201240010A（上巻）

厚生労働科学研究費補助金
難病・がん等の疾患分野の医療の実用化研究事業
（肝炎関係分野）

慢性肝炎・肝硬変・肝癌の病態解明と各病態および都市形態別で求められる医療を考慮したクリティカルパスモデルの開発のための研究

平成24年度　総括研究報告書　上巻

研究代表者　泉　並木

平成25（2013）年　３月

【主任研究者】

【分担研究者】

泉　並木

武蔵野赤十字病院　副院長・消化器科部長

Hepatology Research 2010; 40: 251–260 doi: 10.1111/j.1872-034X.2009.00607.x

Original Article

A predictive model of response to peginterferon ribavirin in chronic hepatitis C using classification and regression tree analysis

Masayuki Kurosaki,[1] Kotaro Matsunaga,[2] Itsuko Hirayama,[1] Tomohiro Tanaka,[1] Mitsuaki Sato,[1] Yutaka Yasui,[1] Nobuharu Tamaki,[1] Takanori Hosokawa,[1] Ken Ueda,[1] Kaoru Tsuchiya,[1] Hiroyuki Nakanishi,[1] Hiroki Ikeda,[1] Jun Itakura,[1] Yuka Takahashi,[1] Yasuhiro Asahina,[1] Megumu Higaki,[4] Nobuyuki Enomoto[3] and Namiki Izumi[1]

[1]*Division of Gastroenterology and Hepatology and* [2]*Division of Pathology, Musashino Red Cross Hospital, Tokyo,* [3]*First Department of Internal Medicine, University of Yamanashi, Yamanashi, and* [4]*Department of Medical Science, Jikei Medical University, Tokyo, Japan*

Aim: Early disappearance of serum hepatitis C virus (HCV) RNA is the prerequisite for achieving sustained virological response (SVR) in peg-interferon (PEG-IFN) plus ribavirin (RBV) therapy for chronic hepatitis C. This study aimed to develop a decision tree model for the pre-treatment prediction of response.

Methods: Genotype 1b chronic hepatitis C treated with PEG-IFN alpha-2b and RBV were studied. Predictive factors of rapid or complete early virological response (RVR/cEVR) were explored in 400 consecutive patients using a recursive partitioning analysis, referred to as classification and regression tree (CART) and validated.

Results: CART analysis identified hepatic steatosis (<30%) as the first predictor of response followed by low-density-lipoprotein cholesterol (LDL-C) (≥100 mg/dL), age (<50 and <60 years), blood sugar (<120 mg/dL), and gamma-glutamyltransferase (GGT) (<40 IU/L) and built decision tree model. The model consisted of seven groups with variable response rates from low (15%) to high (77%). The reproducibility of the model was confirmed by the independent validation group ($r^2 = 0.987$). When reconstructed into three groups, the rate of RVR/cEVR was 16% for low probability group, 46% for intermediate probability group and 75% for high probability group.

Conclusions: A decision tree model that includes hepatic steatosis, LDL-C, age, blood sugar, and GGT may be useful for the prediction of response before PEG-IFN plus RBV therapy, and has the potential to support clinical decisions in selecting patients for therapy and may provide a rationale for treating metabolic factors to improve the efficacy of antiviral therapy.

Key words: data mining, decision tree, HCV, low-density-lipoprotein-cholesterol, steatosis

INTRODUCTION

COMBINATION THERAPY WITH pegylated interferon (PEG-IFN) and ribavirin (RBV) is now recognized as a standard treatment for patients with chronic hepatitis C.[1] However, the rate of sustained virological response (SVR) to 48 weeks of PEG-IFN RBV combination therapy is only 50% in patients with hepatitis C virus (HCV) genotype 1b and high HCV RNA titer, so called difficult to treat chronic hepatitis C patients.[2,3] Within this difficult to treat group, the response to treatment sometimes can be highly heterogeneous for cases which are apparently equivalent in HCV RNA titer, making the prediction of response before treatment a difficult task. It has been suggested that early virological response (EVR), defined as either undetectable HCV RNA or a 2 log drop in HCV RNA at week 12, is a reliable means to predict SVR.[2,4] More recently, it has been suggested that patients with a rapid virological response (RVR: undetectable HCV RNA at week 4) and a complete EVR (cEVR: undetectable HCV RNA at week 12)

Correspondence: Dr Namiki Izumi, Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, 1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610, Japan. Email: nizumi@musashino.jrc.or.jp
Received 26 May 2009; revision 25 August 2009; accepted 26 August 2009.

achieve high SVR rates, while patients with a partial EVR (pEVR: 2 log drop in HCV RNA but still detectable at week 12) have lower rates of SVR.[5] Since PEG-IFN RBV combination therapy is costly and accompanied by potential adverse effects, the ability to predict the possibility of RVR or cEVR before therapy and identifying curable patients may significantly influence the selection of patients for therapy. Moreover, identification of baseline predictors of poor response is particularly important to establish a rationale for identifying therapeutic targets to improve the efficacy of antiviral therapy.

Data mining is a method of predictive analysis which explores tremendous volumes of data to discover hidden patterns and relationships in highly complex datasets and enables the development of predictive models. The classification and regression tree (CART) analysis is a core component of the decision tree tool for data mining and predictive modeling,[6] is deployed to decision makers in various fields of business, and currently is being used in the area of biomedicine.[7–13] The results of CART analysis are presented as a decision tree, which is intuitive and facilitates the allocation of patients into subgroups by following the flow-chart form.[14] CART has been shown to be competitive with other traditional statistical techniques such as logistic regression analysis.[15]

In the present study, we used the CART analysis to explore baseline predictors of response to PEG-IFN plus RBV therapy among clinical, biochemical, virological and histological pretreatment variables and to define a pre-treatment algorithm to discriminate chronic hepatitis C patients who are likely to respond to PEG-IFN plus RBV therapy.

MATERIALS AND METHODS

Patients

A TOTAL OF 419 chronic hepatitis C patients were treated with PEG-IFN alpha-2b and RBV at Musashino Red Cross Hospital between December 2001 and December 2007. Among them, 400 patients who fulfilled the following inclusion criteria were enrolled in the present study. (i) infection by genotype 1b (ii) HCV RNA higher than 100 KIU/mL by quantitative PCR (Cobas Amplicor HCV Monitor, Roche Diagnostic systems, CA) which is usually used for the definition of high viral load in Japan (iii) lack of co-infection with hepatitis B virus or human immunodeficiency virus (iv) lack of other causes of liver disease such as autoimmune hepatitis, primary biliary cirrhosis, or alcohol intake of more than 20 g per day, and (v) having completed at least 12 weeks of therapy with an early virological response that could be evaluated. Patients received PEG-IFN alpha-2b (1.5 microgram/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) which is the recommended dosage in Japan. Data from two third of patients (269 patients) were used for the model building set and the remaining one third of patients (131 patients) were used as a validation set. Consent in writing was obtained from each patient and the study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the institutional review committee.

Laboratory tests

Blood samples were obtained before therapy, and at least once every month during therapy and analyzed for hematologic tests, blood chemistries, and HCV RNA. In the present study, RVR and cEVR was defined as undetectable HCV RNA by qualitative PCR with a lower detection limit of 50 IU/mL (Amplicor, Roche Diagnostic systems, CA) at week 4 and 12, respectively. SVR was defined as undetectable HCV RNA at week 24 after the completion of therapy.

Histological examination

For all patients, liver biopsy specimens were obtained before therapy and were evaluated independently by three pathologists who were blinded to the clinical details. If there was a disagreement, the scores assigned by the majority of pathologists were used for the analysis. Fibrosis and activity were scored according to the METAVIR scoring system.[16] Fibrosis was staged on a scale of 0–4: 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). Activity of necroinflammation was graded on a scale of 0–3: A0 (no activity), A1 (mild activity), A2 (moderate activity) and A3 (severe activity). Percentage of steatosis was quantified by determining the average proportion of hepatocytes affected by steatosis and graded on a scale of 0–3: grade 0 (no steatosis), grade 1 (0–9%), grade 2 (10–29%), and grade 3 (over 30%) as we reported previously.[17]

Database for analysis

A pretreatment database of 72 variables was created containing histological findings (grade of fibrosis, activity, and steatosis), laboratory tests including the quantity of HCV RNA by Cobas Amplicor, and clinical information (age, gender, body weight, and body mass index).

The baseline characteristics and test results are listed in Table 1. The overall rate of RVR/cEVR was 43% in the model building set and 48% in the validation set. There were no significant differences in the clinical backgrounds between these two groups. Hepatitis C viral mutations, such as mutations in interferon-sensitivity determining region or core amino acid residues 70 and 91, were not included in the present analysis. The dataset of laboratory tests was based on the digitized records in this hospital. Continuous data was split into categorized data by increment of 10; For example, age was categorized into <30, 30–39, 40–49, 50–59, 60–69, and ≥70.

Statistical analysis

Based on this database, the recursive partitioning analysis algorithm referred to as CART was implemented to define meaningful subgroups of patients with respect to the possibility of achieving RVR/cEVR. The CART belongs to a family of nonparametric regression methods based on binary recursive partitioning of data. The software automatically explore the data to search for optimal split variables, builds a decision tree structure and finally classifies all subjects into particular subgroups that are homogeneous with respect to the outcome of interest.[18] During the CART analysis, first, the entire study population, and thereafter, all newly defined subgroups, were investigated at every step of the analysis to determine which variable at what cut-off point yielded the most significant division into two prognostic subgroups that were as homogeneous as possible with respect to estimates of RVR/cEVR possibilities. This algorithm uses the impurity function (Gini criterion function) for splitting.[19] A restriction was imposed on the tree construction such that terminal subgroups resulting from any given split must have at least 20 patients. The CART procedure stopped when either no additional significant variable was detected or when the sample size was below 20. The resulting final subgroups were most homogeneous with respect to the probability of achieving RVR/cEVR. For this analysis, data mining software Clementine version 12.0 (SPSS Inc, Chicago, IL) was utilized. SPSS 15.0 (SPSS Inc, Chicago, IL) was used for logistic regression analysis.

RESULTS

Factors associated with RVR/cEVR by standard statistical analysis

WE FIRST ANALYZED 72 variables by univariate and multivariate logistic regression analysis to find factors associated with RVR/cEVR (Table 2). Patients with RVR/cEVR were significantly younger than those without. Among histological findings, grade of steatosis and stage of fibrosis was significantly lower in RVR/cEVR. Among hematologic tests, hemoglobin and hematcrit was significantly higher in RVR/cEVR. Among blood chemistry tests, creatinine and low-density lipoprotein cholesterol (LDL-C) was significantly higher and gamma-glutamyltransferase (GGT), low-density-lipoprotein cholesterol (LDL-C), and blood sugar were significantly lower in RVR/cEVR. The level of HCV RNA was significantly lower in RVR/cEVR. There were no significant differences in other tests.

Multivariate logistic regression analysis was performed on age, fibrosis stage, steatosis, HCVRNA, creatinine, hemoglobin, GGT, LDL-C, and blood sugar: hematcrit was not included since it is closely associated with hemoglobin. On multivariate analysis, age, grade of steatosis, level of HCV RNA, creatinine, hemoglobin, GGT, and LDL-cholesterol remained significant whereas stage of fibrosis, hemoglobin and blood sugar were not.

The CART analysis

The CART analysis was carried out on the model building set of 269 patients using the same variables as logistic regression analysis. Figure 1 shows the resulting decision tree. The CART analysis automatically selected five predictive variables to produce a total of seven subgroups of patients. The grade of steatosis was selected as the variable of initial split with an optimal cut-off of 30%. The possibility of achieving RVR/cEVR was only 18% for patients with hepatic steatosis of 30% or more compared to 47% for patients with hepatic steatosis of less than 30%. Among patients with hepatic steatosis of less than 30%, the level of serum LDL-C, with an optimal cut-off of 100 mg/dL, was selected as the variable of second split. Patients with higher LDL-C level had the higher probability of RVR/cEVR (57% vs. 32%). Among patients with LDL-C of less than 100 mg/dL, age, with an optimal cut-off of 60, was selected as the third variable of split. Younger patients had the higher probability of RVR/cEVR (49% vs. 15%). Among patients younger than 60, the blood sugar, with an optimal cut-off of 120 mg/dL, was selected as the forth variable of split. Patients with lower blood sugar level had the higher probability of RVR/cEVR (71% vs. 31%). Among patients with hepatic steatosis of less than 30% and LDL-C of 100 mg/dL or more, age, with an optimal cut-off of 50, was selected as the third variable of split, younger being the predictor of higher RVR/cEVR probability (77% vs. 50%). Among patients older than 50,

Table 1 Clinical characteristics of patients

	Model set $n = 269$	Validation set $n = 131$	*P*-value
Sex (M/F)	127/142	55/76	0.325
Age (years)	57.7 ± 10.1	57.6 ± 10.0	0.932
Body weight (kg)	59.6 ± 11.0	57.5 ± 9.5	0.094
Body mass index (kg/m^2)	23.2 ± 3.1	23.3 ± 3.8	0.934
Total protein (g/dL)	7.6 ± 0.5	7.7 ± 0.6	0.558
Albumin (g/dL)	4.2 ± 0.3	4.2 ± 0.3	0.349
Globulin (g/dL)	3.4 ± 0.5	3.4 ± 0.6	0.989
Aspartate aminotransferase (IU/L)	58.1 ± 43.1	55.8 ± 37.5	0.601
Alanine aminotransferase (IU/L)	70.9 ± 49.2	66.4 ± 52.6	0.462
Gamma-glutamyltransferase (IU/L)	49.6 ± 44.0	45.2 ± 34.4	0.33
Lactate dehydrogenase (IU/L)	289.3 ± 112.3	301.5 ± 109.3	0.417
Total bilirubin (mg/dL)	0.71 ± 0.28	0.69 ± 0.23	0.317
Direct bilirubin (mg/dL)	0.23 ± 0.12	0.25 ± 0.10	0.147
Indirect bilirubin (mg/dL)	0.48 ± 0.21	0.44 ± 0.16	0.064
Alkaline phosphatase (IU/L)	290.9 ± 107.6	292.5 ± 107.6	0.917
Leucine aminopeptidase (IU/L)	64.3 ± 14.3	65.5 ± 12.3	0.543
Thymol turbidity test (KU)	7.1 ± 3.4	8.0 ± 3.7	0.062
Zinc sulfate turbidity test (KU)	15.4 ± 4.9	16.3 ± 5.4	0.188
Choline esterase (IU/L)	318.1 ± 81.7	321.1 ± 78.1	0.798
Ammonia (microg/dL)	39.7 ± 20.2	45.0 ± 15.6	0.668
Blood sugar (mg/dL)	125.9 ± 41.1	117.4 ± 47.9	0.081
Glycohemoglobin (%)	5.6 ± 1.6	5.4 ± 1.2	0.797
Total cholesterol (mg/dL)	170.8 ± 33.9	175.6 ± 36.8	0.170
Low-density-lipoprotein-cholesterol (mg/dL)	96.5 ± 25.2	100.9 ± 28.5	0.153
High-density-lipoprotein-cholesterol (mg/dL)	54.2 ± 15.9	55.2 ± 17.4	0.612
Triglyceride (mg/dL)	108.5 ± 47.8	102.8 ± 46.4	0.306
Creatinine (mg/dL)	0.72 ± 0.15	0.74 ± 0.17	0.236
Urea nitrogen (mg/dL)	14.1 ± 3.4	14.9 ± 3.9	0.123
Uric acid (mg/dL)	5.3 ± 1.2	5.2 ± 1.2	0.715
Sodium (mEq/L)	142.2 ± 2.0	142.4 ± 2.0	0.471
Potassium (mEq/L)	4.3 ± 0.3	4.3 ± 0.4	0.578
Chloride (mEq/L)	104.0 ± 2.2	104.0 ± 2.6	0.905
Calcium (mg/dL)	9.1 ± 0.4	9.2 ± 0.4	0.479
Phosphorus (mg/dL)	3.5 ± 0.5	3.5 ± 0.6	0.814
Magnesium (mg/dL)	2.2 ± 0.2	2.3 ± 0.3	0.390
Amylase (IU/L)	178.7 ± 125.8	175.1 ± 133.1	0.118
Creatine kinase (IU/L)	114.9 ± 147.6	119.3 ± 73.7	0.849
Iron (microg/dL)	104.7 ± 53.2	109 ± 37	0.726
Ferritin (ng/mL)	111.3 ± 103.3	59.7 ± 118.5	0.405
C-reactive peptide (mg/dL)	0.2 ± 1.1	0.1 ± 0.1	0.586
Immunoglobulin G (mg/dL)	1849 ± 426	1988 ± 525	0.129
Immunoglobulin M (mg/dL)	141 ± 69	205 ± 106	0.200
Immunoglobulin A (mg/dL)	323 ± 675	291 ± 81	0.784
Triiodothyronine (pg/mL)	2.3 ± 0.3	2.2 ± 0.3	0.358
Thyroxin (ng/dL)	0.9 ± 0.1	0.9 ± 0.1	0.872
Thyroid stimulating hormone (micro IU/mL)	1.8 ± 1.4	1.7 ± 0.7	0.939
White blood cell count (/microl)	5243 ± 1591	5286 ± 1101	0.843
Segmented neutrophils (%)	55.4 ± 10.8	57.0 ± 10.0	0.297
Band neutrophils (%)	1.5 ± 1.6	0.5 ± 0.6	0.250
Eosinophils (%)	2.9 ± 2.3	2.4 ± 1.4	0.127

Table 1 *Continued*

	Model set $n = 269$	Validation set $n = 131$	*P*-value
Basophiles (%)	0.6 ± 0.4	0.6 ± 0.3	0.727
Lymphocytes (%)	34.6 ± 9.6	34.0 ± 9.3	0.682
Monocytes (%)	6.6 ± 2.2	6.2 ± 2.6	0.149
Red blood cell count (10^4/microl)	458 ± 43	455 ± 47	0.643
Hemoglobin (g/dL)	14.4 ± 1.5	14.5 ± 1.5	0.618
Hematcrit (%)	42.7 ± 4.0	42.9 ± 4.4	0.717
Reticulocytes (%)	1.4 ± 0.4	1.4 ± 0.4	0.762
Mean corpuscular volume (fL)	93.3 ± 4.5	93.8 ± 5.41	0.466
Mean corpuscular hemoglobin concentration (pg)	31.5 ± 1.9	31.7 ± 2.3	0.583
Mean corpuscular hemoglobin concentration (g/dL)	33.8 ± 0.9	33.7 ± 1.3	0.910
Platelets (10^4/microl)	16.8 ± 5.4	16.3 ± 4.5	0.480
Prothrombin time (s)	11.7 ± 1.2	11.7 ± 0.9	0.762
Prothrombin time (activity %)	104.6 ± 14.4	102.6 ± 14.8	0.363
Prothrombin time (international normalized ratio)	1.0 ± 0.1	1.0 ± 0.1	0.387
Thrombin time (%)	97.2 ± 31.3	109 ± 31.5	0.231
Activated partial thromboplastin time (s)	29.7 ± 4.4	29.1 ± 2.7	0.260
Hepaplastin test (%)	97.8 ± 20.3	95.4 ± 19.4	0.523
Fibrinogen (%)	237 ± 44	225 ± 45	0.069
Hepatitis C virus RNA (<850/≥850 KIU/mL)	130/139	70/61	0.394
Histological grade of			
Activity (A1/A2/A3)	138/107/24	62/55/14	0.714
Fibrosis (F1/F2/F3/F4)	135/74/57/3	58/40/27/6	0.131
Steatosis (0%/1–9%/10–29%/30%≦)	89/109/37/34	49/45/21/16	0.643
Hepatitis C virus RNA negative at week 12 (yes/no)	116/153	63/68	0.349

the level of GGT, with an optimal cutoff of 40 U/L, were then selected as the fourth level of split, low levels being the predictor of higher RVR/cEVR probability (60% vs. 35%).

All five factors selected as significant variables in the CART analysis were also significantly associated with RVR/cEVR by univariate analysis (Table 2). In addition, steatosis, LDL-C, age and GGT were also independently

Table 2 Factors associated with rapid or complete early virological response by univariate and multivariate logistic regression analysis

Parameter	Category	Univariate			Multivariate		
		Odds	95% CI	*P*-value	Odds	95% CI	*P*-value
Age (years)	<50 vs. ≥50	2.65	1.51–4.65	<0.001	2.03	1.04–3.97	0.039
Fibrosis stage	F1-2 vs. F3-4	2.47	1.31–4.66	0.005	1.77	0.85–3.68	0.120
Steatosis (%)	<30 vs. ≥30	4.11	1.64–10.29	0.003	2.88	1.07–7.79	0.037
Hepatitis C virus RNA (KIU/mL)	<850 vs. ≥850	1.97	1.21–3.22	0.007	1.93	1.09–3.43	0.025
Creatinine (mg/dL)	≥0.8 vs. <0.8	3.30	1.96–5.56	<0.001	3.54	1.88–6.67	<0.001
Hemoglobin (g/dL)	≥14.5 vs. <14.5	1.76	1.08–2.87	0.023	1.38	0.74–2.57	0.320
Hematcrit (%)	≥43 vs. <43	1.75	1.07–2.84	0.003			
Gamma-glutamyltransferase (IU/L)	<40 vs. ≥40	2.06	1.26–3.37	0.004	2.45	1.32–4.56	0.005
Low-density-lipid cholesterol (mg/dL)	≥100 vs. <100	2.71	1.61–4.55	<0.001	2.21	1.21–4.06	0.010
Blood sugar (mg/dL)	<120 vs. ≥120	2.00	1.02–3.95	0.045	1.42	0.64–3.13	0.390

CI, confidence interval.

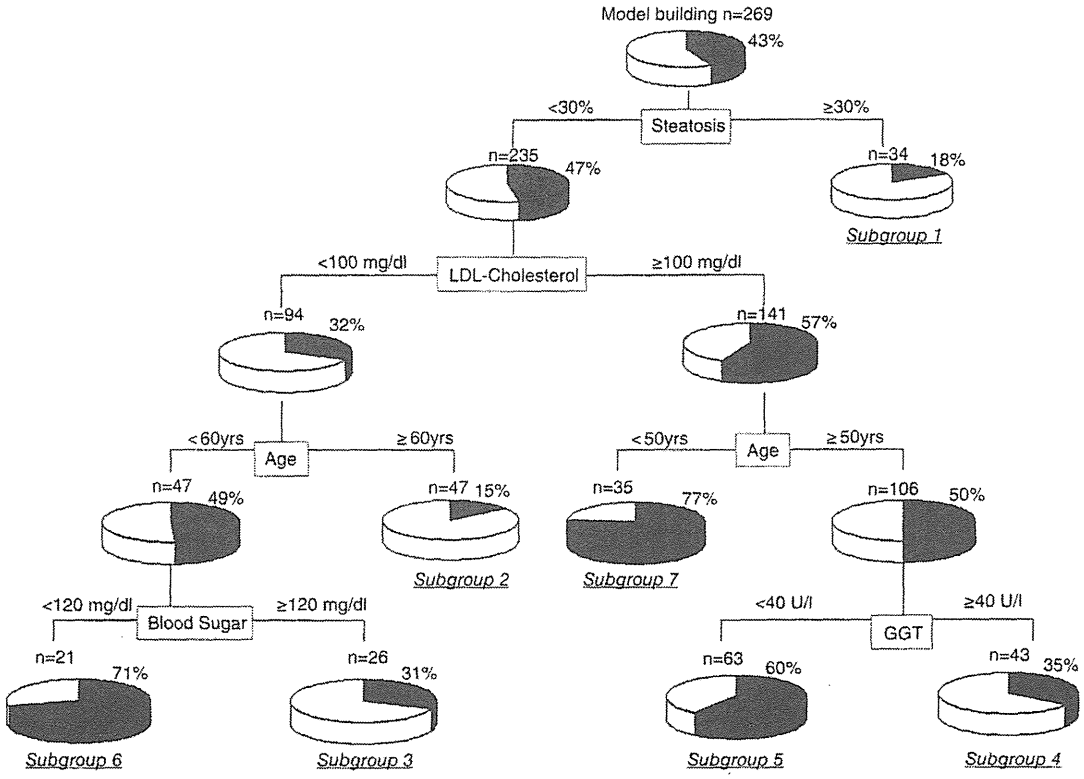

Figure 1 Classification and regression tree analysis. Boxes indicate the factors used for splitting and the cut-off value for the split. Pie charts indicate the rate of RVR/cEVR for each group of patients after splitting. Terminal subgroups of patients discriminated by the analysis are numbered from one to seven. GGT, gamma-glutamyltransferase; LDL, low-density-lipoprotein.

associated with RVR/cEVR by multivariate logistic regression analysis while blood sugar was not (Table 2). On the other hand, HCVRNA and creatinine which were significantly associated with RVR/cEVR by multivariate analysis were not selected as significant variables in CART analysis.

The probabilities of RVR/cEVR for the seven subgroups derived by this process were highly variable. The subgroup whose hepatic steatosis was less than 30%, serum LDL-C was 100 mg/dL or more and of an age less than 50 years (subgroup 7) showed the highest probability of RVR/cEVR (77%), while the subgroup whose hepatic steatosis more than 30% (subgroup 1) and the subgroup whose hepatic steatosis was less than 30% but serum LDL-C was less than 100 mg/dL and of an age greater than 60 years (subgroup 2) showed the lowest probability of RVR/cEVR (18% and 15%, respectively).

Validation of the CART analysis

The results of the CART analysis were validated with a validation dataset of 131 cases which is independent of the model building dataset. Each patient in the validation set was allocated to subgroups 1–7 using the flowchart form of the CART tree. The rates of RVR/cEVR were 20% for subgroups 1 and 2, 29% for subgroups 3, 38% for subgroup 4, 59% for subgroup 5, 71% for subgroup 6, and 85% for subgroups 7. The rates of RVR/cEVR for each subgroup of patients were closely correlated between the model building dataset and the validation dataset (Fig. 2).

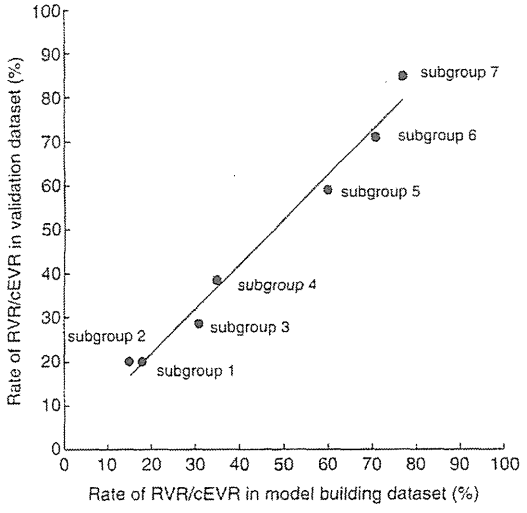

Figure 2 Validation of the classification and regression tree (CART) analysis: Subgroup stratified comparison of the rate of rapid or complete early virological response (RVR/cEVR) between the model building and validation datasets. Each patient in the validation set was allocated to subgroups 1–7 by following the flow-chart form of the CART tree and the rates of RVR/cEVR were calculated. The rate of RVR/cEVR in each subgroup was plotted. The *x*-axis represents the rate of RVR/cEVR in the model building datasets and the y-axis represents the rate of RVR/cEVR in the validation datasets. The rates of achieving RVR/cEVR in each subgroup of patients closely correlated between the model building dataset and the validation dataset ($r^2 = 0.987$).

Construction of 3 groups according to the probability of RVR/cEVR

If the seven subgroups were reconstructed into three groups according to their rate of RVR/cEVR, the rate of RVR/cEVR was 16% for low probability group (subgroup 1 and 2), 46% for intermediate probability group (subgroup 3, 4, and 5) and 75% for high probability group (subgroup 6 and 7; $P < 0.0001$).

Effect of adherence

Adherence of PEG-IFN and RBV was not included as a variable of analysis since the present study aimed to develop a pre-treatment model for the prediction of response. To analyze the possible effect of adherence on the result of CART analysis, three groups of patients divided by CART (low, intermediate and high probability group) were further stratified according to adherence of PEG-IFN and RBV. Poor adherence was defined as taking less than 80% planned dose of PEG-IFN or RBV at 12 weeks, and good adherence was defined as taking more than 80% planned dose of both PEG-IFN and RBV at 12 weeks. The result is shown in Figure 3. Among patients with good adherence, the rate of RVR/cEVR was 19% for low probability group, 52% for intermediate probability group and 77% for high probability group. Among poor adherence group, the rate of RVR/cEVR was 13% for low probability group, 41% for intermediate probability group and 73% for high probability group. Collectively, even after adjustment for adherence, 3 groups of patients divided by CART analysis still had low, intermediate and high probability of achieving RVR/cEVR, respectively.

DISCUSSION

IN THE PRESENT study, we performed the CART analysis and built a simple decision tree model for the pre-treatment prediction of response to PEG-IFN plus

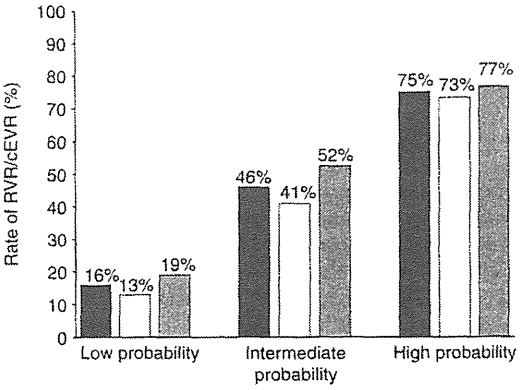

Figure 3 The rate of rapid or complete early virological response (RVR/cEVR) between the classification and regression tree (CART) groups stratified by adherence. The three groups of patients divided by CART (low, intermediate and high probability group) were further stratified according to adherence of peg-interferon (PEG-IFN) plus ribavirin (RBV). Black, white and gray boxes in the bar chart indicate total patients, patients with poor adherence (taking less than 80% planned dose of PEG-IFN or RBV at 12 weeks), and good adherence (taking more than 80% planned dose of both PEG-IFN and RBV at 12 weeks), respectively. Even after adjustment for adherence, 3 groups of patients divided by CART analysis still had low, intermediate and high probability of achieving RVR/cEVR, respectively.

RBV therapy. The analysis highlighted five host variables relevant to response: steatosis, LDL-C, age, blood sugar and GGT. Classification of patients based on these variables identified subgroups of patients with high probabilities of achieving RVR/cEVR among difficult to treat chronic hepatitis C patients. The reproducibility of the model was confirmed by the independent validation datasets. According to the result of the CART, patients were categorized into 3 groups: the rate of RVR/cEVR was 16% for low probability group, 46% for intermediate probability group and 75% for high probability group. The result of the CART analysis could be readily applicable to clinical practice because patients could be allocated to specific subgroups with a defined rate of response simply by following the flow-chart form. Although an early disappearance of serum HCV RNA is the prerequisite for achieving SVR, no reliable baseline predictors of response to PEG-IFN plus RBV therapy are established to date. Thus, this model may have the potential to support decisions in patient selection for PEG-IFN plus RBV therapy or to tailor treatment strategies for individual patients. Moreover, our result may provide a rationale for treating metabolic factors to improve the efficacy of antiviral therapy.

Among variables relevant to the prediction of RVR/cEVR, the grade of hepatic steatosis was selected as the variable of the first split. Previous studies suggested that steatosis induces resistance to IFN and RBV combination therapy[20,21] along with underlining metabolic factors such as insulin resistance or obesity.[21–24] In the present study, the grade of steatosis correlated positively with BMI and serum glucose level (data not shown) suggesting the etiologic role of metabolic factors. In addition, serum glucose level was selected as a predictor of RVR/cEVR at the fourth level of split. Serum GGT, which is associated with obesity,[25] insulin resistance[26] and response to IFN therapy,[27–30] was also selected as a predictor of RVR/cEVR at fourth level of splitting which may emphasize the importance of metabolic factors in therapeutic resistance. These findings raise the possibility that treatment of these metabolic factors may improve the virological response to the PEG-IFN plus RBV therapy. This hypothesis should be examined by a prospective study.

We and others have reported that steatosis, obesity and insulin resistance are associated with the progression of fibrosis,[17,31–33] which can interfere indirectly with the effect of IFN on hepatocytes. Other possible mechanisms of resistance by steatosis or metabolic factors include dysregulation of adipocytokines[34] or oxidative stress which may inhibit intracellular IFN signaling pathway.[35] Despite these findings, the precise mechanism of resistance is not established and further investigation is needed.

Another factor relevant in the prediction of RVR/cEVR was LDL-C. LDL-C was selected as the second factor for splitting by CART, and was an independent predictor of RVR/cEVR by logistic regression analysis. LDL-C recently has attracted attention as a novel predictor of response to IFN or PEG-IFN plus RBV.[30,36,37] Since *in vitro* study showed that LDL-C receptor acts as a receptor for HCV and LDL-C competitively inhibit the binding of HCV,[38] high level of serum LDL-C may inhibit HCV entry to hepatocytes and attenuate replication. LDL-C and its receptor may be a future therapeutic target.

Not all factors selected as significant variables in the CART analysis were also significantly associated with response by standard statistical analysis: blood sugar was associated with response by univariate analysis but not by multivariate logistic regression analysis On the other hand, HCVRNA and creatinine which were significantly associated with RVR/cEVR by multivariate analysis were not selected as significant variables in CART analysis. These differences may indicate both the unique feature and the limitations of the CART analysis. To note, blood sugar was significantly associated with RVR/cEVR within specialized subgroups of patients defined by the CART analysis: in subgroup of patients with steatosis <30%, LDL-C <100 mg/dL and younger than 60, which indicate the unique feature of the CART analysis that it could visualize significant predictors that specifically apply to selected patients. The limitation is that not all significant factors may be adopted in the decision tree since we applied the rule to stop CART procedure when the sample size was below 20. This rule was applied to avoid the generation of over-fit model which may lack universality. Therefore, it is possible that HCVRNA or creatinine may become a significant variable in the CART analysis if larger number of patients were included in the analysis. Stage of fibrosis was significantly associated with response to therapy by univariate analysis but not by multivariate analysis and not selected as a significant variable in the CART analysis. The possible reason is that advanced fibrosis is associated with older age as a confounding factor.

CART analyses are gaining acceptance in medical research in addition to biomedical field. Recent publications include the prediction of aggressive prostate cancer,[8] diabetic vascular complications,[19] prognosis of melanoma,[7,39] response to preoperative radiotherapy for rectal tumor,[9] prognostic groups in colorectal carcinoma,[12] and outcome after liver failure.[11] An advantage

of CART over traditional regression models is that it can identify prognostic subgroups that are useful in clinical practice. Because the results of CART analysis are presented as a decision tree, which is intuitive, they can be readily interpreted by medical professionals without any specific knowledge of statistics. The most important consideration is that five variables used in the decision tree were clinical parameters that are readily available by the usual work-up of patients before therapy. Especially, glucose, GGT and LDL-C are simple biochemical markers that are easily measured at a low cost. Using this model, we can rapidly develop an estimate of the response before treatment, which may facilitate clinical decision making.

In conclusion, we built a pre-treatment model for the prediction of virological response in PEG-IFN plus RBV therapy. Because this decision tree model was made up of simple host factors such as steatosis, LDL-C, age, blood sugar and GGT, it can be easily applied to clinical practice. This model may have the potential to support decisions in patient selection for PEG-IFN plus RBV therapy based on the possibility of response against a potential risk of adverse events or costs, and may provide a rationale for treating metabolic factors to improve the efficacy of antiviral therapy.

ACKNOWLEDGEMENTS

THIS STUDY WAS supported by a grant-in-aid from Ministry of Health, Labor and Welfare, Japan. There exist no conflicts of interest.

REFERENCES

1 Strader DB, Wright T, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C. *Hepatology* 2004; 39: 1147–71.
2 Fried MW, Shiffman ML, Reddy KR *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 2002; 347: 975–82.
3 Manns MP, McHutchison JG, Gordon SC *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: 958–65.
4 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: 645–52.
5 Lee SS, Ferenci P. Optimizing outcomes in patients with hepatitis C virus genotype 1 or 4. *Antivir Ther* 2008; 13 (Suppl 1): 9–16.
6 Breiman LJH, Friedman RA, Olshen CJ, Stone CM. *Classification and Regression Trees*. Calif: Wadsworth, 1980.
7 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. *Surg* 2002; 132: 589–602.
8 Garzotto M, Beer TM, Hudson RG *et al.* Improved detection of prostate cancer using classification and regression tree analysis. *J Clin Oncol* 2005; 23: 4322–9.
9 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: 5440–3.
10 Jin H, Lu Y, Harris ST *et al.* Classification algorithms for hip fracture prediction based on recursive partitioning methods. *Med Decis Making* 2004; 24: 386–98.
11 Baquerizo A, Anselmo D, Shackleton C *et al.* Phosphorus ans an early predictive factor in patients with acute liver failure. *Transplantation* 2003; 75: 2007–14.
12 Valera VA, Walter BA, Yokoyama N *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: 34–40.
13 Martin MA, Meyricke R, O'Neill T, Roberts S. Mastectomy or breast conserving surgery? Factors affecting type of surgical treatment for breast cancer – a classification tree approach. *BMC Cancer* 2006; 6: 98.
14 LeBlanc M, Crowley J. A review of tree-based prognostic models. *Cancer Treat Res* 1995; 75: 113–24.
15 Costanza MC, Paccaud F. Binary classification of dyslipidemia from the waist-to-hip ratio and body mass index: a comparison of linear, logistic, and CART models. *BMC Med Res Methodol* 2004; 4: 7.
16 Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology* 1996; 24: 289–93.
17 Kurosaki M, Matsunaga K, Hirayama I *et al.* The presence of steatosis and elevation of alanine aminotransferase levels are associated with fibrosis progression in chronic hepatitis C with non-response to interferon therapy. *J Hepatol* 2008; 48: 736–42.
18 Segal MR, Bloch DA. A comparison of estimated proportional hazards models and regression trees. *Stat Med* 1989; 8: 539–50.
19 Miyaki K, Takei I, Watanabe K, Nakashima H, Omae K. Novel statistical classification model of type 2 diabetes mellitus patients for tailor-made prevention using data mining algorithm. *J Epidemiol* 2002; 12: 243–8.
20 Akuta N, Suzuki F, Tsubota A *et al.* Efficacy of interferon monotherapy to 394 consecutive naive cases infected with hepatitis C virus genotype 2a in Japan: therapy efficacy as consequence of tripartite interaction of viral, host and interferon treatment-related factors. *J Hepatol* 2002; 37: 831–6.

21 Poynard T, Ratziu V, McHutchison J *et al.* Effect of treatment with peginterferon or interferon alfa-2b and ribavirin on steatosis in patients infected with hepatitis C. *Hepatology* 2003; 38: 75–85.
22 Bressler BL, Guindi M, Tomlinson G, Heathcote J. High body mass index is an independent risk factor for nonresponse to antiviral treatment in chronic hepatitis C. *Hepatology* 2003; 38: 639–44.
23 Romero-Gomez M, Del Mar Viloria M, Andrade RJ *et al.* Insulin resistance impairs sustained response rate to peginterferon plus ribavirin in chronic hepatitis C patients. *Gastroenterology* 2005; 128: 636–41.
24 Konishi I, Horiike N, Hiasa Y *et al.* Diabetes mellitus reduces the therapeutic effectiveness of interferon-alpha2b plus ribavirin therapy in patients with chronic hepatitis C. *Hepatol Res* 2007; 37: 331–6.
25 Marchesini G, Avagnina S, Barantani EG *et al.* Aminotransferase and gamma-glutamyltranspeptidase levels in obesity are associated with insulin resistance and the metabolic syndrome. *J Endocrinol Invest* 2005; 28: 333–9.
26 Fraser A, Ebrahim S, Smith GD, Lawlor DA. A comparison of associations of alanine aminotransferase and gamma-glutamyltransferase with fasting glucose, fasting insulin, and glycated hemoglobin in women with and without diabetes. *Hepatology* 2007; 46: 158–65.
27 Mazzella G, Salzetta A, Casanova S *et al.* Treatment of chronic sporadic-type non-A, non-B hepatitis with lymphoblastoid interferon: gamma GT levels predictive for response. *Dig Dis Sci* 1994; 39: 866–70.
28 Villela-Nogueira CA, Perez RM, de Segadas Soares JA, Coelho HS. Gamma-glutamyl transferase (GGT) as an independent predictive factor of sustained virologic response in patients with hepatitis C treated with interferon-alpha and ribavirin. *J Clin Gastroenterol* 2005; 39: 728–30.
29 Berg T, Sarrazin C, Herrmann E *et al.* Prediction of treatment outcome in patients with chronic hepatitis C: significance of baseline parameters and viral dynamics during therapy. *Hepatology* 2003; 37: 600–9.
30 Akuta N, Suzuki F, Kawamura 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: 403–10.
31 Adinolfi LE, Gambardella M, Andreana A, Tripodi MF, Utili R, Ruggiero G. Steatosis accelerates the progression of liver damage of chronic hepatitis C patients and correlates with specific HCV genotype and visceral obesity. *Hepatology* 2001; 33: 1358–64.
32 Ortiz V, Berenguer M, Rayon JM, Carrasco D, Berenguer J. Contribution of obesity to hepatitis C-related fibrosis progression. *Am J Gastroenterol* 2002; 97: 2408–14.
33 Muzzi A, Leandro G, Rubbia-Brandt L *et al.* Insulin resistance is associated with liver fibrosis in non-diabetic chronic hepatitis C patients. *J Hepatol* 2005; 42: 41–6.
34 Charlton MR, Pockros PJ, Harrison SA. Impact of obesity on treatment of chronic hepatitis C. *Hepatology* 2006; 43: 1177–86.
35 Di Bona D, Cippitelli M, Fionda C *et al.* Oxidative stress inhibits IFN-alpha-induced antiviral gene expression by blocking the JAK-STAT pathway. *J Hepatol* 2006; 45: 271–9.
36 Minuk GY, Weinstein S, Kaita KD. Serum cholesterol and low-density lipoprotein cholesterol levels as predictors of response to interferon therapy for chronic hepatitis C. *Ann Intern Med* 2000; 132: 761–2.
37 Gopal K, Johnson TC, Gopal S *et al.* Correlation between beta-lipoprotein levels and outcome of hepatitis C treatment. *Hepatology* 2006; 44: 335–40.
38 Agnello V, Abel G, Elfahal M, Knight GB, Zhang QX. Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proc Natl Acad Sci USA* 1999; 96: 12766–71.
39 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: 3660–7.

Association of Gene Expression Involving Innate Immunity and Genetic Variation in Interleukin 28B With Antiviral Response

Yasuhiro Asahina,[1] Kaoru Tsuchiya,[1] Masaru Muraoka,[1,2] Keisuke Tanaka,[1,2] Yuichiro Suzuki,[1,2] Nobuharu Tamaki,[1] Yoshihide Hoshioka,[1] Yutaka Yasui,[1] Tomoji Katoh,[1] Takanori Hosokawa,[1] Ken Ueda,[1] Hiroyuki Nakanishi,[1] Jun Itakura,[1] Yuka Takahashi,[1] Masayuki Kurosaki,[1] Nobuyuki Enomoto,[2] Sayuri Nitta,[3] Naoya Sakamoto,[3] and Namiki Izumi[1]

Innate immunity plays an important role in host antiviral response to hepatitis C viral (HCV) infection. Recently, single nucleotide polymorphisms (SNPs) of *IL28B* and host response to peginterferon α (PEG-IFNα) and ribavirin (RBV) were shown to be strongly associated. We aimed to determine the gene expression involving innate immunity in *IL28B* genotypes and elucidate its relation to response to antiviral treatment. We genotyped *IL28B* SNPs (rs8099917 and rs12979860) in 88 chronic hepatitis C patients treated with PEG-IFNα-2b/RBV and quantified expressions of viral sensors (*RIG-I, MDA5*, and *LGP2*), adaptor molecule (*IPS-1*), related ubiquitin E3-ligase (*RNF125*), modulators (*ISG15* and *USP18*), and *IL28* (*IFNλ*). Both *IL28B* SNPs were 100% identical; 54 patients possessed rs8099917 TT/rs12979860 CC (*IL28B* major patients) and 34 possessed rs8099917 TG/rs12979860 CT (*IL28B* minor patients). Hepatic expressions of viral sensors and modulators in *IL28B* minor patients were significantly up-regulated compared with that in *IL28B* major patients (≈3.3-fold, $P < 0.001$). However, expression of *IPS-1* was significantly lower in *IL28B* minor patients (1.2-fold, $P = 0.028$). Expressions of viral sensors and modulators were significantly higher in nonvirological responders (NVR) than that in others despite stratification by *IL28B* genotype (≈2.6-fold, $P < 0.001$). Multivariate and ROC analyses indicated that higher *RIG-I* and *ISG15* expressions and *RIG-I/IPS-1* expression ratio were independent factors for NVR. IPS-1 down-regulation in *IL28B* minor patients was confirmed by western blotting, and the extent of IPS-1 protein cleavage was associated with the variable treatment response. *Conclusion:* Gene expression involving innate immunity is strongly associated with *IL28B* genotype and response to PEG-IFNα/RBV. Both *IL28B* minor allele and higher *RIG-I* and *ISG15* expressions and *RIG-I/IPS-1* ratio are independent factors for NVR. (HEPATOLOGY 2012;55:20-29)

Infection with hepatitis C virus (HCV) is a common cause of chronic hepatitis, which progresses to liver cirrhosis and hepatocellular carcinoma in many patients.[1] Pegylated interferon α (PEG-IFNα) and ribavirin (RBV) combination therapy has been used to treat chronic hepatitis C (CH-C) to alter the natural course of this disease. However, 20% patients are nonvirological responders (NVR) whose HCV-RNA does not become negative during the 48 weeks of PEG-IFNα/RBV combination therapy.[2] In a recent genome-wide association study, single nucleotide polymorphisms (SNPs) located near interleukin 28B

Abbreviations: CH-C, chronic hepatitis C; γ-GTP, γ-glutamyl transpeptidase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HCV, hepatitis C virus; HMBS, hydroxymethylbilane synthase; IL28, interleukin 28; IPS-1, IFNβ promoter stimulator 1; ISG15, interferon-stimulated gene 15; MDA5, melanoma differentiation associated gene 5; NVR, nonvirological responders; PEG-IFNα, pegylated interferonα; SNP, single nucleotide polymorphism; RIG-I, retinoic acid-inducible gene I; RBV, ribavirin; RNF125, ring-finger protein 125; ROC, receiver operator characteristic; SVR, sustained viral responder; TVR, transient virological responder; USP18, ubiquitin-specific protease 18; VR, virological responder.

From the [1]Department of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan; [2]First Department of Internal Medicine, Faculty of Medicine, University of Yamanashi, Yamanashi, Japan; [3]Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan.

Received May 14, 2011; accepted August 16, 2011.

Supported by grants from the Japanese Ministry of Education, Culture, Sports, Science and Technology and the Japanese Ministry of Welfare, Health and Labor. The funding source had no role in the collection, analysis, or interpretation of the data, or in the decision to submit the article for publication.

(*IL28B*) that encodes for type III IFNλ3 were shown to be strongly associated with a virological response to PEG-IFNα/RBV combination therapy.[3-5] In particular, the rs8099917 TG and GG genotypes were shown to be strongly associated with a null virological response to PEG-IFNα/RBV.[3] However, mechanisms involving resistance to PEG-IFNα/RBV have not been completely elucidated.

The innate immune system has an essential role in host antiviral defense against HCV infection.[6] The retinoic acid-inducible gene I (RIG-I), a cytoplasmic RNA helicase, and related melanoma differentiation associated gene 5 (MDA5) play essential roles in initiating the host antiviral response by detecting intracellular viral RNA.[7,8] The IFNβ promoter stimulator 1 (IPS-1)—also called the caspase-recruiting domain adaptor inducing IFNβ, mitochondrial antiviral signaling protein, or virus-induced signaling adaptor—is an adaptor molecule. IPS-1 connects RIG-I sensing to downstream signaling, resulting in *IFNβ* gene activation.[9-12] RIG-I sensing of incoming viral RNA has been shown to be modified by LGP2,[8,13] a helicase related to RIG-I and MDA5 lacking caspase-recruiting domain. The ubiquitin ligase ring-finger protein 125 (RNF125) has been shown to conjugate ubiquitin to RIG-I, MDA5, and IPS-1 and this suppresses the functions of these proteins.[14] Further, these molecules are ISGylated by the IFN-stimulated gene 15 (ISG15), a ubiquitin-like protein,[15] and ISG15 is specifically removed from ISGylated protein by ubiquitin-specific protease 18 (USP18) to regulate the RIG-I/IPS-1 system.[16,17] Moreover, the NS3/4A protease of HCV specifically cleaves IPS-1 as part of its immune-evasion strategy.[9,18] Therefore, the RIG-I/IPS-1 system and its regulatory systems have essential roles in the innate antiviral response.

Recently, we demonstrated that baseline intrahepatic gene expression levels of the RIG-I/IPS-1 system were prognostic biomarkers of the final virological outcome in CH-C patients who were treated with PEG-IFNα/RBV combination therapy.[19] We found that up-regulation of *RIG-I* and *ISG15* and a higher expression ratio of *RIG-I/IPS-1* could predict NVR for subsequent treatment with PEG-IFNα/RBV combination therapy.[19] However, association of gene expression involving innate immunity and genetic variation of *IL28B* has not yet been elucidated. Hence, the aim of this study was to determine gene expression involving the innate immune system in different genetic variations of *IL28B* and elucidate the relation of gene expression to final virological outcome of PEG-IFNα/RBV combination therapy in CH-C patients.

Patients and Methods

Patients. Among histologically proven CH-C patients admitted at the Musashino Red Cross Hospital, 88 patients with HCV genotype 1b and a high viral load (>5 log IU/mL by TaqMan HCV assay; Roche Molecular Diagnostics, Tokyo, Japan) were included in the present study (Table 1). Patients with decompensated liver cirrhosis, autoimmune hepatitis, or alcoholic liver injury were excluded. No patient had tested positive for hepatitis B surface antigen or anti-human immunodeficiency virus antibody or had received immunomodulatory therapy before enrollment. Forty-two patients had been enrolled in a previous study that determined hepatic gene expression involving innate immunity.[19] Written informed consent was obtained from all patients and the study was approved by the Ethical Committee of Musashino Red Cross Hospital in accordance with the Declaration of Helsinki.

Treatment Protocol. The patients were administered subcutaneous injections of PEG-IFNα-2b (PegIntron, MSD, Whitehouse Station, NJ) at a dose of 1.5 $\mu g\ kg^{-1}\ week^{-1}$ for 48 weeks. RBV (Rebetol, MSD) was administered concomitantly over this treatment period, administered orally twice daily at 600 mg/day for patients who weighed less than 60 kg and 800 mg/day for patients who weighed between 60-80 kg. The dose of PEG-IFNα-2b was reduced to 0.75 $\mu g\ kg^{-1}\ week^{-1}$ when either neutrophil count was less than $750/mm^3$ or platelet count was less than $80 \times 10^3/mm^3$. The dose of RBV was reduced to 600 mg/day when the hemoglobin concentration decreased to 10 g/dL. More than 80% adherence was achieved in all patients.

Measurement of Hepatic Gene Expression. Liver biopsy was performed immediately before initiating

Address reprint requests to: Namiki Izumi, M.D., Ph.D., Chief, Department of Gastroenterology and Hepatology, Musashino Red Cross Hospital, 1-26-1 Kyonancho 1-26-1, Musashinoshi, Tokyo 180-8610, Japan. E-mail: nizumi@musashino.jrc.or.jp; fax: +81-422-32-9551.

View this article online at wileyonlinelibrary.com.
DOI 10.1002/hep.24623
Potential conflict of interest: Nothing to report.
Additional Supporting Information may be found in the online version of this article.

Table 1. Patient Characteristics and *IL28B* Genotype

	IL28B Major*	*IL28B* Minor†	*P*-value‡
Patients, n	54	34	
Age (SD), year	58.8 (10.0)	59.1 (10.3)	0.918§
Sex, n (%)			0.051‖
Male	13 (24.1)	15 (44.1)	
Female	41 (75.9)	19 (55.9)	
BMI (SD), kg/m^2	22.7 (3.5)	23.5 (3.6)	0.193§
ALT (SD), IU/L	61.3 (50.7)	62.4 (44.7)	0.962§
γ-GTP (SD), IU/L	36.7 (25.9)	57.3 (52.4)	0.010§
LDL-cholesterol (SD), mg/dL	103.3 (29.8)	91.8 (26.9)	0.067§
Hemoglobin (SD), g/dL	14.1 (1.4)	14.4 (1.3)	0.186§
Platelet count (SD), ×10^{3}3/μL	161 (6.4)	163 (4.4)	0.489§
Fibrosis stage, n (%)			0.532‖
F1, 2	38 (70.4)	26 (76.5)	
F3, 4	16 (29.6)	8 (23.5)	
Viral load (SD), ×10^{63} IU/mL	1.7 (1.4)	1.9 (2.0)	0.788§
%HCV core 70 & 91 a.a. double mutation¶	8.9	43.5	0.001‖
%ISDR wild**	43.5	51.7	0.486‖
Viral response, n (%)			<0.001‖
SVR	17 (31.5)	13 (38.2)	
TVR	26 (48.1)	3 (8.8)	
NVR	11 (20.4)	18 (52.9)	

Unless otherwise indicated, data are given as mean (SD).

*rs8099917 TT and rs12979860 CC.

†rs8099917 TG and rs12979860 CT.

BMI, body mass index; ALT, alanine aminotransferase; γ-GTP, γ-glutamyl transpeptidase; LDL-C, low-density lipoprotein cholesterol; HCV, hepatitis C virus; ISDR, interferon sensitivity determining region; SVR, sustained virological response; TVR, transient virological response; NVR, nonvirological response.

‡Comparison between *IL28B* major and minor genotypes.

§Mann-Whitney *U* test.

‖Chi-square test.

¶HCV core mutation was determined in 68 patients.

**ISDR was determined in 75 patients.

the therapy. After extraction of total RNA from liver biopsy specimens, the messenger RNA (mRNA) expression of the positive and negative cytoplasmic viral sensor (*RIG-I, MDA5,* and *LGP2*), the adaptor molecule (*IPS-1*), the related ubiquitin E3-ligase (*RNF125*), the modulators of these molecules (*ISG15* and *USP18*), and *IFNλ* (*IL28A/B*) was quantified by real-time quantitative polymerase chain reaction (PCR) using target gene-specific primers. In brief, total RNA was extracted by the acid-guanidinium-phenol-chloroform method using Isogen reagent (Nippon Gene, Toyama, Japan) from the liver biopsy specimen, which was 0.2-0.4 cm in length and 13G in diameter. Complementary DNA (cDNA) was transcribed from 2 μg of total RNA template in a 140-μL reaction mixture using the SYBR RT-PCR Kit (Takara Bio, Otsu, Japan) with random hexamer. Real-time quantitative PCR was performed using Smart Cycler version II (Takara Bio) with the SYBR RT-PCR Kit (Takara Bio) according to the manufacturer's instructions. Assays were performed in duplicate and the expression levels of target genes were normalized to the expressions of glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) gene and hydroxymethylbilane synthase (*HMBS*), an enzyme that is stable in the liver, as quantified using real-time quantitative PCR as internal controls. For accurate normalization, a set of two housekeeping genes was used in the present study. Sequences of the primer sets were as follows: *RIG-I*, 5′-AAAGCATGCA TGGTGTTCCAGA-3′, 5′-TCATTCGTGCATGCTC ACTGATAA-3′; *MDA5*, 5′-ACATAACAGCAACATG GGCAGTG-3′, 5′-TTTGGTAAGGCCTGAGCTGG AG-3′; *LGP2*, 5′-ACAGCCTTGCAAACAGTACAAC CTC-3′, 5′-GTCCCAAATTTCCGGCTCAAC-3′; *IPS-1*, 5′-GGTGCCATCCAAAGTGCCTACTA-3′, 5′-CAGC ACGCCAGGCTTACTCA-3′; *RNF125*, 5′-AGGGCA CATATTCGGACTTGTCA-3′, 5′-CGGGTATTAAAC GGCAAAGTGG-3′; *ISG15*, 5′-AGCGAACTCATCT TTGCCAGTACA-3′, 5′-CAGCTCTGACACCGACA TGGA-3′; *USP18*, 5′-TGGTTCTGCTTCAATGACT CCAATA-3′, 5′-TTTGGGCATTTCCATTAGCACT C-3′; *IFNλ*: 5′-CAGCTGCAGGTGAGGGA-3′, 5′-G GTGGCCTCCAGAACCTT-3′; *GAPDH*, 5′-GCACC GTCAAGGCTGAGAAC-3′, 5′-ATGGTGGTGAAGA CGCCAGT-3′; *HMBS*, 5′-AAGCGGAGCCATGTCT GGTAAC-3′, 5′-GTACCCACGCGAATCACTCTCA-3′.

Genotyping for* IL28B *(rs8099917 and rs12979860) Polymorphism. Genetic polymorphism in a tagged SNP located near the *IL28B* gene (rs8099917 and rs12979860) was determined by direct sequencing of PCR-amplified DNA. In brief, after extraction from whole blood samples, genomic DNA was amplified by PCR. Sequences of the primer sets were: rs8099917, 5′-ATCCTCCTCTCATCCCTCA TC-3′, 5′-GGTATCAACCCCACCTCAAAT-3′; rs129 79860, 5′-GGACGAGAGGGCGTTAGAG-3′, 5′-AG GGACCGCTACGTAAGTCAC-3′.

Both strands of the PCR products were sequenced by the dye terminator method using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Chiba, Japan); nucleotide sequences were determined by a capillary DNA sequencer ABI3730xl (Applied Biosystems). Homozygosity (rs8099917 GG and rs12979860 TT) or heterozygosity (rs8099917 TG and rs12979860 CT) of the minor sequence was defined as having the *IL28B* minor allele, whereas homozygosity for the major sequence (rs8099917 TT and rs12979860 CC) was defined as having the *IL28B* major allele.

Western Blotting. Western blotting was performed using samples from 14 patients (six from *IL28B* major patients and eight from *IL28B* minor patients) as described.[19] In brief, liver biopsy specimens of

approximately 10 mg were homogenized in 100 μL of Complete Lysis-M (Roche Applied Science, Penzberg, Germany). Next, 30 μg of protein was separated by NuPAGE 4%-12% Bis-Tris gels (Invitrogen, Carlsbad, CA) and blotted on polyvinylidene difluoride membranes. The membranes were immunoblotted with anti-RIG-I (Cell Signaling Technology, Danvers, MA) or anti-IPS-1 (Enzo Life Science, Farmingdale, NY), followed by anti-β-actin (Sigma Aldrich, St. Louis, MO). After immunoblotting with horseradish peroxidase-conjugated secondary antibody, signals were detected by chemiluminescence (BM Chemiluminescence Blotting Substrate, Roche Applied Science, Mannheim, Germany). Optical densitometry was performed using ImageJ software (NIH, Bethesda, MD). Naive Huh7 cells were used for a positive control for full-length IPS-1, and cells transfected with HCV-1b subgenomic replicon[20] were used for a positive control for cleaved IPS-1.

Definitions of Response to Therapy. A patient negative for serum HCV-RNA during the first 6 months after completing PEG-IFNα-2b/RBV combination therapy was defined as a sustained viral responder (SVR), and a patient for whom HCV-RNA became negative at the end of therapy and reappeared after completion of therapy was defined as a transient virological responder (TVR). A patient for whom HCV-RNA became negative at the end of therapy (SVR + TVR) was defined as a virological responder (VR). A patient whose HCV-RNA did not become negative during the course of therapy was defined as an NVR. HCV-RNA was determined by TaqMan HCV assay (Roche Molecular Diagnostics).

Statistical Analysis. Categorical data were compared using the chi-square test and Fisher's exact test. Distributions of continuous variables were analyzed by the Mann-Whitney *U* test for two groups. All tests of significance were two-tailed and $P < 0.05$ was considered statistically significant.

Results

Patient Characteristics and* IL28B *Genotype. Table 1 shows patient characteristics according to *IL28B* genotype. SNPs at rs8099917 and rs12979860 were 100% identical; 54 patients were identified as having the major alleles (rs8099917 TT/rs12979860 CC; *IL28B* major patients) and the remaining 34 had the minor alleles (rs8099917 TG/rs12979860 CT; *IL28B* minor patients). Patients having a minor homozygote (rs8099917 GG or rs12979860 TT) were not found in this study, which is consistent with a recent report of the rarity of a minor homozygote in Japanese patients.[3] *IL28B* minor patients were significantly associated with a higher γ-glutamyl transpeptidase (γ-GTP) level and higher frequency of mutations at amino acid positions 70 and 91 of the HCV core region (glutamine or histidine mutation at amino acid position 70; methionine mutation at amino acid position 91). NVR rate was significantly higher in *IL28B* minor patients than in *IL28B* major patients.

Gene Expression Involving Innate Immunity and* IFNλ *in the Liver. Hepatic expression levels of cytoplasmic viral sensors (*RIG-I, MDA5*, and *LGP2*) were significantly higher in *IL28B* minor patients than in *IL28B* major patients (Fig. 1). Similarly, expressions of *ISG15* and *USP18* were significantly higher in *IL28B* minor patients than in *IL28B* major patients (Fig. 1). In contrast, the hepatic expression of the adaptor molecule (*IPS-1*) was significantly lower in *IL28B* minor patients than that in *IL28B* major patients (Fig. 1). Hepatic expression of *RNF125* was similar among *IL28B* genotypes (Fig. 1). *IFNλ (IL28A/B)* expression was higher in *IL28B* minor patients, but not statistically significant (Fig. 1). Because expression of *RIG-I* and *IPS-1* were negatively correlated, the expression ratio of *RIG-I/IPS-1* in *IL28B* minor patients was significantly higher than in *IL28B* major patients (Fig. 1).

Next, to assess the relationship between baseline hepatic gene expression and treatment efficacy, we compared levels of gene expression involving innate immunity and *IFNλ* based on the final virological response (Fig. 2). Overall, hepatic expressions of cytoplasmic viral sensors and the ISG15/USP18 system in NVR patients were significantly higher than those in VR patients. In a similar but opposite manner, hepatic expressions of *IPS-1* and *RNF125* in NVR patients were significantly lower than that in VR patients, and the expression of *IFNδ* was higher in NVR patients, but the differences were not statistically significant. Expression ratio of *RIG-I/IPS-1* was significantly higher in NVR patients than that in VR patients.

Because hepatic expressions of the RIG-I/IPS-1 and ISG15/USP18 systems were significantly related both to *IL28B* minor and NVR patients, *RIG-I* and *ISG15* expression levels and the *RIG-I/IPS-1* ratio between VR and NVR patients were further stratified by *IL28B* genotype (Fig. 3). Even in the subgroup of *IL28B* minor patients, the expressions of *RIG-I* and *ISG15* were significantly higher in NVR patients than those in VR patients. Similar tendencies were observed in a subgroup of *IL28B* major patients, in whom the *RIG-I/IPS-1* expression ratio was significantly higher in

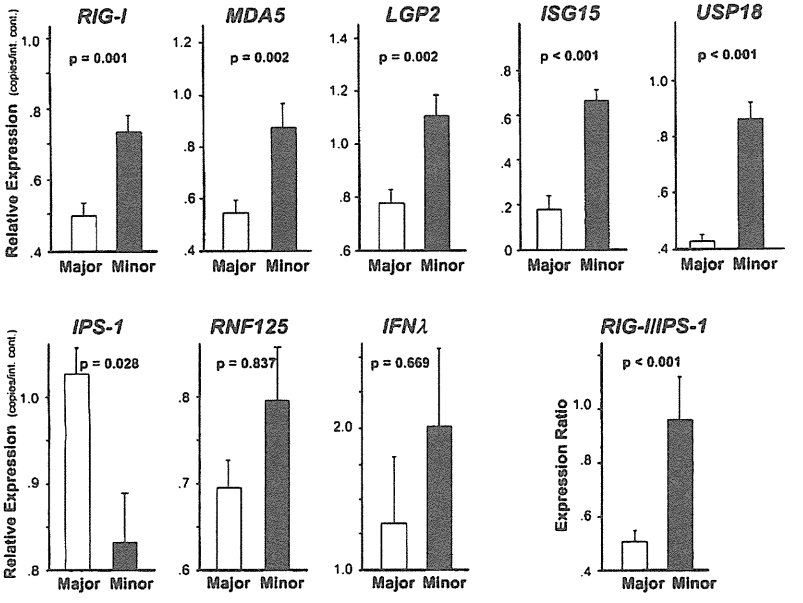

Fig. 1. Comparison of hepatic gene expression levels between *IL28B* major (rs8099917 TT/rs12979860 CC, n = 54) and *IL28B* minor patients (rs8099917 TG/rs12979860 CT, n = 34). Expression levels of cytoplasmic viral sensors (*RIG-I*, *MDA5*, and *LGP2*), modulators (*ISG15* and *USP18*), an adaptor (*IPS-1*), negative regulators (*RNF125*) and *IFNλ*, and expression ratio of the *RIG-I/IPS-1* are shown. Error bars indicate standard error. The *P*-values were determined by the Mann-Whitney *U* test.

NVR patients than in VR patients. However, in patients of the same virological response subgroup, *RIG-I* and *ISG15* expression levels and *RIG-I/IPS-1* ratio were higher in *IL28B* minor patients, and the difference in *ISG15* expression in subgroup of VR and NVR patients and that in *RIG-I/IPS-1* ratio in subgroup of VR patients was statistically significant between *IL28B* genotypes (Fig. 3).

Receiver Operator Characteristic (ROC) Analysis. To determine the usefulness of these gene quantifications and *IL28B* genotyping as predictors of NVR, an ROC analysis was conducted (Fig. 4A). The area under the ROC curve for *RIG-I* and *ISG15* expressions and *RIG-I/IPS-1* expression ratio was 0.712, 0.782, and 0.732, respectively, suggesting that quantification of these gene transcripts is useful for

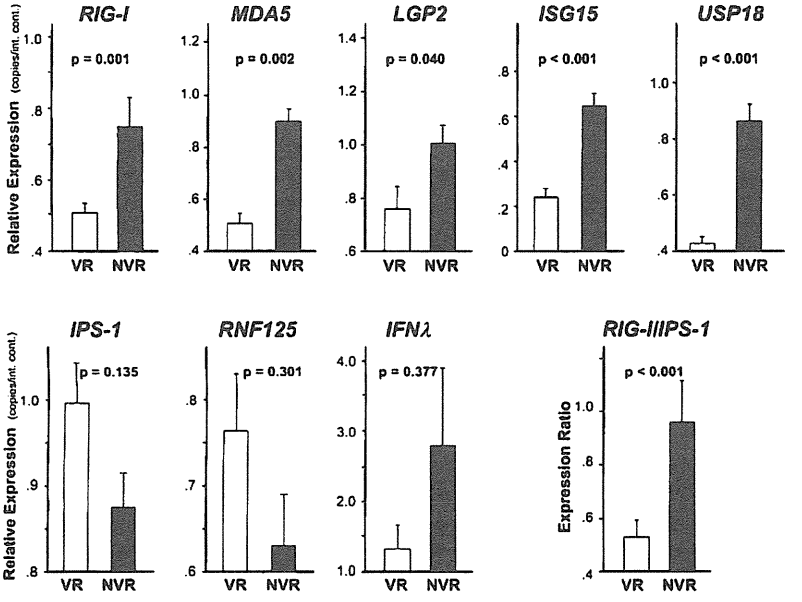

Fig. 2. Comparison of hepatic gene expression levels between virological responders (VR, n = 60) and nonvirological responders (NVR, n = 28). Expression levels of cytoplasmic viral sensors (*RIG-I*, *MDA5*, and *LGP2*), modulators (*ISG15* and *USP18*), an adaptor (*IPS-1*), negative regulators (*RNF125*) and *IFNλ*, and *RIG-I/IPS-1* expression ratio are shown. Error bars indicate standard error. The *P*-values were determined by the Mann-Whitney *U* test.

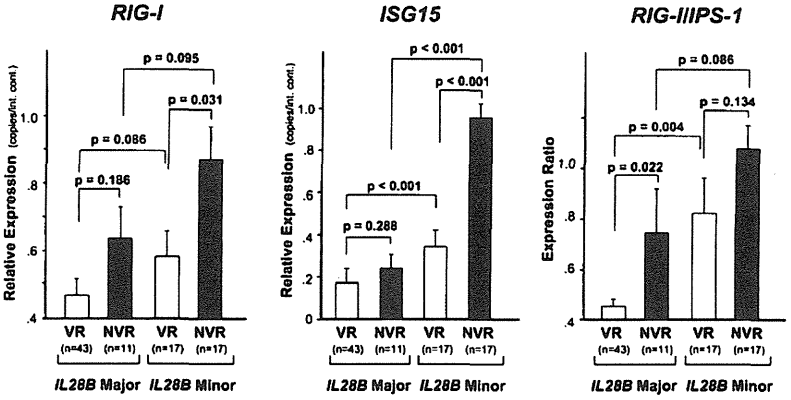

Fig. 3. Comparison of hepatic gene expression levels between virological responders (VR) and nonvirological responders (NVR) in subgroups of the *IL28B* genotype (*IL28B* Major, rs8099917 TT/rs12979860 CC; *IL28B* Minor, rs8099917 TG/rs12979860 CT). Expressions of *RIG-I* and *ISG15* as well as the *RIG-I/IPS-1* expression ratio are shown. Error bars indicate standard error. The numbers of patients in each subgroup are shown in the bottom of the figure.

prediction of NVR (Table 2). The area under the ROC curve for *IL28B* genotype was 0.662, which was lower compared with that for *RIG-I* and *ISG15* expressions and *RIG-I/IPS-1* ratio.

When we stratified the patients by the cutoff value for *RIG-I* and *ISG15* expressions and *RIGI/IPS-1* ratio, no statistically significant difference was found in NVR rates among *IL28B* genotypes within the same subgroup (Fig. 4B).

Factors Associated with NVR. In univariate analysis, age, platelet counts, double mutation at amino acid positions 70 and 91 of the HCV core region, *IL28B* minor allele, and hepatic expressions of *RIG-I, MDA5, LGP2, ISG15,* and *USP18,* and *RIG-I/IPS-1* ratio were significantly

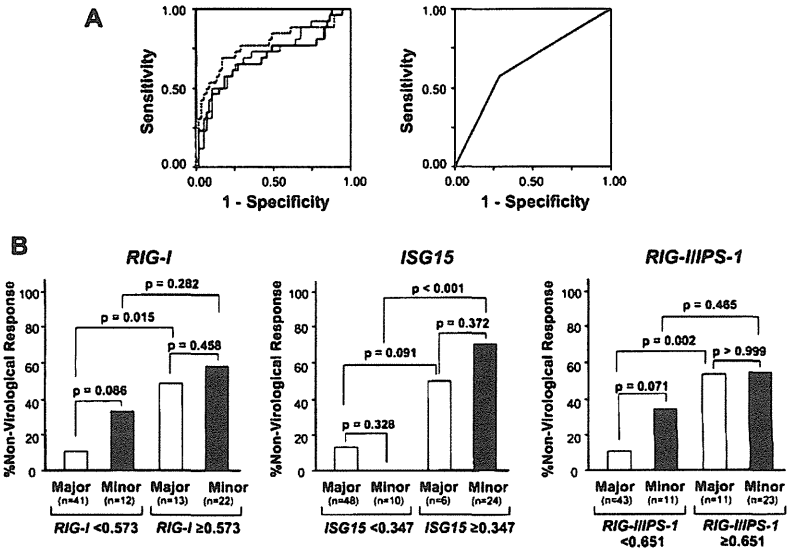

Fig. 4. (A) Receiver operator characteristics (ROC) curve for prediction of nonvirological response. ROC curves were generated to compare *RIG-I* (black line), *ISG15* (dotted line), and *RIG-I/IPS-1* ratio (gray line) (all in the left panel), and *IL28B* genotype (in the right panel). (B) Nonvirological response rate in *IL28B* major (rs8099917 TT/rs12979860 CC) and minor patients (rs8099917 TG/rs12979860 CT) in subgroups divided by the cutoff value of *RIG-I* and *ISG15* expression and the *RIG-I/ISG15* ratio determined by ROC analysis. Cutoff values of *RIG-I* and *ISG15* expression are expressed as expression copy number normalized to the expression of an internal control. The numbers of patients in each subgroup are shown in the bottom of the figure.

Table 2. Area Under the ROC Curves, Sensitivity, Specificity, and Negative as Well as Positive Predictive Values of Nonvirological Responses

Variables	AUC	95% CI	Cutoff	Sensitivity	Specificity	NPV	PPV
RIG-I (copies/int. control)	0.712	0.584-0.840	0.573	0.679	0.733	0.830	0.543
ISG15 (copies/int. control)	0.782	0.666-0.899	0.347	0.714	0.833	0.862	0.667
RIG-I/IPS-1 (copies/int. control)	0.732	0.611-0.852	0.651	0.679	0.750	0.833	0.559
IL28B genotype	0.662	0.537-0.787	TG*/CT†	0.607	0.717	0.796	0.500

AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.
*Genotype at rs8099917.
†Genotype at rs12979860.

associated with NVR (Table 3). Among these, multivariate analysis identified old age, HCV core double mutant, and higher hepatic expressions of *RIG-I* and *ISG15* as factors independently associated with NVR (Table 3).

IPS-1 and RIG-I Protein Expression in the Liver. Western blotting revealed that full-length and cleaved IPS-1 were variably present in all the samples from CH-C patients (Fig. 5A). Similar to mRNA

Table 3. Factors Associated with Nonvirological Response

	Univariate Analysis		Multivariate Analysis*	
Factors	Risk Ratio (95% CI)	*P*-value	Risk Ratio (95% CI)	*P*-value
Age (by every 10 year)	1.84 (1.10-3.14)	0.027	3.76 (1.19-11.7)	0.023
Sex				
Male	1			
Female	1.62 (0.59-4.42)	0.350		
BMI (by every 5 kg/m^2)	0.87 (0.46-1.65)	0.672		
Fibrosis stage				
F1/F2	1			
F3/F4	1.82 (0.69-4.85)	0.228		
Degree of steatosis				
<10%	1			
≥10%	1.46 (0.43-5.03)	0.544		
Albumin (by every 1 g/dL)	0.41 (0.11-1.56)	0.190		
AST (by every 40 IU/L)	0.89 (0.53-1.56)	0.681		
ALT (by every 40 IU/L)	0.85 (0.57-1.32)	0.481		
γ-GTP (by every 40 IU/L)	1.32 (0.82-2.07)	0.235		
Fasting blood sugar (by every 100 mg/dL)	1.35 (0.74-2.45)	0.340		
Hemoglobin (by every 1 g/dL)	0.93 (0.67-1.31)	0.683		
Platelet counts (by every 10^4/μL)	0.90 (0.82-0.99)	0.037	0.92 (0.78-1.08)	0.296
HCV load (by every 100 KIU/mL)	1.00 (1.00-1.00)	0.688		
Core 70 & 91 double mutation				
Wild	1		1	
Mutant	3.92 (1.14-13.5)	0.030	11.1 (1.40-88.7)	0.023
ISDR				
Nonwildtype	1			
Wildtype	1.38 (0.13-3.61)	0.513		
IL28B genotype				
Major allele†	1		1	
Minor allele‡	3.91 (1.52-10.0)	0.005	1.53 (0.20-11.9)	0.684
Hepatic gene expression (by every 0.1 copy/*int. control*)				
RIG-I	1.28 (1.10-1.50)	0.002	1.53 (1.07-2.22)	0.021
MDA5	1.53 (1.12-2.00)	0.001		
LGP2	1.34 (1.04-1.74)	0.026		
IPS-1	0.90 (0.78-1.04)	0.143		
RNF125	0.93 (0.83-1.04)	0.204		
ISG15	1.37 (1.16-1.62)	<0.001	1.28 (1.04-1.58)	0.021
USP18	1.67 (1.27-2.20)	<0.001		
IFNλ	1.02 (0.99-1.05)	0.170		
RIG-I/IPS-1 ratio (by every 0.1)	1.21 (1.07-1.36)	0.002		

Risk ratios for nonvirological response were calculated by the logistic regression analysis. BMI, body mass index; AST, aspartate aminotransferase; ALT, alanine aminotransferase; γ-GTP, gamma-glutamyl transpeptidase; HCV, hepatitis C virus; ISDR, IFN sensitivity determining region.
*Multivariate analysis was performed with factors significantly associated with nonvirological response by univariate analysis except for *MDA5*, *LGP2*, *USP18*, and *RIG-I/IPS-1* ratio, which were significantly correlated with *RIG-I* and *ISG15*.
†rs8099917 TT and rs12979860 CC.
‡rs8099917 TG and rs12979860 CT.