again, when we use a cutoff of 0.35 for the hypoxia score, the Kaplan-Meier curve shows a significant difference in early recurrence ($P = 0.005$; Fig. 3B).

We also conducted a multivariate statistical analysis to determine the performance of the hypoxia score compared with other variables, such as AFP level (≤300 versus >300), tumor size (≤5 versus >5 cm), and differentiation grade (1-2 versus 3-4; Table 4). For the 135 samples, it was already shown that the clustering (A versus B and hepatocyte versus hepatoblast) has a very strong correlation with survival and recurrence (both $P < 0.001$; ref. 15). Multivariate

---

Table 2. List of the seven hypoxia-related prognostic genes in HCC

Genes in the hypoxia signature

Gene	Full name	Response to hypoxia
CCNG2	Cyclin G2	Upregulation
EGLN3	Egl nine homologue 1 (*C. elegans*)	Upregulation
ERO1L	Endoplasmic reticulum oxidoreductin-1 L	Upregulation
WDR45L	WDR45-like	Upregulation
FGF21	Fibroblast growth factor 21	Downregulation
MAT1A	Methionine adenosyltransferase Iα	Downregulation
RCL1	RNA terminal phosphate cyclase-like 1	Downregulation

NOTE: *CCNG2*, *EGLN3*, *ERO1L*, and *WDR45L* will be upregulated under hypoxia in HCC, and *FGF21*, *MAT1A*, and *RCL1* will be downregulated under hypoxia in HCC.

Table 3. Characteristics of the 135 patients

	Total	Hypoxia score <0.35	Hypoxia score >0.35
No. of patients	135	93	42
Age (y) ± SD	56 ± 14	56 ± 14	56 ± 12
Sex (M/F)	98/37	67/26	31/11
Etiology HBV/HCV/(N)ASH/crypto/other/unknown	55/14/13/14/17/22	29/11/11/12/14/16	26/3/2/2/3/6
Cirrhosis (no/yes/NA)	67/67/1	45/47/1	22/20/0
Tumor size (<5 cm/≥5 cm/NA)	55/79/1	41/51/1	14/28/0
Histologic grade (1/2/3/4)	2/55/72/6	2/43/43/5	0/12/29/1
Vascular invasion (no/yes/NA)	27/48/60	24/37/32	3/11/28
AFP level (<300/≥300/NA)	70/54/11	52/24/7	18/20/4
Subgroup (B/A/HB)	74/38/23	67/21/5	7/17/18

NOTE: Baseline characteristics of the 135 patients and per subgroup based on the hypoxia score. Histologic grades were based on the Edmondson and Steiner classification and AFP levels were measured in ng/mL.

Abbreviations: HBV, hepatitis B virus; HCV, hepatitis C virus; (N)ASH, (non)alcoholic steatohepatitis; crypto, cryptogenic; NA, not available.

analysis with Cox proportional hazards regression, including the hypoxia score, again showed that clustering based on A/B/hepatoblast is the only significant predictor of both poor survival and early recurrence ($P = 0.009$ and $P < 0.001$). However, the differentiation between A and B is based on 406 genes and the differentiation between a hepatoblast or hepatocyte phenotype is based on 941 genes. Furthermore, different samples are needed to perform clustering based on these genes. In contrast, our hypoxia score is based on seven genes and can be calculated for an individual patient. We therefore explored the performance of the hypoxia score without integration of clustering. The hypoxia score is the only significant variable for both survival ($P = 0.02$) and early recurrence ($P = 0.008$); this correlation was, however, not as strong as that found for the original clustering (A/B/hepatoblast), but required a much larger set of genes and is only valid in that single study.

Molecular classification

Both in the study of Boyault et al. as well as in the study by Chiang et al., the authors divided their patients into different molecular subgroups. Using their classifications, we found that the hypoxia score corresponded with the subgroups that had a poor prognosis (Fig. 4A and B). This subgroup is associated with proliferation and mTOR signaling and chromosomal instability and is seen more frequently in patients with advanced HCC.

Discussion

The diagnosis for HCC is currently based on a combination of imaging techniques, serum markers, and histology (28, 29); and for a reliable diagnosis, a multidisciplinary approach is required. There are, however, no clear criteria available that could accurately predict the prognosis and

several approaches have been attempted to develop rational criteria. Great progress has been made developing clinically relevant cancer signatures especially for breast cancer (30, 31) and, to a lesser extent, for colon cancer (32, 33). For patients with HCC, a similar method has not been developed until now. Recently, several studies have tried to identify gene sets with prognostic or diagnostic relevance by microarray analysis (34–36). Prognosis is a key factor in the selection of a therapy and therefore molecular classification has become increasingly important for the development of targeted therapy such as the multikinase inhibitor sorafenib (37). Although a lot of progress has been made in the molecular classification of HCC over recent years, it is difficult to transpose the results from one study to another study or patients.

The rationale behind this study was to start from the hypothesis that during the development of a solid tumor such as HCC, there is a chronic hypoxic condition which is different from acute hypoxia. Hypoxia is a well-known characteristic of solid tumors and has an effect on the aggressiveness of tumors, induces neoangiogenesis, anaerobic metabolism, and promotes invasiveness.

Central to this study is the working hypothesis that chronic exposure to hypoxia leads to an adaptive gene expression profile which influences the aggressive behavior of the tumor cells. To test our hypothesis independently of patient selection and variability, we decided to start from cell cultures. Here, we did not want a concomitant role for mutations in the p53 gene or liver cell lines that express fragments of hepatitis B or C (38). Therefore, we selected the human HepG2 cell line, although it has hepatoblastoma characteristics (39). The second selection was made to study hypoxia at 2% O_2, which is well below the oxygen tension found in normal livers (40). Additional experiments at different oxygen tensions support the selection of 2% O_2 for 72 hours *in vitro* (Supplementary

Figs. S1 and S2; Table 4). Subsequently, the gene expression in this model was compared with gene expression determined by microarrays in different groups of HCC patients to identify the genes relevant for HCC and not solely for hepatoblastoma.

In contrast with some of the previous studies on HCC, we did not limit the number of genes we wanted to investigate by a priori selection, and used the Agilent 44 K microarray, which covers all the known genes. Although the dynamics of gene expression indicate that after an adaptation period of 72 hours, the gene expression is not as strongly altered as during the first 24 hours, we still found that 8% of the genes were significantly changed at 72 hours.

Starting with the group of 265 highly significant genes extracted from the microarray study of HepG2 cells, we followed a sequence of analytic steps and developed a very robust seven-gene prognostic signature using the method of Goeman et al. (19). This seven gene prognostic set was applied to four data sets of HCC patients. It could significantly divide patients with different known risk factors into the three studies (training sets) used to develop the seven-gene prognostic signature. Furthermore, in the fourth study (17), this set could significantly separate the BCLC group 0/A/B from C or BCLC group 0/A from B/C. More importantly, the hypoxia score that can be obtained from the seven-gene set significantly correlates with survival and early recurrence.

Two other important microarray studies on HCC with prognostic value have also been reported (34, 41). These studies, however, concentrated on the surrounding liver tissue and showed a correlation with inflammatory pathways and tumor recurrence, whereas our seven-gene set is applicable on the tumor tissue itself. Indirect evidence suggests that our seven-gene set is liver-specific. First of all, the pattern of gene expression in the liver of these seven genes was different compared with other tissues (Appendix: Fig. 3). Second, the hypoxia score based on our seven-gene set had a poor performance in data sets from other malignancies. Third, if we repeat the same experimental sequence with gastrointestinal cancer cell lines, such as from colon and pancreas, we identified other sets of differently expressed genes (Supplementary Fig. S4).

All seven genes are, to a certain extent, related to either hypoxia, proliferation, or metabolism (see Supplementary Materials, Functions and interactions of the seven genes in relation to hypoxia). CCNG2 is a regulator of the cell cycle and is upregulated in response to DNA damage (42). Also WDR45L and RCL1 play a role in proliferation and RNA preprocessing, although the exact function of these two genes are unknown (43). EGLN3 and ERO1L are both upregulated in response to hypoxia. EGLN3 has been described as the main actor in the response to chronic hypoxia and the regulation of Hif1α (44). ERO1L is necessary in the proper protein folding under hypoxic
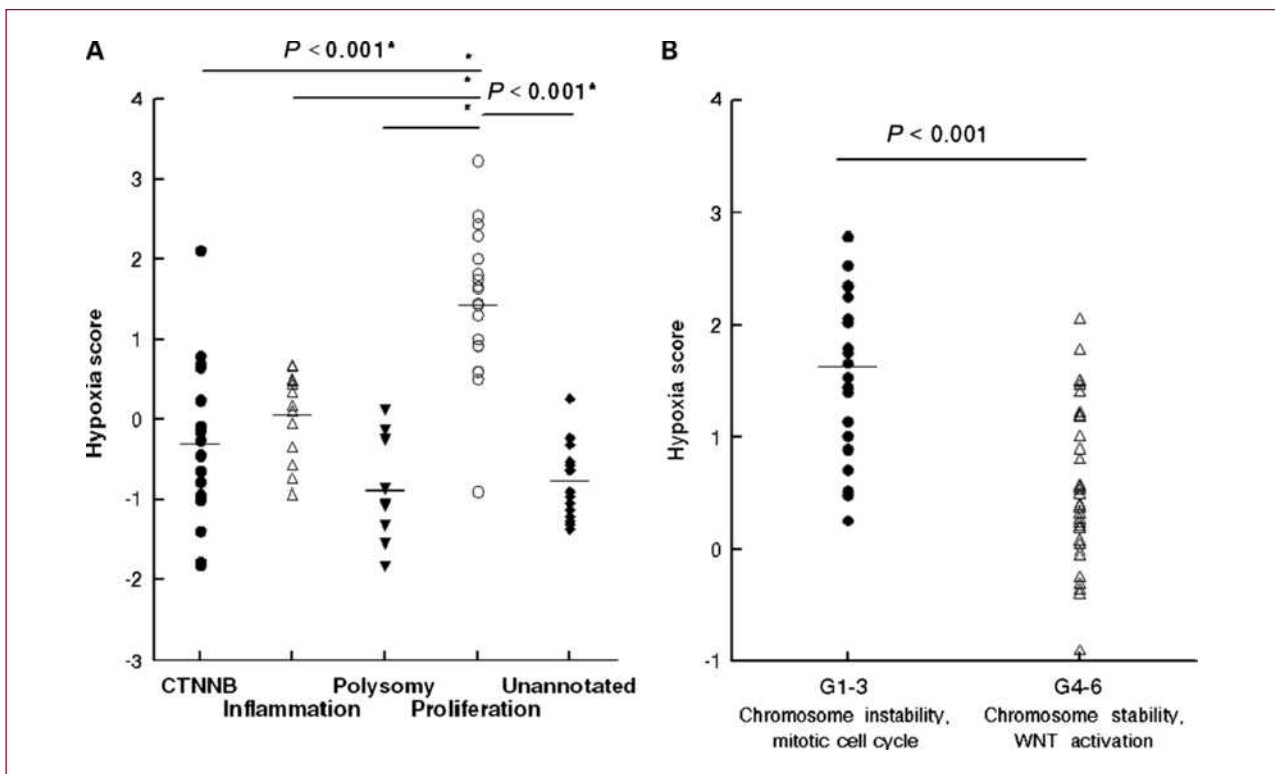

Fig. 4. Hypoxia score and molecular classification. A, hypoxia score applied to the clusters used by Chiang. B, hypoxia score applied to the clusters used by Boyault (G1-3, including proliferation; G4-6, including Wnt signaling). CTNNB: catenin (cadherin-associated protein), β1, 88 kDa (*Homo sapiens*).

Table 4. Multivariate analysis of the 135 patients

Multivariate analysis in 135 patients

	Hazard ratio (95% confidence interval)*	*P*
OS		
HB subgroup	2.24 (1.23-4.06)	0.009
B subgroup	0.58 (0.34-1.00)	0.05
OS analysis without subgroups		
Hypoxia score >0.35	2.57 (1.14-5.76)	0.02
ER		
HB subgroup	5.07 (2.06-12.48)	<0.001
ER analysis without subgroups		
Age >65 y	2.16 (0.99-4.74)	0.06
Hypoxia score >0.35	2.90 (1.32-6.36)	0.008

NOTE: We present only the (nearly) significant variables.

Abbreviations: OS, overall survival; ER, early recurrence; HB, hepatoblastoma.

*The hazard ratios for overall survival and early recurrence for patients with or without a hepatoblastoma genotype, with or without a B genotype, with a hypoxia score > or <0.35, and age > or <65 y. Because the subgroups published in Lee were determined on large gene sets and patient clustering, we also studied the performance of our seven-gene set without the subgroups. Vascular invasion was not included in the multivariate analysis due to the large amount of missing values.

conditions (45). FGF21 and MAT1A are downregulated in response to hypoxia and are involved in metabolic processes. FGF21 is involved in energy homeostasis, it has an endocrine function regulating glucose uptake in adipocytes (46). MAT1A is unbearable in normal liver function because it regulates most methylation processes (47).

As a clinical tool, the hypoxia score seems promising. It shows good correlation with survival as well as early recurrence in patients. It is a good alternative for the classification by Lee et al. (14, 15), with the advantage of requiring a much smaller gene set. In multivariate analyses, our hypoxia score had a better performance than, for example, tumor size and differentiation grade—two variables that are currently used to estimate the aggressiveness of tumors. The identification of these seven genes is the next step in predicting prognosis, and although preliminary, it might influence therapeutic decision-making in the future. Besides, it can be used in randomized trials for risk stratification.

Our findings might also have important implications for the use of antiangiogenic therapies. Nowadays, both chemoembolization and the anti–vascular endothelial growth factor receptor kinase sorafenib are standard treatments for patients with intermediate or advanced HCC. It has been previously noted that resistance emerges after antiangiogenic therapies, leading to untreatable aggressive tumors (48). Our findings on the role of hypoxia lead to the reconsideration of treatment strategies, but more importantly, our gene set could help in the search for new therapeutic targets.

In the last few years, the molecular classification of HCC has attracted a lot of attention. Based on gene expression, patients could be classified into several subgroups such as the β-catenin subgroup, the proliferation subgroup, and the inflammation subgroup. The exact prognostic and therapeutic implications of this categorization are still unclear. We analyzed our hypoxia signature in the studies of Chiang and Boyault, and there was a clear correlation with the proliferation cluster. This cluster consists of genes related to the mTOR pathway and several cell cycle genes, such as cyclins. In the proliferation cluster defined by Chiang, there was a significant increase in pRPS6 staining, indicating aberrant mTOR signaling. The mTOR signaling pathway regulates cell growth, cell proliferation, protein transcription, and survival by orchestrating several upstream signals. Recently, an important role for the mTOR pathway in HCC was shown (49, 50). Also, several studies have shown an interaction between mTOR and hypoxia (or HIF1A). There is evidence for an oxygen-independent induction of HIF1A by mTOR signaling and upregulation of targets such as vascular endothelial growth factor (51, 52). These facts warrant further examinations of the hypoxia score in combination with targeting the mTOR pathway in human studies.

In conclusion, we have identified a seven-gene set that correlates with prognosis in HCC patients. We were able to extract this very small gene set by combining four available microarray data and because we had identified in advance an *in vitro* gene set.

Our findings have potential implications in several areas:

1. Because our prognostic signature relies on a limited number of genes, its application should be broadly possible using generally available techniques such as real time PCR. However, it will be necessary to

determine the criteria which a sample should fulfill and we recommend further cross-validation versus histology.

2. As a prognostic indicator, it can influence the therapeutic options that are available for a patient. Therefore, this signature needs to be further validated in new prospective studies to confirm its use.

3. The seven-gene set we have identified is based on an *in vitro* selection of genes that were altered during chronic hypoxia. These results suggest a role for hypoxia in the aggressive behavior of HCCs, although it still has to be proven that the gene expression of these seven genes in aggressive tumors is the direct consequence of chronic hypoxia.

4. The hypoxia score could possibly be further improved with genes that can be identified with the use of other cell lines.

5. The method we used to identify this limited gene set —the combination of a cell culture model and the global test method—could also be applied to other tumors. With this hypothesis-driven method, it is easier to extract the most important genes out of the large amount of information from the microarray technique, even when the number of patients that can be recruited into a single study is limited, as for HCC patients.

Although *in vitro* studies are never fully representative for the situation as it develops in an organ, the validation using four clinical data sets proves the value of our study beyond theoretical objections. These findings have prognostic implications for patients with HCC, and therefore, could be incorporated into the molecular classification of HCC.

Disclosure of Potential Conflicts of Interest

Acknowledgments

Grant Support

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.
2. Llovet JM, Burroughs A, Bruix J. Hepatocellular carcinoma. Lancet 2003;362:1907–17.
3. Verslype C, Libbrecht L. The multidisciplinary management of gastrointestinal cancer. The diagnostic and therapeutic approach for primary solid liver tumours in adults. Best Pract Res Clin Gastroenterol 2007;21:983–96.
4. Thorgeirsson SS, Lee JS, Grisham JW. Functional genomics of hepatocellular carcinoma. Hepatology 2006;43:S145–50.
5. Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. Proc Natl Acad Sci U S A 2006;103:5923–8.
6. Gort EH, Groot AJ, van der Wall E, van Diest PJ, Vooijs MA. Hypoxic regulation of metastasis via hypoxia-inducible factors. Curr Mol Med 2008;8:60–7.
7. Semenza GL. Targeting HIF-1 for cancer therapy. Nat Rev Cancer 2003;3:721–32.
8. Brown JM, Wilson WR. Exploiting tumour hypoxia in cancer treatment. Nat Rev Cancer 2004;4:437–47.
9. Brown JM, Giaccia AJ. The unique physiology of solid tumors: opportunities (and problems) for cancer therapy. Cancer Res 1998;58:1408–16.
10. Kim KR, Moon HE, Kim KW. Hypoxia-induced angiogenesis in human hepatocellular carcinoma. J Mol Med 2002;80:703–14.
11. Folkman J, Hahnfeldt P, Hlatky L. Cancer: looking outside the genome. Nat Rev 2000;1:76–9.
12. Brahimi-Horn MC, Chiche J, Pouyssegur J. Hypoxia and cancer. J Mol Med 2007;85:1301–7.
13. Boyault S, Rickman DS, de Reynies A, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology 2007;45:42–52.
14. Lee JS, Chu IS, Heo J, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology 2004;40:667–76.
15. Lee JS, Heo J, Libbrecht L, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med 2006;12:410–6.
16. Wurmbach E, Chen YB, Khitrov G, et al. Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology 2007;45:938–47.
17. Chiang DY, Villanueva A, Hoshida Y, et al. Focal gains of VEGFA and molecular classification of hepatocellular carcinoma. Cancer Res 2008;68:6779–88.
18. Chang HY, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proc Natl Acad Sci U S A 2005;102:3738–43.
19. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004;20:93–9.
20. Tian L, Greenberg SA, Kong SW, Altschuler J, Kohane IS, Park PJ. Discovering statistically significant pathways in expression profiling studies. Proc Natl Acad Sci U S A 2005;102:13544–9.
21. Dvorchik I, Schwartz M, Fiel MI, Finkelstein SD, Marsh JW. Fractional allelic imbalance could allow for the development of an equitable transplant selection policy for patients with hepatocellular carcinoma. Liver Transpl 2008;14:443–50.
22. Iizuka N, Oka M, Yamada-Okabe H, et al. Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. Lancet 2003;361:923–9.
23. Wang SM, Ooi LL, Hui KM. Identification and validation of a novel gene signature associated with the recurrence of human hepatocellular carcinoma. Clin Cancer Res 2007;13:6275–83.
24. Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis 1999;19:329–38.
25. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995;57:289–300.
26. Cucchetti A, Piscaglia F, Caturelli E, et al. Comparison of recurrence of hepatocellular carcinoma after resection in patients with cirrhosis to its occurrence in a surveilled cirrhotic population. Ann Surg Oncol 2009;16:413–22.

27. Imamura H, Matsuyama Y, Tanaka E, et al. Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. J Hepatol 2003;38:200–7.

28. Colli A, Fraquelli M, Casazza G, et al. Accuracy of ultrasonography, spiral CT, magnetic resonance, and α-fetoprotein in diagnosing hepatocellular carcinoma: a systematic review. Am J Gastroenterol 2006;101:513–23.

29. Libbrecht L, Severi T, Cassiman D, et al. Glypican-3 expression distinguishes small hepatocellular carcinomas from cirrhosis, dysplastic nodules, and focal nodular hyperplasia-like nodules. Am J Surg Pathol 2006;30:1405–11.

30. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002; 347:1999–2009.

31. Bueno-de-Mesquita JM, van Harten WH, Retel VP, et al. Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). Lancet Oncol 2007;8:1079–87.

32. Arango D, Laiho P, Kokko A, et al. Gene-expression profiling predicts recurrence in Dukes' C colorectal cancer. Gastroenterology 2005; 129:874–84.

33. Lin YH, Friederichs J, Black MA, et al. Multiple gene expression classifiers from different array platforms predict poor prognosis of colorectal cancer. Clin Cancer Res 2007;13:498–507.

34. Hoshida Y, Villanueva A, Kobayashi M, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N Engl J Med 2008;359:1995–2004.

35. Katoh H, Ojima H, Kokubu A, et al. Genetically distinct and clinically relevant classification of hepatocellular carcinoma: putative therapeutic targets. Gastroenterology 2007;133:1475–86.

36. Chen X, Cheung ST, So S, et al. Gene expression patterns in human liver cancers. Mol Biol Cell 2002;13:1929–39.

37. Llovet JM, Bruix J. Molecular targeted therapies in hepatocellular carcinoma. Hepatology 2008;48:1312–27.

38. Bressac B, Galvin KM, Liang TJ, Isselbacher KJ, Wands JR, Ozturk M. Abnormal structure and expression of p53 gene in human hepatocellular carcinoma. Proc Natl Acad Sci U S A 1990;87:1973–7.

39. Lopez-Terrada D, Cheung SW, Finegold MJ, Knowles BB. Hep G2 is a hepatoblastoma-derived cell line. Hum Pathol 2009;40:1512–5.

40. Jiang J, Nakashima T, Liu KJ, Goda F, Shima T, Swartz HM. Measurement of PO2 in liver using EPR oximetry. J Appl Physiol 1996;80:552–8.

41. Budhu A, Forgues M, Ye QH, et al. Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. Cancer Cell 2006;10:99–111.

42. Murray JI, Whitfield ML, Trinklein ND, Myers RM, Brown PO, Botstein D. Diverse and specific gene expression responses to stresses in cultured human cells. Mol Biol Cell 2004;15:2361–74.

43. Billy E, Wegierski T, Nasr F, Filipowicz W. Rcl1p, the yeast protein similar to the RNA 3′-phosphate cyclase, associates with U3 snoRNP and is required for 18S rRNA biogenesis. EMBO J 2000; 19:2115–26.

44. Ginouves A, Ilc K, Macias N, Pouyssegur J, Berra E. PHDs overactivation during chronic hypoxia "desensitizes" HIFα and protects cells from necrosis. Proc Natl Acad Sci U S A 2008;105:4745–50.

45. Wouters BG, Koritzinsky M. Hypoxia signalling through mTOR and the unfolded protein response in cancer. Nat Rev Cancer 2008;8:851–64.

46. Kharitonenkov A, Shiyanova TL, Koester A, et al. FGF-21 as a novel metabolic regulator. J Clin Invest 2005;115:1627–35.

47. Mato JM, Corrales FJ, Lu SC, Avila MA. S-Adenosylmethionine: a control switch that regulates liver function. FASEB J 2002;16:15–26.

48. Schmidt C. Why do tumors become resistant to antiangiogenesis drugs? J Natl Cancer Inst 2009;101:1530–2.

49. Villanueva A, Chiang DY, Newell P, et al. Pivotal role of mTOR signaling in hepatocellular carcinoma. Gastroenterology 2008;135: 1972–83, 83 e1–11.

50. Newell P, Toffanin S, Villanueva A, et al. Ras pathway activation in hepatocellular carcinoma and anti-tumoral effect of combined sorafenib and rapamycin in vivo. J Hepatol 2009;51:725–33.

51. Zhong H, Chiles K, Feldser D, et al. Modulation of hypoxia-inducible factor 1α expression by the epidermal growth factor/phosphatidylinositol 3-kinase/PTEN/AKT/FRAP pathway in human prostate cancer cells: implications for tumor angiogenesis and therapeutics. Cancer Res 2000;60:1541–5.

52. Land SC, Tee AR. Hypoxia-inducible factor 1α is regulated by the mammalian target of rapamycin (mTOR) via an mTOR signaling motif. J Biol Chem 2007;282:20534–43.