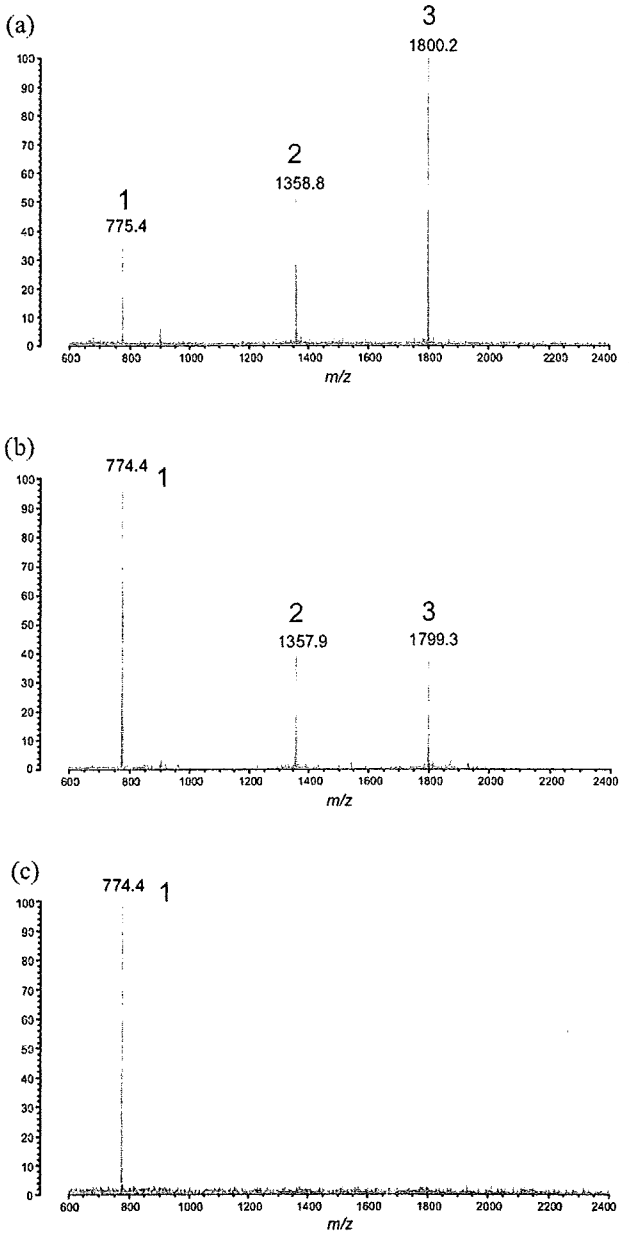

Figure 3. MALDI-TOF mass spectra of three-peptide mixture (a), after the transamination reaction (b) and after treatment with DITC glass (c). The three peptides are as follows: (1) VYIHPF ($[M+H]^+$: 775.4), (2) MHRQETVDCLK-NH_2 ($[M+H]^+$: 1358.7), and (3) AAKIQASFRGHMARKK ($[M+H]^+$: 1800.0). An aliquot (ca.10 pmol) of mixture was loaded on the target plate.

Modification of isolated C-terminal peptides

This method has the advantage that the isolated C-terminal peptide is open to introduction of virtually any type of functionality because it incorporates an α-ketocarbonyl group at its N-terminus that can accept nucleophilic attack.

As an example, we applied DNPH modification to the isolated C-terminal peptide of α_{S2}-casein (Fig. 5). The isolated C-terminal peptide was successfully modified with DNPH (Fig. 5(b)) and sequenced by MS (Fig. 5(c)).

DISCUSSION

It has been reported that some, but not all, N-terminal residues can be transaminated.[11,13,16] We examined the reactivity towards transamination of all 20 amino acid residues by using commercially available peptides. Sixteen of twenty residues (glycine, alanine, valine, leucine, isoleucine, phenylalanine, tyrosine, arginine, glutamic acid, glutamine, serine, threonine, asparagine, methionine, cysteine* (*: the SH group in the N-terminal cysteine residue was alkylated with iodoacetamide), and lysine) were transaminated, and peaks with a decrease of 1 *m/z* unit were detected (data not shown). For peptides having threonine, asparagine, cysteine*, or lysine residues at their N-termini, the signals of transaminated peptides were weak. Peptides incorporating histidine, tryptophan or proline residues at their N-termini afforded no transaminated product. However, such incompleteness of the reaction for peptides incorporating one of those seven residues (threonine, asparagine, cysteine*, lysine, histidine, tryptophan, or proline) at their N-termini does not affect the effectiveness of the method itself, because all peptides other than C-terminal peptides of proteins, which have lysine residues at their C-termini after LysC digestion, are depleted by DITC glass treatment whether or not their transamination reactions have proceeded to completion (data not shown). However, if a target C-terminal fragment of a protein has one of the three residues (threonine, asparagine or cysteine*: lysine residue is not placed at its N-terminus in this protocol) at its N-terminus, the signal in the mass spectrum can be detected but is weak after DITC depletion. For a C-terminal peptide having histidine, tryptophan or proline at its N-terminus, an orthogonal method should be employed.[5–7]

Optimized conditions or alternative routes to yield an α-ketocarbonyl group for peptides with these no-yield and low-yield residues should be exploited. For threonine and serine residues, there is an alternative technique that employs periodate oxidation.[22]

Dangers of side reactions have been mentioned in previous reports.[11,13,16,17] Peaks with a 56 and 74 Da increase were mainly detected among by-product peaks in the mass spectra. These by-products can generally form from almost all amino acid residues except for proline residues (data not shown). The mass increase of 56 Da may indicate the formation of a Schiff base with glyoxylate, and the increase of 74 Da can be explained by the addition of glyoxylate (before dehydration to the Schiff base). For the peptides containing aspartic acid at their N-terminus, the transamination reaction was incomplete although a 45-Da lighter species was detected as the principal species, which was attributed to the decarboxylated product.[13] This peak corresponds to that with an alanine residue at its N-terminus, so care should be taken of this special case.

The complexity of the mass spectrum that arose from the formation of these by-products was alleviated by DNPH modification. Figure 6 presents an example. A peptide incorporating an N-terminal serine residue (SFLLR) produced several by-products under this protocol (Fig. 6(a)). However, after DNPH modification, the modified peptide was selectively detected and the spectrum was simplified (Fig. 6(b)).

DOI: 10.1002/rcm

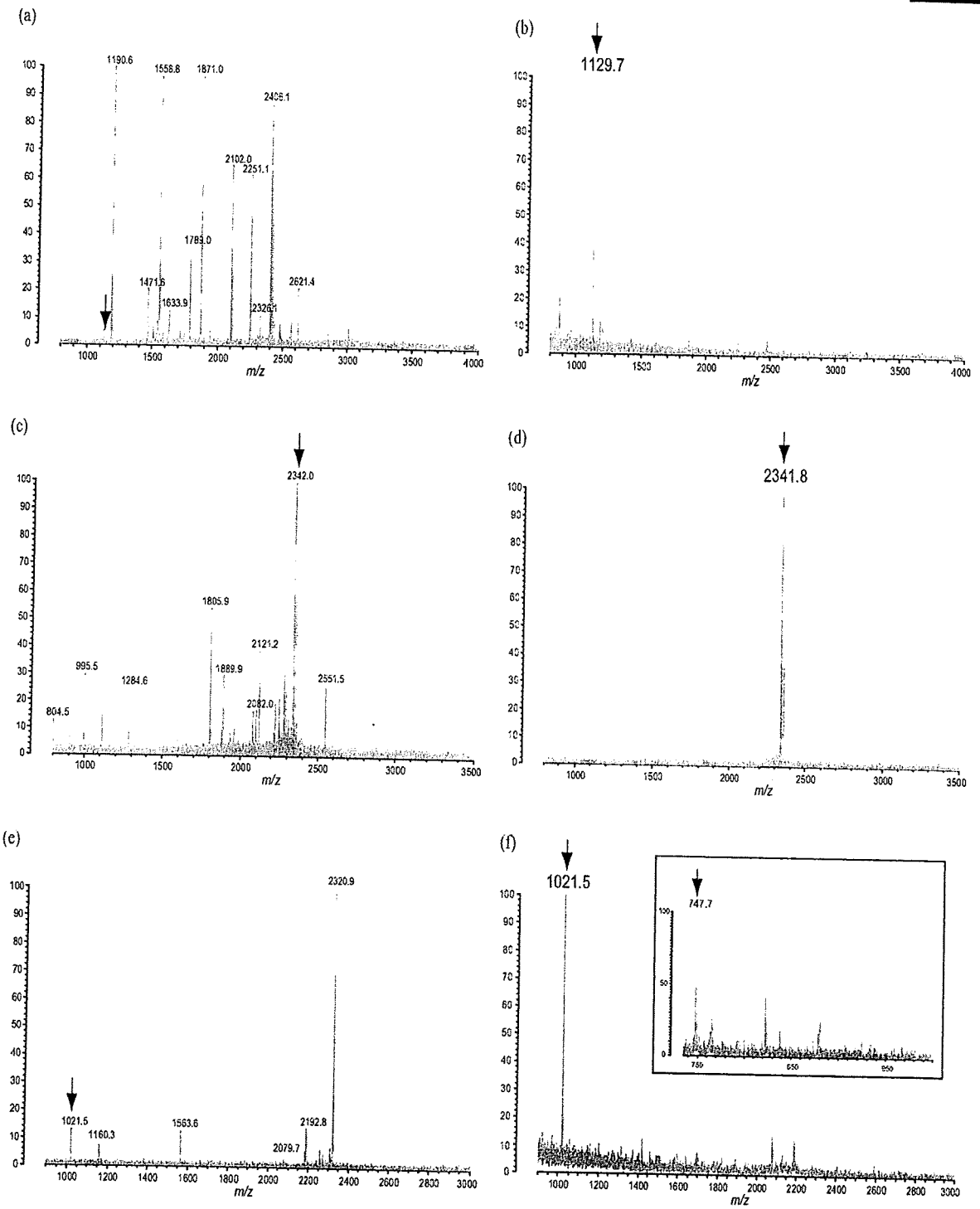

Figure 4. MALDI-TOF mass spectra of LysC digests after the transamination reaction (a, c, e) and after treatment with DITC glass (b, d, f). Three model proteins were used: pig albumin (a, b), bovine cathepsin B (c, d), and bovine α-casein (e, f). For bovine α-casein, two C-terminal peptides were isolated ($[M+H]^+$: 1021.06 from S2 and $[M+H]^+$: 747.4 from S1 (f: inset)). Arrows indicate the peaks of C-terminal peptides. An aliquot (ca. 3 pmol of digest) was loaded on the target plate.

It was reported that DNPH modification of peptides and use of DNPH as a matrix for MALDI-MS increase sensitivity.[20] With the experiment using the peptide (SFLLR, $[M+H]^+$: 635.4), the signal intensity of the modified peptide was enhanced. Successful application of the DNPH modification technique to transaminated peptides was thus demonstrated. This is partly explained by the higher affinity of DNPH-derivatized peptides for the DNPH matrix and structural similarity between derivatized peptides and the matrix.[20] It was reported that a nitrobenzene derivative functions as a matrix in MALDI-MS for selectively detecting 2-nitrobenzenesulfenyl-modified peptides and nitrotyrosine-containing peptides.[23]

DOI: 10.1002/rcm

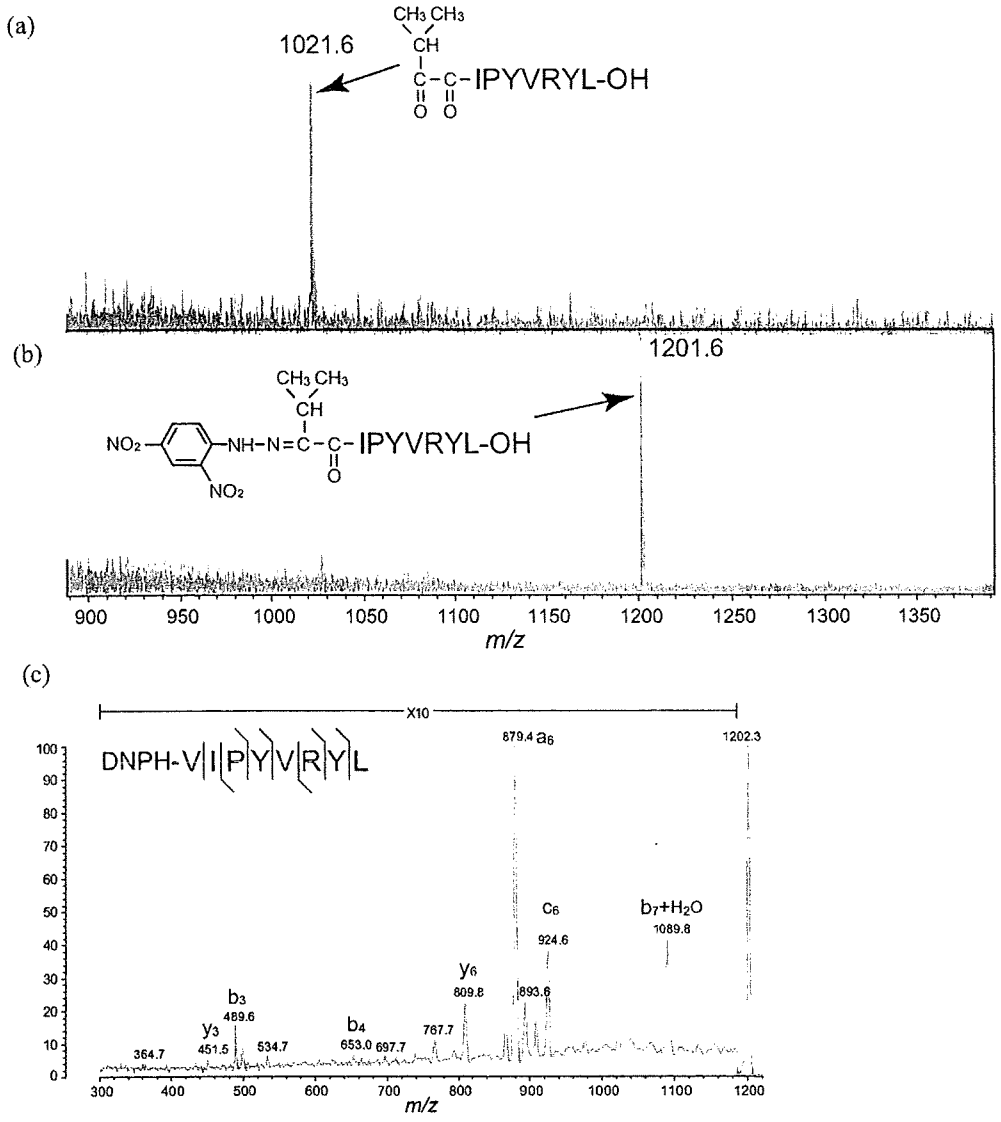

Figure 5. MALDI-TOF mass spectra of the isolated C-terminal peptide of α_{S2}-casein (a), after modification with DNPH reagent (b), and tandem mass spectrum (PSD mode) (c). An aliquot (ca. 3 pmol of digest) was loaded on the target plate. The b_7+H_2O ion arises from the loss of the C-terminal amino acid residue.[32]

The transamination reaction introduces a reactive carbonyl group into the peptide. It is beneficial for following chemical modifications because the resulting carbonyl group serves as a nucleophilic acceptor, whereas it increases the risk of undesired reactions, such as aldol reaction,[16,17] during subsequent steps. These undesired reactions can be minimized by the hydrazone or oxime formation just after the transamination reaction. This improved protocol is now under investigation.

In cases where proteins have a C-terminal lysine residue, the mass spectrum shows no peptide-derived signal because their C-terminal peptides, as well as internal peptides, are bound to DITC glass. This, in turn, suggests that the C-terminal residue can be lysine if no signal is observed in mass analysis. As mentioned already, the transamination reaction gives a couple of by-products which mainly derive from the addition reaction between the α-amino group of the peptide and the aldehyde group of glyoxylate. However, a C-terminal peptide whose N-terminus is proline gives no transaminated product and no adducts, which leaves the secondary amino group at the α-position intact (data not shown). Hence, in this case, the mass spectrum shows no peptide-derived signal as well. Therefore, if no signal is observed in the mass spectrum after the protocol, the C-terminal residue of a protein of interest can be lysine, or a C-terminal peptide whose N-terminus is proline.

For identification of proteins, N- or C-terminal analysis gives more reliable results than conventional methods such as peptide mass fingerprinting (PMF) or peptide fragment fingerprinting (PFF).[24,25] Hence, if the C-terminal analysis fails, N-terminal analysis can be complementarily used for this purpose.[8,26–28]

The N-terminus of a protein plays an important role as well.[4] As with C-terminal analysis, N-terminal structural determination serves for the analysis of post-translational modification and protein identification. We developed an N-terminal analysis method[8] by slightly modifying the method

DOI: 10.1002/rcm

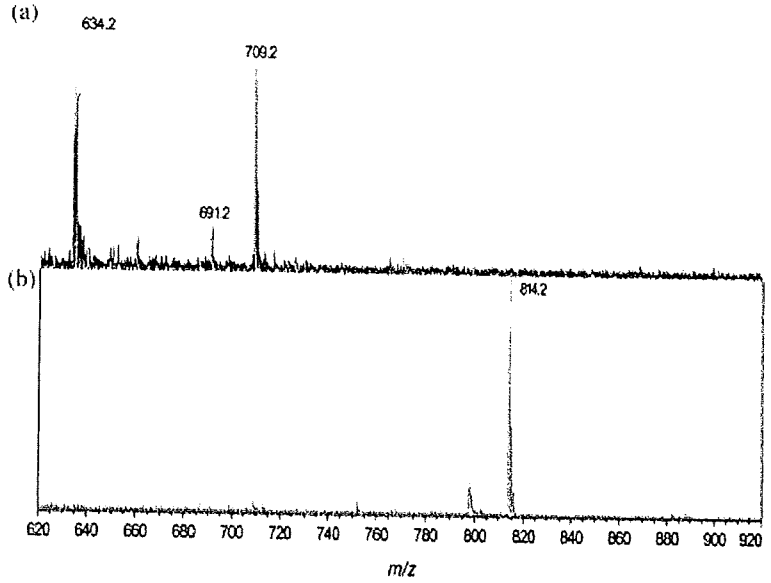

Figure 6. MALDI-TOF mass spectra of peptide SFLLR (10 pmol, $[M+H]^+$: 635.4) after the transamination reaction. The transaminated product (10 pmol, $[M+H]^+$: 634.4) was detected along with accompanying by-products (a), whereas the DNPH-modified product (10 pmol, $[M+H]^+$: 814.4) was mainly detected after the reaction (b).

for C-terminal analysis by just switching protease from LysC to *Grifola frondosa* metalloendopeptidase[29–31] (LysN). We are now studying N-terminal-specific isolation and sequencing by the transamination reaction (to be published elsewhere).

CONCLUSIONS

Isolation and sequencing of C-terminal peptides of proteins and their modification with DNPH were successfully performed. As far as we know, this is the first report of application of the transamination reaction to C-terminal analysis of protein by MS.

Owing to the α-diketone moiety, virtually any functionality can be introduced into the isolated C-terminal peptide. Therefore, C-terminal peptides isolated by this method can serve for a wide range of spectroscopic detection methodology.

REFERENCES

1. Chung JJ, Shikano S, Hanyu Y, Li M. *Trends Cell Biol.* 2002; **12**: 146.
2. Han X, Jin M, Breuker K, McLafferty FW. *Science* 2006; **314**: 109.
3. Suckau D, Resemann A. *Anal. Chem.* 2003; **75**: 5817.
4. Nakazawa T, Yamaguchi M, Okamura T, Ando E, Nishimura O, Tsunasawa S. *Proteomics* 2008; **8**: 673.
5. Samyn B, Sergeant K, Castanheira P, Faro C, Beeumen JV. *Nat. Methods* 2005; **2**: 193.
6. Panchaud A, Guillaume E, Affolter M, Robert F, Moreillon P, Kussmann M. *Rapid Commun. Mass Spectrom.* 2006; **20**: 1585.
7. Kuyama H, Shima K, Sonomura K, Yamaguchi M, Ando E, Nishimura O, Tsunasawa S. *Proteomics* 2008; **8**: 1539.
8. Kuyama H, Sonomura K, Nishimura O, Tsunasawa S. *Anal. Biochem.* 2008; **380**: 291.
9. Nishimura O, Suenaga M, Ohmae H, Tsuji S, Fujino M. *Chem. Commun.* 1998; 1135.
10. Dixon HBF. *Biochem. J.* 1964; **92**: 661.
11. Dixon HBF, Fields R. *Methods Enzymol.* 1972; **25**: 409.
12. Duan X, Zhao Z, Ye J, Ma H, Xia A, Yang G, Wang CC. *Angew. Chem. Int. Ed.* 2004; **43**: 4216.
13. Gilmore JM, Scheck RA, Esser-Kahn AP, Joshi NS, Francis MB. *Angew. Chem. Int. Ed.* 2006; **45**: 5307.
14. Scheck RA, Francis MB. *ACS Chem. Biol.* 2007; **2**: 247.
15. Sasaki T, Kodama K, Suzuki H, Fukuzawa S, Tachibana K. *Bioorg. Med. Chem. Lett.* 2008; **18**: 4550.
16. Papanikos A, Rademann J, Meldal M. *J. Am. Chem. Soc.* 2001; **123**: 2176.
17. Kawakami T, Hasegawa K, Teruya K, Akaji K, Horiuchi M, Inagaki F, Kurihara Y, Uesugi S, Aimoto S. *J. Peptide Sci.* 2001; **7**: 474.
18. Wachter E, Machleidt W, Hofner H, Otto J. *FEBS Lett.* 1973; **35**: 97.
19. Kuyama H, Sonomura K, Nishimura O. *Rapid Commun. Mass Spectrom.* 2008; **22**: 1109.
20. Fenaille F, Tabet JC, Guy PA. *Anal. Chem.* 2004; **76**: 867.
21. Fields R, Dixon HBF. *Biochem. J.* 1971; **121**: 587.
22. Dixon HBF, Weitkamp LR. *Biochem. J.* 1962; **84**: 462.
23. Matsuo E, Toda C, Watanabe M, Ojima N, Izumi S, Tanaka K, Tsunasawa S, Nishimura O. *Proteomics* 2006; **6**: 2042.
24. Wilkins MR, Gasteiger E, Tonella L, Ou K, Tyler M, Sanchez JC, Gooley AA, Walsh BJ, Bairoch A, Appel RD, Williams KL, Hochstrasser DF. *J. Mol. Biol.* 1998; **278**: 599.
25. Schandorff S, Olsen JV, Bunkenborg J, Blagoev B, Zhang Y, Andersen JS, Mann M. *Nat. Methods* 2007; **4**: 465.
26. Yamaguchi M, Obama T, Kuyama H, Nakayama D, Ando E, Okamura T, Ueyama N, Nakazawa T, Norioka S, Nishimura O, Tsunasawa S. *Rapid Commun. Mass Spectrom.* 2007; **21**: 3329.
27. Yamaguchi M, Nakayama D, Shima K, Kuyama H, Ando E, Okamura T, Ueyama N, Nakazawa T, Norioka S, Nishimura O, Tsunasawa S. *Rapid Commun. Mass Spectrom.* 2008; **22**: 3313.
28. Guillaume E, Panchaud A, Affolter M, Desvergnes V, Kussmann M. *Proteomics* 2006; **6**: 2338.
29. Nonaka T, Ishikawa H, Tsumuraya Y, Hashimoto Y, Dohmae N, Takio K. *J. Biochem.* 1995; **118**: 1014.
30. Nonaka T, Hashimoto Y, Takio K. *J. Biochem.* 1998; **124**: 157.
31. Hori T, Kumasaka T, Yamamoto M, Nonaka T, Tanaka N, Hashimoto Y, Ueki T, Takio K. *Acta Crystallogr.* 2001; **D57**: 361.
32. Gonzalez J, Besada V, Garay H, Reyes O, Padron G, Tambara Y, Takao T, Shimonishi Y. *J. Mass. Spectrom.* 1996; **31**: 150.

EUROPEAN JOURNAL OF CANCER 45 (2009) 881–889

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Molecular prediction of early recurrence after resection of hepatocellular carcinoma

Shinichi Yoshioka[a], *Ichiro Takemasa*[a,*], *Hiroaki Nagano*[a], *Nobuyoshi Kittaka*[a], *Takehiro Noda*[a], *Hiroshi Wada*[a], *Shogo Kobayashi*[a], *Shigeru Marubashi*[a], *Yutaka Takeda*[a], *Koji Umeshita*[a], *Keizo Dono*[a], *Kenichi Matsubara*[b], *Morito Monden*[a]

[a]*Department of Gastroenterological Surgery, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka E-2, Suita, 565-0871 Osaka, Japan*
[b]*DNA Chip Industry Co., Ltd., Tokyo, Japan*

ARTICLE INFO

Article history:
Received 16 October 2008
Received in revised form
11 December 2008
Accepted 15 December 2008
Available online 23 January 2009

Keywords:
Hepatocellular carcinoma
Gene expression profile
Intrahepatic recurrence

ABSTRACT

The prognosis of hepatocellular carcinoma (HCC) remains poor. Vascular invasion, tumour multiplicity and large tumour size are the conventional poor prognostic indicators related to early recurrence. However, it is difficult to predict prognosis of each HCC in the absence of these indicators. The purpose of this study is to predict early recurrence of HCC after radical resection based on whole human gene expression profiling. Microarray analyses were performed in 139 HCC primary tumours. A total of 88 cases lacking the conventional poor prognostic indicators were analysed to establish a molecular prediction system characteristic for early recurrence in 42 training cases with two polarised prognoses, and to test its predictive performance in 46 independent cases (group C). Subsequently, this system was applied to another 51 independent cases with some poor prognostic indicators (group D). The molecular prediction system accurately differentiated HCC cases into poor and good prognoses in both the independent group C (disease-free survival [DFS]: $p = 0.029$, overall survival [OS]: $p = 0.0043$) and independent group D (DFS: $p = 0.0011$, OS, $p = 0.035$). Multivariate Cox regression analysis indicated that the clinical value of molecular prediction system was an independent prognostic factor ($p < 0.0001$, hazard ratio = 3.29). Gene expression pattern related to early intrahepatic recurrence inherited in the primary HCC tumour can be useful for the prediction of prognosis.

1. Introduction

Hepatocellular carcinoma (HCC) is a common malignancy worldwide and is currently the third major cause of cancer-related deaths in Japan.[1] Recent progress in diagnostic and treatment technologies has improved the long-term survival of patients with HCC, but the prognosis remains unfavourable. Surgical resection has been one of the mainstays in curative treatment of HCC. However, even after curative resection, 80% of patients develop intrahepatic recurrence and 50% die within 5 years.[2,3]

Some patients who have undergone curative resection suffer an unpredictable early fulminant recurrence in the remnant liver, and this is associated with dismal prognosis. Detection of cases with early recurrence at the time of resection is beneficial for better decision making for treatment. In this regard, a staging system for HCC according to clinicopathological findings has been applied to assess the risk of recurrence following resection.[4]

Vascular invasion, tumour multiplicity and large tumour size (tumours measuring more than 5 cm in diameter) are poor prognostic indicators of HCC,[2,5–7] and it is difficult to

* *Corresponding author*: Tel.: +81 6 6879 3251; fax: +81 6 6879 3259.
E-mail address: alfa-t@sf6.so-net.ne.jp (I. Takemasa).
0959-8049/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2008.12.019

predict the prognosis of each case of HCC in the absence of these conventional indicators. However, the above-mentioned poor prognostic indicators are insufficient to predict the recurrence of HCC patients who undergo curative resection[21], thus new indicators are sought to help predict early intrahepatic recurrence developing after surgery in these patients.

Carcinogenesis is regulated by various changes on a genetic level, and several studies have discussed the phenomenon of cancer metastasis based on the analysis of various molecules. While it is useful to understand cancer progression, it is difficult to predict early recurrence with the analysis of a single molecule. The reason is that recurrence might be regulated by multiple molecular changes and interactions, and it might be difficult to explain the phenomenon of recurrence of HCC by a single molecule.[8–10] Therefore, it is important to conduct a comprehensive analysis of these molecules. The approach of microarray technology provides considerable information on cancer features and behaviour in individuals in several malignant tumours.[11–14] Several molecular and genetic studies have been reported on the progression of HCC and prediction of response of chemotherapy,[15–18] and some concluded that the specific gene expression patterns in HCC cancerous tissues could predict early intrahepatic recurrence.[19–22] However, it is still challenging to detect early recurrence tumours at the time of resection due to the complex pathogenesis of HCC. A recent study suggested that the strict selection of a homogeneous training set of patients in building the classifiers is essential to improve the predictability, reproducibility and validity of classifiers.[23]

In the present study, whole gene analysis was performed using a more clearly and strictly defined design set taking account of the complex pathogenic process of HCC, which reflected the prognosis more directly than previous reports with larger number of analyses.[24]

2. Materials and methods

2.1. Patients

A total of 139 HCC patients who had undergone hepatectomy at the Osaka University Hospital were enrolled in this study. All patients were followed up after resection for at least 3 months and the median follow-up time of survival cases in this study was 36 months (range, 12–87 months). Informed consent was obtained from all patients to use their surgical specimens and the clinicopathological data for research purposes. Histological classification was based on the Edmondson grading system and clinical stage was determined according to the Cancer of the Liver Italian Programme (CLIP). A mixture of RNA from the normal parts of liver specimens of seven patients with liver metastases from intestinal carcinomas was used as a reference for microarray analysis. None of the reference cases had hepatitis B or C (HBV or HCV, respectively) infection and their liver function tests were within normal values. All tissues were snap-frozen into liquid nitrogen and were stored at −80 °C.

2.2. Experimental design

Fig. 1 illustrates schematically our experimental design. Prediction of early recurrence in patients lacking the above-mentioned conventional poor prognostic indicators is clinically beneficial. In our study, we analysed patients lacking the aforementioned poor prognostic indicators to solve such a problem. To select the informative genes that are related to the phenomenon of early recurrence, we used two groups with polarised time course during the training phase. One group (group A) comprised cases with poor prognosis ($n = 21$), representing patients who developed multiple

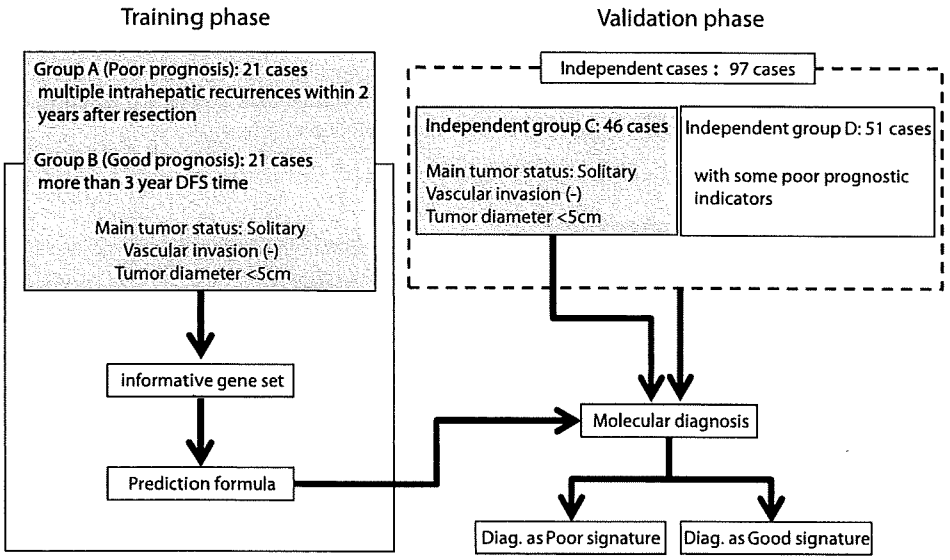

Fig. 1 – Schematic diagram of the experimental protocol. A molecular prediction system was constructed in the training phase. In the next step (validation phase), we applied this system to the independent group C ($n = 46$) and the entire group of independent cases ($n = 97$) comprising group C ($n = 46$) and group D ($n = 51$). Cases in grey coloured zones (Groups A, B and C) had similar clinicopathological conditions.

intrahepatic recurrences within 2 years after resection of the primary HCC. The second group (group B) comprised patients with satisfactory prognosis ($n = 21$), defined as more than 3-year disease-free survival (DFS) time. Table 1 summarises the clinicopathological features of patients of the two groups during the training phase. There were no differences between the two groups with regard to liver function tests and other clinicopathological variables except for the range of protein induced by vitamin K absence or antagonist II (PIVKA-II).

Based on the studies conducted in the training phase, a molecular prediction system was constructed using a set of informative genes. In the next step, we applied this system to another (independent) group C without any poor prognostic indicators as well ($n = 46$). The prediction system classified patients of group C into a 'poor signature' group (gene expression pattern resembled that of cases with poor prognosis) and a 'good signature' group (gene expression pattern resembled that of cases with good prognosis). Subsequently, we applied the prediction system to the independent group D ($n = 51$), which was composed of cases with positive status of some poor prognostic indicators. Finally, the independence of the diagnostic value of the molecular prediction results was verified by univariate and multivariate analyses using the whole independent cases, comprising patients of groups C and D.

2.3. *Microarray analysis*

Total RNA was extracted using TRIzol agent (Invitrogen, Carlsbad, CA), according to the instructions supplied by the manufacturer. Next, 2 μg of total RNA was used to synthesise double-strand cDNA that contained a promoter for T7 RNA polymerase. Amplified antisense RNA was synthesised by *in vitro* transcription of the cDNA templates by using the Amino Allyl MessageAmp aRNA kit (Ambion, Austin, TX). The reference and test sample were labelled with Cy3 and Cy5, mixed and hybridised on a microarray, AceGene Human oligo chip (DNA chip Research and Hitachi Software, Yokohama, Japan) DNA microarray. DNA microarray was used according to the instructions provided by the manufacturer (http://www.dna-chip.co.jp/thesis/AceGeneProtocol.pdf).

2.4. *Data analysis for postscanning*

The microarrays were scanned using ScanArray Lite and signal values were calculated using DNASIS array software (Hitachi Software Engineering Co., Yokohama, Japan). The local background was subtracted from each spot, and the ratio of the intensity of fluorescence from the Cy5 channel to the intensity of fluorescence from the Cy3 channel was calculated

Table 1 – Clinicopathological variables during the training phase.

	Poor prognosis group A ($n = 21$)	Good prognosis group B ($n = 21$)	P Value
Sex			
M	18	15	
F	3	6	0.452
Age, years			
<65	10	11	>0.999
⩾65	11	10	
HB infection			
+ve	8	11	0.535
–ve	13	10	
HC infection			
+ve	12	15	0.520
–ve	9	6	
Liver status			
Child A	19	15	0.239
Child B	2	6	
Tumour diameter, cm; mean (SD)	2.7 (0.8)	2.7 (1.1)	0.849•
AFP			
<400 ng/ml	18	18	>0.999
⩾400 ng/ml	3	3	
PIVKA-II			
<45 AU/ml	14	20	0.049
⩾45 AU/ml	7	1	
Capsule formation			
–ve	6	11	0.209
+ve	15	10	
Edmondson Grade			
I/II	13	16	0.504
III/IV	8	5	

P Values were calculated by the chi-square test, or by •Student *t*-test.

for each spot. Spots with intensity levels below the limit value were omitted. The ratio of expression level of each gene was converted to a logarithmic scale (base 2), and the data matrix was normalised to a median of 0 by standardising each sample.

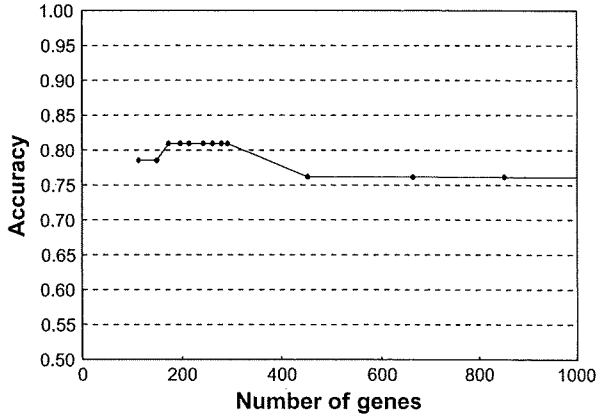

Fig. 2 – The accuracy curve based on weighted-voting algorithm with a leave-one-out cross validation. The accuracies in diagnosis of groups [ordinate] were plotted against the degree of *p*-value [abscissa]. The 172-gene set [P = 0.0004] marked the top accuracy. The accuracy was 80.2%. The P value was calculated by 10,000 times permutation test.

Genes with more than 15% missing data values in all samples in the training phase were excluded from the analysis. Missing data were compensated by averaging the expression data of 42 cases in the training phase.

2.5. *Construction of prediction system using gene expression patterns*

To detect the significant genes for prediction, we used permutation testing.[20] The original score of each gene (signal-to-noise ratio, $Si = (\mu A - \mu B)/(\sigma A + \sigma B)$, where μ and σ represent the mean and standard deviation of expression for each class, respectively) was calculated without permuting labels (responder or non-responder). The labels were randomly swapped and the values of S2N were calculated for the two groups. Repetition of this permutation 10,000 times provided a data matrix nearly the same as normal distribution. For each gene, the P value was calculated for the original S2N ratio with reference to the distribution of permuted data matrix. This model was evaluated by leave-one-out cross validation and the accuracy of each gene set was calculated based on the P value of the genes. As a supervised classification method, we adopted a weighted-voting (WV) algorithm.[13,14,19–22,25] We determined the optimal P value of the genes and classifier and constructed the prediction formula.

2.6. *Statistical analysis*

Clinicopathological indicators were compared using chi-square test and continuous variables were compared using

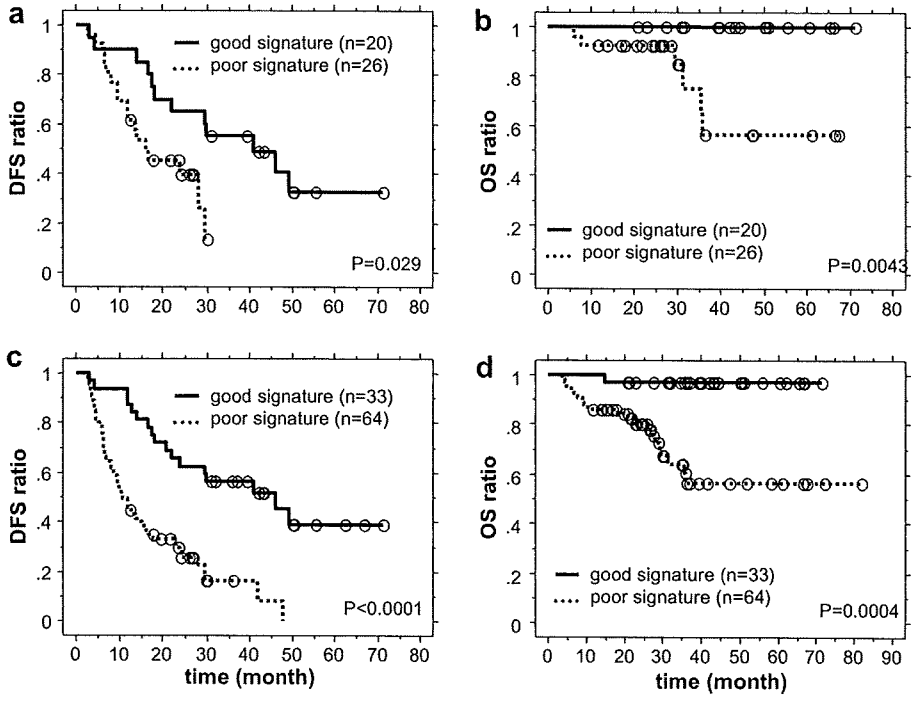

Fig. 3 – Disease-free survival curves and overall survival curves calculated using the Kaplan–Meier method for the independent cases. (a) DFS curves and (b) OS curves of the independent group C (*n* = 46). (c) DFS curves and (d) OS curves of the entire group of independent cases (*n* = 97) composed of groups C and D. Differences in survival curves were estimated by the log-rank test.

the Student t-test. Survival curves were computed using the Kaplan–Meier method, and differences between survival curves were compared using the log-rank test. To evaluate the risk associated with the prognostic variables, the Cox model with determination of the hazard ratio was applied; a 95% confidence interval was adopted. Statistical analyses were conducted using the SPSS software (version 11.0.1 J, SPSS Inc., Chicago, IL). We also performed network analysis using the Ingenuity Pathways Analysis (Ingenuity systems, Mountain View, CA; http://www.ingenuity.com), a web-based application.

Table 2 – Univariate analysis of the independent group C.

Parameter	Independent group C ($n = 46$)	P Value*	
		DFS	OS
Sex			
M	37	0.147	0.878
F	9		
Age, years			
<65	24	0.781	0.589
⩾65	22		
HB infection			
–ve	26	0.791	0.776
+ve	20		
HC infection			
–ve	18	0.467	0.980
+ve	28		
PIVKA-II			
<45 AU/ml	36	0.646	0.170
⩾45 AU/ml	10		
Capsule formation			
–ve	13	0.199	0.942
+ve	33		
Edmondson Grade			
I/II	27	0.479	0.479
III/IV	19		
CLIP score			
0–1	44	0.874	0.141
2–	2		
Liver status			
Child A	38	0.920	0.530
Child B	8		
AFP			
<400 ng/ml	38	0.724	0.374
⩾400 ng/ml	8		
Tumour diameter			
<5 cm	46	–	–
⩾5 cm	0		
Vascular invasion			
–ve	46	–	–
+ve	0		
Tumour multiplicity			
Single	46	–	–
Multiple	0		
Molecular-based diagnosis			
Poor prognosis	26	0.029	0.0043
Good prognosis	22		
Follow-up, months (median)	30 (4–81)		

* P Value was calculated by log-rank test according to the result of molecular diagnosis for DFS time.

Table 3 – Univariate analysis of the entire group of independent cases of independent groups C and D.

Parameter	Entire group ($n = 97$)	P Value*	
		DFS	OS
Sex			
M	80	0.604	0.582
F	17		
Age, years			
<65	42	0.892	0.850
⩾65	55		
HB infection			
–ve	51	0.584	0.416
+ve	46		
HC infection			
–ve	39	0.963	0.653
+ve	58		
PIVKA-II			
<45 AU/ml	73	0.897	0.387
⩾45 AU/ml	24		
Capsule formation			
–ve	19	0.730	0.748
+ve	78		
Edmondson grade			
I/II	49	0.015	0.169
III/IV	48		
CLIP score			
0–1	69	0.009	0.0024
2–	28		
Liver status			
Child A	81	0.229	0.032
Child B	16		
AFP			
<400 ng/ml	63	0.103	0.021
⩾400 ng/ml	34		
Tumour diameter			
<5 cm	72	0.062	0.021
⩾5 cm	25		
Vascular invasion			
–ve	82	0.187	0.0058
+ve	15		
Tumour multiplicity			
Single	65	0.0046	0.0033
Multiple	32		
Molecular-based diagnosis			
Poor prognosis	64	<0.0001	<0.0001
Good prognosis	33		
Follow-up, months (median)	30 (4–81)		

* P Value was calculated by log-rank test according to the result of molecular diagnosis for DFS time.

3. Results

3.1. *Differentially regulated genes during the training phase*

In the training phase, we examined the accuracy of prediction of HCC recurrence using full genes based on a WV algorithm with a leave-one-out cross validation approach. The accuracy of each gene set is shown in Fig. 2. The gene set of 0.0004% of P value using permutation test with 10,000 random trials marked the highest accuracy. We defined these differentially expressed 172 genes ($P = 0.0004\%$) as the informative gene set. Supplementary Table 1 provides a list of the informative genes. The results of molecular-based diagnosis system were correct in 34 of 42 cases. When compared with each annotated group, this system correctly classified 18 of 21 cases with poor prognosis and 16 of 21 cases with good prognosis in this set.

3.2. *Results of molecular diagnosis of the independent group C*

We adopted the prediction system constructed during the training phase to the independent group C, and compared DFS and overall survival (OS) of the patients between the two diagnosis groups (Fig. 3A). Both the DFS and OS ratios were significantly lower in patients diagnosed as 'poor signature'. The DFS curves showed significant difference between the two groups (log-rank test: $P = 0.029$) and all the seven patients who died of cancer were diagnosed as poor signature ($P = 0.0043$) (Fig. 3B). To compare other clinicopathological indicators with DFS and OS, we performed univariate analysis. Only molecular diagnosis was significantly different (Table 2).

3.3. *Results of molecular diagnosis of independent group D*

For cases of the independent group D, the DFS ratio and OS ratio were significantly lower in cases diagnosed as 'poor signature'. The log-rank test indicated that the DFS ratio ($P = 0.0011$) and OS ratio ($P = 0.035$) were significantly different between the 'poor signature' and 'good signature' groups (Figure not shown).

3.4. *Results of whole independent cases and evaluation of prediction ability of molecular diagnosis relative to other conventional poor prognostic indicators*

Our prediction system was further tested in the entire group of 97 cases (groups C and D). These cases were divided into 64 cases with poor signature and 33 cases with good signature based on the prediction system. Fig. 3C and D show the DFS and OS curves, respectively, for the two groups, according to the results of the prediction system. Kaplan–Meier survival estimates showed that DFS ratio was significantly lower in cases diagnosed as 'poor signature' than in patients diagnosed as 'good signature' ($P < 0.0001$). Twenty of 21 patients who died of cancer were of the 'poor signature' group and their OS curves were statistically different ($P = 0.0004$).

To compare our molecular prediction system with other conventional clinicopathological indicators, we performed univariate and multivariate analyses for DFS and OS. Univariate analysis of each factor for DFS time showed nearly significant differences with regard to Edmondson grade, AFP, tumour diameter, vascular invasion, tumour multiplicity and the result of molecular diagnosis (Table 3). To test the independence of the molecular prediction system, we performed multivariate Cox analysis. The result of the molecular prediction system was an independent factor ($P < 0.0001$), with a hazard ratio of 3.29 (95% CI 1.83–5.91) for the DFS ratio (Table 4). As for the OS ratio, the result of the molecular prediction system was also an independent factor ($P = 0.013$), with a hazard ratio of 13.28 (95% CI 1.72–102.63) (Table 4).

4. Discussion

The major finding of the present study was that early intrahepatic recurrence in patients who had undergone curative resection of HCC can be predicted accurately using our anal-

Table 4 – Results of multivariate analysis of the entire group of independent cases.

Variables	Hazard ratio	95% CI	P Value
Multivariate analysis of the entire group of independent cases (n = 97, DFS)			
Molecular diagnosis: poor signature (*versus* good signature)	3.29	1.83–5.91	<0.0001
Tumour multiplicity: multiple (*versus* single)	2.21	1.34–3.65	0.002
Edmondson grade: III/IV (*versus* I/II)	1.88	1.11–3.18	0.018
Tumour diameter: ⩾5 cm (*versus* <5 cm)	1.40	0.78–2.52	0.26
AFP: ⩾400 ng/ml (*versus* <400 ng/ml)	1.17	0.60–2.20	0.62
Vascular invasion: +ve (*versus* –ve)	0.93	0.46–1.87	0.84
Multivariate analysis of the entire group of independent cases (n = 97, OS)			
Molecular diagnosis: poor signature (*versus* good signature)	13.28	1.72–102.63	0.013
Tumour multiplicity: multiple (*versus* single)	3.06	1.16–8.06	0.024
Liver status: Child B (*versus* Child A)	2.38	0.90–6.29	0.08
Vascular invasion: +ve (*versus* –ve)	2.20	0.73–6.67	0.16
Tumour diameter: ⩾5 cm (*versus* <5 cm)	2.02	0.67–6.05	0.21
AFP: ⩾400 ng/ml (*versus* <400 ng/ml)	1.52	0.48–4.83	0.47
Edmondson grade: III/IV (*versus* I/II)	0.97	0.35–2.71	0.96

ysis system of gene expression patterns. Characteristic genes were selected by comparing the gene expression pattern between cases with multiple intrahepatic recurrences within 2 years and cases without recurrence over 3 years during the system training phase. The molecular prediction system accurately detected the high-risk group for early recurrence. Multivariate analysis identified molecular diagnosis, tumour multiplicity, and Edmondson grade as the independent factors. Taking into consideration that the majority of the patients who undergo curative resection become negative for the conventional poor prognostic indicators, molecular diagnosis could be potentially useful clinically for detecting patients at high-risk for early recurrence.

To improve the predictive accuracy, it is essential to clear the criteria of a homogeneous training set.[23] Our definition of the two groups was based on a study reported on the analysis of DFS ratio in HCC patients.[24] The DFS curve is composed of two regression lines. The majority of patients who developed recurrence within 2 years and who formed the first regression line were considered to have poor prognosis. On the other hand, the recurrence ratio of patients who showed no recurrence over a 3-year follow-up was almost the same as the annual relapse ratio of HCC in patients with hepatitis and their prognosis was better. This constant decrease in DFS ratio in the late recurrence cases is not usually observed in hepatectomised patients with liver metastasis from intestinal cancer.[26,27]

Recurrence of HCC is based on residual intrahepatic recurrence (IM) or multicentric metastasis (MC). IM is thought to originate from the primary cancer, while MC is considered to reflect a significant influence of the underlying liver status.[27,28] The two recurrence patterns are clinically important in patients with HCC where intrahepatic metastatic spread carries in general a poorer prognosis than that with multicentric nodules.[24] However, the conventional approach of histopathological examination is limited with regard to the differentiation of recurrence patterns as IM or MC.[29] With regard to the results of the validation phase, 17 patients survived for more than 3 years and only three of these 17 were diagnosed as poor signature and one of three cases was considered to have recurrence by metastasis from the primary tumour. On the other hand, 59 of 97 patients had intrahepatic recurrence within 2 years. This prediction system diagnosed these samples into 46 cases of poor signature and 13 cases of good signature. All 13 patients did not undergo a repeat resection, and thus pathological examination of recurrence pattern could not be conducted. However, as for the overall survival time in these 13 patients, only one died of cancer at 14 months postoperatively, while the remaining 12 patients remain alive for more than 21 months after surgery (range 21–48 months, median: 35 month). About half of the 13 patients had long survival though they had early recurrence. When we consider the relationship between study design and these results, the two groups diagnosed by our molecular-based diagnosis system may represent two recurrence patterns. The poor signature group may represent cases with recurrence due to IM, and the good signature group may represent cases with recurrence due to MC. This study may be clinically meaningful and helpful to solve the mechanism of recurrence patterns.

The prognoses of 42 patients during the training phase were polarised and those of the remaining 46 of the independent cases were intermediate. The 2-year survival ratio of the good signature independent cases was 65%, which was not as good as the annual relapse ratio of HCC. However, it is meaningful that the independent group C without any poor prognostic indicators could be divided into two groups of different prognoses. The reason for the discrepancy between the DFS ratio of cases diagnosed as good signature and annual relapse ratio is probably due to the fact that the independent group C did not include cases of extremely poor prognosis with early fulminant recurrence or cases of extremely good prognosis without long-term intrahepatic recurrence. Further analysis of cases with natural distribution of clinical status may help in moving the result of cases with good signature towards the annual relapse ratio.

In the conventional theory of metastasis, it is thought that tumours acquire the metastatic potential based on their progression and that metastasis occurs in the late phase. Based on this theory, recurrence could not be predicted by the analysis of the primary tumour. This theory was challenged recently by a new paradigm, which argues that the metastatic potential is not acquired in proportion to cancer progression but is already encoded in the primary tumour. Ramaswamy and colleagues[30] reported that a gene expression programme peculiar to metastasis may already be present in the bulk of some primary tumours and that a predictive diagnosis for metastasis was possible based on the analysis of the primary tumour profile. Several studies suggested that the molecular programme of primary tumour is generally retained in its metastasis.[31–33] Interestingly, Hoshida and colleagues[34] reported that the gene expression profiles in early-stage HCC tumours were highly associated with late recurrence (more than 2 years after resection) in the surrounding non-tumoural liver tissue but not in the tumoural tissue, indicating that environmental exposure leads to an increased potential of future malignant transformation. In this study, we evaluated the predictability of early recurrence using gene expression profiles of whole tumour tissue, based on the assumption that IM related to early recurrence might originate from the primary cancer.[27,28] For the entire group of independent cases, 78% of the recurrent cases within 2 years were diagnosed as poor signature. Some metastatic events may occur according to tumour progression, but cases with metastasis via the new paradigm should exist. Application of the theory of this paradigm may lead to the design of new diagnostic methods for cases in whom conventional clinicopathological parameters could not predict the prognosis.

Among the informative gene set, various genes correlate with cancer progression and carcinogenesis. PPARBP is regulated by RB18A and acts as a transcription cofactor by regulating the activity of p53wt transactivation on physiological promoters. Furthermore, downregulation of RB18A results in p53wt-dependent apoptosis.[35] RREB-1, a novel zinc finger protein, is involved in the differentiation response to Ras.[36] The Ras family is thought to be particularly important determinant of tumour initiation and progression.[37] BCL2 is one of the well-known tumour suppressor genes and is associated with recurrence and survival of HCC patients.[38,39] HDAC1 is reported to induce hyperacetylation of nucleosomal histones

in tumour cells, resulting in the expression of repressed genes that cause growth arrest, terminal differentiation and apoptosis.[40] The expression of HDAC1 is associated with prognosis of various carcinomas. BCOR is BCL6 co-receptor and is regulated by p53 and its characteristic expression is reported in various cancer cells.[41] This gene contributes to carcinogenesis in various malignancies such as B cell lymphoma and breast cancer.[42] BIRC5 is one of the major apoptosis regulators and is reported to be a prognostic marker of urothelial carcinomas and breast cancer.[43,44] These genes may serve as diagnostic markers for the development of HCC and help in the resolution of molecular mechanism of recurrence of HCC. To gain biological insights from these informative gene sets, we also used network analysis using Ingenuity Pathway Analysis. This analysis revealed that a few canonical signalling pathways, P38 MAPK signalling and PPARa/RXRa Activation signalling that are reported to be related to metastasis in human cancer,[45,46] harboured many of the upregulated informative genes (Supplementary Figure 1).

Our group has investigated the prediction of recurrence of various malignancies by gene expression profiling.[12,13,15,20,33,47] Kurokawa and colleagues[20] also reported a prediction model for HCC recurrence using a small-scale PCR-array system. They reported that early recurrence (within 2 year) in HCC patients could be predicted using 20-gene set after comparing cases with recurrence within 2 years and cases without recurrence over 2 years from 3072 primers.[20] In our study, cases with early intrahepatic recurrence within 2 years and reference cases without recurrence over 3 years during the training phase were defined based on DFS time of the characteristic recurrence patterns.[24] Furthermore, cases with common clinical background were analysed during the training phase using more strict criteria than previously reported using whole gene analysis, and accordingly, our results should be more reliable.

The report of Iizuka and colleagues[21] showed a correlation between gene expression, using a predictive system consisting of 12 genes, with early (within 1 year) post-hepatectomy intrahepatic recurrence, with a prediction accuracy of 89.3%. In their study, the DFS time of the reference group was more than one year and it is possible that characteristic recurrence patterns coexisted in the reference group. The study of Ho and colleagues[22] identified a molecular signature associated with vascular invasion (VI) in HCC and concluded that the signature could serve as a surrogate marker for predicting early recurrence after surgical resection.[22] Conventional prognostic indicators for early intrahepatic recurrence are not limited to vascular invasion only. A more direct approach should be considered for the prediction of early intrahepatic recurrence. While we did not analyse the reasons for the discrepancy in the prediction genes among the reported studies, we suspect that differences in clinical end-point may affect the results of the analysis and that accumulation of subtle differences in the dynamic range due to the platform of array might influence selection of the prediction genes.

In conclusion, the results of the present study showed that a characteristic gene expression pattern for early intrahepatic recurrence is encoded in primary HCC tumour and that gene profiling can be potentially helpful in predicting the prognosis of patients. Prediction of early recurrence of HCC may allow tailored treatment of individual patients and improvement of prognosis.

Conflicts of interest statement

None declared.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejca.2008.12.019.

REFERENCES

1. Yang Y, Nagano H, Ota H, Morimoto O, Nakamura M, Wada H, et al. Patterns and clinicopathologic features of extrahepatic recurrence of hepatocellular carcinoma after curative resection. *Surgery* 2007;**141**:196–202.
2. Tung-Ping Poon R, Fan ST, Wong J. Risk factors, prevention, and management of postoperative recurrence after resection of hepatocellular carcinoma. *Ann Surg* 2000;**232**:10–24.
3. Taura K, Ikai I, Hatano E, Fujii H, Uyama N, Shimahara Y. Implication of frequent local ablation therapy for intrahepatic recurrence in prolonged survival of patients with hepatocellular carcinoma undergoing hepatic resection: an analysis of 610 patients over 16 years old. *Ann Surg* 2006;**244**:265–73.
4. Park JH, Koh KC, Choi MS, Lee JH, Yoo BC, Paik SW, et al. Analysis of risk factors associated with early multinodular recurrences after hepatic resection for hepatocellular carcinoma. *Am J Surg* 2006;**192**:29–33.
5. Kosuge T, Makuuchi M, Takayama T, Yamamoto J, Shimada K, Yamasaki S. Long-term results after resection of hepatocellular carcinoma: experience of 480 cases. *Hepatogastroenterology* 1993;**40**:328–32.
6. Arii S, Tanaka J, Yamazoe Y, Minematsu S, Morino T, Fujita K, et al. Predictive factors for intrahepatic recurrence of hepatocellular carcinoma after partial hepatectomy. *Cancer* 1992;**69**:913–9.
7. Di Carlo V, Ferrari G, Castoldi R, Nadalin S, Marenghi C, Molteni B, et al. Surgical treatment and prognostic variables of hepatocellular carcinoma in 122 cirrhotics. *Hepatogastroenterology* 1995;**42**:222–9.
8. Armengol C, Boix L, Bachs O, Sole M, Fuster J, Sala M, et al. P27(Kip1) is an independent predictor of recurrence after surgical resection in patients with small hepatocellular carcinoma. *J Hepatol* 2003;**38**:591–7.
9. Kobayashi T, Sugawara Y, Shi YZ, Makuuchi M. Telomerase expression and p53 status in hepatocellular carcinoma. *Am J Gastroenterol* 2002;**97**:3166–71.
10. Kondo K, Chijiiwa K, Makino I, Kai M, Maehara N, Ohuchida J, et al. Risk factors for early death after liver resection in patients with solitary hepatocellular carcinoma. *J Hepatobiliary Pancreat Surg* 2005;**12**:399–404.
11. Arango D, Laiho P, Kokko A, Alhopuro P, Sammalkorpi H, Salovaara R, et al. Gene-expression profiling predicts recurrence in Dukes' C colorectal cancer. *Gastroenterology* 2005;**129**:874–84.
12. Takemasa I, Higuchi H, Yamamoto H, Sekimoto M, Tomita N, Nakamori S, et al. Construction of preferential cDNA microarray specialized for human colorectal carcinoma: molecular sketch of colorectal cancer. *Biochem Biophys Res Commun* 2001;**285**:1244–9.

13. Motoori M, Takemasa I, Yano M, Saito S, Miyata H, Takiguchi S, et al. Prediction of recurrence in advanced gastric cancer patients after curative resection by gene expression profiling. *Int J Cancer* 2005;**114**:963–8.
14. Hannemann J, Oosterkamp HM, Bosch CA, Velds A, Wessels LF, Loo C, et al. Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* 2005;**23**:3331–42.
15. Kurokawa Y, Matoba R, Nagano H, Sakon M, Takemasa I, Nakamori S, et al. Molecular prediction of response to 5-fluorouracil and interferon-alpha combination chemotherapy in advanced hepatocellular carcinoma. *Clin Cancer Res* 2004;**10**:6029–38.
16. 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**:2129–37.
17. Honda M, Kaneko S, Kawai H, Shirota Y, Kobayashi K. Differential gene expression between chronic hepatitis B and C hepatic lesion. *Gastroenterology* 2001;**120**:955–66.
18. Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J, et al. Gene expression patterns in human liver cancers. *Mol Biol Cell* 2002;**13**:1929–39.
19. Cheung ST, Chen X, Guan XY, Wong SY, Tai LS, Ng IO, et al. Identify metastasis-associated genes in hepatocellular carcinoma through clonality delineation for multinodular tumor. *Cancer Res* 2002;**62**:4711–21.
20. Kurokawa Y, Matoba R, Takemasa I, Nagano H, Dono K, Nakamori S, et al. Molecular-based prediction of early recurrence in hepatocellular carcinoma. *J Hepatol* 2004;**41**:284–91.
21. Iizuka N, Oka M, Yamada-Okabe H, Nishida M, Maeda Y, Mori N, et al. Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. *Lancet* 2003;**361**:923–9.
22. Ho MC, Lin JJ, Chen CN, Chen CC, Lee H, Yang CY, et al. A gene expression profile for vascular invasion can predict the recurrence after resection of hepatocellular carcinoma: a microarray approach. *Ann Surg Oncol* 2006;**13**:1474–84.
23. Dupuy A, Simon RM. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J Natl Cancer Inst* 2007;**99**:147–57.
24. Sakon M, Umeshita K, Nagano H, Eguchi H, Kishimoto S, Miyamoto A, et al. Clinical significance of hepatic resection in hepatocellular carcinoma: analysis by disease-free survival curves. *Arch Surg* 2000;**135**:1456–9.
25. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;**286**:531–7.
26. Ueno H, Mochizuki H, Hatsuse K, Hase K, Yamamoto T. Indicators for treatment strategies of colorectal liver metastases. *Ann Surg* 2000;**231**:59–66.
27. Kemeny N, Huang Y, Cohen AM, Shi W, Conti JA, Brennan MF, et al. Hepatic arterial infusion of chemotherapy after resection of hepatic metastases from colorectal cancer. *N Engl J Med* 1999;**341**:2039–48.
28. Morimoto O, Nagano H, Sakon M, Fujiwara Y, Yamada T, Nakagawa H, et al. Diagnosis of intrahepatic metastasis and multicentric carcinogenesis by microsatellite loss of heterozygosity in patients with multiple and recurrent hepatocellular carcinomas. *J Hepatol* 2003;**39**:215–21.
29. Kumada T, Nakano S, Takeda I, Sugiyama K, Osada T, Kiriyama S, et al. Patterns of recurrence after initial treatment in patients with small hepatocellular carcinoma. *Hepatology* 1997;**25**:87–92.
30. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. *Nat Genet* 2003;**33**:49–54.
31. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;**406**:747–52.
32. van't Veer LJ, Dai H, Van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;**415**:530–6.
33. Yamasaki M, Takemasa I, Komori T, Watanabe S, Sekimoto M, Doki Y, et al. The gene expression profile represents the molecular nature of liver metastasis in colorectal cancer. *Int J Oncol* 2007;**30**:129–38.
34. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med* 2008;**359**:1995–2004.
35. Frade R, Balbo M, Barel M. RB18A regulates p53-dependent apoptosis. *Oncogene* 2002;**21**:861–6.
36. Thiagalingam A, De Bustros A, Borges M, Jasti R, Compton D, Diamond L, et al. RREB-1, a novel zinc finger protein, is involved in the differentiation response to Ras in human medullary thyroid carcinomas. *Mol Cell Biol* 1996;**16**:5335–45.
37. Qin LX, Tang ZY. The prognostic molecular markers in hepatocellular carcinoma. *World J Gastroenterol* 2002;**8**:385–92.
38. Ikeguchi M, Hirooka Y, Kaibara N. Quantitative analysis of apoptosis-related gene expression in hepatocellular carcinoma. *Cancer* 2002;**95**:1938–45.
39. Farinati F, Marino D, De Giorgio M, Baldan A, Cantarini M, Cursaro C, et al. Diagnostic and prognostic role of alpha-fetoprotein in hepatocellular carcinoma: both or neither? *Am J Gastroenterol* 2006;**101**:524–32.
40. Pathil A, Armeanu S, Venturelli S, Mascagni P, Weiss TS, Gregor M, et al. HDAC inhibitor treatment of hepatoma cells induces both TRAIL-independent apoptosis and restoration of sensitivity to TRAIL. *Hepatology* 2006;**43**:425–34.
41. Huynh KD, Fischle W, Verdin E, Bardwell VJ. BCoR, a novel corepressor involved in BCL-6 repression. *Genes Dev* 2000;**14**:1810–23.
42. Phan RT, Dalla-Favera R. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature* 2004;**432**:635–9.
43. Yin W, Chen N, Zhang Y, Zeng H, Chen X, He Y, et al. Survivin nuclear labeling index: a superior biomarker in superficial urothelial carcinoma of human urinary bladder. *Mod Pathol* 2006;**19**:1487–97.
44. Span PN, Tjan-Heijnen VC, Heuvel JJ, de Kok JB, Foekens JA, Sweep FC. Do the survivin (BIRC5) splice variants modulate or add to the prognostic value of total survivin in breast cancer? *Clin Chem* 2006;**52**:1693–700.
45. Hsieh YH, Wu TT, Huang CY, Hsieh YS, Hwang JM, Liu JY. P38 mitogen-activated protein kinase pathway is involved in protein kinase Calpha-regulated invasion in human hepatocellular carcinoma cells. *Cancer Res* 2007;**67**:4320–7.
46. Knapp P, Jarzabek K, Blachnio A, Zbroch T. The role of peroxisome proliferator-activated receptors (PPAR) in carcinogenesis. *Ginekol Pol* 2006;**77**:643–51.
47. Motoori M, Takemasa I, Doki Y, Saito S, Miyata H, Takiguchi S, et al. Prediction of peritoneal metastasis in advanced gastric cancer by gene expression profiling of the primary site. *Eur J Cancer* 2006;**42**:1897–903.

British Journal of Cancer (2009) 100, 1647–1658

www.bjcancer.com

Activation of Wnt/β-catenin signalling pathway induces chemoresistance to interferon-α/5-fluorouracil combination therapy for hepatocellular carcinoma

T Noda[1], H Nagano*[,1], I Takemasa[1], S Yoshioka[1], M Murakami[1], H Wada[1], S Kobayashi[1], S Marubashi[1], Y Takeda[1], K Dono[1], K Umeshita[2], N Matsuura[3], K Matsubara[4], Y Doki[1], M Mori[1] and M Monden[1]

[1]Department of Surgery, Graduate School of Medicine and Health Science, Osaka University, Osaka, Japan; [2]Department of Health Science, Graduate School of Medicine and Health Science, Osaka University, Osaka, Japan; [3]Department of Molecular Pathology, Graduate School of Medicine and Health Science, Osaka University, Osaka, Japan; [4]DNA Chip Research Inc., Kanagawa, Japan

Type I IFN receptor type 2 (IFNAR2) expression correlates significantly with clinical response to interferon (IFN)-α/5-fluorouracil (5-FU) combination therapy for hepatocellular carcinoma (HCC). However, some IFNAR2-positive patients show no response to the therapy. This result suggests the possibility of other factors, which would be responsible for resistance to IFN-α/5-FU therapy. The aim of this study was to examine the mechanism of anti-proliferative effects of IFN-α/5-FU therapy and search for a biological marker of chemoresistance to such therapy. Gene expression profiling and molecular network analysis were used in the analysis of non-responders and responders with IFNAR2-positive HCC. The Wnt/β-catenin signalling pathway contributed to resistance to IFN-α/5-FU therapy. Immunohistochemical analysis showed positive epithelial cell adhesion molecule (Ep-CAM) expression, the target molecule of Wnt/β-catenin signalling, only in non-responders. *In vitro* studies showed that activation of Wnt/β-catenin signalling by glycogen synthesis kinase-3 inhibitor (6-bromoindirubin-3′-oxime (BIO)) induced chemoresistance to IFN-α/5-FU. BrdU-based cell proliferation ELISA and cell cycle analysis showed that concurrent addition of BIO and IFN-α/5-FU significantly to hepatoma cell cultures reduced the inhibitory effects of the latter two on DNA synthesis and accumulation of cells in the S-phase. The results indicate that activation of Wnt/β-catenin signalling pathway induces chemoresistance to IFN-α/5-FU therapy and suggest that Ep-CAM is a potentially useful marker for resistance to such therapy, especially in IFNAR2-positive cases.
British Journal of Cancer (2009) **100,** 1647–1658. doi:10.1038/sj.bjc.6605064 www.bjcancer.com
Published online 28 April 2009

Keywords: hepatocellular carcinoma; combination therapy; interferon-α; 5-fluorouracil; chemoresistance; Wnt signalling

Interferon (IFN) is a regulatory cytokine with various cellular activities, such as anti-proliferative, immunomodulatory and anti-angiogenic activities (Baron and Dianzani, 1994; Gutterman, 1994). Several studies emphasised the strong anti-tumour activity of IFN in hepatocellular carcinoma (HCC), when used in combination with other chemotherapeutic agents (Patt *et al*, 1993; Obi *et al*, 2006). We also reported the clinical efficacy of IFN-α/5-fluorouracil (5-FU) combination therapy for advanced HCC (Miyamoto *et al*, 2000; Sakon *et al*, 2002; Ota *et al*, 2005; Nagano *et al*, 2007a, b) and the mechanism of its anti-tumour effects (Eguchi *et al*, 2000; Yamamoto *et al*, 2004; Kondo *et al*, 2005; Nakamura *et al*, 2007; Wada *et al*, 2007). Further studies showed that the expression of IFN receptor type 2 (IFNAR2) in HCC tissue samples correlates significantly with clinical response to IFN-α/5-FU combination therapy (Ota *et al*, 2005; Nagano *et al*, 2007a). In an earlier study, we reported that 66% of those who responded to such treatment were IFNAR2-positive, but half of these 20 patients showed no clinical response (Nagano *et al*, 2007a). Therefore, finding novel biological markers of resistance to IFN-α/5-FU combination therapy is desirable, not only so that non-responders receive other potentially more successful treatments, but also to avoid their suffering caused by debilitating side effects.

Development of microarray technology has facilitated analysis of genome-wide expression profiles (Zembutsu *et al*, 2002; Yang *et al*, 2005). It can generate a large body of information concerning genetic networks related to pathological subtype, prognosis and resistance to anticancer drugs of neoplasm. We have reported many studies using PCR array or oligonucleotide microarray system in various human malignancies, particularly in gastrointestinal and HCCs (Komori *et al*, 2008; Kurokawa *et al*, 2004a, b; Motoori *et al*, 2005, 2006). To understand the complex biological processes, such as cancer progression and drug resistance, it is also important to consider differential gene expression by the gene network analysis (Kittaka *et al*, 2008). A detailed human interactive network that captures the entire cellular network would be invaluable in interpreting cancer signatures (Calvano *et al*, 2005; Rhodes and Chinnaiyan, 2005).

*Correspondence: Dr H Nagano, Department of Surgery, Graduate School of Medicine, Osaka University, 2-2, Yamadaoka E-2, Suita, Osaka 565-0871 Japan; E-mail: hnagano@gesurg.med.osaka-u.ac.jp
Received 5 January 2009; revised 9 March 2009; accepted 30 March 2009; published online 28 April 2009

In this study, we applied the methods of oligonucleotide microarray system and gene network analysis to identify informative gene set(s) and signalling pathway(s) related to resistance to IFN-α/5-FU combination therapy, especially in patients with IFNAR2-positive HCC. The results showed that Wnt/β-catenin signalling influenced resistance to IFN-α/5-FU combination therapy. The study also investigated the potential importance of epithelial cell adhesion molecule (Ep-CAM), which is encoded by the TACSTD1 gene and confirmed as one of the target genes of Wnt/β-catenin signalling (Yamashita *et al*, 2007), as a biological marker for resistance to IFN-α/5-FU combination therapy.

MATERIALS AND METHODS

Cell lines

The human HCC cell lines, PLC/PRF/5, HuH7, HLE, HLF and HepG2, were purchased from the Japanese Cancer Research Resources Bank (Tokyo, Japan). The Hep3B cell line was obtained from the Institute of Development, Aging and Cancer, Tohoku University (Sendai, Japan). They were maintained in Dulbecco's Modified Eagle Medium supplemented with 10% foetal bovine serum, 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin at 37°C in a humidified incubator with 5% CO_2 in air.

Drugs and reagents

Purified human IFN-α was kindly supplied by Otsuka Pharmaceutical Co. (Tokyo, Japan) and 5-FU was obtained from Kyowa Hakko Co. (Tokyo, Japan). The small molecule of 6-bromoindirubin-3′-oxime (BIO), a specific inhibitor of glycogen synthesis kinase-3 (GSK-3), activating the Wnt/β-catenin signalling pathway (Sato *et al*, 2004), was purchased from Calbiochem (San Diego, CA, USA) and was dissolved in dimethyl sulphoxide (DMSO) (Wako Pure Chemical Industries, Osaka, Japan). We used the following antibodies for immunohistochemistry and western blot analysis: monoclonal mouse anti-human Ep-CAM antibody (Abcam, Cambridge, UK), polyclonal rabbit anti-human c-MYC antibody (Cell Signaling Technology, Beverly, MA, USA) and polyclonal rabbit anti-human β-actin (Sigma, St Louis, MO, USA).

Patients and specimens

In total, 30 patients with multiple liver tumours spreading to both lobes with tumour thrombi in the major branches of the portal vein, underwent palliative reduction surgery at the Osaka University Hospital as described in our earlier report (Nagano *et al*, 2007a). All 30 patients had visible tumours in the remnant liver, and received combination chemotherapy with 5-FU and IFN-α as described earlier (Sakon *et al*, 2002; Ota *et al*, 2005). The chemotherapeutic response was evaluated clinically according to the criteria of the Eastern Cooperative Oncology Group (Oken *et al*, 1982). In this study, responders were defined as patients with complete response or partial response; non-responders were defined as patients with stable disease or progressive disease. All aspects of our study protocol were approved by the Human Ethics Committee of Graduate School of Medicine, Osaka University, Japan. Surgical specimens were obtained from these patients. Appropriate informed consent was obtained from all patients.

For microarray analysis, we used samples of 18 cases that were positive for IFNAR2 expression, whereas no samples were available from 2 cases with insufficient quality of RNA. For reference in microarray experiment, we obtained a mixture of RNA from normal parts of the liver specimens of seven patients with liver metastases from intestinal carcinomas who were free of HBV and HCV infections. All tissues were snap-frozen into liquid nitrogen and stored at −80°C. Other samples were fixed in 10% buffered formalin, embedded in paraffin and stained with haematoxylin–eosin to study the pathological features of HCC in accordance with the classification proposed by the Liver Cancer Study Group of Japan.

Microarray experiments

The microarray experiments were conducted according to the method described earlier (Kittaka *et al*, 2008). In brief, total RNA was purified by TRIzol agent (Invitrogen, San Diego, CA, USA), according to the instructions provided by the manufacturer. The integrity of RNA was assessed by Agilent 2100 Bioanalyzer and RNA 6000 LabChip kits (Yokokawa Analytical Systems, Tokyo, Japan). Only high-quality RNA was used for analysis. For control reference, RNAs from normal liver tissues were mixed. The reference and sample were mixed and hybridised on a microarray covering 30 336 human probes (AceGene Human 30K; DNA Chip Research Inc. and Hitachi Software Engineering Co., Kanagawa, Japan). The ratio of expression level of each gene was converted to a logarithmic scale (base 2) and the data matrix was normalised. Genes with >10% missing data values in all samples were excluded from analysis; a total of 14 473 genes out of 30 336 were available for analysis.

To detect the significant genes for resistance, we used permutation testing (Kurokawa *et al*, 2004b). The original score of each gene (signal-to-noise ratio (S2N), $Si = (\mu A - \mu B)/(\sigma A + \sigma B)$, where μ and σ represent the mean and standard deviation of expression for each class, was calculated without permuting labels (responder or non-responder). The labels were randomly swapped and the values of S2N were calculated between two groups. Repetition of this permutation 10 000 times provided data matrix that was nearly the same as normal distribution. For each gene, the *P*-value was calculated from the original S2N ratio with reference to the distribution of permuted data matrix. We determined the optimal *P*-value and the informative gene set.

Pathway analysis

We further analysed the significant genes for resistance by the Ingenuity Pathways Analysis (Ingenuity Systems, Mountain View, CA, USA; http://www.ingenuity.com). The Ingenuity Pathway Knowledge Base (IPKB) is a database of earlier published findings on mammalian biology. Canonical pathways analysis identifies the pathways that were statistically significant from the submitted data matrix from the canonical pathways of IPKB. The *P*-value of each canonical pathway is calculated using Fischer's exact test determining the probability that the association between the genes in the data set and the canonical pathway is because of chance alone.

Network analysis was conducted as described earlier (Calvano *et al*, 2005). In brief, the submitted genes that were mapped to the corresponding gene objects in the IPKB were called 'focus genes'. The focus genes were used to generate biological networks. The Ingenuity software queries the IPKB for interactions between focus genes and then generates a set of networks. The *P*-value of each network is calculated according to the fit of the user's set of significant genes. The score of a network is displayed as a negative log of the *P*-value, indicating the probability that a collection of genes equal to or greater than the number in a network could be achieved by chance alone.

RT–PCR analysis

Complementary DNA was generated from 1 μg RNA with avian myeloblastosis virus reverse transcriptase (Promega, Madison, WI, USA) as described earlier (Damdinsuren *et al*, 2006). Quantitative real-time PCR (qRT–PCR) assays were carried out using the Light Cycler (Roche Diagnostics, Mannheim, Germany), as described

earlier (Ogawa *et al*, 2004). Gene expression was measured in duplicate. The conditions set for qRT-PCR for TACSTD1, TCF3, AXIN2, MYC, CCND1 and β-actin were one cycle of denaturing at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, 60°C for 15 s and 72°C for 35 s, and final extension at 72°C for 10 min (or annealing at 58°C for β-actin). The primer sequences were as follows: TACSTD1 forward primer, 5′-TCCAGAAAGAAGAGA ATGGCA-3′; TACSTD1 reverse primer, 5′-AAAGATGTCTTCGT CCCACG-3′; TCF3 forward primer, 5′-ATCTGTGTCCCATGTCCC AG-3′; TCF3 reverse primer, 5′-CCAGGGTAGGAGACTTGCAG-3′; AXIN2 forward primer, 5′-GGTGTTTGAGGAGATCTGGG-3′; AXIN2 reverse primer, 5′-TGCTCACAGCCAAGACAGTT-3′; MYC forward primer, 5′-AAGAGGACTTGTTGCGGAAA-3′; MYC reverse primer, 5′-CTCAGCCAAGGTTGTGAGGT-3′; CCND1 forward primer, 5′-AAGGCCTGAACCTGAGGAG-3′; CCND1 reverse primer, 5′-CTTGACTCCAGCAGGGCTT-3′; β-actin forward primer, 5′-GA AAATCTGGCACCACACCTT-3′; and β-actin reverse primer, 5′-G TTGAAGGTAGTTTCGTGGAT-3′.

Immunohistochemical staining

For immunohistochemical staining of Ep-CAM expression, we used the method described earlier (Kondo *et al*, 1999) with minor modifications. Briefly, formalin-fixed, paraffin-embedded 4-μm thick sections were deparaffinised, then treated with an antigen retrieval procedure and incubated in methanol containing 0.3% hydrogen peroxide to block endogenous peroxidase. The sections were incubated with normal protein block serum solution, and the biotin-blocking solution (Wako) was used as recommended by the manufacturer. Then, the sections were incubated overnight at 4°C with anti-Ep-CAM antibody as the primary antibody. After washing in phosphate-buffered saline (PBS), the sections were incubated with a biotin-conjugated secondary antibody (horse anti-mouse antibody for Ep-CAM) and with peroxidase-conjugated streptavidin. The peroxidase reaction was then developed with 0.02% 3, 30-diaminobenzidine tetrachloride (Wako) solution with 0.03% hydrogen peroxide. Finally, the sections were counterstained with Meyer's haematoxylin. For negative controls, sections were treated the same way except that they were incubated with Tris-buffered saline instead of the primary antibody.

Ep-CAM expression was assessed by two investigators (TN and NM) independently without knowledge of the corresponding clinicopathological data. Antigen expression was defined as the presence of specific staining on the surface membrane of tumour cells. Ep-CAM expression was evaluated by calculating the total immunostaining score, representing the product of the proportion score and the intensity score, as described earlier (Gastl *et al*, 2000). In brief, the proportion score described the estimated fraction of positively stained tumour cells (0, none; 1, $<10\%$; 2, 10–50%; 3, 50–80% and 4, $\geqslant 80\%$). The intensity of Ep-CAM expression was compared with staining of normal bile duct epithelium present in each section of positive control. The intensity score represented the estimated staining intensity (0, no staining; 1, weak; 2, moderate and 3, strong). The total score ranged from 0 to 12. Ep-CAM-positive cases represented those with a total score >4.

Western blot analysis

The cells were washed with PBS and collected with a rubber scraper. After centrifugation, the cell pellets were resuspended in RIPA buffer (25 mM Tris (pH 7.5), 50 mM NaCl, 0.5% sodium deoxycholate, 2% Nonidet P-40, 0.2% sodium dodecyl sulphate, 1 mM phenylmethylsulphonyl fluoride and 500 KIE ml^{-1} 'Trasylol' proteinase inhibitor (Bayer LeverKusen, Germany)) with phosphatase inhibitor (Sigma). The extracts were centrifuged and the supernatant fraction was collected. Western blot analysis was carried out as described earlier (Kondo *et al*, 2005).

Luciferase reporter assay

The reporter assay kit was purchased from SA Biosciences (Frederick, MD, USA) to evaluate TCF/LEF transcriptional activity and is used according to the instructions provided by the manufacturer. In brief, 2×10^4 cells per well were added in triplicate to a 96-well microplate, and 24 h later, cells were transiently transfected with the transcription factor-responsive reporter or negative control by the Lipofectamine 2000 reagent (Invitrogen). Culture media were changed 16 h after transfection, and the transfected cells were treated with various concentrations of BIO (0–5 nM). After 24 h treatment, luciferase activity was measured with the Dual-Luciferase Assay System (Promega) using microplate luminometer, microlumat LB96P (Berthold Technologies, Calmbacher, Germany). The Firefly luciferase activity, indicating TCF-dependent transcription, was normalised to the *Renilla* luciferase activity as an internal control to obtain the relative luciferase activity.

Growth-inhibitory assays with 5-FU and IFN-α

The growth inhibitory assay was assessed by the 3-(4-, 5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) (Sigma) assay as described earlier (Eguchi *et al*, 2000). The tested concentrations of 5-FU were 0.05, 0.5 and 5 μg ml^{-1}, and those of IFN-α were 50, 500 and 5000 U ml^{-1}. The cells were incubated in a medium containing variable concentrations of 5-FU and IFN-α with DMSO or 5 nM BIO for 48 h. The proportion of cells incubated without drugs was defined as 100% viability.

DNA synthesis-inhibition assay

DNA synthesis inhibition was assessed by bromodeoxyuridine (BrdU) incorporation rate using the Cell Proliferation enzyme-linked immunosorbent assay (ELISA)-Chemiluminescent kit (Roche Applied Science, Indianapolis, IN, USA) according to the protocol provided by the manufacturer. In brief, HuH7 cells (1×10^4 per well) were seeded in triplicate into 96-well microplate. After treatment with control, 5-FU alone (5 μg ml^{-1}), IFN-α alone (5000 U ml^{-1}) and combination of 5-FU and IFN-α, with or without BIO (5 nM), the plates were incubated at 37°C under 5% CO_2 for 24 h. Then cells were labelled for 2 h with BrdU. Chemiluminescent signals were detected on the microplate luminometer (microlumat LB96P, Berthold Technologies).

Cell cycle analysis

Flow cytometric analysis was carried out as described earlier (Eguchi *et al*, 2000). In brief, cells were washed with PBS and then fixed in 70% cold ethanol. Propidium iodide (Sigma) and RNase (Sigma) were added for 30 min at 37°C. Samples were filtered, and data were acquired with a FACSort (Becton Dickinson Immunocytometry Systems, San Jose, CA, USA). Analysis of the cell cycle was carried out using ModFIT software (Becton Dickinson Immunocytometry Systems).

Statistical analysis

Clinicopathological indicators were compared using χ^2-test, whereas continuous variables were compared using the Student's *t*-test. Survival curves were computed using the Kaplan–Meier method, and differences between survival curves were compared using the log-rank test. To evaluate the risk associated with the prognostic variables, the Cox model with determination of the hazard ratio was applied; a 95% confidence interval was adopted. All statistical analyses were calculated using the SPSS software

(version 11.0.1 J, SPSS Inc., Chicago, IL, USA), and P-value <0.05 was considered statistically significant.

RESULTS

Patients' characteristics

The characteristics of the 30 HCC patients are shown in Table 1. A total of 10 patients were IFNAR2-negative and 20 patients were IFNAR2-positive. In 20 cases with positive IFNAR2, 10 patients were classified as responders; the remaining 10 patients were classified as non-responders. All patients with negative IFNAR2 were non-responders. We have earlier reported that IFNAR2 expression correlated significantly with the response to IFN-α/5-FU therapy (Ota *et al*, 2005; Nagano *et al*, 2007a). A larger proportion of responders were infected with HCV than non-responders. But all other analysed parameters were comparable among these groups and there were no significant differences in these parameters.

Table 1 Clinicopathological characteristics of responders and non-responders

	IFNAR2-positive		IFNAR2-negative	
	Responders ($n=10$)	Non-responders ($n=10$)	Non-responders ($n=10$)	*P*-value
Age (year)				NS
<60	6	7	5	
⩾60	4	3	5	
Sex				NS
Male	9	9	9	
Female	1	1	1	
HBV infection				NS
Present	6	8	7	
Absent	4	2	3	
HCV infection				0.0180
Present	7	1	3	
Absent	3	9	7	
Child-pugh score				NS
A	7	7	5	
B, C	3	3	5	
Liver cirrhosis				NS
Present	4	7	3	
Absent	6	3	7	
α-fetoprotein (ng ml^{-1})				NS
<300	5	1	3	
⩾300	5	9	7	
Tumour size (cm)				NS
<5	2	3	2	
⩾5	8	7	8	
Histological grade				NS
Moderately differentiated	1	0	0	
Poorly differentiated	9	8	9	
Undifferentiated	0	2	1	
IFNAR2 expression				<0.0001
0	0	0	10	
1	8	10	0	
2	2	0	0	

HBV = hepatitis B virus; HCV = hepatitis C virus; IFNAR2 = type I interferon receptor 2.

Microarray analysis and pathway analysis

Genes with significant P-values ($P<0.001$) were defined by the random permutation test. These differentially expressed 161 genes were selected as informative gene set and are listed in Table 2. The status of gene expression was defined as expression in non-responders compared with responders. Of the total, 98 genes were relatively upregulated in the responder group and 63 genes were relatively downregulated.

Then we carried out the canonical pathway analysis of the 161 genes using the software Ingenuity. Eight canonical pathways were identified as pathways that significantly influenced the resistance of IFN-α/5-FU combination therapy in 161 informative genes (Table 3). We also simultaneously carried out network analysis of the same informative genes set. A total of 14 networks were identified, and these networks were ranked by the score on a P-value calculation, which ranged from 2 to 55. Then, we selected one network with the highest score. The network with the highest score consisted of 35 molecules in 25 focus molecules and 11 interconnecting molecules (Figure 1A). This network included AXIN2, TCF3, RARA, CREBBP and TACSTD1, which were all associated with Wnt/β-catenin signalling identified by the canonical pathway analysis. In recent reports, Wnt/β-catenin signalling has been shown to mediate radiation resistance and chemotherapy resistance of various malignancies. In the Wnt/β-catenin signalling-related genes, TACSTD1 was most highly upregulated in the non-responders at the level of transcription.

TACSTD1 expression by RT-PCR and correlation with microarray data

Next, we examined the correlation between the expression data of gene expression and qRT-PCR of TACSTD1 to verify the microarray expression data. qRT-PCR analysis was carried out on 13 HCC tissue samples with positive expression of IFNAR2. Individual mRNA levels were normalised to β-actin and expressed relative to those in a mixture of seven normal livers. In the 13 IFNAR2-positive samples, TACSTD1 expression correlated significantly with the microarray data (Figure 1B). The Pearson correlation coefficient (P-value) for TACSTD1 was 0.668 ($P=0.0107$). We then analysed TACSTD1 expression according to the clinical response to IFN-α/5-FU combination therapy. TACSTD1 expression was higher in several non-responders with IFNAR2-positive HCC or IFNAR2-negative HCC, compared with responders with IFNAR2-positive HCC (Figure 1C). Using a cut-off value of 10 for TACSTD1 expression ratio, it was possible to exclude some non-responders from patients with IFNAR2-positive HCC.

Immunohistochemical staining for Ep-CAM

We examined the Ep-CAM expression in 30 HCC patients who underwent palliative reduction surgery. In tumour lesions, Ep-CAM staining was specifically observed on the plasma membrane of cancer cells. In Figure 1D (left), strong Ep-CAM expression was noted in 80% of cancerous tissue in the representative case of non-responders with IFNAR2-negative HCC. On the other hand, no Ep-CAM expression was evident in the representative case of IFNAR2-positive responders (Figure 1D, right). Among the 30 patients examined, Ep-CAM expression was observed in six (20%). It is important that Ep-CAM expression was associated with resistance to IFN-α/5-FU therapy, and no Ep-CAM expression was noted in the responders (Table 4). However, the difference in the expression

Table 2 List of informative 161 genes defining responders and non-responders

Rank	Status	Gene symbol	Gene name	Ref Seq ID
1	Down	—	dj329I24.3 (member of mcm2/3/5 family)	—
2	Down	CREBBP	CREB-binding protein (Rubinstein-Taybi syndrome)	NM_004380
3	Down	C20orf116	Chromosome 20 open reading frame 116	NM_023935
4	Down	—	Ensembl genscan prediction	—
5	Down	SLC25A40	Solute carrier family 25, member 40	NM_018843
6	Up	ZNF598	Zinc-finger protein 598	NM_178167
7	Down	RCOR1	REST corepressor 1	NM_015156
8	Up	TACSTD1	Tumour-associated calcium signal transducer 1	NM_002354
9	Down	ZNF397	Zinc-finger protein 397	NM_032347
10	Down	RTP3	Receptor (chemosensory) transporter protein 3	NM_031440
11	Up	AARS2	Alanyl-tRNA synthetase 2, mitochondrial (putative)	NM_020745
12	Down	RARA	Retinoic acid receptor, alpha	NM_000964
13	Up	—	DNA segment on chromosome 4 (unique) 234 expressed sequence	NM_014392
14	Up	CASP7	Caspase 7, apoptosis-related cysteine peptidase	NM_033339
15	Up	AZIN1	Antizyme inhibitor 1	NM_148174
16	Down	—	Ensembl prediction	—
17	Down	—	Ensembl genscan prediction	—
18	Up	PABPN1	Poly(A)-binding protein, nuclear 1	—
19	Down	NDUFS7	NADH dehydrogenase (ubiquinone) Fe-S protein 7, 20 kDa (NADH-coenzyme Q reductase)	NM_024407
20	Down	C9orf142	Chromosome 9 open reading frame 142	NM_183241
21	Down	—	Ensembl genscan prediction	—
22	Up	SNX21	Sorting nexin family member 21	NM_001042632
23	Down	C12orf47	Chromosome 12 open reading frame 47	XR_017874
24	Down	CCRL1	Chemokine (C–C motif) receptor-like 1	NM_178445
25	Down	GLG1	Golgi apparatus protein 1	NM_012201
26	Down	MRPS21	Mitochondrial ribosomal protein S21	NM_018997
27	Up	ABCA2	ATP-binding cassette, sub-family A (ABC1), member 2	NM_001606
28	Up	AXIN2	Axin 2 (conductin, axil)	NM_004655
29	Down	—	Ensembl genscan prediction	—
30	Down	NOXA1	NADPH oxidase activator 1	NM_006647
31	Down	COX4I1	Cytochrome c oxidase subunit IV isoform 1	NM_001861
32	Up	PCSK7	Proprotein convertase subtilisin/kexin type 7	XM_001128785
33	Up	FABP3	Fatty acid-binding protein 3, muscle and heart (mammary-derived growth inhibitor)	NM_004102
34	Down	C20orf111	Chromosome 20 open reading frame 111	NM_016470
35	Down	CAST	Calpastatin	NM_173061
36	Down	C12orf47	Chromosome 12 open reading frame 47	XR_017874
37	Up	—	Hypothetical protein xp_032244	—
38	Down	—	Ensembl genscan prediction	—
39	Up	PCDHA1	Protocadherin alpha 1	NM_018900
40	Down	MATK	Megakaryocyte-associated tyrosine kinase	NM_002378
41	Up	—	Hypothetical protein xp_051475	—
42	Up	UBE2Q1	Ubiquitin-conjugating enzyme E2Q (putative) 1	NM_017582
43	Up	GPATCH4	G patch domain containing 4	NM_182679
44	Down	PARP2	Poly (ADP-ribose) polymerase family, member 2	NM_005484
45	Down	HAL	Histidine ammonia-lyase	NM_002108
46	Up	ASCC3	Activating signal cointegrator 1 complex subunit 3	NM_006828
47	Down	KRTAP9-8	Keratin-associated protein 9–8	NM_031963
48	Up	MAGED4B	Melanoma antigen family D, 4B	NM_030801
49	Down	—	Hypothetical LOC339123	NM_001005920
50	Down	SPHK1	Sphingosine kinase 1	NM_021972
51	Up	—	Partial ighv ig h-chain v-region clone a81	—
52	Up	CCDC109A	Coiled-coil domain containing 109A	NM_138357
53	Up	GPR139	G protein-coupled receptor 139	NM_001002911
54	Up	C1orf78	Chromosome 1 open reading frame 78	NM_018166
55	Down	LRRC50	Leucine rich repeat containing 50	—
56	Down	FAM125B	Family with sequence similarity 125, member B	NM_033446
57	Down	IFT52	Intraflagellar transport 52 homolog (Chlamydomonas)	NM_016004
58	Down	C3orf36	Chromosome 3 open reading frame 36	NM_025041
59	Up	GUCA1B	Guanylate cyclase activator 1B (retina)	NM_002098
60	Down	EDF1	Endothelial differentiation-related factor 1	NM_003792
61	Down	CCDC69	Coiled-coil domain containing 69	NM_015621
62	Down	NDUFS6	NADH dehydrogenase (ubiquinone) Fe-S protein 6, 13 kDa (NADH-coenzyme Q reductase)	NM_004553
63	Up	CD93	CD93 molecule	NM_012072
64	Down	—	Ensembl genscan prediction	—
65	Down	ENO2	Enolase 2 (gamma, neuronal)	NM_001975
66	Down	CDCP2	CUB domain-containing protein 2	—
67	Down	—	Ensembl genscan prediction	—
68	Up	—	Similar to helicase-like protein nhl	—
69	Down	FOXN3	Forkhead box N3	NM_005197
70	Down	DEF8	Differentially expressed in FDCP 8 homolog (mouse)	NM_207514
71	Up	—	Hypothetical protein xp_034013	—

Table 2 (*Continued*)

Rank	Status	Gene symbol	Gene name	Ref Seq ID
72	Down	TNS3	Tensin 3	NM_022748
73	Up	FAM40A	Family with sequence similarity 40, member A	NM_033088
74	Down	PRDM7	PR domain containing 7	NM_052996
75	Up	—	Hypothetical protein xp_039419	—
76	Up	NNMT	Nicotinamide N-methyltransferase	NM_006169
77	Up	RPL9	Ribosomal protein L9	—
78	Up	ITM2C	Integral membrane protein 2C	NM_030926
79	Up	—	Ensembl genscan prediction	—
80	Up	BEX4	BEX family member 4	XM_936467
81	Up	CSNK1G1	Casein kinase 1, gamma 1	NM_022048
82	Up	EEF2K	Eukaryotic elongation factor-2 kinase	NM_013302
83	Down	DNAJC8	DnaJ (Hsp40) homolog, subfamily C, member 8	—
84	Up	GP5	Glycoprotein V (platelet)	NM_004488
85	Down	DPH3	DPH3, KTI11 homolog (*S. cerevisiae*)	NM_001047434
86	Down	—	Ensembl genscan prediction	—
87	Up	RPS21	Ribosomal protein S21	—
88	Down	—	Kiaa1658 protein	—
89	Down	ADCYAP1R1	Adenylate cyclase activating polypeptide 1 (pituitary) receptor type 1	NM_001118
90	Down	C5orf25	Chromosome 5 open reading frame 25	XR_015120
91	Down	—	PRO0132 protein	NR_002763
92	Down	FGF3	Fibroblast growth factor 3 (murine mammary tumour virus integration site (v-int-2) oncogene homolog)	NM_005247
93	Up	FABP7	Fatty acid-binding protein 7, brain	NM_001446
94	Down	HSD11B1	Hydroxysteroid (11-beta) dehydrogenase 1	NM_181755
95	Up	—	Chondroitin sulphate glucuronyltransferase	NM_019015
96	Down	OPTN	Optineurin	NM_001008213
97	Up	—	Erythroid differentiation-related factor 2	—
98	Down	—	Truncated alpha ig h-chain of disease patient har	—
99	Down	WT1	Wilms tumour 1	NM_024425
100	Down	C8G	Complement component 8, gamma polypeptide	NM_000606
101	Down	—	pro1454	—
102	Down	CADM1	Cell adhesion molecule 1	NM_001098517
103	Down	GH1	Growth hormone 1	—
104	Down	DNPEP	Aspartyl aminopeptidase	NM_012100
105	Up	—	Actin-like gene	—
106	Down	—	Ensembl genscan prediction	—
107	Down	—	Ensembl genscan prediction	—
108	Up	INTS4	Integrator complex subunit 4	NM_033547
109	Down	SRA1	Steroid receptor RNA activator 1	NM_001035235
110	Down	—	Ensembl genscan prediction	—
111	Up	RPS25	Ribosomal protein S25	NM_001028
112	Down	—	KIAA1450 protein	NM_020840
113	Up	SH2D3C	SH2 domain containing 3C	NM_170600
114	Down	NDUFA12	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 12	NM_018838
115	Down	—	NEFA-interacting nuclear protein NIP30	NM_024946
116	Down	TDRD1	tudor domain containing 1	NM_198795
117	Down	—	14a9ct dna sequence	—
118	Up	FBXW11	F-box and WD repeat domain containing 11	NM_012300
119	Down	CBFA2T3	Core-binding factor, runt domain, alpha subunit 2; translocated to, 3	NM_005187
120	Down	TCF3	Transcription factor 3 (E2A immunoglobulin enhancer-binding factors E12/E47)	NM_003200
121	Down	LASS5	LAG1 homolog, ceramide synthase 5	NM_147190
122	Down	—	Ensembl genscan prediction	—
123	Up	ACTR1B	ARP1 actin-related protein 1 homolog B, centractin beta (yeast)	NM_005735
124	Down	—	Hypothetical protein mgc5566	—
125	Up	RPS4X	Ribosomal protein S4, X-linked	XR_019325
126	Down	CDK6	Cyclin-dependent kinase 6	NM_001259
127	Up	AVIL	Advillin	—
128	Down	—	Hypothetical protein xp_043732	—
129	Down	C1orf136	Chromosome 1 open reading frame 136	—
130	Down	—	Hypothetical protein xp_043783	—
131	Down	—	Ews-fli-1	—
132	Up	CDC42BPG	CDC42-binding protein kinase gamma (DMPK-like)	NM_017525
133	Down	—	Ensembl genscan prediction	—
134	Down	GDAP1L1	Ganglioside-induced differentiation-associated protein 1-like 1	NM_024034
135	Up	C12orf4	Chromosome 12 open reading frame 4	NM_020374
136	Up	KIAA0415	KIAA0415	NM_014855
137	Down	PDLIM2	PDZ and LIM domain 2 (mystique)	NM_198042
138	Down	KHK	Ketohexokinase (fructokinase)	NM_006488
139	Down	SLC36A1	Solute carrier family 36 (proton/amino acid symporter), member 1	NM_078483
140	Up	—	Hypothetical protein xp_050311	—
141	Up	TBRG4	Transforming growth factor β regulator 4	NM_199122
142	Down	—	Rearranged vk3 of Hodgkin cell line	—

Table 2 (*Continued*)

Rank	Status	Gene symbol	Gene name	Ref Seq ID
143	Up	CD59	CD59 molecule, complement regulatory protein	NM_203330
144	Down	PEX26	Peroxisome biogenesis factor 26	—
145	Up	VEGFC	Vascular endothelial growth factor C	NM_005429
146	Down	DTX2	Deltex homolog 2 (*Drosophila*)	XM_941785
147	Up	ELAVL3	ELAV (embryonic lethal, abnormal vision, *Drosophila*)-like 3 (Hu antigen C)	NM_032281
148	Up	BSDC1	BSD domain containing 1	NM_018045
149	Down	FUBP3	Far upstream element (FUSE)-binding protein 3	XM_001128545
150	Down	CCDC48	Coiled-coil domain containing 48	NM_024768
151	Down	EPHA6	EPH receptor A6	NM_001080448
152	Down	ST8SIA1	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1	NM_003034
153	Down	MKKS	McKusick-Kaufman syndrome	NM_018848
154	Down	MGA	MAX gene associated	NM_001080541
155	Up	—	Hypothetical protein xp_040140	—
156	Down	—	Hypothetical protein xp_043452	—
157	Up	MMP20	Matrix metallopeptidase 20 (enamelysin)	NM_004771
158	Up	SLC23A2	Solute carrier family 23 (nucleobase transporters), member 2	NM_005116
159	Up	GABARAPL1	GABA(A) receptor-associated protein like 1	NM_031412
160	Down	—	Ensembl genscan prediction	—
161	Up	PKLR	Pyruvate kinase, liver and RBC	NM_000298

Ranking was according to absolute value of signal-to-noise ratio. Status was defined as expression in non-responders compared with responders.

Table 3 List of significant pathways from 161 informative genes by canonical pathway analysis

Pathway	*P*-value
Ubiquinone biosynthesis	0.0004
Oxidative phosphorylation	0.0074
Mitochondrial dysfunction	0.0095
FXR/RXR activation	0.0162
Wnt/β-catenin signalling	0.0170
Complement system	0.0191
Histidine metabolism	0.0263
Sphingolipid metabolism	0.0389

rate was not significant probably because of the small sample size ($P=0.0528$). In non-tumour lesions, Ep-CAM staining was observed in a few scattered cells and proliferating bile duct epithelium showed positive expression.

Analysis of the degree of Ep-CAM expression in tumour lesions showed five (16.7%) samples negative for Ep-CAM expression (score 0), 19 (63.3%) with weak expression (score 1–4), four (13.3%) stained moderately (score 6–8) and two (6.7%) samples exhibited strong Ep-CAM expression (score 9–12). These results suggest that Ep-CAM expression in advanced HCC could be a potentially useful marker for resistance to IFN-α/5-FU combination therapy.

Ep-CAM expression and activation of Wnt/β-catenin signalling by BIO

We analysed the protein expression level of Ep-CAM in hepatoma cell lines. Western blot analysis using an anti-Ep-CAM antibody confirmed the positive expression of Ep-CAM in three of the six cell lines (HuH7, HepG2 and Hep3B), whereas PLC/PRF/5, HLE and HLF were negative (Figure 2A). We earlier reported strong IFNAR2 expression in PLC/PRF/5 cells and weak IFNAR2 expression in HuH7 cells (Eguchi *et al*, 2000). We transfected a TCF/LEF reporter into PLC/PRF/5, HuH7 and HepG2 cells to evaluate TCF/LEF transcriptional activity, representing the activity of Wnt/β-catenin signalling pathway. We found that the luciferase activities were high in both Ep-CAM-positive HuH7 cells and HepG2 cells, whereas very low in Ep-CAM-negative PLC/PRF/5 cells (Figure 2B). And, HepG2 cell line was reported to have mutation and activated β-catenin (de La Coste *et al*, 1998). For these reasons, we used the cell line HuH7 to investigate how Wnt/β-catenin signalling affected on the growth-inhibitory effect of IFN-α/5-FU. In the next step, we examined whether Wnt/β-catenin signalling can be activated in HuH7 cells when treated with various concentrations of specific GSK-3 inhibitor, BIO. HuH7 cells treated with BIO showed a substantial, dose-dependent increase in TCF/LEF reporter activity. Consequently, treatment with BIO at 0.5, 1 and 5 nM induced 8.6-, 29.1-, and 48.6-fold increases in relative luciferase activity compared with HuH7 cells treated by DMSO, respectively (Figure 2C). To examine the effects of BIO on the expression of Wnt/β-catenin signalling targeted genes, qRT–PCR analysis of five targeted genes (TACSTD1, AXIN2, MYC, TCF3 and CCND1) was carried out in HuH7 cells after 24 h treatment with BIO (5 nM). The concentration of BIO was selected on the basis of the results of luciferase reporter assay. Treatment with BIO increased the mRNA expression of targeted genes from 1.3-fold to 7.6-fold compared with cells treated with DMSO (Figure 2D). In western blot analysis, the expression levels of Ep-CAM and c-MYC increased in a BIO dose-dependent manner in HuH7 cells, but not in PLC/PRF/5 cells (Figure 2E).

Growth inhibition assay and reduction of growth-inhibitory effect of 5-FU and/or IFN-α treatment

Next, we investigated the role of activation of Wnt/β-catenin signalling in the reduction of the growth-inhibitory effect of IFN-α/5-FU. The growth of HCC cells (PLC/PRF/5 and HuH7 cell lines) was suppressed by 5-FU and IFN-α in a dose-dependent manner. Concurrent addition of BIO and IFN-α/5-FU to the cell cultures significantly reduced the growth-inhibitory effects of IFN-α/5-FU in HuH7 cells. In HuH7 cells, the growth inhibitory effects of IFN-α/5-FU without BIO were $22.3\pm2.8\%$ at $0.5\,\mu g\,ml^{-1}$ 5-FU and 500 $U\,ml^{-1}$ IFN-α, and $44.6\pm0.9\%$ at $5\,\mu g\,ml^{-1}$ and $5000\,U\,ml^{-1}$. Concurrent addition of BIO decreased the growth inhibitory effect to $8.6\pm3.9\%$ ($P=0.0012$; $0.5\,\mu g\,ml^{-1}$ of 5-FU and 500 $U\,ml^{-1}$ for IFN-α) and $29.0\pm2.0\%$ ($P<0.0001$, $5\,\mu g\,ml^{-1}$ for 5-FU and 5000 $U\,ml^{-1}$ for IFN-α) of control cells. In contrast, no change in the growth-inhibitory effect was found in PLC/PRF/5 cell line (Figure 3A). We also investigated the effects of BIO when combined with 5-FU alone or IFN-α alone in HuH7 cells. The