

- [32] Asahina Y, Tsuchiya K, Tamaki N, Hirayama I, Tanaka T, Sato M, et al. Effect of aging on risk for hepatocellular carcinoma in chronic hepatitis C virus infection. *Hepatology* 2010;52:518-527.
- [33] Ganne-Carrie N, Chastang C, Chapel F, Munz C, Pateron D, Sibony M, et al. Predictive score for the development of hepatocellular carcinoma and additional value of liver large cell dysplasia in Western patients with cirrhosis. *Hepatology* 1996;23:1112-1118.
- [34] Nishiguchi S, Kuroki T, Nakatani S, Morimoto H, Takeda T, Nakajima S, et al. Randomised trial of effects of interferon-alpha on incidence of hepatocellular carcinoma in chronic active hepatitis C with cirrhosis. *Lancet* 1995;346:1051-1055.
- [35] Shiratori Y, Ito Y, Yokosuka O, Imazeki F, Nakata R, Tanaka N, et al. Antiviral therapy for cirrhotic hepatitis C: association with reduced hepatocellular carcinoma development and improved survival. *Ann Intern Med* 2005;142:105-114.
- [36] Hezode C, Forestier N, Dusheiko G, Ferenci P, Pol S, Goefer T, et al. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N Engl J Med* 2009;360:1839-1850.
- [37] 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;360:1827-1838.

Please cite this article in press as: Kurosaki M et al. Data mining model using simple and readily available factors could identify patients at high risk for hepatocellular carcinoma in chronic hepatitis C. *J Hepatol* (2011), doi:10.1016/j.jhep.2011.09.011

## Original article

# Relationship between polymorphisms of the inosine triphosphatase gene and anaemia or outcome after treatment with pegylated interferon and ribavirin

Masayuki Kurosaki<sup>1</sup>, Yasuhito Tanaka<sup>2</sup>, Keisuke Tanaka<sup>1</sup>, Yuichiro Suzuki<sup>1,3</sup>, Yoshihide Hoshioka<sup>1</sup>, Nobuharu Tamaki<sup>1</sup>, Tomoji Kato<sup>1</sup>, Yutaka Yasui<sup>1</sup>, Takanori Hosokawa<sup>1</sup>, Ken Ueda<sup>1</sup>, Kaoru Tsuchiya<sup>1</sup>, Teiji Kuzuya<sup>1</sup>, Hiroyuki Nakanishi<sup>1</sup>, Jun Itakura<sup>1</sup>, Yuka Takahashi<sup>1</sup>, Yasuhiro Asahina<sup>1</sup>, Kentaro Matsuura<sup>2</sup>, Fuminaka Sugauchi<sup>2</sup>, Nobuyuki Enomoto<sup>3</sup>, Nao Nishida<sup>4</sup>, Katsushi Tokunaga<sup>4</sup>, Masashi Mizokami<sup>5</sup>, Namiki Izumi<sup>1\*</sup>

<sup>1</sup>Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan

<sup>2</sup>Department of Virology, Liver Unit, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan

<sup>3</sup>First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan

<sup>4</sup>Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

<sup>5</sup>Research Center for Hepatitis and Immunology, International Medical Center of Japan Konodai Hospital, Ichikawa, Japan

\*Corresponding author e-mail: nizumi@musashino.jrc.or.jp

**Background:** A genome-wide association study revealed an association between variants of the inosine triphosphatase (*ITPA*) gene and ribavirin (RBV)-induced anaemia. The aim of this study was to replicate this finding in an independent Japanese cohort and to define a method to allow pretreatment prediction of anaemia in combination with other factors.

**Methods:** Genotype 1b chronic hepatitis C patients ( $n=132$ ) treated with pegylated interferon (PEG-IFN)- $\alpha$  and RBV for 48 weeks were genotyped for *ITPA* rs1127354 and examined for anaemia and treatment outcome.

**Results:** Variants of the *ITPA* gene protected against severe anaemia throughout the 48-week treatment period and were associated with lower incidence of anaemia-related RBV dose reduction. A combination of the *ITPA* genotype with baseline haemoglobin (Hb)

and creatinine clearance (CLcr) levels predicted severe anaemia with high accuracy (90% sensitivity and 62% specificity). Among a subset of patients with the *IL28B* genotype of TT at rs8099917, patients with variants of the *ITPA* gene were associated with a higher rate of receiving >80% of the expected RBV dose, a higher rate of sustained virological response (SVR), and a lower rate of relapse.

**Conclusions:** The variants of the *ITPA* gene, which could protect against haemolytic anaemia and RBV dose reduction, were associated with a high rate of SVR by standard PEG-IFN and RBV therapy in a subset of Japanese patients with the favourable TT genotype at rs8099917 of *IL28B*. A combination of *ITPA* genetic polymorphisms with baseline Hb and CLcr levels further improves the predictive accuracy of severe anaemia.

## Introduction

Treatment with pegylated interferon (PEG-IFN) combined with ribavirin (RBV) is the most effective standard treatment for chronic HCV infection. Successful eradication of HCV is associated with a reduced risk of developing hepatocellular carcinoma. However, the rate of sustained virological response (SVR) is approximately 50% in patients with HCV genotype 1 [1,2]. The probability of SVR decreases when the patients become intolerant to therapy and receive <80% of the planned dose of PEG-IFN and/or RBV [3]. One of the major reasons

for intolerance to therapy is severe haemolytic anaemia induced by RBV [1]. The degree of haemolytic anaemia caused by RBV varies among individuals, and no reliable baseline predictors exist for this severe anaemia.

Recently, a genome-wide association study revealed that a single nucleotide polymorphism (SNP) at rs6051702 is strongly associated with RBV-induced haemolytic anaemia at week 4 of treatment [4]. This SNP was linked to two functional SNPs (rs1127354 and rs7270101) in the inosine triphosphatase (*ITPA*)

gene on chromosome 20, which had previously been well-characterized in studies of patients with ITPase deficiency [5–8]. Subsequent studies confirmed independently that variants of the *ITPA* gene are protective against haemolytic anaemia during the early weeks of treatment [9,10]. Furthermore, Thompson *et al.* [9] showed that the variants are protective against anaemia over the entire 48-week course of therapy and are associated with reduced requirement for an anaemia-related dose reduction of RBV. Notably, despite these protective effects, variants in the *ITPA* gene were not associated with treatment outcome [4,9] or showed only a marginal association [10].

In the present study, we aimed to replicate the association between *ITPA* genetic polymorphisms and RBV-induced anaemia in the early weeks, as well as throughout the entire course, of therapy in an independent Japanese cohort. In addition, for the general application of these genetic associations in clinical practice, we aimed to define a pretreatment prediction for severe anaemia in combination with other clinical covariates.

## Methods

### Patients

Data were collected retrospectively from a total of 132 genotype 1b chronic hepatitis C patients who were treated with PEG-IFN- $\alpha$  and RBV at Musashino Red Cross Hospital (Tokyo, Japan) and at Nagoya City University Graduate School of Medical Sciences (Nagoya, Japan). The inclusion criteria were: genotype 1b, HCV RNA titre >100 KIU/ml by quantitative PCR (Cobas Amplicor HCV Monitor version 2.0; Roche Diagnostic Systems, Indianapolis, IN, USA), no coinfection with HBV or HIV, no other causes of liver disease such as autoimmune hepatitis and primary biliary cirrhosis, and availability of DNA for the analysis of the genetic polymorphism of *ITPA*. Patients received PEG-IFN- $\alpha$ 2a (180  $\mu$ g) and - $\alpha$ 2b (1.5  $\mu$ g/kg) subcutaneously every week and were administered a daily weight-adjusted dose of RBV (600 mg for patients weighing <60 kg, 800 mg for patients weighing 60–80 kg, and 1,000 mg for patients weighing >80 kg) for 48 weeks. Dose reduction of RBV was considered by physicians based on the clinical conditions of the individual patients or the recommendations on the package inserts: dose reduction from 800 mg and 1,000 mg to 600 mg or from 600 mg to 400 mg for haemoglobin levels <10 g/dl and drug discontinuation when haemoglobin levels drop to <8.5 g/dl. No patient received erythropoietin or other growth factors for the treatment of anaemia. PEG-IFN and RBV was stopped prematurely in 22 patients: in 15 patients due to non-virological response and in 7 patients due to adverse events. 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 committees.

### Laboratory and histological tests

Blood samples were obtained before therapy and at 1, 2, 4, 6, 8, 12, 16, 20, 24, 36 and 48 weeks after the start of therapy, and were analysed for haematological tests, blood chemistry and HCV RNA. Genetic polymorphisms in an SNP located in exon 2 (rs1127354) and in intron 2 (rs7270101) of the *ITPA* gene were determined using ABI TaqMan Probes (Applied Biosystems, Carlsbad, CA, USA) [4]. Since a recent paper studying Japanese patients showed no variants in rs7270101 [10] and our preliminary genotyping data for 100 Japanese patients also showed no variations in rs7270101, rs1127354 was used for further analysis (major allele =C and minor allele =A). Genetic polymorphisms in the *IL28B* gene (rs8099917), an SNP recently identified to be associated with hepatitis C treatment response [11–14], was also determined by a DigiTag2 assay [15]. Viral factors affecting therapeutic efficacy was determined. A stretch of 40 amino acids in the NS5A region of HCV, designated as the interferon sensitivity-determining region (ISDR) [16,17] and amino acid substitutions at positions 70 of the core region (Core70) [18] were determined by direct sequencing after amplification by reverse transcription and PCR as reported previously. Arginine at Core70 was defined as the wild type, and glutamine or histidine was defined as the mutant type. Baseline creatinine clearance (CLcr) levels were calculated using the formula of Cockcroft and Gault [19]: for males,  $CLcr = ([140 - \text{age in years}] \times \text{body weight in kg}) / (72 \times \text{serum creatinine in mg/dl})$  and for females,  $CLcr = 0.85 \times ([140 - \text{age in years}] \times \text{body weight in kg}) / (72 \times \text{serum creatinine in mg/dl})$ . Fibrosis was evaluated on a scale of 0–4: F0 indicates no fibrosis, F1 indicates mild fibrosis, F2 indicates moderate fibrosis, F3 indicates severe fibrosis and F4 indicates cirrhosis according to the Metavir scoring system [20]. The end of treatment response was defined as an undetectable HCV RNA level by qualitative PCR with a lower detection limit of 50 IU/ml (Amplicor; Roche Diagnostic Systems) at the end of therapy. SVR was defined as an undetectable HCV RNA level 24 weeks after the completion of therapy. A relapse was defined as the reappearance of HCV RNA after the completion of therapy.

### Statistical analysis

We analysed the association between an SNP of the *ITPA* gene (rs1127354) and the following: the incidence of haemoglobin (Hb) reduction of >3.0 g/dl at week 4 and the incidence of severe anaemia (Hb<10 g/dl) at week 4 or at any time point during the therapy; the time-dependent decrease in Hb levels throughout

the treatment period; the time-dependent requirement for RBV dose reduction throughout the treatment period; and the rate of virological response or relapse. Associations between pretreatment variables and anaemia were analysed by multivariable regression. The association between the *ITPA* polymorphisms and anaemia or treatment outcome was analysed by Fisher's exact test. The association between the *ITPA* polymorphisms and the time-dependent reduction in Hb levels or the requirement for RBV dose reduction was analysed by Kaplan–Meier survival analysis. SPSS software version 15.0 (SPSS Inc., Chicago, IL, USA) was used for these analyses.

Table 1. Clinical characteristics of the study population

Characteristic	Value
Age, years	57.5 (±9.5)
Sex, male/female	50/82
Baseline platelet count, 10 <sup>9</sup> /l	150.4 (±55.8)
Baseline Hb, g/dl	14.0 (±1.5)
Baseline creatinine clearance, ml/min	94.8 (±24.1)
Baseline liver fibrosis, F0–2/F3–4	102/30
Initial ribavirin dose	
600 mg/day, n (%)	91 (69)
800 mg/day, n (%)	38 (29)
1,000 mg/day, n (%)	3 (2)
Dose reduction of ribavirin, n (%)	58 (43)
Hb reduction at week 4, g/dl	2.2 (±1.4)
Hb reduction >3.0 g/dl at week 4, n (%)	37 (28)
Severe anaemia at week 4, n (%) <sup>a</sup>	21 (16)
Severe anaemia at any time point, n (%) <sup>a</sup>	57 (43)
<i>ITPA</i> rs1127354, AA/CA/CC	4/33/95
ISDR mutation ≤1, n/total n (%)	96/114 (84)
Core70 mutant type, n/total n (%)	42/105 (40)

Continuous variables were described as mean (±sd) and categorical variables were described as frequency and percentage. <sup>a</sup>Severe anaemia defined as haemoglobin (Hb) <10 g/dl. Core70, amino acid substitutions at position 70 of the core region; ISDR, interferon sensitivity-determining region; *ITPA*, inosine triphosphatase gene.

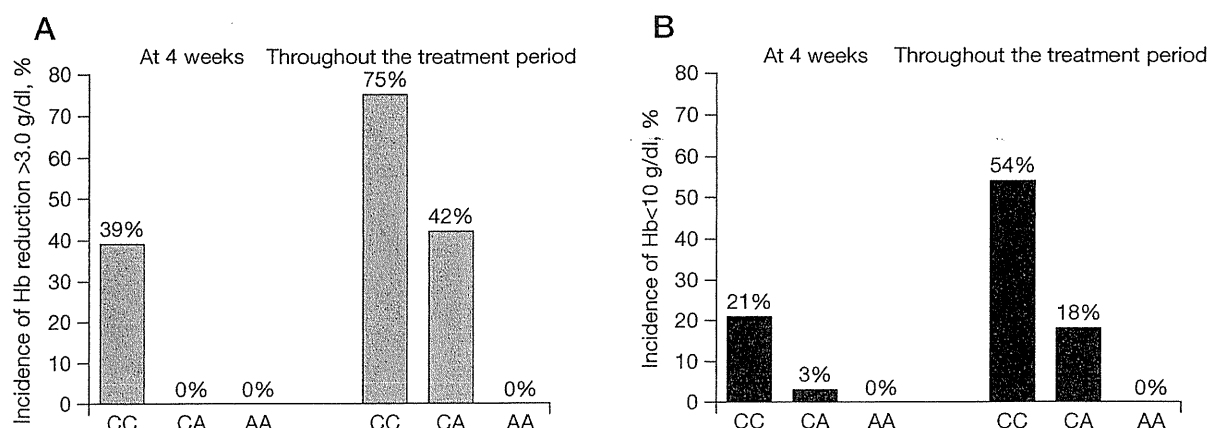
## Results

*ITPA* rs1127354 minor genotype alleles AA and CA were protective for anaemia during drug therapy. The baseline characteristics are listed in Table 1. Genotyping of rs1127354 revealed that 4 patients were homozygous for the minor allele (AA), 95 were homozygous for the major allele (CC) and 33 were heterozygous (CA). The frequency of the minor allele A was 0.16. The *ITPA* genotype was not associated with any baseline factors including age, gender, Hb levels, CLcr, platelet counts, liver fibrosis, mutations in the ISDR and Core70 (Table 2). The mean value of Hb reduction at week 4 was 2.2 g/dl and a reduction of >3.0 g/dl developed in 37 patients (28%) at week 4. Severe anaemia (Hb <10 g/dl) developed in 21 (16%) patients at week 4 of therapy and in 57 (43%) patients at any time point during the entire 48 weeks of therapy. Figure 1A and 1B shows the percentages of patients with anaemia according to the rs1127354 genotypes. At week 4, Hb reduction of >3.0 g/dl developed in 37 patients (39%) with the CC genotype, which is in contrast to 0 patients with the CA or AA genotypes (Figure 1A). Severe anaemia developed in 20 (21%) patients with the CC genotype, which is in contrast to only 1 (3%) patient with the CA genotype and 0 patients with the AA genotype (CC versus AA/CA,  $P=0.008$ ; Figure 1B). Throughout the course of the 48-week therapy, Hb reduction of >3.0 g/dl developed in 71 (75%) patients with the CC genotype in contrast to 14 (42%) patients with the CA genotype and 0 patients with the AA genotype (CC versus AA/CA,  $P=0.0001$ ). Severe anaemia was observed in 51 (54%) patients with the CC genotype, which is in contrast to 6 (18%) patients with the CA genotype and 0 patients with the AA genotype (CC versus AA/CA,  $P<0.0001$ ). The mean reduction of Hb levels and the time course of therapy are shown in Figure 2. Patients with genotypes AA and CA showed less Hb reduction at weeks 2, 4, 6, 8 and 12 during drug therapy compared to those with the

Table 2. Clinical characteristics of patients according to *ITPA* genotype

Characteristic	rs1127354		P-value
	AA/CA	CC	
Age, n (%)	56.0 (10.9)	58.1 (8.8)	0.316
Sex, male/female	17/20	33/62	0.239
Baseline platelet count, 10 <sup>9</sup> /l	153.3 (±48.5)	149.2 (±58.5)	0.711
Baseline Hb, g/dl	14.3 (±1.4)	13.8 (±1.5)	0.132
Baseline creatinine clearance, ml/min	93.4 (±23.3)	95.3 (±24.5)	0.692
Baseline liver fibrosis, F0–2/F3–4	33/4	69/26	0.063
ISDR mutation ≤1, n/total n (%)	26/30 (87)	70/84 (83)	0.777
Core70 mutant type, n/total n (%)	11/27 (41)	31/78 (40)	1.000

Continuous variables were described as mean (±sd) and categorical variables were described as frequency and percentage. Core70, amino acid substitutions at position 70 of the core region; Hb, haemoglobin; ISDR, interferon sensitivity-determining region.

Figure 1. *ITPA* rs1127354 genotypes and anaemia during drug therapy

The percentage of patients with (A) haemoglobin (Hb) reduction of >3.0 g/dl or (B) Hb concentrations of <10 g/dl at week 4 and at any time point throughout the treatment period is shown for rs1127354 genotypes. Severe anaemia was less frequent in patients with the rs1127354 genotypes AA and CA (Hb reduction >3.0 g/dl at any time point: CC versus AA/CA,  $P=0.0001$ ; Hb concentrations <10 g/dl at week 4: CC versus AA/CA,  $P=0.008$ ; and Hb concentrations <10 g/dl at any time point: CC versus AA/CA,  $P<0.0001$ ). *ITPA*, inosine triphosphatase gene.

CC genotype ( $P<0.0001$  for weeks 2, 4 and 6;  $P=0.02$  for weeks 8 and 12). These results show that the AA and CA genotypes are significantly associated with less absolute reduction in Hb levels, especially during the early weeks of therapy, and are protective against the development of severe anaemia. The sensitivity and specificity of the *ITPA* genotype for the prediction of severe anaemia (Hb<10 g/dl) throughout the course of treatment was 89% (51/57) and 41% (31/75), respectively.

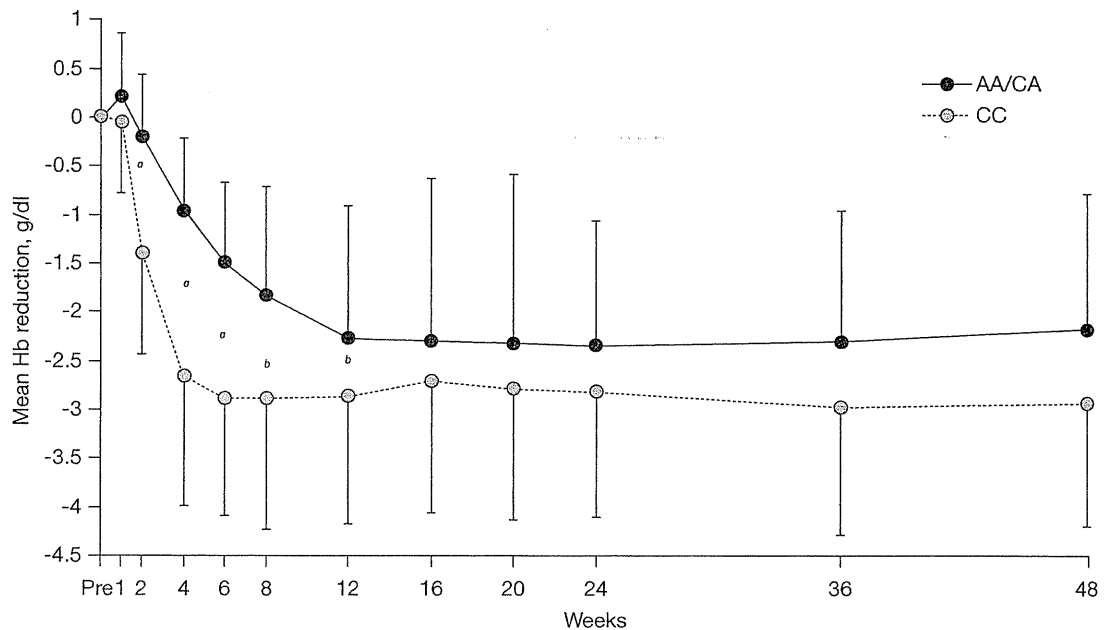
*ITPA* rs1127354 minor genotypes AA and CA were protective against the requirement for RBV dose reduction

The dose of RBV was reduced in 58 (43%) patients. Severe anaemia was the indication for dose reduction in 45 of the 58 (78%) patients. In the remaining 13 patients, the RBV dose was reduced because of other adverse events such as fatigue, skin eruption or loss of appetite. Figure 3 shows the time to the first RBV dose reduction during the 48 weeks of therapy. A dose reduction of RBV for any reason was less frequent and delayed in patients with the AA and CA genotypes compared to those with the CC genotype (Figure 3A;  $P=0.048$ ). The difference was more significant for anaemia-related RBV dose reduction (Figure 3B;  $P=0.004$ ).

Other factors associated with severe anaemia during therapy

Since 18% of the patients with the protective *ITPA* genotype of CA developed severe anaemia, we analysed the patients for other predictive factors of severe

anaemia. By univariable analysis, the rs1127354 CC genotype, female gender, older age, and lower baseline Hb levels, platelet counts and CLcr levels were associated with severe anaemia. Next, multivariable regression models with backward selection were used to identify the independent predictors of severe anaemia. Covariates included age, sex, fibrosis stage, baseline Hb levels, CLcr levels and platelet counts, and the rs1127354 genotype. The multivariable regression analysis showed that the rs1127354 CC genotype, a baseline Hb of <14 g/dl and a baseline CLcr of  $\leq 95$  ml/min were independent predictors of severe anaemia at week 4 and at any time point during the 48 weeks of therapy (Table 3). Figure 4 shows the percentage of patients with Hb concentrations of <10 g/dl at any time point during therapy for the subgroups of patients stratified by rs1127354 genotype, baseline Hb levels and baseline CLcr levels. Among patients with the rs1127354 CC genotype, the risk of developing severe anaemia was more prominent in those with a baseline Hb<14 g/dl and a baseline CLcr $\leq 95$  ml/min (88%) compared to those with a baseline Hb $\geq 14$  g/dl and a baseline CLcr >95 ml/min ( $P<0.0001$ ) or those with a baseline Hb<14 g/dl or a baseline CLcr $\leq 95$  ml/min ( $P=0.0036$ ). Notably, the incidence of severe anaemia was only 12% in patients with the rs1127354 CC genotype if the baseline Hb was  $\geq 14$  g/dl and the CLcr was >95 ml/min. By contrast, there was a moderate risk of severe anaemia (33%) even in patients with the rs1127354 protective genotypes AA or CA when the baseline Hb was <14 g/dl and the baseline CLcr was  $\leq 95$  ml/min. Thus, patients who have >30%

Figure 2. *ITPA* rs1127354 genotypes and the quantitative Hb reduction from baseline

The mean reduction of haemoglobin (Hb) levels along the time points of treatment is shown for the rs1127354 genotypes. Solid and dotted lines indicate patients with the AA/CA and CC genotypes, respectively. The error bars indicate standard deviation. The AA/CA genotype had less of a reduction in the mean Hb levels at weeks 2–12 during therapy compared to the CC genotype. \* $P < 0.001$ ; <sup>b</sup> $P = 0.02$ . *ITPA*, inosine triphosphatase gene; Pre, pretreatment.

risk of severe anaemia had the following characteristics: rs1127354 CC genotype, baseline Hb < 14 g/dl and CLCr ≤ 95 ml/min; rs1127354 CC genotype and baseline Hb < 14 g/dl or CLCr ≤ 95 ml/min; and rs1127354 AA or CA genotype, baseline Hb < 14 g/dl and CLCr ≤ 95 ml/min. The sensitivity and specificity of the combination of these three factors for the prediction of severe anaemia (Hb < 10 g/dl) throughout the course of treatment was 89% (51/57) and 64% (48/75). Compared to the *ITPA* genotype alone, specificity improved from 41% to 64% with the same sensitivity (89%), indicating that the combination of the *ITPA* genotype, baseline Hb levels and baseline CLCr levels could improve the prediction accuracy. The AA/CA genotypes of rs1127354 were protective against the requirement for RBV dose reduction even after standardization by baseline Hb and CLCr (Figure 3C). The predictive model for anaemia and recommendations for monitoring and treatment were made for clinical practice application (Table 4).

*ITPA* rs1127354 minor genotypes AA and CA were associated with higher adherence to RBV, higher rate of SVR and lower rate of relapse

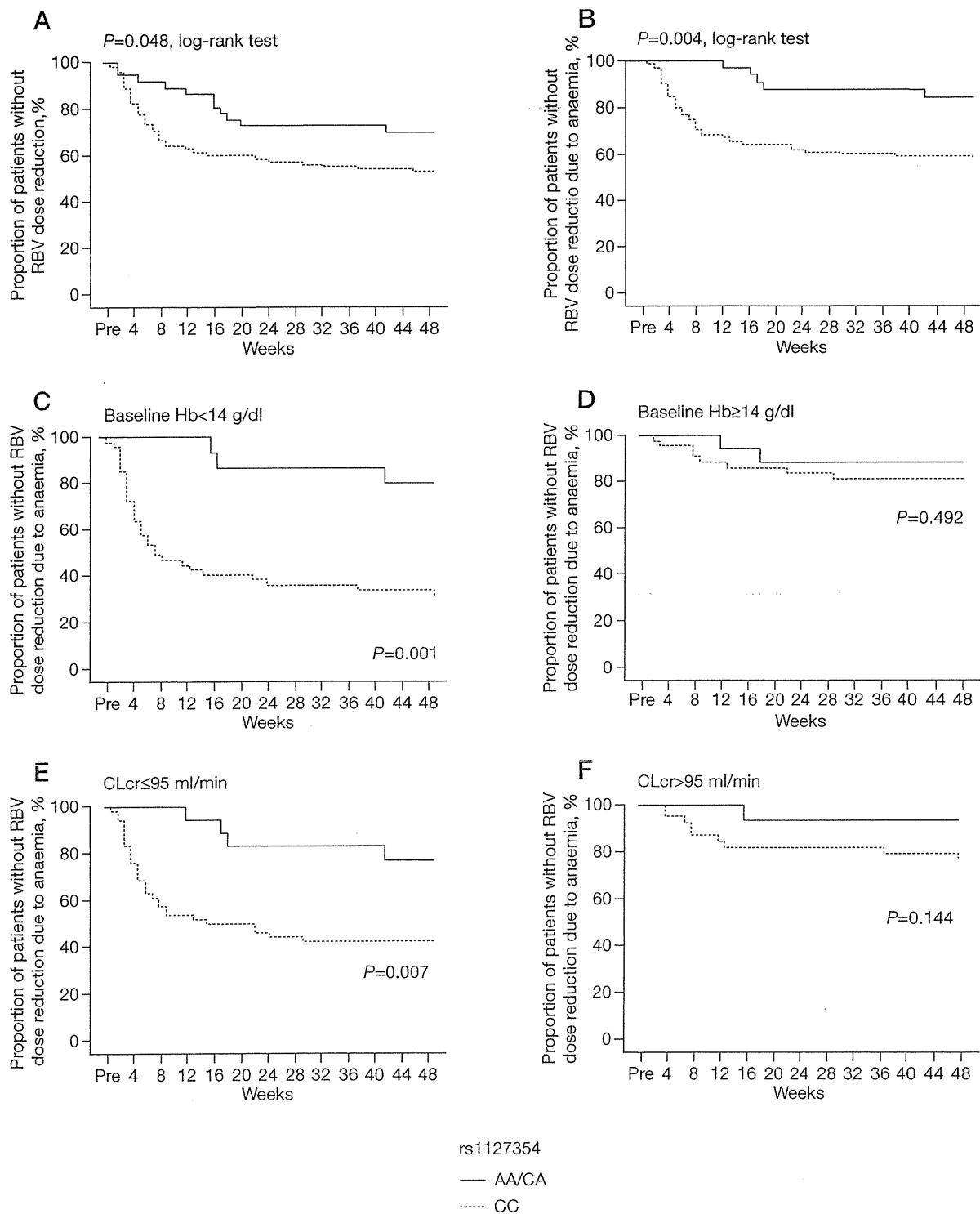
The association of the rs1127354 genotype with the adherence to RBV or treatment outcome was analysed. When analysed in the entire population, the percentage

of patients receiving >80% of the expected RBV dose, which was reported to be a threshold for an enhanced response to therapy [3], was not significantly different among the rs1127354 genotypes. Treatment outcomes such as the end-of-treatment response, SVR and relapse were also not different among the rs1127354 genotypes (Table 5). By contrast, SVR was closely associated with the *IL28B* genotype [11–14,21]: the rate of SVR was 0% (0/51) for *IL28B* minor type (TG/GG genotype at rs8099917) and 48% (39/81) for *IL28B* major type (TT genotype at rs8099917). This finding confirms that *IL28B* genotype is a significant factor for the prediction of SVR. Thus, we performed a subset analysis on subgroup of patients with the favourable *IL28B* genotype (TT at rs8099917). As a result, patients with the rs8099917 TT genotype and the rs1127354 AA or CA genotypes had a significantly higher rate of receiving >80% of the expected RBV dose ( $P = 0.016$ ), a higher rate of SVR ( $P = 0.031$ ), as well as a lower rate of relapse ( $P = 0.046$ ) compared to patients with the rs8099918 TT and rs1127354 CC genotype (Table 5).

## Discussion

In the present study, we confirmed that variants of the *ITPA* gene protect against severe haemolytic anaemia not

Figure 3. *ITPA* rs1127354 genotypes and the time-dependent incidence of RBV dose reduction



The time to the first reduction of the ribavirin (RBV) dose (A) due to any reason or (B) due to anaemia is shown stratified by the *rs1127354* genotypes. Solid and broken lines indicate patients with the AA/CA and CC genotypes, respectively. The AA/CA genotype protected against the requirement for RBV dose reduction. (C-F) Patients were standardized according to the baseline haemoglobin (Hb) and creatinine clearance (CLcr). Even after standardization by baseline Hb and CLcr, the AA/CA genotype protected against the requirement for RBV dose reduction. *ITPA*, inosine triphosphatase gene; Pre, pretreatment.

only at the early stage of treatment, but also throughout the 48-week course of treatment in a Japanese cohort of genotype 1b chronic hepatitis C patients treated with PEG-IFN and RBV. We also replicated a previous study [9] that showed that the *ITPA* genotype is significantly associated with a time-dependent reduction of the RBV dose. Furthermore, we found that a combination of the

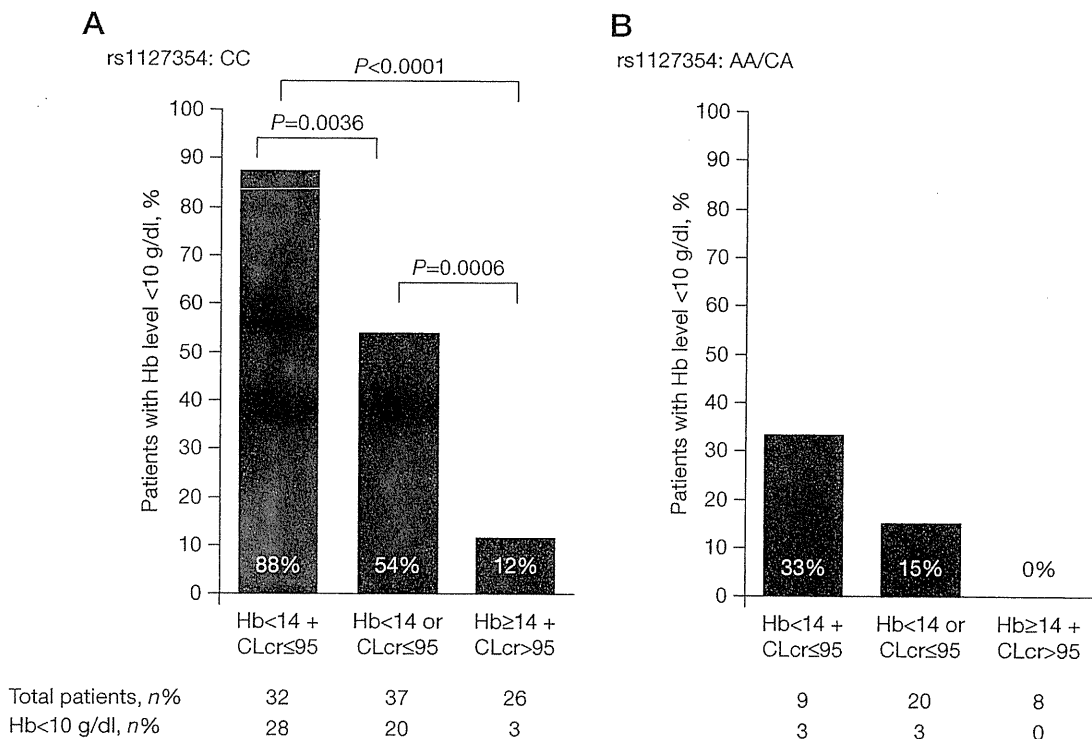
*ITPA* genotype and the baseline Hb and CLcr levels improve the accuracy of predicting RBV-induced severe anaemia. Previous reports on the IDEAL [4] or Vira-Hep-C [9] studies did not find any association between the *ITPA* genotype and treatment outcome; however, we were able to demonstrate the association of the *ITPA* genotype with a higher adherence to RBV, a higher rate

Table 3. Multivariable regression analysis of factors associated with severe anaemia during therapy<sup>a</sup>

Predictor	OR	95% CI	P-value
<b>At week 4</b>			
Baseline Hb<14 g/dl	7.18	1.90–27.09	0.004
Baseline creatinine clearance ≤95 ml/min	5.30	1.39–20.26	0.015
<i>ITPA</i> rs1127354: CC	10.17	1.25–82.85	0.030
<b>At any time point</b>			
Baseline Hb<14 g/dl	7.67	3.07–19.12	<0.0001
Baseline creatinine clearance ≤95 ml/min	5.51	2.21–13.73	<0.0001
<i>ITPA</i> rs1127354: CC	9.66	3.11–29.95	<0.0001

<sup>a</sup>Severe anaemia was defined as haemoglobin (Hb)<10 g/dl. *ITPA*, inosine triphosphatase gene.

Figure 4. Combination of the *ITPA* rs1127354 genotype, baseline Hb level and baseline CLcr level is predictive of severe anaemia during the therapy



Patients with rs1127354 genotype (A) CC and (B) AA/CA were further stratified by the baseline haemoglobin (Hb) and creatinine clearance (CLcr) levels. The percentage of patients with Hb concentrations of <10 g/dl (severe anaemia) at any time point during therapy is shown for the subgroups of patients. Patients with baseline Hb levels of <14 g/dl and CLcr levels of <95 ml/min had a higher incidence of severe anaemia among patients with the rs1127354 genotype CC (Hb<14 g/dl and CLcr≤95 ml/min versus Hb≥14 g/dl and CLcr>95 ml/min,  $P<0.0001$ ; Hb<14 g/dl and CLcr≤95 ml/min versus Hb<14 g/dl or CLcr≤95 ml/min,  $P=0.0036$ ). *ITPA*, inosine triphosphatase gene.



Table 4. Prediction model for severe anaemia and recommendation for monitoring and treatment

<i>ITPA</i> genotype (rs1127354)	Baseline Hb and CLcr	Risk of anaemia	Recommendation	
			Monitoring	Treatment option
CC	Hb<14 g/dl and CLcr≤95 ml/min	High	Intensive	Consider erythropoietin
	Hb<14 g/dl or CLcr≤95 ml/min	Intermediate	Intensive	Early dose reduction of RBV
	Hb≥14 g/dl and CLcr>95 ml/min	Low	As usual	–
AA/CA	Hb<14 g/dl and CLcr≤95 ml/min	Intermediate	Intensive	Early dose reduction of RBV
	Hb<14 g/dl or CLcr≤95 ml/min	Low	As usual	–
	Hb≥14 g/dl and CLcr>95 ml/min	Absent	As usual	May consider higher RBV dose

CLcr, creatinine clearance; Hb, haemoglobin; *ITPA*, inosine triphosphatase gene; RBV, ribavirin.

Table 5. Treatment response and ribavirin adherence in terms of *ITPA* rs1127354 genotype

Response	rs1127354		<i>P</i> -value
	AA/CA, n/total n (%)	CC, n/total n (%)	
<b>All patients</b>			
Ribavirin adherence >80%	19/37 (51)	40/95 (42)	0.436
End-of-treatment response	19/37 (51)	58/95 (61)	0.332
Sustained virological response	13/37 (35)	26/95 (27)	0.401
Relapse	6/19 (32)	32/58 (55)	0.112
<b>Subgroup of patients with <i>IL28B</i> rs8099917 TT</b>			
Ribavirin adherence >80%	14/18 (78)	28/63 (49)	0.016
End of treatment response	16/18 (89)	50/63 (79)	0.501
Sustained virological response	13/18 (79)	26/63 (41)	0.031
Relapse	3/16 (19)	24/50 (48)	0.046

*ITPA*, inosine triphosphatase gene.

of SVR and a lower rate of relapse among a subset of Japanese patients with the favourable *IL28B* genotype (TT at rs8099917).

Haemolytic anaemia induced by RBV is one of the major adverse events of PEG-IFN and RBV therapy leading to dose reduction of RBV or premature termination of therapy [1]. RBV is essential for improving SVR by prevention of relapses and a breakthrough [22], and a reduction of the RBV dose can lower the response rates considerably. It was reported that the maintenance of >80% of the expected RBV dose is associated with an increased SVR [23]. Thus, the prediction and prevention of RBV-induced haemolytic anaemia is clinically important. Previously, no reliable means were available to predict RBV-induced anaemia before therapy, but a recent genome-wide association study identified a strong association between two functional SNPs (rs1127354 and rs7270101) in the *ITPA* gene on chromosome 20 [4] and severe anaemia at week 4 of treatment. This genetic association has been replicated recently by two studies [9,10]. However, the effect of these variants on the long-term development of anaemia or on the requirement for RBV dose reduction has been reported by only one study to date [9]. Therefore, validation of these results by an independent cohort with respect to different geographical areas,

age, gender or race is needed. Although the clinical background of our cohort was different from that of the US cohort [9], such as their race, older age (mean age of 57.5 years versus the median age of 48.5 years), and higher predominance of females (62% versus 35%), we were still able to replicate the results that the rs1127354 genotypes AA and CA are protective against anaemia throughout the 48-week course of treatment, especially within the 12 weeks following the initial treatment. We also replicated the association of this genotype with less requirement for RBV dose reduction. These results indicate that the *ITPA* genotype is universally an important determinant of RBV-induced haemolytic anaemia.

For the general application of these genetic associations in clinical practice, we aimed to further improve the accuracy of prediction by combining other clinical covariates. Among the patients with the rs1127354 CC genotype, the risk of developing severe anaemia was as high as 88% in those with baseline Hb levels of <14 g/dl and baseline CLcr levels of ≤95 ml/min, which is in contrast to only 12% in patients with Hb levels of ≥14 g/dl and CLcr levels of >95 ml/min. The rs1127354 AA and CA genotypes were protective against anaemia, but an exception occurred when patients (33%) with a baseline Hb level of <14 g/dl and a CLcr level of ≤95 ml/min developed severe

anaemia. The combination of these three factors may therefore be useful in clinical practice, since it improved the specificity of prediction from 41% to 64% with the same sensitivity (89%) compared to examining just the *ITPA* genotype. These findings may have the potential to support individualized treatment strategies. Patients with the rs1127354 CC genotype, especially those with a baseline Hb level of <14 g/dl and a baseline CLcr level of  $\leq 95$  ml/min, require intensive monitoring for anaemia during therapy, and an early dose reduction of RBV or support by erythropoietin may be indicated for safety. By contrast, patients with the AA and CA genotypes, excluding those with a baseline Hb level of <14 g/dl and a baseline CLcr level of  $\leq 95$  ml/min, may be candidates for therapy with a higher RBV dose, which may lead to higher rates of SVR. The prediction of RBV-induced anaemia will remain an important issue even in the near future, since direct antiviral agents require RBV and PEG-IFN in combination in order to achieve higher SVR rates for genotype 1 [24,25] and this combination will remain a standard therapy for other genotypes.

In a previous study, there was no clear association between ITPase deficiency and treatment outcome [4,9,10], even after a detailed subset analysis that excluded patients in whom RBV had been reduced for indications other than anaemia or after stratification by the *IL28B* genotype [9]. Thompson *et al.* [9] speculated that the lack of association may derive from several reasons such as an underpowered error due to the small number of patients, a high incidence of RBV dose reduction unrelated to anaemia, and the possibility that the ITPase deficiency may reduce antiviral efficacy. In the present study, we also failed to show associations between the *ITPA* genotype and treatment outcomes among the entire cohort. However, when patients were stratified by the *IL28B* genotype, which is now recognized as the major determinant of treatment outcome [11–14,21], the AA and CA genotypes at rs1127354 were linked to a higher adherence to RBV, a lower rate of relapse and a significantly higher rate of SVR. One of the reasons for this discrepancy may be the lower incidence of anaemia-unrelated RBV dose reduction in our study compared to the participants of the Vira-Hep-C study (22% versus 48%) [9]. The effect of the *ITPA* genotype on RBV adherence and treatment outcome may be less apparent in patients who reduced their RBV dose in the absence of anaemia. Another possibility is that the difference in mean age may have some effect on this association between the *ITPA* genotype and treatment outcome since older age has been reported to compromise drug adherence or treatment outcomes [26,27]. Our results indicated that, although *IL28B* genotype is the major determinant of SVR, the *ITPA* genotype may be used supplementary to predict the treatment outcome in patients with a favourable *IL28B* genotype (TT at

rs8099917), as long as the RBV dose is not reduced in the absence of anaemia. Further studies involving larger populations in different geographical areas or races may be necessary to confirm this speculation.

In conclusion, variants of the *ITPA* gene, which could protect against haemolytic anaemia and RBV dose reduction, were associated with a high rate of SVR by standard PEG-IFN and RBV therapy in a subset of Japanese patients with the favourable *IL28B* genotype. A combination of the *ITPA* genetic polymorphism with baseline Hb and CLcr levels further improved the predictive accuracy of severe anaemia. These findings may have the potential to support selection of the optimum and personalized treatment strategy for individual patients.

### Acknowledgements

This study was supported by a grant-in-aid from the Ministry of Health, Labor and Welfare, Japan.

### Disclosure statement

The authors declare no competing interests.

### References

1. Fried MW, Shiffman ML, Reddy KR, *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 2002; 347:975–982.
2. 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; 358:958–965.
3. McHutchison JG, Manns M, Patel K, *et al.* Adherence to combination therapy enhances sustained response in genotype-1-infected patients with chronic hepatitis C. *Gastroenterology* 2002; 123:1061–1069.
4. Fellay J, Thompson AJ, Ge D, *et al.* *ITPA* gene variants protect against anaemia in patients treated for chronic hepatitis C. *Nature* 2010; 464:405–408.
5. Sumi S, Marinaki AM, Arenas M, *et al.* Genetic basis of inosine triphosphate pyrophosphohydrolase deficiency. *Hum Genet* 2002; 111:360–367.
6. Maeda T, Sumi S, Ueta A, *et al.* Genetic basis of inosine triphosphate pyrophosphohydrolase deficiency in the Japanese population. *Mol Genet Metab* 2005; 85:271–279.
7. Shipkova M, Lorenz K, Oellerich M, Wieland E, von Ahsen N. Measurement of erythrocyte inosine triphosphate pyrophosphohydrolase (*ITPA*) activity by HPLC and correlation of *ITPA* genotype-phenotype in a Caucasian population. *Clin Chem* 2006; 52:240–247.
8. Atanasova S, Shipkova M, Svinarov D, *et al.* Analysis of *ITPA* phenotype-genotype correlation in the Bulgarian population revealed a novel gene variant in exon 6. *Ther Drug Monit* 2007; 29:6–10.
9. Thompson AJ, Fellay J, Patel K, *et al.* Variants in the *ITPA* gene protect against ribavirin-induced hemolytic anemia and decrease the need for ribavirin dose reduction. *Gastroenterology* 2010; 139:1181–1189.
10. Ochi H, Maekawa T, Abe H, *et al.* Inosine triphosphate pyrophosphatase polymorphism affects ribavirin-induced anemia and outcomes of therapy – a genome-wide study of Japanese hepatitis C virus patients. *Gastroenterology* 2010; 139:1190–1197.

11. Ge D, Fellay J, Thompson AJ, *et al.* Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009; 461:399–401.
12. Tanaka Y, Nishida N, Sugiyama M, *et al.* Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* 2009; 41:1105–1109.
13. Suppiah V, Moldovan M, Ahlenstiel G, *et al.* IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 2009; 41:1100–1104.
14. Rauch A, Kuralik Z, Descombes P, *et al.* Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* 2010; 138:1338–1345.
15. 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; 364:78–85.
16. Enomoto N, Sakuma I, Asahina Y, *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; 334:77–82.
17. Kurosaki M, Enomoto N, Murakami T, *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; 25:750–753.
18. Akuta N, Suzuki F, Sezaki H, *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; 48:372–380.
19. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron* 1976; 16:31–41.
20. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology* 1996; 24:289–293.
21. Tanaka Y, Nishida N, Sugiyama M, Tokunaga K, Mizokami M.  $\lambda$ -Interferons and the single nucleotide polymorphisms: a milestone to tailor-made therapy for chronic hepatitis C. *Hepatol Res* 2010; 40:449–460.
22. Bronowicki JP, Ouzan D, Asselah T, *et al.* Effect of ribavirin in genotype 1 patients with hepatitis C responding to pegylated interferon alfa-2a plus ribavirin. *Gastroenterology* 2006; 131:1040–1048.
23. McHutchison JG, Lawitz EJ, Shiffman ML, *et al.* Peginterferon alfa-2b or alfa-2a with ribavirin for treatment of hepatitis C infection. *N Engl J Med* 2009; 361:580–593.
24. McHutchison JG, Everson GT, Gordon SC, *et al.* Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *N Engl J Med* 2009; 360:1827–1838.
25. Kwo PY, Lawitz EJ, McCone J, *et al.* Efficacy of boceprevir, an NS3 protease inhibitor, in combination with peginterferon alfa-2b and ribavirin in treatment-naive patients with genotype 1 hepatitis C infection (SPRINT-1): an open-label, randomised, multicentre Phase 2 trial. *Lancet* 2010; 376:705–716.
26. Hiramatsu N, Oze T, Tsuda N, *et al.* Should aged patients with chronic hepatitis C be treated with interferon and ribavirin combination therapy? *Hepatol Res* 2006; 35:185–189.
27. Iwasaki Y, Ikeda H, Araki Y, *et al.* Limitation of combination therapy of interferon and ribavirin for older patients with chronic hepatitis C. *Hepatology* 2006; 43:54–63.

---

Accepted 28 November 2010; published online 18 May 2011

# Sequences in the Interferon Sensitivity-Determining Region and Core Region of Hepatitis C Virus Impact Pretreatment Prediction of Response to PEG-Interferon Plus Ribavirin: Data Mining Analysis

Masayuki Kurosaki,<sup>1</sup> Naoya Sakamoto,<sup>2</sup> Manabu Iwasaki,<sup>3</sup> Minoru Sakamoto,<sup>4</sup> Yoshiyuki Suzuki,<sup>5</sup> Naoki Hiramatsu,<sup>6</sup> Fuminaka Sugauchi,<sup>7</sup> Akihiro Tamori,<sup>8</sup> Mina Nakagawa,<sup>2</sup> and Namiki Izumi, MD<sup>1\*</sup>

<sup>1</sup>Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan

<sup>2</sup>Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan

<sup>3</sup>Department of Computer and Information Science, Seikei University, Tokyo, Japan

<sup>4</sup>First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan

<sup>5</sup>Department of Hepatology, Toranomon Hospital, Tokyo, Japan

<sup>6</sup>Department of Gastroenterology and Hepatology, Osaka University Graduate School of Medicine, Osaka, Japan

<sup>7</sup>Department of Gastroenterology and Metabolism, Nagoya City University Graduate School of Medical Sciences, Aichi, Japan

<sup>8</sup>Department of Hepatology, Osaka City University Medical School, Osaka, Japan

The aim of the present study was to clarify the significance of viral factors for pretreatment prediction of sustained virological response to pegylated-interferon (PEG-IFN) plus ribavirin (RBV) therapy for chronic hepatitis C using data mining analysis. Substitutions in the IFN sensitivity-determining region (ISDR) and at position 70 of the HCV core region (Core70) were determined in 505 patients with genotype 1b chronic hepatitis C treated with PEG-IFN plus RBV. Data mining analysis was used to build a predictive model of sustained virological response in patients selected randomly ( $n = 304$ ). The reproducibility of the model was validated in the remaining 201 patients. Substitutions in ISDR (odds ratio = 9.92,  $P < 0.0001$ ) and Core70 (odds ratio = 1.92,  $P = 0.01$ ) predicted sustained virological response independent of other covariates. The decision-tree model revealed that the rate of sustained virological response was highest (83%) in patients with two or more substitutions in ISDR. The overall rate of sustained virological response was 44% in patients with a low number of substitutions in ISDR (0–1) but was 83% in selected subgroups of younger patients (<60 years), wild-type sequence at Core70, and higher level of low-density lipoprotein cholesterol (LDL-C) ( $\geq 120$  mg/dl). Reproducibility of the model was validated ( $r^2 = 0.94$ ,  $P < 0.001$ ). In conclusion, substitutions in ISDR and Core70 of

HCV are significant predictors of response to PEG-IFN plus RBV therapy. A decision-tree model that includes these viral factors as predictors could identify patients with a high probability of sustained virological response. **J. Med. Virol.** 83:445–452, 2011.

© 2011 Wiley-Liss, Inc.

**KEY WORDS:** data mining; decision-tree model; ISDR; core region; PEG-interferon

## INTRODUCTION

The combination of pegylated-interferon (PEG-IFN) plus ribavirin (RBV) is currently the most effective therapy for chronic hepatitis C, but the rate of sustained virological response after 48 weeks of therapy is about 50% in patients with HCV genotype 1b and a high HCV

Grant sponsor: Ministry of Health, Labor and Welfare, Japan. The authors report no conflicts of interest.

\*Correspondence to: Namiki Izumi, MD, Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, 1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610, Japan. E-mail: nizumi@musashino.jrc.or.jp

Accepted 26 October 2010

DOI 10.1002/jmv.22005

Published online in Wiley Online Library (wileyonlinelibrary.com).

RNA titer [Manns et al., 2001; Fried et al., 2002]. The most reliable means to predict sustained virological response is to monitor the viral response during the early weeks of treatment. The early virological response, defined as undetectable HCV RNA at week 12, is associated with a high rate of sustained virological response [Davis et al., 2003; Lee and Ferenci, 2008]. The rapid virological response, defined as undetectable HCV RNA at week 4 of therapy, is even more predictive of sustained virological response than the early virological response [Jensen et al., 2006; Yu et al., 2008; Izumi et al., 2010]. However, there is no established means that predicts the virological response before commencing treatment. Recent reports have revealed that single nucleotide polymorphisms located near the *IL28B* gene show a strong association with the response to PEG-IFN plus RBV therapy [Ge et al., 2009; Suppiah et al., 2009; Tanaka et al., 2009; Kurosaki et al., 2010c]. These findings indicate that the host factor is an important determinant of the treatment response. On the other hand, the present study's authors have reported that a stretch of 40 amino acids in the NS5A region of HCV, designated as the interferon sensitivity-determining region (ISDR), has a close association with the virological response to interferon mono-therapy [Enomoto et al., 1995, 1996; Kurosaki et al., 1997]. More recently, amino acid substitutions at positions 70 and 91 of the core region have been reported to be associated with response to PEG-IFN plus RBV combination therapy [Akuta et al., 2005, 2007a]. The impact of these HCV substitutions on treatment response is yet to be validated.

Decision-tree analysis is a core component of data mining analysis that can be used to build predictive models [Breiman et al., 1980]. This method has been used to define prognostic factors in various diseases such as prostate cancer [Garzotto et al., 2005], diabetes [Miyaki et al., 2002], melanoma [Averbook et al., 2002; Leiter et al., 2004], colorectal carcinoma [Zlobec et al., 2005; Valera et al., 2007], and liver failure [Baquerizo et al., 2003]. The major advantage of decision-tree analysis over logistic regression analysis is that the results of analysis are easy to understand. The simple allocation of patients into subgroups by following the flowchart form could define the predicted possibility of outcome [LeBlanc and Crowley, 1995].

Decision-tree analysis was used for the prediction of early virological response (undetectable HCV RNA within 12 weeks of therapy) to PEG-IFN and RBV combination therapy in chronic hepatitis C [Kurosaki et al., 2010a], and more recently for the pretreatment prediction of sustained virological response [Kurosaki et al., 2010b]. In the latter model, simple and noninvasive standard tests were used as parameters; specialized tests such as viral mutations and host genetics, or invasive tests such as liver histology, were not included because the aim of that model was for use in general medical practice, especially in some countries or areas where resources are limited. Thus, the impact of viral mutations or liver histology was not considered in that model.

The present study examined whether including viral substitutions in ISDR and the core region of HCV in the decision-tree model could improve its predictive accuracy over the previous model to identify chronic hepatitis C patients who are likely to respond to PEG-IFN plus RBV therapy.

## MATERIALS AND METHODS

### Patients

This multicenter retrospective cohort study included 505 chronic hepatitis C patients who were treated with PEG-IFN alpha-2b and RBV at Musashino Red Cross Hospital, Toranomon Hospital, Tokyo Medical and Dental University, Osaka University, Nagoya City University Graduate School of Medical Sciences, Yamanashi University, Osaka City University, and their related hospitals. The inclusion criteria were: (1) genotype 1b, (2) HCV RNA titer higher than 100 kIU/ml by quantitative PCR (Cobas Amplicor HCV Monitor v 2.0, Roche Diagnostic Systems, Pleasanton, CA), (3) no coinfection with hepatitis B virus or human immunodeficiency virus, (4) no other causes of liver disease, (5) patients having undergone liver biopsy prior to IFN treatment, (6) number of substitutions in ISDR having been determined, (7) substitutions in the amino acid positions 70 and 91 of the core region having been determined, and (8) completion of at least 12 weeks of therapy. Patients were treated with PEG-IFN alpha-2b (1.5 µg/kg) weekly plus RBV. The daily dose of RBV was adjusted by weight: 600 mg for <60 kg, 800 mg for 60–80 kg, and 1,000 mg for >80 kg. For the analysis, patients were assigned randomly to either the model building (304 patients) or validation (201 patients) groups. There were no significant differences in the clinical backgrounds between these two groups (Table I). Informed consent was obtained from each patient. The study protocol conformed to the ethical guidelines of the Declaration of Helsinki and was approved by the institutional review committees of all concerned hospitals.

### Laboratory Tests

Hematological tests, blood chemistry, and HCV RNA titer were analyzed before therapy and at least once every month during therapy. 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. At position 70 of the core region (Core70), arginine was defined as the wild type, and glutamine or histidine was defined as the mutant type. At position 91 of the core region, leucine was defined as the wild type and methionine was defined as the mutant type, as described previously [Akuta et al., 2005]. Fibrosis and activity were scored according to the METAVIR scoring system [Bedossa and Poynard, 1996]. Fibrosis was staged on a scale of 0–4: F0 (no fibrosis), F1 (mild fibrosis), F2 (moderate fibrosis), F3 (severe fibrosis), and F4 (cirrhosis). Activity of necroinflammation was graded on a scale of

TABLE I. Comparison of Pretreatment Factors Between Model Building and Validation Patients

	Model (n = 304)	Validation (n = 201)	P-value
Age (years)	55.6 (9.4)	56.0 (12.2)	0.80
Male (%)	53 (%)	55 (%)	0.13
Body mass index (kg/m <sup>2</sup> )	23.1 (3.1)	23.1 (4.0)	0.99
Albumin (g/dl)	4.0 (0.3)	4.0 (0.3)	0.47
Creatinine (mg/dl)	0.72 (0.15)	0.72 (0.14)	0.62
AST (IU/L)	63.3 (45.6)	58.9 (46.4)	0.91
ALT (IU/L)	78.7 (58.6)	74.5 (67.5)	0.68
GGT (IU/L)	53.2 (49.1)	57.4 (63.5)	0.43
Total cholesterol (mg/dl)	170.9 (32.6)	169.4 (34.1)	0.33
Triglyceride (mg/dl)	107.0 (44.7)	105.7 (48.0)	0.90
LDL-C (mg/dl)	95.5 (28.0)	96.4 (28.8)	0.34
White blood cell count (/μl)	4,902 (1,489)	4,906 (1,319)	0.86
Hemoglobin (g/dl)	14.1 (1.3)	14.3 (1.4)	0.09
Platelets (10 <sup>9</sup> /L)	164 (56)	172 (55)	0.68
HCV RNA (10 <sup>3</sup> IU/ml)	1,859 (1,468)	2,021 (1,393)	0.09
ISDR mutations: ≥2 (%)	15 (%)	20 (%)	0.11
Core70: mutant (%)	36 (%)	29 (%)	0.22
Core91: mutant (%)	40 (%)	36 (%)	0.20
Fibrosis: F2-4 (%)	49 (%)	48 (%)	0.36
Activity: A2-3 (%)	42 (%)	34 (%)	0.10

AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase; LDL-C, low-density-lipoprotein-cholesterol; ISDR, interferon sensitivity-determining region. Data expressed as mean (SD).

0-3: A0 (no activity), A1 (mild activity), A2 (moderate activity), and A3 (severe activity). Sustained virological response was defined as undetectable HCV RNA by qualitative PCR with a lower detection limit of 50 IU/ml (Amplicor, Roche Diagnostic Systems) at week 24 after the completion of therapy.

### Statistical Analysis

A database of pretreatment variables included hematological tests (hemoglobin level, white blood cell count, and platelet count), blood chemistry tests (serum levels of creatinine, albumin, aspartate aminotransferase, alanine aminotransferase (ALT), gamma-glutamyltransferase (GGT), total cholesterol, triglyceride, and low-density lipoprotein cholesterol (LDL-C)), viral factors (HCV RNA titer, number of substitutions in ISDR, substitutions in the amino acid positions 70 and 91 of the core region), histological findings (stage of fibrosis and grade of activity) and patient characteristics (age, sex, and body mass index). Based on this database, decision-tree analysis was used to define a predictive model for sustained virological response.

Student's *t*-test was used for the univariable comparison of quantitative variables and Fisher's exact test was used for the comparison of qualitative variables. For the multivariable analysis for factors associated with sustained virological response, logistic regression models with backward selection were used to identify independent predictors of sustained virological response. Variables that showed significant association with sustained virological response by univariable analysis were included in the multivariable analysis. IBM-SPSS software v.15.0 (SPSS, Inc., Chicago, IL) was used for these analyses. For the decision-tree analysis [Segal and

Bloch, 1989], the data mining software IBM SPSS Modeler 13 (IBM SPSS, Inc.) was used, as reported previously [Kurosaki et al., 2010a,b]. In brief, the software searched for the optimal split variables to build a decision-tree structure. The entire study population was first evaluated to determine the variables and cut-off points for the most significant division into two subgroups having different probabilities of sustained virological response. Thereafter, analysis was repeated on all subgroups in the same way until either no additional significant variable was detected or the sample size was below 20.

## RESULTS

### Generation of the Decision-Tree Model

The decision-tree analysis selected five predictive variables to produce six subgroups of patients (Fig. 1). The number of substitutions in ISDR was selected as the best predictor of sustained virological response. The possibility of achieving sustained virological response was 83% for patients with two or more substitutions in ISDR compared with 44% for patients with a single or no substitution. Among patients with a single or no substitution in ISDR, age, with an optimal cut-off of 60 years, was selected as the variable of second split. Patients younger than 60 had the higher probability of sustained virological response (55%) compared with those older than 60 years (31%). Among younger patients, amino acid substitution at Core70 was selected as the third variable of split—wild-type sequence being the predictor of favorable response compared with the mutant type (65% vs. 36%). Among patients with wild-type Core70, the level of serum LDL-C was selected as the fourth variable of split, with an optimal cutoff of

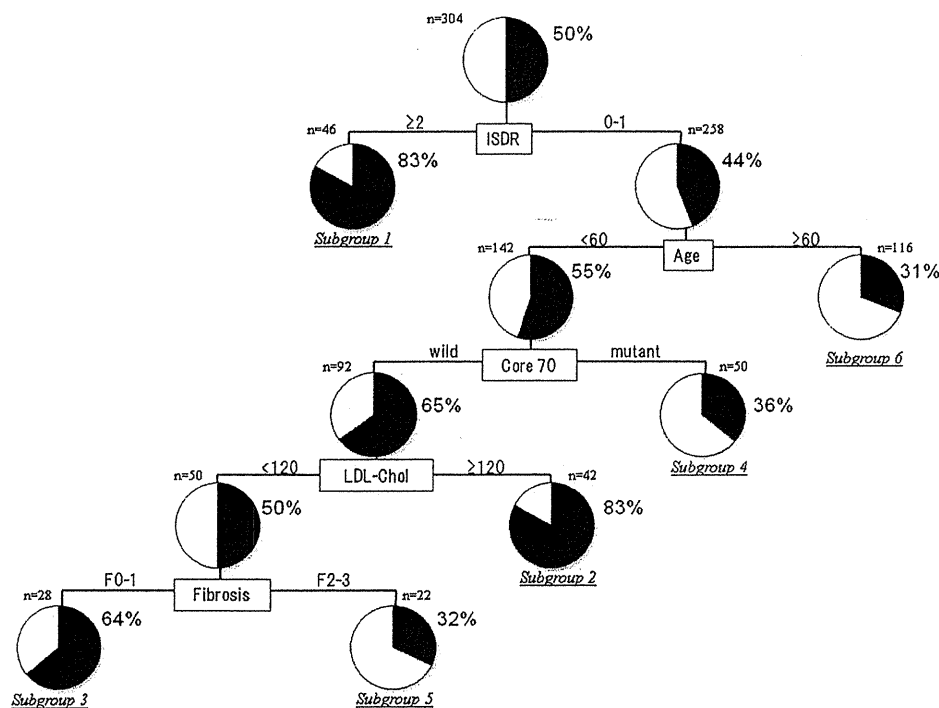


Fig. 1. Decision-tree model. Boxes indicate the factors used for splitting and the cutoff value for the split. Pie charts indicate the rate of sustained virological response for each group of patients after splitting. Terminal subgroups of patients discriminated by the analysis are numbered from 1 to 7. The rate of sustained virological response was >80% in subgroups 1 and 2, 64% in subgroup 3, and 31–36% in subgroups 4, 5, and 6. LDL-C represents low-density lipoprotein cholesterol and Core70 represents amino acid substitution at position 70 of the core region.

120 mg/dl. Patients with higher LDL-C level had the higher probability of sustained virological response (83% vs. 50%). The stage of fibrosis was selected as the final variable of split, with significant fibrosis (F2–4) being the predictor of lower sustained virological response probability (64% vs. 32%).

Among the six subgroups derived by this decision tree, the subgroup of patients with two or more substitutions in ISDR (subgroup 1) or with a single or no substitution in ISDR but younger than 60 years of age, having the wild-type Core70 and high serum level of LDL-C ( $\geq 120$  mg/dl) (subgroup 2) showed the highest probability of sustained virological response (83%).

#### Validation of the Decision-Tree Model

The decision-tree model was validated using a validation dataset of 201 cases that were not included in the model-building dataset. Each patient in the validation set was allocated to subgroups 1–6 using the flowchart form of the decision tree. The rates of sustained virological response were 75% for subgroup 1, 73% for subgroup 2, 65% for subgroup 3, 41% for subgroup 4, 46% for subgroup 5, and 33% for subgroup 6. The rates of sustained virological response for each subgroup of patients were correlated closely between the model building dataset and the validation dataset ( $r^2 = 0.94$ ) (Fig. 2).

*J. Med. Virol.* DOI 10.1002/jmv

The six subgroups were reconstructed into three groups according to their rate of sustained virological response: the high-probability group consisted of subgroups 1 and 2, the intermediate-probability group consisted of subgroup 3, and the low-probability group consisted of subgroups 4, 5, and 6. The rate of sustained virological response in the high-probability group was high on a consistent basis: 83% for model-building patients and 74% for validation patients. The rate of sustained virological response in the intermediate-probability group was 64% for model building patients and 65% for internal validation patients. The rate of sustained virological response in the low-probability group was low on a consistent basis: 32% for model-building patients and 36% for internal validation patients (Fig. 3). Thirty percent of the patients were classified into the high-probability group and 10% of the patients were classified into intermediate-probability group, which means that about 40% of patients with higher than average probability of achieving sustained virological response were identified.

#### Effect of Dose Reductions of PEG-IFN and RBV

The possible effect of drug reductions was analyzed in the three groups of patients divided by decision tree (low-, intermediate-, and high-probability groups)

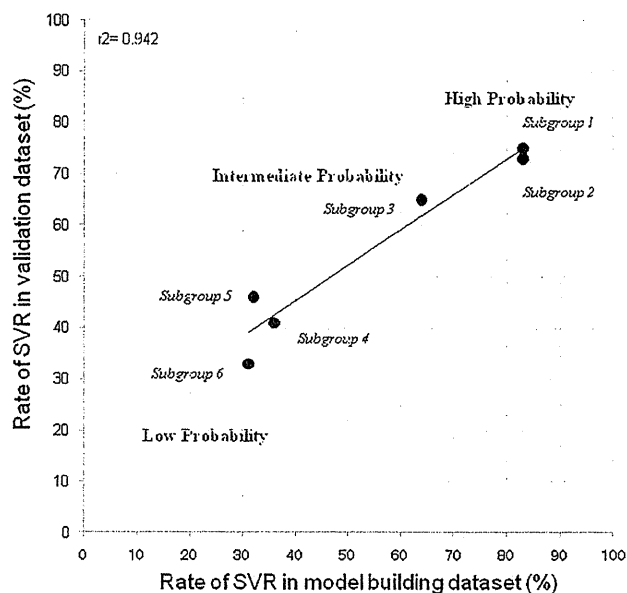


Fig. 2. Validation of the decision-tree analysis: Subgroup-stratified comparison of the rate of sustained virological response. Each patient in the validation set was allocated to subgroups 1–6 by following the flowchart form of the decision tree, and the rates of sustained virological response were then calculated and plotted for each subgroup. The x-axis represents the rate of sustained virological response in the model-building datasets and the y-axis represents the rate of sustained virological response in the validation datasets. The rates of achieving sustained virological response in each subgroup of patients correlated closely between the model-building dataset and the validation dataset (correlation coefficient:  $r^2 = 0.94$ ).

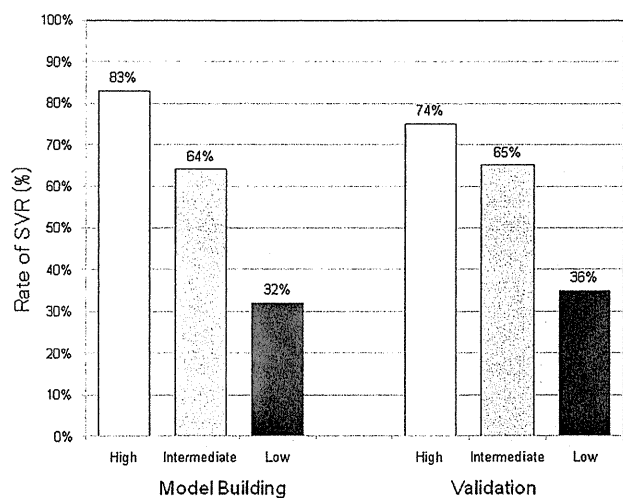


Fig. 3. Comparison of sustained virological response rates between groups divided by the decision tree. The rate of sustained virological response was compared between three groups of patients as divided by the decision-tree analysis. Black, gray, and white boxes indicate the low-probability group (subgroup 4, 5, and 6), intermediate-probability group (subgroup 3), and high-probability group (subgroup 1 and 2), respectively. The rate of sustained virological response showed significant difference between the three groups.

(Fig. 4). Patients were stratified according to the cumulative drug exposure with PEG-IFN and RBV: the good adherence group consisted of patients who took  $\geq 80\%$  planned doses of both PEG-IFN and RBV; the poor adherence group consisted of patients who took  $< 80\%$  of planned doses of both PEG-IFN and RBV. Even after adjustment for drug adherence, the three groups of patients divided by decision-tree analysis still had low, intermediate, and high probability of achieving sustained virological response, respectively, indicating that this model predicts sustained virological response independent of drug exposure.

### Multivariable Logistic Regression Analysis

Age, sex, serum levels of creatinine, ALT, GGT, LDL-C, hemoglobin, platelet count, HCV RNA titer, ISDR substitution, substitution at Core70, substitution at Core91, histological stage of fibrosis, and grade of activity were found to be associated with sustained virological response by standard univariable analysis. Multivariable analysis including these factors showed that age, sex, LDL-C levels, GGT levels, platelet count, ISDR substitution, and substitution at Core70 showed independent associations with sustained virological response (Table II). Substitution in ISDR had the highest odds ratio, at 9.92. Fibrosis, which was selected as a significant predictor of response in the decision-tree analysis, was not found to be an independent predictor of response in standard multivariable analysis, indicating that the decision-tree analysis could identify significant predictors that would apply specifically to selected patients.

### DISCUSSION

The present study revealed that viral factors such as substitutions in ISDR and Core70 are significant and independent predictors of sustained virological response to PEG-IFN plus RBV in chronic hepatitis C. In a decision-tree model for the pretreatment prediction of sustained virological response, the number of substitutions in ISDR was the best predictor of sustained virological response, followed by younger age, wild-type sequence at Core70, higher level of LDL-C, and absent fibrosis. This decision-tree model could identify patients with high probability of sustained virological response (83%) among difficult-to-treat genotype 1b chronic hepatitis C patients. Using this model, rapid estimates of the response before treatment can be made by allocating patients to specific subgroups with a defined rate of response simply by following the flowchart form. Because more potent therapy, such as a combination of protease inhibitor, PEG-IFN, and RBV, is under clinical trial and may become available in the near future [Hezode et al., 2009; McHutchison et al., 2009], pretreatment prediction of the likelihood of sustained virological response may be useful for both patients and physicians to support clinical decisions whether to start current standard therapy or to wait for emerging new therapies.



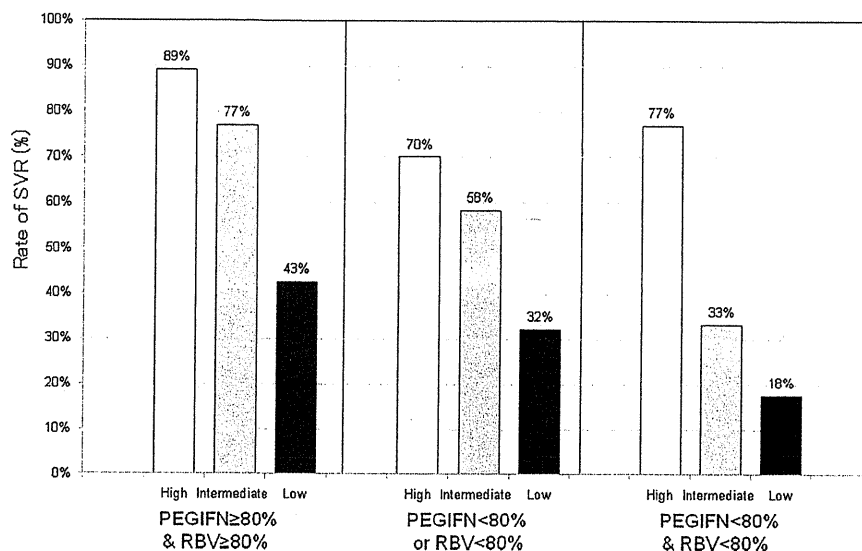


Fig. 4. Comparison of the rate of sustained virological response between the decision-tree groups stratified by drug adherence. The three groups of patients divided by the decision tree (black, gray, and white boxes indicating the low-, intermediate-, and high-probability groups, respectively) were further stratified according to cumulative drug exposure to PEG-IFN and RBV.

Two or more substitutions in ISDR had a strong impact on sustained virological response, because this factor was selected as a top variable in decision-tree analysis and had the highest odds ratio in multivariable analysis. Moreover, even among patients with unfavorable ISDR (0 or 1 mutation), younger patients (<60 years) with the wild-type sequence at Core70 and high level of LDL-C ( $\geq 120$  mg/dl) had a high rate of sustained virological response. The sustained virological response rate of these two subgroups of patients was 83% in the model-building patients and 75% in the validation patients. Thus, patients with high possibility of sustained virological response could be extracted by the combined analysis of ISDR and Core70. These patients may be the best-suited candidates for treatment with the current combination therapy. Conversely, the following patients with 0–1 mutation in ISDR had a low probability of sustained virological response (32–35%): (1) older (>60 years); or (2) younger (<60 years) patients but having mutant-type sequence at Core70; or (3) younger (<60 years) patients having a wild-type sequence at Core70, but having a low level of LDL-C (<120 mg/dl) and advanced fibrosis. These patients may

be advised to wait for a more effective therapy. Decision may be made on a case-by-case basis, taking into account the potential risk of disease progression while waiting.

In a previous decision-tree model using simple and noninvasive standard tests that are available readily worldwide [Kurosaki et al., 2010b], the rate of sustained virological response was at most 65–76% among those in the high-probability group. That model focused on use by general physicians in routine general practice, especially where specialized resources, such as liver biopsy or determination of viral sequences, are not available. In that model, younger age, male sex, higher platelet counts, lower alpha-fetoprotein (AFP) levels, and lower GGT levels were identified as favorable predictive parameters. Higher AFP levels and lower platelet counts that are hallmarks of advanced fibrosis [Shiratori and Omata, 2000; Akuta et al., 2007b] were associated with low probability of sustained virological response in that model. On the other hand, the present analysis aimed to clarify the significance of viral factors for pretreatment prediction of sustained virological response, and to build an advanced model that may be used by specialist physicians engaged in the

TABLE II. Multivariable Logistic Regression Analysis for Factors Associated With SVR

Parameter		Odds	95% CI	P-value
Age (years)	<60 vs. $\geq 60$	2.28	1.31–3.94	0.003
Sex	Male vs. female	3.36	1.87–5.99	<0.0001
GGT (IU/L)	<40 vs. $\geq 40$	2.65	1.45–4.85	0.002
LDL-C (mg/dl)	$\geq 120$ vs. <120	1.79	0.91–3.53	0.094
Platelets (10 <sup>9</sup> /L)	$\geq 120$ vs. <120	2.69	1.22–5.90	0.014
ISDR mutations	$\geq 2$ vs. 0–1	9.92	3.71–26.54	<0.0001
Core70	Wild vs. mutant	1.92	1.07–3.47	0.030

GGT, gamma-glutamyltransferase; LDL-C, low-density-lipoprotein-cholesterol; ISDR, interferon sensitivity-determining region.

treatment of hepatitis. In the present model, stage of fibrosis was selected as a predictive factor, but at lower level of significance than HCV mutations. The predicted rate of sustained virological response in the high-probability group of the present model is higher than that in the previous model (75–83% vs. 65–76%). These results indicate that substitutions in ISDR and Core70 were important pretreatment predictors of sustained virological response. Determination of these viral factors is not available readily in clinical practice, but is of value for improving the accuracy of pretreatment prediction of sustained virological response.

Substitutions in ISDR and Core70 have been reported previously to be associated with efficacy of IFN therapy. The association between the number of substitutions in ISDR and response to therapy was demonstrated originally in patients treated with IFN mono-therapy [Enomoto et al., 1995, 1996; Kurosaki et al., 1997], but recent studies have reported a positive correlation with PEG-IFN and RBV combination therapy as well [Munoz de Rueda et al., 2008; Shirakawa et al., 2008; Ikeda et al., 2009]. Another important viral factor relevant to treatment response is amino acid substitution in Core70. The sequence of this amino acid was reported originally to be associated with nonresponse to therapy [Akuta et al., 2005], but subsequent studies confirmed the positive correlation of a wild-type Core70 with sustained virological response [Akuta et al., 2009]. The multiple logistic regression analysis showed that ISDR and Core70 were independent factors associated with sustained virological response along with host factors. How these important viral factors and other host factors can be combined to predict response to PEG-IFN plus RBV is an important clinical question. Decision-tree modeling can make the response probability apparent by combining all these factors. Some factors that may be associated with treatment outcome, such as levels of ferritin or homocysteine, were not included. This may be a potential limitation of the present study.

It is of interest that a recent study by Li et al. [2010] has shown that a high serum level of LDL-C is linked to the *IL28B* major allele (CC in rs12979860). In that study, a high serum level of LDL-C was associated with sustained virological response, 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 in the present study may reflect the underlining link of LDL cholesterol levels to the *IL28B* genotype. Recent reports indicate that the *IL28B* genotype and HCV substitutions are correlated closely [Akuta et al., 2010; Kurosaki et al., 2010c]. Still, Core70 [Akuta et al., 2010] or ISDR [Kurosaki et al., 2010c] were predictors of response to therapy independent of *IL28B* genotype. Future study is needed to elucidate the possible mechanisms underlying the association between HCV sequences and host genetic factors, and also the role of host and viral factors for the prediction of treatment response.

In conclusion, a data mining analysis emphasized the impact of substitutions in ISDR and Core70 on pretreatment prediction of sustained virological response to PEG-IFN plus RBV therapy. A decision-tree model that includes substitutions in ISDR and Core70 of HCV could identify patients with high probability of sustained virological response, and could thereby improve the predictive accuracy over predictions that are based on standard tests.

## REFERENCES

- Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, Kobayashi M, Saitoh S, Watahiki S, Sato J, Matsuda M, Arase Y, Ikeda K, Kumada H. 2005. 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* 48:372–380.
- Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Arase Y, Ikeda K, Kumada H. 2007a. 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* 46:403–410.
- Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Arase Y, Ikeda K, Kumada H. 2007b. Predictors of viral kinetics to peginterferon plus ribavirin combination therapy in Japanese patients infected with hepatitis C virus genotype 1b. *J Med Virol* 79:1686–1695.
- Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Saitoh S, Arase Y, Ikeda K, Kumada H. 2009. A matched case-controlled study of 48 and 72 weeks of peginterferon plus ribavirin combination therapy in patients infected with HCV genotype 1b in Japan: Amino acid substitutions in HCV core region as predictor of sustained virological response. *J Med Virol* 81:452–458.
- Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Saitoh S, Arase Y, Ikeda K, Chayama K, Nakamura Y, Kumada H. 2010. Amino acid substitution in hepatitis C virus core region and genetic variation near the interleukin 28B gene predict viral response to telaprevir with peginterferon and ribavirin. *Hepatology* 52:421–429.
- Averbook BJ, Fu P, Rao JS, Mansour EG. 2002. 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* 132:589–602.
- Baquerizo A, Anselmo D, Shackleton C, Chen TW, Cao C, Weaver M, Gornbein J, Geevarghese S, Nissen N, Farmer D, Demetriou A, Busuttill RW. 2003. Phosphorus an early predictive factor in patients with acute liver failure. *Transplantation* 75:2007–2014.
- Bedossa P, Poynard T. 1996. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology* 24:289–293.
- Breiman L, Friedman RA, Olshen CJ, Stone CM. 1980. Classification and regression trees. CA: Wadsworth.
- Davis GL, Wong JB, McHutchison JG, Manns MP, Harvey J, Albrecht J. 2003. Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in patients with chronic hepatitis C. *Hepatology* 38:645–652.
- Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, Izumi N, Marumo F, Sato C. 1995. 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* 96:224–230.
- Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, Ogura Y, Izumi N, Marumo F, Sato C. 1996. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *N Engl J Med* 334:77–81.
- Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncalves FL, Haussinger D, Diago M, Carosi G, Dhumeaux D, Craxi A, Lin A, Hoffman J, Yu J. 2002. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 347:975–982.

- Garzotto M, Beer TM, Hudson RG, Peters L, Hsieh YC, Barrera E, Klein T, Mori M. 2005. Improved detection of prostate cancer using classification and regression tree analysis. *J Clin Oncol* 23:4322–4329.
- Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB. 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 461:399–401.
- Hezode C, Forestier N, Dusheiko G, Ferenci P, Pol S, Goeser T, Bronowicki JP, Bourliere M, Gharakhanian S, Bengtsson L, McNair L, George S, Kieffer T, Kwong A, Kauffman RS, Alam J, Pawlotsky JM, Zeuzem S. 2009. Telaprevir and peginterferon with or without ribavirin for chronic HCV infection. *N Engl J Med* 360:1839–1850.
- Ikeda H, Suzuki M, Okuse C, Yamada N, Okamoto M, Kobayashi M, Nagase Y, Takahashi H, Matsunaga K, Matsumoto N, Itoh F, Yotsuyanagi H, Koitabashi Y, Yasuda K, Iino S. 2009. Short-term prolongation of pegylated interferon and ribavirin therapy for genotype 1b chronic hepatitis C patients with early viral response. *Hepatology* 49:753–759.
- Izumi N, Nishiguchi S, Hino K, Suzuki F, Kumada Y, Itoh Y, Asahina Y, Tamori A, Hiramatsu N, Hayashi N, Kudo M. 2010. Management of hepatitis C; Report of the consensus meeting at the 45th annual meeting of the Japan society of hepatology (2009). *Hepatology* 40:347–368.
- Jensen DM, Morgan TR, Marcellin P, Pockros PJ, Reddy KR, Hadziyannis SJ, Ferenci P, Ackrill AM, Willems B. 2006. Early identification of HCV genotype 1 patients responding to 24 weeks peginterferon alpha-2a (40 kd)/ribavirin therapy. *Hepatology* 43:954–960.
- Kurosaki M, Enomoto N, Murakami T, Sakuma I, Asahina Y, Yamamoto C, Ikeda T, Tozuka S, Izumi N, Marumo F, Sato C, Ogura Y. 1997. 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* 25:750–753.
- Kurosaki M, Matsunaga K, Hirayama I, Tanaka T, Sato M, Yasui Y, Tamaki N, Hosokawa T, Ueda K, Tsuchiya K, Nakanishi H, Ikeda H, Itakura J, Takahashi Y, Asahina Y, Higak M, Enomoto N, Izumi N. 2010a. A predictive model of response to peginterferon ribavirin in chronic hepatitis C using classification and regression tree analysis. *Hepatology* 40:251–260.
- Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, Hiramatsu N, Sugauchi F, Yatsushashi H, Izumi N. 2010b. Pretreatment prediction of response to peginterferon plus ribavirin therapy in genotype 1 chronic hepatitis C using data mining analysis. *J Gastroenterol* DOI: 10.1007/s00535-010-0322-5.
- Kurosaki M, Tanaka Y, Nishida N, Sakamoto N, Enomoto N, Honda M, Sugiyama M, Matsuura K, Sugauchi F, Asahina Y, Nakagawa M, Watanabe M, Sakamoto M, Maekawa S, Sakai A, Kaneko S, Ito K, Masaki N, Tokunaga K, Izumi N, Mizokami M. 2010c. Pretreatment prediction of response to pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in *IL28B* and viral factors. *J Hepatol* DOI: 10.1016/j.jhep.2010.07.037.
- LeBlanc M, Crowley J. 1995. A review of tree-based prognostic models. *Cancer Treat Res* 75:113–124.
- Lee SS, Ferenci P. 2008. Optimizing outcomes in patients with hepatitis C virus genotype 1 or 4. *Antivir Ther* 13:9–16.
- Leiter U, Buettner PG, Eigentler TK, Garbe C. 2004. Prognostic factors of thin cutaneous melanoma: An analysis of the central malignant melanoma registry of the German Dermatological Society. *J Clin Oncol* 22:3660–3667.
- Li JH, Lao XQ, Tillmann HL, Rowell J, Patel K, Thompson A, Suchindran S, Muir AJ, Guyton JR, Gardner SD, McHutchison JG, McCarthy JJ. 2010. Interferon-lambda genotype and low serum low-density lipoprotein cholesterol levels in patients with chronic hepatitis C infection. *Hepatology* 51:1904–1911.
- Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, Goodman ZD, Koury K, Ling M, Albrecht JK. 2001. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: A randomised trial. *Lancet* 358:958–965.
- McHutchison JG, Everson GT, Gordon SC, Jacobson IM, Sulkowski M, Kauffman R, McNair L, Alam J, Muir AJ. 2009. Telaprevir with peginterferon and ribavirin for chronic HCV genotype 1 infection. *N Engl J Med* 360:1827–1838.
- Miyaki K, Takei I, Watanabe K, Nakashima H, Omae K. 2002. Novel statistical classification model of type 2 diabetes mellitus patients for tailor-made prevention using data mining algorithm. *J Epidemiol* 12:243–248.
- Munoz de Rueda P, Casado J, Paton R, Quintero D, Palacios A, Gila A, Quiles R, Leon J, Ruiz-Extremera A, Salmeron J. 2008. Mutations in E2-PePHD, NS5A-PKRBD, NS5A-ISDR, and NS5A-V3 of hepatitis C virus genotype 1 and their relationships to pegylated interferon-ribavirin treatment responses. *J Virol* 82:6644–6653.
- Segal MR, Bloch DA. 1989. A comparison of estimated proportional hazards models and regression trees. *Stat Med* 8:539–550.
- Shirakawa H, Matsumoto A, Joshita S, Komatsu M, Tanaka N, Umemura T, Ichijo T, Yoshizawa K, Kiyosawa K, Tanaka E. 2008. Pretreatment prediction of virological response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. *Hepatology* 48:1753–1760.
- Shiratori Y, Omata M. 2000. Predictors of the efficacy of interferon therapy for patients with chronic hepatitis C before and during therapy: How does this modify the treatment course? *J Gastroenterol Hepatol* 15:E141–E151.
- Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, Bassendine M, Spengler U, Dore GJ, Powell E, Riordan S, Sheridan D, Smedile A, Fragomeli V, Muller T, Bahlo M, Stewart GJ, Booth DR, George J. 2009. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 41:1100–1104.
- Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, Hino K, Hige S, Ito Y, Mita E, Tanaka E, Mochida S, Murawaki Y, Honda M, Sakai A, Hiasa Y, Nishiguchi S, Koike A, Sakaida I, Imamura M, Ito K, Yano K, Masaki N, Sugauchi F, Izumi N, Tokunaga K, Mizokami M. 2009. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* 41:1105–1109.
- Valera VA, Walter BA, Yokoyama N, Koyama Y, Iiai T, Okamoto H, Hatakeyama K. 2007. Prognostic groups in colorectal carcinoma patients based on tumor cell proliferation and classification and regression tree (CART) survival analysis. *Ann Surg Oncol* 14:34–40.
- Yu ML, Dai CY, Huang JF, Chiu CF, Yang YH, Hou NJ, Lee LP, Hsieh MY, Lin ZY, Chen SC, Wang LY, Chang WY, Chuang WL. 2008. Rapid virological response and treatment duration for chronic hepatitis C genotype 1 patients: A randomized trial. *Hepatology* 47:1884–1893.
- Zlobec I, Steele R, Nigam N, Compton CC. 2005. A predictive model of rectal tumor response to preoperative radiotherapy using classification and regression tree methods. *Clin Cancer Res* 11:5440–5443.

## Pre-treatment prediction of response to pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in *IL28B* and viral factors

Masayuki Kurosaki<sup>1</sup>, Yasuhito Tanaka<sup>2</sup>, Nao Nishida<sup>3</sup>, Naoya Sakamoto<sup>4</sup>, Nobuyuki Enomoto<sup>5</sup>, Masao Honda<sup>6</sup>, Masaya Sugiyama<sup>2</sup>, Kentaro Matsuura<sup>2</sup>, Fuminaka Sugauchi<sup>2</sup>, Yasuhiro Asahina<sup>1</sup>, Mina Nakagawa<sup>4</sup>, Mamoru Watanabe<sup>4</sup>, Minoru Sakamoto<sup>5</sup>, Shinya Maekawa<sup>5</sup>, Akito Sakai<sup>6</sup>, Shuichi Kaneko<sup>6</sup>, Kiyooki Ito<sup>7</sup>, Naohiko Masaki<sup>7</sup>, Katsushi Tokunaga<sup>3</sup>, Namiki Izumi<sup>1,\*</sup>, Masashi Mizokami<sup>2,7</sup>

<sup>1</sup>Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan; <sup>2</sup>Department of Virology, Liver Unit, Nagoya City University, Graduate School of Medical Sciences, Nagoya, Japan; <sup>3</sup>Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; <sup>4</sup>Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan; <sup>5</sup>First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan; <sup>6</sup>Department of Gastroenterology, Kanazawa University, Graduate School of Medicine, Kanazawa, Japan; <sup>7</sup>Research Center for Hepatitis and Immunology, International Medical Center of Japan, Konodai Hospital, Ichikawa, Japan

**Background & Aims:** Pegylated interferon and ribavirin (PEG-IFN/RBV) therapy for chronic hepatitis C virus (HCV) genotype 1 infection is effective in 50% of patients. Recent studies revealed an association between the *IL28B* genotype and treatment response. We aimed to develop a model for the pre-treatment prediction of response using host and viral factors.

**Methods:** Data were collected from 496 patients with HCV genotype 1 treated with PEG-IFN/RBV at five hospitals and universities in Japan. *IL28B* genotype and mutations in the core and IFN sensitivity determining region (ISDR) of HCV were analyzed to predict response to therapy. The decision model was generated by data mining analysis.

**Results:** The *IL28B* polymorphism correlated with early virological response and predicted null virological response (NVR) (odds ratio = 20.83,  $p < 0.0001$ ) and sustained virological response (SVR) (odds ratio = 7.41,  $p < 0.0001$ ) independent of other covariates. Mutations in the ISDR predicted relapse and SVR independent of *IL28B*. The decision model revealed that patients with the minor *IL28B* allele and low platelet counts had the highest NVR (84%) and lowest SVR (7%), whereas those with the major *IL28B* allele and mutations in the ISDR or high platelet counts had the lowest NVR (0–17%) and highest SVR (61–90%). The model had high reproducibility and predicted SVR with 78% specificity and 70% sensitivity.

**Conclusions:** The *IL28B* polymorphism and mutations in the ISDR of HCV were significant pre-treatment predictors of response to PEG-IFN/RBV. The decision model, including these host and viral factors may support selection of optimum treatment strategy for individual patients.

© 2010 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved.

### Introduction

Hepatitis C virus (HCV) infection is the leading cause of cirrhosis and hepatocellular carcinoma worldwide [1]. The successful eradication of HCV, defined as a sustained virological response (SVR), is associated with a reduced risk of developing hepatocellular carcinoma. Currently, pegylated interferon (PEG-IFN) plus ribavirin (RBV) is the most effective standard of care for chronic hepatitis C but the rate of SVR is around 50% in patients with HCV genotype 1 [2,3], the most common genotype in Japan, Europe, the United States, and many other countries. Moreover, 20–30% of patients with HCV genotype 1 have a null virological response (NVR) to PEG-IFN/RBV therapy [4]. The most reliable method for predicting the response is to monitor the early decline of serum HCV-RNA levels during treatment [5] but there is no established method for prediction before treatment. Because PEG-IFN/RBV therapy is costly and often accompanied by adverse effects such as flu-like symptoms, depression and hematological abnormalities, pre-treatment predictions of those patients who are unlikely to benefit from this regimen enables ineffective treatment to be avoided.

Recently, it has been reported through a genome-wide association study (GWAS) of patients with genotype 1 HCV that single nucleotide polymorphisms (SNPs) located near the *IL28B* gene are strongly associated with a response to PEG-IFN/RBV therapy in

Keywords: *IL28B*; ISDR; Peg-interferon; Ribavirin; Data mining; Decision tree.  
Received 14 March 2010; received in revised form 22 June 2010; accepted 7 July 2010; available online 19 September 2010

\* Corresponding author. Address: Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, 1-26-1 Kyonan-cho, Musashino-shi, Tokyo 180-8610, Japan. Tel.: +81 422 32 3111; fax: +81 422 32 9551.

E-mail address: nizumi@musashino.jrc.or.jp (N. Izumi).



ELSEVIER