

To date, both ultrasonography and serum tumor markers such as the alpha fetoprotein (AFP), and des- γ -carboxy prothrombin (DCP) assay are the principle methods for screening and detecting HCC. Routine screening is the best method to detect early HCC and improve patient survival; however, elevated serum AFP and DCP levels have insufficient sensitivity and specificity, respectively. The sensitivity and specificity of serum elevated AFP levels were reported to range from 39–64% and 76–91%, while those of the serum elevated DCP levels were 41–77% and 72–98%, respectively [4–9]. In addition, it was recently reported that only a small percentage of small HCC tumors were diagnosed based on AFP and DCP [6, 10]. The lens culinaris agglutinin-reactive fraction of AFP (lectin-bound AFP or AFP-L3) has been reported to be elevated in the serum of HCC patients. Although AFP-L3 has a high range of specificity for detecting HCC, the sensitivity is low [11, 12]. The ability to detect early HCC, prior to the onset of clinical symptoms, leads to curative treatment and significantly improves the disease prognosis. Thus, additional biochemical markers are necessary for the specific detection of early HCC.

Serum profiling using a proteomic approach is thought to be a useful technique to detect or predict early HCC in chronic liver disease patients. Studies using the Protein-Chip SELDI system, which is a powerful tool to discover new biomarkers, have shown that this method may be successfully used to diagnose HCC. Zinkin et al. [13], Schweigler et al. [14] and our research group [15] previously detected early HCC using the profile of several protein peaks that were identified by the ProteinChip SELDI system. Paradis et al. [16] reported the highest discriminating peak (8900 Da), which was identified as the V10 fragment of vitronectin. Furthermore, Lee et al. [17] described complement 3a, which had a molecular weight of approximately 8900 Da, as a novel marker of HCC. Therefore, using this proteomic approach to identify specific proteins may not only help establish simple methods to detect HCC, but also further our understanding of the molecular mechanisms of hepatocarcinogenesis and facilitate the development of novel cancer therapies. Therefore, this study assessed and compared the protein expression profiles in the sera of HCC patients in order to identify a more useful biomarker of HCC-associated HCV infection using proteomic approach.

Materials and methods

Samples

Eighty-seven patients [45 HCC patients and 42 patients with chronic liver diseases without HCC (CLD)] with

Table 1 Patient characteristics

	HCC ^a	CLD ^b	<i>p</i> value
Patients (male/female)	45 (40/5)	42 (40/2)	–
Age	73.6 [63–85]	61.8 [41–83]	<0.0001
PLT ^c ($\times 10^4$ /ul)	12.5 \pm 5.8	8.4 \pm 4.6	0.001
Albumin (g/dl)	3.8 \pm 0.8	4.2 \pm 1.6	0.8
ALT ^d (IU/l)	57.7 \pm 28.3	52.8 \pm 37.5	0.7
AFP ^e (ng/ml)	311 \pm 1144	51.6 \pm 36.1 (38)	0.008
DCP ^f (mAU/ml)	235 \pm 605 (44)	37.1 \pm 59.8 (39)	<0.0001
HA ^g (ng/ml)	388 \pm 446 (40)	280 \pm 272 (27)	0.6
Diameter of HCC (mm)	23.2 [10–40]	–	–
TNM stage ^h (I/II/III/IV)	24/18/3/0	–	–

Data are shown as the means \pm SD or means [range] (numbers)

^a Hepatocellular carcinoma

^b Chronic liver disease

^c Platelet counts

^d Alanine aminotransferase

^e Alpha fetoprotein

^f Des- γ -carboxy prothrombin

^g Hyaluronic acid

^h TNM; primary tumor/lymph node/distant metastasis

HCV infection were selected to participate in this study (Table 1). These patients provided informed consent. Serum samples were collected by the Faculty of Medicine, University of Miyazaki (Miyazaki, Japan), and some patients were in a hyperendemic HCV area with a cohort study in Miyazaki [18]. The sera of all patients with and without HCC, which was confirmed by abdominal ultrasonography or computed tomography, were obtained prior to treatment. All of the sera samples from HCV-infected patients were analyzed in a previous study [15]. In addition, sera from 10 HCV-HCC patients who were diagnosed with HCC within 1 or 2 years and sera from five patients who had received radiofrequency ablation (RFA), percutaneous ethanol injection therapy (PEIT) and/or transarterial chemoembolization (TACE) for HCC were collected through a cohort study in Miyazaki. We also analyzed the sera of 21 healthy volunteers without HCC as controls. After freezing and thawing once, all samples were separated into 50–100 μ l aliquots and refrozen at -80°C . The study protocol was approved by the Ethics Committee of the Faculty of Medicine, University of Miyazaki, Kagoshima University Graduate School of Medical and Dental Sciences, and Harvard School of Public Health and Boston University School of Public Health.

SELDI-TOF/MS analysis of sera

Expression difference mapping analysis profiles of the samples were obtained using weak cation-exchange (CM10) ProteinChip Arrays (Bio-Rad Laboratories). Arrays were analyzed by ProteinChip reader as previously reported [15]. In addition, the laser intensity ranged from 220 to 245, with a detector sensitivity of 8, and spectra ranging from 1300 to 150000 *m/z* were selected for analysis in this study.

Separation of candidate biomarker (8.1 k *m/z*)

The purification strategy was determined by the ProteinChip Arrays. Two hundred microliters of sera from HCV-HCC patients were diluted 5-fold into 50 mM Naphosphate buffer, pH 7.0, and loaded onto a CM-Ceramic HyperD F spin column (Bio-Rad Laboratories). After equilibrating with the same buffer, the samples were eluted with a stepwise sodium chloride gradient from 0, 200, 300, and 1000 mM. The elution was desalinated and concentrated using a centrifugal concentrator (VIVA-SPIN, Vivascience, Hannover, Germany), and the purification progress was monitored using NP20 arrays. The flow-through fraction was dialyzed and then separated by 16.5% tricine one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The SDS-PAGE samples were run in tricine sodium dodecyl sulfate buffer according to the manufacturer's instructions and then stained with Coomassie brilliant blue (CBB).

Identification of the candidate biomarker (8.1 k *m/z*)

Gel pieces containing the target 8.1 k *m/z* protein were excised. The excised bands were reduced and alkylated for 30 min at room temperature, and then digested with trypsin (Modified Sequence Grade, Roche Diagnostics, Basel, Switzerland) in Tris-HCl, pH 8.0, for 20 h at 35°. The reaction solution was applied to NP20 arrays and allowed to air dry. To identify the protein, the digested peptides were purified by high-performance liquid chromatography (HPLC; MAGIC 2002; Michrom Biore-sources Inc., Auburn, CA) and analyzed by Q-Tof2 (Micromass; Waters Ltd., Hertsfordshire, UK). The HPLC solvent consisted of solvent A (2% acetonitrile/0.1% formic acid) and B (90% acetonitrile/0.1% formic acid). The digested peptides were separated with a linear gradient from 10 to 50% solvent B with a flow rate of 400 nl/min using HPLC [19]. Mass spectral data were searched with Mascot (<http://www.matrixscience.com>) to identify proteins based on the peptide mass [20, 21].

Immunodepletion assay

For immunodepletion, serum samples were prepared as follows. Sera (250 μ l) from HCC patients were diluted 5-fold in 50 mM Tris-HCl buffer, pH 8.0, and loaded onto a CM-Sepharose Fast Flow spin column (GE Healthcare Bio-Sciences Corp., NJ). After equilibration with the same buffer, the samples were eluted with a stepwise sodium chloride gradient from 0, 500, and 1000 mM. The elution from each NaCl concentration was monitored using NP20 arrays. To prepare the antibodies for immunodepletion, 6 μ l anti-human C3 antibody, which detected C3 and C3a expression, or anti-C4a antibody (Santa Cruz Biotechnology, Santa Cruz, CA) was incubated with 20 μ l Interaction Discovery Mapping (IDM) affinity beads (Bio-Rad Laboratories) and Protein A (Sigma Chemical Co, St. Louis, MO) over night at 4° with shaking. These beads were centrifuged, and the supernatant was discarded. The beads were washed with 50 mM phosphate buffer (pH 7.0), and 3 μ l of the prepared serum sample was incubated with 15 μ l IDM affinity beads with shaking for 2 h at 4°. As a negative control, 3 μ l sample was incubated with IDM affinity beads and Protein A with an anti-C4a antibody or without antibody. After the incubation, the samples were cleared by centrifugation, and 5 μ l of each supernatant was analyzed on NP20 ProteinChip arrays in a PBS II reader.

Cell culture and SELDI-TOF/MS analysis of culture supernatants

The human hepatocarcinoma cell line HuH-7 and human hepatoblastoma cell line HepG2 were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS), 100 IU/ml penicillin G, and 100 mg/ml streptomycin sulfate (Invitrogen, Carlsbad, CA). Before starting the experiments, the cells were cultured on 96-well microplates in medium without FBS for 24 h. After washing with FBS-free media, the cells were cultured for 24 h with FBS-free media with or without 500 μ g/ml of C3a (Calbiochem, San Diego, CA). The supernatants were collected by centrifugation and analyzed for the expression of 8.1 k *m/z* using the ProteinChip system.

Statistical analysis

Values are shown as the means \pm SD. Statistical differences, including laboratory data and individual peaks in SELDI TOF/MS, were determined using the Mann-Whitney *U* test. Values of *p* < 0.05 were considered statistically significant. The discriminatory power for each putative marker was described via receiver operating characteristics

(ROC) area under the curve (AUC). These statistical analyses were performed using STATVIEW 4.5 software (Abacus Concepts, Berkeley, CA), SPSS software (SPSS Inc., Chicago, IL), JMP software, or CIPHERGEN ProteinChip Software, version 3.0.2.

Results

Profiling sera from HCC patients and healthy controls

We analyzed the sera of all patients with HCV-HCC or HCV-CLD and healthy controls without HCC using the CM10 ProteinChip array to identify the most differential protein peak. Peaks were automatically detected using the CIPHERGEN ProteinChip Software 3.0.2. following baseline subtraction as described previously [15, 22]. This analysis identified 178 protein peak clusters, as seen in the spectrum representations from the three groups (HCV-HCC, HCV-CLD, and healthy control) in the 3000- to 15000-*m/z* range. Peak expressions were increased for 18 proteins and decreased for 14 proteins in sera from HCV-HCC patients compared to HCV-CLD patients. Compared to healthy subjects, 68 protein peaks were increased, and 16 protein peak intensities were decreased in the sera of HCV-HCC patients. Five protein peaks (4067, 4470, 7564, 7929, and 8130 *m/z*) had a *p*-value less than 1×10^{-7} and were significantly increased in the sera of HCC patients compared to the sera of HCV-CLD patients and healthy volunteers. In particular, an 8130 *m/z* peak was the most

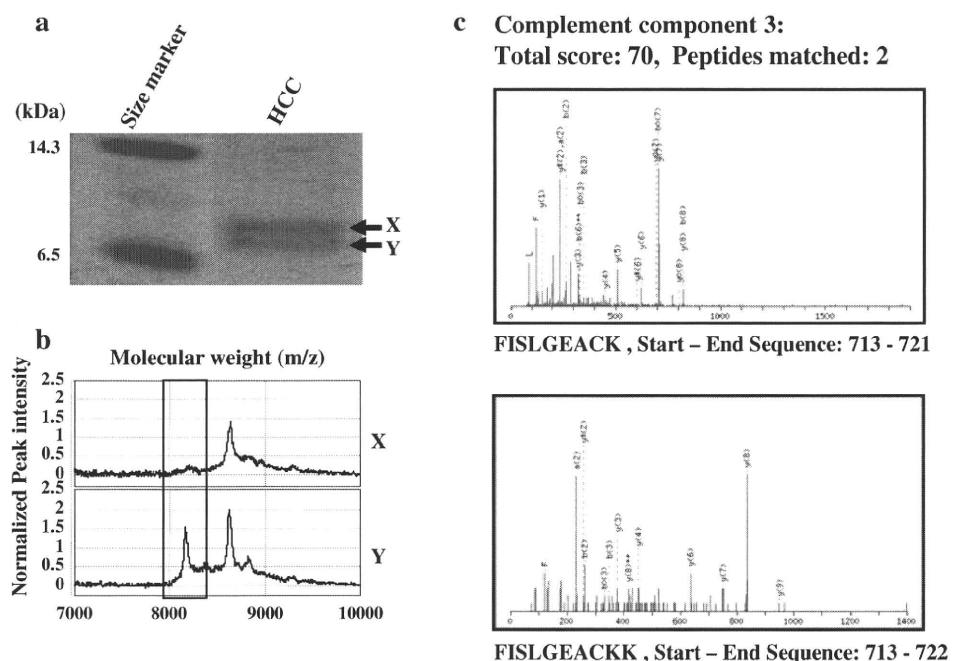
significantly different peak and had the most differential expression profile between patients with HCV-HCC and with HCV-CLD.

Purification and identification of the 8.1 k *m/z* peak

We optimized the adsorption and desorption conditions on the arrays using an HCV-HCC patient serum sample and healthy volunteer serum sample in order to determine a procedure to purify the target 8.1 k *m/z* protein. The optimal pH for retention of the 8.1 k *m/z* protein was a *pI* value of approximately 7.0 on the CM10 arrays, which indicates that weak cation-exchange sorbents and buffer pH should be fixed for further experiments. The target protein was eluted by increasing the sodium chloride concentrations in a Na-phosphate buffer and was eluted in the 1000 mM sodium chloride fraction. The concentrated serum protein that was eluted with 1000 mM sodium chloride was applied to SDS-PAGE for further separation. The 8.1 k *m/z* protein was identified and excised by in-gel trypsin digestion for identification. The peptide sequences were analyzed using liquid chromatography (LC)-MS/MS and then examined by a database search with Mascot. The digested peptides matched human complement C3a (Fig. 1).

After reacting the HCC sera with anti-complement C3a or anti-C4 antibodies or without antibody, the supernatants were analyzed by the SELDI ProteinChip system for immunodepletion. Analysis of the supernatant showed that only the 8.1 k *m/z* peak corresponding to complement C3a

Fig. 1 **a** Partially purified proteins were separated by SDS-PAGE using serum samples from HCV-HCC patients. The Coomassie-stained SDS-PAGE gel shows two clear bands at approximately 8 kDa (X and Y). **b** After each band (X and Y) was excised from the gel, the proteins were extracted and analyzed using the ProteinChip system. The target protein in the excised band was detected, and the 8.1 k *m/z* peak corresponded only to the “Y” band contained in gel. **c** The excised “Y” band was alkylated and digested using trypsin. The peptides were collected and subjected to LC-MS/MS analysis. The proteins, which were derived from complement C3a, were identified using a database search



was reduced. On the other hand, immunodepletion with a control anti-C4 antibody or without antibody did not reduce the 8.1 k m/z peak (Fig. 2).

Profiling the C3a of sera from patients with HCC and without HCC

The 8.1 k m/z peak was confirmed as the complement C3a fragment using an immunodepletion assay. However, C3a was stabilized as C3adesArg with a molecular weight of approximately 8.9 k m/z. Figure 3a, b compares the expression of the 8.1 k m/z peak in the sera of HCV-HCC or HCV-CLD patients and healthy controls. The intensities

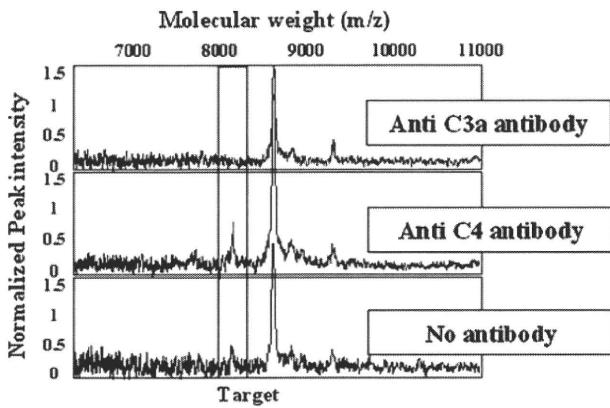
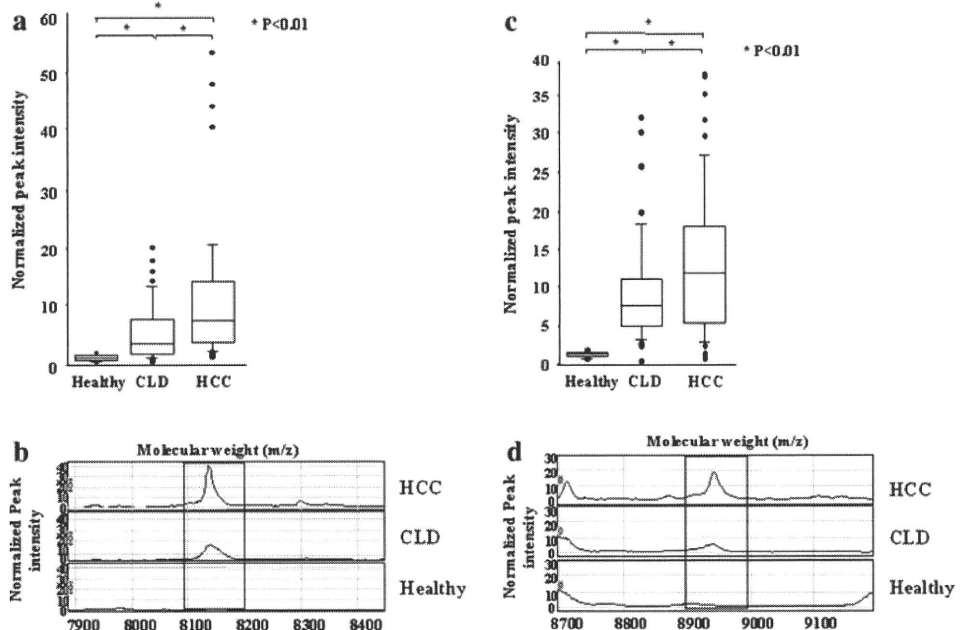


Fig. 2 Immunodepletion assay of the C3a fragment. Analysis of supernatant that had been immunodepleted with an anti-C3a antibody showed that only the 8.1 k m/z peak corresponding to complement C3a was reduced. Supernatants that had been immunodepleted with either a control anti-C4 antibody or without antibody did not have reduced 8.1 k m/z peaks by the ProteinChip system

Fig. 3 a and c Comparisons of the expression profiles of the 8.1 and 8.9 k m/z peaks in HCV-HCC, HCV-CLD, and healthy sera. Boxes indicate the median \pm 25th percentile. The lower and upper bars represent the 10th and 90th percentiles, respectively. **b and d** Representative spectra of the 8.1 and 8.9 k m/z peaks from patients in each group. The horizontal axis indicates the protein molecular weight, while the vertical axis designates the relative intensity



in HCC patient sera were significantly higher than those in the HCV-CLD patients or healthy controls. The expression of the 8.9 k m/z peak in HCV-HCC patients was also higher than that in HCV-CLD patients or healthy controls (Fig. 3c, d). Although the 8.9 k m/z peak was not identified as C3adesArg, it is possible that both the 8.1 and 8.9 k m/z peaks were specific tumor markers for HCC. Furthermore, we analyzed sera from 10 HCV-HCC patients who were diagnosed with HCC within 1 or 2 years and sera from five patients who had received curative treatments using RFA, PEIT, and TACE for HCC. The 8.1 k m/z C3a fragment in the HCV-HCC patients was significantly increased in the year of disease onset compared to the pre-onset year. After treatment, expression of the C3a fragment significantly decreased in all five of the patients who had measurable samples after treatment (Fig. 4a). In contrast, the 8.9 k m/z peak did not change regardless of the occurrence of HCC over time (Fig. 4b). Thus, the 8.1 k m/z C3a fragment appears to be the most discriminatory tumor marker for HCV-HCC.

Relationship between the C3a fragment and other tumor markers

AFP and DCP levels were measured in sera from 83 of 87 patients with HCV-associated liver disease. The recommended cutoff levels for these tumor markers, AFP and DCP, are 20 ng/ml and 40 mAU/ml, respectively. AFP-L3 in 26 patients with HCV-associated liver disease was also investigated among measurable samples in which AFP in a total 35 patients was higher than 20 ng/ml. The cutoff level of AFP-L3 was set at 10%. When samples from patients

Fig. 4 Comparisons of the expression profiles of the 8.1 k m/z (a) and 8.9 k m/z (b) peaks in sera from HCV-HCC patients before diagnosis, during disease onset, and after treatment. The samples in the before diagnosis group included sera collected 1 or 2 years before the onset of HCC. Boxes indicate the median \pm 25th percentile, the lower bar indicates the 10th percentile and the upper bar indicates the 90th percentile

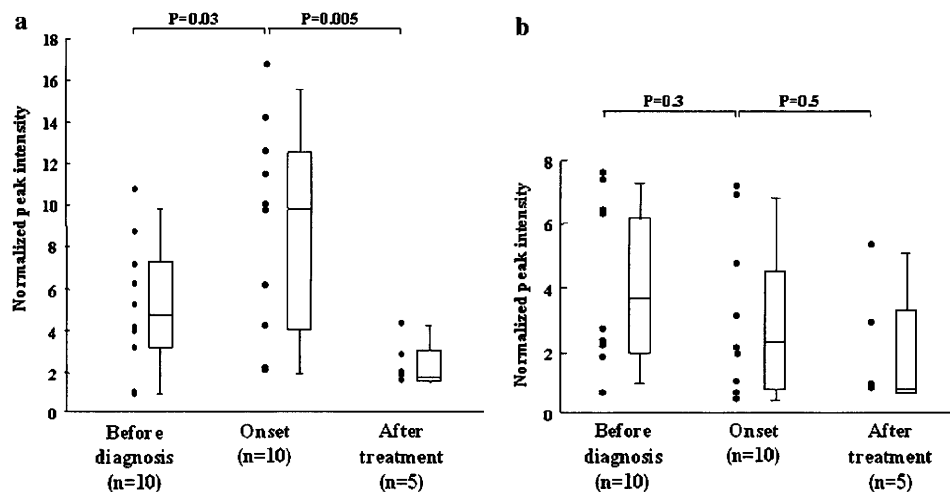


Table 2 Diagnostic rates for hepatocellular carcinoma in the HCV infected patients

Markers	Sensitivity (%)	Specificity (%)	ROC AUC
AFP ^a (>20 ng/ml)	38 (17/45)	47 (18/38)	0.53
DCP ^b (>40 mAU/ml)	45 (20/44)	74 (29/39)	0.68
AFP-L3 ^c (>10%)	58 (8/14)	50 (6/12)	0.58
C3a fragment (>3.5)	78 (37/45)	52 (22/42)	0.70
C3a fragment + AFP	91 (41/45)	26 (10/38)	0.72
C3a fragment + DCP	93 (41/44)	33 (13/39)	0.77
AFP + DCP	64 (28/44)	34 (12/35)	0.70
C3a fragment + AFP + DCP	98 (43/44)	20 (7/35)	0.80

^a Alpha fetoprotein

^b Des- γ -carboxy prothrombin

^c Alpha fetoprotein, lectin lens culinaris agglutinin-bound fraction

with HCV-HCC and HCV-CLD without HCC were compared, the sensitivity and specificity of AFP were 38 and 47%, whereas those of DCP were 45 and 74% and those of AFP-L3 were 58 and 50%, respectively. When the cutoff level for the relative intensity of the C3a fragment was set at 3.5, the sensitivity and specificity were 78 and 52%, respectively; the C3a fragment had the most sensitivity for the diagnosis of HCC. Furthermore, the ROC AUC of the C3a fragment, AFP, DCP, and AFP-L3 was 0.70, 0.53, 0.68, and 0.58, respectively (Table 2). There was no relationship between the C3a fragment and several other tumor and inflammation markers [AFP, DCP, AFP-L3, alanine aminotransferase (ALT), and high-sensitivity C-reactive protein (hs-CRP)], and each of these markers was independent of the diameter and number of tumors. The ROC AUC using AFP and DCP was highly similar to the ROC AUC with the C3a fragment alone. In addition, we investigated a combination assay that included the C3a fragment, AFP and DCP. This combination test, in which at

least AFP, DCP, or the C3a fragment was positive, had a positive identification rate of 98%, although the specificity of this assay was too low at 20%. The ROC AUC of the combination test using AFP, DCP, and the C3a fragment was higher than those of any other markers. This result indicates that this combination assay using three markers is more useful than the combination assay using AFP \pm DCP, which are measured worldwide to detect HCC (Table 2).

Profiling C3a expression in culture medium

C3a reacted with HCC cell lines, and the C3a peak in the culture medium was monitored by the ProteinChip system. The C3a fragment (approximately 8.1 k m/z) was not detected in the supernatants of HuH-7 and HepG2 cell cultures. However, the 8.9 k m/z peak was detected in the culture medium. This 8.9 k m/z peak was considered to be a stabilized form of C3a. This result indicated that the stabilized form of C3a (8.9 k m/z) was not undergoing proteasome-mediated degradation to yield the C3a fragment (8.1 k m/z) in these HCC cell lines.

Discussion

Because the HCC disease-associated mortality rate remains high, it is highly important to develop early diagnostic tools and treatments for HCC. Our study indicates that an 8.1 k m/z peak, which was identified as the C3a fragment by both peptide sequencing and an immunoassay, is up-regulated in the serum of HCC patients, 93% (42/45) of whom were TNM stage I or II. The C3a fragment in some HCC cases was also significantly higher in the year of HCC onset compared to the pre-onset year and decreased after curative treatment. Therefore, the C3a fragment appears to

be a promising simple tumor marker for the diagnosis of early HCC. In addition, a combination serum HCC diagnostic test that included AFP, DCP, and the C3a fragment had higher sensitivity than each individual marker. These results suggest that this combination test may be a useful HCC screening method, although the low specificity may pose challenges. Further examinations are needed to determine whether the C3a fragment or a combination test can be used to detect early HCC.

The results of our study demonstrated that the C3a fragment (8.1 k m/z) is a highly expressed novel tumor marker that is abundant in the sera of early HCC patients but not in the sera of healthy volunteers or HCV-CLD patients. A similar study by Lee et al. [17] used the ProteinChip SELDI system to show that C3a is a potential candidate biomarker for HCV-HCC. However, Lee et al. found that the molecular weight of C3a was represented by an approximately 8.9 k m/z peak. C3a has a very short half-life and is immediately cleaved into the more stable C3adesArg (8.9 k m/z), which is the anaphylatoxin C3a that lacks the C-terminal arginine and is stable state in the serum [23]. In our study, the 8.9 k m/z peak was also significantly different among HCV-HCC patients, HCV-CLD patients, and healthy volunteers (Fig. 3c, d). However, the discriminatory power of the 8.9 k m/z peak (ROC AUC was 0.60) was lower than the 8.1 k m/z peak (ROC AUC was 0.70) to distinguish between HCV-HCC and HCV-CLD. In addition, unlike the 8.1 k m/z peak, the levels of the 8.9 k m/z peak did not significantly increase with time as HCC progressed in 10 HCV-HCC cases (Fig. 4b). In contrast, Li et al. identified two proteins (8926 m/z and 8116 m/z) as complement component C3adesArg and a C-terminal truncated form of C3adesArg; the latter was a C-terminal truncation of C3adesArg that lacked the C-terminal sequence RASHLGLA (referred to as C3adesArg Δ 8) in breast cancer patients [24]. However, these two biomarkers cannot be used to discriminate between breast cancers and benign tumors, and there were minimal differences in the peak intensities between breast cancer patients and healthy controls. Therefore, the C3a fragment with a molecular weight of 8.1 k m/z appears to be a potential diagnostic marker for HCC, although we cannot explain why the 8.1 k m/z fragment of C3a is overexpressed in HCC patients and did not confirm whether our C3a fragment (8.1 k m/z) is C3adesArg Δ 8.

C3a, including C3adesArg, was also previously identified as a tumor marker for lymphoid malignancies, breast and colorectal cancers using the ProteinChip SELDI system [24–26]. Complement activation and subsequent deposition of complement components on tumor tissues has been demonstrated in cancer patients [27]. Malignant ovarian cells isolated from ascitic fluid samples had C3 activation products deposited on their cell surface [28].

Complement components are important mediators of inflammation and help regulate the immune response. C3a is biologically active and binds to mast cells and basophils, triggering the release of their vasoactive contents [29]. We investigated C3a expression by immunochemical examination of HCC tissues and Western blot analysis of proteins extracted from human HCC cell lines, including HepG2 and HuH-7. However, specific C3a expression, including the C3a fragment (8.1 k m/z), was not detected.

The complement system can be activated after exposure to tumor antigens [30]. It is speculated that small tumors can trigger a systematic reaction. Therefore, elevated C3a (8.9 k m/z) levels in the serum of HCV-HCC patients may reflect both a systematic immune response to HCV infection and non-specific tumor antigens rather than a specific immune response to HCC [24–26, 31]. In contrast, it is possible that overexpression of the C3a fragment (8.1 k m/z) is specific for HCC in addition to non-specific C3 activation.

In contrast to our results, Steel et al. [32] searched for HCC biomarkers using HCC-associated HBV-infected patient sera and found that the C-terminal fragment of complement C3 was down-regulated. Kawakami et al. [33] searched for characteristic alterations in the sera of HBV- and HCV-HCC-infected patients who had undergone curative radiofrequency ablation treatment and showed that C3 was up-regulated after treatment. In these studies, C3 was separated and identified using 2-DE of a mixture of proteins from a small number of patient sera samples, and this process identified various molecular weights for C3. In addition, we analyzed the sera of 25 patients with HCC-associated HBV infections, and the profile of several proteins was different between HCV- and HBV-infected patients. Although 35 protein peaks, including the C3a fragment, were overexpressed in the sera of both HCV-HCC and HBV-HCC patients compared to sera from healthy volunteers, the C3a fragment (8.1 k m/z) was particularly overexpressed in the sera of HCV-HCC patients and was not significantly different between HBV-HCC patients and HCV-CLD patients without HCC (data not shown). The biologic and pathogenic activities of HCV and HBV are different, and the molecular mechanisms underlying the development of hepatitis and hepatocarcinogenesis may differ between HBV and HCV infections [34–36]. Although the number of samples, cause of liver disease, and method of protein identification may affect these results, we speculate that the C3a fragment with a molecular weight of 8.1 k m/z is a candidate tumor marker for HCV-HCC but not HBV-HCC.

AFP, which is a commonly used HCC tumor marker, is elevated not only during HCC, but also during hepatocyte regeneration following liver damage. Previous reports revealed that AFP was abnormally elevated in the sera of patients with acute hepatitis, chronic hepatitis, and liver

cirrhosis. This lack of specificity for HCC means that AFP has a comparatively high false-positive rate [37]. The C3a fragment may also be elevated during hepatocyte regeneration following liver damage [38], and early diagnosis of small HCC tumors may be difficult with one marker alone. Therefore, the false-positive rates for HCC must be carefully considered [39–41]. Also, a combination of markers, including AFP, DCP, and the C3a fragment, in the serum should be verified to improve the diagnostic rate.

The ProteinChip SELDI system can separate and partially characterize multiple proteins in tissue and serum samples. Our previous report used a panel of proteins to diagnose early HCC with the ProteinChip SELDI system [15]. This panel diagnosis of seven protein peaks included a discriminant peak of 4060 m/z. This 4060 m/z peak may be a double-charged 8130 m/z peak, although the C3a fragment (8130 m/z) was not used to develop this diagnostic method. These results suggest that the C3a fragment is a useful HCC biomarker, regardless of whether this fragment carries a single or double charge. In addition, the panel diagnosis method is more useful than measuring the C3a fragment alone to diagnose and predict the occurrence of HCC. However, this method must be performed using the ProteinChip SELDI system, which is expensive and does not detect putative interactions between various proteins. Identifying a specific HCC protein such as the C3a fragment will also further our understanding of the molecular mechanisms of hepatocarcinogenesis. Therefore, the C3a fragment should not only be considered a simple HCC tumor marker, but should also be evaluated for its contribution to HCC carcinogenesis.

In conclusions, serum profiling with the ProteinChip SELDI system may be used to distinguish HCC from chronic liver disease without HCC and to detect early HCC in HCV-infected patients. Because we identified the C3a fragment (8.1 k m/z) in serum samples from HCC patients, the C3a fragment is a promising marker that can be used to screen for HCV-HCC and to develop new therapeutic targets.

Acknowledgments We thank Mr. Hiroyuki Nakao for assistance with the statistical analyses. We also thank Ms. Yuko Morinaga for her technical assistance. This work was supported in part by a grant-in-aid from the Collaboration of Regional Entities for the Advancement of Technological Excellence (CREATE) from the Japan Science and Technology Agency, a grant (no. CA87982) from the United States National Institutes of Health, and a grant-in-aid (Research on Hepatitis and BSE) from the Ministry of Health, Labour and Welfare of Japan.

References

1. El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. *N Engl J Med*. 1999;340:745–50.
2. Robert GG. Hepatocellular carcinoma: overcoming challenges in disease management. *Clin Gastroenterol Hepatol*. 2006;4:252–61.
3. Okuda K. Hepatocellular carcinoma. *J Hepatol*. 2000;32:225–37.
4. Oka H, Tamori A, Kuroki T, Kobayashi K, Yamamoto S. Prospective study of alpha-fetoprotein in cirrhotic patients monitored for development of hepatocellular carcinoma. *Hepatology*. 1994;19:61–6.
5. Ishii M, Gama H, Chida N, Ueno Y, Shinzawa H, Takagi T, et al. Simultaneous measurements of serum alpha-fetoprotein and protein induced by vitamin K absence for detecting hepatocellular carcinoma. South Tohoku District Study Group. *Am J Gastroenterol*. 2000;95:1036–40.
6. Okuda H, Nakanishi T, Takatsu K, Saito A, Hayashi N, Takasaki K, et al. Serum levels of des-gamma-carboxy prothrombin measured using the revised enzyme immunoassay kit with increased sensitivity in relation to clinicopathologic features of solitary hepatocellular carcinoma. *Cancer*. 2000;88:544–9.
7. Grazi GL, Mazziotti A, Legnani C, Jovine E, Miniero R, Gallucci A, et al. The role of tumor markers in the diagnosis of hepatocellular carcinoma, with special reference to the des-gamma-carboxy prothrombin. *Liver Transpl Surg*. 1995;1:249–55.
8. Wang CS, Lin CL, Lee HC, Chen KY, Chiang MF, Chen HS, et al. Usefulness of serum des-gamma-carboxy prothrombin in detection of hepatocellular carcinoma. *World J Gastroenterol*. 2005;11:6115–9.
9. Marrero JA, Su GL, Wei W, Emick D, Conjeevaram HS, Fontana RJ, et al. Des-gamma carboxyprothrombin can differentiate hepatocellular carcinoma from nonmalignant chronic liver disease in American patients. *Hepatology*. 2003;37:1114–21.
10. Mita Y, Aoyagi Y, Yanagi M, Suda T, Suzuki Y, Asakura H. The usefulness of determining des-gamma-carboxy prothrombin by sensitive enzyme immunoassay in the early diagnosis of patients with hepatocellular carcinoma. *Cancer*. 1998;82:1643–8.
11. Taketa K, Okada S, Win N, Hlaing NK, Wind KM. Evaluation of tumor markers for the detection of hepatocellular carcinoma in Yangon General Hospital, Myanmar. *Acta Med Okayama*. 2002;56:317–20.
12. Khien VV, Mao HV, Chinh TT, Ha PT, Bang MH, Lac BV, et al. Clinical evaluation of lentil lectin-reactive alpha-fetoprotein-L3 in histology-proven hepatocellular carcinoma. *Int J Biol Markers*. 2001;16:105–11.
13. Zinkin NT, Grall F, Bhaskar K, Otu HH, Spentzos D, Kalmowitz B, et al. Serum proteomics and biomarkers in hepatocellular carcinoma and chronic liver disease. *Clin Cancer Res*. 2008;14:470–7.
14. Schwegler EE, Cazares L, Steel LF, Adam BL, Johnson DA, Semmes OJ, et al. SELDI-TOF MS profiling of serum for detection of the progression of chronic hepatitis C to hepatocellular carcinoma. *Hepatology*. 2005;41:634–42.
15. Kanmura S, Uto H, Kusumoto K, Ishida Y, Hasuike S, Nagata K, et al. Early diagnostic potential for hepatocellular carcinoma using the SELDI ProteinChip system. *Hepatology*. 2007;45:948–56.
16. Paradis V, Degos F, Dargère D, Pham N, Belghiti J, Degott C, et al. Identification of a new marker of hepatocellular carcinoma by serum protein profiling of patients with chronic liver diseases. *Hepatology*. 2005;41:40–7.
17. Lee IN, Chen CH, Sheu JC, Lee HS, Huang GT, Chen DS, et al. Identification of complement C3a as a candidate biomarker in human chronic hepatitis C and HCV-related hepatocellular carcinoma using a proteomics approach. *Proteomics*. 2006;6:2865–73.
18. Uto H, Hayashi K, Kusumoto K, Hasuike S, Nagata K, Kodama M, et al. Spontaneous elimination of hepatitis C virus RNA in individuals with persistent infection in a hyperendemic area of Japan. *Hepatol Res*. 2006;34:28–34.
19. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem*. 1996;68:850–8.
20. Prahalad AK, Hickey RJ, Huang J, Hoelz DJ, Dobrolecki L, Murthy S, et al. Serum proteome profiles identifies parathyroid hormone physiologic response. *Proteomics*. 2006;6:3482–93.

21. Shiwa M, Nishimura Y, Wakatabe R, Fukawa A, Arikuni H, Ota H, et al. Rapid discovery and identification of a tissue-specific tumor biomarker from 39 human cancer cell lines using the SELDI ProteinChip platform. *Biochem Biophys Res Commun.* 2003;309:18–25.
22. Adam BL, Qu Y, Davis JW, Ward MD, Clements MA, Cazares LH, et al. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res.* 2002;62:3609–14.
23. Sahu A, Lambris JD. Structure and biology of complement protein C3, a connecting link between innate and acquired immunity. *Immunol Rev.* 2001;180:35–48.
24. Miguet L, Bogumil R, Decloquement P, Herbrecht R, Potier N, Mauvieux L, et al. Discovery and identification of potential biomarkers in a prospective study of chronic lymphoid malignancies using SELDI-TOF-MS. *J Proteome Res.* 2006;5:2258–69.
25. Ward DG, Suggett N, Cheng Y, Wei W, Johnson H, Billingham LJ, et al. Identification of serum biomarkers for colon cancer by proteomic analysis. *Br J Cancer.* 2006;94:1898–905.
26. Li J, Orlandi R, White CN, Rosenzweig J, Zhao J, Seregini E, et al. Independent validation of candidate breast cancer serum biomarkers identified by mass spectrometry. *Clin Chem.* 2005;51:2229–35.
27. Jurianz K, Ziegler S, Garcia-Schüler H, Kraus S, Bohana-Kashtan O, Fishelson Z, et al. Complement resistance of tumor cells: basal and induced mechanisms. *Mol Immunol.* 1999;36:929–39.
28. Bjørge L, Hakulinen J, Vintermyr OK, Jarva H, Jensen TS, Iversen OE, et al. Ascitic complement system in ovarian cancer. *Br J Cancer.* 2005;92:895–905.
29. Mollnes TE, Garred P, Bergseth G. Effect of time, temperature and anticoagulants on in vitro complement activation: consequences for collection and preservation of samples to be examined for complement activation. *Clin Exp Immunol.* 1988;73:484–8.
30. Verhaegen H, De Cock W, De Cree J, Verbruggen F. Increase of serum complement levels in cancer patients with progressing tumors. *Cancer.* 1976;38:1608–13.
31. Habermann JK, Roblick UJ, Luke BT, Prieto DA, Finlay WJ, Podust VN, et al. Increased serum levels of complement C3a anaphylatoxin indicate the presence of colorectal tumors. *Gastroenterology.* 2006;131:1020–9.
32. Steel LF, Shumpert D, Trotter M, Seeholzer SH, Evans AA, London WT, et al. A strategy for the comparative analysis of serum proteomes for the discovery of biomarkers for hepatocellular carcinoma. *Proteomics.* 2003;3:601–9.
33. Kawakami T, Hoshida Y, Kanai F, Tanaka Y, Tateishi K, Ikenoue T, et al. Proteomic analysis of sera from hepatocellular carcinoma patients after radiofrequency ablation treatment. *Proteomics.* 2005;5:4287–95.
34. Honda M, Kaneko S, Kawai H, Shirota Y, Kobayashi K. Differential gene expression between chronic hepatitis B and C hepatic lesion. *Gastroenterology.* 2001;120:955–66.
35. Kim W, Oe Lim S, Kim JS, Ryu YH, Byeon JY, Kim HJ, et al. Comparison of proteome between hepatitis B virus- and hepatitis C virus-associated hepatocellular carcinoma. *Clin Cancer Res.* 2003;9:5493–500.
36. Koike K. Steatosis, liver injury, and hepatocarcinogenesis in hepatitis C viral infection. *J Gastroenterol.* 2009;44:82–8.
37. Lok AS, Lai CL. Alpha-fetoprotein monitoring in Chinese patients with chronic hepatitis B virus infection: role in the early detection of hepatocellular carcinoma. *Hepatology.* 1989;9:110–5.
38. Markiewski MM, Mastellos D, Tudoran R, DeAngelis RA, Strey CW, Franchini S, et al. C3a and C3b activation products of the third component of complement (C3) are critical for normal liver recovery after toxic injury. *J Immunol.* 2004;173:747–54.
39. Oka H, Kurioka N, Kim K, Kanno T, Kuroki T, Mizoguchi Y, et al. Prospective study of early detection of hepatocellular carcinoma in patients with cirrhosis. *Hepatology.* 1990;12:680–7.
40. Tanaka N, Horiuchi A, Yamaura T, Komatsu M, Tanaka E, Kiyosawa K. Efficacy and safety of 6-month iron reduction therapy in patients with hepatitis C virus-related cirrhosis: a pilot study. *J Gastroenterol.* 2007;42:49–55.
41. Tsamandas AC, Antonacopoulou A, Kalogeropoulou C, Tsota I, Zabakis P, Giannopoulou E, et al. Oval cell proliferation in cirrhosis in rats. An experimental study. *Hepatol Res.* 2007;37:755–64.

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

Masayuki Kurosaki¹, Yasuhito Tanaka², Nao Nishida³, Naoya Sakamoto⁴, Nobuyuki Enomoto⁵, Masao Honda⁶, Masaya Sugiyama², Kentaro Matsuura², Fuminaka Sugauchi², Yasuhiro Asahina¹, Mina Nakagawa⁴, Mamoru Watanabe⁴, Minoru Sakamoto⁵, Shinya Maekawa⁵, Akito Sakai⁶, Shuichi Kaneko⁶, Kiyooki Ito⁷, Naohiko Masaki⁷, Katsushi Tokunaga³, Namiki Izumi^{1,*}, Masashi Mizokami^{2,7}

¹Division of Gastroenterology and Hepatology, Musashino Red Cross Hospital, Tokyo, Japan; ²Department of Virology, Liver Unit, Nagoya City University, Graduate School of Medical Sciences, Nagoya, Japan; ³Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; ⁴Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Tokyo, Japan; ⁵First Department of Internal Medicine, University of Yamanashi, Yamanashi, Japan; ⁶Department of Gastroenterology, Kanazawa University, Graduate School of Medicine, Kanazawa, Japan; ⁷Research 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

* 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

Journal of Hepatology 2010 vol. xxx | xxx–xxx

Please cite this article in press as: Kurosaki M et al. Pre-treatment prediction of response to pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in *IL28B* and viral factors. J Hepatol (2010), doi:10.1016/j.jhep.2010.07.037

Research Article

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

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

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

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

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

This study aimed to define the pre-treatment prediction of response to PEG-IFN/RBV therapy through the integrated analysis of host factors, such as the *IL28B* genetic polymorphism and various clinical covariates, as well as viral factors, such as mutations in the HCV core and ISDR and serum HCV-RNA load. In addition,

for the general application of these results in clinical practice, decision models for the pre-treatment prediction of response were determined by data mining analysis.

Materials and methods

Patients

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

Laboratory and histological tests

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

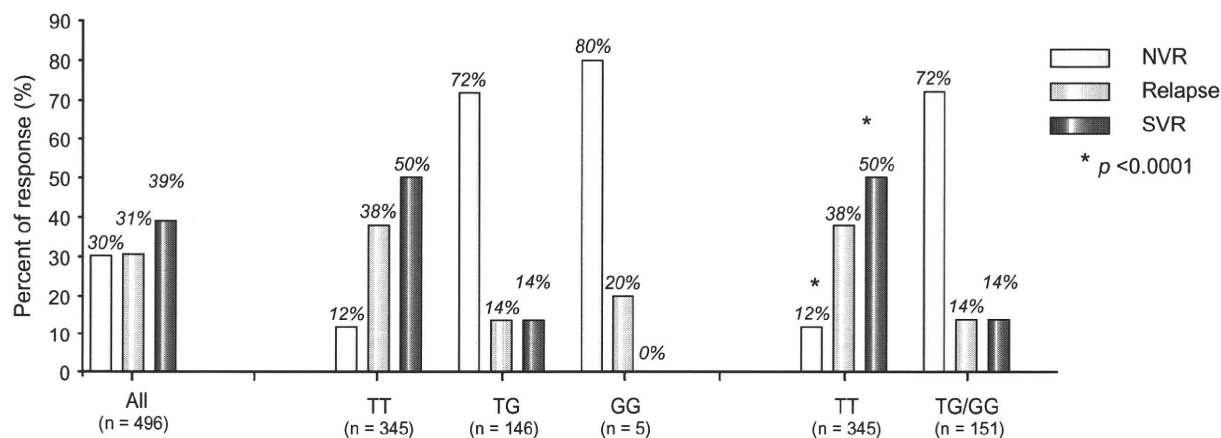


Fig. 1. Association between the *IL28B* genotype (rs8099917) and treatment response. The rates of response to treatment are shown for each rs8099917 genotype. The rate of null virological response (NVR), relapse, and sustained virological response (SVR) is shown. The *p* values are from Fisher's exact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

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

Statistical analysis

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

Results

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

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

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

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

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

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

Factors associated with response by multivariate logistic regression analysis

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

Research Article

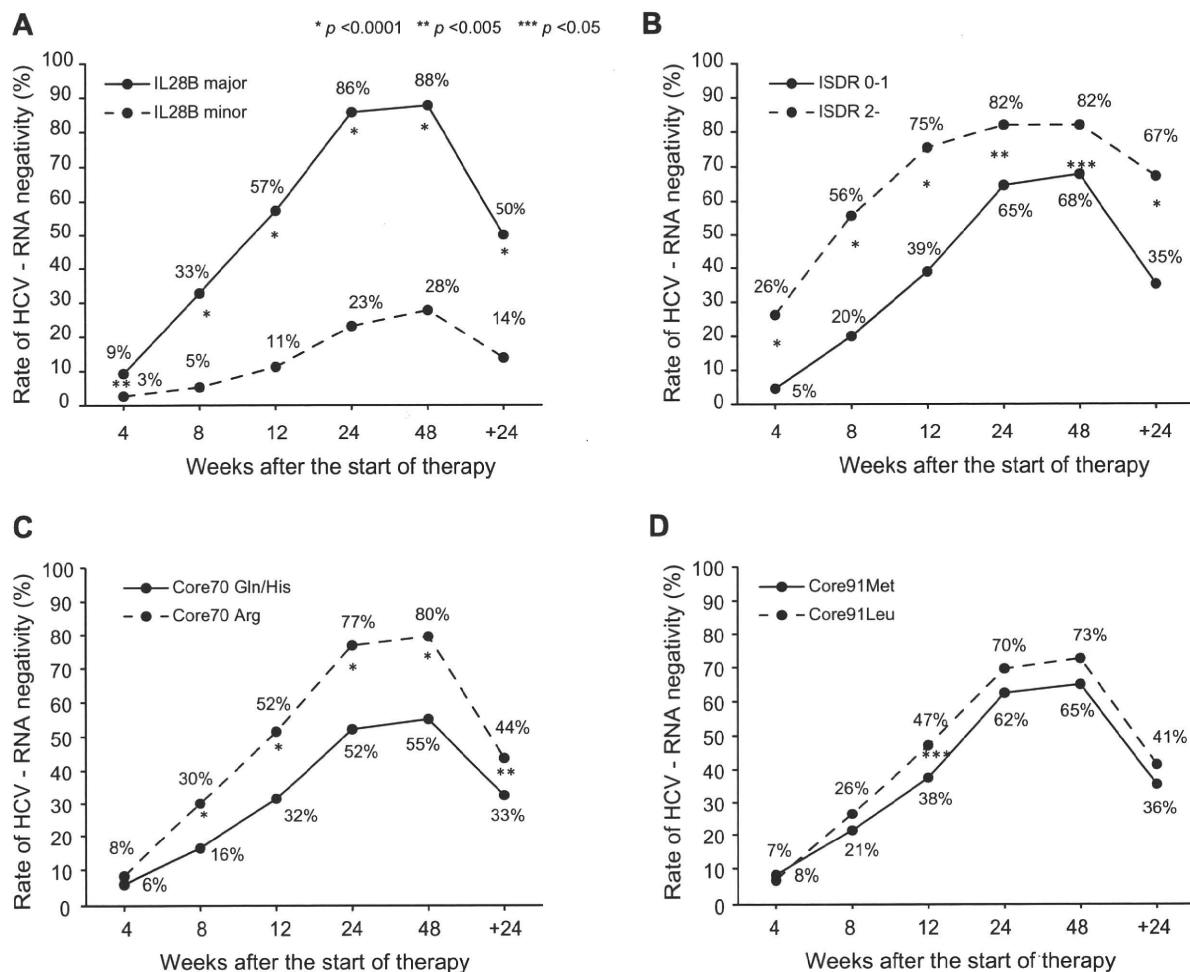


Fig. 2. Effect of *IL28B* mutations in the ISDR, Core70 and Core91 of HCV on time-dependent clearance of HCV. The rate of undetectable HCV-RNA was plotted for serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks) and for 24 weeks after the completion of therapy. Patients were stratified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

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

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

Factors associated with the *IL28B* polymorphism

Patients with the *IL28B* minor allele had significantly higher serum level of gamma-glutamyltransferase (GGT) and a higher

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

Data mining analysis

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

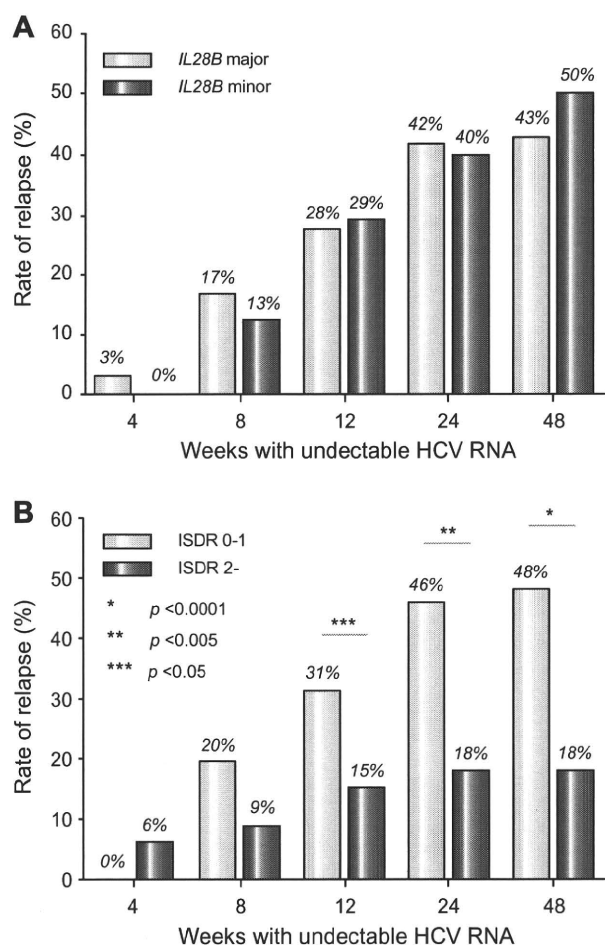


Fig. 3. Association between relapse and the *IL28B* allele or mutations in the ISDR. The rate of relapse was calculated for patients who had undetectable HCV-RNA at serial time points after the start of therapy (4, 8, 12, 24, and 48 weeks). Patients were stratified according to (A) the *IL28B* allele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

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

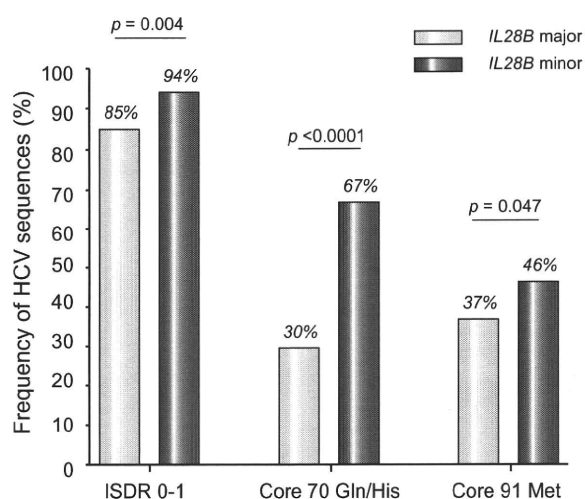


Fig. 4. Associations between the *IL28B* allele and HCV sequences. The prevalence of HCV sequences predicting a resistant phenotype to IFN was higher in patients with the minor *IL28B* allele than those with major allele. (A) 0 or 1 mutation in the ISDR of NS5A, (B) Gln or His at Core70, and (C) Met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

portion of patients with advanced fibrosis (F3–4) was 39% (84/217) in patients with low platelet counts ($<140 \times 10^9/L$) compared to 13% (37/279) in those with high platelet counts ($\geq 140 \times 10^9/L$).

Validation of the data mining analysis

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

Discussion

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

Research Article

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

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

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

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

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

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

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

and whose virological response is determined because the IL28B genotype is not useful for the prediction of relapse. The duration of therapy should be personalized based on the virological response. Future studies need to explore whether the combination of baseline IL28B genotype and response-guided approach further improves the optimization of treatment duration.

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

Table 4. Factors associated with *IL28B* genotype.

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

ALT, alanine aminotransferase; GGT, gamma-glutamyltransferase.

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

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

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

correlation between the SVR and the number of mutations in the ISDR in Japanese as well as in European patients [43] but this correlation was more pronounced in Japanese patients. Thus, geographical factors may account for the different impact of ISDR on treatment response, which may be a potential limitation of our study.

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

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

Research Article

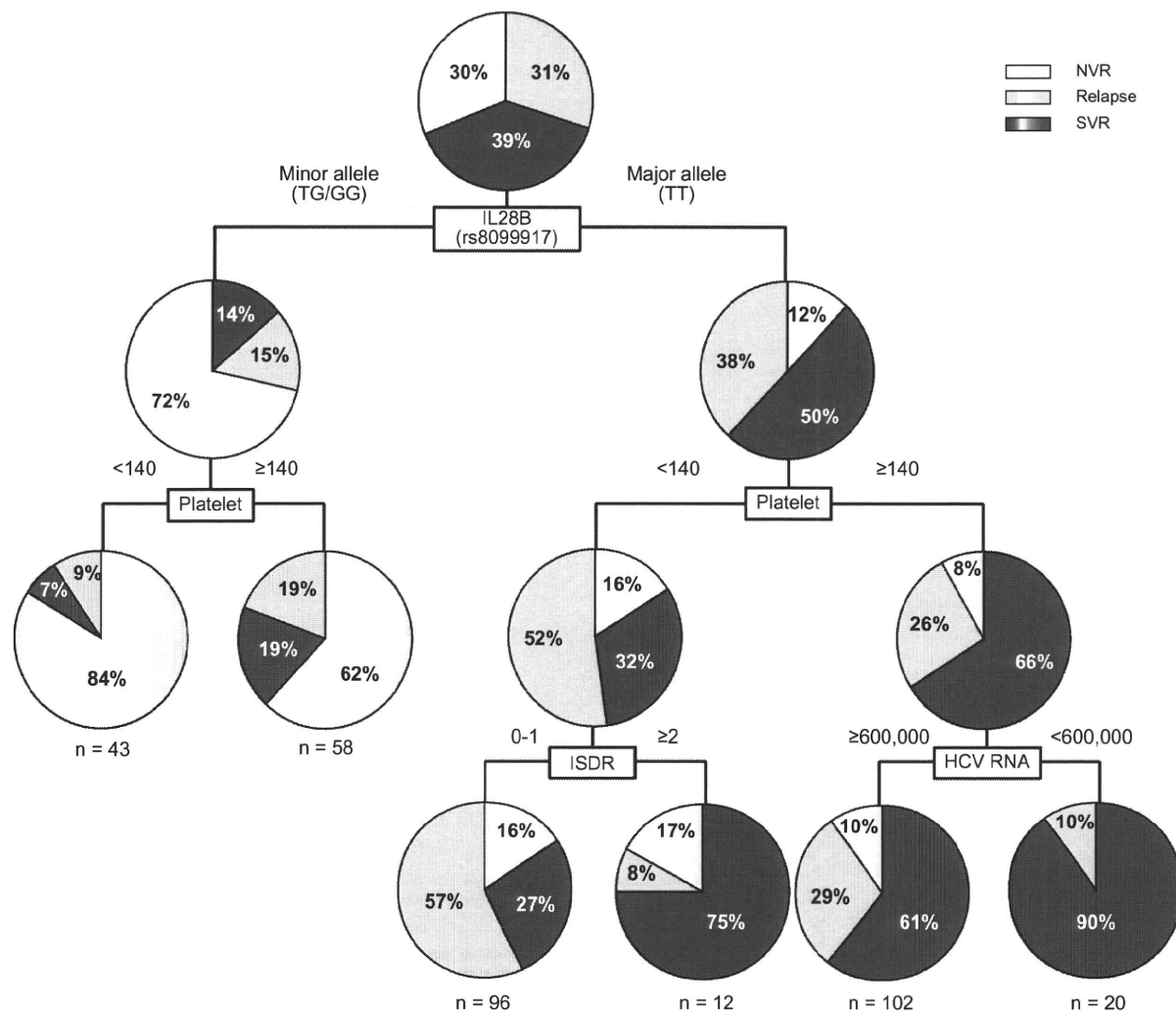


Fig. 5. Decision tree for the prediction of response to therapy. The boxes indicate the factors used for splitting. Pie charts indicate the rate of response for each group of patients after splitting. The rate of null virological response, relapse, and sustained virological response is shown. [This figure appears in colour on the web.]

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

We have previously reported the predictive model of early virological response to PEG-IFN and RBV in chronic hepatitis C

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

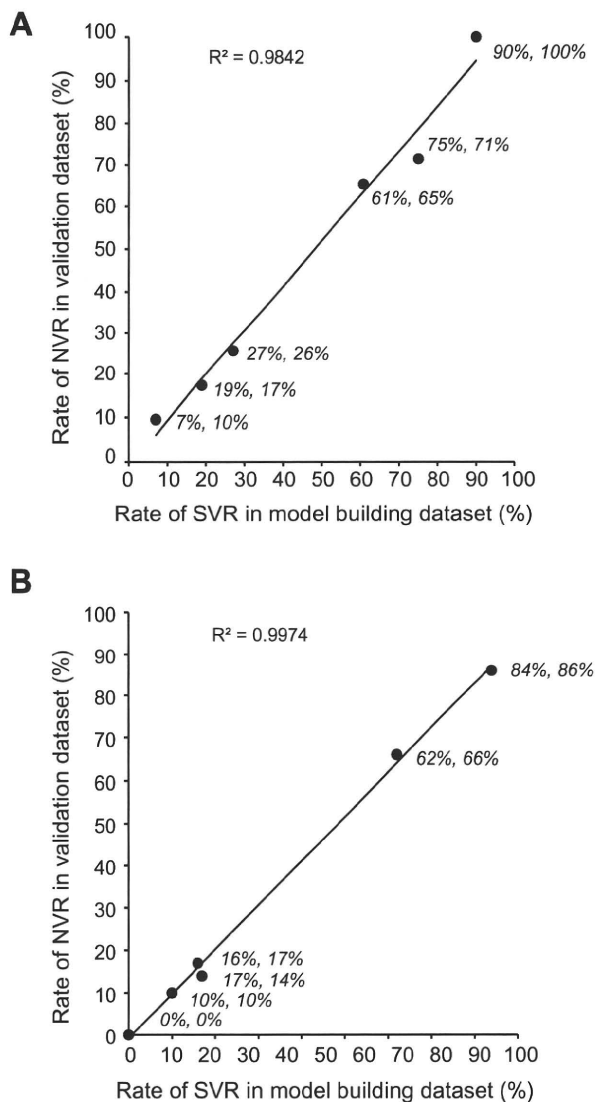


Fig. 6. Validation of the CART analysis. Each patient in the validation group was allocated to one of the six subgroups by following the flow-chart form of the decision tree. The rate of (A) sustained virological response (SVR) and (B) null virological response (NVR) in each subgroup was calculated and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

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

In conclusion, the present study highlighted the impact of the IL28B polymorphism and mutation in the ISDR on the pre-treatment prediction of response to PEG-IFN/RBV therapy. A decision model including these host and viral factors has the potential to

support selection of the optimum treatment strategy for individual patients, which may enable personalized treatment.

Conflicts of interest

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

Financial support

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

References

- [1] Ray Kim W. Global epidemiology and burden of hepatitis C. *Microbes Infect* 2002;4 (12):1219-1225.
- [2] Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncalves Jr FL, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* 2002;347 (13):975-982.
- [3] Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, et al. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* 2001;358 (9286):958-965.
- [4] Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, et al. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. *Intervirology* 2005;48 (6):372-380.
- [5] Davis GL, Wong JB, McHutchison JG, Manns MP, Harvey J, Albrecht J. Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in patients with chronic hepatitis C. *Hepatology* 2003;38 (3):645-652.
- [6] Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* 2009;41:1105-1109.
- [7] Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* 2009;41:1100-1104.
- [8] Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009;461 (7262):399-401.
- [9] Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, Mueller T, et al. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* 2010;138 (4):1338-1345.
- [10] Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *J Clin Invest* 1995;96 (1):224-230.
- [11] Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *N Engl J Med* 1996;334 (2):77-81.
- [12] Kurosaki M, Enomoto N, Murakami T, Sakuma I, Asahina Y, Yamamoto C, et al. Analysis of genotypes and amino acid residues 2209 to 2248 of the NS5A region of hepatitis C virus in relation to the response to interferon-beta therapy. *Hepatology* 1997;25 (3):750-753.
- [13] Shirakawa H, Matsumoto A, Joshita S, Komatsu M, Tanaka N, Umemura T, et al. Pretreatment prediction of virological response to peginterferon plus ribavirin therapy in chronic hepatitis C patients using viral and host factors. *Hepatology* 2008;48 (6):1753-1760.
- [14] Akuta N, Suzuki F, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, et al. Predictive factors of early and sustained responses to peginterferon plus ribavirin combination therapy in Japanese patients infected with hepatitis C virus genotype 1b: amino acid substitutions in the core region and low-density lipoprotein cholesterol levels. *J Hepatol* 2007;46 (3):403-410.

Research Article

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

HCV Genetic Elements Determining the Early Response to Peginterferon and Ribavirin Therapy

Nobuyuki Enomoto Shinya Maekawa

First Department of Medicine, Faculty of Medicine, University of Yamanashi, Chuo, Japan

Key Words

Full open reading frame analysis · Hepatitis C virus · Peginterferon/ribavirin therapy

Abstract

The aim of this study was to search hepatitis C virus (HCV) genetic elements determining the early response to peginterferon/ribavirin therapy using HCV genome-wide analysis. From a total of 88 chronic hepatitis C patients with HCV-1b treated with peginterferon/ribavirin, the whole HCV amino acid sequence was determined and analyzed according to the viral response during the treatment. Mutations in NS5A-ISDR (interferon sensitivity-determining region) are associated with rapid viral response at week 4, and the core arginine70glutamine (R70Q) mutation is associated with no early viral response at week 12, revealing that core 70 and NS5A are the most important factors determining the virological kinetics during peginterferon and ribavirin therapy.

Copyright © 2010 S. Karger AG, Basel

Introduction

Hepatitis C virus (HCV) is a major cause of chronic liver diseases, and worldwide 170 million people are infected with HCV. With the introduction of the recent

combination therapy of pegylated-interferon (PEG-IFN) and ribavirin (RBV), half of patients can eradicate the virus (sustained virological response, SVR). The SVR rate of HCV to the PEG-IFN/RBV therapy is dependent on HCV genotypes, and the viral kinetics during the treatment strongly affect the final viral clearance [1, 2]. It is generally considered that HCV structures affect the treatment response since the SVR rate to PEG-IFN/RBV therapy depends upon viral genotypes as described above. However, comprehensive analysis of the contribution of HCV structures to different responses has not yet been conducted. In the present study, in order to clarify the relationship between HCV sequences and viral responses, we have determined the complete HCV open reading frame sequences obtained from pretreatment patients' serum, and investigated their response by searching for HCV genetic elements determining the early response to PEG-IFN/RBV therapy using HCV genome-wide analysis.

Methods

A total of 88 chronic hepatitis C patients with HCV-1b treated with PEG-IFN/RBV were studied. From pretreatment sera, the whole HCV deduced amino acid sequence (3,010 amino acids) was determined in each patient by direct RT-PCR.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2010 S. Karger AG, Basel
0300-5526/10/0531-0066\$26.00/0

Accessible online at:
www.karger.com/int

Nobuyuki Enomoto, MD
First Department of Medicine, Faculty of Medicine
University of Yamanashi
1110 Shimokato, Chuo, Yamanashi 409-3898 (Japan)
Tel. +81 55 273 9584, Fax +81 55 273 6748, E-Mail enomoto@yamanashi.ac.jp

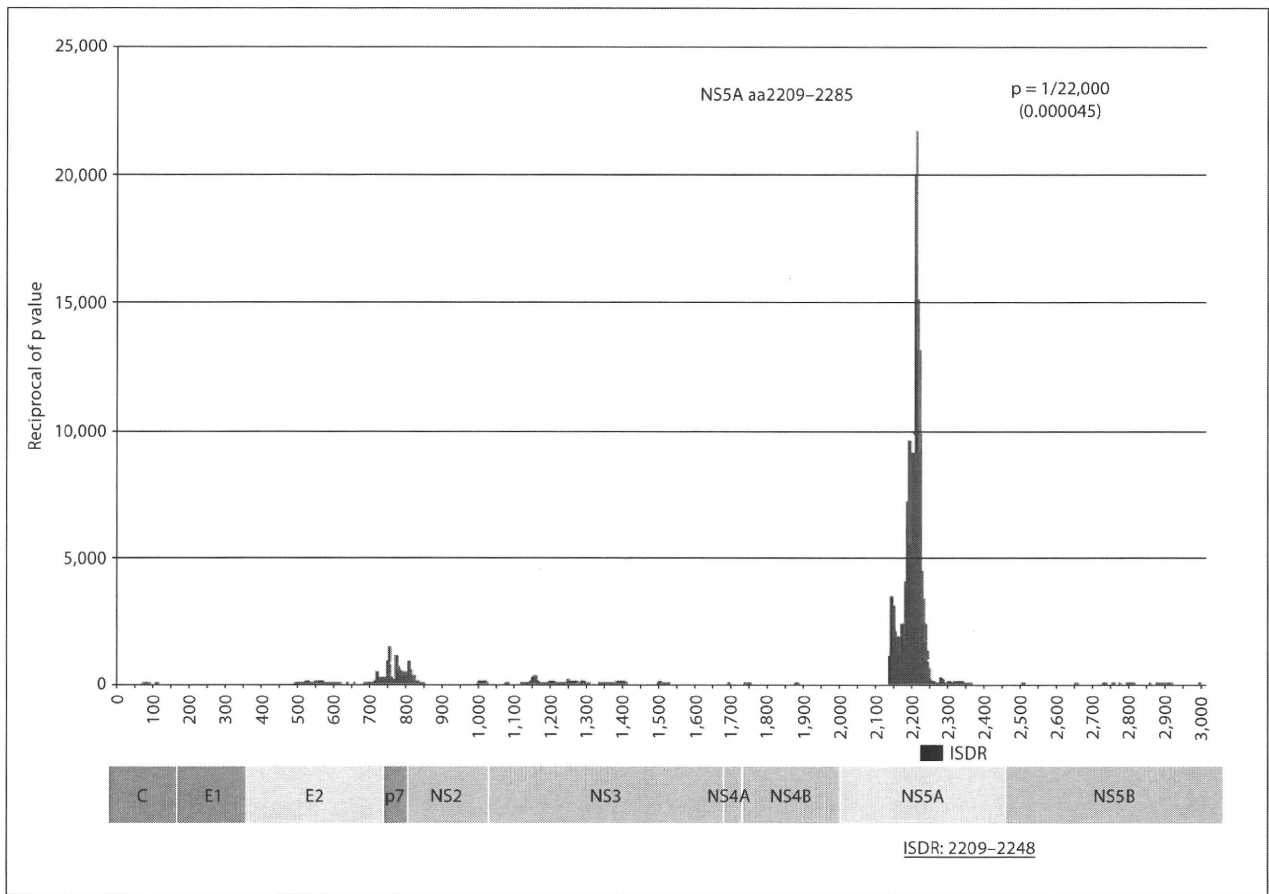


Fig. 1. Reciprocal of p value for sliding window analysis with 77 amino acid width for RVR versus others.

Amino acid usage of each of the 3,010 positions was compared according to the different virological response in order to identify the single amino acid differences determining the virological response. In addition, sliding window analyses were carried out in order to identify the amino acid region associated with the virological response. The number of the amino acid changes in the fixed stretch of the sequence (window: 2–100 amino acids) were compared according to the virological response, scanning the whole HCV amino acid sequence by sliding this window one by one.

Results

Of 88 patients studied, 9 showed rapid viral response (RVR; HCV-RNA undetectable at week 4) and 71 showed early viral response (EVR; over 2-log drop of HCV-RNA at week 12). The other 17 patients showed no EVR, indicating these patients are highly resistant to the treatment.

Mutations in the region overlapping NS5A-ISDR (interferon sensitivity-determining region, aa2209–2248) are associated with the good response to PEG-IFN/RBV therapy as shown in sliding window analysis comparing RVR patients at week 4 and others (fig. 1). In contrast, the core R(arginine)70Q (glutamine) mutation is associated with a poor response resulting in no EVR at week 12 by single amino difference analysis comparing non-EVR patients and the others (fig. 2).

Discussion

In the present study, using a sliding window analysis comparing all HCV amino acids, the amino acid region located in ISDR was extracted as the most significant region discriminating the RVR and non-RVR patients. By