Table 1. Baseline characteristics of all patients, and patients assigned to the model building or validation groups.

	All patients n = 496	Model group n = 331	Validation group n = 165
Gender: male	250 (50%)	170 (51%)	80 (48%)
Age (years)	57.1 ± 9.9	56.8 ± 9.7	57.5 ± 10.2
ALT (IU/L)	78.6 ± 60.8	78.1 ± 61.4	79.7 ± 59.6
GGT (IU/L)	59.3 ± 63.6	58.9 ± 62.0	60.2 ± 66.9
Platelets (10^9/L)	154 ± 53	153 ± 52	154 ± 56
Fibrosis: F3-4	121 (24%)	80 (24%)	41 (25%)
HCV-RNA: >600,000 IU/ml	409 (82%)	273 (82%)	136 (82%)
ISDR mutation: ≤1	220 (88%)	290 (88%)	145 (88%)
Core 70 (Arg/Gln or His)	293 (59%)/203 (41%)	197 (60%)/134 (40%)	96 (58%)/69 (42%)
Core 91 (Leu/Met)	299 (60%)/197 (40%)	200 (60%)/131 (40%)	99 (60%)/66 (40%)
IL28B: Minor allele	151 (30%)	101 (31%)	50 (30%)
SVR	194 (39%)	129 (39%)	65 (39%)
Relapse	152 (31%)	103 (31%)	49 (30%)
NVR	150 (30%)	99 (30%)	51 (31%)

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Arg, arginine; Gln, glutamine; His, histidine; Leu, leucine; Met, methionine; Minor, heterozygote or homozygote of minor allele; SVR, sustained virological response; NVR, null virological response.

Japanese [6], European [7], and a multi-ethnic population [8,9]. The last three studies focused on the association of SNPs in the *IL28B* region with SVR [7–9] but we found a stronger association with NVR [6]. In addition to these host genetic factors, we have reported that mutations within a stretch of 40 amino acids in the NS5A region of HCV, designated as the IFN sensitivity determining region (ISDR), are closely associated with the virological response to IFN therapy: a lower number of mutations is associated with treatment failure [10–13]. Amino acid substitutions at positions 70 and 91 of the HCV core region (Core70, Core91) also have been reported to be associated with response to PEG-IFN/RBV therapy: glutamine (Gln) or histidine (His) at Core70 and methionine (Met) at Core91 are associated with treatment resistance [4,14]. The importance of substitutions in the HCV core and ISDR was confirmed recently by a Japanese multicenter study [15]. How these viral factors contribute to response to therapy is yet to be determined. For general application in clinical practice, host genetic factors and viral factors should be considered together.

Data mining analysis is a family of non-parametric regression methods for predictive modeling. Software is used to automatically explore the data to search for optimal split variables and to build a decision tree structure [16]. The major advantage of decision tree analysis over logistic regression analysis is that the results of the analysis are presented in the form of flow chart, which can be interpreted intuitively and readily made available for use in clinical practice [17]. The decision tree analysis has been utilized to define prognostic factors in various diseases [18–25]. We have reported recently its usefulness for the prediction of an early virological response (undetectable HCV-RNA within 12 weeks of therapy) to PEG-IFN/RBV therapy in chronic hepatitis C [26].

This study aimed to define the pre-treatment prediction of response to PEG-IFN/RBV therapy through the integrated analysis of host factors, such as the *IL28B* genetic polymorphism and various clinical covariates, as well as viral factors, such as mutations in the HCV core and ISDR and serum HCV-RNA load. In addition, for the general application of these results in clinical practice, decision models for the pre-treatment prediction of response were determined by data mining analysis.

Materials and methods

Patients

This was a multicentre retrospective study supported by the Japanese Ministry of Health, Labor and Welfare. Data were collected from a total of 496 chronic hepatitis C patients who were treated with PEG-IFN alpha and RBV at five hospitals and universities throughout Japan. Of these, 98 patients also were included in the original GWAS analysis [6]. The inclusion criteria in this study were as follows (1) infection by genotype 1b, (2) lack of co-infection with hepatitis B virus or human immunodeficiency virus, (3) lack of other causes of liver disease, such as autoimmune hepatitis, and primary biliary cirrhosis, (4) completion of at least 24 weeks of therapy, (5) adherence of more than 80% to the planned dose of PEG-IFN and RBV for the NVR patients, (6) availability of DNA for the analysis of the genetic polymorphism of *IL28B*, and (7) availability of serum for the determination of mutations in the ISDR and substitutions of Core70 and Core91 of HCV. Patients received PEG-IFN alpha-2a (180 μg) or 2b (1.5 μg/kg) subcutaneously every week and were administered a weight adjusted dose of RBV (600 mg for <60 kg, 800 mg for 60–80 kg, and 1000 mg for >80 kg daily) which is the recommended dosage in Japan. Written informed consent was obtained from each patient and the study protocol conformed to the ethical guidelines of the Declaration of Helsinki and was approved by the institutional ethics review committee. The baseline characteristics are listed in Table 1. For the data mining analysis, 67% of the patients (331 patients) were assigned randomly to the model building group and 33% (165 patients) to the validation group. There were no significant differences in the clinical backgrounds between these two groups.

Laboratory and histological tests

Blood samples were obtained before therapy and were analyzed for hematologic tests and for blood chemistry and HCV-RNA. Sequences of ISDR and the core region of HCV were determined by direct sequencing after amplification by reverse-transcription and polymerase chain reaction as reported previously [4,11]. Genetic polymorphism in one tagging SNP located near the *IL28B* gene (rs8099917) was determined by the GWAS or DigiTag2 assay [27]. Homozygosity (GG) or heterozygosity (TG) of the minor sequence was defined as having the *IL28B* minor allele, whereas homozygosity for the major sequence (TT) was

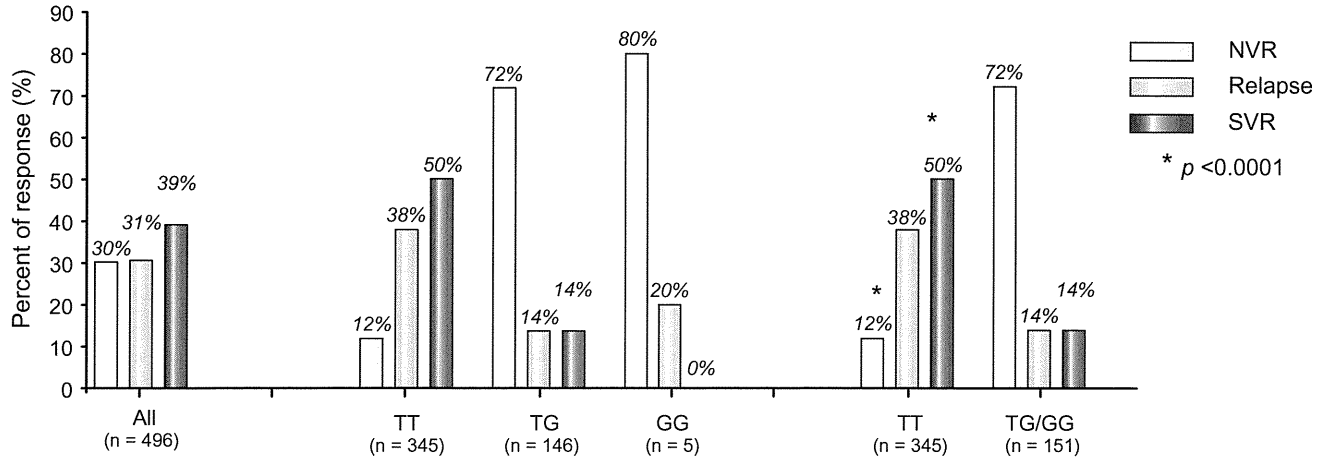

Fig. 1. Association between the IL28B genotype (rs8099917) and treatment response. The rates of response to treatment are shown for each rs8099917 genotype. The rate of null virological response (NVR), relapse, and sustained virological response (SVR) is shown. The *p* values are from Fisher's exact test. The rate of NVR was significantly higher ($p < 0.0001$) and the rate of SVR was significantly lower ($p < 0.0001$) in patients with the *IL28B* minor allele compared to those with the major allele.

defined as having the *IL28B* major allele. In this study, NVR was defined as a less than 2 log reduction of HCV-RNA at week 12 and detectable HCV-RNA by qualitative PCR with a lower detection limit of 50 IU/ml (Amplicor, Roche Diagnostic systems, CA) at week 24 during therapy. RVR (rapid virological response) and complete early virological response (cEVR) were defined as undetectable HCV-RNA at 4 weeks and 12 weeks during therapy and SVR was defined as undetectable HCV-RNA 24 weeks after the completion of therapy. Relapse was defined as reappearance of HCV-RNA after the completion of therapy. The stage of liver fibrosis was scored according to the METAVIR scoring system: F0 (no fibrosis), F1 (mild fibrosis: portal fibrosis without septa), F2 (moderate fibrosis: few septa), F3 (severe fibrosis: numerous septa without cirrhosis) and F4 (cirrhosis). Percentage of steatosis was quantified in 111 patients by determining the average proportion of hepatocytes affected by steatosis.

Statistical analysis

Associations between pre-treatment variables and treatment response were analyzed by univariate and multivariate logistic regression analysis. Associations between the *IL28B* polymorphism and sequences of HCV were analyzed by Fisher's exact test. SPSS software v.15.0 (SPSS Inc., Chicago, IL) was used for these analyses. For the data mining analysis, IBM-SPSS Modeler version 13.0 (IBM-SPSS Inc., Chicago, IL) software was utilized as reported previously [26]. The patients used for model building were divided into two groups at each step of the analysis based on split variables. Each value of each variable was considered as a potential split. The optimum variables and cut-off values were determined by a statistical search algorithm to generate the most significant division into two prognostic subgroups that were as homogeneous as possible for the probability of SVR. Thereafter, each subgroup was evaluated again and divided further into subgroups. This procedure was repeated until no additional significant variable was detected or the sample size was below 15. To avoid over-fitting, 10-fold cross validation was used in the tree building process. The reproducibility of the resulting model was tested with the data from the validation patients.

Results

Association between the IL28B (rs8099917) genotype and the PEG-IFN/RBV response

The rs8099917 allele frequency was 70% for TT ($n = 345$), 29% for TG ($n = 146$), and 1% for GG ($n = 5$). We defined the *IL28B* major allele as homozygous for the major sequence (TT) and the *IL28B* minor allele as homozygous (GG) or heterozygous (TG) for the minor sequence. The rate of NVR was significantly higher (72% vs. 12%, $p < 0.0001$) and the rate of SVR was significantly lower (14% vs. 50%, $p < 0.0001$) in patients with the *IL28B* minor allele compared to those with the major allele (Fig. 1).

Effect of the IL28B polymorphism, substitutions in the ISDR, Core70, and Core91 of HCV on time-dependent clearance of HCV

Patients were stratified according to their *IL28B* allele type, the number of mutations in the ISDR, the amino acid substitutions in Core70 and Core91, and the rate of undetectable HCV-RNA at 4, 8, 12, 24, and 48 weeks after the start of therapy were analyzed (Fig. 2A–D). The rate of undetectable HCV-RNA was significantly higher in patients with the *IL28B* major allele than the minor allele, in patients with two or more mutations in the ISDR compared to none or only one mutation, in patients with arginine (Arg) at Core70 rather than Gln/His, and in patients with leucine (Leu) at Core91 rather than Met. The difference was most significant when stratified by the *IL28B* allele type. The rate of RVR and cEVR was significantly more frequent in patients with the *IL28B* major allele compared with those with the *IL28B* minor allele: 9% vs. 3% for RVR ($p < 0.005$) and 57% vs. 11% for cEVR ($p < 0.0001$). These findings suggest that *IL28B* has the greatest impact on early virological response to therapy.

Association between substitutions in the ISDR and relapse after the completion of therapy

Patients were stratified according to the *IL28B* allele, number of mutations in the ISDR, and amino acid substitutions of Core70 and Core91, and the rate of relapse was analyzed (Fig. 3A and B). Among patients who achieved cEVR, the rate of relapse was significantly lower in patients with two or more mutations in the ISDR compared to those with only one or no mutations (15% vs. 31%, $p < 0.005$) (Fig. 3 B). On the other hand, the relapse rate was not different between the *IL28B* major and minor alleles within patients who achieved RVR (3% vs. 0%) or cEVR (28% vs. 29%) (Fig. 3A). Amino acid substitutions of Core70 and Core91 were not associated with the rate of relapse (data not shown).

Factors associated with response by multivariate logistic regression analysis

By univariate analysis, the minor allele of *IL28B* ($p < 0.0001$), one or no mutations in the ISDR ($p = 0.03$), high serum level of

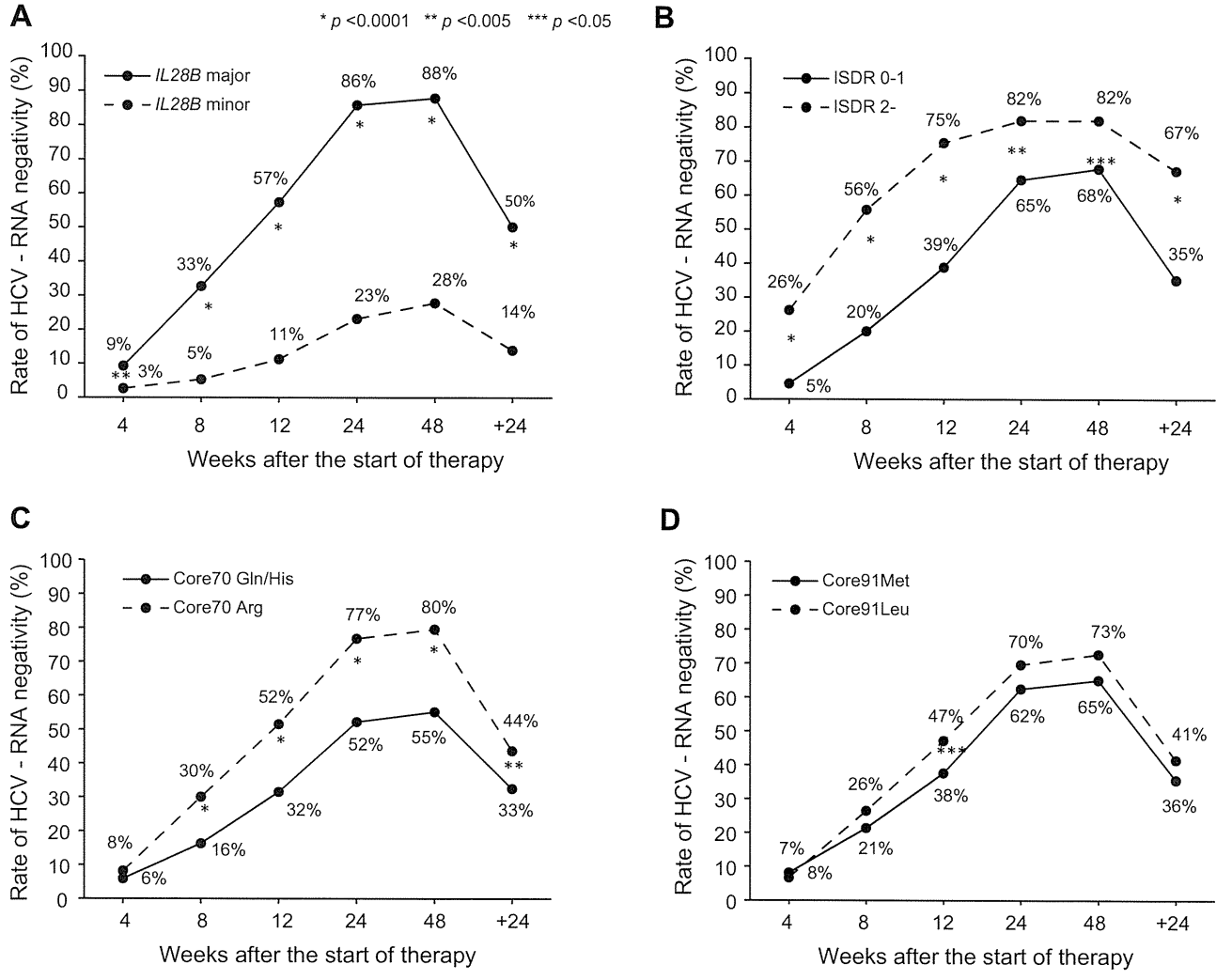

Fig. 2. Effect of *IL28B* mutations in the ISDR, Core70, and Core91 of HCV on time-dependent clearance of HCV. The rate of undetectable HCV-RNA was plotted for serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks) and for 24 weeks after the completion of therapy. Patients were stratified according to (A) the *IL28B* allele (minor allele vs. major allele), (B) the number of mutations in the ISDR (0–1 mutation vs. 2 or more mutations), amino acid substitutions of (C) Core70 (Gln/His vs. Arg), and (D) Core91 (Met vs. Leu). The *p* values are from Fisher's exact test.

HCV-RNA ($p = 0.035$), Gln or His at Core70 ($p < 0.0001$), low platelet counts ($p = 0.009$), and advanced fibrosis ($p = 0.0002$) were associated with NVR. By multivariate analysis, the minor allele of *IL28B* (OR = 20.83, 95%CI = 11.63–37.04, $p < 0.0001$) was associated with NVR independent of other covariates (Table 2). Notably, mutations in the ISDR ($p = 0.707$) and at amino acid Core70 ($p = 0.207$) were not significant in multivariate analysis due to the positive correlation with the *IL28B* polymorphism ($p = 0.004$ for ISDR and $p < 0.0001$ for Core70, Fig. 4).

Genetic polymorphism of *IL28B* also was associated with SVR (OR = 7.41, 95% CI = 4.05–13.57, $p < 0.0001$) independent of other covariates, such as platelet counts, fibrosis, and serum levels of HCV-RNA. Mutation in the ISDR was an independent predictor of SVR (OR = 2.11, 95% CI = 1.06–4.18, $p = 0.033$) but the amino acid at Core70 was not (Table 3).

Factors associated with the IL28B polymorphism

Patients with the *IL28B* minor allele had significantly higher serum level of gamma-glutamyltransferase (GGT) and a higher frequency of hepatic steatosis (Table 4). When the association between the *IL28B* polymorphism and HCV sequences was analyzed, Gln or His at Core70, that is linked to resistance to PEG-IFN and RBV therapy [4,14,15], was significantly more frequent in patients with the minor *IL28B* allele than in those with the major allele (67% vs. 30%, $p < 0.0001$) (Fig. 4). Other HCV sequences with an IFN resistant phenotype also were more prevalent in patients with the minor *IL28B* allele than those with the major allele: Met at Core91 (46% vs. 37%, $p = 0.047$) and one or no mutations in the ISDR (94% vs. 85%, $p = 0.004$) (Fig. 4).

Data mining analysis

Data mining analysis was performed to build a model for the prediction of SVR and the result is shown in Fig. 5. The analysis selected four predictive variables, resulting in six subgroups of patients. Genetic polymorphism of *IL28B* was selected as the best predictor of SVR. Patients with the minor *IL28B* allele had a lower probability of SVR and a higher probability of NVR than those with the major *IL28B* allele (SVR: 14% vs. 50%, NVR: 72% vs.

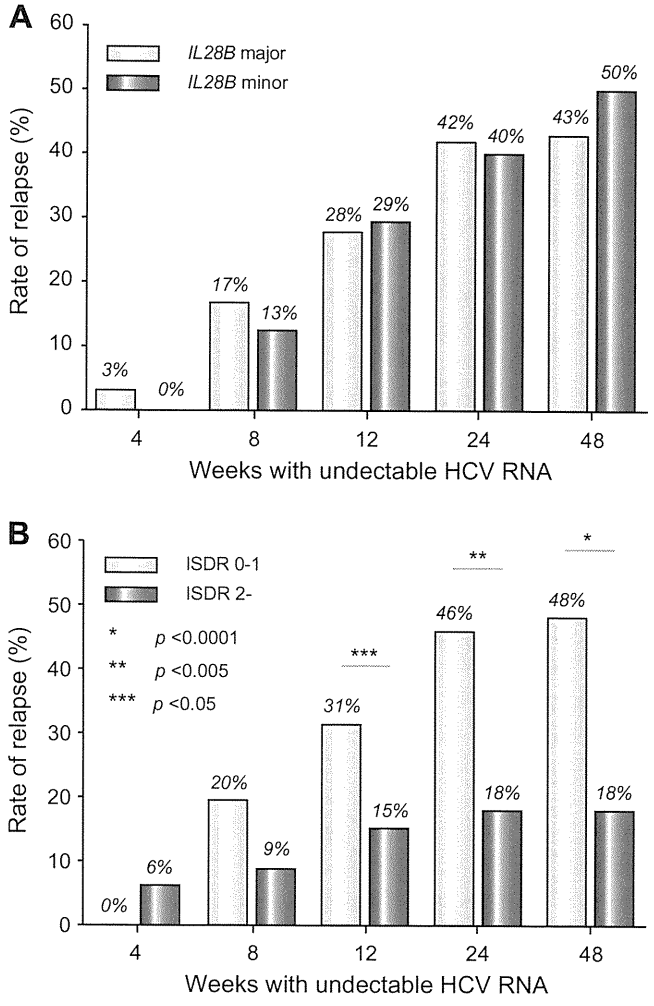

Fig. 3. Association between relapse and the *IL28B* allele or mutations in the ISDR. The rate of relapse was calculated for patients who had undetectable HCV-RNA at serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks). Patients were stratified according to (A) the *IL28B* allele (minor allele vs. major allele) and (B) the number of mutations in the ISDR (0–1 mutation vs. 2 or more mutations). The *p* values are from Fisher's exact test.

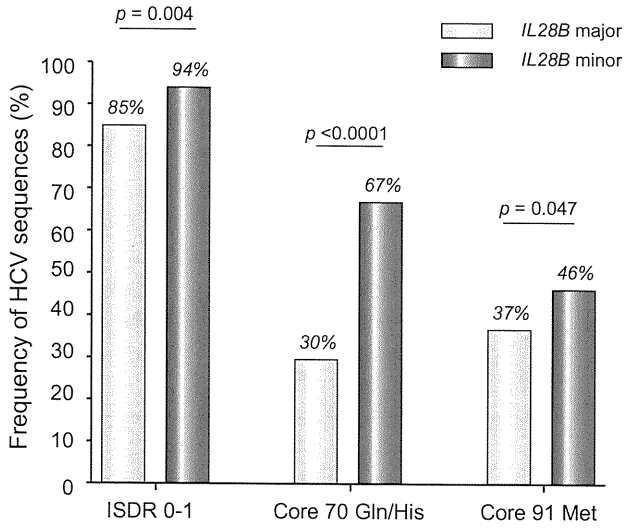

Fig. 4. Associations between the *IL28B* allele and HCV sequences. The prevalence of HCV sequences predicting a resistant phenotype to IFN was higher in patients with the minor *IL28B* allele than those with major allele. (A) 0 or 1 mutation in the ISDR of NS5A, (B) Gln or His at Core70, and (C) Met at Core91. *p* values are from Fisher's exact test.

12%). After stratification by the *IL28B* allele, patients with low platelet counts (<140 × 10^9/L) had a lower probability of SVR and higher probability of NVR than those with high platelet counts (⩾140 × 10^9/L): for the minor *IL28B* allele, SVR was 7% vs. 19%, and NVR was 84% vs. 62%, and for the major *IL28B* allele, SVR was 32% vs. 66% and NVR was 16% vs. 8%. Among patients with the major *IL28B* allele and low platelet counts, those with two or more mutations in the ISDR had a higher probability of SVR and lower probability of relapse than those with one or no mutations in the ISDR (SVR: 75% vs. 27%, and relapse: 8% vs. 57%). Among patients with the major *IL28B* allele and high platelet counts, those with a low HCV-RNA titer (<600,000 IU/ml) had a higher probability of SVR and lower probability of NVR and relapse than those with a high HCV-RNA titer (SVR: 90% vs. 61%, NVR: 0% vs. 10%, and relapse: 10% vs. 29%). The sensitivity and specificity of the decision tree were 78% and 70%, respectively. The area under the receiver operating characteristic (ROC) curve of the model was 0.782 (data not shown). The proportion of patients with advanced fibrosis (F3-4) was 39% (84/217) in patients with low platelet counts (<140 × 10^9/L) compared to 13% (37/279) in those with high platelet counts (⩾140 × 10^9/L).

Validation of the data mining analysis

The results of the data mining analysis were validated with 165 patients who differed from those used for model building. Each patient was allocated to one of the six subgroups for the validation using the flow-chart form of the decision tree. The rate of SVR and NVR in each subgroup was calculated. The rates of SVR and NVR for each subgroup of patients were closely correlated between the model building and the validation patients (r^2 = 0.99 and 0.98) (Fig. 6).

Discussion

The rate of NVR after 48 weeks of PEG-IFN/RBV therapy among patients infected with HCV of genotype 1 is around 20–30%. Previously, there have been no reliable baseline predictors of NVR or SVR. Because more potent therapies, such as protease and polymerase inhibitor of HCV [28,29] and nitazoxianide [30], are in clinical trials and may become available in the near future, a pre-treatment prediction of the likelihood of response may be helpful for patients and physicians, to support clinical decisions about whether to begin the current standard of care or whether to wait for emerging therapies. This study revealed that the *IL28B* polymorphism was the overwhelming predictor of NVR and is independent of host factors and viral sequences reported previously. The *IL28B* encodes a protein also known as IFN-lambda 3, which is thought to suppress the replication of various viruses including HCV [31,32]. The results of the current study and the findings of the GWAS studies [6–9] may provide the rationale for developing diagnostic testing or an IFN-lambda based therapy for chronic hepatitis C in the future.

Table 2. Factors associated with NVR analyzed by univariate and multivariate logistic regression analysis.

	Univariate			Multivariate		
	Odds ratio	95%CI	*p* value	Odds ratio	95%CI	*p* value
Gender: female	0.98	0.67-1.45	0.938	1.29	0.75-2.23	0.363
Age	1.01	0.97-1.01	0.223	0.99	0.97-1.02	0.679
ALT	1.00	1.00-1.00	0.867	1.00	0.99-1.00	0.580
GGT	1.004	1.00-1.01	0.029	1.00	1.00-1.00	0.715
Platelets	0.95	0.91-0.99	0.009	0.92	0.87-0.98	0.006
Fibrosis: F3-4	2.23	1.46-3.42	0.0002	1.97	1.09-3.57	0.025
HCV-RNA: ≥600,000 IU/ml	1.83	1.05-3.19	0.035	2.49	1.17-5.29	0.018
ISDR mutation: ≤1	2.14	1.08-4.22	0.030	0.96	0.78-1.18	0.707
Core 70 (Gln/His)	3.23	2.16-4.78	<0.0001	1.41	0.83-2.42	0.207
Core 91 (Met)	1.39	0.95-2.06	0.093	1.21	0.72-2.04	0.462
IL28B: Minor allele	19.24	11.87-31.18	<0.0001	20.83	11.63-37.04	<0.0001

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Gln, glutamine; His, histidine; Met, methionine; Minor allele, heterozygote or homozygote of minor allele.

Table 3. Factors associated with SVR analyzed by univariate and multivariate logistic regression analysis.

	Univariate			Multivariate		
	Odds ratio	95%CI	*p* value	Odds ratio	95%CI	*p* value
Gender: female	0.81	0.56-1.16	0.253	0.86	0.55-1.35	0.508
Age	0.97	0.95-0.99	0.0003	0.99	0.96-1.01	0.199
ALT	1.00	1.00-1.00	0.337	1.00	1.00-1.01	0.108
GGT	1.00	1.00-1.00	0.273	1.00	1.00-1.00	0.797
Platelets	1.12	1.01-116	<0.0001	1.13	1.08-1.19	<0.0001
Fibrosis: F0-2	2.64	1.65-4.22	<0.0001	1.87	1.07-3.28	0.029
HCV-RNA: <600,000 IU/ml	2.49	1.55-3.98	0.0001	2.75	1.55-4.90	0.001
ISDR mutation: 2≤	3.78	2.14-6.68	<0.0001	2.11	1.06-4.18	0.033
Core 70 (Arg)	1.61	1.11-2.28	0.012	0.84	0.52-1.35	0.470
Core 91 (Leu)	1.28	0.88-1.85	0.185	1.26	0.81-1.96	0.300
IL28B: Major allele	6.21	3.75-10.31	<0.0001	7.41	4.05-13.57	<0.0001

ALT, alanine aminotransferase; GGT, Gamma-glutamyltransferase; ISDR, interferon sensitivity determining region; Arg, arginine; Leu, leucine; Major allele, homozygote of major allele.

Among baseline factors, *IL28B* was the most significant predictor of NVR and SVR. Moreover, the *IL28B* allele type was also correlated with early virological response: the rate of RVR and cEVR was significantly high for the *IL28B* major allele compared to the *IL28B* minor allele: 9% vs. 3% for RVR and 57% vs. 11% for cEVR (Fig. 2). On the other hand, the relapse rate was not different between the *IL28B* genotypes within patients who achieved RVR or cEVR (Fig. 3). We believe that optimal therapy should be based on baseline features and a response-guided approach. Our findings suggest that the *IL28B* genotype is a useful baseline predictor of virological response which should be used for selecting the treatment regimen: whether to treat patients with PEG-IFN and RBV or to wait for more effective future therapy including direct acting antiviral drugs. On the other hand, baseline *IL28B* genotype might not be suitable for determining the treatment duration in patients who started PEG-IFN/RBV therapy and whose virological response is determined because the *IL28B* genotype is not useful for the prediction of relapse. The duration of therapy should be personalized based on the virological response. Future studies need to explore whether the combination of baseline *IL28B* genotype and response-guided approach further improves the optimization of treatment duration.

The SVR rate in patients having the *IL28B* minor allele was 14% in the present study while it was 23% in Caucasians and 9% in African Americans in a study by McCarthy et al. [33]. On the other hand, the SVR rate in patients having the *IL28B* minor allele was 28% in genotypes 1/4 compared to 80% in genotypes 2/3 in a study by Rauch et al. [9]. These data imply that the impact of the *IL28B* polymorphism on response to therapy may be different in terms of race, geographical areas, or HCV genotypes, and that our data need to be validated in future studies including different populations and geographical areas before generalization.

Table 4. Factors associated with IL28B genotype.

	IL28B major allele n = 345	*IL28B* minor allele n = 151	*p* value
Gender: male	166 (48%)	84 (56%)	0.143
Age (years)	57 ± 10	57 ± 10	0.585
ALT (IU/L)	79 ± 60	78 ± 62	0.842
Platelets (10^9/L)	153 ± 54	155 ± 52	0.761
GGT (IU/L)	51 ± 45	78 ± 91	0.001
Fibrosis: F3-4	76 (22%)	45 (30%)	0.063
Steatosis:			
>10%	16/88 (18%)	13/23 (57%)	0.024
>30%	6/88 (7%)	6/23 (26%)	0.017
HCV-RNA: >600,000 IU/ml	284 (82%)	125 (83%)	1.000

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase.

Four GWAS studies have shown the association between a genetic polymorphism near the *IL28B* gene and response to PEG-IFN plus RBV therapy. The SNPs that showed significant association with response were rs12979860 [8] and rs8099917 [6,7,9]. There is a strong linkage-disequilibrium (LD) between these two SNPs as well as several other SNPs near the *IL28B* gene in Japanese patients [34] but the degree of LD was weaker in Caucasians and Hispanics [8]. Thus, the combination of SNPs is not useful for predicting response in Japanese patients but may improve the predictive value in patients other than Japanese who have weaker LD between SNPs.

Other significant predictors of response independent of *IL28B* genotype were platelet counts, stage of fibrosis, and HCV RVA load. A previous study reported that platelet count is a predictor of response to therapy [35], and the lower platelet count was related with advanced liver fibrosis in the present study. The association between response to therapy and advanced fibrosis independent of the *IL28B* polymorphism is consistent with a recent study by Rauch et al. [9].

There is agreement that the viral genotype is significantly associated with the treatment outcome. Moreover, viral factors such as substitutions in the ISDR of the NS5A region [10] or in the amino acid sequence of the HCV core [4] have been studied in relation to the response to IFN treatment. The amino acid Gln or His at Core70 and Met at Core91 are repeatedly reported to be associated with resistance to therapy [4,14,15] in Japanese patients but these data wait to be validated in different populations or other geographical areas. In this study, we confirmed that patients with two or more mutations in the ISDR had a higher rate of undetectable HCV-RNA at each time point during therapy. In addition, the rate of relapse among patients who achieved cEVR was significantly lower in patients with two or more mutations in ISDR compared to those with only one or no mutations (15% vs. 31%, $p < 0.05$). Thus, the ISDR sequence may be used to predict a relapse among patients who achieved virological response during therapy, while the *IL28B* polymorphism may be used to predict the virological response before therapy. A higher number of mutations in the ISDR are reported to have close association with SVR in Japanese [11–13,15,36] or Asian [37,38] populations but data from Western countries have been controversial [39–42]. A meta-analysis of 1230 patients including 525 patients from Europe has shown that there was a positive correlation between the SVR and the number of mutations in the ISDR in Japanese as well as in European patients [43] but this correlation was more pronounced in Japanese patients. Thus, geographical factors may account for the different impact of ISDR on treatment response, which may be a potential limitation of our study.

To our surprise, these HCV sequences were associated with the *IL28B* genotype: HCV sequences with an IFN resistant phenotype were more prevalent in patients with the minor IL28B allele than those with the major allele. This was an unexpected finding, as we initially thought that host genetics and viral sequences were completely independent. A recent study reported that the *IL28B* polymorphism (rs12979860) was significantly associated with HCV genotype: the *IL28B* minor allele was more frequent in HCV genotype 1-infected patients compared to patients infected with HCV genotype 2 or 3 [33]. Again, patients with the *IL28B* minor allele (IFN resistant genotype) were infected with HCV sequences that are linked to an IFN resistant phenotype. The mechanism for this association is unclear, but may be related to an interaction between the *IL28B* genotype and HCV sequences in the development of chronic HCV infection as discussed by McCarthy et al., since the *IL28B* polymorphism was associated with the natural clearance of HCV [44]. Alternatively, the HCV sequence within the patient may be selected during the course of chronic infection [45,46]. These hypotheses should be explored through prospective studies of spontaneous HCV clearance or by testing the time-dependent changes in the HCV sequence during the course of chronic infection.

How these host and viral factors can be integrated to predict the response to therapy in future clinical practice is an important question. Because various host and viral factors interact in the same patient, predictive analysis should consider these factors in combination. Using the data mining analysis, we constructed a simple decision tree model for the pre-treatment prediction of SVR and NVR to PEG-IFN/RBV therapy. The classification of patients based on the genetic polymorphism of *IL28B*, mutation in the ISDR, serum levels of HCV-RNA, and platelet counts, identified subgroups of patients who have the lowest probabilities of NVR (0%) with the highest probabilities of SVR (90%) as well as those who have the highest probabilities of NVR (84%) with the lowest probability of SVR (7%). The reproducibility of the model was confirmed by the independent validation based on a second group of patients. Using this model, we can rapidly develop an

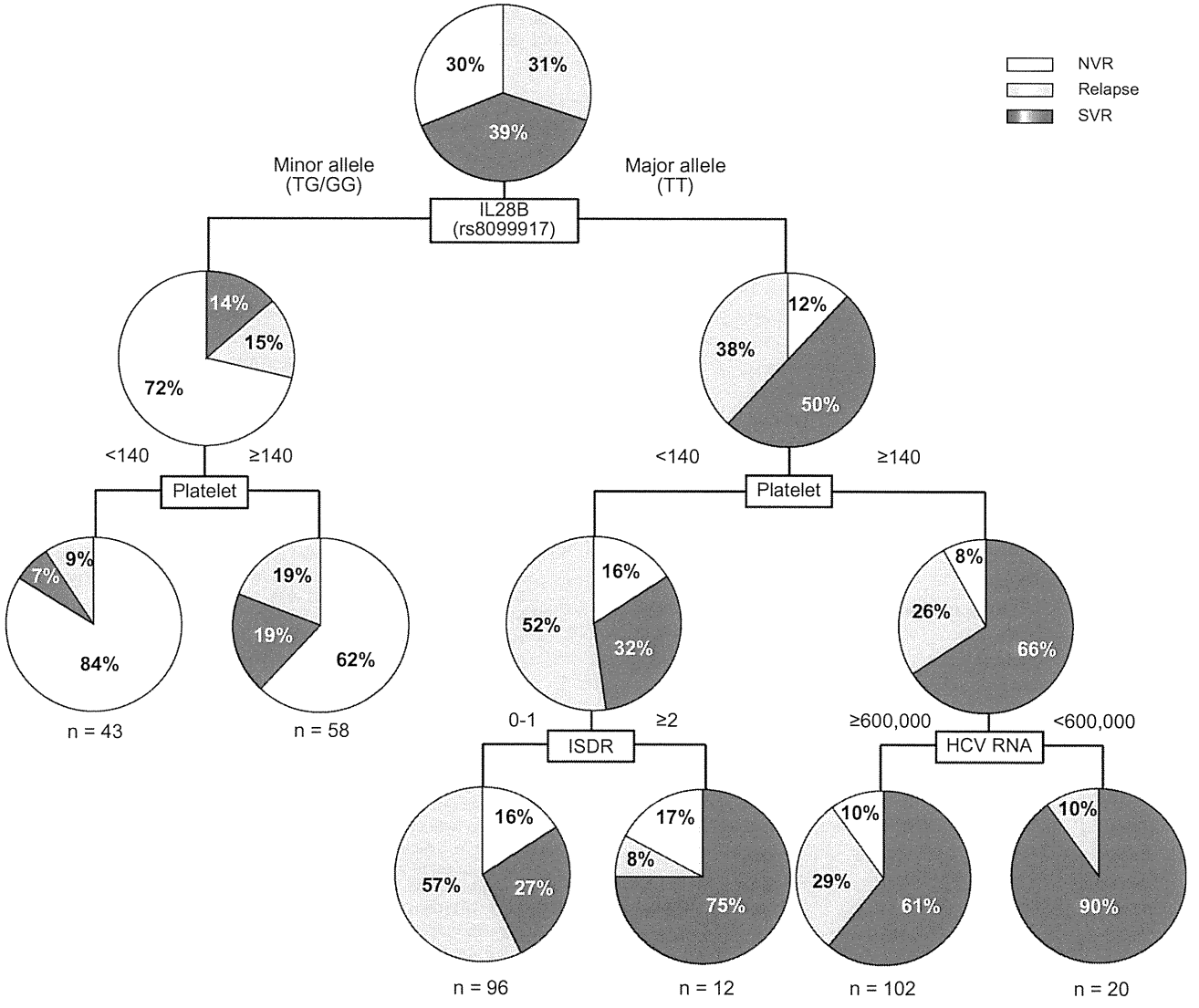

Fig. 5. Decision tree for the prediction of response to therapy. The boxes indicate the factors used for splitting. Pie charts indicate the rate of response for each group of patients after splitting. The rate of null virological response, relapse, and sustained virological response is shown.

estimate of the response before treatment, by simply allocating patients to subgroups by following the flow-chart form, which may facilitate clinical decision making. This is in contrast to the calculating formula, which was constructed by the traditional logistic regression model. This was not widely used in clinical practice as it is abstruse and inconvenient. These results support the evidence based approach of selecting the optimum treatment strategy for individual patients, such as treating patients with a low probability of NVR with current PEG-IFN/RBV combination therapy or advising those with a high probability of NVR to wait for more effective future therapies. Patients with a high probability of relapse may be treated for a longer duration to avoid a relapse. Decisions may be based on the possibility of a response against a potential risk of adverse events and the cost of the therapy, or disease progression while waiting for future therapy.

We have previously reported the predictive model of early virological response to PEG-IFN and RBV in chronic hepatitis C [26]. The top factor selected as significant was the grade of steatosis, followed by serum level of LDL cholesterol, age, GGT, and blood sugar. The mechanism of association between these factors and treatment response was not clear at that time. To our interest, a recent study by Li et al. [47] has shown that high serum level of LDL cholesterol was linked to the *IL28B* major allele (CC in rs12979860). High serum level of LDL cholesterol was associated with SVR but it was no longer significant when analyzed together with the *IL28B* genotype in multivariate analysis. Thus, the association between treatment response and LDL cholesterol levels may reflect the underlining link of LDL cholesterol levels to *IL28B* genotype. Steatosis is reported to be correlated with low lipid levels [48] which suggest that *IL28B* genotypes may be also associated with steatosis. In fact, there were significant correlations between the *IL28B* genotype and the presence of steatosis in the present study (Table 4). In addition, the serum level of GGT, another predictive factor in our previous study, was signif-

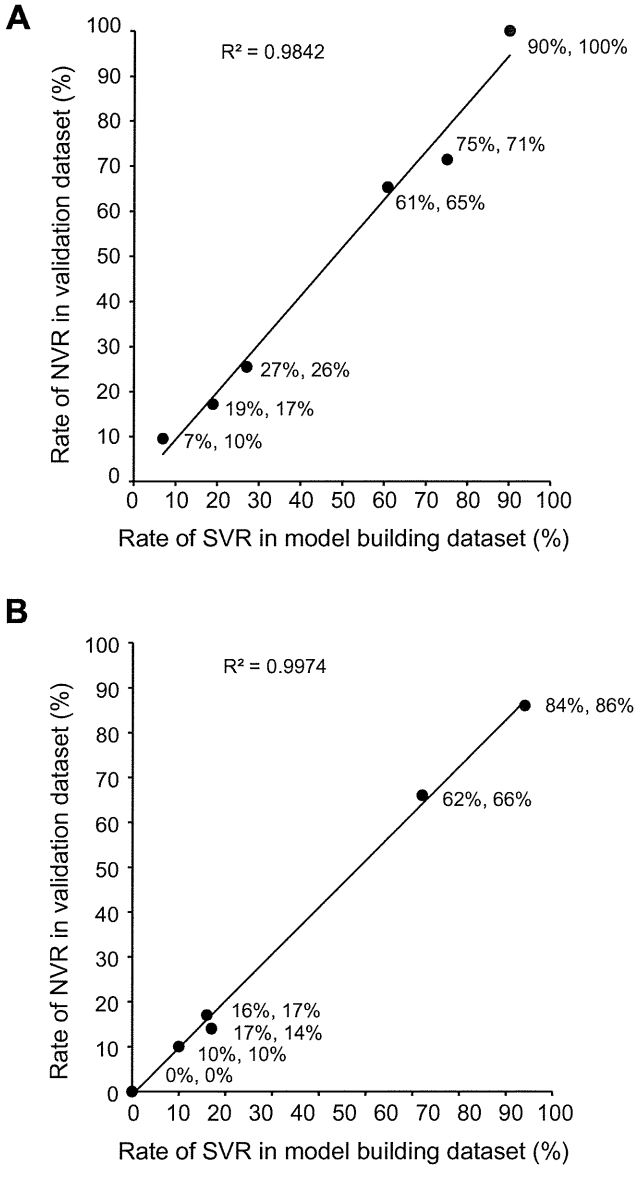

Fig. 6. Validation of the CART analysis. Each patient in the validation group was allocated to one of the six subgroups by following the flow-chart form of the decision tree. The rate of (A) sustained virological response (SVR) and (B) null virological response (NVR) in each subgroup was calculated and plotted. The *X*-axis represents the rate of SVR or NVR in the model building patients and the *Y*-axis represents those in the validation patients. The rate of SVR and NVR in each subgroup of patients is closely correlated between the model building and the validation patients (correlation coefficient: $r^2 = 0.98–0.99$).

icantly associated with *IL28B* genotype in the present study (Table 4). The serum level of GGT was significantly associated with NVR when examined independently but was no longer significant when analyzed together with the *IL28B* genotype. These observations indicate that some of the factors that we have previously identified may be associated with virological response to therapy through the underlining link to the *IL28B* genotype.

In conclusion, the present study highlighted the impact of the *IL28B* polymorphism and mutation in the ISDR on the pre-treatment prediction of response to PEG-IFN/RBV therapy. A decision model including these host and viral factors has the potential to support selection of the optimum treatment strategy for individual patients, which may enable personalized treatment.

Conflict of interest

The authors who have taken part in this study declare that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Financial support

This study was supported by a grant-in-aid from the Ministry of Health, Labor and Welfare, Japan, (H19-kannen-013), (H20-kannen-006).

References

[1] Ray Kim W. Global epidemiology and burden of hepatitis C. Microbes Infect 2002;4 (12):1219–1225.
[2] Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncales Jr FL, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 2002;347 (13):975–982.
[3] Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, et al. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet 2001;358 (9286):958–965.
[4] Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. Intervirology 2005;48 (6):372–380.
[5] Davis GL, Wong JB, McHutchison JG, Manns MP, Harvey J, Albrecht J. Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in patients with chronic hepatitis C. Hepatology 2003;38 (3):645–652.
[6] Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 2009;10:1105–1109.
[7] Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 2009;10:1100–1104.
[8] Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 2009;461 (7262):399–401.
[9] Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. Gastroenterology 2010;138 (4):1338–1345.
[10] Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. J Clin Invest 1995;96 (1):224–230.
[11] Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. N Engl J Med 1996;334 (2):77–81.
[12] Kurosaki M, Enomoto N, Murakami T, Sakuma I, Asahina Y, Yamamoto C, et al. Analysis of genotypes and amino acid residues 2209 to 2248 of the NS5A region of hepatitis C virus in relation to the response to interferon-beta therapy. Hepatology 1997;25 (3):750–753.
[13] Shirakawa H, Matsumoto A, Joshita S, Komatsu M, Tanaka N, Umemura T, et al. Pretreatment prediction of virological response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. Hepatology 2008;48 (6):1753–1760.
[14] Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, et al. Predictive factors of early and sustained responses to peginterferon plus ribavirin combination therapy in Japanese patients infected with hepatitis C virus genotype 1b: amino acid substitutions in the core region and low-density lipoprotein cholesterol levels. J Hepatol 2007;46 (3):403–410.

[15] Okanoue T, Itoh Y, Hashimoto H, Yasui K, Minami M, Takehara T, et al. Predictive values of amino acid sequences of the core and NS5A regions in antiviral therapy for hepatitis C: a Japanese multi-center study. J Gastroenterol 2009;44 (9):952–963.
[16] Segal MR, Bloch DA. A comparison of estimated proportional hazards models and regression trees. Stat Med 1989;8 (5):539–550.
[17] LeBlanc M, Crowley J. A review of tree-based prognostic models. Cancer Treat Res 1995;75:113–124.
[18] Garzotto M, Beer TM, Hudson RG, Peters L, Hsieh YC, Barrera E, et al. Improved detection of prostate cancer using classification and regression tree analysis. J Clin Oncol 2005;23 (19):4322–4329.
[19] Averbook BJ, Fu P, Rao JS, Mansour EG. A long-term analysis of 1018 patients with melanoma by classic Cox regression and tree-structured survival analysis at a major referral center: implications on the future of cancer staging. Surgery 2002;132 (4):589–602.
[20] Leiter U, Buettner PG, Eigentler TK, Garbe C. Prognostic factors of thin cutaneous melanoma: an analysis of the central malignant melanoma registry of the german dermatological society. J Clin Oncol 2004;22 (18):3660–3667.
[21] Valera VA, Walter BA, Yokoyama N, Koyama Y, Iiai T, Okamoto H, et al. Prognostic groups in colorectal carcinoma patients based on tumor cell proliferation and classification and regression tree (CART) survival analysis. Ann Surg Oncol 2007;14 (1):34–40.
[22] Zlobec I, Steele R, Nigam N, Compton CC. A predictive model of rectal tumor response to preoperative radiotherapy using classification and regression tree methods. Clin Cancer Res 2005;11 (15):5440–5443.
[23] Thabane M, Simunovic M, Akhtar-Danesh N, Marshall JK. Development and validation of a risk score for post-infectious irritable bowel syndrome. Am J Gastroenterol 2009;104 (9):2267–2274.
[24] Wu BU, Johannes RS, Sun X, Tabak Y, Conwell DL, Banks PA. The early prediction of mortality in acute pancreatitis: a large population-based study. Gut 2008;57 (12):1698–1703.
[25] Fonarow GC, Adams Jr KF, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. Jama 2005;293 (5):572–580.
[26] Kurosaki M, Matsunaga K, Hirayama I, Tanaka T, Sato M, Yasui Y, et al. A predictive model of response to peginterferon ribavirin in chronic hepatitis C using classification and regression tree analysis. Hepatol Res 2010;40 (3):251–260.
[27] Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K. Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. Anal Biochem 2007;364 (1):78–85.
[28] Hezode C, Forestier N, Dusheiko G, Ferenci P, Pol S, Goeser T, et al. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. N Engl J Med 2009;360 (18):1839–1850.
[29] McHutchison JG, Everson GT, Gordon SC, Jacobson IM, Sulkowski M, Kauffman R, et al. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. N Engl J Med 2009;360 (18):1827–1838.
[30] Rossignol JF, Elfert A, El-Gohary Y, Keeffe EB. Improved virologic response in chronic hepatitis C genotype 4 treated with nitazoxanide, peginterferon, and ribavirin. Gastroenterology 2009;136 (3):856–862.
[31] Marcello T, Grakoui A, Barba-Spaeth G, Machlin ES, Kotenko SV, MacDonald MR, et al. Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. Gastroenterology 2006;131 (6):1887–1898.
[32] Robek MD, Boyd BS, Chisari FV. Lambda interferon inhibits hepatitis B and C virus replication. J Virol 2005;79 (6):3851–3854.
[33] McCarthy JJ, Li JH, Thompson A, Suchindran S, Lao XQ, Patel K, et al. Replicated association between an IL28B Gene Variant and a Sustained Response to Pegylated Interferon and Ribavirin. Gastroenterology 2010;138:2307–2314.
[34] Tanaka Y, Nishida N, Sugiyama M, Tokunaga K, Mizokami M. Λ-interferons and the single nucleotide polymorphisms: a milestone to tailor-made therapy for chronic hepatitis C. Hepatol Res 2010;40:449–460.
[35] Backus LI, Boothroyd DB, Phillips BR, Mole LA. Predictors of response of US veterans to treatment for the hepatitis C virus. Hepatology 2007;46 (1):37–47.
[36] Mori N, Imamura M, Kawakami Y, Saneto H, Kawaoka T, Takaki S, et al. Randomized trial of high-dose interferon-alpha-2b combined with ribavirin in patients with chronic hepatitis C: correlation between amino acid substitutions in the core/NS5A region and virological response to interferon therapy. J Med Virol 2009;81 (4):640–649.
[37] Hung CH, Lee CM, Lu SN, Lee JF, Wang JH, Tung HD, et al. Mutations in the NS5A and E2-PePHD region of hepatitis C virus type 1b and correlation with the response to combination therapy with interferon and ribavirin. J Viral Hepat 2003;10 (2):87–94.
[38] Yen YH, Hung CH, Hu TH, Chen CH, Wu CM, Wang JH, et al. Mutations in the interferon sensitivity-determining region (nonstructural 5A amino acid 2209–2248) in patients with hepatitis C-1b infection and correlating response to combined therapy of pegylated interferon and ribavirin. Aliment Pharmacol Ther 2008;27 (1):72–79.
[39] Zeuzem S, Lee JH, Roth WK. Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon alfa. Hepatology 1997;25 (3):740–744.
[40] Squadrito G, Leone F, Sartori M, Nalpas B, Berthelot P, Raimondo G, et al. Mutations in the nonstructural 5A region of hepatitis C virus and response of chronic hepatitis C to interferon alfa. Gastroenterology 1997;113 (2):567–572.
[41] Sarrazin C, Berg T, Lee JH, Teuber G, Dietrich CF, Roth WK, et al. Improved correlation between multiple mutations within the NS5A region and virological response in European patients chronically infected with hepatitis C virus type 1b undergoing combination therapy. J Hepatol 1999;30 (6):1004–1013.
[42] Murphy MD, Rosen HR, Marousek GI, Chou S. Analysis of sequence configurations of the ISDR, PKR-binding domain, and V3 region as predictors of response to induction interferon-alpha and ribavirin therapy in chronic hepatitis C infection. Dig Dis Sci 2002;47 (6):1195–1205.
[43] Pascu M, Martus P, Hohne M, Wiedenmann B, Hopf U, Schreier E, et al. Sustained virological response in hepatitis C virus type 1b infected patients is predicted by the number of mutations within the NS5A-ISDR: a meta-analysis focused on geographical differences. Gut 2004;53 (9):1345–1351.
[44] Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. Nature 2009;461 (7265):798–801.
[45] Kurosaki M, Enomoto N, Marumo F, Sato C. Evolution and selection of hepatitis C virus variants in patients with chronic hepatitis C. Virology 1994;205 (1):161–169.
[46] Enomoto N, Kurosaki M, Tanaka Y, Marumo F, Sato C. Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis. J Gen Virol 1994;75 (Pt 6):1361–1369.
[47] Li JH, Lao XQ, Tillmann HL, Rowell J, Patel K, Thompson A, et al. Interferon-lambda genotype and low serum low-density lipoprotein cholesterol levels in patients with chronic hepatitis C infection. Hepatology 1904;51 (6):1904–1911.
[48] Serfaty L, Andreani T, Giral P, Carbonell N, Chazouilleres O, Poupon R. Hepatitis C virus induced hypobetalipoproteinemia: a possible mechanism for steatosis in chronic hepatitis C. J Hepatol 2001;34 (3):428–434.

Journal of Gastroenterology and Hepatology

JGHF

doi:10.1111/j.1440-1746.2011.06723.x

MECHANISMS OF GASTROINTESTINAL, PANCREATIC AND LIVER DISEASES

Molecular mechanisms of hepatocarcinogenesis in chronic hepatitis C virus infection

Taro Yamashita, Masao Honda and Shuichi Kaneko

Department of Gastroenterology, Kanazawa University Graduate School of Medical Science, Kanazawa, Ishikawa, Japan

Key words
hepatitis C virus, hepatocellular carcinoma, transcriptome.

Accepted for publication 10 March 2011.

Correspondence
Professor Shuichi Kaneko, Department of Gastroenterology, Kanazawa University Graduate School of Medical Science, 13-1 Takara-Machi, Kanazawa, Ishikawa 920-8641, Japan. Email: skaneko@m-kanazawa.jp

Abstract

Hepatitis C virus (HCV) infection is a major cause of hepatocellular carcinoma (HCC) and chronic liver disease worldwide. Recent developments and advances in HCV replication systems *in vitro* and *in vivo*, transgenic animal models, and gene expression profiling approaches have provided novel insights into the mechanisms of HCV replication. They have also helped elucidate host cellular responses, including activated/inactivated signaling pathways, and the relationship between innate immune responses by HCV infection and host genetic traits. However, the mechanisms of hepatocyte malignant transformation induced by HCV infection are still largely unclear, most likely due to the heterogeneity of molecular paths leading to HCC development in each individual. In this review, we summarize recent advances in knowledge about the mechanisms of hepatocarcinogenesis induced by HCV infection.

Introduction

Hepatocellular carcinoma (HCC) is the fifth most common malignancy and the third leading cause of cancer death worldwide.[1] The majority of HCCs arise from a background of chronic liver diseases caused by infection with hepatitis B virus (HBV) or hepatitis C virus (HCV).[2] Although both viruses are hepatotropic and regarded as causative agents of HCC, the underlying mechanisms of hepatocarcinogenesis are considered to be largely different, partly due to differences in the nature of DNA viruses (with an integration capacity for the host genome) and RNA viruses (with no genome integration capacity).

Hepatitis C virus is an RNA virus that is unable to integrate into the host genome but, instead, its proteins interact with various host proteins and induce host responses that potentially contribute to the malignant transformation of cells. In addition, HCC usually develops in the setting of liver cirrhosis after long-term continuous inflammation/regeneration processes; these accelerate the turnover of hepatocytes with increased risk of replication errors and DNA damage. Furthermore, recent genome-wide association studies have suggested that the natural course of HCV infection might be modified by the genetic background of the host.[3,4] Thus, both host and virus factors are considered to affect the process of hepatocarcinogenesis in a complex manner.

In this review, we summarize the current knowledge of the mechanisms of hepatocarcinogenesis induced by HCV infection. We also focus on recent findings of transcriptomic characteristics of HCV-related HCC and summarize the potential signaling pathways that are altered in this condition.

Epidemiology

Chronic HCV infection is a major risk factor for the development of HCC worldwide. According to the World Health Organization (WHO), approximately 170 million people are chronically infected with HCV. Although epidemiological evidence has suggested a clear, close relationship between HCV infection and HCC,[5,6] the prevalence of HCV infection in HCC patients differs noticeably between geographical regions. Thus, HCV infection is found in 70–80% of HCC patients in Japan, 70% in Egypt, 40–50% in Italy and Spain, about 20% in the United States (US), and less than 10% in China.[7–9] In industrialized countries including the US, a recent increase in HCC incidence and mortality has been observed, potentially due to the rising incidence of HCV infection transmitted through contaminated blood.[10]

Hepatitis C virus increases the risk of HCC by promoting inflammation and fibrosis of the infected liver that eventually results in liver cirrhosis. Once HCV-related cirrhosis is established, HCC develops at an annual rate of about 4–7%.[11] Other factors including alcohol intake, diabetes, and obesity have also been reported to increase the risk of HCC development by about two- to fourfold, indicating a strong life-style effect on the process of hepatocarcinogenesis.[12,13] Age and male gender are also contributing risk factors for HCV-related HCC, although the detailed mechanisms are still debatable.

Virus proteins and host responses

Hepatitis C virus belongs to the Flaviviridae family. It has a positive-stranded linear RNA genome of about 9.6-kb containing a

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

single large open reading frame encoding three structural (core, E1, and E2) and seven non-structural (p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B) proteins.[14] The structural proteins form the HCV virions, whereas non-structural proteins are involved in the processes of viral replication, assembly, and maturation. HCV proteins are known to be processed by host and viral proteases. Both structural and non-structural proteins can interact with various host cellular proteins to potentially promote the malignant transformation of hepatocytes (see recent reviews[7,15,16]). In this review, because of space limitations, we focus on the findings of core and NS5A proteins in terms of host responses potentially evoked during the process of HCV-related hepatocarcinogenesis.

Core protein

Hepatitis C virus core is a 21-kDa nucleocapsid protein with an RNA-binding capacity. In addition to its function in regulating HCV-RNA translation and HCV particle assembly, core protein is known to be involved in mediating the alteration of various host cell signaling pathways, transcriptional activation, modulation of immune responses, apoptosis, oxidative stress, and lipid metabolism.[7] Several recent studies have indicated the statistically significant high frequency of mutations in the *core* gene in HCV-infected patients who developed HCC.[17,18] However, the functional relevance of mutant core proteins on the malignant transformation of hepatocytes or the HCV life cycle has yet to be clarified.

Evidence of core protein as a causative agent of HCC was initially obtained from the transgenic mice model in which *core* gene overexpression, under the regulation of the HBV regulatory element used as a promoter, resulted in steatosis of mouse livers in early life, with subsequent development of adenoma and HCC.[19] However, another mouse model using a different promoter and of a different strain background resulted only in steatosis or different phenotypes without HCC development.[20,21] Similar controversial findings were reported in transgenic mice expressing HCV polyprotein or structural protein with regards to the development of HCC.[22,23] Thus, the role of core protein alone in the development of HCC remains unclear in transgenic mouse models.

Although the direct role of core protein in the malignant transformation of hepatocytes is still under investigation, it seems to be related to the development of hepatic steatosis.[19,24] Indeed, steatosis is one of the risk factors for the development of HCV-related HCC,[25,26] and activation of the lipogenic pathway has been reported in a subset of HCC cases.[27] Core protein is associated with the surface of lipid droplets in infected cells and might be directly related to steatosis through several factors responsible for lipid biogenesis and degradation, including peroxisome proliferator-activated receptor alpha and sterol-regulatory element binding protein-1.[21,28–30] Furthermore, core protein is reported to interact with endoplasmic reticulum (ER) or mitochondrial outer membranes and induce ER stress by perturbation of protein folding or by the accumulation of reactive oxygen species (ROS) through mitochondrial dysfunction.[31,32] ROS produced in this way might result in DNA damage to the host genome and accelerate the process of hepatocarcinogenesis. Increased hepatic iron deposition may also induce oxidative stress and lipid peroxidation, thus increasing the risk of HCC development in HCV polyprotein transgenic mice.[33]

Since the discovery of HCV, various studies have investigated the role of core on host cells. Its effects have been demonstrated on signaling pathways responsible for the cell cycle, and apoptosis through interaction with several tumor suppressors including p53, p73, and p21[34–39] as well as apoptosis regulators such as tumor necrosis factor-α (TNF-α) signaling or Bcl-2 members.[40–42] However, the data obtained from these studies are relatively inconsistent with each other and have varied across experimental models. Core protein may influence the growth and proliferation of host cells through activation of signaling pathways such as Raf/mitogen activated protein kinase (MAPK),[43] Wnt/beta catenin,[16] and transforming growth factor-β (TGF-β).[15,44] These pathways are known to be activated in HCC.[45] The findings therefore indicate a potential role for core in cell proliferation or suppression of apoptosis during malignant transformation of hepatocytes in the liver of chronic hepatitis C, where chronic inflammation and regeneration of hepatocytes continuously occurs.

NS5A protein

NS5A is a 56–58-kDa protein phosphorylated at serine residues by serine-threonine kinase[46] and is essential for replication of the HCV genome. NS5A protein forms part of the viral replicase complex and is localized mainly in the cytoplasm of infected cells in association with the ER. NS5A can become a lower molecular weight protein through post-translational modification, after which it can undergo translocation to the nucleus where it acts as a transcriptional activator. High frequencies of wild-type *NS5A* genes were reported to be dominant in liver cirrhosis patients who finally developed HCC compared with those who did not,[47] but the mechanistic significance of the *NS5A* wild/mutant genotypes in the process of HCV-related hepatocarcinogenesis remains uncertain.

NS5A protein has been suggested to interact with various signaling pathways including cell cycle/apoptosis[48] and lipid metabolism[28,49,50] in host cells and shares some signaling targets with core protein. NS5A is recognized as a transcriptional activator for many target genes[51] including p53 and its binding protein, TATA binding protein (TBP). Transcription factor IID activities were reported to be modified by NS5A in the suppression of p53-dependent transcriptional transactivation and apoptosis.[52,53] NS5A may also interact with pathways such as Bcl2,[54] phosphatidylinositol 3-kinase (PI3-K),[55] Wnt/beta catenin signaling,[56] and mTOR[57] to activate cell proliferation signaling and inhibit apoptosis.

Taken together, intriguing data concerning the function of core and NS5A proteins on host cell signaling pathways, transcriptional activation, apoptosis, oxidative stress, and lipid metabolism described above suggest a diverse role for HCV proteins in the pathophysiology of chronic hepatitis C that leads to malignant transformation in infected hepatocytes. Key findings and present concepts are summarized in Figure 1.

Transcriptomic characteristics of HCV-related HCC

As described above, HCV proteins can evoke various host responses in infected cells at transcriptional/translational/post-translational levels. Furthermore, enhanced cell death/regeneration processes are considered to induce DNA damage and accelerate replication errors that cause frequent mutations and genomic alter-

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

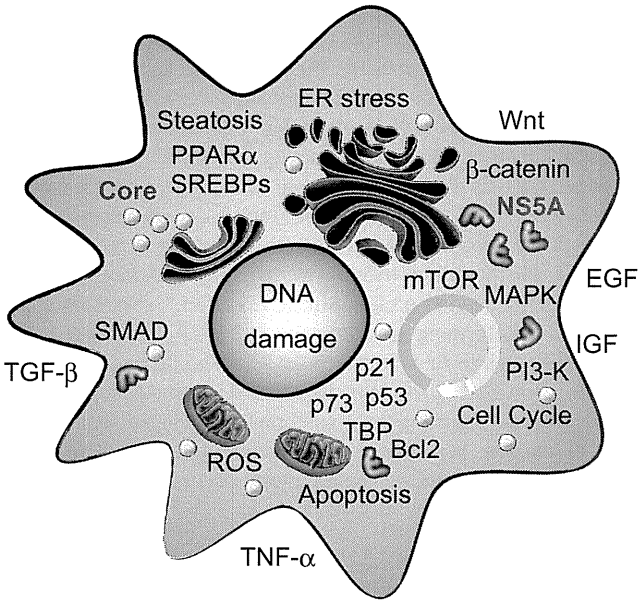

Figure 1 Signaling pathways potentially affected by hepatitis C virus (HCV) proteins. EGF, epidermal growth factor; ER, endoplasmic reticulum; IGF, insulin-like growth factor; MAPK, mitogen activated protein kinase; mTOR, mammalian target of rapamycin; PI3-K, phosphatidylinositol 3-kinase; PPAR, peroxisome proliferator-activated receptor; ROS, reactive oxygen species; SREBP, sterol-regulatory element binding protein; TBP, TATA binding protein.

ation in the host genome. The central dogma is defined as the flow of genetic information from DNA to mRNA and then to protein, so genetic/genomic alterations and transcriptional/translational modifications are ultimately considered to affect the cellular signaling pathway at the transcriptional level.

Over the past decade, several methods (including differential display, serial analysis of gene expression [SAGE], and microarray) have been developed to allow comparative studies of gene expression between normal and cancer cells on a genome-wide scale,[58] and the analysis of a set of all RNA molecules (mainly indicating mRNAs) is termed as whole transcriptome analysis. Extensive transcriptome analysis of HCC and corresponding noncancerous livers has been performed, and the results have greatly increased our knowledge about the transcriptome characteristics of HCV-related HCC.

Early microarray and SAGE studies investigating the gene expression patterns of chronic hepatitis B and C tissues indicated that these two chronic hepatitis tissues had distinct gene expression profiles; the genes activated in chronic hepatitis C were correlated with signaling pathways associated with apoptosis, oxidative stress responses, and Th1 cytokine signaling.[59,60] An early study comparing genes activated in HCV-related and HBV-related HCCs showed that the genes associated with xenobiotic metabolism were more abundantly expressed in HCV-related HCC,[61] suggesting a detoxification role, which is potentially induced by chronic inflammation and generation of ROS resulting from HCV infection. In contrast, HBV-related HCC might closely correlate with the activation of imprint genes, including insulin-like growth factor-II (IGF-II) as investigated by oligo-DNA microarray,[62] suggesting a role of de-differentiation or epigenetic alteration of the host genome in HBV-related HCC. Activation of genes associated with interferon, oxidative stress, apoptosis, and lipid metabolism signaling was detected in HCV-related HCC and chronic hepatitis C specimens,[27,60,63] consistent with numerous functional studies that have investigated the host response evoked by HCV structural and non-structural proteins.[48]

Transcriptome analysis has also recently shed new light on the transcriptional alteration events occurring in early stages of HCV-related hepatocarcinogenesis. *GPC3* (encoding Glypican 3) was identified as one of the most activated transcripts in the early stage of hepatocarcinogenesis,[60,64] while several recent studies showed that gene signatures including *GPC3* can successfully discriminate HCCs from pre-malignant dysplastic nodules and cirrhosis nodules.[65,66] Close examination of genes differentially expressed among cirrhotic nodules, dysplastic nodules, and early and advanced HCV-related HCC tissues has also suggested roles for Toll-like receptor signaling, Wnt signaling, bone morphogenetic protein (BMP)/TGF-β signaling, JAK-STAT signaling, and DNA repair/cell cycle responses in each step of the malignant transformation processes.[67] These processes might therefore provide candidate molecular targets for the chemoprevention of HCV-related HCC.

Recent advances in transcriptome analysis have also provided detailed information on the status of small noncoding RNAs, microRNAs, that can regulate the expression of target genes and viral replication in normal and cancer tissues. Expression of microRNAs including miR-122 and −199a has been reported to modulate HCV replication,[68–70] and miR-122 expression can be regulated by host interferon signaling and responses.[71] HCV protein expression in turn could induce miRNAs and might affect the tumor suppressor DLC1 and the chemosensitivity of malignantly transformed cells.[72,73] Several microRNAs were also differentially expressed between HCV-related and HBV-related HCCs as well as their corresponding non-cancerous liver tissues. The candidate signaling pathways potentially altered by microRNAs in HCV-related tissues were those associated with antigen presentation, cell cycle, and lipid metabolism,[74] consistent with the mRNA microarray data described above. MicroRNAs have also recently been reported to successfully discriminate between HCC and cirrhotic liver tissues,[75] implicating their role in the early stages of malignant transformation. These data suggest that microRNAs may be good targets for the eradication of HCC as well as hepatocytes infected with HCV.

Conclusion

The heterogeneity of genetic/transcriptomic/proteomic events observed in hepatocytes or cell lines expressing HCV proteins and HCV-related HCCs reported thus far has suggested that complex mechanisms underlie malignant transformation induced by HCV infection. These potentially act through convoluted virus-host interactions including HCV replication with host cell cycles, apoptosis, proliferation, quality control of protein synthesis, lipid metabolism, and DNA damage responses. Indeed, HCC is a heterogeneous disease in terms of drug sensitivity, metastatic capacity, and clinical outcome. The heterogeneity of HCV-related HCC may closely correlate with the origin of malignantly transformed cells where multifaceted cellular reactions including apoptosis and

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

cell proliferation are induced by HCV infection. An in-depth understanding of these molecular complexities associated with HCV-related HCC may provide the opportunity for effective chemoprevention of HCC among those with HCV-cirrhosis, and to design tailor-made treatment options for HCV-related HCC patients in the future.

References

1 Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J. Clin.* 2005; **55**: 74–108.
2 Yang JD, Roberts LR. Hepatocellular carcinoma: a global view. *Nat. Rev. Gastroenterol. Hepatol.* 2010; **7**: 448–58.
3 Thomas DL, Thio CL, Martin MP *et al.* Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* 2009; **461**: 798–801.
4 Tillmann HL, Thompson AJ, Patel K *et al.* A polymorphism near IL28B is associated with spontaneous clearance of acute hepatitis C virus and jaundice. *Gastroenterology* 2010; **139**: 1586–92. 92 e1.
5 Bruix J, Barrera JM, Calvet X *et al.* Prevalence of antibodies to hepatitis C virus in Spanish patients with hepatocellular carcinoma and hepatic cirrhosis. *Lancet* 1989; **2**: 1004–6.
6 Colombo M, Kuo G, Choo QL *et al.* Prevalence of antibodies to hepatitis C virus in Italian patients with hepatocellular carcinoma. *Lancet* 1989; **2**: 1006–8.
7 Liang TJ, Heller T. Pathogenesis of hepatitis C-associated hepatocellular carcinoma. *Gastroenterology* 2004; **127**: S62–71.
8 Yoshizawa H. Hepatocellular carcinoma associated with hepatitis C virus infection in Japan: projection to other countries in the foreseeable future. *Oncology* 2002; **62** (Suppl. 1): 8–17.
9 Yuen MF, Hou JL, Chutaputti A. Hepatocellular carcinoma in the Asia pacific region. *J. Gastroenterol. Hepatol.* 2009; **24**: 346–53.
10 El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 2007; **132**: 2557–76.
11 Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular carcinoma in cirrhosis: incidence and risk factors. *Gastroenterology* 2004; **127**: S35–50.
12 Yu MC, Yuan JM. Environmental factors and risk for hepatocellular carcinoma. *Gastroenterology* 2004; **127**: S72–8.
13 Kawaguchi T, Sata M. Importance of hepatitis C virus-associated insulin resistance: therapeutic strategies for insulin sensitization. *World J. Gastroenterol.* 2010; **16**: 1943–52.
14 Penin F, Dubuisson J, Rey FA, Moradpour D, Pawlotsky JM. Structural biology of hepatitis C virus. *Hepatology* 2004; **39**: 5–19.
15 Tsai WL, Chung RT. Viral hepatocarcinogenesis. *Oncogene* 2010; **29**: 2309–24.
16 Levrero M. Viral hepatitis and liver cancer: the case of hepatitis C. *Oncogene* 2006; **25**: 3834–47.
17 Akuta N, Suzuki F, Kawamura Y *et al.* Amino acid substitutions in the hepatitis C virus core region are the important predictor of hepatocarcinogenesis. *Hepatology* 2007; **46**: 1357–64.
18 Fishman SL, Factor SH, Balestrieri C *et al.* Mutations in the hepatitis C virus core gene are associated with advanced liver disease and hepatocellular carcinoma. *Clin. Cancer Res.* 2009; **15**: 3205–13.
19 Moriya K, Fujie H, Shintani Y *et al.* The core protein of hepatitis C virus induces hepatocellular carcinoma in transgenic mice. *Nat. Med.* 1998; **4**: 1065–7.
20 Okuda M, Li K, Beard MR *et al.* Mitochondrial injury, oxidative stress, and antioxidant gene expression are induced by hepatitis C virus core protein. *Gastroenterology* 2002; **122**: 366–75.
21 Perlemuter G, Sabile A, Letteron P *et al.* Hepatitis C virus core protein inhibits microsomal triglyceride transfer protein activity and very low density lipoprotein secretion: a model of viral-related steatosis. *FASEB J.* 2002; **16**: 185–94.
22 Kawamura T, Furusaka A, Koziel MJ *et al.* Transgenic expression of hepatitis C virus structural proteins in the mouse. *Hepatology* 1997; **25**: 1014–21.
23 Wakita T, Taya C, Katsume A *et al.* Efficient conditional transgene expression in hepatitis C virus cDNA transgenic mice mediated by the Cre/loxP system. *J. Biol. Chem.* 1998; **273**: 9001–6.
24 Lerat H, Kammoun HL, Hainault I *et al.* Hepatitis C virus proteins induce lipogenesis and defective triglyceride secretion in transgenic mice. *J. Biol. Chem.* 2009; **284**: 33466–74.
25 Ohata K, Hamasaki K, Toriyama K *et al.* Hepatic steatosis is a risk factor for hepatocellular carcinoma in patients with chronic hepatitis C virus infection. *Cancer* 2003; **97**: 3036–43.
26 Kurosaki M, Hosokawa T, Matsunaga K *et al.* Hepatic steatosis in chronic hepatitis C is a significant risk factor for developing hepatocellular carcinoma independent of age, sex, obesity, fibrosis stage and response to interferon therapy. *Hepatol. Res.* 2010; **40**: 870–7.
27 Yamashita T, Honda M, Takatori H *et al.* Activation of lipogenic pathway correlates with cell proliferation and poor prognosis in hepatocellular carcinoma. *J. Hepatol.* 2009; **50**: 100–10.
28 Dharancy S, Malapel M, Perlemuter G *et al.* Impaired expression of the peroxisome proliferator-activated receptor alpha during hepatitis C virus infection. *Gastroenterology* 2005; **128**: 334–42.
29 Waris G, Felmlee DJ, Negro F, Siddiqui A. Hepatitis C virus induces proteolytic cleavage of sterol regulatory element binding proteins and stimulates their phosphorylation via oxidative stress. *J. Virol.* 2007; **81**: 8122–30.
30 Tanaka N, Moriya K, Kiyosawa K, Koike K, Gonzalez FJ, Aoyama T. PPARalpha activation is essential for HCV core protein-induced hepatic steatosis and hepatocellular carcinoma in mice. *J. Clin. Invest.* 2008; **118**: 683–94.
31 Korenaga M, Wang T, Li Y *et al.* Hepatitis C virus core protein inhibits mitochondrial electron transport and increases reactive oxygen species (ROS) production. *J. Biol. Chem.* 2005; **280**: 37481–8.
32 Li Y, Boehning DF, Qian T, Popov VL, Weinman SA. Hepatitis C virus core protein increases mitochondrial ROS production by stimulation of Ca^{2+} uniporter activity. *FASEB J.* 2007; **21**: 2474–85.
33 Furutani T, Hino K, Okuda M *et al.* Hepatic iron overload induces hepatocellular carcinoma in transgenic mice expressing the hepatitis C virus polyprotein. *Gastroenterology* 2006; **130**: 2087–98.
34 Alisi A, Giambartolomei S, Cupelli F *et al.* Physical and functional interaction between HCV core protein and the different p73 isoforms. *Oncogene* 2003; **22**: 2573–80.
35 Honda M, Kaneko S, Shimazaki T *et al.* Hepatitis C virus core protein induces apoptosis and impairs cell-cycle regulation in stably transformed Chinese hamster ovary cells. *Hepatology* 2000; **31**: 1351–9.
36 Kao CF, Chen SY, Chen JY, Lee YHW. Modulation of p53 transcription regulatory activity and post-translational modification by hepatitis C virus core protein. *Oncogene* 2004; **23**: 2472–83.
37 Otsuka M, Kato N, Lan K *et al.* Hepatitis C virus core protein enhances p53 function through augmentation of DNA binding affinity and transcriptional ability. *J. Biol. Chem.* 2000; **275**: 34122–30.
38 Ray RB, Steele R, Meyer K, Ray R. Hepatitis C virus core protein represses p21WAF1/Cip1/Sid1 promoter activity. *Gene* 1998; **208**: 331–6.
39 Yamanaka T, Kodama T, Doi T. Subcellular localization of HCV core protein regulates its ability for p53 activation and p21 suppression. *Biochem. Biophys. Res. Commun.* 2002; **294**: 528–34.

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

40 Lee SK, Park SO, Joe CO, Kim YS. Interaction of HCV core protein with 14-3-3epsilon protein releases Bax to activate apoptosis. *Biochem. Biophys. Res. Commun.* 2007; **352**: 756–62.
41 Mohd-Ismail NK, Deng L, Sukumaran SK, Yu VC, Hotta H, Tan YJ. The hepatitis C virus core protein contains a BH3 domain that regulates apoptosis through specific interaction with human Mcl-1. *J. Virol.* 2009; **83**: 9993–10006.
42 Saito K, Meyer K, Warner R, Basu A, Ray RB, Ray R. Hepatitis C virus core protein inhibits tumor necrosis factor alpha-mediated apoptosis by a protective effect involving cellular FLICE inhibitory protein. *J. Virol.* 2006; **80**: 4372–9.
43 Tsutsumi T, Suzuki T, Moriya K *et al.* Hepatitis C virus core protein activates ERK and p38 MAPK in cooperation with ethanol in transgenic mice. *Hepatology* 2003; **38**: 820–8.
44 Matsuzaki K, Murata M, Yoshida K *et al.* Chronic inflammation associated with hepatitis C virus infection perturbs hepatic transforming growth factor beta signaling, promoting cirrhosis and hepatocellular carcinoma. *Hepatology* 2007; **46**: 48–57.
45 Wang XW, Hussain SP, Huo TI *et al.* Molecular pathogenesis of human hepatocellular carcinoma. *Toxicology* 2002; **181–182**: 43–7.
46 Tanji Y, Kaneko T, Satoh S, Shimotohno K. Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. *J. Virol.* 1995; **69**: 3980–6.
47 De Mitri MS, Cassini R, Bagaglio S *et al.* Evolution of hepatitis C virus non-structural 5A gene in the progression of liver disease to hepatocellular carcinoma. *Liver Int.* 2007; **27**: 1126–33.
48 Kasprzak A, Adamek A. Role of hepatitis C virus proteins (C, NS3, NS5A) in hepatic oncogenesis. *Hepatol. Res.* 2008; **38**: 1–26.
49 Benga WJ, Krieger SE, Dimitrova M *et al.* Apolipoprotein E interacts with hepatitis C virus nonstructural protein 5A and determines assembly of infectious particles. *Hepatology* 2010; **51**: 43–53.
50 Kim K, Kim KH, Ha E, Park JY, Sakamoto N, Cheong J. Hepatitis C virus NS5A protein increases hepatic lipid accumulation via induction of activation and expression of PPARgamma. *FEBS Lett.* 2009; **583**: 2720–6.
51 Kato N, Lan KH, Ono-Nita SK, Shiratori Y, Omata M. Hepatitis C virus nonstructural region 5A protein is a potent transcriptional activator. *J. Virol.* 1997; **71**: 8856–9.
52 Lan KH, Sheu ML, Hwang SJ *et al.* HCV NS5A interacts with p53 and inhibits p53-mediated apoptosis. *Oncogene* 2002; **21**: 4801–11.
53 Majumder M, Ghosh AK, Steele R, Ray R, Ray RB. Hepatitis C virus NS5A physically associates with p53 and regulates p21/waf1 gene expression in a p53-dependent manner. *J. Virol.* 2001; **75**: 1401–7.
54 Chung YL, Sheu ML, Yen SH. Hepatitis C virus NS5A as a potential viral Bcl-2 homologue interacts with Bax and inhibits apoptosis in hepatocellular carcinoma. *Int. J. Cancer* 2003; **107**: 65–73.
55 He Y, Nakao H, Tan SL *et al.* Subversion of cell signaling pathways by hepatitis C virus nonstructural 5A protein via interaction with Grb2 and P85 phosphatidylinositol 3-kinase. *J. Virol.* 2002; **76**: 9207–17.
56 Park CY, Choi SH, Kang SM *et al.* Nonstructural 5A protein activates beta-catenin signaling cascades: implication of hepatitis C virus-induced liver pathogenesis. *J. Hepatol.* 2009; **51**: 853–64.
57 Peng L, Liang D, Tong W, Li J, Yuan Z. Hepatitis C virus NS5A activates the mammalian target of rapamycin (mTOR) pathway, contributing to cell survival by disrupting the interaction between FK506-binding protein 38 (FKBP38) and mTOR. *J. Biol. Chem.* 2010; **285**: 20870–81.
58 Yamashita T, Honda M, Kaneko S. Application of serial analysis of gene expression in cancer research. *Curr. Pharm. Biotechnol.* 2008; **9**: 375–82.
59 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.
60 Yamashita T, Kaneko S, Hashimoto S *et al.* Serial analysis of gene expression in chronic hepatitis C and hepatocellular carcinoma. *Biochem. Biophys. Res. Commun.* 2001; **282**: 647–54.
61 Okabe H, Satoh S, Kato T *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.
62 Iizuka N, Oka M, Yamada-Okabe H *et al.* Comparison of gene expression profiles between hepatitis B virus- and hepatitis C virus-infected hepatocellular carcinoma by oligonucleotide microarray data on the basis of a supervised learning method. *Cancer Res.* 2002; **62**: 3939–44.
63 Honda M, Yamashita T, Ueda T, Takatori H, Nishino R, Kaneko S. Different signaling pathways in the livers of patients with chronic hepatitis B or chronic hepatitis C. *Hepatology* 2006; **44**: 1122–38.
64 Capurro M, Wanless IR, Sherman M *et al.* Glypican-3: a novel serum and histochemical marker for hepatocellular carcinoma. *Gastroenterology* 2003; **125**: 89–97.
65 Jia HL, Ye QH, Qin LX *et al.* Gene expression profiling reveals potential biomarkers of human hepatocellular carcinoma. *Clin. Cancer Res.* 2007; **13**: 1133–9.
66 Llovet JM, Chen Y, Wurmbach E *et al.* A molecular signature to discriminate dysplastic nodules from early hepatocellular carcinoma in HCV cirrhosis. *Gastroenterology* 2006; **131**: 1758–67.
67 Wurmbach E, Chen YB, Khitrov G *et al.* Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma. *Hepatology* 2007; **45**: 938–47.
68 Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* 2005; **309**: 1577–81.
69 Murakami Y, Aly HH, Tajima A, Inoue I, Shimotohno K. Regulation of the hepatitis C virus genome replication by miR-199a. *J. Hepatol.* 2009; **50**: 453–60.
70 Sarasin-Filipowicz M, Krol J, Markiewicz I, Heim MH, Filipowicz W. Decreased levels of microRNA miR-122 in individuals with hepatitis C responding poorly to interferon therapy. *Nat. Med.* 2009; **15**: 31–3.
71 Pedersen IM, Cheng G, Wieland S *et al.* Interferon modulation of cellular microRNAs as an antiviral mechanism. *Nature* 2007; **449**: 919–22.
72 Braconi C, Valeri N, Gasparini P *et al.* Hepatitis C virus proteins modulate microRNA expression and chemosensitivity in malignant hepatocytes. *Clin. Cancer Res.* 2010; **16**: 957–66.
73 Banaudha K, Kaliszewski M, Korolnek T *et al.* MicroRNA silencing of tumor suppressor DLC-1 promotes efficient hepatitis C virus replication in primary human hepatocytes. *Hepatology* 2011; **53**: 53–61.
74 Ura S, Honda M, Yamashita T *et al.* Differential microRNA expression between hepatitis B and hepatitis C leading disease progression to hepatocellular carcinoma. *Hepatology* 2009; **49**: 1098–112.
75 Wong QW, Lung RW, Law PT *et al.* MicroRNA-223 is commonly repressed in hepatocellular carcinoma and potentiates expression of Stathmin1. *Gastroenterology* 2008; **135**: 257–69.

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

uicc
global cancer control

IJC
International Journal of Cancer

Identification of a secretory protein c19orf10 activated in hepatocellular carcinoma

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

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

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

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

Currently, imaging diagnostic techniques such as ultrasonography, computed tomography, magnetic resonance image and angiography are the gold standards for the early detection of HCC.[13,14] In addition, tumor markers such as alpha-fetoprotein (AFP) and des-gamma carboxyl prothrombin (DCP) have been used for the screening of HCC,[15–18] although their sensitivity and specificity are not sufficiently high. Recently, a gene expression profiling approach shed new light on Glypican 3, a heparin sulfate proteoglycan anchored to the plasma membrane, as a potential HCC marker, and its clinical usefulness as a molecular target as well as a tumor marker is presently under investigation.[19]

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

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

Key words: hepatocellular carcinoma, serial analysis of gene expression, *c19orf10*
Additional Supporting Information may be found in the online version of this article
DOI: 10.1002/ijc.25830
History: Received 14 Mar 2010; Accepted 15 Nov 2010; Online 2 Dec 2010
Correspondence to: Shuichi Kaneko, Kanazawa University Hospital, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan,
Tel.: +81-76-265-2233, Fax: +81-76-234-4250,
E-mail: skaneko@m-kanazawa.jp

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

Material and Methods

SAGE and HCC samples

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

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

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

Laser capture microdissection and RNA isolation

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

Construction of C19ORF10 expression plasmid and recombinant adenovirus vector

PCR was performed on a Marathon cDNA library from Huh7 cells using the following primers: sense primers: 5′-GACCCTAGTCCAACATGGCGGCGCCC-3′ (the first PCR), 5′-ATGGCGGCGCCCAGCGGAGGGTGGAACGGC-3′ (the nested second PCR) and antisense primers: 5′-CACCGGA GATGAGAAGGTGCCACCCGC-3′ (the first PCR), 5′-CAG GGCTGCTGGTCACAGCTCAGTGCGCG-3′ (the nested second PCR). The 5′ and 3′ends of the cDNA were isolated using a SMART RACE cDNA Amplification kit (Clontech, Mountain View, CA) according to the manufacturer's recommendations. The PCR products were cloned into a TA vector (Invitrogen, Carlsbad, CA) to generate the pcDNA3.1-*c19orf10* expression plasmid. Using this plasmid, a C-terminally FLAG-tagged construct of *c19orf10* was generated and inserted in a pSI mammalian expression vector (Promega, Madison, WI), which was driven by the SV40 promoter (pSI-*c19orf10*).

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

RTD-PCR

RTD-PCR was performed as previously described.[31] Briefly, template cDNA was synthesized from 1 μg of total RNA using SuperScript™ II RT (Invitrogen). RTD-PCR of *c19orf10* (Hs. 00384077_m1), *AFP* (Hs00173490_m1), *GPC3* (Hs01018938_m1), *KRT19* (Hs00761767_s1) and the *ACTB* internal control (Hs99999903_m1) was performed using a TaqMan® Gene Expression Assay kit (Applied Biosystems, Foster City, CA). The expression of selected genes was measured in triplicate by ΔΔCT method using the 7900 Sequence Detection System (Applied Biosystems).

Cell lines and transfection of plasmids

Human liver cancer cell lines HuH1, Huh7, Hep3B, HLE and HLF as well as HEK293 and NIH3T3 were cultured in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% heat-inactivated fetal bovine serum (Invitrogen) in 5% CO_2 at 37°C. Transfection of plasmids was performed using FuGENE™ 6 (Roche Diagnostics, Indianapolis, IN) according to the manufacturer's instruction. Briefly, 5×10^5 cells were seeded in a six-well plate 12 hr before transfection, and 3 μg of plasmid DNA was used for each transfection. All experiments were repeated at least twice.

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

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

Table 1. ESTs overexpressed in the HCC library

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

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

Silencing gene expression by short interfering RNA

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

Cell proliferation assay

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

Cell cycle analysis

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

Western blotting

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

Statistical analyses

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

Results

Identification of *C19ORF10* overexpression in HCC by SAGE

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

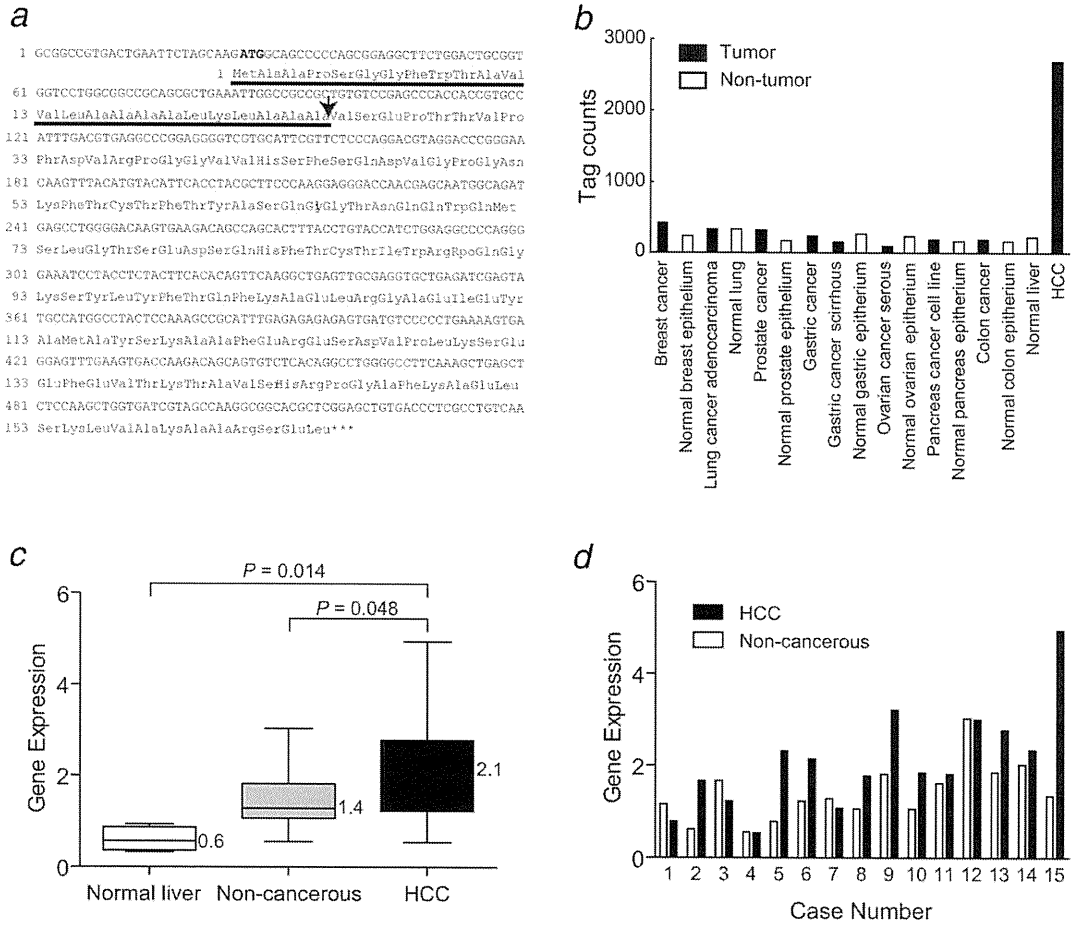

Figure 1. (*a*) Structure of a *c19orf10* gene and a c19orf10 protein. The DNA sequence of *c19orf10* and amino acid alignment of the encoded c19orf10 protein are shown. C19orf10 is predicted to have a molecular weight of 17 kDa and contain a signal peptide cleavage site (indicated as a black arrow).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

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

When we examined the expression profiles of *c19orf10* using retrieved SAGE data from various cancers and their normal counterparts, we identified that *c19orf10* was abundantly expressed in human HCC (Fig. 1*b*). We further examined the publicly available EST profiles of *c19orf10* (http://www.ncbi.nlm.nih.gov/unigene) and confirmed its tendency to be overexpressed in HCC compared to the normal liver (data not shown). We validated the overexpression of *c19orf10* in 15 independent HCC tissues and adjacent noncancerous liver tissues by RTD-PCR. Gene expression of *c19orf10* was significantly higher in the HCC tissues than in the normal liver tissues and adjacent noncancerous liver tissues ($p = 0.014$ and 0.048, respectively; Fig. 1*c*). *C19orf10* expression was elevated in HCC tissues compared to the adjacent noncancerous liver tissues in 10 of 15 patients (66.7%; Fig. 1*d*).

Overexpression of *C19ORF10* in AFP-positive HCC

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

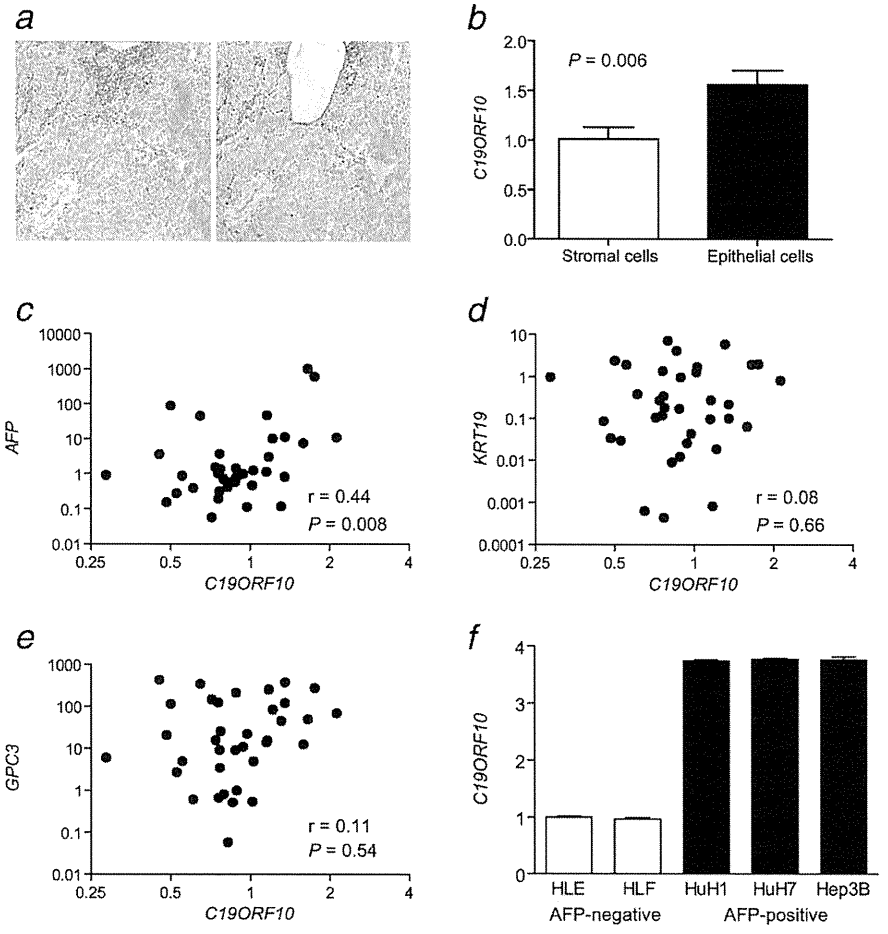

Figure 2. (*a*) Representative photomicrographs of an HCC tissue used for LCM (toluidine blue staining). Inflammatory mononuclear cells and stromal cells were separately captured (left: Pre-LCM, right: Post-LCM). (*b*) RTD-PCR analysis of *c19orf10* expression in inflammatory mononuclear cells and tumor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

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

To explore the relationship of *c19orf10* with other established HCC markers, we investigated the gene expression of *c19orf10*, *AFP* (alpha-fetoprotein), *KRT19* (cytokeratin 19) and *GPC3* (glypican 3). Because only 1 of 15 HCC tissues analyzed above (Fig. 1*d*) was AFP positive (data not shown), we further investigated the expression of *c19orf10* in an additional 36 HCC tissues using RTD-PCR. Interestingly, *c19orf10* expression was significantly positively correlated with *AFP* ($r = 0.44$, $p = 0.008$), but not with *KRT19* ($r = 0.08$, $p = 0.66$) nor *GPC3* ($r = 0.11$, $p = 0.54$) (Figs. 2*c*–2*e*). Furthermore, when we examined the expression of *c19orf10* in AFP-positive (HuH1, HuH7 and Hep3B) and -negative (HLE and HLF) HCC cell lines, we identified the overexpression of *c19orf10* in AFP-positive HCC cell lines (Fig. 2*f*). These data suggested that *c19orf10* is overexpressed and may play some role in AFP-positive HCCs.

C19orf10 regulates MAPK/Akt pathways and activates cell proliferation

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

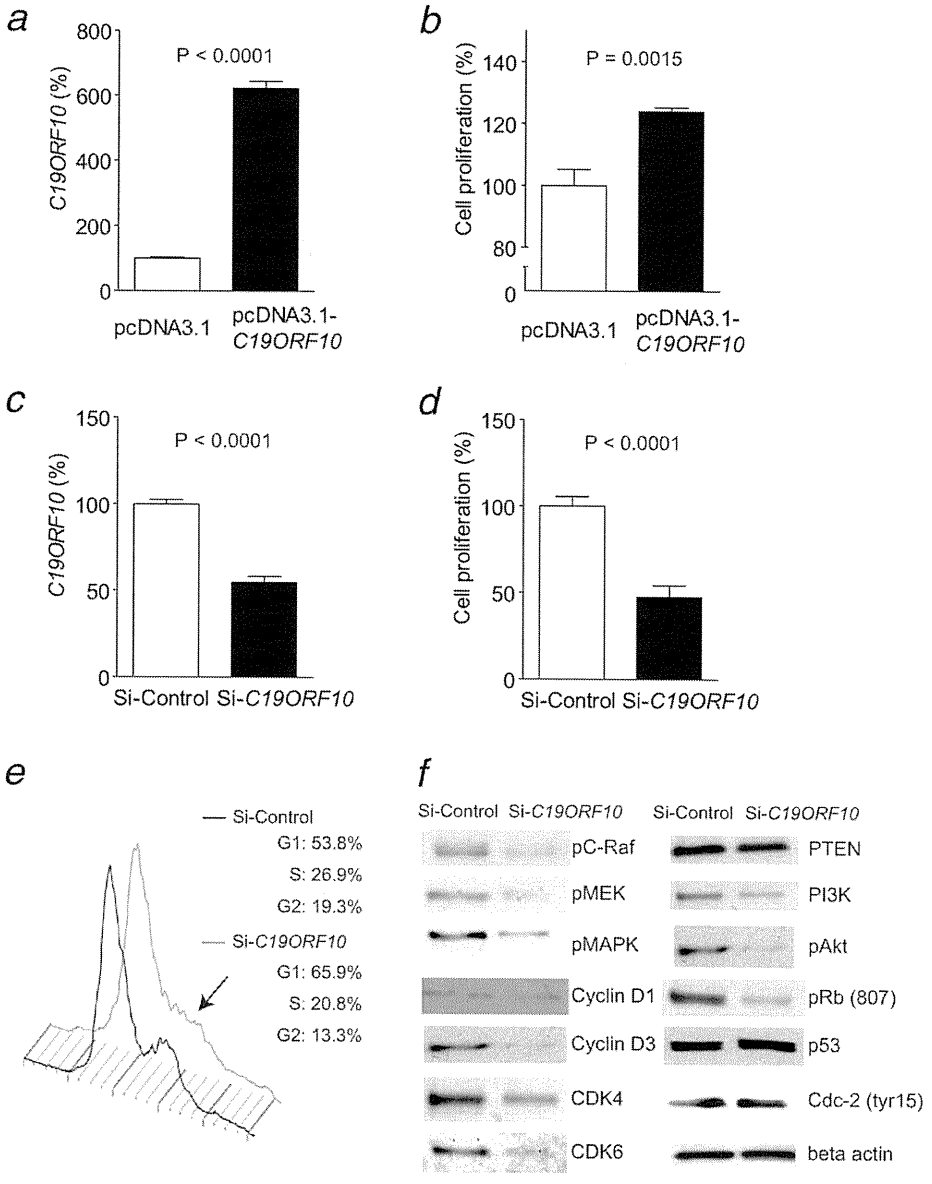

Figure 3. (*a*) RTD-PCR analysis of *c19orf10* expression in HLE cells transfected with pcDNA3.1 or pcDNA3.1-*c19orf10* plasmids. (*b*) Cell proliferation assay of HLE cells transfected with pcDNA3.1 or pcDNA3.1-*c19orf10* plasmids. Cell proliferation was evaluated 72 hr after each plasmid transfection. (*c*) RTD-PCR analysis of *c19orf10* expression in Hep3B cells transfected with Si-*Control* or Si-*c19orf10*. Gene expression was measured in triplicates 48 hr after transfection. (*d*) Cell proliferation assay of Hep3B cells transfected with Si-*Control* or Si-*c19orf10*. Cell proliferation was evaluated 72 hr after siRNA transfection. (*e*) Cell cycle analysis of HuH7 cells transfected with Si-*Control* or Si-*c19orf10*. Cell cycle was evaluated 72 hr after siRNA transfection. A black arrow indicates the G2 phase peak. (*f*) Western blotting analysis of Huh7 cells transfected with Si-*Control* or Si-*c19orf10*. Cells were lysed by RIPA buffer 72 hr after siRNA transfection.

pcDNA3.1 or pcDNA3.1-*c19orf10* plasmids, we identified an approximately sixfold overexpression of *c19orf10* when compared to the control 48 hr after transfection ($p < 0.0001$) (Fig. 3*a*). Interestingly, cell proliferation was modestly, but significantly, enhanced compared to the control 72 hr after transfection ($p = 0.0015$) (Fig. 3*b*).

We also transfected siRNAs targeting an irrelevant sequence (Si-*Control*) or *c19orf10* (Si-*c19orf10*) in Hep3B and