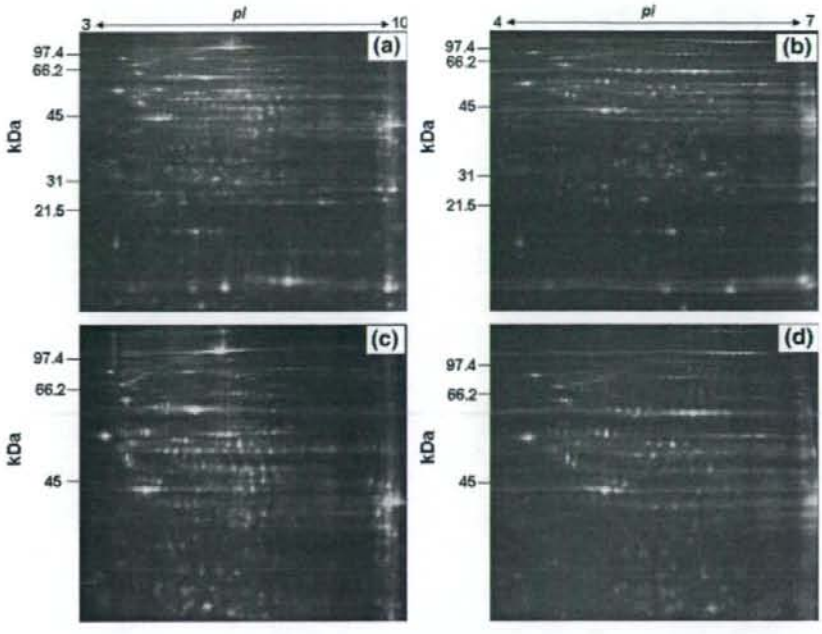

Fig. 2 Merged view of 2D-DIGE images of non HCC and HCC tissues samples. Protein samples (50 μg) extracted from non HCC labeled with Cy3 and HCC labeled with Cy5 were electrofocused on a 24 cm pH 3–10 or pH 4–7 IPG strips and then separated by SDS-PAGE on a 12.5% gel over night. The more electrophoresis was performed to extend high molecular region. Red spots and green spots represent abundant and less abundant proteins in HCC, respectively. (**a**) pH 3–10, (**b**) pH 4–7, (**c**) pH 3–10 extend, (**d**) pH 4–7 extend

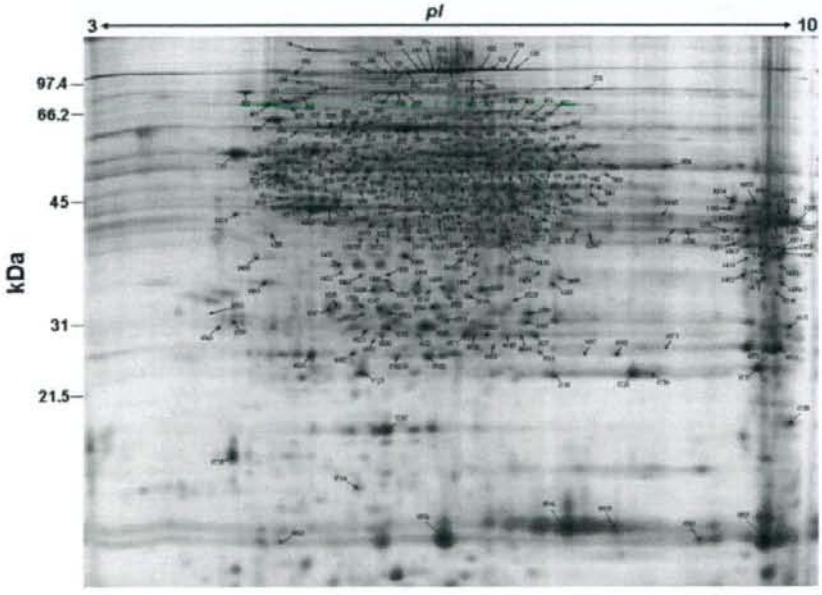

Fig. 3 2-DE map of non HCC proteins (pH 3–10). Proteins were determined through PMF using preparative gels run in parallel with analytical gels. Identified spots were projected onto the analytical image of non HCC sample. Annotations in the gel refer to the Spot ID in Supplementary Table 1

the identified proteins were mitochondrion (40%), cytosol (20%), nucleus (10%), and endoplasmic reticulum (10%; Fig. 4b). It was noticed that ribosome and plasma membrane were major cellular components of identified mRNAs (Fig. 4d), whereas identified proteins localized at ribosome and plasma membrane were rare (Fig. 4b).

3.2 Protein Expression Analysis

Using fluorescence images of proteins from non HCC and HCC samples, the normalized spot volume of protein abundance for each spot was calculated relative to the internal standard using DeCyder software. We performed

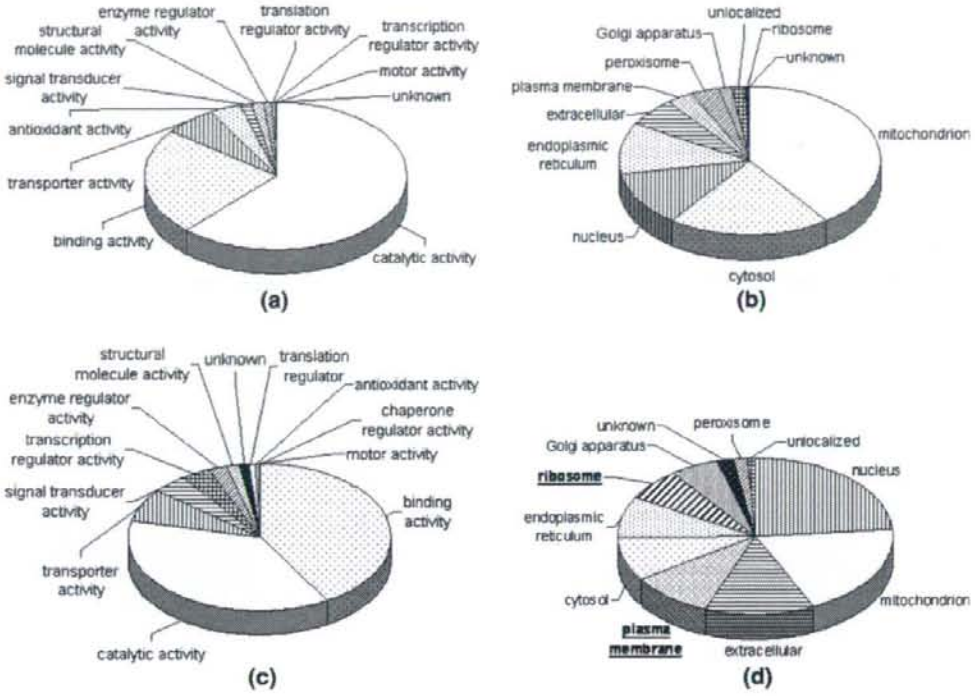

Fig. 4 Functional categorization and cellular components of identified proteins from 2-DE and identified genes from SAGE. The molecular functions and cellular components classification were performed according to criteria described by Gene Ontology Consortium using BioCompass. (**a**) molecular functions of identified proteins, (**b**) cellular components of identified proteins, (**c**) molecular functions of identified genes, (**d**) cellular components of identified genes

Student's paired *t*-test ($p \leq 0.05$) to filter the protein spots which differentially expressed between non HCC and HCC, using 2-DE gel images run in triplicate.

Protein identification revealed that a number of proteins were expressed by multiple spots. Most multi-spotted proteins showed little spot-to-spot variations in averaged HCC/non HCC ratio. Although we do not know at present how these multiple spots were generated, some of them might be caused by conformational equilibrium of proteins rather than caused by any post-translational modifications [2]. On the other hand, the HCC/non HCC expression ratios of nine multi-spotted proteins were varied from spot to spot and we categorized these proteins into "variable" group (Supplementary Table 1, Fig. 5). For example, the expression levels of two spots, #1682 and #1693, identified as apolipoprotein A1 (APOA1) seemed to be regulated oppositely between HCC and non HCC. While the expression of spot #1682 was decreased in HCC, the expression of spot #1693 was increased in HCC (Fig. 5a, b). The spots #302 and #377, both identified as heat shock protein 90 kDa beta (GRP94), showed different molecular mass and pI. The expression levels of these spots seemed to be differently regulated (Fig. 5c, d). These variable group proteins might be caused by multiple isoforms or post-transcriptional modifications. For example, the spots #644, #654, and #655 were identified as phosphoglucomutase 1 (PGM1). The spot of #644 was slightly up-regulated in HCC (average HCC/non HCC ratio: 1.55), and #654 and #655 were down-regulated and unchanged (average ratio: −1.8, −1.1, each) in HCC, respectively. Both #654 and #655 were positively stained with ProQ Diamond which has been shown to detect phosphoserine-, phosphothreonine-, and phosphotyrosine-containing proteins [12, 19], whereas the #644 was negative for ProQ (Fig. 6). The integrated spot volume of #644, #654, and #655 were unchanged between non HCC and HCC (average ratio: 1.03). These observations suggest that the PGM1 does not show the different expression level between non HCC and HCC, but exhibits altered phosphorylation level in HCC.

3.3 Comparison of Protein and mRNA Expression Profile

The expression pattern of 188 proteins was compared with that of mRNA expression pattern obtained from the same tissue samples, and a total of 164 proteins had the

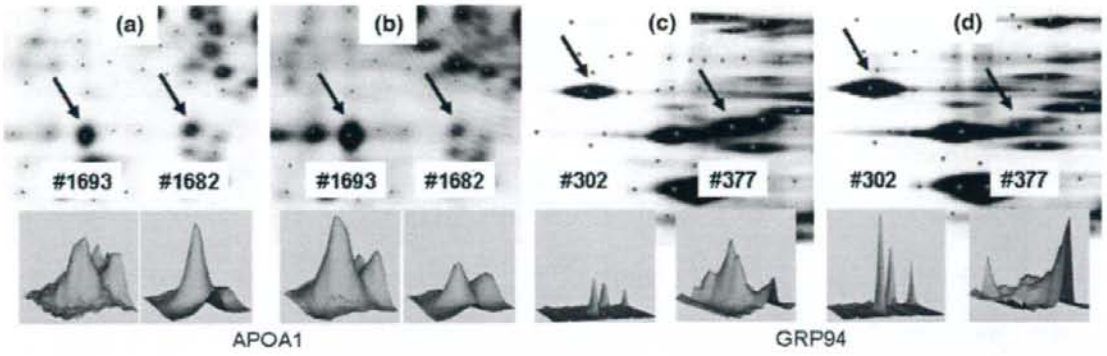

Fig. 5 Protein expressed as multiple spots. Apolipoprotein A1 (APOA1) in non HCC (**a**), in HCC (**b**), and heat shock protein 90 kDa beta (GRP94) in non HCC (**c**), in HCC (**d**) were detected as multi-spotted protein. Magnified images of selected spots and 3-D views of selected spots were shown

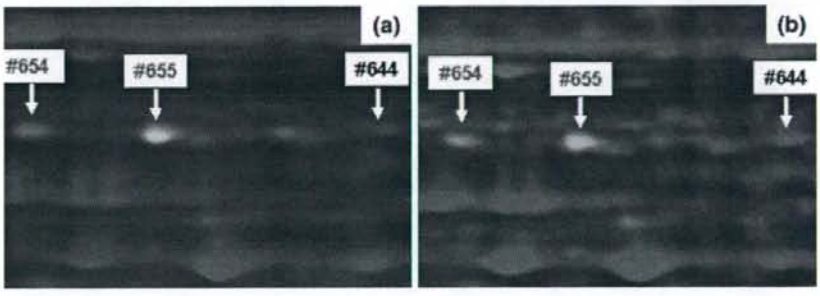

Fig. 6 Detection of phosphoglucomutase 1 with Ruby and ProQ Diamond. Phosphoglucomutase 1 (PGM1) was detected as three spots (#644, #654, and #655). The spot #644 was stained with Ruby (yellow) only, and the spot 654 and 655 were stained with Ruby and ProQ Diamonds (blue green) in non HCC (**a**) and HCC (**b**) gels

corresponding mRNA expression data. The protein expression abundance between non HCC and HCC was calculated using the average spot volume, and the statistical differences were analyzed by the Student's paired *t*-test using DeCyder BVA. The spot volume of a multi-spotted protein was indicated as a total volume by integrating the intensities of multiple spots. The mRNA expression abundance was calculated the ratio of the SAGE tag counts, and the SAGE tag counts were evaluated by Monte Carlo analysis whether each tag ratio between HCC and non HCC was statistically significant or not. Consequently, the mRNA levels for forty proteins out of 164 were shown to be altered significantly ($p \leq 0.05$) between non HCC and HCC (Table 1). These proteins were used to clarify the difference of transcriptional and translational regulation in HCC. The HCC/non HCC expression ratio for each protein was plotted against the mRNA ratio in log scale in Fig. 7a, where a positive value indicates increased expression in HCC and a negative ratio indicates reduced expression. We defined the protein and mRNA showed same (increase or reduce) expression change as similar group, and showed opposite expression as opposite group, respectively. The protein expression did not show significant change ($p > 0.05$) was defined as different group. Above that classification, 24 proteins showed similar protein and mRNA expression tendency between non HCC and HCC (Fig. 7a, open circles).

On the other hand, 11 proteins showed the opposite (Fig. 7a, closed circles) and 5 proteins showed the different (Fig. 7a, half-closed circles) expression patterns, respectively. Thus, about 60% of 40 proteins classified as similar group, 30% as opposite group, and 10% as different group (Fig. 7b). The similar and the opposite group proteins were classified according to their molecular functions and the cellular components (Fig. 8). The ratio of members categorized into each function differs slightly between the similar (Fig. 8a) and the opposite (Fig. 8c) group proteins. On the other hand, the result of classification by cellular components were apparently different each other (Fig. 8b, d).

4 Discussion

In the present study, a paired sample of HCC and non HCC from the single patient at the proteome and transcriptome level were compared. We used 188 identified proteins expression data and 4235 unique mRNA expression data,

Table 1 Proteins of significant mRNA expression abundance

Protein name	Theoretical	mRNA		Protein		Group[e]
	MW (kDa)/pI	Ratio[a]	*p* value[b]	Ratio[c]	*p* value[d]	
Carbamoyl-phospahte synthase 1	164.94/6.29	0.25	0.054	1.4	0.00046	opposite
Actinin, alpha4 (ACTN4)	105.29/5.3	2.1	0.014	1.2	0.013	similar
Formyltetrahydrofolate dehydrogenase	98.83/5.63	9	0.011	0.99	>0.05	different
Protein disulfide isomerase A4 (PDIA4)	72.93/4.96	0.2	0.029	2.5	0.00011	opposite
Heat shock 70 kDa protein 5 (HSPA5)	72.33/5.07	0.3	0.046	2.1	0.000013	opposite
Heat shock 70 kDa protein 8	71.08/5.28	1.9	0.013	1.3	0.0001	similar
Carboxylesterase 1 (CES1)	62.38/6.15	0.048	0.000013	0.55	0.00005	similar
Aldehyde dehydrogenase 4A1	61.72/8.25	3.5	0.016	1.01	>0.05	different
Glutamate dehydrogenase 1 (GLUD1)	61.4/7.66	0.1	0.000063	0.21	0.00000054	similar
Proline 4-hydroxylase (P4HB)	57.11/4.76	1.8	0.033	2.1	0.00000059	similar
ATP synthase H+ transporting	56.56/5.26	3	0.010	1.00	>0.05	different
Aldehyde dehydrogenase 9A1	54.7/5.7	0.2	0.029	0.84	0.00078	similar
Cytokeratin 8 (KRT8)	53.67/5.3	0.11	0.00011	0.41	0.00017	similar
Endothelial cell growth factor 1	49.96/5.36	3.5	0.016	0.43	0.00064	opposite
Heterogeneous nuclear ribonucleoprotein	46.04/5.38	0.17	0.017	0.98	>0.05	different
Betaine-homocysteine methyltransferase	45.43/6.41	10	0.0066	0.88	0.0091	opposite
Acyl-CoA dehydrogenase, C-2 to C-3	44.3/7.96	0.28	0.0012	0.51	0.00031	similar
Alanine-glyoxylate aminotransferase	43.39/8.61	0.33	0	0.78	0.044	similar
Acetyl-CoA acyltransferase 2	42.04/8.52	0.18	0	0.52	0.0007	similar
Actin beta (ACTB)	41.35/5.56	0.067	0.00029	0.74	0.000013	similar
Alcohol dehydrogenase 1B	39.84/8.53	0.18	0	0.44	0.000012	similar
Glyoxylate reductase	37.31/5.95	3.5	0.016	0.63	0.041	opposite
Aldo-keto reductase 1C3	36.85/8.06	3.3	0.047	1.5	0.025	similar
Malate dehydrogenase 1	36.62/5.92	0.33	0.023	1.2	0.00047	opposite
Apolipoprotein E (APOE)	36.25/5.6	7	0.037	1.9	0.013	similar
Lectin, galactoside-binding 4	36.03/9.21	10	0.0066	2.5	0.00045	similar
Malate dehydrogenase 2	35.5/8.92	0.3	0.046	0.73	0.00047	similar
Sulfotransferase family, cytosolic, 2A	33.95/5.71	0.25	0.0019	0.85	0.015	similar
Thiosulfate sulfurtransferase	33.43/6.77	0.38	0.027	0.97	>0.05	different
Peroxiredoxin 4 (PRDX4)	30.77/5.86	0.13	0.0039	1.2	0.032	opposite
Carbonyl reductase 1	30.37/8.55	4.5	0.032	1.5	0.000027	similar
Phosphoglycerate mutase 1	28.92/6.67	0.22	0.032	0.52	0.00016	similar
Proteasome activator subunit 1	28.72/5.78	0.14	0.037	1.4	0.0017	opposite
Glutathione-S-transferase omega 1	27.83/6.2	0.22	0.032	0.64	0.0000004	similar
Quinoid dihydropteridine reductase	26.02/6.9	0.2	0.029	0.78	0.00049	similar
Proteasome subunit alpha 2	25.99/6.9	0.17	0.000073	1.4	0.00035	opposite
Hemoglobin beta	16/6.74	0.23	0.00026	0.40	0.000034	similar
Hemoglobin alpha	15.17/8.7	0.4	0.019	0.44	0.000092	similar
Fatty acid binding protein 1	14.26/6.6	0.39	0	0.83	0.00027	similar
Cytochrome b5A (CYB5A)	11.27/5.02	0.15	0.0039	1.3	0.000062	opposite

[a] mRNA ratio represents SAGE tag counts ratio of HCC/non HCC

[b] SAGE *p* values were calculated by Monte Carlo analysis using SAGE 2000 software

[c] Protein ratio represents the average spot volume ratio of HCC/non HCC

[d] Protein *p* values were analyzed by the Student's paired *t*-test using DeCyder BVA software

[e] Protein and mRNA expression change showed same tendency represents as similar. Protein and mRNA expression change showed opposite tendency represents as opposite. Protein expression showed no significant change represents as different

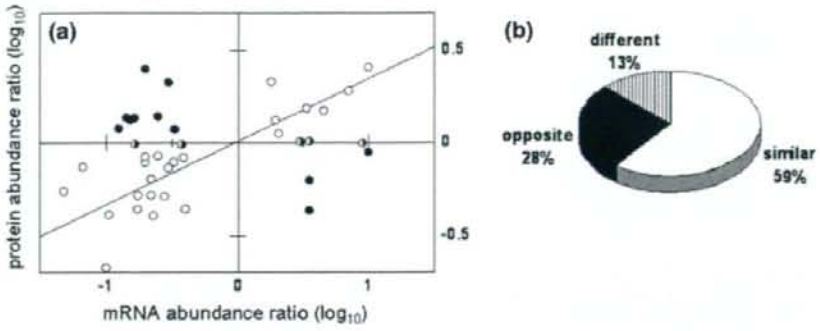

Fig. 7 Comparative analysis of protein and mRNA expression. (**a**) The log_{10} values of the HCC/non HCC 2D-DIGE gel spot volume ratio were plotted against that of the HCC/non HCC SAGE tag counts ratio which showed statistical significance of obtaining a difference in expression. The open circles (○) represent similar expression tendency of protein and mRNA between non HCC and HCC. The closed circles (●) showed opposite tendency, and the half-closed circles (◑) showed different tendency which represent SAGE ratio showed significant change but 2D-DIGE ratio did not between non HCC and HCC. (**b**) The pie chart represents the percentage of similar, opposite, and different tendency of 40 proteins

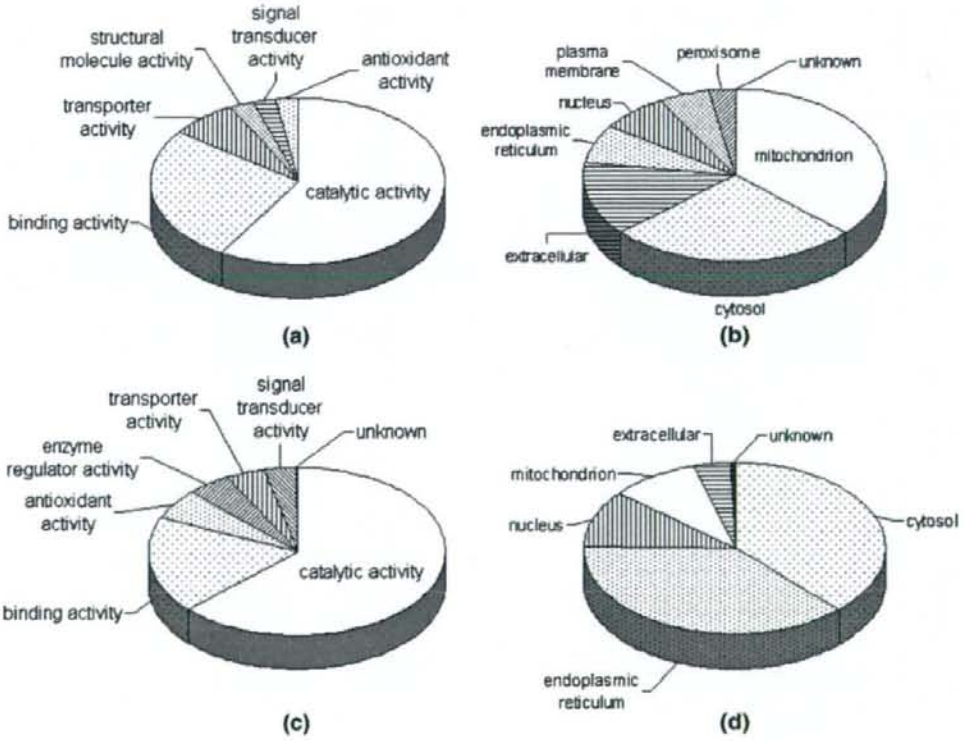

Fig. 8 Molecular functions and cellular components classification of the similar and the opposite proteins The molecular functions and cellular components classification were performed according to criteria described by Gene Ontology Consortium using BioCompass. (**a**) molecular functions of the similar proteins, (**b**) cellular components of the similar proteins, (**c**) molecular functions of the opposite proteins, (**d**) cellular component of the opposite proteins

eliminating multiple gene match tag, and protein data were linked to the corresponding mRNA data using GenBank accession, UniGene accession, and protein name. The mRNA data for 24 proteins were not found in the SAGE data set (Supplementary Table 1). These mRNA data might be excluded as ambiguous identification of SAGE tag

sequence. We classified identified proteins and genes according to their molecular functions and the cellular components (Fig. 4). It was noticed that many genes encoding ribosomal and plasma membrane proteins were identified in SAGE analysis (Fig. 4d), whereas those proteins were rarely identified in 2-DE analysis (Fig. 4b). Since proteins with extremely high pI or low pI as well as hydrophobic proteins are difficult to be detected by 2-DE in general, obtained expression data were biased in favor of cytoplasmic and soluble proteins.

To evaluate the protein expression, we used the ratio of average spot volume of 2D-DIGE between non HCC and HCC, and the 164 proteins expression were compared to corresponding mRNA expression using the SAGE tag count ratio between non HCC and HCC. The SAGE tag counts were tested by Monte Carlo analysis to determine the statistical significance of expression difference, and a total 40 of 164 proteins showed different mRNA expression significantly ($p \leq 0.05$) between non HCC and HCC. The expression changes of 24 proteins out of 40 between non HCC and HCC seem to be correlated with mRNA expression changes shown as SAGE tag frequency ratio (Fig. 7a, correlation coefficient $r = 0.86$). For example, both mRNA and protein expression of several enzymes abundant in liver, such as alcohol dehydrogenase, acetyl-CoA acyltransferase, and carboxylesterase were reduced (Table 1). Thus, these results suggest that the reduction of these proteins in HCC is regulated transcriptionally. This observation suggests that about 60% of the protein expression difference between non HCC and HCC could be explained by transcriptional regulation. Similar phenomenon of mRNA and protein expression abundance between the different cells was observed in mouse tissue samples [14].

On the other hand, about 30% of the proteins showed opposite expression pattern to the mRNA, and the cellular localization of these opposite group proteins were distinctly different to the similar group proteins. Many opposite group proteins were localized in endoplasmic reticulum, while the similar group proteins were localized in mitochondria (Fig. 8b, d). For example, protein disulfide isomerase A4 and heat shock 70 kDa protein 5 were opposite group proteins (Table 1), and were known as the member of multi-protein complex in mammalian endoplasmic reticulum [20]. Our result suggests that the expression of these proteins was not controlled at transcriptional level in HCC.

Nine multi-spotted proteins showed spot-to-spot variable expression patterns, and some of them were found to be differently modified post-translationally. Indeed, two out of three PGM1 spots were positive for ProQ staining, indicating that these two spots were phosphorylated (Fig. 6). One phosphorylated spot volume was reduced in HCC, whereas non phosphorylated spot volume was increased. The integrated three spot volume of PGM1 did not change between non HCC and HCC, and mRNA expression of PGM1 did not show the significant change between non HCC and HCC. Thus, protein and mRNA expression level of PGM1 per se did not differ between non HCC and HCC, though its phosphorylation level was clearly altered. Recently, PGM1 which is a critical regulator of cellular glucose utilization was shown to be regulated through phosphorylation by p21-activated kinase (Pak1) in a leukemia cell line [5].

The post-translational modification, such as phosphorylation, plays critical roles in cellular signaling, and its detection is expected to provide us to understand molecular basis for cellular functions. Since it is difficult to obtain the information about post-translational modification by mRNA analysis, the combined analysis with both proteome and transcriptome approaches will be important. The protein analysis based on 2-DE technology, which can detect and compare post-translationally modified proteins easily as shown here, is useful for this purpose.

In this study we compared the proteome and the transcriptome profiles from matched HCC and non HCC tissue samples from a single patient. Most of protein expression changes between non HCC and HCC seem to be explained by transcriptional regulation. Further research with a more protein and mRNA expression data of non HCC and HCC samples may yield a general view of the molecular event of HCC.

Acknowledgments We are grateful to Dr. K. Miyazaki and Dr. R. Teramoto for fruitful comments on this work, and N. Tetsura for technical assistance.

References

1. Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, Currie I (2003) Proteomics 1:36–44
2. Berven FS, Karisen OA, Murrell JC, Jensen HB (2003) Electrophoresis 24:757–761
3. Dash S, Saxena R, Myung J, Rege T, Tsuji H, Gaglio P, Garry RF, Thung SN, Gerber MA (2000) J Virol Methods 90:15–23
4. Griffin TJ, Gygi SP, Ideker T, Rist B, Eng J, Hood L, Aebersold R (2002) Mol Cell Proteomics 4:323–333
5. Gururaj A, Barnes CJ, Vadlamudi RK, Kumar R (2004) Oncogene 23:8118–8127
6. Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Mol Cell Biol 3:1720–1730
7. Iizuka N, Oka M, Yamada-Okabe H, Mori N, Tamesa T, Okada T, Takemoto N, Hashimoto K, Tangoku A, Hamada K, Nakayama H, Miyamoto T, Uchimura S, Hamamoto Y (2003) Oncogene 22:3007–3014
8. Kawai HF, Kaneko S, Honda M, Shirota Y, Kobayashi K (2001) Hepatology 3:676–691
9. Lee IN, Chen CH, Sheu JC, Lee HS, Huang GT, Yu CY, Lu FJ, Chow LP (2005) J Proteome Res 4:2062–2069
10. Liang CR, Leow CK, Neo JC, Tan GS, Lo SL, Lim JW, Seow TK, Lai PB, Chung MC (2005) Proteomics 8:2258–2271

11. Lopez LJ, Marrero JA (2004) Curr Opin Gastroenterol 3:248–253
12. Martin K, Steinberg TH, Cooley LA, Gee KR, Beechem JM, Patton WF (2003) Proteomics 3:1244–1255
13. Midorikawa Y, Makuuchi M, Tang W, Aburatani H (2007) World J Gastroenterol 13:1487–1492
14. Mijalski T, Harder A, Halder T, Kersten M, Horsch M, Strom TM, Liebscher HV, Lottspeich F, de Angelis MH, Beckers J (2005) Proc Natl Acad Sci USA 102:8621–8626
15. Saha S, Sparks AB, Rago C, Akmaev V, Wang CJ, Vogelstein B, Kinzler KW, Velculescu VE (2002) Nat Biotechnol 5:508–512
16. Seow TK, Liang RCMY, Leow CK, Chung MCM (2001) Proteomics 1:1249–1263
17. Shackel NA, Seth D, Haber PS, Gorrell MD, McCaughan GW (2006) Comp Hepatol 5:6
18. Stahl S, Ittrich C, Marx-Stoelting P, Köhle C, Altug-Teber O, Riess O, Bonin M, Jobst J, Kaiser S, Buchmann A, Schwarz M (2005) Hepatology 2:353–361
19. Steinberg TH, Agnew BJ, Gee KR, Leung WY, Goodman T, Schulenberg B, Hendrickson J, Beechem JM, Haugland RP, Patton WF (2003) Proteomics 3:1128–1144
20. Stockton JD, Merkert MC, Kellaris KV (2003) Biochemistry 42:12821–12834
21. Unlu M, Morgan ME, Minden JS (1997) Electrophoresis 11:2071–2077
22. Yamashita T, Kaneko S, Hashimoto S, Sato T, Nagai S, Toyoda N, Suzuki T, Kobayashi K, Matsushima K (2001) Biochem Biophys Res Commun 282:647–654
23. Zanella I, Rossini A, Domenighini D, Albertini A, Cariani E (2002) J Med Virol 68:494–499

ELSEVIER

Journal of Hepatology 50 (2009) 100–110

Journal of Hepatology

www.elsevier.com/locate/jhep

Activation of lipogenic pathway correlates with cell proliferation and poor prognosis in hepatocellular carcinoma☆

Taro Yamashita[1], Masao Honda[1,2], Hajime Takatori[1], Ryuhei Nishino[1], Hiroshi Minato[3], Hiroyuki Takamura[4], Tetsuo Ohta[4], Shuichi Kaneko[1,*]

[1]*Department of Gastroenterology, Kanazawa University Graduate School of Medical Science, 13-1 Takara-Machi, Kanazawa 920-8641, Japan*
[2]*Department of Advanced Medical Technology, Kanazawa University School of Health Sciences, 13-1 Takara-Machi, Kanazawa 920-8641, Japan*
[3]*Pathology Section, Kanazawa University Hospital, 13-1 Takara-Machi, Kanazawa 920-8641, Japan*
[4]*Department of Gastroenterologic Surgery, Kanazawa University Graduate School of Medical Science, 13-1 Takara-Machi, Kanazawa 920-8641, Japan*

***Background/Aims*: Metabolic dysregulation is one of the risk factors for the development of hepatocellular carcinoma (HCC). We investigated the activated metabolic pathway in HCC to identify its role in HCC growth and mortality.**

***Methods*: Gene expression profiles of HCC tissues and non-cancerous liver tissues were obtained by serial analysis of gene expression. Pathway analysis was performed to characterize the metabolic pathway activated in HCC. Suppression of the activated pathway by RNA interference was used to evaluate its role in HCC *in vitro*. Relation of the pathway activation and prognosis was statistically examined.**

***Results*: A total of 289 transcripts were up- or down-regulated in HCC compared with non-cancerous liver ($P < 0.005$). Pathway analysis revealed that the lipogenic pathway regulated by sterol regulatory element binding factor 1 (*SREBF1*) was activated in HCC, which was validated by real-time RT-PCR. Suppression of *SREBF1* induced growth arrest and apoptosis whereas overexpression of *SREBF1* enhanced cell proliferation in human HCC cell lines. *SREBF1* protein expression was evaluated in 54 HCC samples by immunohistochemistry, and Kaplan–Meier survival analysis indicated that *SREBF1*-high HCC correlated with high mortality.**

***Conclusions*: The lipogenic pathway is activated in a subset of HCC and contributes to cell proliferation and prognosis.**

***Keywords*: Hepatocellular carcinoma; Serial analysis of gene expression; Lipogenesis; Gene expression profiling; Sterol regulatory element binding factor 1**

Received 26 May 2008; received in revised form 1 July 2008; accepted 23 July 2008; available online 12 October 2008
Associate Editor: J.M. Llovet

☆ The authors who have taken part in the research of this paper declared that they do not have a relationship with the manufacturers of the materials involved either in the past or present and they did not receive funding from the manufacturers to carry out their research.

* Corresponding author. Tel.: +81 76 265 2231; fax: +81 76 234 4250.
E-mail address: skaneko@m-kanazawa.jp (S. Kaneko).

Abbreviations: HCC, hepatocellular carcinoma; *SREBF1*, sterol regulatory element binding factor 1; HBV, hepatitis B virus; HCV, hepatitis C virus; SAGE, serial analysis of gene expression; RT-PCR, reverse transcription-polymerase chain reaction; IHC, immunohistochemistry; FADS1, fatty acid desaturase 1; SCD, stearoyl CoA desaturase; FASN, fatty acid synthase; si-RNA, short interfering-RNA; CLD, chronic liver disease; PCNA, proliferating cell nuclear antigen; IGF, insulin-like growth factor.

1. Introduction

Hepatocellular carcinoma (HCC) is one of the most frequently occurring malignancies in the world [1]. The major risk factors associated with HCC include chronic infection with hepatitis B virus (HBV) and hepatitis C virus (HCV), alcohol abuse, and exposure to aflatoxin B1 [2]. HCC usually develops from liver cirrhosis, which involves continuous inflammation and hepatocyte regeneration, suggesting that reactive oxygen species and DNA damage are involved in the process of hepatocarcinogenesis [3].

The development of gene expression profiling technologies including DNA microarrays and serial analysis

0168-8278/$34.00
doi:10.1016/j.jhep.2008.07.036

of gene expression (SAGE) have enhanced our ability to identify inventory transcripts and global genetic alterations in HCC [4–10]. In general, these methods have demonstrated that transcripts associated with cell growth are up-regulated, whereas those related to inhibition of cell growth are down-regulated, in HCC [11]. It is difficult, however, to decipher molecular pathways activated during hepatocarcinogenesis.

Epidemiological studies suggest that metabolic dysregulation in the liver increases the risk of HCC development. For example, diabetes is associated with a 2-fold increase in the risk of HCC [12]. Obesity and hepatic steatosis also increase the risk of HCC [13–15]. Furthermore, recent studies indicate that HCV infection provokes hepatic steatosis, which may be a vulnerable factor for liver inflammation and HCC development [16,17]. Thus, dysregulation of a metabolic pathway may play a crucial role to promote HCC growth, but the molecular mechanism is still obscure. In this study, we have utilized SAGE [18,19], which enables us to monitor the differential expression of all genes, to determine the global changes in gene expression that occur during hepatocarcinogenesis.

2. Materials and methods

2.1. Tissue samples

All HCC tissues, adjacent non-cancerous liver tissues, and normal liver tissues were obtained from 69 patients who underwent hepatectomy from 1997 to 2005 in Kanazawa University Hospital. Normal liver tissue samples were obtained from patients undergoing surgical resection of the liver for treatment of metastatic colon cancer. HCC and surrounding non-cancerous liver samples were obtained from patients undergoing surgical resection of the liver for the treatment of HCC. The samples used for SAGE, real-time reverse-transcription (RT)-PCR analysis, and immunohistochemistry (IHC) are listed in Supplemental Table 1. All samples used for SAGE and real-time RT-PCR analysis were snap-frozen in liquid nitrogen. Four normal liver tissues and 20 HCCs and their corresponding non-cancerous liver tissues were used for real-time RT-PCR analysis; seven of these HCC samples, along with 47 additional HCC samples, were formalin-fixed paraffin-embedded and used for IHC. HCC and adjacent non-cancerous liver were histologically characterized as described [20].

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, and each patient provided written informed consent.

2.2. SAGE

Total RNA was purified from each homogenized tissue sample using a ToTally RNA extraction kit (Ambion, Inc., Austin, TX), and polyadenylated RNA was isolated using a MicroPoly (A) Pure kit (Ambion). A total of 2.5 μg mRNA per sample was analyzed by SAGE [18]. SAGE libraries were randomly sequenced at the Genomic Research Center (Shimadzu-Biotechnology, Kyoto, Japan), and the sequence files were analyzed 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 by SAGE 2000 software. Monte Carlo simulation was used to select genes with significant differences in expression between two libraries without multiple hypothesis testing correction ($P < 0.005$) [21]. Each SAGE tag was annotated using a gene-mapping web site (http://www.ncbi.nlm.nih.gov/SAGE/index.cgi).

2.3. Analysis of signaling networks

Ingenuity Pathways Analysis software (Ingenuity® Systems, www.ingenuity.com) was used to investigate the molecular pathways activated in an HCC SAGE library compared with an adjacent non-cancerous liver SAGE library. All reliable transcripts statistically up-regulated in HCC were investigated and annotated with biological processes, protein-protein interactions, and gene regulatory networks, using a reference-based data file with statistical significance. All identified pathways were screened individually. MetaCore™ software (GeneGo Inc., St. Joseph, MI) was used to evaluate candidate transcription factors responsible for up-regulation of transcripts in HCC.

2.4. RT-PCR

A 1-μg aliquot of each total RNA was reverse-transcribed using SuperScript II reverse-transcriptase (Invitrogen, Carlsbad, CA). Real-time RT-PCR analysis was performed using ABI PRISM 7900 Sequence Detection System (Applied Biosystems, Foster City, CA). Using the standard curve method, quantitative **PCR** was performed in triplicate for each sample-primer set. Each sample was normalized relative to β-actin. The assay IDs used were Hs00231674_m1 for sterol regulatory element binding factor 1 (*SREBF1*); Hs00203685_m1 for fatty acid desaturase 1 (*FADS1*) ;Hs00748952_s1 for stearoyl CoA desaturase (*SCD*); Hs00188012_m1 for fatty acid synthase (*FASN*); and Hs99999_m1 for β-actin. *SREBF1a* and *SREBF1c* mRNA levels were assayed by semi-quantitative RT-PCR [22].

2.5. RNA Interference targeting SREBF1

Si-RNAs targeting *SREBF1* were constructed using a *Silencer*™ SiRNA Construction kit (Ambion) according to the manufacturer's protocol. We constructed two different si-RNAs, targeting different sites of *SREBF1* (*SREBF1*-1; CAGTGGCACTGACTCTTCC, *SREBF1*-2; TCTACGACCAGTGGGACTG). Control si-RNA duplexes targeting scramble sequences were also synthesized (Dharmacon Research, Inc., Lafayette, CO). Lipofectamine 2000™ reagent (Invitrogen) was used for transfection according to the manufacturer's instructions.

2.6. Cell proliferation assay

Cell proliferation assays were performed using a Cell Titer96 Aqueous kit (Promega, Madison, WI). Results are expressed as the mean optical density (OD) of each five-well set. All experiments were repeated at least twice.

2.7. Soft agar assay

To each well of a six-well plate, containing a base layer of 0.72% agar in growth medium, was added 1×10^4 cells, suspended in 2 ml of 0.36% agar with growth medium (DMEM supplemented with 10% FBS), and the plates were incubated at 37 °C in a 5% CO_2 incubator for 2 weeks. The numbers of colonies in each well were counted as previously described [23].

2.8. TUNEL assay

A DeadEnd™ Colorimetric TUNEL System (Promega) was used to measure nuclear DNA fragmentation as described previously [24].

2.9. Annexin V staining

To evaluate apoptotic cell death, Annexin V binding to cell membranes was evaluated using Annexin V-FITC antibodies and FACSCalibur flow cytometer (BD Biosciences, Franklin Lakes, NJ), as described by the manufacturer.

2.10. Focus assay

HuH7 cells and Hep3B cells were transiently transfected with pCMV7 or pCMV7-*SREBF1c* vectors (kindly provided by Dr. Hitoshi Shimano) using Lipofectamine 2000™ reagent (Invitrogen), as described by the manufacturer. A total of 2×10^3 cells were seeded on six-well plates 48 h after transfection, and cultured in usual media with 400 ng/ml of Geneticin for 9 days. The foci were fixed with ice-cold 100% methanol and stained with 0.5% crystal violet solution. All experiments were performed in triplicates.

2.11. Western blotting

Whole cell lysates were prepared using RIPA lysis buffer. Antibodies used were rabbit polyclonal antibodies to phospho-GSK-3β (ser9) (Cell Signaling Technology Inc., Danvers, MA), rabbit anti-sterol regulatory element binding protein-1 (encoded by *SREBF1*) polyclonal antibody H-160 (Santa Cruz Biotechnology, Inc., Santa Cruz, CA), and β-actin (Sigma–Aldrich Japan K.K., Tokyo, Japan). Immune complexes were visualized by enhanced chemiluminescence (Amersham Biosciences Corp., Piscataway, NJ) as described in the manufacturer's protocol.

2.12. Immunohistochemistry

Rabbit anti-*SREBF1* polyclonal antibody H-160 (Santa Cruz Biotechnology, Inc.) and mouse anti-proliferating cell nuclear antigen (PCNA) monoclonal antibody PC10 (Calbiochem, San Diego, CA) were used to evaluate the immunoreactivity of HCC samples, using a DAKO EnVision+™ Kit, as described by the manufacturer. The signal intensity of *SREBF1* was scored as negative, low, or high determined by the representative staining of the normal liver tissue and cirrhotic liver tissue (Supplemental Fig. 1). HCC was referred as *SREBF1*-high if *SREBF1* expression in the tumor was higher than that in the cirrhotic liver tissue. PCNA index was evaluated as previously described [25].

2.13. Statistical analysis

Kruskal–Wallis test was used to compare the differentially expressed genes, as shown by real-time PCR, among normal liver, CLD, and HCC tissues. Mann–Whitney U test was also used to evaluate the statistical significance of differences of gene expression between CLD and HCC tissues. Spearman's correlation coefficient was used to assess correlations between the expression levels of *SREBF1*, *FADS1*, *SCD*, and *FASN*. Univariate Cox proportional hazards regression analysis was used to evaluate the association of gene expression and clinicopathologic parameters with patient outcomes. All statistical analyses were performed using SPSS software (SPSS software package; SPSS Inc., Chicago, IL) and GraphPad Prism software (GraphPad Software Inc., La Jolla, CA).

3. Results

3.1. Gene expression profiling of HCC

We constructed two SAGE libraries from a HCC–HBV tissue and a corresponding non-cancerous tissue (chronic liver disease (CLD)–HBV). We also used two previously described SAGE libraries, from an HCC–HCV sample and a corresponding non-cancerous tissue sample (CLD–HCV) [4]. After excluding tags detected only once in each library, to avoid the contamination of tags derived from sequence errors, we selected 105,288 tags corresponding to the 9731 genes in all libraries. Using Monte Carlo simulation, we compared the differentially expressed transcripts in HCC and corresponding CLD libraries. Compared with their corresponding CLD libraries, there were statistically significant increases or decreases in 140 transcripts in the HCC–HBV library and in 197 transcripts in the HCC–HCV library ($P < 0.005$).

The HCC–HBV library contained one SAGE tag encoding the HBV–X region, which was increased more than 35-fold compared with its expression in the corresponding CLD–HBV library (Supplemental Table 2). We identified two additional SAGE tags, encoding unknown genes (GTTCTAAAGG, GCATTATGAT), which were expressed more than 10-fold in the HCC–HBV library than in the corresponding CLD–HBV library. The HCC–HBV library also contained tags associated with lipogenesis, at greater than 10-fold abundance, in the HCC–HBV library; these including tags for steroyl-CoA desaturase, fatty acid synthase, and fatty acid desaturase 1.

In contrast, SAGE tags associated with the immune response were up-regulated in the HCC–HCV library. These included tags for Th1-type chemokines, including chemokine ligand 10 (C–X–C motif), chemokine ligand 9 (C–X–C motif), and major histocompatibility complex classes IA and IB (Supplemental Table 3). In addition, tags associated with lipogenesis were increased in the HCC–HCV library, including tags for 3-hydroxy-3-methylglutaryl-coenzyme A synthase 1 and cytochrome P450, family 51, subfamily A, polypeptide 1. Taken together, the differential gene expression patterns may exist in HCC–HBV and HCC–HCV. HBV–X and lipogenesis-related genes are activated in HCC–HBV, whereas genes associated with inflammation as well as lipogenesis are activated in HCC–HCV.

3.2. Analysis of molecular pathways activated in HCC

To further characterize the gene expression patterns of HCC–HBV and HCC–HCV, we performed pathway analysis on SAGE data. Using MetaCore™ software, we found that the candidate transcription factors activated were distinct in each HCC library (Table 1). Several of these transcription factors, including NF-κB, c-Myc, c-Jun, and HNF4-α, have been reported to be activated in HCC [26–29]. In addition, our findings indicated that the transcription factor *SREBF1* may be activated in both HCC–HBV and HCC–HCV (to avoid a confusion, we use HUGO symbol *SREBF1* to indicate both gene/protein name).

Table 1
Candidate transcription factors that regulate molecular pathways activated in HCC.

SAGE library	Transcription factor	Molecular processes	*P*-value
HCC–HCV	NF-κB	Antigen presentation Antigen processing Defense response Immune response	0.004
	SREBF1	Cholesterol biosynthesis Lipid biosynthesis β-Glucoside transport Negative regulation of lipoprotein metabolism	0.05
	SP1	Electron transport; drug metabolism Oxygen and reactive oxygen species metabolism Cell-substrate junction assembly; wound healing	0.05
	IRF1	Immune response Antigen presentation; antigen processing Defense response; positive regulation of cell	0.05
HCC–HBV	HNF4-α	Lipid transport Fatty acid metabolism Smooth muscle cell proliferation	0.002
	HNF1	Acute-phase response; lipid transport Negative regulation of lipid catabolism β-Glucoside transport Negative regulation of lipoprotein metabolism	0.01
	SP1	Zinc ion homeostasis; response to biotic stimulus Nitric oxide mediated signal transduction Copper ion homeostasis; fatty acid biosynthesis	0.01
	c-Jun	Progesterone catabolism; progesterone metabolism Regulation of lipid metabolism; Prostaglandin metabolism	0.03
	C/EBP-α	Lipid transport; negative regulation of lipid catabolism Negative regulation of lipoprotein metabolism β-Glucoside transport Positive regulation of interleukin-8 biosynthesis	0.03
	SREBF1	Lipid biosynthesis; fatty acid biosynthesis Fatty acid metabolism Negative regulation of lipid catabolism Negative regulation of lipoprotein metabolism	0.03
	c-Myc	Fatty acid biosynthesis; fatty acid metabolism Fatty acid desaturation; Activation of pro-apoptotic gene products Release of cytochrome c from mitochondria	0.03
	USF1	Fatty acid metabolism Smooth muscle cell proliferation	0.03
	PPAR-α	Fatty acid metabolism Smooth muscle cell proliferation	0.03
	COUP-TFI	Lipid transport Smooth muscle cell proliferation	0.03
	C/EBP-β	Acute-phase response Regulation of interleukin-6 biosynthesis Fat cell differentiation Inflammatory response	0.03

These findings were evaluated by other pathway analysis software, Ingenuity Pathways Analysis (IPA). We applied the signaling network analysis to the transcripts up-regulated in the HCC libraries ($P < 0.005$). We found that the top signaling network activated in HCC–HBV contained several pathways involved in ERK/MAPK signaling, PPAR signaling, linoleic acid metabolism, and fatty acid metabolism (Supplemental Fig. 2A). Similarly, pathways involved in interferon signaling, NF-κB signaling, antigen presentation, PPAR signaling, linoleic acid metabolism, and fatty acid metabolism were included in the top signaling network activated in HCC–HCV (Supplemental Fig. 2B). Consistent with the results of transcription factor analysis by MetaCore™, pathway analysis indicated that *SREBF1* participates in the lipogenesis pathway in both HCC–HBV and HCC–HCV (blue nodes in Supplemental Fig. 2A and B). *SREBF1*, a major regulator of the lipogenesis pathway, binds to sterol regulatory elements on the genome [30], but less is known about its role in

HCC [31]. We therefore focused on the role of *SREBF1* signaling in HCC.

3.3. Validation of SAGE and signaling network analysis

We performed real-time RT-PCR analysis of *SREBF1* and three representative target genes (*SCD*, *FADS1*, and *FASN*) [20] on 44 samples not used for SAGE. We found that the levels of *SREBF1*, *SCD*, and *FASN* mRNAs were higher in HCC tissues and CLD tissues compared with normal liver, and that these differences were statistically significant (Fig. 1A). We further compared the expression of *SREBF1*, *FADS1*, and *FASN* between HCC and non-cancerous liver tissues, and identified the overexpression of *SREBF1* in HCC with statistical significance (Supplemental Fig. 3). Scatter plot analysis showed that the expression levels of *SREBF1* were correlated with those of *FADS1* ($R = 0.57$, $P < 0.0001$), *SCD* ($R = 0.82$, $P < 0.0001$), and *FASN* ($R = 0.74$, $P < 0.0001$) (Fig. 1B).

Since the mammalian genome encodes two *SREBF1* isoforms, *SREBF1a* and *SREBF1c* [22], we performed semi-quantitative RT-PCR with isoform specific primers to determine which of these isoforms was up-regulated in HCC. We found that *SREBF1c* mRNA, but not *SREBF1a* mRNA, was up-regulated in HCC compared with adjacent non-cancerous liver and normal liver tissues (Supplemental Fig. 4A).

3.4. Functional assay of the lipogenesis pathway in cell lines

Although genome-wide expression profiling showed that the lipogenesis pathway was activated in HCC possibly through up-regulation of *SREBF1*, it was not clear that this pathway played a role in HCC growth. To investigate the role of lipogenesis in HCC cell proliferation, we transfected two short interfering (si)-RNAs (*SREBF1*-1 and *SREBF1*-2) targeting *SREBF1* into the HuH7 and Hep3B cells. These cell lines have no chromosome amplification or deletion on 17p11, on which *SREBF1* is located [32]. Transfection of the si-RNA constructs for *SREBF1*-1 or *SREBF1*-2 decreased expression of *SREBF1* 90% and 70%, respectively, and the expression of both *SCD* and *FADS1* 70% and 60%, respectively (Fig. 2A). Because differences in *SREBF1c* and *SREBF1a* *sequence alignments* are very small, we could not design si-RNAs specifically targeting *SREBF1c*. We therefore checked the effect of si-RNAs on the expression of the *SREBF1* isoforms. We found that the expression of *SREBF1c* was relatively more suppressed than that of *SREBF1a* (Supplemental Fig. 4B), which may have been associated with the higher expression of *SREBF1a* than *SREBF1c* in cultured cell lines [25].

We found that the growth of these transfected cells was significantly inhibited at 72 h compared with mock transfected cells (Fig. 2B and Supplemental Fig.5A). Examination of anchorage independent cell growth showed strong suppression by deactivation of the lipogenesis pathway (Fig. 2C). Because insulin-like growth factor (IGF) is known to induce cancer cell proliferation through activation of PI3-kinase signaling followed by *SREBF1* induction, we investigated the effect of *SREBF1* knockdown on IGF2 mediated cell proliferation. Interestingly, *SREBF1* knockdown abrogated the IGF2 dependent cell proliferation (Supplemental Fig. 5B). Moreover, both the TUNEL assay and annexin V staining showed that transfection of *SREBF1* si-RNAs increased apoptosis compared with mock transfected cells (Fig. 2D and E).

We further investigated the role of *SREBF1* overexpression on cell growth *in vitro*. We transiently transfected control pCMV7 plasmids or pCMV7-*SREBF1c* plasmids (Fig. 3A), and cell proliferation was enhanced in *SREBF1* overexpressing cells compared with the control in both HuH7 and Hep3B cells evaluated by focus assay (Fig. 3B and supplemental Fig. 6). Furthermore, overexpression of *SREBF1* intensified the phosphorylation of GSK-3β, one of the major kinase phosphorylated by the activation of IGF signaling, in a dose-dependent manner (Fig. 3C).

3.5. SREBF1 Expression and prognosis

Since the above results indicated that *SREBF1* signaling may play an important role on tumor cell growth, we investigated the relationship between *SREBF1* expression and mortality in 54 HCC patients by IHC. When we examined the expression of *SREBF1* in HCC tissues and adjacent non-cancerous liver tissues, we identified the increase of the cytoplasmic *SREBF1* staining in a subset of HCC (Fig. 4A). We evaluated the expression of *SREBF1* in HCC and classified 4, 30, and 20 HCCs as *SREBF1*-negative, *SREBF1*-low, and *SREBF1*-high HCC, respectively (Fig. 4B and Supplemental Fig. 1). We could not detect any differences of clinico-pathological characteristics between *SREBF1*-high HCC and *SREBF1*-low/-negative HCC including histological steatosis (Supplemental Table4). Since the seven of these HCC samples were also used for real-time RT-PCR analysis, we investigated the relation of *SREBF1* RNA and protein expression (Fig. 4C). *SREBF1*RNA expression was significantly higher in *SREBF1*-high HCC than in *SREBF1*-low/-negative HCC with statistical significance ($P = 0.03$). Then we examined the cell proliferation of these HCC samples by PCNA staining. Notably, PCNA indexes were significantly higher in *SREBF1*-high HCC than *SREBF1*-low/-negative HCC with statistical significance ($P < 0.001$) (Fig. 4D). We further investigated the relationship between *SREBF1*

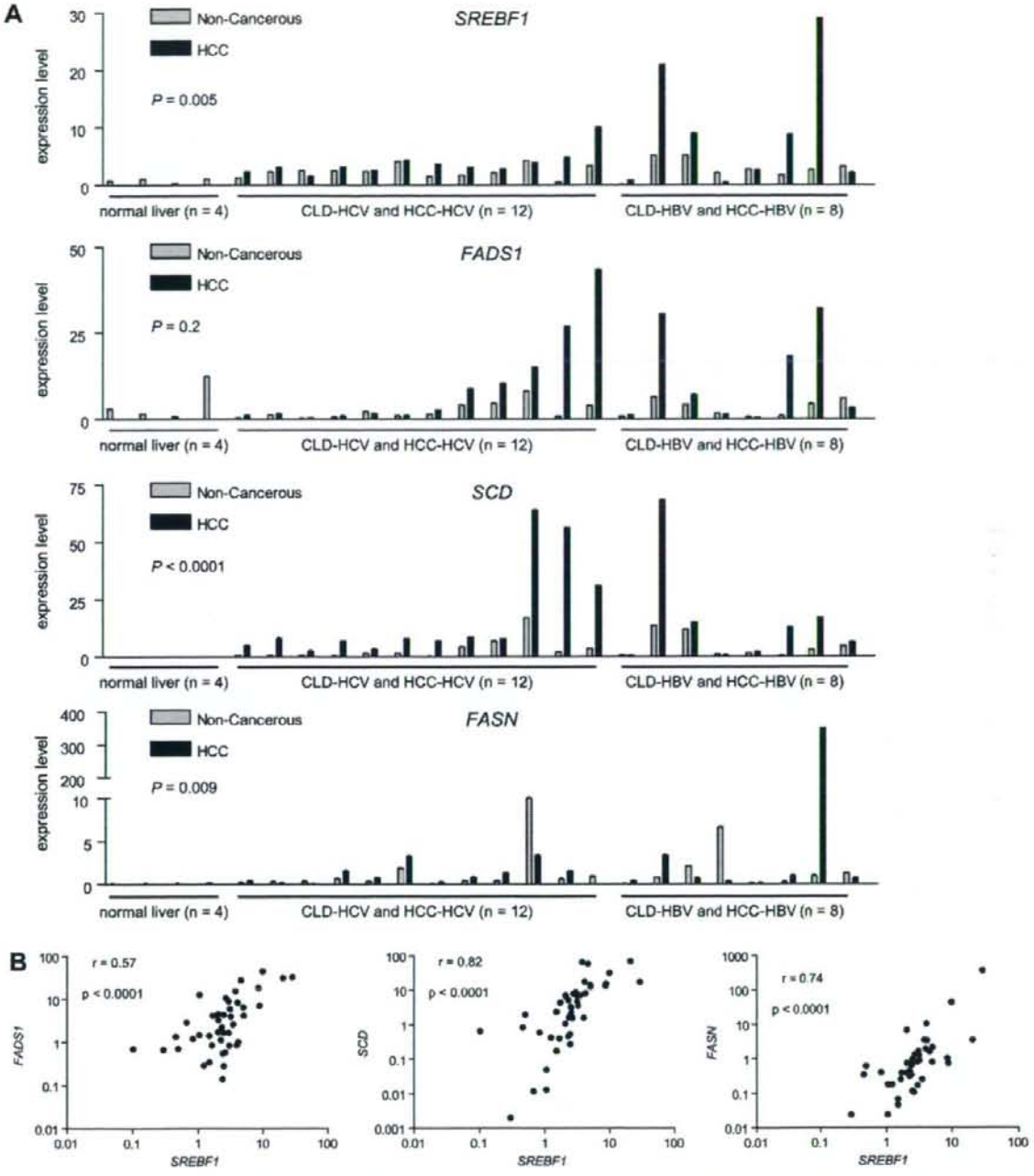

Fig. 1. (A) Real-time quantitative RT-PCR analysis. RNA was isolated from 44 tissue samples: 20 HCC, 20 corresponding CLD, and four normal liver samples. Differential expression of each gene among normal liver tissues, CLD tissues, and HCC tissues was examined by Kruskal–Wallis tests. (B) Scatter plot analysis. Gene expression levels of *FADS1*, *SCD* and *FASN* were well-correlated with those of *SREBF1*, as shown by Spearman's correlation coefficients.

protein expression and prognosis. Kaplan–Meier survival analysis showed a significant relationship between poor survival and high *SREBF1* protein expression ($P = 0.04$; Fig. 4E). Univariate Cox regression analysis showed a correlation between high *SREBF1* protein expression and high risk of mortality with statistical

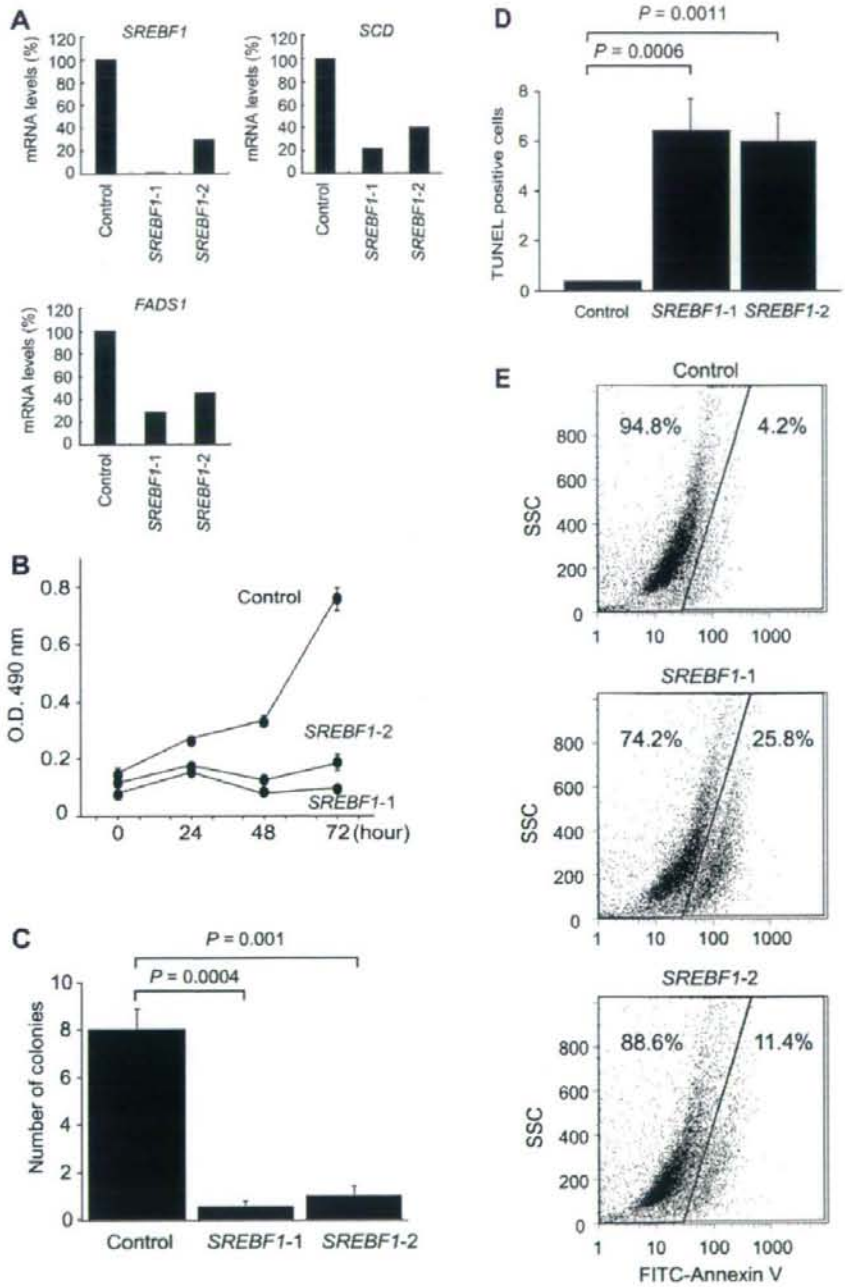

Fig. 2. (A) Effect of RNA interference targeting *SREBF1* in HuH7 cells. Expression levels of *SREBF1* mRNA were reduced by si-RNAs targeting different exons in *SREBF1*. Transcripts of *FADS1* and *SCD* were also down-regulated, showing transcriptional deactivation of the lipogenesis pathway. (B) Cell proliferation assay. Deactivation of the lipogenesis pathway severely reduced cell growth in HuH7 cells. (C) Soft agar assay. Deactivation of the lipogenesis pathway inhibited anchorage independent cell growth in HuH7 cells. (D) TUNEL assay. Deactivation of the lipogenesis pathway significantly increased the number of TUNEL-positive cells in HuH7 cells. (E) Annexin V staining evaluated by flow cytometer. Deactivation of the lipogenesis pathway significantly increased the number of annexin V positive cells in HuH7 cells.

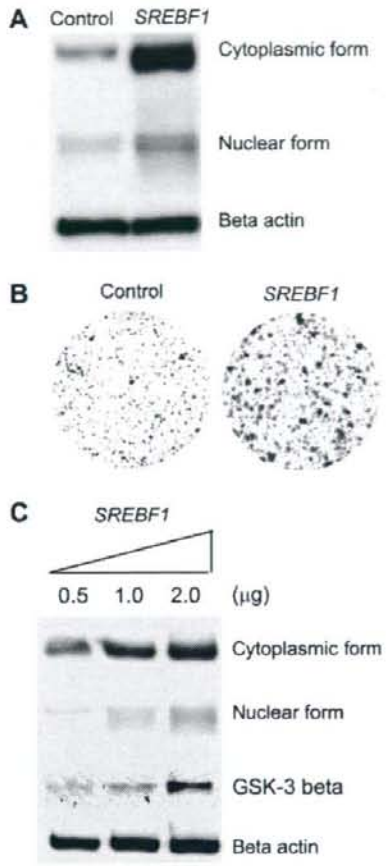

Fig. 3. (A) Western blot analysis of *SREBF1* protein expression in HuH7 cells transfected with control pCMV7 plasmids or pCMV7-*SREBF1c* plasmids. Both cytoplasmic and nuclear forms of *SREBF1* protein expression were increased by pCMV7-*SREBF1c* overexpression. (B) Focus assay of HuH7 cells transfected with control pCMV7 plasmids or pCMV7-*SREBF1c* plasmids. (C) Western blot analysis of *SREBF1* and phospho-GSK-3β protein expression in HuH7 cells transfected with indicated amounts of pCMV7-*SREBF1c* plasmids.

significance (HR, 3.7; 95% CI, 1.0–13.7; $P = 0.05$; Table 2).

4. Discussion

Using large-scale gene expression profiling, we have shown that the lipogenesis pathway is transcriptionally activated in HCC. Our SAGE profiles will be available on our homepage (http://www.intmedkanazawa.jp/) and will be submitted to the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/).

We found that the levels of expression of *FADS1*, *SCD*, and *FASN* were each correlated with those of *SREBF1*, suggesting that *SREBF1* is one of the main factors involved in the activation of lipogenesis in HCC. Activation of growth signaling pathways, such as the PI 3-kinase and mitogen-activated protein kinase pathways, has been shown to induce up-regulation of *SREBF1* in prostate and breast cancer cells [33,34]. We have observed induction of *SREBF1* protein expression by IGF2 in HuH7 cells (data not shown). Furthermore, we have identified that *SREBF1* overexpression results in the activation of cell proliferation and PI 3-kinase signaling, whereas expression inhibition of *SREBF1* abrogated the IGF2 induced cell proliferation. Although detailed mechanisms should be clarified in future, our results suggest that *SREBF1* is a key component of PI 3-kinase signaling in HCC.

SREBF1 is induced by alcohol [35], insulin, and fat [30,36], and plays a central role in the mechanism of hepatic steatosis [37]. Interestingly, these *SREBF1* inducers are risk factors for HCC [12,13,38,14]. Strikingly, two recent studies have shown that HBV and HCV infection may also induce hepatic steatosis through activation of *SREBF1* [39,40]. Furthermore, a recent report revealed the activation of *SREBF1* signaling in cancer by hypoxia [41]. Thus, these pathologic conditions such as chronic viral hepatitis, alcohol abuse, obesity, diabetes, and local hypoxia may up-regulate the expression of *SREBF1*, which, in turn, may contribute to an increased risk of hepatocarcinogenesis. Transgenic mice overexpressing *SREBF1* in the liver exhibited hepatic steatosis and hepatomegaly, suggesting the role of *SREBF1* on lipid metabolism and cell proliferation. However, it should be noted that no transgenic mice overexpressing *SREBF1* have been reported to have the risk of HCC development thus far. Interestingly, a recent report indicated that HCV core transgenic mice known to develop HCC showed coordinated activation of lipogenic pathway genes and *SREBF1* [42]. Although further studies are clearly required, we speculate that the activation of *SREBF1* may contribute to promote the development of HCC in already-initiated hepatocytes but not in normal hepatocytes.

Recently, Yahagi et al. reported the activation of lipogenic enzyme related genes in HCC [31]. In that paper, the authors suggested that *SREBF1* expression was not correlated with the expression of other lipogenic genes by Northern blotting, inconsistent with our current data. One possible explanation of these discrepancies might be the different methods for quantitation of mRNA, and we believe that real-time RT-PCR method used in our study would be more accurate. In addition, we evaluated the expression of *SREBF1* and lipogenic genes using more samples (a total of 44 liver and HCC tissues) than Yahagi et al did (10 HCC tissues). Furthermore, a recent paper indicated the coordinated activation of *SREBF1* and lipogenic genes in HCC

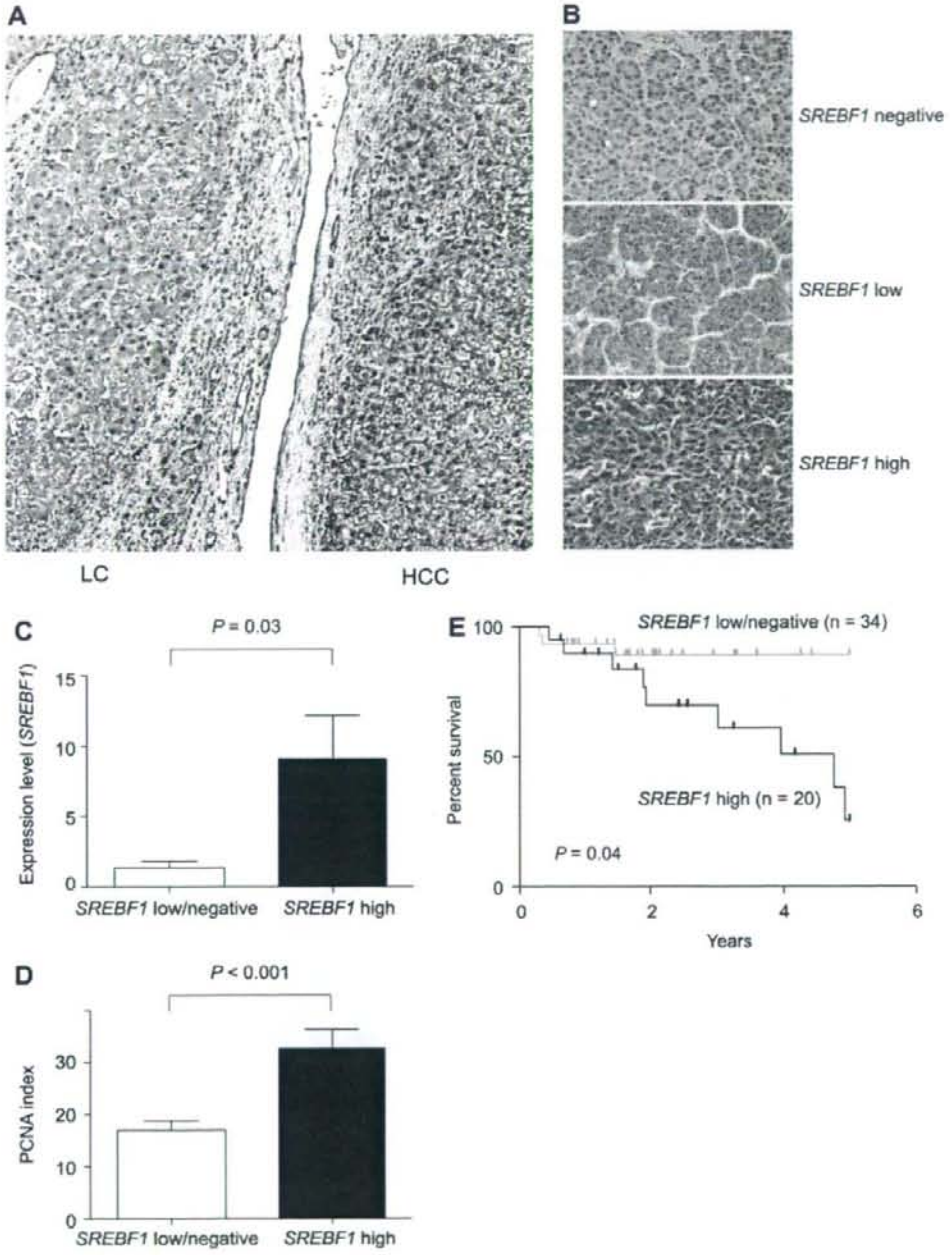

Fig. 4. (A) A photomicrograph of an HCC with adjacent non-cancerous cirrhotic liver stained with anti-*SREBF1* antibodies. (B) Representative photomicrographs of *SREBF1*-negative-, *SREBF1*-low-, and *SREBF1*-high-HCC tissues stained with anti- *SREBF1* antibodies. (C) *SREBF1* gene expression by real-time RT-PCR according to protein expression status assessed by IHC. *SREBF1* was highly expressed in *SREBF1*-high HCC ($P = 0.03$). (D) SREBF1 expression and cell proliferation in HCC. PCNA indexes in *SREBF1*-high HCC were higher than those in *SREBF1*-low/-negative HCC with statistical significance ($P < 0.001$). (E) Kaplan–Meier plots of 54 HCC patients analyzed by immunohistochemistry. The differences between *SREBF1*-high and -low/-negative HCC were analyzed by log-rank test.

developed in the liver of HCV core transgenic mice [42], strongly support our data. Although further studies using large numbers of HCC tissues may be required, these data suggest that the lipogenic gene activation seems to be mediated, at least in part, by *SREBF1* expression in HCC.

Table 2
Univariate Cox regression analysis of survival relative to *SREBF1* protein expression and clinicopathological parameters.

Variables (*n*)	HR (95% CI)	*P*-value
SREBF1 and mortality (n = 54)		
Tumor size		
<3 cm (*n* = 37)	1	
⩾3 cm (*n* = 17)	2.2 (0.6–8.3)	0.2
pTNM stage		
I, II (*n* = 45)	1	
III, IV (*n* = 9)	2.0 (0.4–9.4)	0.4
Serum AFP		
<20 ng/ml (*n* = 35)	1	
⩾20 ng/ml (n = 19)	1.5 (0.4–5.4)	0.5
SREBF1		
Low (*n* = 34)	1	
High (*n* = 20)	3.7 (1.0–13.7)	0.05

Because the majority of our HCC patients analyzed had Child–Pugh class A scores and about 70% had tumors less than 3 cm in diameter, all were expected to have a good prognosis. Indeed, patient survival in this cohort was not segregated by tumor size or pTNM stage (Table 2). Although the sample size was relatively small, we found that enhanced expression of *SREBF1* was a prognostic factor for mortality in HCC possibly due to the highly proliferative nature. Activation of lipogenesis pathways, as shown by overexpression of *FASN*, has been found to correlate with high mortality in breast, prostate, and lung cancer [43], suggesting that activation of lipogenesis may be a fundamental characteristic of cancer with poor prognosis. Thus, *SREBF1* expression may be a good biomarker for HCC classification, a finding that should be validated in a large scale cohort. Because deactivation of the lipogenesis pathway by inhibition of *SREBF1* gene expression could inhibit HCC cell growth *in vitro*, *SREBF1* may be a good target for pharmaceutical intervention in these tumors.

In conclusion, our genome-wide gene expression profiling analyses found that the lipogenesis pathway was activated in a subset of HCC. *SREBF1*, which activates the lipogenesis pathway, may be a good biomarker for HCC prognosis and may be a good target for therapeutic intervention.

Acknowledgements

We are grateful to the members of The Liver Disease Center at Kanazawa University Hospital for providing data of human liver tissue samples. We would also like to thank Dr. Hitoshi Shimano for providing invaluable reagents.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhep.2008.07.036.

References

[1] El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med 1999;340:745–750.
[2] Bosch FX, Ribes J, Diaz M, Cleries R. Primary liver cancer: worldwide incidence and trends. Gastroenterology 2004;127:S5–S16.
[3] Wang XW, Hussain SP, Huo TI, Wu CG, Forgues M, Hofseth LJ, et al. Molecular pathogenesis of human hepatocellular carcinoma. Toxicology 2002;181-182:43–47.
[4] Yamashita T, Kaneko S, Hashimoto S, Sato T, Nagai S, Toyoda N, et al. Serial analysis of gene expression in chronic hepatitis C and hepatocellular carcinoma. Biochem Biophys Res Commun 2001;282:647–654.
[5] Shirota Y, Kaneko S, Honda M, Kawai HF, Kobayashi K. Identification of differentially expressed genes in hepatocellular carcinoma with cDNA microarrays. Hepatology 2001;33:832–840.
[6] Okabe H, Satoh S, Kato T, Kitahara O, Yanagawa R, Yamaoka Y, et al. Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. Cancer Res 2001;61:2129–2137.
[7] Xu XR, Huang J, Xu ZG, Qian BZ, Zhu ZD, Yan Q, et al. Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver. Proc Natl Acad Sci USA 2001;98:15089–15094.
[8] Iizuka N, Oka M, Yamada-Okabe H, Mori N, Tamesa T, Okada T, et al. Comparison of gene expression profiles between hepatitis B virus- and hepatitis C virus-infected hepatocellular carcinoma by oligonucleotide microarray data on the basis of a supervised learning method. Cancer Res 2002;62:3939–3944.
[9] Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet 2002;31:339–346.
[10] Lee JS, Thorgeirsson SS. Genome-scale profiling of gene expression in hepatocellular carcinoma: classification, survival prediction, and identification of therapeutic targets. Gastroenterology 2004;127:S51–S55.
[11] Suriawinata A, Xu R. An update on the molecular genetics of hepatocellular carcinoma. Semin Liver Dis 2004;24:77–88.
[12] El-Serag HB, Tran T, Everhart JE. Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma. Gastroenterology 2004;126:460–468.
[13] Hassan MM, Hwang LY, Hatten CJ, Swaim M, Li D, Abbruzzese JL, et al. Risk factors for hepatocellular carcinoma: synergism of alcohol with viral hepatitis and diabetes mellitus. Hepatology 2002;36:1206–1213.
[14] Ohata K, Hamasaki K, Toriyama K, Matsumoto K, Saeki A, Yanagi K, et al. Hepatic steatosis is a risk factor for hepatocellular carcinoma in patients with chronic hepatitis C virus infection. Cancer 2003;97:3036–3043.
[15] Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of US adults. N Engl J Med 2003;348:1625–1638.
[16] Walsh MJ, Vanags DM, Clouston AD, Richardson MM, Purdie DM, Jonsson JR, et al. Steatosis and liver cell apoptosis in chronic hepatitis C: a mechanism for increased liver injury. Hepatology 2004;39:1230–1238.

[17] Powell EE, Jonsson JR, Clouston AD. Steatosis: co-factor in other liver diseases. Hepatology 2005;42:5–13.
[18] Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science 1995;270:484–487.
[19] Yamashita T, Hashimoto S, Kaneko S, Nagai S, Toyoda N, Suzuki T, et al. Comprehensive gene expression profile of a normal human liver. Biochem Biophys Res Commun 2000;269:110–116.
[20] Desmet VJ, Gerber M, Hoofnagle JH, Manns M, Scheuer PJ. Classification of chronic hepatitis: diagnosis, grading and staging. Hepatology 1994;19:1513–1520.
[21] Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B. A model for p53-induced apoptosis. Nature 1997;389:300–305.
[22] Yokoyama C, Wang X, Briggs MR, Admon A, Wu J, Hua X, et al. SREBP-1, a basic-helix–loop-helix–leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. Cell 1993;75:187–197.
[23] Wang HC, Chang WT, Chang WW, Wu HC, Huang W, Lei HY, et al. Hepatitis B virus pre-S2 mutant upregulates cyclin A expression and induces nodular proliferation of hepatocytes. Hepatology 2005;41:761–770.
[24] Takeba Y, Kumai T, Matsumoto N, Nakaya S, Tsuzuki Y, Yanagida Y, et al. Irinotecan activates p53 with its active metabolite, resulting in human hepatocellular carcinoma apoptosis. J Pharmacol Sci 2007;104:232–242.
[25] Closset J, Van de Stadt J, Delhaye M, El Nakadi I, Lambilliotte JP, Gelin M. Hepatocellular carcinoma: surgical treatment and prognostic variables in 56 patients. Hepatogastroenterology 1999;46:2914–2918.
[26] Arsura M, Cavin LG, Calvisi DF, Thorgeirsson SS, Eferl R, Ricci R, et al. Nuclear factor-kappaB and liver carcinogenesis. Cancer Lett 2005;229:157–169.
[27] Calvisi DF, Thorgeirsson SS. Molecular mechanisms of hepatocarcinogenesis in transgenic mouse models of liver cancer. Toxicol Pathol 2005;33:181–184.
[28] Eferl R, Ricci R, Kenner L, Zenz R, David JP, Rath M, et al. Liver tumor development. c-Jun antagonizes the proapoptotic activity of p53. Cell 2003;112:181–192.
[29] Xu L, Hui L, Wang S, Gong J, Jin Y, Wang Y, et al. Expression profiling suggested a regulatory role of liver-enriched transcription factors in human hepatocellular carcinoma. Cancer Res 2001;61:3176–3181.
[30] Horton JD, Goldstein JL, Brown MS. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. J Clin Invest 2002;109:1125–1131.
[31] Yahagi N, Shimano H, Hasegawa K, Ohashi K, Matsuzaka T, Najima Y, et al. Co-ordinate activation of lipogenic enzymes in hepatocellular carcinoma. Eur J Cancer 2005;41:1316–1322.
[32] Kawaguchi K, Honda M, Yamashita T, Shirota Y, Kaneko S. Differential gene alteration among hepatoma cell lines demonstrated by cDNA microarray-based comparative genomic hybridization. Biochem Biophys Res Commun 2005;329:370–380.
[33] Van de Sande T, De Schrijver E, Heyns W, Verhoeven G, Swinnen JV. Role of the phosphatidylinositol 3′-kinase/PTEN/Akt kinase pathway in the overexpression of fatty acid synthase in LNCaP prostate cancer cells. Cancer Res 2002;62:642–646.
[34] Yang YA, Han WF, Morin PJ, Chrest FJ, Pizer ES. Activation of fatty acid synthesis during neoplastic transformation: role of mitogen-activated protein kinase and phosphatidylinositol 3-kinase. Exp Cell Res 2002;279:80–90.
[35] You M, Fischer M, Deeg MA, Crabb DW. Ethanol induces fatty acid synthesis pathways by activation of sterol regulatory element-binding protein (SREBP). J Biol Chem 2002;277:29342–29347.
[36] Muller-Wieland D, Kotzka J. SREBP-1: gene regulatory key to syndrome X? Ann NY Acad Sci 2002;967:19–27.
[37] Sekiya M, Yahagi N, Matsuzaka T, Najima Y, Nakakuki M, Nagai R, et al. Polyunsaturated fatty acids ameliorate hepatic steatosis in obese mice by SREBP-1 suppression. Hepatology 2003;38:1529–1539.
[38] Marrero JA, Fontana RJ, Su GL, Conjeevaram HS, Emick DM, Lok AS. NAFLD may be a common underlying liver disease in patients with hepatocellular carcinoma in the United States. Hepatology 2002;36:1349–1354.
[39] Kim KH, Shin HJ, Kim K, Choi HM, Rhee SH, Moon HB, et al. Hepatitis B virus X protein induces hepatic steatosis via transcriptional activation of SREBP1 and PPARgamma. Gastroenterology 2007;132:1955–1967.
[40] Waris G, Felmlee DJ, Negro F, Siddiqui A. Hepatitis C virus induces proteolytic cleavage of sterol regulatory element binding proteins and stimulates their phosphorylation via oxidative stress. J Virol 2007;81:8122–8130.
[41] Furuta E, Pai SK, Zhan R, Bandyopadhyay S, Watabe M, Mo YY, et al. Fatty acid synthase gene is up-regulated by hypoxia via activation of Akt and sterol regulatory element binding protein-1. Cancer Res 2008;68:1003–1011.
[42] Tanaka N, Moriya K, Kiyosawa K, Koike K, Gonzalez FJ, Aoyama T. PPARalpha activation is essential for HCV core protein-induced hepatic steatosis and hepatocellular carcinoma in mice. J Clin Invest 2008;118:683–694.
[43] Kuhajda FP. Fatty acid synthase and cancer: new application of an old pathway. Cancer Res 2006;66:5977–5980.

Application of Serial Analysis of Gene Expression in Cancer Research

T. Yamashita, M. Honda, and S. Kaneko*

Department of Disease Control and Homeostasis, Graduate School of Medical Science, Kanazawa University, Kanazawa, Ishikawa, Japan

Abstract: It is now widely believed that tumors originate from normal cells as a result of accumulated genetic/epigenetic changes. These alterations affect the signaling pathways at transcriptional and post-transcriptional level that drive cells into uncontrolled cell division, growth, and migration. Recent advancement of molecular technologies have yielded comprehensive gene expression profiling techniques that have successfully provided candidate diagnostic and prognostic markers in human cancers. Serial Analysis of Gene Expression (SAGE) is a technology to facilitate the measurement of mRNA transcripts of normal and malignant tissues in a non-biased and highly accurate and quantitative manner. SAGE produces a comprehensive gene expression portrait without *a priori* gene sequence information, leading to the identification of novel transcripts potentially involved in the pathogenesis of human cancer. In this review, we provide a brief outline of SAGE to underscore the advantages of the method relative to the other gene expression profiling approaches in cancer research. We also summarize the progression of recent gene expression profiling studies and discuss the current topics of SAGE analysis in cancer research for the development of novel therapeutic interventions.

Keywords: SAGE, gene expression profiling, transcriptome, pathway analysis, biomarker, prognosis, metastasis, chemo sensitivity.

1. INTRODUCTION

Tumors originally develop from normal cells that gain the ability to grow aberrantly and metastasize to the distant organs. These malignant transformations are considered to be induced by the accumulation of multiple genetic/epigenetic changes [1, 2]. Although a classical genetic strategy such as gene knockout has been a powerful technique to explore genetic diseases with the gain or loss of a single gene, this strategy is less effective to obtain mechanistic insights of diseases affected by multiple genes such as cancer [3]. Complete sequences of the 3×10^9 base-pair human genome (as a result of the Human Genome Project (HGP)) have provided the basic framework of the genome-wide study in human diseases [4, 5]. Furthermore, genomic sequencing studies of various cancers have been reported or currently underway to reveal the genetic alterations on a genome-wide scale (The Cancer Genome Atlas; http://cancergenome.nih.gov/, The Cancer Genome Project; http://www.sanger.ac.uk/genetics/CGP/) [6, 7]. Comprehension of these framework studies will shed new lights on the genetic traits how cancer cells evolved from normal cells.

The central dogma is defined as the flow of genetic information from DNA to messenger RNA and then to protein, and the genetic/genomic alterations are considered to affect the cellular signaling pathway at transcriptional and post-transcriptional level. It is expected that the cataloging of tumor specific changes compared with normal counterparts will provide mechanistic insights into the process of carcinogenesis and potential tumor markers [8]. In the past, gene expression was investigated by the simple imaging limited to gene-by-gene approaches such as Northern blotting and reverse transcription-polymerase chain reaction (RT-PCR). Over the past decade, several methods have been developed to allow comparative studies of gene (and protein) expression between normal and cancer cells at genome-wide scale upon requests in the post-HGP era [3]. In particular, two main technologies that allow the genome-wide detection of differences in gene expression, microarray and serial analysis of gene expression (SAGE), have been used to analyze gene expression in cancer research [9]. These technologies have successfully provided the candidate gene sets used for the early detection or prognosis prediction of cancer [10-13]. Furthermore, these technologies have provided the candidate signaling pathways activated and therefore can be molecularly targeted for elimination of tumors [14-16]. In this review, we provide examples of SAGE application for discovering tumor markers and subclasses for stratification of tumors with distinct clinical outcomes. We also discuss the future direction on the gene expression profiling approaches to better understand the cancer biology.

2. SAGE AND OTHER GENE EXPRESSION PROFILING TECHNOLOGIES

In early 1990s, Craig Venter and his group pioneered the novel approach named expressed sequence tags (ESTs), which was the first transcriptomic approach for characterization of a tissue [17-19]. ESTs are randomly selected clones sequenced from cDNA libraries at large scale. Each cDNA library is constructed from poly-A RNA derived from a tis-

*Address correspondence to this author at the Department of Disease Control and Homeostasis, Graduate School of Medical Science, Kanazawa University, Kanazawa, Ishikawa, Japan; Tel: +81-76-265-2230; Fax: +81-76-265-2916; E-mail: skaneko@m-kanazawa.jp

1389-2010/08 $55.00+.00

sue, and thus the library represents genes expressed in the original population. An EST consists of approximately 300 base pairs of DNA as a single run of a sequence that is sufficiently long to establish the identity of the expressed gene. Importantly, ESTs have provided novel genes as well as new members of gene families that are expressed in a tissue. The analysis of EST data revealed fundamental characteristics of gene expression such as which genes are abundantly expressed and which genes are differentially expressed in each tissue. Development of high-throughput sequencing technologies has enabled the acceleration of ESTs collection since then, and comprehensive samples of human ESTs have been compiled in a public domain dbEST (http://www.ncbi.nlm.nih.gov/dbEST/). More than eight million human ESTs sequences are currently available.

Although ESTs collection have provided huge amounts of gene expression information in human tissues including various tumors, deposited sequence information was derived from various institutes/laboratories with different methods to obtain cDNA libraries and their sequence data. Strong bioinformatics supports were required to quantitatively compare the gene expression changes between tissues in a non-biased manner using public ESTs database. To overcome these limitations of cataloging ESTs, Victor Velculescu reported a novel elegant gene expression profiling approach to effectively collect short EST sequences in 1995, called SAGE [20]. The novelty of SAGE technique is based on two points: 1) collection of a short sequence called a SAGE tag from a defined mRNA position that is enough to identify the genes, and 2) concatenation of SAGE tags to increase the efficiencies of sequence-based short ESTs collection. Development of SAGE has enabled to analyze tens of thousands transcripts expressed in a tissue in a highly quantitative and qualitative manner by one laboratory, and the results can be easily comparable among any SAGE libraries reported by others. More than 400 SAGE libraries derived from human tissues are currently available in a public domain GEO (http://www.ncbi.nlm.nih.gov/geo/).

Around the same time, Patrick Brown and David Botstein's group as well as Affymetrix team independently pioneered the innovative technologies called cDNA microarrays and GeneChip arrays, respectively, to measure the expression of genes simultaneously in a tissue [21-23]. Although the methods to immobilize the probes on glass slides as well as to label dyes in samples are different from each other, they share the basic concept to enable broad-scale gene expression profiling based on the sample-probe hybridization [24, 25]. The major advantage of using array-based technologies compared to SAGE is the high-throughput performance to obtain gene expression data of large numbers of samples. On the other hand, the disadvantages of array-based gene expression profiling are difficulties to compare data in different platforms and to obtain absolute quantity of mRNA expressed in a tissue. Additionally, whereas array-based profiling technologies are closed format and can only provide expression data of known genes, SAGE is an open format and can be utilized to discover novel genes or novel splice variants expressed in a tissue. Thus, SAGE and array-based technologies have both advantages and disadvantages, and the appropriate method should be chosen to fit the objectives in each experiment. Besides, recent studies implicated the utility of combination of SAGE and DNA array technologies for target identification and validation studies to take advantages of both methods [26-31]. The outlines of brief SAGE and microarray protocol for cancer target identification are depicted in Fig. (**1**).

The gene signatures or biomarkers can be explored to discriminate or predict different clinical outcomes such as prognosis and chemotherapeutic sensitivities by pattern discovery and pattern prediction [9]. However, large-scale gene expression profiling data, especially these generated by SAGE, often provide false-positive candidates even after the rigorous statistical selection. Therefore, the results should be validated using different platforms such as real-time RT-PCR and/or tissue microarrays [32]. Furthermore, it is better to evaluate the significance of the identified gene signatures or biomarkers using an independent dataset, because the first data set used for the identification of biomarkers may not be representative of the diversity of patients for whom biomarkers are in broad clinical use [33].

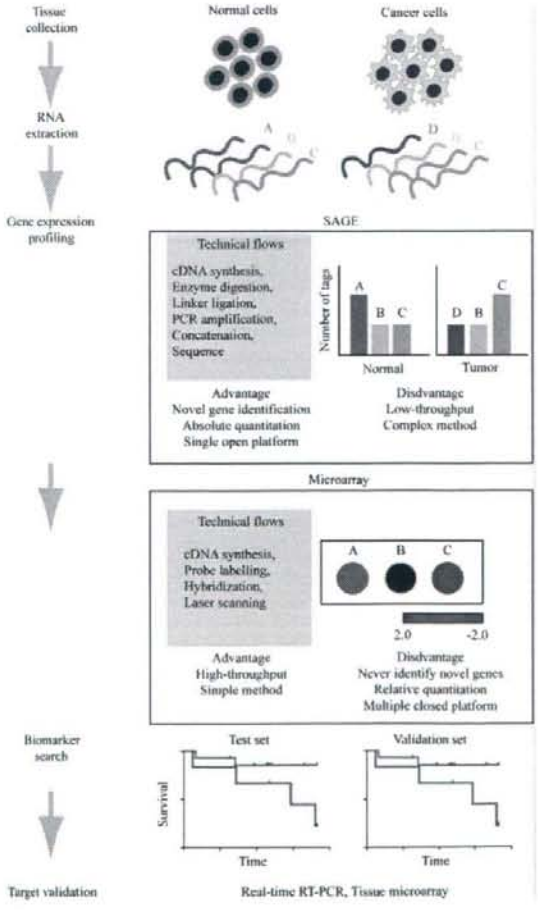

Fig. (1). The outlines of SAGE and microarray protocol for cancer target identification.