

patients (13). Thus, drug sensitivity appears to be one of the major determinants of the prognosis of advanced HCC patients treated with chemotherapy. Therefore, a hallmark of successful treatment would be the identification of useful biomarkers for determining the survival benefits offered by each treatment strategy.

In this study, we investigated the gene expression profiles of HCCs using serial analysis of gene expression (SAGE) to identify novel molecular markers or targets for the treatment of HCC (14–18). Here, we identified the upregulation of the *DUT* gene that encodes dUTP pyrophosphatase (dUTPase) in HCC. Markedly, HCC with a high nuclear dUTPase expression correlated with a poorly differentiated morphology and a poor prognosis. *DUT* gene knockdown not only suppressed cell proliferation but also sensitized HuH7 cells to low-dose 5-FU.

## Materials and methods

### Samples

All HCC tissues, adjacent non-cancerous liver tissues and normal liver tissues were obtained from 110 patients undergoing a hepatectomy between 1997 and 2006 in Kanazawa University Hospital, Kanazawa, Japan. Five normal liver tissue samples were obtained from patients undergoing surgical resection of the liver for the treatment of metastatic colon cancer. These samples were snap-frozen in liquid nitrogen immediately after resection. One hundred and five HCC and surrounding non-cancerous liver samples were obtained from patients undergoing surgical resection of the liver for HCC treatment, and part of these samples were used for the recent study (19). Three HCC and adjacent non-cancerous liver tissue samples were snap-frozen in liquid nitrogen and later used for SAGE. Twenty HCC tissues and their corresponding non-cancerous liver tissues were also snap-frozen and later used for real-time reverse transcription-polymerase chain reaction (RT-PCR) analysis, as described previously (19). Eighty-two additional HCC samples were formalin-fixed, paraffin-embedded and used for immunohistochemistry (IHC). HCC and adjacent non-cancerous liver tissues were histologically characterized, as reported elsewhere (19).

All strategies used for gene expression analysis as well as tissue acquisition processes were approved by the Ethics Committee and the Institutional Review Board of Kanazawa University Hospital. All procedures and risks were explained verbally to each patient, who then provided written informed consent.

### Serial analysis of gene expression

Total RNA was purified from each homogenized tissue sample using a ToTally RNA extraction kit (Ambion Inc., Austin, TX, USA), and polyadenylated RNA was isolated using a MicroPoly (A) Pure kit (Ambion). A total of 2.5 µg of mRNA per sample was analysed by SAGE (20, 21). SAGE libraries were randomly sequenced at the

Genomic Research Center (Shimadzu-Biotechnology, Kyoto, Japan), and the sequence files were analysed with SAGE 2000 software. The size of each SAGE library was normalized to 300 000 transcripts per library, and the abundance of transcripts was compared with SAGE 2000 software. Monte Carlo simulation was used for selecting genes whose expression levels were significantly different between the two libraries (22). Each SAGE tag was annotated using a gene-mapping website SAGE Genie database (<http://cgap.nci.nih.gov/SAGE/>) and the Source database (<http://smd.stanford.edu/cgi-bin/source/sourceSearch>), as described previously (23).

### Quantitative reverse transcription-polymerase chain reaction

A 1 µg aliquot of each total RNA was reverse-transcribed using SuperScript II reverse-transcriptase (Invitrogen, Carlsbad, CA, USA). Real-time RT-PCR analysis was performed using the ABI PRISM 7700 sequence detection system (Applied Biosystems, Foster City, CA, USA). Using the standard curve method, quantitative PCR was performed in duplicate for each sample-primer set. Each sample was normalized relative to β actin. The assay IDs used were Hs00798995\_s1 for dUTPase and Hs99999903\_m1 for β actin.

### RNA interference targeting *DUT*

Small interfering RNAs (siRNAs) targeting *DUT* or control (scrambled sequence) were synthesized by Dharmacon (Dharmacon Research Inc., Lafayette, CO, USA). The target sequences of *DUT* are 5'-AAGUUGU GAAAACGGACAUUC-3' (*DUT*1) and 5'-CGGACAUU CAGAUAGCGCUTT-3' (*DUT*2). Lipofectamine 2000™ reagent (Invitrogen) was used for transfection according to the manufacturer's instructions.

### Cell proliferation assay, soft agar assay and matrigel invasion assay

Cell proliferation assays were performed using a Cell Titer96 Aqueous kit in quintuplicate (Promega, Madison, WI, USA). For the soft agar assay,  $1 \times 10^4$  cells were suspended in 2 ml of 0.36% agar with growth medium and added in each well of a six-well plate containing a base layer of 0.72% agar. The plates were incubated at 37 °C in a 5% CO<sub>2</sub> incubator for 2 weeks. Matrigel invasion assays were performed using BD BioCoat™ Matrigel Matrix Cell Culture Inserts and Control Inserts (BD Biosciences, San Jose, CA, USA), as described in the manufacturer's instruction. 5-FU was obtained from Kyowa Kirin (Kyowa Kirin, Tokyo, Japan). All experiments were repeated at least twice.

### Immunohistochemistry

Mouse monoclonal anti-dUTPase antibody M01 (Abnova Corporation, Taipei, Taiwan) and mouse antiproliferating

cell nuclear antigen (PCNA) monoclonal antibody PC10 (Calbiochem, San Diego, CA, USA) were used to evaluate the immunoreactivity of HCC and adjacent non-cancerous liver samples using a Dako EnVision+™ kit (Dako, Carpinteria, CA, USA), according to the manufacturer's instruction. Immunoreactivity was evaluated by determining the percentage of cells expressing dUTPase in the examined fields, graded as low (0–50%) or high (> 50%). The PCNA index was evaluated as described previously (19).

#### Statistical analysis

Student's *t*-test was used to determine the statistical significance of the differences in cell viability between the two groups. The Mann–Whitney *U*-test was used for the analysis of gene expression between chronic liver disease (CLD) and HCC tissues. The  $\chi^2$ -test was used to evaluate the correlation between clinicopathological characteristics and dUTPase expression status. Univariate and multivariate Cox proportional hazards regression analysis was used to evaluate the association of dUTPase expression and clinicopathological parameters with patient outcome. All statistical analyses were performed using SPSS software (SPSS software package; SPSS Inc., Chicago, IL, USA) and GRAPHPAD PRISM software (GraphPad Software Inc., La Jolla, CA, USA).

## Results

### Gene expression profiling identified the overexpression of *DUT* in hepatocellular carcinoma

To overcome the considerable individual variability of transcriptomic characteristics, we constructed a SAGE library of normal human liver using RNAs derived from five normal liver tissues. In addition, we constructed two SAGE libraries derived from three HCC tissues or corresponding non-cancerous liver tissues from patients who developed HCC with a history of chronic hepatitis C. We detected a total of 226 267 tags corresponding to 45 746 unique tags from these SAGE libraries (supporting information Table S1). After excluding the tags detected only once in each library, we selected 15 333 reliable unique transcripts expressed in at least one of the SAGE libraries to avoid contamination of tags derived from sequence errors. Then, we annotated these transcripts using SAGE Genie database and the Source database to identify the potential subcellular localization of transcripts categorized into eight groups in each SAGE library.

The number of nuclear component-related transcripts was increased in the HCC library compared with the normal liver and non-cancerous liver libraries, whereas the other cellular component-related transcripts did not show this tendency (supporting information Fig. S1). Because nuclear component-related genes may closely correlate with cancer cell proliferation and chemosensitivity (24), we further investigated the expression of nuclear component-related tags in

each library, and identified 10 transcripts associated with nucleotide/nucleoside metabolism that are overexpressed in HCC (Table 1). Using Monte Carlo simulation, we evaluated the significance of differentially expressed transcripts in HCC and corresponding CLD libraries or in HCC and normal liver libraries. We identified a *DUT* gene encoding dUTPase (dUTPase) whose expression was significantly altered ( $P=0.01$ ). We also identified a *TS* gene encoding thymidylate synthase in the list, but the difference did not reach statistical significance.

dUTPase is a phosphatase known to maintain a dUMP pool by catalysing the hydrolysis of dUTP to dUMP, and thus provides a substrate of thymidylate synthase. Its role in HCC is unknown; therefore, we examined *DUT* expression in 20 independent HCC and corresponding non-cancerous liver tissues, and identified significant overexpression of *DUT* in HCC tissue ( $P=0.0015$ ) (Fig. 1A). Moreover, we detected more than a two-fold increase in *DUT* expression in 70% of HBV-related and HCV-related HCC cases (14 of 20 HCCs) compared with the non-cancerous liver tissues (Fig. 1B). We further examined the expression of *DUT* in 238 HCC tissues compared with the non-cancerous liver tissues using publicly available microarray data (GSE5975) (Fig. S2). Consistent with the SAGE data, *DUT* was overexpressed more than two-fold in 121 of 238 HCC tissues (median: 2.03), whereas *TS* was overexpressed more than two-fold in 54 of 238 HCC tissues (median: 1.41) compared with the non-cancerous liver tissues.

### Pivotal role of dUTP pyrophosphatase expression in cell proliferation in hepatocellular carcinoma cell lines

In general, cancer gene signatures discovered by comparison between tumour and non-tumour tissues are more likely to reflect the differences in the control of cell proliferation and growth (25). Accordingly, we investigated the function of dUTPase in cell proliferation in HuH7 cells by *DUT* gene knockdown. *DUT* expression was decreased by 60–70% following the transfection of the siRNA constructs specifically targeting *DUT* 48 h after transfection (*DUT*1 in Fig. 2A and *DUT*2 in Fig. S3A), and cell growth was significantly inhibited compared with the control 72 h after transfection (Fig. 2B and Fig. S3B). Anchorage-independent cell growth was also significantly impaired by *DUT* gene knockdown 14 days after transfection (Fig. 2C). Furthermore, *DUT* gene knockdown decreased the numbers of both migrating and invading cells 72 h after transfection (Fig. 2D and E).

dUTPase is known to be associated with thymidylate synthesis (26), and thus we evaluated the effects of 5-FU, a thymidylate synthase inhibitor, on dUTPase expression in HCC cell lines *in vitro*. When we treated HuH7 cells with low-dose 5-FU (0.25 mg/ml), we could not detect any growth-inhibitory effects (Fig. 2F). Based on this condition, we evaluated the effect of *DUT* gene knockdown on 5-FU sensitivity 72 h after transfection.

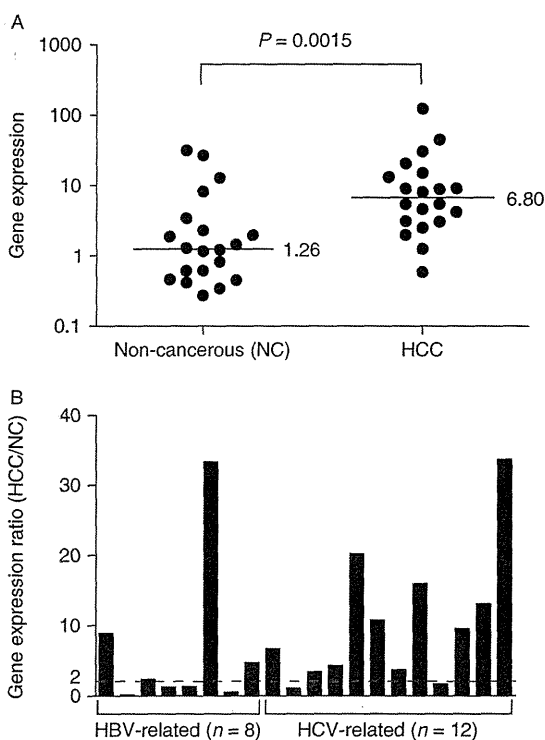
**Table 1.** Genes associated with nucleic acid metabolism overexpressed in hepatocellular carcinoma

Tag sequence	Normal liver	Non-cancerous liver	HCC	Fold*	Gene	P-value†
CAGCTCCGCT	0	2	11	5.5	dUTP pyrophosphatase	0.010
AAAGGATAAT	0	0	3	> 3	General transcription factor II H, polypeptide 2	0.127
ACGGTCCAGG	0	0	3	> 3	Cytidine deaminase	0.127
ATGTAGAGTG	0	0	3	> 3	Thymidylate synthase	0.127
TGGGGATTAC	1	0	3	> 3	Zinc ribbon domain containing, 1	0.127
CACCCTGTAC	2	2	6	3	Solute carrier family 29	0.147
GAACGCCTAA	1	1	3	3	Dihydropyrimidinase-like 2	0.308
GCGCTGGTAC	0	1	3	3	2'-5'-oligoadenylate synthetase 3	0.308
CTTAGTGCAA	0	2	4	2	3'-phosphoadenosine 5'-phosphosulphate synthase 2	0.335
TTGITACATC	0	2	3	1.5	Phosphoribosyl pyrophosphatase synthetase-associated protein 1	0.506

\*Fold increase was calculated by dividing the number of tags in HCC by that of tags in non-cancerous liver. To avoid division by 0, a tag value of 1 was used for any tag that was not detectable in one sample.

†Statistical significance of differentially expressed genes between two groups (HCC and non-cancerous liver libraries) was calculated using Monte Carlo simulation.

HCC, hepatocellular carcinoma.

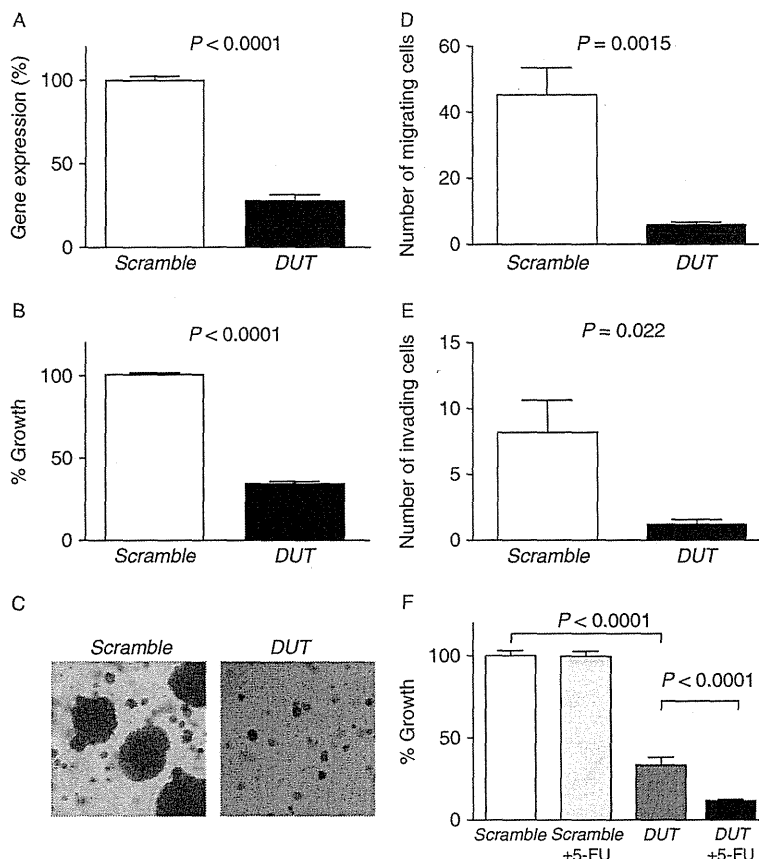


**Fig. 1.** (A) Quantitative reverse transcription-polymerase chain reaction analysis of *DUT* expression in hepatocellular carcinoma (HCC) and corresponding non-cancerous liver tissues. *DUT* was significantly activated in HCC tissues compared with non-cancerous liver tissues ( $P=0.0015$ ). A median value in each group is indicated. (B) *DUT* gene expression ratios of HCC and corresponding non-cancerous liver tissues. Fourteen of 20 HCC tissues expressed *DUT* more than two-fold compared with the background non-cancerous liver tissues. HBV, hepatitis B virus; HCV, hepatitis C virus.

Interestingly, *DUT* gene knockdown not only suppressed cell proliferation but also sensitized HuH7 cells to low-dose 5-FU (Fig. 2F and Fig. S3B). These data suggest that dUTPase overexpression in HCC tissues may be associated with enhanced cell proliferation and 5-FU resistance.

Intense dUTP pyrophosphatase expression is correlated with a poor prognosis in hepatocellular carcinoma patients

To characterize the clinicopathological characteristics of dUTPase expression in HCC, we performed IHC using an additional independent HCC cohort. Accordingly, we explored the dUTPase expression in HCC using 82 formalin-fixed paraffin-embedded HCC specimens. All HCC tissues were surgically resected at the Liver Disease Center of Kanazawa University Hospital with full clinical information, and their immunoreactivity to anti-dUTPase antibodies was evaluated by IHC. We noticed that anti-dUTPase antibodies reacted to both nuclear (red arrows) and cytoplasmic (blue arrows) isoforms of dUTPase, as described previously (26) (Fig. 3A and B). We therefore evaluated the nuclear and cytoplasmic expression of dUTPase separately. We stratified HCC tissues and evaluated the dUTPase expression status based on the percentages of dUTPase-positive cells. The frequency of nuclear or cytoplasmic dUTPase-positive cells was highly variable in each HCC tissue, and we defined HCCs with nuclear or cytoplasmic dUTPase expressed in  $\geq 50\%$  of tumour cells as nuclear or cytoplasmic dUTPase-high HCC (Fig. 3C). Nuclear dUTPase overexpression was detected in 36.6% (30 of 82), whereas cytoplasmic dUTPase overexpression was detected in 67.1% (55 of 82) of HCC tissues compared with the corresponding non-cancerous liver tissues

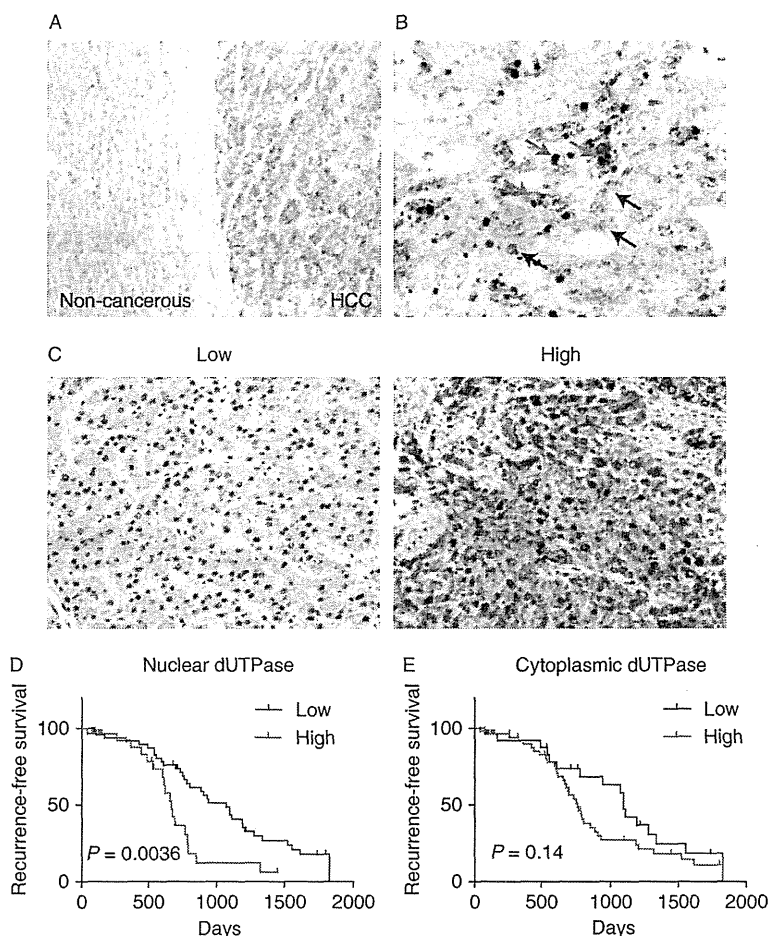


**Fig. 2.** (A) Transfection of small interfering RNAs targeting *DUT* (*DUT1*) decreased *DUT* expression compared with the control (scrambled sequence). Gene expression was evaluated in triplicate 72 h after transfection (mean  $\pm$  SD). (B) *DUT* gene knockdown significantly suppressed cell proliferation ( $P < 0.0001$ ). Cell viability was evaluated in triplicate 72 h after transfection (mean  $\pm$  SD). (C) Soft agar assay. *DUT* gene knockdown suppressed anchorage-independent cell growth. (D and E) Matrigel invasion assay. *DUT* gene knockdown decreased the numbers of both migrating and invading cells. Experiments were performed in triplicate (mean  $\pm$  SD). (F) *DUT* gene knockdown sensitized HuH7 cells to low-dose 5-fluorouracil (5-FU) (0.25  $\mu$ g/ml), which had no effect on the cell proliferation in the control (mean  $\pm$  SD).

(Table 2). In general, non-cancerous hepatocytes rarely expressed nuclear dUTPase (Fig. 3A).

We investigated the clinicopathological characteristics of nuclear or cytoplasmic dUTPase in low/high HCC cases (Table 2). The expression status of nuclear dUTPase showed no correlation with age, gender, virus, presence of cirrhosis,  $\alpha$ -fetoprotein value, tumour size and TNM stages. However, nuclear dUTPase expression was significantly correlated with the histological grades of HCC ( $P = 0.0099$ ), and high frequencies of nuclear dUTPase-positive cells were associated with poorly differentiated cell morphology in the HCC tissue. In contrast, cytoplasmic dUTPase expression was not correlated with the histological grades of HCC ( $P = 0.077$ ). We examined the cell proliferation of these HCC samples by PCNA staining, and PCNA indexes were significantly higher in nuclear dUTPase high HCC than low HCC with statistical significance ( $P = 0.01$ ) (Fig. S4).

We further investigated the prognostic significance of dUTPase expression in HCC. Strikingly, high nuclear dUTPase expression in HCC tissue correlated with a poor survival outcome compared with low nuclear dUTPase expression ( $P = 0.0036$ ), whereas high cytoplasmic dUTPase expression had little effects when evaluated by recurrence-free survival (Fig. 3D). Furthermore, univariate Cox regression analysis showed a significant correlation between high nuclear dUTPase expression and a high risk of mortality (HR, 2.47; 95% CI, 1.08–5.66;  $P = 0.032$ ; Table 3). By multivariate Cox regression analysis, TNM stage (HR, 2.75; 95% CI, 1.11–6.79;  $P = 0.027$ ) and nuclear dUTPase (HR, 2.61; 95% CI, 1.13–6.05;  $P = 0.024$ ) were independent prognostic factors associated with a high risk of mortality, and other clinicopathological features did not add independent prognostic information. These data indicate a significant correlation between the malignant potential of



**Fig. 3.** Immunohistochemistry analysis of dUTP pyrophosphatase (dUTPase) expression in hepatocellular carcinoma (HCC). (A) A representative photomicrograph of dUTPase staining in an HCC and adjacent non-cancerous liver tissue. (B) A representative photomicrograph of dUTPase staining in an HCC. Both nuclear (red arrows) and cytoplasmic (blue arrows) forms of dUTPase were detected. (C) Representative photomicrographs of HCC tissues with low (0–50%) and high ( $\geq 50\%$ ) frequencies of nuclear and cytoplasmic dUTPase-positive cells. (D and E) Kaplan–Meier survival analysis of HCC tissues with nuclear (D) or cytoplasmic (E) dUTPase expression. High percentages of nuclear dUTPase-positive tumour cells significantly correlated with poor clinical outcome in recurrence-free survival.

HCC and nuclear dUTPase expression, implicating the potential effectiveness of nuclear dUTPase level as a biomarker for predicting the survival of HCC patients after surgical resection.

### Discussion

Here, using a global gene expression profiling approach (18), we have identified the activation of the nucleotide/nucleoside metabolism-related gene *DUT* (encoding dUTPase) in HCC. Notably, an intense dUTPase expression was detected in a subset of HCC with a poor prognosis. To the best of our knowledge, this is the first

report describing the correlation between dUTPase activation and poor survival outcome in HCC patients.

In normal cells, dUTPase is known to catalyse the hydrolysis of dUTP to dUMP in order to maintain the dUMP pool at a certain level for thymidylate synthesis (26). Interestingly, dUTPase mutations in *Escherichia coli* increased dUTP levels, leading to dUTP misincorporation into DNA during replication, which resulted in DNA fragmentation and apoptosis (27). Furthermore, introduction of *E. coli* dUTPase into human tumour cells resulted in the induction of resistance to fluorodeoxyuridine cytotoxicity (28), suggesting a pivotal role of dUTPase in the prevention of DNA damage. Thus, dUTPase activation in the nucleus appears to be critical

**Table 2.** Clinicopathological characteristics and dUTP pyrophosphatase expression in hepatocellular carcinoma (n = 82)

dUTPase expression (nuclear)	Low (n = 52)	High (n = 30)	P-value
Age (< 60 years/≥ 60 years)	19/33	8/22	0.36
Sex (male/female)	36/16	23/7	0.47
Virus (HBV/HCV/B+C/NBNC)	15/33/1/3	10/20/0/0	0.48
Cirrhosis (yes/no)	33/19	22/8	0.36
AFP (< 20 ng/ml/≥ 20 ng/ml)	32/20	15/15	0.31
Histological grade*			
I-II	14	3	
II-III	36	20	
III-IV	2	7	0.0099
Tumour size (< 3 cm/≥ 3 cm)	31/21	19/11	0.74
TNM classification† (I, II/III, IV)	43/9	25/5	0.94

dUTPase expression (cytoplasmic)	Low (n = 27)	High (n = 55)	P-value
Age (< 60 years/≥ 60 years)	10/17	17/38	0.58
Sex (male/female)	19/8	40/15	0.82
Virus (HBV/HCV/B+C/NBNC)	8/17/1/1	17/36/0/2	0.56
Cirrhosis (yes/no)	17/10	38/17	0.58
AFP (< 20 ng/ml/≥ 20 ng/ml)	16/11	31/24	0.80
Histological grade*			
I-II	7	10	
II-III	20	36	
III-IV	0	9	0.077
Tumour size (< 3 cm/≥ 3 cm)	17/10	33/22	0.80
TNM classification† (I, II/III, IV)	21/6	47/8	0.39

\*Edmondson–Steiner grades.

†UICC TNM classification of liver cancer, 6th edition (2002).

AFP, α-fetoprotein; dUTPase, dUTP pyrophosphatase; HBV, hepatitis B virus; HCV, hepatitis C virus.

for preventing DNA damage possibly at the S phase. Specifically, this activation may prevent dUTP misincorporation in various cancers and thus avert DNA damage and apoptosis induction. Indeed, dUTPase activation has recently been reported in colorectal and brain cancer (29, 30), and dUTPase accumulation might correlate with 5-FU-based chemotherapy resistance and poor prognosis in colorectal cancer (26).

If dUTPase activation plays a central role in the development of resistance to thymidylate synthase inhibitors in order to prevent a DNA damage response, dUTPase inhibition may facilitate the eradication of cancer cells by sensitizing these cells to such inhibitors. Indeed, a recent study suggested a drastic sensitization of colon cancer cells to 5-FU by siRNAs-mediated dUTPase suppression (31, 32), which is consistent with our current observation. Because all HCC samples used in this study were surgically resected, we could not evaluate the effect of dUTPase expression on clinical HCC patients' outcome in relation to chemosensitivity to thymidylate synthase inhibitors. Nevertheless, intense nuclear dUTPase expression may be a good biomarker

**Table 3.** Cox regression analysis of recurrence-free survival rate relative to dUTP pyrophosphatase expression and clinicopathological parameters (n = 82)

Variables (n)	Univariate		Multivariate	
	HR (95% CI)	P-value	HR (95% CI)	P-value
Child–Pugh				
A	1			
B	1.73 (0.50–5.97)	0.38		
Tumour size				
< 3 cm (n = 50)	1			
≥ 3 cm (n = 32)	1.58 (0.69–3.63)	0.28		
TNM stage*				
I, II (n = 68)	1		1	
III, IV (n = 14)	2.57 (1.05–6.29)	0.039	2.75 (1.11–6.79)	0.027
Serum AFP				
< 20 ng/ml (n = 49)	1			
≥ 20 ng/ml (n = 38)	1.54 (0.66–3.56)	0.31		
Microvascular invasion				
No	1			
Yes	1.98 (0.89–4.44)	0.095		
BCLC stage				
A	1			
B/C	2.16 (0.93–5.00)	0.07		
Cytoplasmic dUTPase				
Low (n = 27)	1			
High (n = 55)	1.15 (0.50–2.62)	0.73		
Nuclear dUTPase				
Low (n = 52)	1		1	
High (n = 30)	2.47 (1.08–5.66)	0.032	2.61 (1.13–6.05)	0.024

\*UICC TNM classification of liver cancer, 6th edition (2002).

AFP, α-fetoprotein; CI, confidence intervals; dUTPase, dUTP pyrophosphatase; HR, hazard ratio.

for predicting the response to thymidylate synthase inhibitors, and its usefulness should be further evaluated in the future.

In conclusion, comprehensive gene expression profiling shed new light on the role of dUTPase in HCC. Nuclear dUTPase accumulation is potentially a good biomarker for predicting poor prognosis in HCC patients, and the development of a dUTPase inhibitor may promote the possibility of tumour eradication in HCC patients.

### Acknowledgements

The authors would like to thank Ms Masayo Baba and Nami Nishiyama for technical assistance. This research was supported in part by a Grant-in-Aid for Special Purposes from the Ministry of Education, Culture, Sports, Science and Technology, Japan (no. 20599005).

**Grant support:** Grant-in-Aid for Special Purposes from the Ministry of Education, Culture, Sports, Science and Technology, Japan (no. 20599005).

## References

- Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005; **55**: 74–108.
- El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 2007; **132**: 2557–76.
- Farazi PA, Depinho RA. Hepatocellular carcinoma pathogenesis: from genes to environment. *Nat Rev Cancer* 2006; **6**: 674–87.
- Roessler S, Budhu A, Wang XW. Future of molecular profiling of human hepatocellular carcinoma. *Future Oncol* 2007; **3**: 429–39.
- El-Serag HB, Marrero JA, Rudolph L, Reddy KR. Diagnosis and treatment of hepatocellular carcinoma. *Gastroenterology* 2008; **134**: 1752–63.
- Llovet JM, Bruix J. Novel advancements in the management of hepatocellular carcinoma in 2008. *J Hepatol* 2008; **48**(Suppl. 1): S20–37.
- Poon RT, Fan ST, Lo CM, Liu CL, Wong J. Long-term survival and pattern of recurrence after resection of small hepatocellular carcinoma in patients with preserved liver function: implications for a strategy of salvage transplantation. *Ann Surg* 2002; **235**: 373–82.
- Friedman MA. Primary hepatocellular cancer—present results and future prospects. *Int J Radiat Oncol Biol Phys* 1983; **9**: 1841–50.
- Lin DY, Lin SM, Liaw YF. Non-surgical treatment of hepatocellular carcinoma. *J Gastroenterol Hepatol* 1997; **12**: S319–28.
- Nagano H, Miyamoto A, Wada H, *et al.* Interferon-alpha and 5-fluorouracil combination therapy after palliative hepatic resection in patients with advanced hepatocellular carcinoma, portal venous tumor thrombus in the major trunk, and multiple nodules. *Cancer* 2007; **110**: 2493–501.
- Patt YZ, Hassan MM, Lozano RD, *et al.* Phase II trial of systemic continuous fluorouracil and subcutaneous recombinant interferon alfa-2b for treatment of hepatocellular carcinoma. *J Clin Oncol* 2003; **21**: 421–7.
- Urabe T, Kaneko S, Matsushita E, Unoura M, Kobayashi K. Clinical pilot study of intrahepatic arterial chemotherapy with methotrexate, 5-fluorouracil, cisplatin and subcutaneous interferon-alpha-2b for patients with locally advanced hepatocellular carcinoma. *Oncology* 1998; **55**: 39–47.
- Obi S, Yoshida H, Toune R, *et al.* Combination therapy of intraarterial 5-fluorouracil and systemic interferon-alpha for advanced hepatocellular carcinoma with portal venous invasion. *Cancer* 2006; **106**: 1990–7.
- Honda M, Yamashita T, Ueda T, *et al.* Different signaling pathways in the livers of patients with chronic hepatitis B or chronic hepatitis C. *Hepatology* 2006; **44**: 1122–38.
- Nishino R, Honda M, Yamashita T, *et al.* Identification of novel candidate tumour marker genes for intrahepatic cholangiocarcinoma. *J Hepatol* 2008; **49**: 207–16.
- Yamashita T, Honda M, Takatori H, *et al.* Genome-wide transcriptome mapping analysis identifies organ-specific gene expression patterns along human chromosomes. *Genomics* 2004; **84**: 867–75.
- Yamashita T, Kaneko S, Hashimoto S, *et al.* Serial analysis of gene expression in chronic hepatitis C and hepatocellular carcinoma. *Biochem Biophys Res Commun* 2001; **282**: 647–54.
- Yamashita T, Honda M, Kaneko S. Application of serial analysis of gene expression in cancer research. *Curr Pharm Biotechnol* 2008; **9**: 375–82.
- Yamashita T, Honda M, Takatori H, *et al.* Activation of lipogenic pathway correlates with cell proliferation and poor prognosis in hepatocellular carcinoma. *J Hepatol* 2009; **50**: 100–10.
- Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. *Science* 1995; **270**: 484–7.
- Yamashita T, Hashimoto S, Kaneko S, *et al.* Comprehensive gene expression profile of a normal human liver. *Biochem Biophys Res Commun* 2000; **269**: 110–6.
- Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B. A model for p53-induced apoptosis. *Nature* 1997; **389**: 300–5.
- Misu H, Takamura T, Matsuzawa N, *et al.* Genes involved in oxidative phosphorylation are coordinately upregulated with fasting hyperglycaemia in livers of patients with type 2 diabetes. *Diabetologia* 2007; **50**: 268–77.
- Longley DB, Harkin DP, Johnston PG. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* 2003; **3**: 330–8.
- Whitfield ML, George LK, Grant GD, Perou CM. Common markers of proliferation. *Nat Rev Cancer* 2006; **6**: 99–106.
- Ladner RD, Lynch FJ, Groshen S, *et al.* dUTP nucleotidohydrolase isoform expression in normal and neoplastic tissues: association with survival and response to 5-fluorouracil in colorectal cancer. *Cancer Res* 2000; **60**: 3493–503.
- El-Hajj HH, Zhang H, Weiss B. Lethality of a dut (deoxyuridine triphosphatase) mutation in *Escherichia coli*. *J Bacteriol* 1988; **170**: 1069–75.
- Canman CE, Radany EH, Parsels LA, *et al.* Induction of resistance to fluorodeoxyuridine cytotoxicity and DNA damage in human tumor cells by expression of *Escherichia coli* deoxyuridinetriphosphatase. *Cancer Res* 1994; **54**: 2296–8.
- Fleischmann J, Kremmer E, Muller S, *et al.* Expression of deoxyuridine triphosphatase (dUTPase) in colorectal tumours. *Int J Cancer* 1999; **84**: 614–7.
- Romeike BF, Bockeler A, Kremmer E, *et al.* Immunohistochemical detection of dUTPase in intracranial tumors. *Pathol Res Pract* 2005; **201**: 727–32.
- Koehler SE, Ladner RD. Small interfering RNA-mediated suppression of dUTPase sensitizes cancer cell lines to thymidylate synthase inhibition. *Mol Pharmacol* 2004; **66**: 620–6.

32. Wilson PM, Fazzone W, Labonte MJ, *et al.* Novel opportunities for thymidylate metabolism as a therapeutic target. *Mol Cancer Ther* 2008; 7: 3029–37.

### Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Fig. S1.** Subcellular localization of genes detected in each SAGE library.

**Fig. S2.** Microarray analysis of *DUT* and *TS* gene expression in 238 HCC cases publicly available (GSE5975). *DUT* was overexpressed more than 2-fold in 121 of 238 HCC tissues (median: 2.03), whereas *TS* was overexpressed more than 2-fold in 54 of 238 HCC tissues (median: 1.41) compared with the non-cancerous liver tissues.

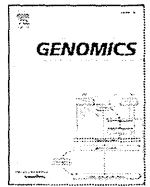
**Fig. S3.** (A) Transfection of siRNAs targeting *DUT* (*DUT2*) decreased *DUT* expression compared with the control (scrambled sequence). Gene expression was evaluated in triplicates 72 hours after transfection (mean  $\pm$  SD). (B) *DUT* gene knockdown sensitized HuH7 cells to low-dose 5-FU (0.25 mg/ml) (mean  $\pm$  SD).

**Fig. S4.** Nuclear and cytoplasmic dUTPase expression and cell proliferation in HCC. PCNA indexes in nuclear dUTPase-high HCC were higher than those in low HCC with statistical significance ( $P = 0.01$ ). Cytoplasmic dUTPase expression was not associated with PCNA indexes in HCC.

**Table S1.** A summary of constructed SAGE libraries.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.





## Comprehensive gene expression analysis of 5'-end of mRNA identified novel intronic transcripts associated with hepatocellular carcinoma

Yuji Hodo<sup>a</sup>, Shin-ichi Hashimoto<sup>b</sup>, Masao Honda<sup>a</sup>, Taro Yamashita<sup>a</sup>, Yutaka Suzuki<sup>c</sup>, Sumio Sugano<sup>c</sup>, Shuichi Kaneko<sup>a,\*</sup>, Kouji Matsushima<sup>b</sup>

<sup>a</sup> Department of Gastroenterology, Kanazawa University Graduate School of Medical Science, 13-1 Takara-Machi, Kanazawa, Ishikawa 920-8641, Japan

<sup>b</sup> Department of Molecular Preventive Medicine, School of Medicine, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

<sup>c</sup> Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa, Chiba 277-8562, Japan

### ARTICLE INFO

#### Article history:

Received 1 June 2009

Accepted 14 January 2010

Available online 21 January 2010

#### Keywords:

5'-end serial analysis of gene expression

Transcriptional start site

Acyl-coenzyme A oxidase 2

Intron

Hepatocellular carcinoma

### ABSTRACT

To elucidate the molecular feature of human hepatocellular carcinoma (HCC), we performed 5'-end serial analysis of gene expression (5'SAGE), which allows genome-wide identification of transcription start sites in addition to quantification of mRNA transcripts. Three 5'SAGE libraries were generated from normal human liver (NL), non-B, non-C HCC tumor (T), and background non-tumor tissues (NT). We obtained 226,834 tags from these libraries and mapped them to the genomic sequences of a total of 8,410 genes using RefSeq database. We identified several novel transcripts specifically expressed in HCC including those mapped to the intronic regions. Among them, we confirmed the transcripts initiated from the introns of a gene encoding acyl-coenzyme A oxidase 2 (ACOX2). The expression of these transcript variants were up-regulated in HCC and showed a different pattern compared with that of ordinary ACOX2 mRNA. The present results indicate that the transcription initiation of a subset of genes may be distinctively altered in HCC, which may suggest the utility of intronic RNAs as surrogate tumor markers.

© 2010 Elsevier Inc. All rights reserved.

### Introduction

Hepatocellular carcinoma (HCC) is the fifth most common cancer worldwide and the third most common cause of cancer mortality. HCC usually develops in patients with virus-induced (e.g., hepatitis B virus (HBV) and hepatitis C virus (HCV)) chronic inflammatory liver disease [1]; however, non-B, non-C HCC has been reported in patients negative for both HBV and HCV [2]. HCC development is a multistep process involving changes in host gene expression, some of which are correlated with the appearance and progression of a tumor. Multiple studies linking hepatitis viruses and chemical carcinogens with hepatocarcinogenesis have provided insights into tumorigenesis [1,3]. Nevertheless, the genetic events that lead to HCC development remain unknown, and the molecular pathogenesis of HCC in most patients is still unclear. Therefore, elucidation of the genetic changes specific to the pathogenesis of non-B, non-C HCC may be useful to reveal the molecular features of HCCs irrelevant to viral infection.

Gene expression profiling, either by cDNA microarray [4] or serial analysis of gene expression (SAGE) [5], is a powerful molecular technique that allows analysis of the expression of thousands of

genes. In particular, SAGE enables the rapid, quantitative, and simultaneous monitoring of the expression of tens of thousands of genes in various tissues [6,7]. Although numerous studies using cDNA microarrays and SAGE have been performed to clarify the genomic and molecular alterations associated with HCC [6,8–10], most expression data have been derived from the 3'-end region of mRNA. Recent advances in molecular biology have enabled genome-wide analysis of the 5'-end region of mRNA that revealed the variation in transcriptional start sites [11,12] and the presence of a large number of non-coding RNAs [13]. These approaches might be useful for identifying the unique and undefined genes associated with HCC not identified by the analysis of the 3'-end region of mRNA. SAGE based on the 5'-end (5'SAGE), a recently developed technique, allows for a comprehensive analysis of the transcriptional start site and quantitative gene expression [14]. This article is to elucidate the molecular carcinogenesis of non-B, non-C HCCs, while those heterogeneous entities are supposed not to share the same etiology, by using 5'SAGE.

### Results

#### Annotation of the 5'SAGE tags to the human genome

We characterized a total of 226,834 tags from three unique 5'SAGE libraries (75,268 tags from the normal liver (NL) library, 75,573 tags from the non-tumor tissue (NT) library, and 75,993 tags from the tumor (T) library) and compared them against the human genome

Abbreviations: 5'SAGE, 5'-end serial analysis of gene expression; HCC, hepatocellular carcinoma; ACOX2, acyl-coenzyme A oxidase 2.

\* Corresponding author. Fax: +81 76 234 4250

E-mail address: [skaneko@m-kanazawa.jp](mailto:skaneko@m-kanazawa.jp) (S. Kaneko).

sequence. A total of 211,818 tags matched genomic sequences, representing 104,820 different tags in the three libraries (Table 1). About 60–65% of these tags mapped to a single locus in the genome in each library. Then, we mapped these single-matched tags to the well-annotated genes using RefSeq database ([www.ncbi.nlm.nih.gov/RefSeq/](http://www.ncbi.nlm.nih.gov/RefSeq/), reference sequence database developed by NCBI). A total of 45,601 tags from the NL library, 39,858 from the NT library, and 41,265 from the T library were successfully mapped to 8410 unique genes (4397 genes detected in the NL library, 5194 genes in the NT library, and 6304 genes in the T library).

#### Gene expression profiling of non-B, non-C HCC

Abundantly expressed transcripts in the NL library and their corresponding expression in the NT and T libraries are shown in Table 2. The most abundant transcript in all three libraries was encoded by the *albumin (ALB)* gene. Transcripts encoding apolipoproteins were also abundantly expressed in each library, suggesting the preservation of hepatocytic gene expression patterns in HCC. Of note, the expression of *haptoglobin (HP)* (NL: 631, NT: 329, T: 57) and *metallothionein 1G (MT1G)* (NL: 392, NT: 169, T: 2) was decreased in the NT library and more in T library compared with NL library. Furthermore, the expression of *metallothionein 2A (MT2A)* (NL: 1027, NT: 872, T: 19), *metallothionein 1X (MT1X)* (NL: 547, NT: 644, T: 11), and *metallothionein 1E (MT1E)* (NL: 275, NT: 340, T: 2) was decreased almost fifty-fold or more in the T library compared with the NL and NT libraries. In contrast, the expression of *ribosomal protein S29 (RPS29)* (NL: 372, NT: 1011, T: 1768) was increased in the NT library and more in T library compared with NL library. Thus, transcripts associated with a certain liver function including xenobiotic metabolism might be suppressed whereas those associated with protein synthesis might be expressed in non-B, non-C HCC, similar to that observed in HCV-HCC [15].

We then investigated the characteristics of gene expression patterns in non-B, non C HCC. Two hundred fifty-four and 172 genes were up- or down-regulated in the T library more than five-fold compared with the NL library (data not shown). The top 10 genes are listed in Table 3a, and we identified several novel genes not yet reported to be differentially expressed in non-B, non-C HCC. Representative novel gene expression changes identified by 5'SAGE were validated by semi-quantitative reverse transcriptase-polymerase chain reaction (RT-PCR) analysis (Supplemental Fig. 1). RT-PCR results showed that the expression of *galectin 4 (LGALS4)*, *X antigen family, member 1A (XAGE 1A)*, *retinol dehydrogenase 11 (RDH11)*, *hydroxysteroid (17-beta) dehydrogenase 14 (HSD17B14)*, *transmembrane 14A (TMEM14A)*, *stimulated by retinoic acid 13 homolog (STRA13)*, and *dual specificity phosphatase 23 (DUSP23)* was increased, whereas the expression of *C-type lectin superfamily 4 member G (CLEC4G)* was decreased in HCC tissues compared with the non-tumor tissues.

To further characterize the gene expression patterns of non-B, non-C HCC comprehensively, we compared the Gene Ontology process of three types of HCCs (i.e., non-B, non-C HCC; HBV-HCC;

HCV-HCC) based on our previously described data [16]. The pathway analysis using MetaCore™ software showed that the immune related and cell adhesion related pathways were up-regulated in HCV-HCC with statistically significance, and the insulin signaling and angiogenesis related pathways were up-regulated in HBV-HCC with statistically significance, confirming our previous results [16]. Interestingly, genes associated with progesterone signaling were up-regulated in non-B, non-C HCC, while genes associated with proteolysis in the cell cycle, apoptosis and the ESR1-nuclear pathway were up-regulated in all types of HCC (Supplemental Fig. 2).

#### Dynamic alteration of transcription initiation in HCC

Although various transcriptome analyses have discovered considerable gene expression changes in cancer, it is still unclear if transcription is differentially initiated and/or terminated in HCC compared with the non-cancerous liver. We therefore explored the characteristics of transcription initiation and/or termination in HCC using 5'SAGE and 3'SAGE data. Markedly, we observed relevant differences between 5'SAGE and 3'SAGE data derived from the same HCC sample (Tables 3a and b). For example, a gene encoding *coagulation factor XIII, B polypeptide (F13B)* was 13-fold up-regulated at transcription start sites (5'SAGE) but two-fold down-regulated at transcription termination sites (3'SAGE). On the other hand, a gene encoding *adenylate cyclase 1 (ADCY1)* was 50-fold down-regulated at transcriptional termination sites (3'SAGE) but showed no difference at transcriptional start sites (5'SAGE). These data suggest the dramatic alteration of all process of transcription in HCC, and the transcripts initiated at certain sites might be specifically associated with and involved in HCC pathogenesis, which could be a novel marker for HCC diagnosis.

#### Identification of novel intronic transcripts in HCC

Recent lines of evidence suggest that the majority of sequences of eukaryotic genomes may be transcribed, not only from known transcription start sites but also from intergenic regions and introns [17,18]. Introns are recognized as a significant source of functional non-coding RNAs (ncRNAs) including microRNAs (miRNAs) [18]. Moreover, a recent report implied the role of some large intronic RNAs in the pathogenesis of several types of malignancies [19]. Thus, analysis of transcripts originating from introns might be valuable for elucidating the genetic traits of HCC. We therefore focused on the transcriptional start sites potentially initiated from the intron and deregulated in HCC using 5'SAGE data. We identified that 97% of 5'SAGE tags annotated by the RefSeq database matched the sequences in the exons, while 3% matched those in the introns (1257 in the N library, 1225 in the NT library, and 1261 in the T library) (Table 4a). To identify the possible promoter regions located in the intron, we clustered the different SAGE tags to a certain genomic region if these tags positioned within 500 bp intervals (Supplemental Fig. 3), as described previously [12].

**Table 1**  
Experimental matching of 5'SAGE tags to genome.

	Normal liver	Non-tumor	Tumor	Total
All tags	75,268	75,573	75,993	226,834
Tags mapped to genome (%)				
1 locus/genome	51,076 (71.2)	47,200 (68.0)	48,503 (68.5)	146,779 (69.3)
Multiple loci/genome	20,608 (28.8)	22,142 (32.0)	22,289 (31.5)	65,039 (30.7)
Total tags	71,684 (100)	69,342 (100)	70,792 (100)	211,818 (100)
Unique tags mapped to genome (%)				
1 locus/genome	20,736 (65.5)	20,487 (60.2)	23,753 (60.7)	64,976 (62.0)
Multiple loci/genome	10,914 (34.5)	13,548 (39.8)	15,382 (39.3)	39,844 (38.0)
Total tags	31,650 (100)	34,035 (100)	39,135 (100)	104,820 (100)
Total tags to RefSeq	45,601	39,858	41,265	126,724
Unique gene	4397	5194	6304	8410

5'SAGE indicates 5'-end serial analysis of gene expression.

**Table 2**

The highly expressed genes in the NL library and corresponding expression in the NT and T libraries (top 50 from NL library).

Tag count			Ratio		Gene
NL	NT	T	NT/ NL	T/NL	
3731	1716	2328	0.460	0.624	Albumin (ALB)
2484	2146	2042	0.864	0.822	Apolipoprotein C-I (APOC1)
1955	1603	1079	0.820	0.552	Apolipoprotein A-II (APOA2)
1653	1050	828	0.635	0.501	Apolipoprotein A-I (APOA1)
1252	1908	1203	1.524	0.961	Transthyretin (prealbumin, amyloidosis type I) (TTR)
1233	724	220	0.587	0.178	Serpin peptidase inhibitor, clade A, member 1 (SERPINA1)
1027	872	19	0.849	0.019	Metallothionein 2A (MT2A)
755	1144	762	1.515	1.009	Ferritin, light polypeptide (FTL)
713	632	680	0.886	0.954	Alpha-1-microglobulin/bikunin precursor (AMBIP)
635	524	1336	0.825	2.104	Apolipoprotein E (APOE)
631	329	57	0.521	0.090	Haptoglobin (HP)
600	228	212	0.380	0.353	Fibrinogen gamma chain (FGG)
549	395	302	0.719	0.550	Apolipoprotein C-III (APOC3)
547	644	11	1.177	0.020	Metallothionein 1X (MT1X)
479	257	290	0.537	0.605	Tumor protein, translationally-controlled 1 (TPT1)
463	217	53	0.469	0.114	Serpin peptidase inhibitor, clade A, member 3 (SERPINA3)
393	204	206	0.519	0.524	Ribosomal protein L26 (RPL26)
392	169	2	0.431	0.005	Metallothionein 1G (MT1G)
372	1011	1768	2.718	4.753	Ribosomal protein S29 (RPS29)
306	163	223	0.533	0.729	Ribosomal protein S27 (RPS27)
279	135	159	0.484	0.570	Ribosomal protein S16 (RPS16)
275	340	2	1.236	0.007	Metallothionein 1E (MT1E)
269	170	246	0.632	0.914	Ribosomal protein S23 (RPS23)
260	142	92	0.546	0.354	Fibrinogen beta chain (FGB)
260	200	195	0.769	0.750	Aldolase B, fructose-bisphosphate (ALDOB)
255	228	286	0.894	1.122	Ribosomal protein S12 (RPS12)
248	162	198	0.653	0.798	Ribosomal protein S14 (RPS14)
246	175	70	0.711	0.285	Interferon induced transmembrane protein 3 (IFITM3)
239	198	273	0.828	1.142	Ribosomal protein L31 (RPL31)
229	264	0	1.153	0.004	Hepcidin antimicrobial peptide (HAMP)
228	149	156	0.654	0.684	Ribosomal protein S20 (RPS20)
222	191	117	0.860	0.527	Ubiquitin B (UBB)
216	218	352	1.009	1.630	Ribosomal protein L41 (RPL41)
210	150	155	0.714	0.738	Ribosomal protein, large, P1 (RPLP1)
201	110	90	0.547	0.448	Ribosomal protein, large, P2 (RPLP2)
198	102	64	0.515	0.323	Fibrinogen alpha chain (FGA)
196	143	408	0.730	2.082	Ribosomal protein L37 (RPL37)
192	123	56	0.641	0.292	Ribosomal protein L37a (RPL37A)
191	208	346	1.089	1.812	Ribosomal protein L30 (RPL30)
174	109	76	0.626	0.437	Ribosomal protein L35 (RPL35)
169	208	3	1.231	0.018	Cytochrome P450, family 2, subfamily E, polypeptide 1 (CYP2E1)
167	105	300	0.629	1.796	Apolipoprotein H (beta-2-glycoprotein I) (APOH)
162	106	33	0.654	0.204	Serum amyloid A4, constitutive (SAA4)
159	85	157	0.535	0.987	Ribosomal protein L34 (RPL34)
159	113	229	0.711	1.440	Transferrin (TF)
155	84	135	0.542	0.871	Ribosomal protein S11 (RPS11)
152	125	101	0.822	0.664	Ribosomal protein S13 (RPS13)
147	84	1	0.571	0.007	Nicotinamide N-methyltransferase (NNMT)
147	180	35	1.224	0.238	Hemopexin (HPX)
146	89	121	0.610	0.829	Alpha-2-HS-glycoprotein (AHSG)

To avoid division by 0, a tag value of 1 for any tag that was not detectable was used. NL, normal liver; NT, non-tumor; T, tumor.

More than 2 tags were detected in the intronic regions of the 164 genes in the NL, 168 genes in the NT, and 157 genes in the T library, suggesting that these regions might be potential intronic promoter regions (Table 4a). The biological process of these intron-origin transcripts using Human Protein Reference Database (<http://www.hprd.org/>) showed that these were related to basic cellular functions such as signal transduction, transport, and regulation of the nucleobase and nucleotide, suggesting that these intronic transcripts

**Table 3a**

Differently expressed genes in HCC (top 10 from 5'SAGE).

5'SAGE	3'SAGE	5'/3'	Gene
T/NL	T/NL	Ratio	
<i>Up-regulated gene</i>			
19	6	3.17	P antigen family, member 2 (prostate associated) (PAGE2)
18	10	1.8	Lectin, galactoside-binding, soluble, 4 (LGALS4)
16	3	5.33	Choline phosphotransferase 1 (CHPT1)
14	2	7	X antigen family, member 1A (XAGE1A)
14	2	7	Dehydrogenase/reductase (SDR family) member 4 (DHRS4)
14	2	7	Sterol-C5-desaturase-like (SC5DL)
13	0.5	26	Coagulation factor XIII, B polypeptide (F13B)
13	2.33	5.58	Retinol dehydrogenase 11 (all-trans and 9-cis) (RDH11)
13	0.5	26	Transmembrane protein 14A (TMEM14A)
12	1.33	9.02	Dual specificity phosphatase 23 (DUSP23)
<i>Down-regulated gene</i>			
0.00436	0.0137	0.318	Hepcidin antimicrobial peptide (HAMP)
0.0051	ND		Metallothionein 1G (MT1G)
0.0068	0.04	0.17	Nicotinamide N-methyltransferase (NNMT)
0.00727	ND		Metallothionein 1E (functional) (MT1E)
0.0098	0.0526	0.186	C-reactive protein, pentraxin-related (CRP)
0.0145	ND		Metallothionein 1 M (MT1M)
0.0152	ND		Phospholipase A2, group IIA (platelets, synovial fluid) (PLA2G2A)
0.0178	0.111	0.16	Cytochrome P450, family 2, subfamily E, polypeptide 1 (CYP2E1)
0.0185	0.192	0.096	Metallothionein 2A (MT2A)
0.0201	ND		Metallothionein 1X (MT1X)

3'SAGE, 3'-end serial analysis of gene expression; 5'SAGE, 5'-end serial analysis of gene expression; HCC, hepatocellular carcinoma; NL, normal liver; T, tumor.

may play a fundamental role in the liver (data not shown). Among these genes, 12 were differentially expressed between the NL and T libraries more than four-fold (Table 4b). Interestingly, intronic transcripts (determined by 5'SAGE) of genes encoding SAMD3,

**Table 3b**

Differently expressed genes in HCC (top 10 from 3'SAGE).

5'SAGE	3'SAGE	5'/3'	Gene
T/NL	T/NL	Ratio	
<i>Up-regulated gene</i>			
ND	15		Leukocyte immunoglobulin-like receptor, subfamily B, member 1 (LILRB1)
ND	12		Fibroblast growth factor 5 (FGF5)
1	11	0.909	Adenosine deaminase, tRNA-specific 1 (ADAT1)
5	11	0.454	px19-like protein (PRELID1)
4.4	11	0.4	Anaphase promoting complex subunit 11 (ANAPC11)
ND	10.3		Chromosome 21 open reading frame 77 (C21orf77)
ND	10		von Willebrand factor (VWF)
2.333	10	0.233	ATX1 antioxidant protein 1 homolog (yeast) (ATOX1)
18	10	1.8	Lectin, galactoside-binding, soluble, 4 (LGALS4)
ND	9.5		Solute carrier family 26 (sulfate transporter), member 2 (SLC26A2)
<i>Down-regulated gene</i>			
0.5	0.012	41.7	ELL associated factor 1 (EAF1)
0.5	0.0137	36.5	TGF beta-inducible nuclear protein 1 (NSA2)
0.000436	0.0137	0.032	Hepcidin antimicrobial peptide (HAMP)
1	0.0179	55.9	Basic, immunoglobulin-like variable motif containing (BIVM)
ND	0.0182		DNA fragmentation factor, 45 kDa, alpha polypeptide (DFFA)
1	0.0185	54.1	GRIP1 associated protein 1 (GRIPAP1)
ND	0.0189		Nuclear factor of activated T-cells 5, tonicity-responsive (NFAT5)
1	0.0204	49	Adenylate cyclase 1 (ADCY1)
0.333	0.0312	10.7	Dihydroorotate dehydrogenase (DHODH)
0.738	0.0312	23.7	Ribosomal protein, large, P1 (RPLP1)

3'SAGE, 3'-end serial analysis of gene expression; 5'SAGE, 5'-end serial analysis of gene expression; HCC, hepatocellular carcinoma; NL, normal liver; T, tumor.

**Table 4a**  
Number of 5'SAGE tags mapped to intronic region.

	NL	NT	T
Tag mapped to intron	1287	1253	1292
Total promoter region	952	981	1020
(tag number = 1)	788	813	863
(tag number $\geq 2$ )	164	168	157

*ACOX2*, *HGD*, *CYP3A5*, *KNG1* and *AGXT* were increased, while their 3' transcripts (determined by 3'SAGE) were decreased in HCC. In contrast, both 5' intronic transcripts and 3' transcripts encoding *HFM1*, *SERPINA1*, *SUPT3H*, *A2M* and *TMEM176B* were similarly decreased in HCC. Taken together, these data imply that the canonical- and intronic-promoter activities of a subset of genes including *SAMD3*, *ACOX2*, *HGD*, *CYP3A5*, *KNG1* and *AGXT* might be differently regulated in HCC.

#### *ACOX2* as a novel intronic gene deregulated in HCC

A subset of genes listed above may be transcribed from intronic regions specifically in HCC. Among these genes, we focused on the regulation of *ACOX2*, which is reported to be potentially involved in peroxisomal beta-oxidation and hepatocarcinogenesis [20]. The intron-origin expression of *ACOX2* increased six-fold in HCC compared with the NT by 5'SAGE, while the expression based on the 3' end was almost similar between HCC and NT lesions (Table 4b). Close examination of 5'SAGE data identified two potential intron-origin transcripts of *ACOX2* (Supplemental Fig. 4). The first (intronic-*ACOX2-1*) was initiated upstream of the tenth exon, whereas the second (intronic-*ACOX2-2*) was initiated upstream of the twelfth exon of *ACOX2* (Supplemental Fig. 4). The sequence of the intronic part was unique, and the remaining part of the sequence was shared with the canonical transcripts of *ACOX2*.

The expression of canonical *ACOX2* and the two types of intron-origin transcripts was investigated in NL, NT, and T tissues by RT-PCR (Fig. 1A). Although canonical *ACOX2* expression was decreased in T than in NL, the intron-origin transcript, particularly intronic-*ACOX2-1*, was increased in T. Intronic-*ACOX2-2* transcripts also showed a modest increase. We further evaluated the alteration of these

**Table 4b**  
Differentially expressed intronic promoter regions in HCC.

5'SAGE	3'SAGE	5'/3'	Gene
T/NL	T/NL	Ratio	
<i>Up-regulated</i>			
9	1	9.00	Sterile alpha motif domain containing 3 ( <i>SAMD3</i> )
6	0.89	6.74	Acyl-Coenzyme A oxidase 2, branched chain ( <i>ACOX2</i> )
6	0.62	9.68	Homogentisate 1,2-dioxygenase (homogentisate oxidase) ( <i>HGD</i> )
6	0.009	666.67	Cytochrome P450, family 3, subfamily A, polypeptide 5 ( <i>CYP3A5</i> )
5	0.64	7.81	Kininogen 1 ( <i>KNG1</i> )
4	0.36	11.11	Alanine-glyoxylate aminotransferase ( <i>AGXT</i> )
4	1	4.00	Crystallin, alpha A ( <i>CRYAA</i> )
<i>Down-regulated</i>			
0.13	1	0.13	HFM1, ATP-dependent DNA helicase homolog ( <i>S. cerevisiae</i> ) ( <i>HFM1</i> )
0.25	0.51	0.49	Serpin peptidase inhibitor, clade A member 1 ( <i>SERPINA1</i> )
0.25	1	0.25	Suppressor of Ty 3 Homolog ( <i>S. cerevisiae</i> ) ( <i>SUPT3H</i> )
0.25	0.2	1.25	Alpha-2-macroglobulin ( <i>A2M</i> )
0.25	0.083	3.13	Transmembrane protein 176B ( <i>TMEM176B</i> )

3'SAGE, 3'-end serial analysis of gene expression; 5'SAGE, 5'-end serial analysis of gene expression; HCC, hepatocellular carcinoma; NL, normal liver; NT, non-tumor; T, tumor.

transcripts in 19 HBV-HCCs, 20 HCV-HCCs, and 4 non-B, non-C HCCs, and their background liver tissues by canonical *ACOX2* and intronic-*ACOX2* specific real-time detection (RTD)-PCR. Although the expression of canonical *ACOX2* was decreased, the expression of intronic-*ACOX2* was significantly increased (Fig. 1B). Importantly, the gene expression ratios of intronic-to canonical *ACOX2* increased more in moderately differentiated HCCs (mHCC) than in well-differentiated HCCs (wHCC), suggesting the involvement of intronic-*ACOX2* expression on HCC progression.

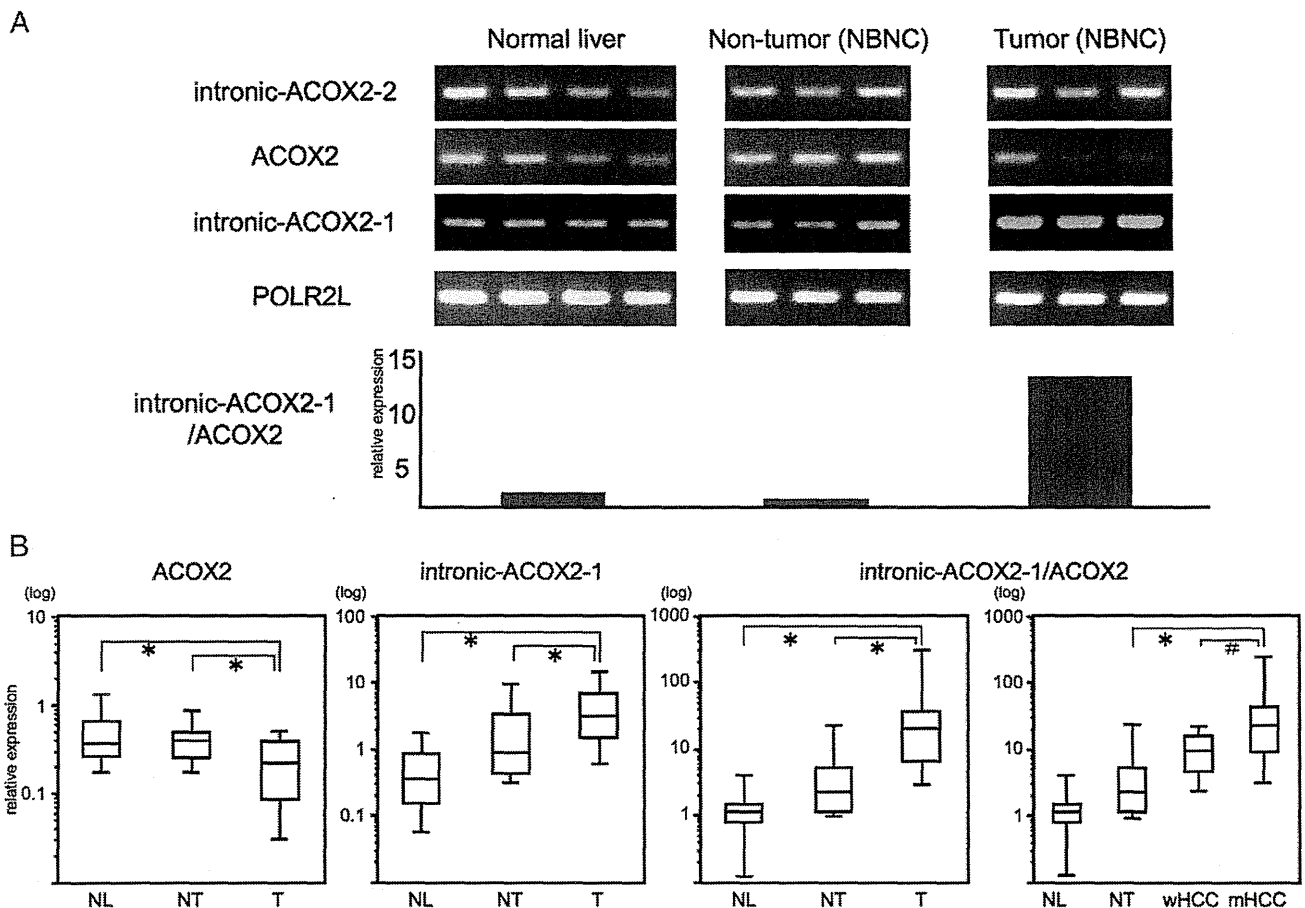
## Discussion

This is the first comprehensive transcriptional analysis of tissue lesions of non-B, non-C HCC, background liver and NL using the 5' SAGE method. Approximately 6.7% of our 5'SAGE tags showed no matching within the human genome, possibly due to the presence of a single nucleotide polymorphism (SNP) in the human genome. Out of the complete matched tags in the genome, 70% were assigned to unique positions and 30% to two or more loci. The tags with multiple matches with genomic loci were largely retrotransposon elements, repetitive sequences, and pseudogenes.

In this study, the analysis of non-B, non-C HCC enabled us to evaluate direct molecular changes associated with HCC without any bias of gene induction by virus infection. The gene expression profile based on our 5'SAGE tags revealed that *albumin* (*ALB*) and apolipoproteins were highly expressed in NL, indicating the massive production of plasma proteins in NL; these results are similar to those of our previous study using 3'SAGE [6]. Other genes such as *aldolase B* (*ALDOB*), *antitrypsin* (*SERPINA1*), and *haptoglobin* (*HP*) were also highly expressed in NL, in both the 5'SAGE and 3'SAGE libraries (Table 2) [6]. Comparison of the expression profiles among NL, background NT and T identified several differentially expressed transcripts in T. *Galectin-4* (*LGALS4*) was up-regulated and *HAMP*, *NNMT*, *CYP2E1*, and *metallothionein* were down-regulated in HCC in accordance with previous findings (Table 3a) [8,9,21]. Moreover, *CLEC4G*, which was predominantly expressed in the sinusoidal endothelial cells of the liver, was down-regulated in HCC. In addition, we first found that *P antigen family, member 2* (*PAGE2*) and *XAGE1A* were up-regulated in HCC (Table 3a, Supplemental Fig. 1). These genes were members of cancer-testis antigen that include *MAGE*-family genes. *MAGE*-family members were originally found to be up-regulated in HCV-related HCC, and reported to be useful as molecular markers and as possible target molecules for immunotherapy in human HCC [22]. In this study, we identified that these members of genes were also up-regulated in non B, non-C HCC. Thus, these genes may be useful as molecular markers and therapeutic targets for the treatment of a certain type of human HCC.

There existed some discrepancy between 5'SAGE and 3'SAGE results, even though they were derived from the same sample. Technical issues such as amplification error, difference of restriction enzyme, and annotation error have been described previously [14]. It is possible that 3' transcripts might be more stable than 5' transcripts by binding of ribosomal proteins during translation. Another possibility is the diversity of the transcriptional start sites and/or termination sites. One of the advantages of 5'SAGE analysis is the potential to determine the transcriptional start sites in each gene. Indeed, a recent study indicated the importance of an insulin splice variant in the pathogenesis of insulinomas [23]. Considering the diversity of 5' ends of genes, it is more appropriate to perform 5'SAGE in combination with 3'SAGE when determining the frequency of gene expression and identifying novel transcript variants.

Here, we were able to identify at least 12 intron-origin transcripts that were differentially expressed in HCC compared with the background liver or NL. These transcripts could not be identified by the 3'SAGE approach. We also performed detailed expression analysis of *ACOX2* that was involved in the beta-oxidation of peroxisome. We



**Fig. 1.** (A) RT-PCR results of *ACOX2* and *ACOX2* intronic RNAs in independent NL, NT (non-B, non-C), and T (non-B, non-C) samples. RT-PCR was performed in triplicate for each sample-primer set from cDNA. The PCR products were semi-quantitatively analyzed with ImageJ software and calculated as levels relative to *polymerase (RNA) II (DNA directed) polypeptide L (POLR2L)*. The bar graph indicates the expression ratio of intronic-*ACOX2-1* to canonical *ACOX2*. The expression pattern of intron 1 was different from that of canonical *ACOX2*. (B) RTD-PCR analysis of *ACOX2* and *ACOX2* intronic RNAs in NL, T (HBV-related, HCV-related, and non-B, non-C), and NT tissues. Quantitative RTD-PCR was performed in duplicate for each sample-primer set from cDNA. Each sample was normalized relative to *POLR2L*. All HCC tissues were pathologically diagnosed as well differentiated HCC (wHCC) or moderately differentiated HCC (mHCC). Kruskal-Wallis tests and Mann-Whitney *U* tests were used for statistical analysis. *ACOX2*, acyl-Coenzyme A oxidase 2; HCC, hepatocellular carcinoma; NL, normal liver; NT, non-tumor; RT-PCR, reverse transcriptase-polymerase chain reaction; RTD-PCR, real-time detection-PCR; T, tumor. \* $P < 0.01$ , # $P < 0.05$ .

were able to clone the intron-origin *ACOX2* RNAs (intronic-*ACOX2-1*, 2) for the first time and found that intronic-*ACOX2-1* was significantly overexpressed in T compared with NT and NL. The ratio of intronic-*ACOX2-1* and canonical *ACOX2* (relative intronic-*ACOX2*) was progressively up-regulated from NL via the background liver to HCC. Importantly, the expression of relative intronic-*ACOX2* was more up-regulated in moderately differentiated HCC than in well-differentiated HCC. The intronic difference in expression might be due to a polymorphism, since the 5'SAGE library for NL and T were from different people. The mechanisms of stepwise increase of intronic-*ACOX2* in the process of hepatocarcinogenesis should be clarified in future.

*ACOX2* is a rate-limiting enzyme of branched-chain acyl-CoA oxidase involved in the degradation of long branched fatty acid and bile acid intermediates in peroxisomes. *ACOX2* expression was associated with the differentiation state of hepatocytes and was repressed under the undifferentiated phase of human hepatoma cell lines [24]. A decreased *ACOX2* expression was also reported in prostate cancer [25]. Here, the expression of canonical *ACOX2* was decreased, while that of intronic-*ACOX2-1* was increased in HCC. The deduced amino acid of intronic-*ACOX2-1* encodes the C-terminal (from 386 to 681 amino acids) of canonical *ACOX2*, lacking the active sites for FAD binding and a fatty acid as the substrate, suggesting that the protein may be functionally departed [26]. The biological role of

the increased intronic-*ACOX2-1* was not clear, but it might be reflected by the activation of peroxisome proliferators-activated receptor alpha (PPARA). It is reported that mice lacking *ACOX1*, another rate-limiting enzyme in peroxisomal straight-chain fatty acid oxidation, developed steatosis and HCC characterized by increased mRNA and protein expression of genes regulated by PPAR $\alpha$  [27]. The importance of PPAR $\alpha$  activation in HCC development has been recently reported using HCV core protein transgenic mice [28]. Moreover, the overexpression of alpha-methylacyl-CoA racemase (AMACR), an enzyme for branched-chain fatty acid beta-oxidation, is reported to be a reliable diagnostic marker of prostate cancer and is associated with the decreased expression of *ACOX2* [25]. Therefore, the expression of intronic-*ACOX2-1* might open the door for further investigations of their potential clinical use, e.g., serving as diagnostic markers of HCC, although the functional relevance of this gene should be further clarified.

In conclusion, we report the first comprehensive transcriptional analysis of non-B, non-C HCC, NT background liver, and NL tissue, based on 5'SAGE. This study offers new insights into the transcriptional changes that occur during HCC development as well as the molecular mechanism of carcinogenesis in the liver. The results suggest the presence of unique intron-origin RNAs that are useful as diagnostic markers and may be used as new therapeutic targets.

## Material and methods

### Samples

Samples were obtained from a 56-year-old man who had undergone surgical hepatic resection for the treatment of solitary HCC. Serological tests for hepatitis B surface (HBs) antigen and anti-HCV antibodies were negative. Tumor (T) and non-tumor (NT) tissue samples were separately obtained from the tumorous parts (diagnosed as moderately differentiated HCC) and non-tumorous parts (diagnosed as mild chronic hepatitis: F1A1) of the resected tissue. We also obtained five normal liver (NL) tissue samples from five patients who had undergone surgical hepatic resection because of metastatic liver cancer. None of the patients was seropositive for both HBs antigen and anti-HCV antibodies. Neither heavy alcohol consumption nor the intake of chemical agents was observed before surgical resection. All laboratory values related to hepatic function were within the normal range. All procedures and risks were explained verbally and provided in a written consent form.

We additionally used independent four NL tissue samples, 19 HBV-HCCs, 20 HCV-HCCs and 4 non-B, non-C HCCs, and their background liver tissue samples for reverse transcriptase-polymerase chain reaction (RT-PCR) and real-time detection (RTD)-PCR (Supplemental Table 1). Four non-B, non-C HCCs were histologically diagnosed as moderately differentiated HCCs, and the adjacent non-cancerous liver tissues were diagnosed as a normal liver, a chronic hepatitis, a pre-cirrhotic liver and a cryptogenic liver cirrhosis, respectively. None of the patients was seropositive for HBs antigen, anti-HBs antibodies, anti-hepatitis B core (HBc) antibodies and anti-HCV antibodies. Neither heavy alcohol consumption nor the intake of chemical agents was observed. Histological grading of the tumor was evaluated by two independent pathologists as described previously [16].

### Generation of the 5' SAGE library

5'SAGE libraries were generated as previously described [14]. Five to ten micrograms of poly(A)+RNA was treated with bacterial alkaline phosphatase (BAP; TaKaRa, Otsu, Japan). Poly(A)+RNA was extracted twice with phenol: chloroform (1:1), ethanol precipitated, and then treated with tobacco acid pyrophosphatase (TAP). Two to four micrograms of the BAP-TAP-treated poly(A)+RNA was divided into two aliquots and an RNA linker containing recognition sites for *EcoRI*/*MmeI* was ligated using RNA ligase (TaKaRa): one aliquot was ligated to a 5'-oligo 1 (5'-GGA UUU GCU GGU GCA GUA CAA CGA AUU CCG AC-3') linker, and the other aliquot was ligated to a 5'-oligo 2 (5'-CUG CUC GAA UGC AAG CUU CUG AAU UCC GAC-3') linker. After removing unligated 5'-oligo, cDNA was synthesized using RNaseH-free reverse-transcriptase (Superscript II, Invitrogen, Carlsbad, CA, USA) at 12 °C for 1 h and 42 °C for the next hour, using 10 pmol of dT adapter-primer (5'-GCG GCT GAA GAC GGC CTA TGT GGC CTT TTT TTT TTT TTT TTT-3'). After first-strand synthesis, RNA was degraded in 15 mM NaOH at 65 °C for 1 h. cDNA was amplified in a volume of 100 µl by PCR with 16 pmol of 5' (5' [biotin]-GGA TTT GCT GGT GCA GTA CAA-3' or 5' [biotin]-CTG CTC GAA TGC AAG CTT CTG-3') and 3' (5'-GCG GCT GAA GAC GGC CTA TGT-3') PCR primers. cDNA was amplified using 10 cycles at 94 °C for 1 min, 58 °C for 1 min, and 72 °C for 2 min. PCR products were digested with the *MmeI* type IIS restriction endonuclease (NEB, Pickering, Ontario, Canada). The digested 5'-terminal cDNA fragments were bound to streptavidin-coated magnetic beads (Dyna, Oslo, Norway). cDNA fragments that bound to the beads were directly ligated together in a reaction mixture containing T4 DNA ligase in a supplied buffer for 2.5 h at 16 °C. The ditags were amplified by PCR using the following primers: 5' GGA TTT GCT GGT GCA GTA CA 3' and 5' CTG CTC GAA TGC AAG CTT CT 3'. The PCR products were analyzed by polyacrylamide gel electrophoresis (PAGE) and digested with *EcoRI*. The region of the gel containing the ditags was excised and the fragments were self-ligated to produce

long concatamers that were then cloned into the *EcoRI* site of pZero 1.0 (Invitrogen). Colonies were screened by PCR using the M13 forward and reverse primers. PCR products containing inserts of more than 600 bp were sequenced with Big Dye terminator ver.3 and analyzed using a 3730 ABI automated DNA sequencer (Applied Biosystems, Foster City, CA, USA). All electrophoretograms were reanalyzed by visual inspection to check for ambiguous bases and to correct misreads. In this study, we obtained 19–20 bp tag information.

### Association of the 5'SAGE tags with their corresponding genes

We attempted to align our 5'tags with the human genome (NCBI build 36, available from <http://www.genome.ucsc.edu/>) using the alignment program ALPS (<http://www.alps.gi.k.u-tokyo.ac.jp/>). Only tags that matched in sense orientation were considered in our analysis. The RefSeq database was searched for transcripts corresponding to the regions adjacent to the alignment location of each 5'tag.

### RT-PCR

Total RNA was extracted using a ToTally RNA extraction kit (Ambion, Inc., Austin, TX, USA). Total RNA (500 ng) was reverse-transcribed in a 100-µl reaction solution containing 240 U of Moloney murine leukemia virus reverse transcriptase (Promega, Madison, WI, USA), 80 U of RNase inhibitor (Promega), 4.6 mM MgCl<sub>2</sub>, 6.6 mM DTT, 1 mM dNTPs, and 2 mM random hexamer (Promega), at 42 °C for 1 h. PCR was performed in a 20-µl volume containing 0.5 U of AmpliTaq DNA polymerase (Applied Biosystems), 16.6 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 67 mM Tris-HCl, 6.7 mM MgCl<sub>2</sub>, 10 mM 2-mercaptoethanol, 1 mM dNTPs, and 1.5 µM sense and antisense primers, using an ABI 9600 thermal cycler (Applied Biosystems). The amplification protocol included 28–30 cycles of 95 °C for 45 s, 58 °C for 1 min, and 72 °C for 1 min. Primer sequences are shown in Supplemental Table 2. RT-PCR was performed in triplicate for each sample-primer set. Each sample was normalized relative to *polymerase (RNA) II (DNA directed) polypeptide L (POLR2L)*. *POLR2L* is a housekeeping gene that showed relatively stable gene expression in various tissues [29]. The PCR products were semi-quantitatively analyzed with ImageJ software (<http://rsb.info.nih.gov/ij/>).

### RTD-PCR

Intron-origin transcript expression was quantified using TaqMan Universal Master Mix (Applied Biosystems). The samples were amplified using an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems). Using the standard curve methods, quantitative PCR was performed in duplicate for each sample-primer set. Each sample was normalized relative to *POLR2L*. The assay IDs used were Hs00185873\_m1 for *ACOX2* and Hs00360764\_m1 for *POLR2L*. The specific primers and probe sequence of intronic-*ACOX2-1* were 5'-TTCATAAAGTTGTGAGCA-GAGGAAA-3' (forward), 5'-TGCACCACTTACTGAGCATCTACTC-3' (reverse), and 5'-ACTTCTACCTCAGAGCTG-3' (probe).

### Analysis of pathway network

MetaCore™ software (GeneGo Inc., St. Joseph, MI) was used to investigate the molecular pathway networks of non-B, non-C HCC, HBV-HCC and HCV-HCC. All genes up-regulated more than five-fold in all HCC libraries subjected to Enrichment analysis in GO process networks by default settings ( $p < 0.05$ ).

### Statistical analysis

Kruskal–Wallis tests were used to compare the expression among normal liver, non-cancerous tissues, and HCC tissues. Mann–Whitney U tests were also used to evaluate the statistical significance of *ACOX2*

gene expression levels between two groups. All statistical analyses were performed using R (<http://www.r-project.org/>).

### Acknowledgments

The authors would like to thank Mr. Shungo Deshimaru and Ms. Keiko Harukawa for technical assistance.

### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2010.01.004.

### References

- [1] H.B. El-Serag, K.L. Rudolph, Hepatocellular carcinoma: epidemiology and molecular carcinogenesis, *Gastroenterology* 132 (2007) 2557–2576.
- [2] Y. Yokoi, S. Suzuki, S. Baba, K. Inaba, H. Konno, S. Nakamura, Clinicopathological features of hepatocellular carcinomas (HCCs) arising in patients without chronic viral infection or alcohol abuse: a retrospective study of patients undergoing hepatic resection, *J. Gastroenterol.* 40 (2005) 274–282.
- [3] R.N. Aravalli, C.J. Steer, E.N. Cressman, Molecular mechanisms of hepatocellular carcinoma, *Hepatology* 48 (2008) 2047–2063.
- [4] D.J. Duggan, M. Bittner, Y. Chen, P. Meltzer, J.M. Trent, Expression profiling using cDNA microarrays, *Nat. Genet.* 21 (1999) 10–14.
- [5] V.E. Velculescu, L. Zhang, B. Vogelstein, K.W. Kinzler, Serial analysis of gene expression, *Science* 270 (1995) 484–487.
- [6] T. Yamashita, S. Hashimoto, S. Kaneko, S. Nagai, N. Toyoda, T. Suzuki, K. Kobayashi, K. Matsushima, Comprehensive gene expression profile of a normal human liver, *Biochem. Biophys. Res. Commun.* 269 (2000) 110–116.
- [7] S. Hashimoto, S. Nagai, J. Sese, T. Suzuki, A. Obata, T. Sato, N. Toyoda, H.Y. Dong, M. Kurachi, T. Nagahata, K. Shizuno, S. Morishita, K. Matsushima, Gene expression profile in human leukocytes, *Blood* 101 (2003) 3509–3513.
- [8] H. Okabe, S. Satoh, T. Kato, O. Kitahara, R. Yanagawa, Y. Yamaoka, T. Tsunoda, Y. Furukawa, Y. Nakamura, 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.* 61 (2001) 2129–2137.
- [9] Y. Shirota, S. Kaneko, M. Honda, H.F. Kawai, K. Kobayashi, Identification of differentially expressed genes in hepatocellular carcinoma with cDNA microarrays, *Hepatology* 33 (2001) 832–840.
- [10] T. Yamashita, M. Honda, S. Kaneko, Application of serial analysis of gene expression in cancer research, *Curr. Pharm. Biotechnol.* 9 (2008) 375–382.
- [11] Y. Suzuki, H. Taira, T. Tsunoda, J. Mizushima-Sugano, J. Sese, H. Hata, T. Ota, T. Isogai, T. Tanaka, S. Morishita, K. Okubo, Y. Sakaki, Y. Nakamura, A. Suyama, S. Sugano, Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites, *EMBO Rep.* 2 (2001) 388–393.
- [12] K. Kimura, A. Wakamatsu, Y. Suzuki, T. Ota, T. Nishikawa, R. Yamashita, J. Yamamoto, M. Sekine, K. Tsuritani, H. Wakaguri, S. Ishii, T. Sugiyama, K. Saito, Y. Isono, R. Irie, N. Kushida, T. Yoneyama, R. Otsuka, K. Kanda, T. Yokoi, H. Kondo, M. Wagatsuma, K. Murakawa, S. Ishida, T. Ishibashi, A. Takahashi-Fujii, T. Tanase, K. Nagai, H. Kikuchi, K. Nakai, T. Isogai, S. Sugano, Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes, *Genome Res.* 16 (2006) 55–65.
- [13] T. Shiraki, S. Kondo, S. Katayama, K. Waki, T. Kasukawa, H. Kawaji, R. Kodzius, A. Watahiki, M. Nakamura, T. Arakawa, S. Fukuda, D. Sasaki, A. Podhajnska, M. Harbers, J. Kawai, P. Carninci, Y. Hayashizaki, Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage, *Proc. Natl. Acad. Sci. U. S. A.* 100 (2003) 15776–15781.
- [14] S. Hashimoto, Y. Suzuki, Y. Kasai, K. Morohoshi, T. Yamada, J. Sese, S. Morishita, S. Sugano, K. Matsushima, 5'-end SAGE for the analysis of transcriptional start sites, *Nat. Biotechnol.* 22 (2004) 1146–1149.
- [15] T. Yamashita, S. Kaneko, S. Hashimoto, T. Sato, S. Nagai, N. Toyoda, T. Suzuki, K. Kobayashi, K. Matsushima, Serial analysis of gene expression in chronic hepatitis C and hepatocellular carcinoma, *Biochem. Biophys. Res. Commun.* 282 (2001) 647–654.
- [16] T. Yamashita, M. Honda, H. Takatori, R. Nishino, H. Minato, H. Takamura, T. Ohta, S. Kaneko, Activation of lipogenic pathway correlates with cell proliferation and poor prognosis in hepatocellular carcinoma, *J. Hepatol.* 50 (2009) 100–110.
- [17] J.S. Mattick, Introns: evolution and function, *Curr. Opin. Genet. Dev.* 4 (1994) 823–831.
- [18] J.S. Mattick, I.V. Makunin, Non-coding RNA, *Hum. Mol. Genet.* 15 (Spec No 1) (2006) R17–29.
- [19] R. Louro, A.S. Smirnova, S. Verjovski-Almeida, Long intronic noncoding RNA transcription: expression noise or expression choice? *Genomics* 93 (2009) 291–298.
- [20] S. Yu, S. Rao, J.K. Reddy, Peroxisome proliferator-activated receptors, fatty acid oxidation, steatohepatitis and hepatocarcinogenesis, *Curr. Mol. Med.* 3 (2003) 561–572.
- [21] N. Kondoh, T. Wakatsuki, A. Ryo, A. Hada, T. Aihara, S. Horiuchi, N. Goseki, O. Matsubara, K. Takenaka, M. Shichita, K. Tanaka, M. Shuda, M. Yamamoto, Identification and characterization of genes associated with human hepatocellular carcinogenesis, *Cancer Res.* 59 (1999) 4990–4996.
- [22] Y. Kobayashi, T. Higashi, K. Nouse, H. Nakatsukasa, M. Ishizaki, T. Kaneyoshi, N. Toshikuni, K. Kariyama, E. Nakayama, T. Tsuji, Expression of MAGE, GAGE and BAGE genes in human liver diseases: utility as molecular markers for hepatocellular carcinoma, *J. Hepatol.* 32 (2000) 612–617.
- [23] A.H. Minn, M. Kayton, D. Lorang, S.C. Hoffmann, D.M. Harlan, S.K. Libutti, A. Shalev, Insulinomas and expression of an insulin splice variant, *Lancet* 363 (2004) 363–367.
- [24] H. Stier, H.D. Fahimi, P.P. Van Veldhoven, G.P. Mannaerts, A. Volk, E. Baumgart, Maturation of peroxisomes in differentiating human hepatoblastoma cells (HepG2): possible involvement of the peroxisome proliferator-activated receptor alpha (PPAR alpha), *Differentiation* 64 (1998) 55–66.
- [25] S. Zha, S. Ferdinandusse, J.L. Hicks, S. Denis, T.A. Dunn, R.J. Wanders, J. Luo, A.M. De Marzo, W.B. Isaacs, Peroxisomal branched chain fatty acid beta-oxidation pathway is upregulated in prostate cancer, *Prostate* 63 (2005) 316–323.
- [26] K. Tokuoka, Y. Nakajima, K. Hirotsu, I. Miyahara, Y. Nishina, K. Shiga, H. Tamaoki, C. Setoyama, H. Tojo, R. Miura, Three-dimensional structure of rat-liver acyl-CoA oxidase in complex with a fatty acid: insights into substrate-recognition and reactivity toward molecular oxygen, *J. Biochem.* 139 (2006) 789–795.
- [27] K. Meyer, Y. Jia, W.Q. Cao, P. Kashireddy, M.S. Rao, Expression of peroxisome proliferator-activated receptor alpha, and PPARalpha regulated genes in spontaneously developed hepatocellular carcinomas in fatty acyl-CoA oxidase null mice, *Int. J. Oncol.* 21 (2002) 1175–1180.
- [28] N. Tanaka, K. Moriya, K. Kiyosawa, K. Koike, F.J. Gonzalez, T. Aoyama, PPARalpha activation is essential for HCV core protein-induced hepatic steatosis and hepatocellular carcinoma in mice, *J. Clin. Invest.* 118 (2008) 683–694.
- [29] C. Rubie, K. Kempf, J. Hans, T. Su, B. Tilton, T. Georg, B. Brittner, B. Ludwig, M. Schilling, Housekeeping gene variability in normal and cancerous colorectal, pancreatic, esophageal, gastric and hepatic tissues, *Mol. Cell. Probes.* 19 (2005) 101–109.

## Identification of a secretory protein *c19orf10* activated in hepatocellular carcinoma

Hajime Sunagozaka, Masao Honda, Taro Yamashita, Ryuhei Nishino, Hajime Takatori, Kuniaki Arai, Tatsuya Yamashita, Yoshio Sakai and Shuichi Kaneko

Department of Gastroenterology, Kanazawa University Hospital, Kanazawa, Ishikawa, Japan

The identification of genes involved in tumor growth is crucial for the development of inventive anticancer treatments. Here, we have cloned a 17-kDa secretory protein encoded by *c19orf10* from hepatocellular carcinoma (HCC) serial analysis of gene expression libraries. Gene expression analysis indicated that *c19orf10* was overexpressed in approximately two-thirds of HCC tissues compared to the adjacent noncancerous liver tissues, and its expression was significantly positively correlated with that of alpha-fetoprotein (AFP). Overexpression of *c19orf10* enhanced cell proliferation of AFP-negative HLE cells, whereas knockdown of *c19orf10* inhibited cell proliferation of AFP-positive Hep3B and HuH7 cells along with G1 cell cycle arrest. Supplementation of recombinant *c19orf10* protein in culture media enhanced cell proliferation in HLE cells, and this effect was abolished by the addition of antibodies developed against *c19orf10*. Intriguingly, *c19orf10* could regulate cell proliferation through the activation of Akt/mitogen-activated protein kinase pathways. Taken together, these data suggest that *c19orf10* might be one of the growth factors and potential molecular targets activated in HCC.

Hepatocellular carcinoma (HCC) is one of the most common cancers with an estimated worldwide incidence of 1,000,000 cases per year.<sup>1</sup> Most HCCs develop as a consequence of chronic liver disease such as chronic viral hepatitis due to hepatitis C virus (HCV) or hepatitis B virus (HBV) infection.<sup>2-7</sup> Liver cirrhosis patients with any etiology are considered to be at an extremely high risk for HCC.<sup>8-10</sup> Indeed, ~7% of liver cirrhosis patients with HCV infection develop HCC annually,<sup>8,11</sup> and the advancement of reliable HCC screening methods for high-risk patients is crucial for the improvement of their overall survival.<sup>12</sup>

Currently, imaging diagnostic techniques such as ultrasonography, computed tomography, magnetic resonance image and angiography are the gold standards for the early detection of HCC.<sup>13,14</sup> In addition, tumor markers such as alpha-fetoprotein (AFP) and des-gamma carboxyl prothrombin (DCP) have been used for the screening of HCC,<sup>15-18</sup> although their sensitivity and specificity are not sufficiently high. Recently, a gene expression profiling approach shed new light on Glypican 3, a heparin sulfate proteoglycan anch-

ored to the plasma membrane, as a potential HCC marker, and its clinical usefulness as a molecular target as well as a tumor marker is presently under investigation.<sup>19</sup>

There are several options available for the treatment of HCC, including surgical resection, liver transplantation, radiofrequency ablation, transcatheter arterial chemoembolization and chemotherapy, while taking the HCC stage and liver function into consideration. Recently, molecular therapy targeting the Raf kinase/vascular endothelial growth factor receptor (VEGFR) kinase inhibitor sorafenib improved the survival of patients with advanced HCC,<sup>20,21</sup> emphasizing the importance of deciphering the molecular pathogenesis of HCC for the development of effective treatment options.

Here, we investigated the gene expression profiles of HCC by serial analysis of gene expression (SAGE) to discover a novel gene activated in HCC.<sup>22-25</sup> We identified a gene, *c19orf10*, overexpressed in HCC and determined that the encoded 17-kDa protein (*c19orf10*) is a secretory protein. Murine *c19orf10* was originally discovered to encode a cytokine interleukin (IL)-25/stroma-derived growth factor (SF20) in 2001.<sup>26</sup> The gene *c19orf10* was mapped in the H2 complex region of mouse chromosome 17 between *C3* and *Ir5*, and the hypothetical protein was predicted as globular protein.<sup>26</sup> However, the subsequent study failed to reproduce its proliferative effect on lymphoid cells, and the paper was retracted by the authors in 2003.<sup>26,27</sup> Nevertheless, independent studies revealed that *c19orf10* was indeed produced by synovocytes, macrophages and adipocytes, although the function of *c19orf10* remained elusive.<sup>28,29</sup> In our study, we identified that *c19orf10* was overexpressed in AFP-positive HCC samples. Our data imply that *c19orf10* could activate the mitogen-activated protein kinase (MAPK)/Akt pathway and

**Key words:** hepatocellular carcinoma, serial analysis of gene expression, *c19orf10*

Additional Supporting Information may be found in the online version of this article

DOI: 10.1002/ijc.25830

History: Received 14 Mar 2010; Accepted 15 Nov 2010; Online 2 Dec 2010

**Correspondence to:** Shuichi Kaneko, Kanazawa University Hospital, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan, Tel.: +81-76-265-2233, Fax: +81-76-234-4250, E-mail: skaneko@m-kanazawa.jp



enhance cell proliferation in HCC cell lines, suggesting that *c19orf10* may be a growth factor produced by tumor epithelial cells and/or stromal cells, and, therefore, would be a good target for the treatment of HCC.

## Material and Methods

### SAGE and HCC samples

HCC and normal liver SAGE libraries that we had constructed were reanalyzed using SAGE 2000 software. The size of each SAGE library was normalized to 300,000 transcripts per library. Monte Carlo simulation was used to select genes whose expression levels were significantly different between the two libraries. Each SAGE tag was annotated using the gene-mapping website SAGE Genie database (<http://cgap.nci.nih.gov/SAGE/>) and the SOURCE database (<http://smd.stanford.edu/cgi-bin/source/sourceSearch>) as previously described.<sup>30</sup> An additional 15 SAGE libraries of normal and cancerous tissues from various organs were retrieved using the National Center for Biotechnology Information SAGEmap (<http://www.ncbi.nlm.nih.gov/SAGE/>).

Fifteen HCC tissues (four HBV-related and 11 HCV-related) and the corresponding noncancerous liver tissues were obtained from HCC patients who received hepatectomy. Four normal liver tissues were obtained from patients undergoing surgical resection of the liver for the treatment of metastatic colon cancer. Additionally, 36 HCC tissues (17 HBV-related and 19 HCV-related) were obtained from HCC patients undergoing hepatectomy. These samples were snap frozen in liquid nitrogen immediately after resection and used for quantitative real-time detection PCR (RTD-PCR). Total RNA was extracted using a ToTALLY RNA<sup>TM</sup> kit (Ambion, Austin, TX).

The study protocol conformed to the ethical guidelines of the Declaration of Helsinki (1975) and was approved by the institutional ethical review board committee. All patients provided written informed consent for the analysis of the specimens.

### Laser capture microdissection and RNA isolation

Laser capture microdissection (LCM) was performed as previously described.<sup>31</sup> Briefly, 20 HCV-related surgically resected HCC tissues were frozen in OCT compound (Sakura Finetech, Torrance, CA).<sup>32</sup> Inflammatory cells and cancerous cells in HCC tissues were separately excised by LCM using a Laser Scissors CRI-337 (Cell Robotics, Albuquerque, NM) under a microscope. Total RNA was isolated from these cells using a microRNA isolation kit (Stratagene, La Jolla, CA) in accordance with the supplied protocol, with slight modifications.<sup>31</sup>

### Construction of C19ORF10 expression plasmid and recombinant adenovirus vector

PCR was performed on a Marathon cDNA library from Huh7 cells using the following primers: sense primers:

5'-GACCCTAGTCCAACATGGCGGCC-3' (the first PCR), 5'-ATGGCGGCCAGCGGAGGGTGGAAACGGC-3' (the nested second PCR) and antisense primers: 5'-CACCGGAGATGAGAAGGTGCCACCCGC-3' (the first PCR), 5'-CAGGGCTGCTGGTCACAGCTCAGTGCGCG-3' (the nested second PCR). The 5' and 3' ends of the cDNA were isolated using a SMART RACE cDNA Amplification kit (Clontech, Mountain View, CA) according to the manufacturer's recommendations. The PCR products were cloned into a TA vector (Invitrogen, Carlsbad, CA) to generate the pcDNA3.1-*c19orf10* expression plasmid. Using this plasmid, a C-terminally FLAG-tagged construct of *c19orf10* was generated and inserted in a pSI mammalian expression vector (Promega, Madison, WI), which was driven by the SV40 promoter (pSI-*c19orf10*).

The replication-incompetent recombinant adenovirus vector expressing FLAG-tagged *c19orf10* (Ad. *c19orf10*-FLAG) was generated by homologous recombination using the AdMax system (Microbix, Toronto, Canada) as previously described.<sup>33</sup> The generated recombinant adenovirus was purified by limiting dilution, and the titer of viral aliquots was determined by the 50% tissue culture infectious dose method as previously described.<sup>34</sup>

### RTD-PCR

RTD-PCR was performed as previously described.<sup>31</sup> Briefly, template cDNA was synthesized from 1  $\mu$ g of total RNA using SuperScript<sup>TM</sup> II RT (Invitrogen). RTD-PCR of *c19orf10* (Hs. 00384077\_m1), *AFP* (Hs00173490\_m1), *GPC3* (Hs01018938\_m1), *KRT19* (Hs00761767\_s1) and the *ACTB* internal control (Hs99999903\_m1) was performed using a TaqMan<sup>®</sup> Gene Expression Assay kit (Applied Biosystems, Foster City, CA). The expression of selected genes was measured in triplicate by  $\Delta\Delta$ CT method using the 7900 Sequence Detection System (Applied Biosystems).

### Cell lines and transfection of plasmids

Human liver cancer cell lines HuH1, Huh7, Hep3B, HLE and HLF as well as HEK293 and NIH3T3 were cultured in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% heat-inactivated fetal bovine serum (Invitrogen) in 5% CO<sub>2</sub> at 37°C. Transfection of plasmids was performed using FuGENE<sup>TM</sup> 6 (Roche Diagnostics, Indianapolis, IN) according to the manufacturer's instruction. Briefly, 5  $\times$  10<sup>5</sup> cells were seeded in a six-well plate 12 hr before transfection, and 3  $\mu$ g of plasmid DNA was used for each transfection. All experiments were repeated at least twice.

### Purification of c19orf10-FLAG fused protein and production of anti-c19orf10 antibody

Approximately 500 ml of culture supernatant obtained from HEK293 cells infected with Ad. *C19ORF10*-FLAG at a multiplicity of infection of 20 was applied to an anti-FLAG affinity gel column (Sigma-Aldrich, St. Louis, MO). The column was

Table 1. ESTs overexpressed in the HCC library

Tag sequence	<i>p</i> value	HCC	Normal liver	T/N ratio	Name	UniGene ID
TGGGCAGGTG	<0.00001	33	0	>33	Chromosome 5 open reading frame 13	Hs.483067
GCAAAATATC	<0.00001	31	2	15.5	Liver cancer-associated noncoding mRNA, partial sequence	Hs.214343
AGCCTGCAGA	0.0002	12	1	12	Chromosome 19 open reading frame 10	Hs.465645
TTGTGCACGT	0.000228	12	1	12	CDNA FLJ45284 fis, clone BRHIP3001964	Hs.514273
ACATTCTTGT	0.000042	12	0	>12	Transcribed locus, strongly similar to XP_496055.1	Hs.76704
ACAAGTACCC	0.001161	10	1	>10	Chromosome 5 open reading frame 13	Hs.483067
GAGGTGAAGG	0.000174	10	0	>10	KIAA1914	Hs.501106
GCTGGAGGAG	0.000114	10	0	>10	Transcribed locus	Hs.520115

subjected to elution by competition with FLAG peptide (5 µg/ml), and each 1 ml fraction of the eluted aliquot was collected to obtain the most concentrated c19orf10-FLAG protein in accordance with the manufacturer's protocol. The anti-c19orf10 antibodies were developed by immunizing rabbits with repeated intradermal injections of purified c19orf10-FLAG. Protein concentration was measured by the Bradford method.

#### Silencing gene expression by short interfering RNA

The selected short interfering RNA (siRNA) targeting *C19ORF10* (Si-*C19ORF10*; Silencer Select siRNAs s31855) and the irrelevant control sequence (Si-*Control*; Silencer Select siRNAs 4390843) was obtained from Applied Biosystems. Transfection of these siRNAs was performed using FuGENE™ 6 (Roche Diagnostics) as previously described.<sup>30</sup> Briefly,  $2 \times 10^5$  cells were seeded in a six-well plate 12 hr before transfection. A total of 100 pmol/l of siRNA duplex was used for each transfection. The experiments were performed at least twice.

#### Cell proliferation assay

Cell proliferation was evaluated in quadruplicate using a Cell Titer 96 MTS Assay kit (Promega). Briefly,  $2 \times 10^3$  HLE or HuH7 cells were harvested in a 96-well plate 12 hr before the transfection or addition of the recombinant proteins. Transfection of siRNAs or plasmids was performed using FuGENE™ 6 (Roche Diagnostics). After incubation with MTS/PMS solution at 37°C for 2 hr, the absorbance at 450 nm was measured. The experiments were performed at least twice.

#### Cell cycle analysis

Cells were fixed using 80% ice-cold ethanol and incubated with propidium iodide for 10 min. DNA content was analyzed using a FACS Caliber flow cytometer (BD Biosciences, San Jose, CA) counting 10,000 stained cells. The distribution of cells in each cell cycle phase was determined using FlowJo software (Tree Star, Ashland, OR).

#### Western blotting

Cells were lysed in radioimmunoprecipitation assay (RIPA) buffer, and the extracts were subsequently electrophoresed on sodium dodecyl sulfate–10% polyacrylamide gels and transferred onto protean nitrocellulose membranes. The blots were then incubated for 1 hr with an appropriate primary monoclonal antibody: phospho-PI3K (#4228), phospho-Akt (#4060), phospho-GSK-3β (#9323), phospho-c-Raf (#9427), phospho-MEK1/2 (#9154), phospho-p44/42 MAPK (Erk1/2) (#4370), Cdk4 (CDK4 (#2906)), Cdk6 (#3136), cyclinD1 (#2926), cyclinD3 (#2936), phospho-Rb (#9308), phospho-P53 (# 9286), phospho-cdc2 (#9111) and β-actin (#4970) (Cell Signaling Technology, Allschwil, Switzerland) and anti-FLAG antibodies (Sigma-Aldrich, St. Louis, MO). The blots were washed and exposed to peroxidase-conjugated secondary antibodies, such as anti-mouse or rabbit IgG antibodies, and visualized using the ECL™ kit (Amersham Biosciences, Piscataway, NJ). All experiments were performed at least twice.

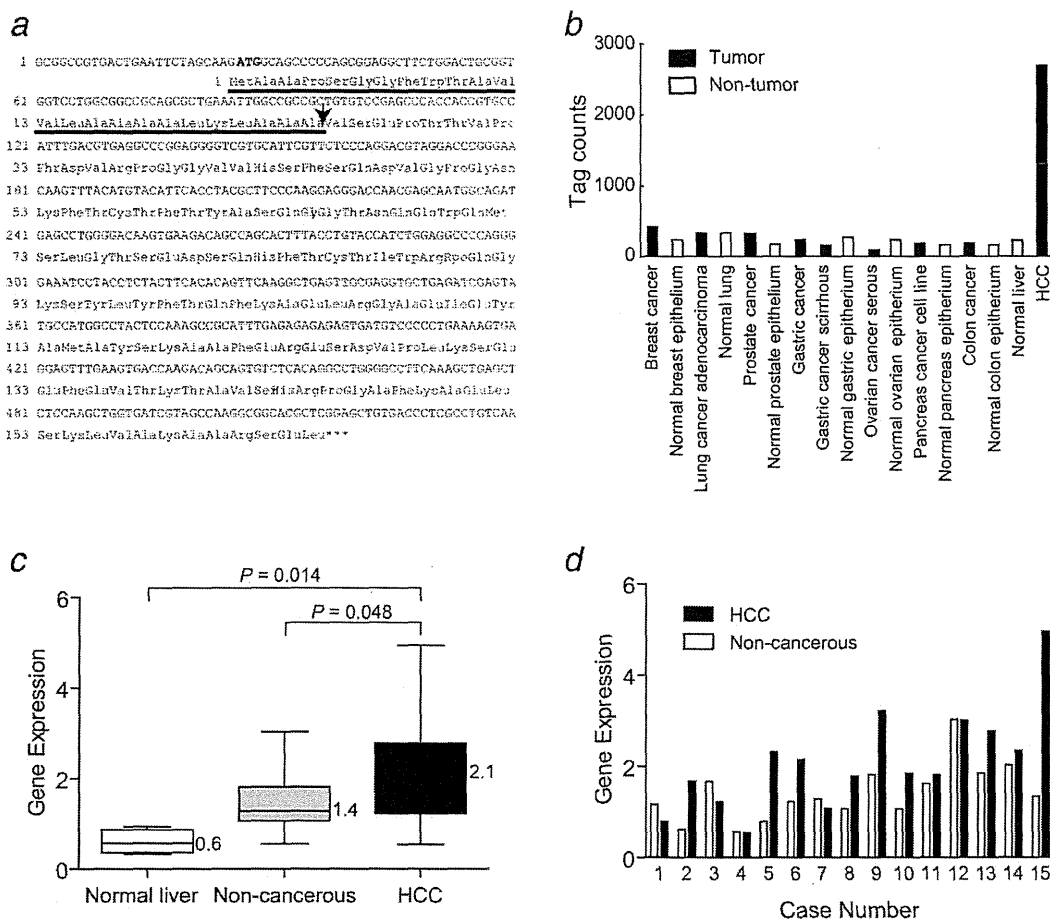
#### Statistical analyses

Unpaired *t*-tests and Kruskal–Wallis tests were performed on the RTD-PCR and cell proliferation data using GraphPad Prism software (www.graphpad.com).

#### Results

##### Identification of *C19ORF10* overexpression in HCC by SAGE

To comprehensively explore the candidate novel genes activated in HCC, we reanalyzed two SAGE libraries derived from HCC tissues and normal liver tissues.<sup>30</sup> After normalization of each SAGE library size to 300,000 tags, we compared the HCC and normal liver libraries to obtain the list of genes overexpressed in HCC. We identified 79 genes significantly overexpressed in the HCC library by more than tenfold when compared to the normal liver library (Supporting Information Table 1). Among them, we explored expressed sequence tags (ESTs) as candidates for novel HCC-related genes to identify eight unique tags corresponding to seven ESTs (Table 1). We especially focused on the EST chromosome 19 open reading frame 10 (*c19orf10*) because the



**Figure 1.** (a) Structure of a *c19orf10* gene and a *c19orf10* protein. The DNA sequence of *c19orf10* and amino acid alignment of the encoded *c19orf10* protein are shown. *C19orf10* is predicted to have a molecular weight of 17 kDa and contain a signal peptide cleavage site (indicated as a black arrow). (b) *C19orf10* gene expression profiles in various tissues by SAGE. Y-axis indicates the number of tags corresponding to *c19orf10* in each tissue. (c, d) RTD-PCR analysis of *c19orf10*. RNA was isolated from 34 tissue samples: 15 HCC, 15 corresponding noncancerous liver samples and four normal liver samples. Differential expression of each gene among normal liver tissues, noncancerous liver tissues and HCC tissues was examined using the Kruskal–Wallis test and unpaired *t*-test. The mean value of gene expression data in each group is indicated (c). *C19orf10* was overexpressed in 10 of 15 examined HCC tissues compared to the noncancerous liver tissues (d).

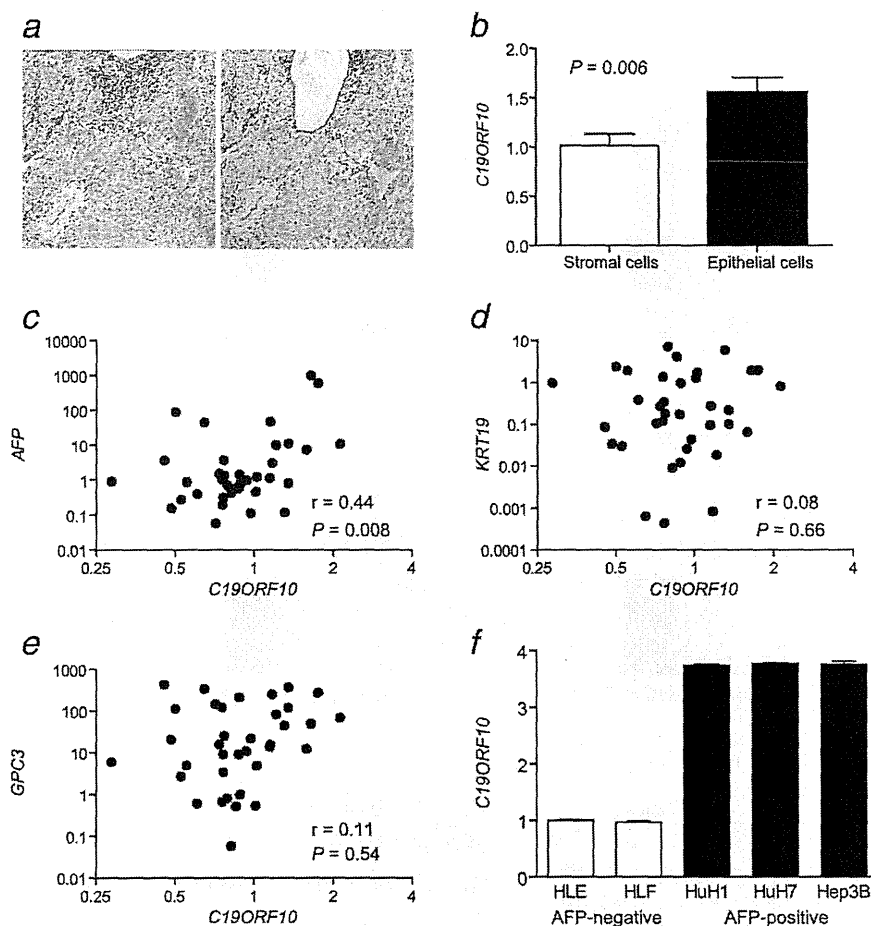
sequence presumably encoded a secretory protein with a signal peptide sequence (Fig. 1a).

When we examined the expression profiles of *c19orf10* using retrieved SAGE data from various cancers and their normal counterparts, we identified that *c19orf10* was abundantly expressed in human HCC (Fig. 1b). We further examined the publicly available EST profiles of *c19orf10* (<http://www.ncbi.nlm.nih.gov/unigene>) and confirmed its tendency to be overexpressed in HCC compared to the normal liver (data not shown). We validated the overexpression of *c19orf10* in 15 independent HCC tissues and adjacent non-cancerous liver tissues by RTD-PCR. Gene expression of *c19orf10* was significantly higher in the HCC tissues than in

the normal liver tissues and adjacent noncancerous liver tissues ( $p = 0.014$  and  $0.048$ , respectively; Fig. 1c). *C19orf10* expression was elevated in HCC tissues compared to the adjacent noncancerous liver tissues in 10 of 15 patients (66.7%; Fig. 1d).

#### Overexpression of *C19ORF10* in AFP-positive HCC

As HCC is a heterogeneous mixture of cancer epithelial cells and stromal cells, and a previous report indicated that *c19orf10* is expressed in fibroblast-like synoviocytes. We, therefore, evaluated the expression of *c19orf10* in tumor epithelial cells and stromal cells separately using LCM and RTD-PCR in 20 HCC tissues (Fig. 2a). Although tumor



**Figure 2.** (a) Representative photomicrographs of an HCC tissue used for LCM (toluidine blue staining). Inflammatory mononuclear cells and stromal cells were separately captured (left: Pre-LCM, right: Post-LCM). (b) RTD-PCR analysis of *c19orf10* expression in inflammatory mononuclear cells and tumor epithelial cells in 20 HCV-related HCC tissues. Tumor-inflammatory mononuclear cells and stromal cells were isolated using LCM. RNAs were isolated from these cells as well as parenchymal tissues from the same liver, followed by RTD-PCR for *c19orf10* gene expression. Expression of the *c19orf10* gene was higher than that observed in HCC-infiltrating inflammatory mononuclear cells. \* $p < 0.05$ . (c–e) Scatter plot analysis of *c19orf10*, *AFP*, *KRT19* and *GPC3* expression in HCC. RNA was isolated from 17 HBV-related HCC and 19 HCV-related HCC. (f) RTD-PCR analysis of *c19orf10* in AFP-negative (HLE and HLF) and -positive (HuH1, HuH7 and Hep3B) liver cancer cell lines.

stromal cells expressed *c19orf10* at some level, the expression levels were significantly higher in tumor epithelial cells than in stromal cells ( $p = 0.006$ ) (Fig. 2b).

To explore the relationship of *c19orf10* with other established HCC markers, we investigated the gene expression of *c19orf10*, *AFP* (alpha-fetoprotein), *KRT19* (cytokeratin 19) and *GPC3* (glypican 3). Because only 1 of 15 HCC tissues analyzed above (Fig. 1d) was AFP positive (data not shown), we further investigated the expression of *c19orf10* in an additional 36 HCC tissues using RTD-PCR. Interestingly, *c19orf10* expression was significantly positively correlated with *AFP* ( $r = 0.44$ ,  $p = 0.008$ ), but not with *KRT19* ( $r = 0.08$ ,  $p = 0.66$ ) nor *GPC3* ( $r = 0.11$ ,  $p = 0.54$ ) (Figs. 2c–2e).

Furthermore, when we examined the expression of *c19orf10* in AFP-positive (HuH1, HuH7 and Hep3B) and -negative (HLE and HLF) HCC cell lines, we identified the overexpression of *c19orf10* in AFP-positive HCC cell lines (Fig. 2f). These data suggested that *c19orf10* is overexpressed and may play some role in AFP-positive HCCs.

#### ***c19orf10* regulates MAPK/Akt pathways and activates cell proliferation**

To explore the functional role of *c19orf10* in HCC, we performed *c19orf10* overexpression and knockdown studies using *c19orf10*-low HLE cells and *c19orf10*-high Hep3B and HuH7 cells, respectively. When we transfected HLE cells with