ORIGINAL ARTICLE

Neoangiogenesis-related genes are hallmarks of fast-growing hepatocellular carcinomas and worst survival. Results from a prospective study

Erica Villa,[1] Rosina Critelli,[1] Barbara Lei,[1] Guido Marzocchi,[2] Calogero Cammà,[3] Gianluigi Giannelli,[4] Patrizia Pontisso,[5] Giuseppe Cabibbo,[3] Marco Enea,[6] Stefano Colopi,[2] Cristian Caporali,[2] Teresa Pollicino,[7] Fabiola Milosa,[1] Aimilia Karampatou,[1] Paola Todesca,[1] Elena Bertolini,[1] Livia Maccio,[8] Maria Luz Martinez-Chantar,[9] Elena Turola,[1] Mariagrazia Del Buono,[1] Nicola De Maria,[1] Stefano Ballestri,[10] Filippo Schepis,[1] Paola Loria,[10] Giorgio Enrico Gerunda,[11] Luisa Losi,[8] Umberto Cillo[12]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2014-308483).

For numbered affiliations see end of article.

Correspondence to
Professor Erica Villa, Department of Internal Medicine, Gastroenterology Unit, Università degli Studi di Modena & Reggio Emilia, and Azienda Ospedaliero-Universitaria di Modena, Via del Pozzo 71, Modena 41124, Italy;
erica.villa@unimore.it

Received 19 September 2014
Revised 2 January 2015
Accepted 16 January 2015

ABSTRACT

Objective The biological heterogeneity of hepatocellular carcinoma (HCC) makes prognosis difficult. We translate the results of a genome-wide high-throughput analysis into a tool that accurately predicts at presentation tumour growth and survival of patients with HCC.

Design Ultrasound surveillance identified HCC in 78 (training set) and 54 (validation set) consecutive patients with cirrhosis. Patients underwent two CT scans 6 weeks apart (no treatment in-between) to determine tumour volumes (V_0 and V_1) and calculate HCC doubling time. Baseline-paired HCC and surrounding tissue biopsies for microarray study (Agilent Whole Human Genome Oligo Microarrays) were also obtained. Predictors of survival were assessed by multivariate Cox model.

Results Calculated tumour doubling times ranged from 30 to 621 days (mean, 107±91 days; median, 83 days) and were divided into quartiles: ≤53 days (n=19), 54–82 days (n=20), 83–110 days (n=20) and ≥111 days (n=19). Median survival according to doubling time was significantly lower for the first quartile versus the others (11 vs 41 months, 42, and 47 months, respectively) (p<0.0001). A five-gene transcriptomic hepatic signature including angiopoietin-2 (*ANGPT2*), delta-like ligand 4 (*DLL4*), neuropilin (NRP)/tolloid (TLL)-like 2 (*NETO2*), endothelial cell-specific molecule-1 (*ESM1*), and nuclear receptor subfamily 4, group A, member 1 (*NR4A1*) was found to accurately identify rapidly growing HCCs of the first quartile (ROC AUC: 0.961; 95% CI 0.919 to 1.000; p<0.0001) and to be an independent factor for mortality (HR: 3.987; 95% CI 1.941 to 8.193, p<0.0001).

Conclusions The hepatic five-gene signature was able to predict HCC growth in individual patient and the consequent risk of death. This implies a role of this molecular tool in the future therapeutic management of patients with HCC.

Significance of this study

What is already known on this subject?

- Hepatocellular carcinoma (HCC) is a cancer particularly difficult to classify because of the highly heterogeneous natural history. Barcelona Clinic Liver Cancer (BCLC) staging score currently is the most used algorithm to choose the therapeutic strategy and predict the clinical outcome.
- Decision-making process in HCC mainly ruled by BCLC classification is mostly based on clinical and imaging characteristics performed at the time of diagnosis.
- The major limit in this daily clinical life is represented by the 'snapshot' quality of this analysis, which overlooks the dynamic progression of the disease.

What are the new findings?

- About 25% of newly diagnosed HCCs identified on surveillance have very rapid growth, with a doubling time <2 months regardless of initial BCLC classification.
- A five-gene transcriptomic hepatic signature including angiopoietin-2 (*ANGPT2*), delta-like ligand 4 (*DLL4*), neuropilin (NRP)/tolloid (TLL)-like 2 (*NETO2*), endothelial cell-specific molecule-1 (*ESM1*), and nuclear receptor subfamily 4, group A, member 1 (*NR4A1*) identifies with high sensitivity and specificity rapidly growing HCCs.
- This signature is also an accurate indicator of survival.

How might it impact on clinical practice in the foreseeable future?

- The addition of the identified molecular signature to the clinical and radiological parameters indicated by current guidelines would have significant implication for the therapeutic management of patients with HCC, allowing a drastic refinement of prognosis.

To cite: Villa E, Critelli R, Lei B, *et al*. *Gut* Published Online First: [*please include* Day Month Year] doi:10.1136/gutjnl-2014-308483

INTRODUCTION

There are several staging systems that have been developed for the classification of patients with hepatocellular carcinoma (HCC).[1–4] However,

none of the proposed staging systems encompass the biological and clinical heterogeneity exhibited by HCCs. First, they include descriptive variables at baseline, which may only be weakly representative of tumour growth behaviour, residual hepatic function and overall health status of the patient. Second, these predictive algorithms consider HCCs to be static rather than dynamic entities. They account for the size and number of neoplastic lesions at the time of presentation, yet do not take into account their growth behaviour during follow-up, such as tumour doubling time (DT). Third, molecular characteristics and tumour pathobiology (eg, local and micro vascular invasion, angiogenesis and vasculogenesis) are not accounted for in scoring systems that only consider clinical features. Indeed, these relevant biological variables may profoundly impact tumour growth, tumour responsiveness to treatment and ultimately patient survival.

Microarray technology has led to the identification of several molecular signatures in HCC, associated with deregulation of specific genes and molecular pathways, including the extracellular matrix, cytoskeleton, oncogenes, tumour suppressor genes, immune response-related genes, apoptosis-related genes and signal transduction/translational regulatory genes (WNT, TGFβ, MAPK, EGFR, IGF-R, and MET/HGF).[5–11] These signatures were able to predict prognosis,[5 6] survival[7 8] and early recurrence after treatment[9–11] in selected HCC cohorts, allowing stratification of HCC into several clinically relevant subgroups, unrecognisable by conventional diagnostic methods.

The goal of this study was to prospectively test the hypothesis, in a consecutive series of patients at first diagnosis of HCC, that HCCs have different growth patterns marked by specific molecular signatures, which can be used to predict tumour progression and patient survival in individual cases at first diagnosis.

METHODS

Patients and samples

A training set consisted of tissue samples obtained from patients with Child–Pugh class A liver cirrhosis of any aetiology who were followed up in our Gastroenterology Unit with ultrasound (US) surveillance at 6-month interval. Patients who received, between September 2008 and December 2010, a new diagnosis of HCC at US surveillance were eligible if they had a clinical condition that allowed a US-guided liver biopsy of a focal lesion, with the largest lesion biopsied in case of multifocality. To further confirm HCC diagnosis, a CT scan was performed. To measure the growth of lesions, a second CT was performed by either of two dedicated radiologists (SC, CC) 6 weeks later. This interval was chosen as previous data from our group indicated that a relevant difference in size could be demonstrated within this very short time interval.[12 13] During the 6-week interval, patients did not undergo any specific treatment. This interval is much shorter than the average time to treatment after HCC diagnosis[14 15] and therefore no ethical issues were raised by IRB. After the second CT, patients were treated according to international guidelines[1] and were monitored every 3 months unless their clinical condition required more frequent monitoring. At the time of diagnosis (eg, baseline), US-guided liver biopsy was performed both inside the lesion and in the surrounding tissue. Tumour and non-tumour (NT) liver samples were collected in cold RNA later (Qiagen, Milan, Italy) and immediately processed for gene expression analysis. Portions of biopsies were also fixed in 10% formaldehyde, paraffin-embedded and stained with H&E. The diagnosis of HCC was based on established histological criteria.[16] The study endpoint was death, liver transplant (LT) or completion of this study. Results were analysed with the intention to treat.

The study was conducted according to the guidelines of the Declaration of Helsinki and Good Clinical Practice in clinical trials. All patients provided written informed consent.

Validation cohort

The validation set consisted of tissue samples from patients with Child–Pugh class A liver cirrhosis of any aetiology enrolled in semi-annual US surveillance who consecutively presented at the Modena Gastroenterology Unit between January 2011 and July 2012 with a new HCC diagnosis. These patients underwent the same imaging protocol with the same two radiologists as the training set. Quantitative reverse transcriptase PCR was performed instead of microarrays to evaluate the five-gene signature.

Radiological protocol

CT scans were performed using a 64-detector machine (Lightspeed VCT, GE Healthcare, Milwaukee, Wisconsin, USA) as detailed in the online supplementary methods_imaging. Postprocessing of CT data was used to obtain the volume of each nodule detected in CT0 and CT1. Each single HCC nodule was reconstructed by the same radiologist (GM) in three orthogonal planes, with two orthogonal diameters drawn on every plane. The six resulting diameters were used to calculate the two volumes using the formula: $4/3\pi r^3$. The DT for each single mass was determined using the following formula:

$$DT = Ti \log 2 / \log(V1/V0)$$

where Ti is the time interval in days, V0 is the volume of the tumour at CT0 and V1 is the volume of the tumour at CT1.[17] Based on these values, tumour growth was classified according to quartiles of the fastest to slowest tumour growth, respectively.

Analysis of gene expression

Microarray experiments and bioinformatic analysis are detailed in the online supplementary methods_gene expression. Total RNA was isolated from NT and tumour (T) liver tissues using Trizol (Invitrogen, Carlsbad, California, USA), according to the manufacturer's directions. The quality and quantity of the RNA samples obtained was checked using an Agilent Model 2100 Bioanalyzer (Agilent Technologies, Palo Alto, California, USA) and an ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, Delaware, USA), respectively.

Upregulated and downregulated genes were identified for the first DT quartile (eg, the rapidly growing tumours) versus the three other quartiles. Only genes with an uncorrected p value that was <0.01 and an at least twofold expression difference were selected (see online supplementary table S2). Gene expression data are available at the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/geo) under the accession number: GSE54236.

Real-time reverse transcription PCR

The microarray data were validated in an independent cohort by real-time reverse transcription PCR (qRT-PCR) as detailed in online supplementary methods_gene expression.[18]

Risk score calculation

Risk score was calculated for both training and validation cohorts as detailed in online supplementary methods_gene expression.[19]

Statistical analysis

Survival at 1 year was chosen as the outcome to calculate sample size. Assuming a 30% difference in the 1-year survival between the fast-growing (20%) and slow-growing HCC (50%), and 5% α error and 20% β error, 39 patients were needed in each group.

Dichotomous and continuous variables were compared using Fisher's exact test and the non-parametric Mann–Whitney U test, respectively. Logistic regression analysis was performed to identify variables independently associated with faster growth. Candidate risk factors for faster growth were sex, age, aetiology of cirrhosis (viral vs non-viral), Edmondson–Steiner grading, presence of macrovascular invasion assessed by CT scan, volume of the tumour at baseline, multifocality at baseline, platelets level, α-fetoprotein levels and identified five-gene signature. The dependent variable (eg, rapid growth) was coded as 1 (present) vs 0 (absent). To visualise the capacity of the risk signature to discriminate between fast-growing and slow-growing HCCs, we summarised the data in a receiver operating characteristic curve.[20]

The Kaplan–Meier method was used to estimate the cumulative probability of overall survival. Patients were censored at the time of LT, death or last available follow-up. Differences in observed probability were assessed using the log-rank test.

The Cox proportional [illegible] used to identify risk factors for overall survival. The same independent variables assessed for growth speed were also used for survival analysis, with albumin, creatinine and bilirubin, used as additional independent variables. To avoid the effect of colinearity in the logistic regression and Cox models, the individual components of the scores evaluated (Child–Pugh; Barcelona Clinic Liver Cancer, BCLC) were included in the multivariate models separately. Variables with a p value <0.10 in univariate analyses were included in the final multivariate model. Internal assessment of the accuracy of the survival prognostic model was performed by data splitting and by bootstrapping.[21]

The PASW Statistics V.20 program (IBM, Armonk, New York, USA) was used for statistical analyses.

RESULTS

A total of 117 consecutive patients, undergoing US surveillance for HCC, had HCC detected. Of these, 39 were excluded (see online supplementary figure S1), resulting in 78 patients enrolled as a training set. Additional 71 consecutive patients with HCC were evaluated for enrolment as the validation set. Seventeen were excluded, resulting in a validation cohort of 54 patients (see online supplementary figure S1). Online supplementary table S1 summarises the clinical characteristics of each cohort at enrolment in surveillance. Data were censored in August 2012 for the training cohort (mean follow-up 24.1 ±12.8 months) and in December 2013 for the validation cohort (mean follow-up 15.6±11.0 months).

Baseline characteristics and growth patterns for HCCs of the training cohort

In the training cohort, the incidence of a single nodule of HCC was 54/78 (69.2%), while the incidence of two nodules was 12/78 (15.4%) and three or more nodules was 12/78 (15.4%). Mean volume at presentation was 12.1±19.7 cm^3 (median, 4.9 cm^3; range, 1–59.2 cm^3) (table 1).

Tumour DT was found to range from 30 to 621 days (mean, 107±91 days; median, 83 days) for the training cohort. Patients were grouped into four quartiles according to increasing DT: ≤53 days (n=19), 54–82 days (n=20), 83–110 days (n=20) and ≥111 days (n=19), respectively. Clinical characteristics of patients and HCC features according to growth speed are summarised in table 2. A representative example of a fast-growing tumour is displayed in online supplementary figure S2. The incidence of multifocal HCC was 9/19 for patients in the first quartile versus 0/20, 1/20 and 2/19 for the other three quartiles, respectively (p=0.001). The size of the individual lesions did not significantly differ between the patients with monofocalities, bifocalities or multifocalities (p=0.717). More patients were categorised as BCLC C in the fast-growing subgroup as a consequence of the significantly higher presence of portal vein thrombosis (5/19 vs 4/59, p=0.034).

Table 1 Characteristics of HCCs at presentation

Variable	Training cohort n=78	Validation cohort n=54	p Value
Tumour volume (cm^3) (M±SD)	12.1±19.7	9.9±14.8	0.465
Multifocality at baseline	12 (15.4)	8 (14.8)	0.999
AFP >400, n (%) (ng/mL)	7 (9.6)	3 (5.5)	0.525
Macrovascular invasion, n (%)	9 (11.5)	4 (7.4)	0.558
BCLC class, n (%)			
A	56 (71.8)	43 (79.6)	
B	13 (16.6)	7 (13.0)	0.690
C	9 (11.5)	4 (7.4)	
Edmondson–Steiner Grade, n (%)			
1	24 (30.8)	23 (42.6)	
2	30 (38.4)	20 (37.0)	0.293
3	24 (30.8)	11 (20.4)	
Treatment, n (%)			
Supportive care	9 (11.5)	6 (11.1)	
Liver transplant	8 (10.3)	4 (7.4)	
Resection	6 (7.7)	7 (13.0)	
TACE	14 (17.9)	13 (24.0)	0.475
RFA	8 (10.3)	5 (9.3)	
Sorafenib	13 (16.6)	4 (7.4)	
Sequential treatments	20 (25.6)	15 (27.8)	

AFP, α-fetoprotein; BCLC, Barcelona Clinic Liver Cancer; HCC, hepatocellular carcinoma; RFA, radiofrequency ablation; TACE, transarterial chemoembolisation.

Expression profile of HCC tissues from the training cohort

To determine whether genes were differentially expressed in relation with tumour growth, a discriminatory gene analysis was performed on normalised log2 gene expression values. Each tumour sample was individually compared with the combined group of NT samples. By doing so, we identified a number of genes that deviated from normality for multiple tumour samples, but not necessarily for all of them. We used these data to perform discrimination between the fastest-growing quartile versus the other slow-growing quartiles of the tumour samples. An in-between groups t test (unpaired, two-tailed, unequal variances) between the fast and the slow groups was performed. This analysis identified 86 genes exhibiting higher levels of expression and 157 genes exhibiting lower levels of expression in rapidly growing tumours relative to slow-growing tumours (see online supplementary table S2).

These genes were ranked on the basis of their predictive power for survival (univariate z score), with a negative score associated with longer overall survival and a positive score associated with shorter overall survival. The genes with an absolute

Table 2 Clinical characteristics of patients and HCC features at presentation according to fast (first quartile) or slow (other quartiles) growth speed

	Training cohort (n=78)			Validation cohort (n=54)		
Variable	Fast growing n=19	Slow growing n=59	p Value	Fast growing n=15	Slow growing n=39	p Value
Male, n (%)	15 (78.9)	46 (78.0)	1.000	9 (60.0)	30 (76.9)	0.214
Median age, y (range)	73 (45–82)	67 (44–88)	0.360	66 (45–76)	67 (32–84)	0.900
Aetiology, n (%)						
HCV	9 (47.4)	35 (59.3)		6 (54.5)	24 (61.6)	
HBV	1 (5.3)	9 (15.3)		1 (9.1)	5 (12.8)	
Alcohol	3 (15.8)	9 (15.3)	0.123	2 (18.2)	4 (10.3)	0.765
Dysmetabolic	6 (31.6)	6 (10.2)		2 (18.2)	6 (15.4)	
AFP >400 ng/mL	4 (23.5)	3 (5.4)	0.027	2 (13.3)	1 (2.6)	0.140
Performance status, n (%)						
0	18 (94.7)	56 (94.9)		14 (93.4)	37 (94.9)	
1	1 (5.3)	3 (5.0)	1.000	1 (6.6)	2 (5.1)	0.939
MELD	9.7±3.1	11.0±3.9	0.141	10.1±2.8	10.5±3.6	0.917
Tumour characteristics						
Mean doubling time (days)	40±6	128±96	<0.0001	39±11	210±167	<0.0001
Tumour volume (cm^3)	24.2±9.5	9.7±11.3	0.649	25.5±5.0	12.4±12.2	0.730
Multifocality at baseline	9 (47.4)	3 (5.1)	<0.0001	6 (54.5)	2 (5.3)	<0.0001
Macrovascular Invasion, n (%)	5 (26.3)	4 (6.8)	0.034	2 (20.0)	2 (5.4)	0.194
Edmondson–Steiner grading, n (%)						
1	1 (5.3)	23 (39.0)		2 (13.3)	21 (53.8)	
2	10 (52.6)	20 (33.9)	0.022	8 (53.3)	12 (30.8)	0.025
3	8 (42.1)	16 (27.1)		5 (33.3)	6 (15.4)	
BCLC class, n (%)						
A	8 (42.1)	48 (81.4)		12 (80.0)	31 (79.5)	
B	6 (31.6)	7 (11.9)	0.004	0	7 (17.9)	0.027
C	5 (26.3)	4 (6.8)		3 (20.0)	1 (2.6)	
Treatment, n (%)						
Supportive care	3 (15.8)	6 (10.2)		2 (14.3)	4 (10.3)	
Liver transplant	0	8 (13.6)		0	4 (10.3)	
Resection	3 (15.8)	3 (5.1)		2 (14.3)	5 (12.8)	
TACE	2 (10.5)	12 (15.3)	0.093	6 (15.8)	7 (17.9)	0.324
RFA	0	8 (13.6)		0	5 (12.8)	
Sorafenib	6 (31.6)	7 (11.9)		2 (5.3)	2 (5.1)	
Sequential treatments	5 (26.3)	15 (25.4)		3 (21.4)	12 (30.8)	
Biochemistry panel, M±SD						
Hb (g%)	12.9±2.1	13.3±1.5	0.466	12.6±1.9	13.2±1.8	0.655
Platelets ($\times10^3/mm^3$)	126±66	106±62	0.264	114±14	150±69	0.374
White blood cells ($\times10^3/mm^3$)	5.4±2.7	5.3±2.2	0.831	5.4±2.8	5.6±1.6	0.867
Blood glucose (mg/dL)	115±32	114±28	0.889	112±10	113±26	0.932
Cholesterol (mg/dL)	149±43	147±44	0.842	138±39	140±31	0.931
Blood iron (ng/mL)	79±46	127±70	0.022	121±92	106±59	0.420
Ferritin (ng/mL)	504±183	305±140	0.443	432±312	273±253	0.204
Albumin (g/dL)	3.6±0.7	3.6±0.5	0.344	3.3±0.6	3.3±0.6	0.905
Creatinine (mg/dL)	0.8±0.24	0.9±0.3	0.394	0.8±0.3	0.9±0.5	0.728
Bilirubin (mg/dL)	1.2±0.9	1.8±1.3	0.191	1.3±0.9	1.5±1.0	0.164
INR	1.2±0.3	1.3±0.1	0.752	1.2±0.1	1.3±0.3	0.337
AST (IU/mL)	75±49	94±79	0.331	106±77	61±27	0.396
ALT (IU/mL)	61±49	72±65	0.470	60±34	51±26	0.661
GGT (IU/mL)	128±138	132±188	0.926	195±205	108±109	0.402
ALP (IU/mL)	147±103	184±156	0.253	147±49	130±71	0.702
Na (mEq/L)	138±3.9	138±3.7	0.767	139±4.8	140±4.5	0.720

AFP, α-fetoprotein; ALP, alkaline phosphatase; ALT, alanine transaminase; AST, aspartate aminotransferase; BCLC, Barcelona Clinic Liver Cancer; GGT, gamma-glutamyl transpeptidase; Hb, haemoglobin; HCC, hepatocellular carcinoma; INR, International Normalised Ratio; MELD, Model For End-Stage Liver Disease; RFA, radiofrequency ablation; TACE, transarterial chemoembolisation.

univariate z score ±2.5 were MCM10, DLL4, NR4A1, NETO2, ANGPT2, ESM1, NCAPH. We then subjected these genes to a multivariate Cox proportional-hazards regression model, with overall survival as the dependent variable. The following genes were independently related to survival: ANGPT2, DLL4, NETO2, NR4A1, ESM1. The risk index was defined as

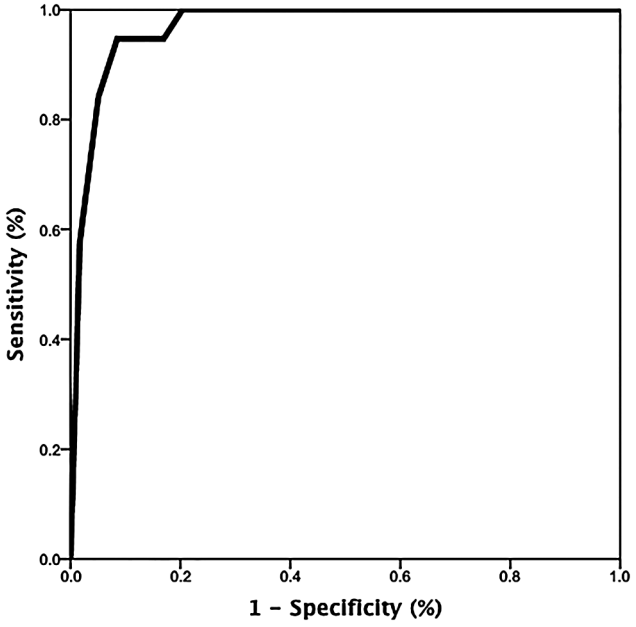

Figure 1 Receiver operating characteristic (ROC) curve analysis to evaluate the diagnostic power of the five-gene signature. The area under the ROC curves (AUCs) was analysed using the Hanley and McNeil method.[20] The AUC value was 0.961 (95% CI 0.919 to 1.000; p<0.0001).

a linear combination of the log2 gene expression values for the top genes identified by univariate Cox proportional hazard regression modelling weighted by their estimated regression coefficients (1.294×ANGPT2; 0.966×DLL4; 0.726×NETO2; 0.624×NR4A1; 0.557×ESM1). The distribution of risk index values calculated in the training set was examined to determine an appropriate cut-point to distinguish high and low risk. We examined a continuum of cut-points (ranging from the 50th to 80th percentile). The 70th percentile cut-point had the best discriminatory power. This five-gene risk signature identified rapidly growing tumours with high specificity and sensitivity (figure 1). Only one case with rapid growth did not bear the complete five-gene signature, while two of the second quartile did (one of these had 54 days as DT). None of the patients in the third and fourth quartile bore the five-gene signature.

Fast growth was associated at univariate analysis with aetiology of cirrhosis, Edmondson–Steiner grading, presence of macrovascular invasion, tumour volume at baseline, multifocality at baseline and five-gene risk signature. Only five-gene risk signature (OR 4.253; 95% CI 2.030 to 8.910, p<0.0001) was independently associated at multivariate logistic regression analysis (table 3).

Prediction of survival for the training cohort

No significant difference was present between quartiles regarding Child–Pugh class (p=0.330) or Model For End-Stage Liver Disease score (p=0.237) at the time of HCC diagnosis.

A total of 40/78 (51.2%) patients died during follow-up: 17/19 (98.4%) patients in the first quartile versus 23/59 (38.9%) patients in the other three quartiles (p<0.0001). The difference was significant for the first quartile versus each of the others (17/19 (89.4%) versus 8/20 (40.0%), 7/19 (36.8%) and 8/19 (42.1%), respectively (p=0.0013, p=0.0005 and p=0.0021, respectively)). Moreover, death due to neoplastic invasion of the liver occurred significantly more often in the first quartile (see online supplementary table S3) (p=0.001). The overall mean and median survival for the training cohort was 24.1±12.8 months and 38 months, respectively. The patients with rapidly growing tumours had mean and median survival periods of 14.0±10.4 months and 11 months, respectively, while the mean and median survival periods for patients in the second, third and fourth quartiles were 25.5±10.9 and 41 months, 26.3±11.1 and 42 months and 30.5±13.4 and 47 months, respectively. Correspondingly, Kaplan–Meier curve analysis of survival showed a significantly lower survival rate for HCC cases stratified by rapid growth (<53 days) compared with the other quartiles (figure 2A; p<0.0001) or by gene risk signature (figure 2C, p<0.0001).

Cox regression analysis identified the five-gene risk signature (HR: 3.987; 95% CI 1.941 to 8.193, p<0.0001), macrovascular invasion (HR: 3.885, 95% CI 1.491 to 10.123, p=0.005), treatment (HR: 0.460, 95% CI 0.213 to 0.997; p=0.049) and serum albumin levels (HR: 0.403, 95% CI 0.211 to 0.769, p=0.006) as independent risk factors for mortality (table 4).

Treatment was performed according to internationally accepted guidelines (1). Only nine patients received the best supportive care (three in the first quartile, one in the second, one in the third and four in the fourth, p=0.153). Treatment had a significant impact on survival for the training cohort (table 4). However, when treatment was stratified according to growth speed, patients with slow growth gained a significant

Table 3 Univariate and multivariate logistic regression analysis of baseline factors associated with rapid growth

	Univariate analysis		Multivariate analysis	
Variables	**OR (95% CI)**	**p Value**	**OR (95% CI)**	**p Value**
Gender*	0.944 (0.267 to 3.337)	0.928		
Age, years	1.029 (0.971 to 1.091)	0.326		
Aetiology	2.640 (0.902 to 7.731)	0.077	1.097 (0.696 to 1.731)	0.690
Edmondson–Steiner grading	2.126 (1.021 to 4.427)	0.044	1.506 (0.412 to 5.509)	0.536
Macrovascular Invasion	4.911 (1.164 to 20.722)	0.030	2.343 (0.136 to 15.763)	0.558
Tumour Volume at Baseline	1.557 (0.855 to 2.835)	0.148		
Multifocality at Baseline	4.865 (1.621 to 14.606)	0.005	3.533 (0.445 to 19.021)	0.232
Platelets (×10^3/mm^3)	1.001 (0.984 to 1.019)	0.880		
AFP (ng/mL)	1.002 (0.990 to 1.020)	0.811		
Five-gene risk signature	5.331 (2.522 to 11.268)	0.001	4.253 (2.030 to 8.910)	<0.0001

*Male gender used as reference.
AFP, α-fetoprotein.

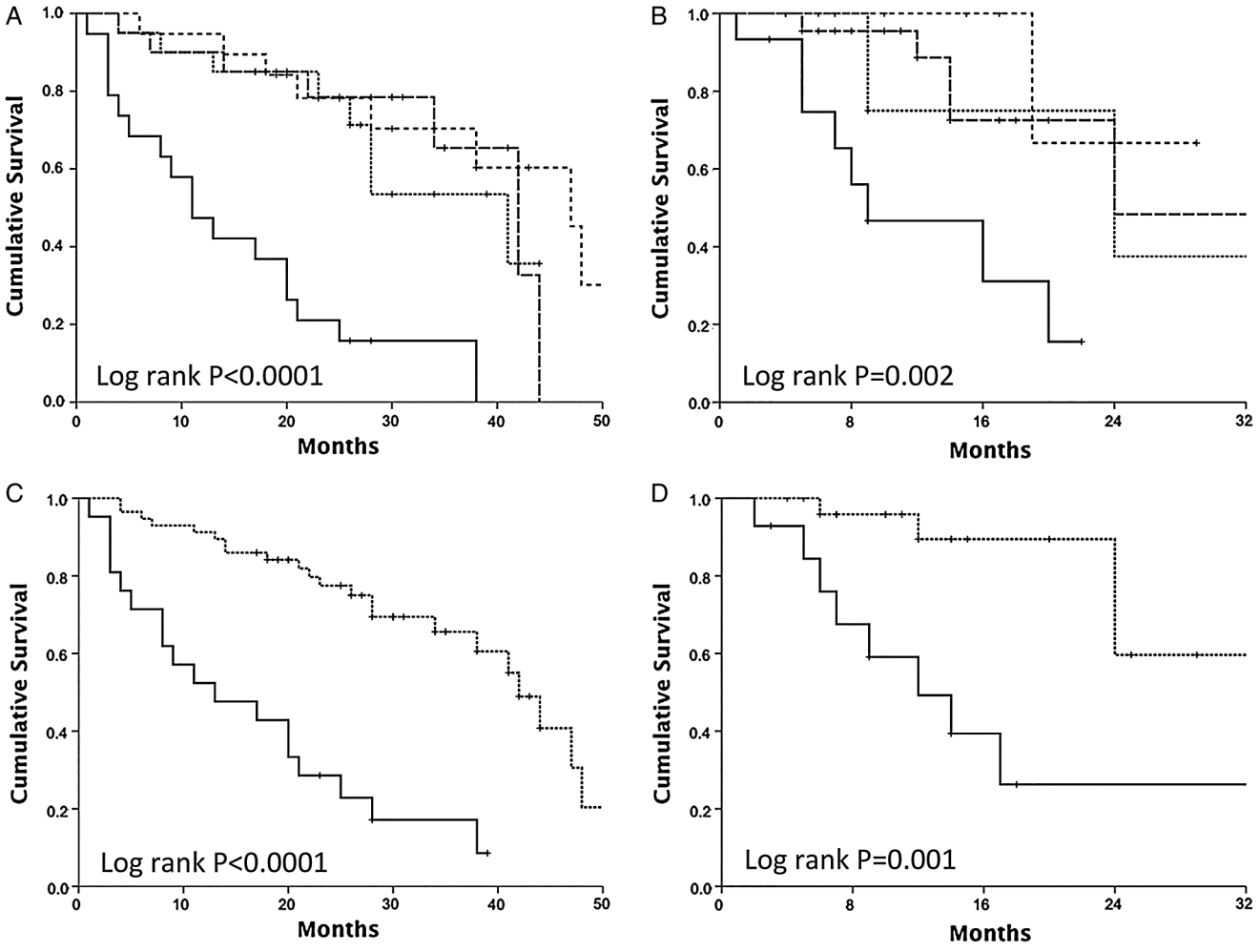

Figure 2 Kaplan–Meier plots for survival of patients with hepatocellular carcinoma (HCC). Patients were grouped according to quartiles of HCC growth (solid line: ≤53 days; dashed-dot line: 54–82 days; dashed line: 83–110 days; dotted line: ≥111 days) (A: training and B: validation cohort, respectively) or according to the five-gene risk signature (C: training and D: validation cohort) (solid line: high-risk signature; dotted line: low-risk signature).

advantage with treatment (median survival of no treatment vs treatment: 13 vs 42 months, respectively; p=0.017). In contrast, patients with fast-growing tumours had a marginal survival advantage that did not reach statistical significance (median survival of no treatment vs treatment: 5 vs 11 months, respectively; p=0.088).

Table 4 Cox regression analysis of baseline factors associated with mortality in the training cohort

	Univariate analysis		Multivariate analysis	
Variables	**HR (95% CI)**	**p Value**	**HR (95% CI)**	**p Value**
Gender*	0.687 (0.302 to 1.564)	0.371		
Age, years	0.989 (0.956 to 1.024)	0.540		
Aetiology	1.587 (0.845 to 2.980)	0.151		
Treatment (yes/no)†	0.919 (0.844 to 1.000)	0.051	0.460 (0. 0.213 to 0.997)	0.049
Edmondson–Steiner grading	1.443 (1.571 to 2.398)	0.036	1.438 (0.882 to 2.345)	0.145
Macrovascular invasion	4.818 (2.140 to 10.846)	<0.0001	3.885 (1.491 to 10.123)	0.005
Tumour volume at baseline	1.003 (0.518 to 1.944)	0.993		
Platelets ($\times 10^3/mm^3$)	0.586 (0.271 to 1.268)	0.586		
Albumin (g/dL)	0.492 (0.253 to 0.882)	0.017	0.403 (0.211 to 0.769)	0.006
Creatinine (mg/dL)	1.111 (0.515 to 2.397)	0.788		
Bilirubin (mg/dL)	1.069 (0.964 to 1.185)	0.207		
AFP (ng/mL)	1.153 (0.776 to 1.712)	0.481		
Five-gene risk signature	1.548 (1.296 to 1.849)	<0.0001	3.987 (1.941 to 8.193)	<0.0001

*Male gender used as reference.
†No treatment used as reference.
AFP, α-fetoprotein.

Validation of the five-gene risk signature

Univariate analysis identified number of nodules at baseline (OR: 3.839; 95% CI 1.363 to 10.814; p=0.011) and the five-gene signature (OR: 2.994; 95% CI 1.465 to 6.122; p=0.003) as independent predictors of rapid tumour growth. At multivariate analysis, only the five-gene signature was identified as an independent factor for rapid tumour growth (OR: 3.467; 95% CI 1.494 to 8.047; p=0.004).

Seventeen patients out of 54 (31.5%) died during follow-up, 8/15 (53.3%) in the fast-growing group and 9/39 (23.0%) in the slow-growing group (p=0.032). Death in the high-risk cluster was, as in the training cohort, more frequently due to HCC progression versus other causes (see online supplementary table S3, p=0.005). The patients with rapidly growing tumours had mean and median survival periods of 12.2±8.1 months and 9 months, respectively, while the mean and median survival periods for patients in the second, third and fourth quartiles were 26.8±10.9 and 24 months, 25.6±8.1 months and 30.3 ±27 and 24 months, respectively (figure 2B, p=0.002). Kaplan–Meier curve analysis of survival showed a significantly lower survival rate for HCC cases when stratified by gene risk signature (p=0.001) (figure 2D).

To assess the internal validity of the survival model the data-splitting and bootstrapping validation methods were performed. Univariate analysis found the significant prognostic factors for survival including Edmondson–Steiner grading (HR: 2.678; 95% CI 1.317 to 5.444, p=0.007), macrovascular invasion (HR: 5.149; 95% CI 1.378 to 19.241; p=0.082), treatment (HR: 0.127; 95% CI 0.034 to 0.474; p=0.002), albumin levels (HR: 0.214; 95% CI 0.62 to 0.742, p=0.015) and the five-gene signature (HR: 6.896; 95% CI:1.820–26.128; p=0.004). At multivariate analysis, only Edmondson–Steiner grading (HR: 4.489; 1.635 to 12.329, p=0.004), identified treatment (HR: 0.197; 95% CI 0.039 to 0.995; p=0.048), albumin levels (HR: 0.161; 95% CI 0.033 to 0.783, p=0.024) and the five-gene signature (HR: 5.798; 95% CI 1.510 to 22.260; p=0.010) were independent factors related to survival. A validation of the survival model carried out by bootstrapping identified the same prognostic factors as the training cohort plus Edmondson–Steiner grading (see online supplementary table S4).

DISCUSSION

In this prospective study of patients with compensated liver cirrhosis on US surveillance at first identification of HCC, we have shown that a five-gene hepatic transcriptomic signature (Angiopoietin-2, NETO2, DLL4, ESM1, NR4A1) is able to identify patients with extremely rapid tumour growth (ie, a DT of <53 days as determined by two CT scans performed at 6-week interval) and ominous prognosis (median survival of 11 months vs more than 41 months in patients with slowly growing tumours).

The biological characterisation of these fast-growing tumours is quite novel in comparison with other reported prognostic signatures.[5–11] Although many of these signatures were very informative in regard to prognosis, recurrence rate and survival, it should be underlined that all these studies used frozen or paraffin-embedded archival samples obtained at resection. This rather limits the generalisability to all HCCs as not more than 5% of them are suitable for resection, as degree of portal hypertension, number and size of lesions, macrovascular involvement, greatly restrict the indications for surgery.

In the present study, although many genes involved in cell cycle control and proliferation were upregulated, five (*ANGPT2*, *NETO2*, *ESM1*, *NR4A1*, and *DLL4*) that have roles in endothelial cell migration, angiogenesis and blood vessel morphogenesis were also related with survival. *ANGPT2* was the most significantly upregulated gene. Its product is secreted by endothelial cells at sites of active vascular remodelling.[22] Levels of ANGPT2 mRNA and protein have also been found to correlate with microvessel density[23] and highly vascular and poorly differentiated HCCs, respectively.[24 25] Correspondingly, the injection of an HCC cell line overexpressing angiopoietin-2 into nude mice resulted in faster tumour growth that was associated with greater vessel density.[26] Experimentally, its blockade provides an effective anti-angiogenic therapy.[27 28] The results of previous studies also suggest that functional relationship occurs among genes characterising the signature identified in this study. For example, *DLL4* is a vascular-specific ligand of Notch and plays a critical role in the angiogenesis of several types of tumours.[29 30] Under ischaemic conditions, upregulation of both *ANGPT-2/Tie2* and *DLL4* may represent a compensatory mechanism for local ischaemia, eventually favouring neo-angiogenesis.[31] *ESM1* is expressed by the vascular endothelium and participates in the regulation of cell adhesion, inflammatory disorders and tumour progression.[32] Moreover, in some experimental models of cancer, *ESM1* expression has been identified as one of the main switches for the induction of a dormant tumour to a rapidly growing tumour with increased angiogenesis.[33] *NR4A1* is a member of the nuclear orphan hormone receptor-1 family, is a direct target of vascular endothelial growth factor and is able to induce endothelial cell proliferation and migration.[34 35] *NETO2* encodes a transmembrane protein that is highly upregulated, along with *ANGPT2*, in proliferating infantile haemangiomas.[36] Thus, we hypothesise that the upregulation of these genes plays a pathogenic role in the rapid growth pattern of the tumours included in the first quartile of this study.

The five-gene signature, specific treatment and serum albumin emerged as independent predictors for survival in patients with liver cirrhosis and HCC both in training and in validation cohorts. Although only one of the three laboratory parameters that compose the Child–Pugh score (eg, albumin, bilirubin and International Normalised Ratio, INR) was identified as a significant factor in a multivariate analysis, this can be explained by the evidence that our patients with HCC represent very early cases identified on surveillance in whom bilirubin and INR values (indicators of more advanced stage of disease) were near normal. These results further confirm that in patients with cirrhosis and HCC, in addition to cancer-related features (eg, the five-gene signature), underlying liver function is relevant in risk modelling.

Impact of treatment on survival was clear-cut only in HCCs with slow growth as for patients with rapidly growing HCC, early detection of the tumour on US surveillance did not improve prognosis. In those with fast growth, further tumour growth was only marginally influenced by therapeutic interventions (p=0.088). None of these patients (irrespective of belonging to training or validation set) had the chance of being transplanted as, in the period between HCC discovery and listing for transplant, HCC had already exceeded transplant criteria. Considering that curative therapies should be offered to the sickest patients, there should be an attempt to verify whether extreme prioritisation for access to transplant for patients with the five-gene signature would be a successful strategy. However, the biological aggressiveness of these tumours could increase the risk of post-transplant recurrence. Indeed, in a retrospective evaluation of HCC recurrence after liver transplantation by Hanouneh *et al*,[37] 58% of patients with rapidly

growing HCC experienced recurrence. The reliability of the prognostic model described is strengthened by the validation performed in an independent series of patients with similar features. Nevertheless, larger prospective studies are still needed to confirm these results in order to obtain more conclusive arguments for treatment recommendation.

In conclusion, growth patterns of HCC were characterised by applying a five-gene transcriptomic signature at presentation. This high-risk signature identified a subgroup of patients with rapidly growing HCCs, which prevented their access to LT and indicated a poor prognosis. With this study, we provided evidence on the importance of evaluating HCCs statically (as simple number and size of nodules) and dynamically, that is, as growing lesions with extremely different growth patterns. Based on these results and in a more general sense for a better personalised management and therapy of HCC,[38] adding the prognostic information of the identified five-gene signature to clinical and radiological parameters already indicated by the current guidelines, would have significant implication for the therapeutic management of patients with HCC.

Author affiliations
[1]Division of Gastroenterology, AOU Modena, Modena, Italy
[2]Institute of Radiology, AOU Modena, Modena, Italy
[3]Division of Gastroenterology, DiBiMIS, University of Palermo, Palermo, Italy
[4]Institute of Internal Medicine, University of Bari, Bari, Italy
[5]Medical Clinic, University of Padua, Padua, Italy
[6]Dipartimento di Scienze Statistiche e Matematiche 'S. Vianelli', University of Palermo, Palermo, Italy
[7]Department of Internal Medicine, Clinical and Molecular Hepatology, University of Messina, Messina, Italy
[8]Department of Pathology, AOU Modena, Modena, Italy
[9]Metabolomic Unit, CicBioGune, Bilbao, Spain
[10]Medicina Metabolica, Nuovo Ospedale S. Agostino, Modena, Italy
[11]Department of Surgery, AOU Modena, Modena, Italy
[12]Liver Transplant, University of Padua, Padua, Italy

Contributors EV: conception and design. RC, ET, FM, LM and MLM-C: development of methodology. BL, SC, SC, SB, MDB and NDM: acquisition of data. EV, GG, GM, CC, GC, LL, NDM, PP, TP and UC: analysis and interpretation of data. ME: bootstrap analysis. EV, GEG, PL and FS: writing, review and/or revision of the manuscript. AK, EB and PT: technical support (constructing/managing database). EV: study supervision.

Funding This work was entirely supported by Regione Emilia-Romagna [Grants "PRU (Programma di ricerca Regione-Università) 2007–2009 and 2010–2012"]. GG was supported by grant number 11389 from Italian Association Cancer Research (AIRC). The opinions, results and conclusions reported in this paper are those of the authors and are independent from the funding sources.

Competing interests None.

Patient consent Obtained.

Ethics approval The Ethics Committee of Azienda Ospedaliero-Universitaria, Modena, approved the study protocol (IRB10/08_CE_UniRer; ClinicalTrials ID: NCT01657695).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement There are no additional data presently available to share; source data are available at the Gene Expression Omnibus website (http://www.ncbi.nlm.nih.gov/geo) under the accession number: GSE54236.

REFERENCES

1 Bruix J, Sherman M; American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma: an update. *Hepatology* 2011;53:1020–2.
2 European Association for the Study of the Liver; European Organisation for Research and Treatment of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2012;56:908–43.
3 Marrero JA, Fontana RJ, Barrat A, *et al*. Prognosis of hepatocellular carcinoma: comparison of 7 staging systems in an American cohort. *Hepatology* 2005;41:707–16.
4 Cammà C, Di Marco V, Cabibbo G, *et al*. Survival of patients with hepatocellular carcinoma in cirrhosis: a comparison of BCLC, CLIP and GRETCH staging systems. *Aliment Pharmacol Ther* 2008;28:62–75.
5 Lee JS, Heo J, Libbrecht L, *et al*. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med* 2006;12:410–16.
6 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.
7 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.
8 Kim SM, Leem SH, Chu IS, *et al*. Sixty-five gene-based risk score classifier predicts overall survival in hepatocellular carcinoma. *Hepatology* 2012;55:1443–52.
9 Woo HG, Park ES, Cheon JH, *et al*. Gene expression-based recurrence prediction of hepatitis B virus-related human hepatocellular carcinoma. *Clin Cancer Res* 2008;14:2056–64.
10 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.
11 Ye QH, Qin LX, Forgues M, *et al*. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nature medicine* 2003;9:416–23.
12 Villa E, Dugani A, Fantoni E, *et al*. Type of estrogen receptor determines response to antiestrogen therapy. *Cancer Res* 1996;56:3883–5.
13 Villa E, Moles A, Ferretti I, *et al*. Natural history of inoperable hepatocellular carcinoma: estrogen receptors' status in the tumor is the strongest prognostic factor for survival. *Hepatology* 2000;32:233–8.
14 Yopp AC, Mansour JC, Beg MS, *et al*. Establishment of a multidisciplinary hepatocellular carcinoma clinic is associated with improved clinical outcome. *Ann Surg Oncol* 2014;21:1287–95.
15 Singal AG, Waljee AK, Patel N, *et al*. Therapeutic delays lead to worse survival among patients with hepatocellular carcinoma. *J Natl Compr Canc Netw* 2013;11:1101–8.
16 International Working Party. Terminology of nodular hepatocellular lesions. *Hepatology* 1995;22:983–93.
17 Schwartz M. A biomathematical approach to clinical tumor growth. *Cancer* 1961;14:1272–94.
18 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PC and the $2^{-\Delta\Delta C}$T method. *Methods* 2001;25:402–8.
19 Lossos IS, Czerwinski DK, Alizadeh AA, *et al*. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *N Engl J Med* 2004;350:1828–37.
20 Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839–43.
21 Efron B. How biased is the apparent error rate of a prediction rule? *J Am Stat Assoc* 1986;81:461–70.
22 Fiedler U, Scharpfenecker M, Koidl S, *et al*. The Tie-2 ligand angiopoietin-2 is stored in and rapidly released upon stimulation from endothelial cell Weibel-Palade bodies. *Blood* 2004;103:4150–6.
23 Moon WS, Rhyu KH, Kang MJ, *et al*. Overexpression of VEGF and angiopoietin 2: a key to high vascularity of hepatocellular carcinoma? *Mod Pathol* 2003;16:552–7.
24 Tanaka S, Mori M, Sakamoto Y, *et al*. Biologic significance of angiopoietin-2 expression in human hepatocellular carcinoma. *J Clin Invest* 1999;103:341–5.
25 Sugimachi K, Tanaka S, Taguchi K, *et al*. Angiopoietin switching regulates angiogenesis and progression of human hepatocellular carcinoma. *J Clin Pathol* 2003;5:854–60.
26 Peng L, Sun J, Wang WD, *et al*. Biological effect of ectopic expression of angiopoietin-1 and -2 in hepatocellular carcinoma cell line. *Hepatobiliary Pancreat Dis Int* 2003;2:94–7.
27 Mazzieri R, Pucci F, Moi D, *et al*. Targeting the ANG2/TIE2 axis inhibits tumor growth and metastasis by impairing angiogenesis and disabling rebounds of proangiogenic myeloid cells. *Cancer Cell* 2011;19:512–26.
28 Lewis CE, Ferrara N. Multiple effects of angiopoietin-2 blockade on tumors. *Cancer Cell* 2011;19:431–3.
29 Indraccolo S, Minuzzo S, Masiero M, *et al*. Cross-talk between tumor and endothelial cells involving the Notch3-Dll4 interaction marks escape from tumor dormancy. *Cancer Res* 2009;69:1314–23.
30 Ziyad S, Iruela-Arispe ML. Molecular mechanisms of tumor angiogenesis. *Genes Cancer* 2011;2:1085–96.
31 Skuli N, Majmundar AJ, Krock BL, *et al*. Endothelial HIF-2α regulates murine pathological angiogenesis and revascularization processes. *J Clin Invest* 2012;122:1427–43.
32 Sarrazin S, Adam E, Lyon M, *et al*. Endocan or endothelial cell specific molecule-1 (ESM-1): a potential novel endothelial cell marker and a new target for cancer therapy. *Biochim Biophys Acta* 2006;1765:25–37.

33 Almog N, Ma L, Raychowdhury R, *et al*. Transcriptional switch of dormant tumors to fast-growing angiogenic phenotype. *Cancer Res* 2009;69:836–44.
34 Vacca M, Murzilli S, Salvatore L, *et al*. Neuron-derived orphan receptor 1 promotes proliferation of quiescent hepatocytes. *Gastroenterology* 2013;144:1518–29.
35 Akeson A, Herman A, Wiginton D, *et al*. Endothelial cell activation in a VEGF-A gradient: relevance to cell fate decisions. *Microvasc Res* 2010;80:65–74.
36 Calicchio ML, Collins T, Kozakewich HP. Identification of signaling systems in proliferating and involuting phase infantile hemangiomas by genome-wide transcriptional profiling. *Am J Pathol* 2009;174:1638–49.
37 Hanouneh IA, Macaron C, Lopez R, *et al*. Rate of tumor growth predicts recurrence of hepatocellular carcinoma after liver transplantation in patients beyond Milan or UCSF criteria. *Transplant Proc* 2011;43:3813–18.
38 Giannelli G, Rani B, Dituri F, *et al*. Moving towards personalised therapy in patients with hepatocellular carcinoma: the role of the microenvironment. *Gut* 2014;63:1668–76.