38. Tong X, Arasappan A, Bennett F, et al. Preclinical characterization of the antiviral activity of SCH 900518 (Narlaprevir), a novel mechanism-based inhibitor of hepatitis C virus NS3 protease. Antimicrob Agents Chemother 2010;54:2365–2370

39. Chase R, Skelton A, Xia E, et al. A novel HCV NS3 protease mutation selected by combination treatment of the protease inhibitor boceprevir and NS5B polymerase inhibitors. Antiviral Res 2009;84:178–184

Hepatol Int
DOI 10.1007/s12072-011-9307-6

ORIGINAL ARTICLE

Analysis of viral amino acids sequences and the IL28B SNP influencing the development of hepatocellular carcinoma in chronic hepatitis C

Mika Miura · Shinya Maekawa · Makoto Kadokura · Ryota Sueki · Kazuki Komase · Hiroko Shindo · Takako Ohmori · Asuka Kanayama · Kuniaki Shindo · Fumitake Amemiya · Yasuhiro Nakayama · Takatoshi Kitamura · Tomoyoshi Uetake · Taisuke Inoue · Minoru Sakamoto · Shunichi Okada · Nobuyuki Enomoto

Received: 23 April 2011 / Accepted: 22 July 2011

Abstract
Background and aims The association between hepatitis C virus (HCV) sequences with interleukin 28B (IL28B) single-nucleotide polymorphism (SNP) in the development of hepatocellular carcinoma (HCC) has not been well clarified.
Methods Complete HCV open-reading frame sequences were determined in 20 patients developing HCC and 23 non-HCC patients with HCV-1b infection in two distant time points. An additional 230 patients were studied cross-sectionally for core and NS5A sequences with HCC development. Among them, 98 patients with available samples were investigated for changes in viral core sequences over time. Finally, IL28B SNPs and HCC development were investigated in 228 patients.
Results During observation period (HCC for 10.8 years, and non-HCC for 11.1 years), changes in core a.a. 70 and three amino acid positions in NS5A were characteristics of the patients developing HCC. In 230 patients, Q (glutamine) or H (histidine) to R (arginine) ratio at core a.a. 70 was significantly higher in the HCC group (HCC group 43:22 vs. non-HCC group 66:99, $p = 0.001$). A change in core R70Q was observed over time in 11 patients associated with a decrease in platelets ($p = 0.005$) and albumin ($p = 0.005$), while a Q70R change was observed in 4 patients without associated changes in platelets (nonsignificant) and albumin (nonsignificant). IL28B SNP showed significant correlation with the core a.a. 70 residue. There was no evident link between IL28B SNPs and the occurrence of HCC.
Conclusions Hepatitis C virus core a.a. 70 residue is associated with liver disease progression and is independent factor for HCC development in genotype-1b infection. IL28B SNPs are related to core a.a. 70 residue, but not to HCC. The functional relevance of core a.a. 70 residue in hepatitis C pathogenesis should be further investigated.

Keywords HCV · HCC · Core · IL28B

Introduction

Hepatitis C virus (HCV) infection is a major risk factor for hepatocellular carcinoma (HCC). Chronic HCV infection can result in liver cirrhosis (LC) and HCC over the course of 20–30 years [1]. However, the rate of progression is variable; some patients remain for a long time with persistently normal ALT values, while others progress rapidly to LC and HCC.

Viral factors, host factors, and their interplay appear to play an important role in determining the progression of chronic hepatitis C to LC and HCC. In terms of viral factors, most previous clinical studies have focused on searching for HCV regions correlated with the response to interferon (IFN)-based therapy. In those analyses, correlation between amino acid substitutions and treatment response have been reported for the IFN sensitivity

M. Miura and S. Maekawa have contributed equally to this study.

Electronic supplementary material The online version of this article (doi:10.1007/s12072-011-9307-6) contains supplementary material, which is available to authorized users.

M. Miura · S. Maekawa (✉) · M. Kadokura · R. Sueki · K. Komase · H. Shindo · T. Ohmori · A. Kanayama · K. Shindo · F. Amemiya · Y. Nakayama · T. Kitamura · T. Uetake · T. Inoue · M. Sakamoto · S. Okada · N. Enomoto
First Department of Internal Medicine, Faculty of Medicine, University of Yamanashi, 1110, Shimokato, Chuo, Yamanashi 409-3898, Japan
e-mail: maekawa@yamanashi.ac.jp

Published online: 17 August 2011

determining region in nonstructural (NS)5A [2], a.a. 70 and 91 in core [3], the PKR/eIF-2alpha phosphorylation homology domain (PePHD) in envelope (E)2 [4], and the IFN–ribavirin resistance determining region (IRRDR) in NS5A [5].

Regarding the viral factors related to disease progression, among the various HCV proteins, the core protein has been thought widely to contribute because it has been shown experimentally to affect multiple cellular functions, in addition to the evidence from clinical studies [6–10]. Core protein modifies cellular apoptosis, oncogenic signaling, reactive oxygen species formation, lipid metabolism, transcriptional activation, transformation, and immune reactivity. Core protein has oncogenic potential in transgenic mice [11]. In contrast, fewer clinical studies to date have systematically investigated the correlation between the variability of HCV regions and disease progression. However, some of those limited clinical studies reported a correlation between amino acid substitutions in core or NS5A with disease progression [12–15]. Despite those reports, few studies to date support the correlation. Moreover, it is unclear whether those viral sequences change during disease progression or how the disease activity is modified by those viral sequences in the long course of chronic hepatitis.

On the other hand, regarding host factors, recent reports disclosed a significant correlation between polymorphisms in the IL28B gene and responses to pegylated-IFN plus ribavirin therapy for HCV patients [16–19]. This single-nucleotide polymorphism (SNP) also showed significant correlation with natural HCV clearance [20]. However, it remains unknown whether the IL28B SNP is related to disease progression or the development of HCC.

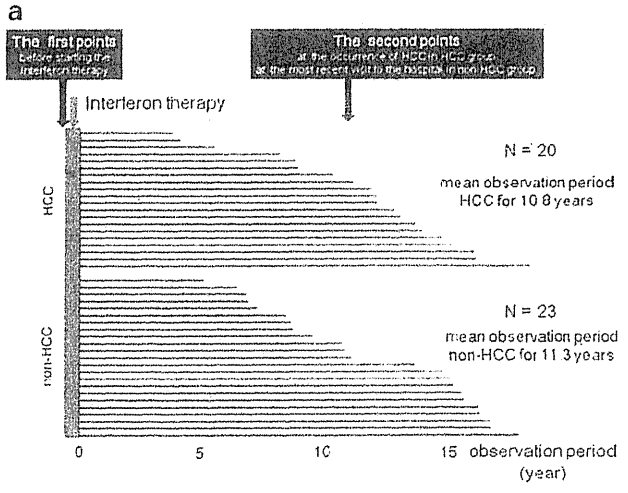
In this study, we first undertook the analysis to identify the viral regions related to disease progression and HCC development through the analysis of complete HCV open-reading frame (ORF) sequences. Because some regions in HCV core and NS5A showed characteristic changes over time in patients developing HCC during the observation period, we proceeded further to analyze the contribution of those regions to disease progression, in association with time and with the IL28B SNP.

Patients and methods

Patients

This study is based on the analysis of two groups of patients, 43 in Group 1 and 230 patients in Group 2.

In the first part, we tried to characterize and extract viral sequences specific to disease progression through the analysis of complete HCV ORFs (Fig. 1a). In particular, we focused our investigation on the changes in viral sequences over time in association with disease progression by comparing HCV sequences of two sufficiently distant time points. With this aim, we determined to investigate patients with a history of IFN therapy, because those patients often were followed long-term with preservation of old and recent sera. However, we excluded sustained



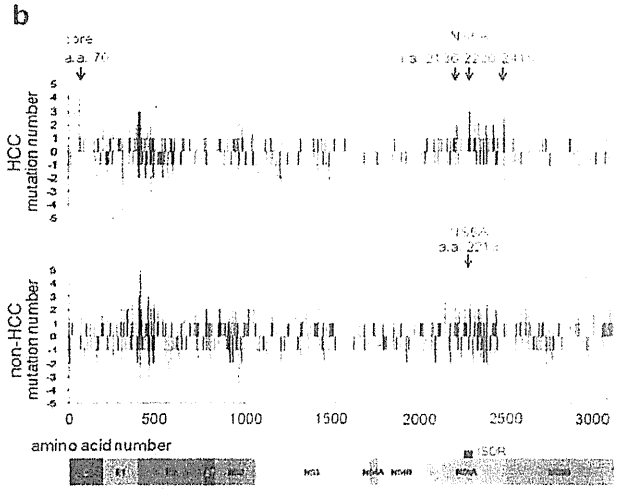


Fig. 1 **a** A total of 43 patients were analyzed for complete HCV ORF sequences. They were all non-responders to the previous IFN therapy. Twenty patients developed HCC during the observation period, while the 23 patients did not. HCV ORF sequences were determined for the paired samples and the predicted amino acid changes were compared in each patient. **b** Specific HCV amino acid changes associated with disease progression was evaluated by the analysis of the full-length viral ORF during the observation period for each patient according to the following rules: *1* +1 point for consensus to non-consensus, *2* −1 point for non-consensus to consensus, *3* 0 point for non-consensus to non-consensus. These points were added together and are shown for HCC and non-HCC patients. Patient with later HCC development (*upper panel*). Non-HCC patients (*lower panel*)

virologic response (SVR) patients because we thought that the viral clearance leads to improvement of the liver disease and, therefore, viral regions influencing the IFN response would be extracted as affecting the course of disease. Between March 1992 and April 2004, 273 consecutive patients with HCV-1b infection were given IFN monotherapy at Yamanashi University Hospital, and 133 were followed long-term. A total of 65 patients showed SVR, while 68 showed non-SVR. Among these 68 non-SVRs, 43 patients were included in the study because laboratory data and sera were available from the two distant time points (Group 1). Twenty patients developed HCC during the observation period, while the remaining 23 did not. Regarding the sera, the first time point for both groups was before starting IFN therapy, while the second points were at the occurrence of HCC for the HCC group and at the most recent visit to the hospital for the non-HCC group (Fig. 1a).

An additional 230 HCV-1b patients were recruited (Group 2) as the second study group. They were mainly outpatients at Yamanashi University Hospital and were selected randomly from those with stored sera at the time of disease diagnosis. Sixty-five had HCC, while the remaining 165 did not. They all were positive for HCV RNA at the time of study, although 74 patients had a history of IFN therapy. Parts of the core and NS5A sequences were determined at HCC onset in 65 HCC patients and at the most recent visit to the hospital in 165 non-HCC patients. Because historical sera around 10 years before were also available for 55 of these 230 patients, HCV sequence analysis was also performed for core and NS5A at those previous time points in those patients.

From these two study groups, 228 patients (68 HCCs and 160 non-HCCs) with available genomic DNAs were examined to determine the IL28B SNP.

All the patients studied all fulfilled following criteria: (1) Negative for hepatitis B surface antigen. (2) No other forms of hepatitis, such as primary biliary cirrhosis, autoimmune liver disease, or alcoholic liver disease. (3) Free of coinfection with human immunodeficiency virus. (4) A signed consent was obtained for the study protocol that had been approved by Human Ethics Review Committee of Yamanashi University Hospital.

Complete and partial HCV ORF sequence determination by direct sequencing from sera

HCV RNA extraction, complementary DNA synthesis and amplification by two-step nested PCR from serum samples were done using the specific primers for full HCV ORF or partial viral regions as described previously [15]. PCR amplicons were sequenced directly by Big Dye Terminator Version 3.1 (ABI, Tokyo, Japan) with universal M13 forward and reverse primers using an ABI prism 3130 sequencer (ABI). Generated sequence files were assembled using Vector NTI software (Invitrogen, Tokyo, Japan) and base-calling errors were corrected following inspection of the chromatogram.

IL28B SNP analysis

Human genomic DNA was extracted from peripheral blood using a blood DNA extraction kit (QIAGEN, Tokyo, Japan) according to the manufacturer's protocol. The allele typing of each DNA sample was performed by real-time PCR with a model 7500 (ABI) using FAM-labeled SNP primer for the locus rs8099917 (ABI).

Statistical analysis

Statistical differences in the parameters, including all available patients' demographic, biochemic, hematologic, and virologic data, were determined between different groups of patients by Student's t test for numerical variables and Fisher's exact probability test for categorical variables. Odds ratios and their 95% confidence intervals were used to quantify the level of association. All p values of <0.05 by the two-tailed test were considered significant throughout. Multiple logistic regression analyses were used to identify the independent variables influencing core a.a. 70 residue and HCC development. Because most variables used for the analyses were generally considered to correlate with the disease progression, we entered all the variables into the multiple logistic regression analysis even if some of them did not reach significant differences in individual univariate analysis.

Results

Comparing complete HCV amino acid sequences between patients with and without HCC

The clinical characteristics of the 43 patients (Group 1) analyzed for HCV ORF changes over time are shown in Table 1. At the start of observation, clinical characteristics did not differ significantly between the HCC group and the non-HCC group. The mean observation period was comparable between the two groups and was 10.8 years for the HCC group and 11.3 years for non-HCC group ($p = 0.745$). On the other hand, platelets ($p < 0.001$), albumin ($p < 0.001$), and AFP ($p = 0.001$) became significantly lower or higher in the HCC group at the end of observation (Table 1).

We proceeded to investigate viral amino acid changes during the course of disease in each patient to determine

Table 1 Patient characteristics in Group 1

	At the start of observation			At the end of observation		
	HCC ($N = 20$)	Non-HCC ($N = 23$)	p value	HCC ($N = 20$)	Non-HCC ($N = 23$)	p value
Observation period (years)				10.8 ± 3.6	11.3 ± 3.8	0.745
Sex (male/female)	11/9	12/11	0.999	11/9	12/11	0.999
Age (years)	51.5 ± 8.0	50.0 ± 9.9	0.604	61.7 ± 10.0*	61.0 ± 10.9*	0.818
Stage of fibrosis (F1/2/3/4)	1/7/6/6	5/11/4/3	0.190	N/A	N/A	–
AST (IU/L)	102 ± 114	74 ± 40	0.695	71 ± 36*	51 ± 30	0.048
ALT (IU/L)	124 ± 86	104 ± 71	0.411	69 ± 47*	52 ± 31*	0.159
Platelets ($10^{-4}/mm^3$)	16.2 ± 4.8	18.3 ± 6.2	0.217	9.7 ± 3.9*	15.3 ± 5.1	<0.001
Albumin (g/dL)	4.1 ± 0.4	4.1 ± 0.2	0.639	3.6 ± 0.4	4.1 ± 0.5	<0.001
γGTP (IU/L)	90 ± 60	71 ± 46	0.275	69 ± 59	45 ± 38*	0.114
T.Chol (mg/dL)	169 ± 28	156 ± 22	0.110	146 ± 21	164 ± 5,108	0.086
Alpha-fetoprotein (ng/mL)	10.5 ± 6.8	9.3 ± 10.8	0.695	42.4 ± 41.1	4.7 ± 2.7	0.001
HCV RNA concentration (kIU/mL)	706 ± 696	614 ± 1,181	0.760	3,325 ± 415*	4,508 ± 5,108*	0.426

* Factors with significant changes over time (<0.05)

whether specific amino acid changes related to disease progression could be identified. First, the consensus amino acid was determined at each amino acid position in the HCV ORF after determination of all sequences in these 43 patients. Amino acid changes were determined according to the following rules to highlight directional changes according to disease progression: When an amino acid changed from the consensus to the non-consensus during the observation period, we scored +1 point. Conversely, a change from the non-consensus to the consensus scored −1 point. We scored 0 point for a change from one non-consensus amino acid to another. As shown in Fig. 1b, directional amino acid changes were observed throughout the HCV genome to some degree both in patients with and without HCC development during the clinical course of almost 10 years, and frequent substitutions in E2 hypervariable region were common in both groups. On the other hand, in patients with HCC development, as many as four directional changes were observed at core a.a. 70 and at three amino acid positions of NS5A (Fig. 1b, upper panel). In contrast, in patients without HCC, the significant change ($n = 4$) was observed at a.a. 2,218 of NS5A when E2 hypervariable region was excluded (Fig. 1b, lower panel).

Core and NS5A sequences in patients with and without HCC

Because the first analysis suggested that the patients with later HCC development might accumulate specific mutations in core and NS5A at the time of HCC occurrence, additional sequences were analyzed from 230 HCV-1b patients to confirm the result. The clinical backgrounds of the additional 230 patients are shown in Table 2 (Group 2).

Table 2 Patient characteristics in Group 2

	HCC ($N = 65$)	Non-HCC ($N = 165$)	p value
Observation period (years)			
Sex (male/female)	42/23	76/89	0.018
Age (years)	68.2 ± 9.2	62.4 ± 11.7	<0.001
AST (IU/L)	66 ± 35	41 ± 21	<0.001
ALT (IU/L)	67 ± 47	44 ± 47	<0.001
Platelets ($10^{-4}/mm^3$)	11.3 ± 5.8	15.3 ± 6.2	<0.001
Albumin (g/dL)	3.6 ± 0.5	4.4 ± 2.9	0.025
γGTP (IU/L)	59 ± 53	38 ± 40	0.001
T.Chol (mg/dL)	153 ± 30	165 ± 31	0.004
Alpha-fetoprotein (ng/mL)	302 ± 1,670	10 ± 25	0.025
HCV RNA concentration (kIU/mL)	5,400 ± 13,574	7,990 ± 8,512	0.104

All patients were positive for HCV RNA. Between the HCC (65 patients) and non-HCC (165 patients) groups, HCC patients were older ($p < 0.001$) and more frequently tended to be males ($p = 0.018$). Moreover, AST, ALT, γGTP, and AFP were significantly higher, and platelets, albumin, and cholesterol were significantly lower in the HCC group. Different predicted amino acids in the core and NS5A regions, between the two groups, are demonstrated in Fig. 2a. The ratio of the core a.a. 70Q (glutamine) or H (histidine) to R (arginine) was significantly higher with the existence of HCC as demonstrated in Fig. 2a (left panel). On the other hand, evident correlations were not confirmed between mutations in NS5A and disease progression (Fig. 2a, right panel). The ratio of Q or H

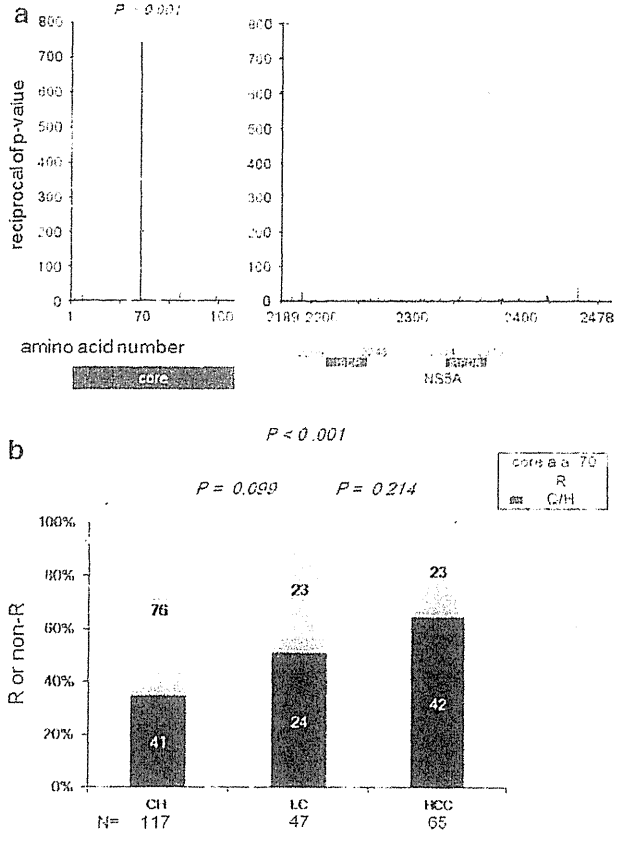


Fig. 2 Core and NS5A sequences in additional patients were studied. **a** Using sera from 230 additional patients at the time of diagnosis, amino acid usage was compared between HCC and non-HCC patients for the part of core and the NS5A region, and this difference is shown as the *bar height* expressed as reciprocal *p* values. **b** In 230 patients, the association between polymorphisms of core a.a. 70 and the state of liver disease (chronic hepatitis, LC, or HCC) is shown

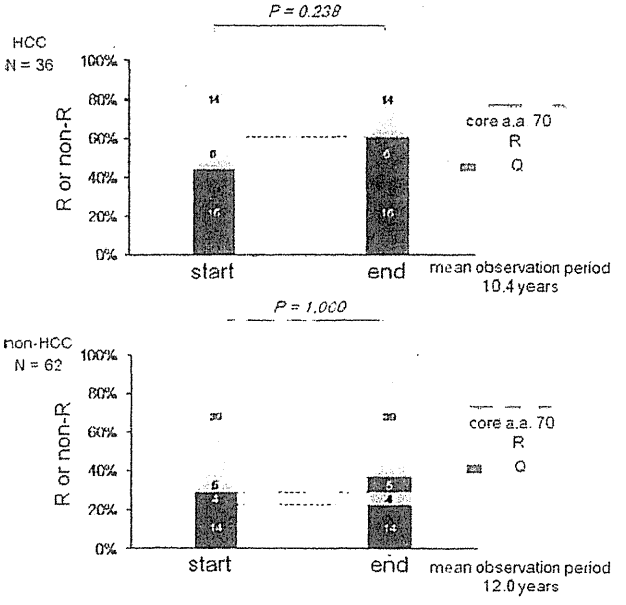


Fig. 3 Changes in core a.a. 70 over time were studied. In 98 patients available for the analysis of 2, sufficiently distant time points, amino acid changes of core a.a. 70 were investigated over time in each patient group. Patients with later HCC development (*upper panel*). Non-HCC patients (*lower panel*)

to R progressively increased in patients in the three major groups of disease activity: chronic hepatitis, cirrhosis, and HCC (Fig. 2b). The association between the disease progression and core a.a. 70 polymorphism also was observed irrespective of IFN-based therapy (data not shown).

Changes in core a.a. 70 over time in patients with and without HCC

We then examined changes in core a.a. 70 over time in association with disease progression (Fig. 3). For this analysis, 55 patients from Group 2, for whom sera from two distant time points were available, were added to the 43 patients in Group 1 and a total of 98 patients were enrolled. When they were classified into two groups according to later HCC onset, the mean observation period was comparable between the groups, 10.4 years for the HCC group and 12 years for the non-HCC group. The occurrence of core a.a. 70Q was 61% (22/36) at the time of HCC onset in the HCC group (Fig. 3, upper panel), but 31% (19/62) for the non-HCC group (Fig. 3, lower panel). In contrast, it was 44% (16/36) at the start of observation in the HCC group (Fig. 3, upper panel) and 29% (18/62) for the non-HCC group (Fig. 3, lower panel). Regarding the core a.a. 70 changes over time, R was dominant throughout the observation period in the non-HCC group (71% at the start and 69% at the end), while the dominant amino acid changed from R (56%) to Q (61%) in the HCC group, so that the core a.a. 70 residues of 17% of the HCC patients changed from R to Q during the course of HCC development. In other words, 6 of 22 patients with 70Q at the onset of HCC had 70R originally (6/22, 27%), while 0 of 14 (0%) with 70R at the most recent observation time had 70Q at the beginning. There were no patients with core a.a. 70H throughout the observation period in this population. The result demonstrates that the relationship between 70Q and HCC development is significant at the time of HCC development. At the start of observation, there was also a tendency that the patients with 70Q compared with 70R develop HCC. However, this difference did not reach statistical significance as shown in Supplementary Fig. 1.

Core a.a. 70 changes over time and their association with disease progression

These 98 patients were classified into four groups according to the pattern of core a.a. 70 change (Table 3) and their

Table 3 Progression of liver disease in 98 patients categorized by core a.a. 70 changes over time

	R → R ($N = 53$)			Q → R ($N = 4$)		
	Start	End	p value	Start	End	p value
HCC rate (HCC/non-HCC)	–	26.4% (14/39)	–	–	0% (0/4)	
Sex (male/female)	–	25/28	–	–	4/0	
Observation period (years)	–	11.1 ± 3.4	–	–	12.9 ± 3.5	
Age (years)	51.3 ± 11.7	62.4 ± 12.1	<0.001	48.0 ± 11.6	61.0 ± 9.1	0.128
AST (IU/L)	68 ± 73	48 ± 26	0.066	56 ± 32	83 ± 61	0.456
ALT (IU/L)	80 ± 71	48 ± 35	0.003	114 ± 71	96 ± 42	0.678
Platelets ($10^{-4}/mm^3$)	17.0 ± 5.8	15.0 ± 6.7	0.104	21.3 ± 3.9	17.2 ± 5.2	0.251
Albumin (g/dL)	4.1 ± 0.4	3.9 ± 0.6	0.225	4.4 ± 0.4	4.3 ± 0.4	0.647
γGTP (IU/L)	56 ± 51	38 ± 40	0.052	95 ± 51	61 ± 46	0.371
T.Chol (mg/dL)	172 ± 36	158 ± 33	0.032	152 ± 14	175 ± 32	0.222
Alpha-fetoprotein (ng/mL)	8.3 ± 9.5	12.5 ± 22.1	0.202	6.0 ± 6.0	5.2 ± 2.2	0.816
HCV RNA concentration (kIU/mL)	4,634 ± 8,509	7,070 ± 14,159	0.291	5,798 ± 7,970	13,676 ± 1,881	0.162
	R → Q ($N = 11$)			Q → Q ($N = 30$)		
	Start	End	p value	Start	End	p value
HCC rate (HCC/non-HCC)		54.5% (6/5)			53.3% (16/14)	
Sex (male/female)		6/5			13/17	
Observation period (years)		13.7 ± 1.65			10.8 ± 3.5	
Age (years)	56.4 ± 7.5	69.3 ± 9.3	0.002	54.6 ± 8.5	64.9 ± 9.9	<0.001
AST (IU/L)	62 ± 47	46 ± 12	0.285	79 ± 51	60 ± 31	0.087
ALT (IU/L)	100 ± 69	37 ± 15	0.008	95 ± 58	59 ± 36	0.006
Platelets ($10^{-4}/mm^3$)	17.7 ± 3.9	11.8 ± 4.8	0.005	16.3 ± 6.5	11.9 ± 5.6	0.007
Albumin (g/dL)	4.2 ± 0.2	3.8 ± 0.4	0.005	4.1 ± 0.3	3.8 ± 0.5	0.009
γGTP (IU/L)	73 ± 53	33 ± 16	0.025	101 ± 55	71 ± 65	0.065
T.Chol (mg/dL)	157 ± 21	144 ± 27	0.245	163 ± 28	150 ± 32	0.100
Alpha-fetoprotein (ng/mL)	7.1 ± 4.3	97.8 ± 63.6	0.267	20.8 ± 50.0	35.1 ± 54.7	0.295
HCV RNA concentration (kIU/mL)	2,415 ± 3,163	2,349 ± 1,851	0.957	2,869 ± 3,984	3,229 ± 4,026	0.731

clinical characteristics were investigated. Significant decreases of platelets ($p = 0.007$) and albumin ($p = 0.009$) were observed in the Q unchanged group during the observation period, but not in the R unchanged group ($p = 0.104$ and 0.225, respectively). Because platelets and albumin are markers of liver disease progression, it was considered that the Q unchanged group progressed rapidly with frequent HCC occurrence (53%, 16/30) while the R unchanged group showed stable disease with less frequent HCC occurrence (26%, 14/53). In contrast, the R to Q group showed progressive disease ($p = 0.005$ and 0.005, respectively) similar to the Q unchanged group, while the Q to R group showed stable disease similar to the R unchanged group ($p = 0.251$ and 0.647, respectively), demonstrating that amino acid changes of core a.a. 70 were significantly associated with disease progression.

The IL28B SNP and its association with core a.a. 70 and disease progression

Next, the association between the state of liver disease and IL28B SNP was analyzed for a total of 228 patients through the analysis of the rs8099917 locus. Among them, 162 patients (71%) had the major homozygous TT alleles, while 66 patients (29%) had the minor homozygous or heterozygous alleles (GG/TG). Although some patients had a history of IFN therapy, all patients were positive for HCV RNA at the time of study. The clinical characteristics related to disease progression were compared, as shown in Table 4. Each group consisted of patients with similar distributions of age and sex. Though most clinical factors showed no evident differences in these groups, γGTP was high ($p = 0.020$) and HCV RNA concentration was apt to be low ($p = 0.085$) in TG/GG group. Moreover, the ratio

Table 4 Patient characteristics classified by IL28B SNPs at the time of diagnosis

	TT ($N = 162$)	TG/GG ($N = 66$)	p value
Sex (male/female)	81/81	32/34	0.951
Age (years)	63.3 ± 10.7	63.8 ± 11.9	0.735
Platelets (10^{-4}/mm^3)	13.8 ± 5.8	14.3 ± 6.6	0.632
Albumin (g/dL)	4.2 ± 3.0	4.0 ± 0.5	0.528
γGTP (IU/L)	41 ± 39	55 ± 49	0.020
T.Chol (mg/dl)	162 ± 31	157 ± 31	0.237
HCV RNA concentration (kIU/ml)	7,576 ± 10,292	5,069 ± 6,701	0.085
Alpha-fetoprotein (ng/ml)	39.0 ± 152.2	27.3 ± 48.3	0.555
AST (IU/L)	49.6 ± 29.6	51.9 ± 30.5	0.607
ALT (IU/IL)	54.5 ± 53.9	51.6 ± 37.7	0.689
Core a.a. 70R/(Q/H)	106/56	17/49	<0.001
Core a.a. 91L/(M/C)	108/54	38/28	0.253
ISDR	1.2 ± 2.0	0.9 ± 1.5	0.164
IRRDR	5.1 ± 2.4	4.7 ± 2.2	0.207
HCC −/+	115/47	45/21	0.788
IFN −/+	95/67	34/32	0.402

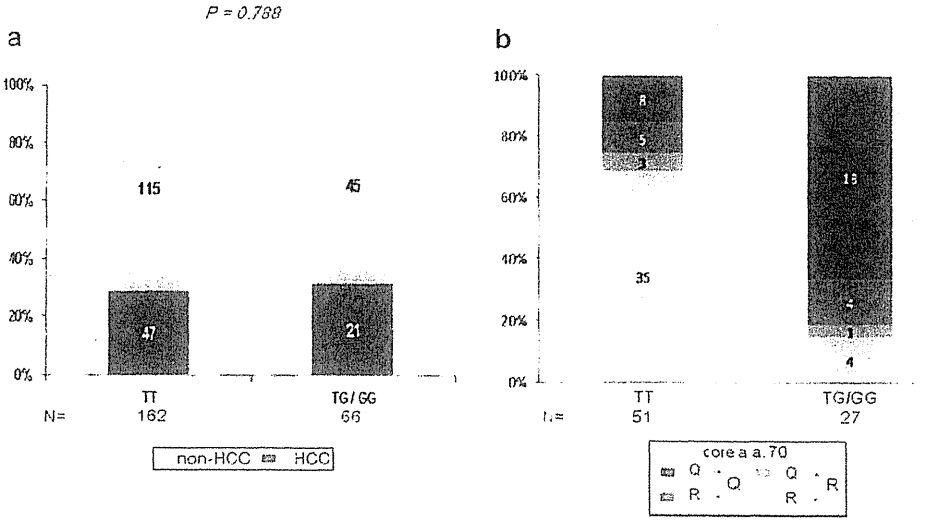


Fig. 4 **a** Association between the state of liver disease and IL28B SNP. **b** Time-dependent core a.a. 70 changes and its relation to IL28B SNP was investigated in 78 patients

of R/(Q/H) at core a.a. 70 was significantly higher in those with the TT alleles than in those with TG/GG ($p < 0.001$). In association of IL28B SNP with HCC development, there was no evident relationship as demonstrated in Fig. 4.

IL28B SNP and time-dependent core a.a. 70 changes

In Fig. 4b, it is demonstrated that the direction of time-dependent core a.a. 70 change was influenced by IL28B SNPs. In IL28B TG/GG patients, 4 (50%) out of 8 patients with the initial core a.a. 70R changed into 70Q, while only 1 (5%) out of 19 patients with the initial core a.a. 70Q changed into 70R, demonstrating that core a.a. 70 tended to change into Q over time in IL28B TG/GG patients ($p = 0.034$). On the other hand, there was no evident changing direction in IL28B TT patients; 5 (13%) out of 40 patients with the initial core a.a. 70R changed into 70Q, while 3 (27%) out of 11 patients with the initial core a.a. 70Q changed into 70R ($p = 0.45$).

Multivariate analysis for independent factors influencing core a.a. 70

To investigate further the relationship between core a.a. 70, the IL28B SNP, and HCC development, we divided the patients according to the specification of core a.a. 70 and

Table 5 Factors related to polymorphism of core a.a. 70

Variables	Univariate analysis ($N = 228$)		Multivariate analysis ($N = 228$)	
	Odds ratio (95% CI)	p value	Odds ratio (95% CI)	p value
Sex				
Female	1	0.415	1	0.812
Male	1.23 (0.74–2.01)		1.08 (0.58–1.99)	
Age (years)				
<65	1	0.216	1	0.855
≥65	1.39 (0.82–2.35)		1.06 (0.57–1.96)	
Platelets (10^{-4}/mm^3)				
>12	1	0.004	1	0.844
<12	1.76 (1.03–2.99)		1.07 (0.53–2.16)	
Albumin (g/dL)				
>4	1	0.002	1	0.300
≤4	2.28 (1.33–3.91)		1.46 (0.71–3.00)	
γGTP (IU/L)				
<41	1	0.003	1	0.299
≥41	2.32 (1.32–4.09)		1.42 (0.73–2.79)	
ALT (IU/L)				
<41	1	0.040	1	0.573
≥41	1.74 (1.03–2.94)		1.22 (0.62–2.39)	
IL28B				
TT	1	<0.001	1	<0.001
TG or GG	5.46 (2.88–10.30)		5.74 (2.91–11.31)	
HCC				
−	1	<0.001	1	0.046
+	2.98 (1.65–5.37)		2.21 (1.01–4.83)	
Previous IFN therapy				
−	1	0.874	1	0.644
+	0.96 (0.57–1.62)		0.87 (0.47–1.59)	

those factors, as well as clinical factors, were compared in univariate and multivariate analyses. In Table 5, it may be seen that platelets, albumin, γGTP, ALT, the IL28B SNP, and number of patients with HCC development differed significantly between the two groups in univariate analysis. In contrast, successive multivariate analysis demonstrated that the number of patients with HCC development ($p = 0.046$) and the IL28B SNP ($p < 0.001$) were extracted as independent variables correlated with the core a.a. 70 residue (Table 5).

Multivariate analysis for independent factors influencing HCC development

To disclose factors influencing HCC development, multivariate analysis was performed. As shown in Table 6, age, albumin, and core a.a. 70 residue were extracted as independent factors. On the other hand, IL28B SNP was not extracted as one of those factors.

Discussion

In this study, we have documented several important findings. Through the investigation of HCV sequences, including complete HCV ORFs analysis, we have shown that the core a.a. 70 residue and its changes over time are associated with the disease progression as well as HCC development in genotype-1b HCV infection. Specifically, core a.a. 70Q/H was associated with HCC development and disease progression; core a.a. 70 often changed with time and R70Q substitutions were associated with progressive disease, while Q70R substitutions were associated with the stable disease. Moreover, we have shown that the IL28B SNP and core a.a. 70 showed significant linkage. In contrast, we have also shown that HCC development and disease progression were not apparently correlated with the IL28B SNP.

Recently, core amino acids have been reported in several studies to be associated with HCC [12, 21–25]. In

Table 6 Factors related to influencing HCC development

Variables	Univariate analysis ($N = 228$)		Multivariate analysis ($N = 228$)	
	Odds ratio (95% CI)	p value	Odds ratio (95% CI)	p value
Sex				
Female	1	0.161	1	0.190
Male	1.50 (0.85–2.67)		1.69 (0.77–3.71)	
Age (years)				
<65	1	0.006	1	0.004
≥65	2.30 (1.28–4.16)		3.26 (1.46–7.25)	
Platelets (10^{-4}/mm^3)				
>12	1	<0.001	1	0.021
≤12	5.82 (3.11–10.88)		2.59 (1.16–5.82)	
Albumin (g/dL)				
>4	1	<0.001	1	<0.001
≤4	13.75 (6.69–28.24)		7.73 (3.53–16.94)	
γGTP (IU/L)				
<41	1	<0.001	1	0.122
≥41	3.09 (1.70–5.62)		1.87 (0.85–4.13)	
ALT (IU/L)				
<41	1	<0.001	1	0.109
≥41	3.88 (2.06–7.31)		1.98 (0.86–4.56)	
IL28B				
TT	1	0.626	1	0.290
TG or GG	1.17 (0.13–2.17)		0.63 (0.27–1.49)	
Core a.a. 70				
R	1	<0.001	1	0.029
Q/H	2.91 (1.61–5.26)		2.44 (1.09–5.44)	
Previous IFN therapy				
−	1	0.949	1	0.331
+	0.98 (0.55–1.74)		1.46 (0.68–3.16)	

these studies, patients with core a.a. 70Q/H frequently developed HCC with exacerbation of liver damage. In this analysis, we confirmed the previous findings. However, because this association might be a reflection of the core-dependent IFN sensitivity differences often reported in recent studies [12, 22, 25], we restricted the analysis to patients, who were unable to clear HCV RNA previously through IFN-based therapy. Moreover, we also confirmed the relationship of the core sequences and disease development among the populations without a previous history of IFN therapy (data not shown). These findings strongly confirmed the role of core a.a. 70 in disease progression, independent of any IFN response.

It is a focus of interest how the core sequence evolves with time or with the course of disease. If the core sequences were fixed throughout the course of disease, HCV with core 70Q might be an "oncogenic" virus, while HCV with core 70R might be "non-oncogenic", and the initial viral sequence might forecast future liver disease. In this study, we have demonstrated that core sequences changed in 15% (15/98) of patients during the observation period of around 10 years. Among these changes, R70Q ($N = 11$) was more common than Q70R ($N = 4$). Interestingly, changes in this region were significantly associated with disease activity or HCC development, although patients with R70Q substitutions were significantly more likely to have exacerbation of the disease and Q70R substitutions were associated with the stable disease. These results demonstrate that the core a.a. 70 residue is not fixed, but often changes with time during the course of disease in close association with disease progression and HCC development. Although the molecular mechanism of their interaction needs further exploration, this result highlights the important clinical and basic implications for the association between host and virus.

The importance of the IL28B SNP has been demonstrated recently in HCV infection in terms of a correlation with treatment outcome of pegylated-IFN plus ribavirin

therapy [16–19]. The contribution of the IL28B SNP to the outcome of therapy was confirmed in successive studies, although the mechanism remains under investigation. On this basis, we sought to investigate the impact of the IL28B SNP on disease progression and HCC development, separate from the IFN-based treatment response. As shown in Table 4, we compared the clinical features between the two groups (IL28B GG/TG vs. IL28B TT). Importantly, this comparison disclosed a significant correlation between the core a.a. 70 polymorphisms and the IL28B SNP ($p < 0.001$) and confirmed the existence of a complex interaction between the host and the virus in chronic HCV infection. According to the result, patients with IL28B TG/GG were more likely to be infected with HCV with core a.a. 70Q/H than with core a.a. 70R and vice versa. Although the molecular mechanisms of their relationship remain unknown, it could be speculated that the IL28B SNP has an influence on the viral core sequences, because the host IL28B SNP remains fixed and cannot be influenced by the viral core sequence.

On the other hand, we observed no evident association between the IL28B SNP and HCC development. This was rather unexpected because it is considered that the IL28B SNP has a significant influence on the core a.a. 70 residue. Therefore, to clarify the correlation among core a.a. 70, IL28B, and HCC development, we undertook multivariate analysis to extract the independent variables affecting the core 70 residue. As demonstrated in Table 5, the IL28B SNP and the development of HCC were extracted as variables independently correlated with the core a.a. 70 residue. The result indicates that the core a.a. 70 residue was not only influenced by the IL28B SNP, but also by factors strongly related to HCC development, independent of the IL28B SNP. When considering the result, it is not strange if there is no direct relationship between IL28B SNP and HCC development. In contrast, multivariate analysis undertaken for disclosing factors influencing HCC development revealed that core a.a. 70 residue was a variable independently associated with HCC development other than age, albumin, or platelets even though the IL28B SNP was not extracted (Table 6). However, further comprehensive studies are warranted to disclose the molecular mechanisms for the complicated relationships among core a.a. 70, IL28B, and HCC development.

In conclusion, we have shown that core a.a. 70 was closely associated with disease progression and, often, changes of that residue were accompanied by temporal changes in liver damage, in close relationship with the IL28B SNP.

Acknowledgements This study was supported in part by a Grant-in-Aid scientific research fund of the Ministry of Education, Science, Sports and Culture number 20390206 and in part by a Grant-in-Aid from the Ministry of Health, Labour, and Welfare of Japan (H19-kanen-002).

References

1. Kiyosawa K, Sodeyama T, Tanaka E, et al. Interrelationship of blood transfusion, non-A, non-B hepatitis and hepatocellular carcinoma: analysis by detection of antibody to hepatitis C virus. Hepatology 1990;12:671–675
2. Enomoto N, Sakuma I, Asahina Y, et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. N Engl J Med 1996;334:77–81
3. Akuta N, Suzuki F, Sezaki H, et al. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. Intervirology 2005;48:372–380
4. Hung CH, Chen CH, Lee CM, et al. Association of amino acid variations in the NS5A and E2-PePHD region of hepatitis C virus 1b with hepatocellular carcinoma. J Viral Hepat 2008;15:58–65
5. El-Shamy A, Nagano-Fujii M, Sasase N, Imoto S, Kim SR, Hotta H. Sequence variation in hepatitis C virus nonstructural protein 5A predicts clinical outcome of pegylated interferon/ribavirin combination therapy. Hepatology 2008;48:38–47
6. Moriya K, Nakagawa K, Santa T, et al. Oxidative stress in the absence of inflammation in a mouse model for hepatitis C virus-associated hepatocarcinogenesis. Cancer Res 2001;61:4365–4370
7. Okuda M, Li K, Beard MR, et al. Mitochondrial injury, oxidative stress, and antioxidant gene expression are induced by hepatitis C virus core protein. Gastroenterology 2002;122:366–375
8. Nishina S, Hino K, Korenaga M, et al. Hepatitis C virus-induced reactive oxygen species raise hepatic iron level in mice by reducing hepcidin transcription. Gastroenterology 2008;134(1):226–238
9. Bartosch B, Thimme R, Blum HE, Zoulim F. Hepatitis C virus-induced hepatocarcinogenesis. J Hepatol 2009;51:810–820
10. Tsai WL, Chung RT. Viral hepatocarcinogenesis. Oncogene 2010;29:2309–2324
11. Moriya K, Fujie H, Shintani Y, et al. The core protein of hepatitis C virus induces hepatocellular carcinoma in transgenic mice. Nat Med 1998;4:1065–1067
12. Akuta N, Suzuki F, Kawamura Y, et al. Amino acid substitutions in the hepatitis C virus core region are the important predictor of hepatocarcinogenesis. Hepatology 2007;46:1357–1364
13. Franco S, Gimenez-Barcons M, Puig-Basagoiti F, et al. Characterization and evolution of NS5A quasispecies of hepatitis C virus genotype 1b in patients with different stages of liver disease. J Med Virol 2003;71:195–204
14. Gimenez-Barcons M, Franco S, Suarez Y, et al. High amino acid variability within the NS5A of hepatitis C virus (HCV) is associated with hepatocellular carcinoma in patients with HCV-1b-related cirrhosis. Hepatology 2001;34:158–167
15. Nagayama K, Kurosaki M, Enomoto N, Miyasaka Y, Marumo F, Sato C. Characteristics of hepatitis C viral genome associated with disease progression. Hepatology 2000;31:745–750
16. Fukuhara T, Taketomi A, Motomura T, et al. Variants in IL28B in liver recipients and donors correlate with response to peginterferon and ribavirin therapy for recurrent hepatitis C. Gastroenterology 2010;139:1577–1585
17. Tanaka Y, Nishida N, Sugiyama M, et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 2009;41:1105–1109

18. Suppiah V, Moldovan M, Ahlenstiel G, et al. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. Nat Genet 2009;41:1100–1104
19. Ge D, Fellay J, Thompson AJ, Simon JS, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 2009;461:399–401
20. Thomas DL, Thio CL, Martin MP, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. Nature 2009;461:798–801
21. Ogura S, Akuta N, Hirakawa M, et al. Virological and biochemical features in elderly HCV patients with hepatocellular carcinoma: amino acid substitutions in HCV core region as predictor of mortality after first treatment. Intervirology 2009;52:179–188
22. Nakamoto S, Imazeki F, Fukai K, et al. Association between mutations in the core region of hepatitis C virus genotype 1 and hepatocellular carcinoma development. J Hepatol 2010;52:72–78
23. Kobayashi M, Akuta N, Suzuki F, et al. Influence of amino-acid polymorphism in the core protein on progression of liver disease in patients infected with hepatitis C virus genotype 1b. J Med Virol 2010;82:41–48
24. Hu Z, Muroyama R, Kowatari N, Chang J, Omata M, Kato N. Characteristic mutations in hepatitis C virus core gene related to the occurrence of hepatocellular carcinoma. Cancer Sci 2009;100:2465–2468
25. Akuta N, Suzuki F, Kawamura Y, et al. Substitution of amino acid 70 in the hepatitis C virus core region of genotype 1b is an important predictor of elevated alpha-fetoprotein in patients without hepatocellular carcinoma. J Med Virol 2008;80:1354–1362

Hepatol Int (2011) 5:789–799
DOI 10.1007/s12072-011-9267-x

ORIGINAL ARTICLE

Analysis of the complete open reading frame of hepatitis C virus in genotype 2a infection reveals critical sites influencing the response to peginterferon and ribavirin therapy

Makoto Kadokura · Shinya Maekawa · Ryota Sueki · Mika Miura · Kazuki Komase · Hiroko Shindo · Fumitake Amemiya · Tomoyoshi Uetake · Taisuke Inoue · Minoru Sakamoto · Mina Nakagawa · Naoya Sakamoto · Mamoru Watanabe · Nobuyuki Enomoto

Received: 25 November 2010 / Accepted: 21 February 2011 / Published online: 20 March 2011

Abstract
Purpose A proportion of patients infected with genotype 2a hepatitis C virus (HCV) cannot achieve a sustained virological response (SVR) to pegylated-interferon plus ribavirin therapy (PEG-IFN/RBV) but the reason remains unclear. The present study aimed to clarify the possible correlation between viral sequence variations and final outcome.
Methods The pretreatment complete open reading frame (ORF) sequences of genotype 2a HCV were determined by direct sequencing for two independent groups of patients (43 patients as test; group 1 and 35 as validation; group 2), and the correlation with the final outcome was explored.
Results Patients with SVR ($n = 58$) and with non-SVR ($n = 20$) differed significantly in pretreatment HCV RNA level ($p = 0.002$), fibrosis score ($p = 0.047$), and cumulative RBV dosage ($p = 0.003$). By comparison of all amino acid positions in the complete HCV ORFs, threonine at amino acid (aa) 110 in the core region was remarkably frequent in SVR ($p = 0.01$ for group 1, $p = 0.004$ for group 2, and $p = 5E{-}05$ for combined). A sliding window analysis revealed that the total number of amino acid variations within the NS5A aa 2258–2306 region were significantly high in SVR compared to non-SVR patients ($p = 0.01$ for group 1, $p = 0.006$ for group 2, and $p = 0.0006$ for combined). Multivariate analyses revealed that core aa 110 ($p = 0.02$), NS5A aa 2258–2306 ($p = 0.03$), and cumulative RBV dosage ($p = 0.02$) were identified as independent variables associated with the final outcome.
Conclusions The outcome of PEG-IFN/RBV therapy is significantly influenced by variation in the core and NS5A regions in genotype 2a HCV infection.

Abbreviations

EVR	Early virological response
IFN	Interferon
IRRDR	Interferon ribavirin resistance determinant region
ISDR	Interferon sensitivity determinant region
ORF	Open reading frame
PEG-IFN	Pegylated-interferon
PePHD	PKR-eIF2 phosphorylation homology domain
PKR-BD	Double-stranded RNA-activated protein Kinase binding domain
RBV	Ribavirin
RVR	Rapid virological response
SVR	Sustained virological response

Introduction

Worldwide, 180 million of people are estimated to be infected with hepatitis C virus (HCV), and HCV is a major cause of chronic hepatitis, liver cirrhosis, and

M. Kadokura · S. Maekawa (✉) · R. Sueki · M. Miura · K. Komase · H. Shindo · F. Amemiya · T. Uetake · T. Inoue · M. Sakamoto · N. Enomoto
First Department of Internal Medicine, Faculty of Medicine, University of Yamanashi, 1110, Shimokato, Chuo, Yamanashi 409-3898, Japan
e-mail: maekawa@yamanashi.ac.jp

M. Nakagawa · N. Sakamoto · M. Watanabe
Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, 1-5-45, Yushima, Bunkyo, Tokyo 113-8510, Japan

Table 1 Baseline characteristics of all patients (groups 1 and 2)

Characteristic	SVR ($n = 58$)			Non-SVR ($n = 20$)			p value$^{\Delta}$
	Group 1 ($n = 36$)	Group 2 ($n = 22$)	Combined ($n = 58$)	Group 1 ($n = 7$)	Group 2 ($n = 13$)	Combined ($n = 20$)	
Gender (male/female)	20/16	9/13	29/29	4/3	5/8	9/11	0.80†
Age (years)	50.0 ± 12.5*	57.3 ± 10.0	52.4 ± 12.1	55.0 ± 9.7	59.8 ± 6.4	58.1 ± 7.8	0.058‡
ALT (IU/l)	86.6 ± 86.6	71.2 ± 50.4	80.5 ± 74.2	52.9 ± 29.3	88.1 ± 90.1	75.8 ± 75.5	0.81‡
Platelet ($\times 10^4$/mm^3)	20.8 ± 6.2	19.0 ± 5.2	20.1 ± 5.8	14.7 ± 7.1	19.1 ± 4.9	17.6 ± 6.0	0.11‡
Fibrosis score (0–2/≥3)§	34/1	19/2	53/3	4/3	11/2	15/5	0.049†
HCV RNA (KIU/ml)	760 (2–3,100)**	340 (54–3,600)	550 (12–3,600)	1,300 (350–30,000)	1,400 (180–5,000)	1,300 (180–30,000)	0.002$^{\parallel}$
IFN dose (≥80%/60–80%)¶	28/4	21/1	49/5	4/3	11/2	15/5	0.12†
Ribavirin dose (≥80%/60–80%)¶	27/5	17/5	44/10	4/3	5/8	9/11	0.003†
RVR rate (%)	87.5	54.5	74.1	33.3	46.1	42.1	0.022†
EVR rate (%)	100	100	100	66.7	100	89.4	0.07†

* Mean ± SD; ** median (range); † Fisher's exact probability test; ‡ Student t test; $^{\parallel}$ Mann–Whitney's U test; $^{\Delta}$ p values between all SVR ($n = 58$) versus all non-SVR ($n = 20$)

Several clinical characteristics listed above were unavailable in some patients. § SVR: $n = 56$ (35 in group 1, 21 in group2), non-SVR: $n = 17$ (7 in group 1, 10 in group 2); ¶ SVR: $n = 54$ (32 in group 1, 22 in group 2)

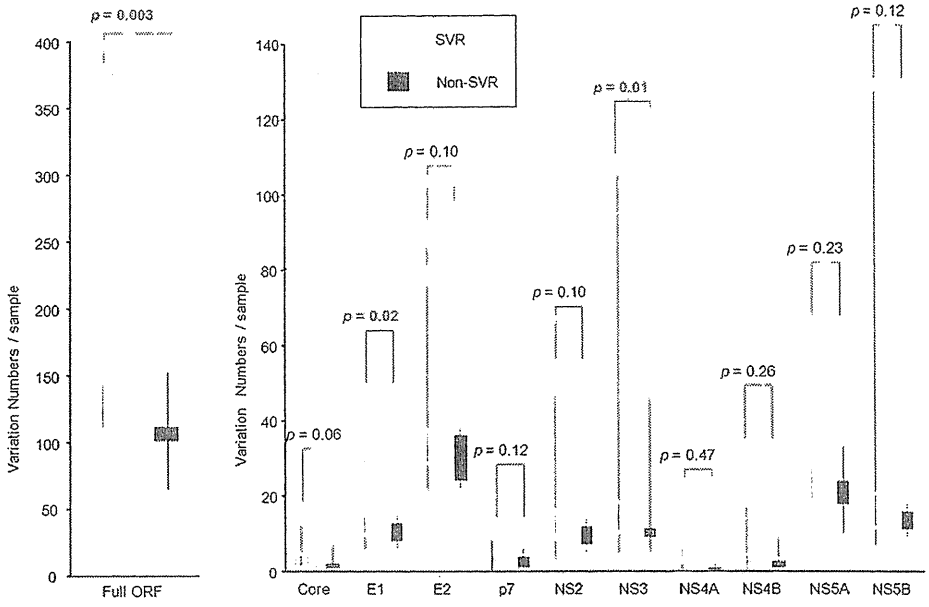


Fig. 1 Number of amino acid substitutions per sample in the sustained viral responders (SVR) and the non-sustained viral responders (non-SVR) group. The numbers of variations, relative to a population consensus, that were unique to either SVR or non-SVR patients are shown for the full ORF (*left*) and for each HCV protein (*right*)

Comparison of HCV sequence variation between the SVR and non-SVR patients at each amino acid position

Next, each amino acid position in the HCV ORF was compared to detect any differences between the SVR and non-SVR patients after determination of the consensus sequence from all 78 patients. In Fig. 2a, the final differences of the two independent studies combined are shown as dots demonstrating –log P values. As shown in the figure, amino acid usage at amino acid 110 in the core region differed strikingly between the two groups ($p = 5E{-}05$). The site was detected in group 1 ($p = 0.01$) and was validated in group 2 ($p = 0.004$) (Table 2), and the final p value became remarkably high, making the p value at this site most significantly low. Variations of aa 773 in p7, aa 2099 in the NS5A, and aa 3013 in NS5B were also shown to differ significantly between the SVR and the non-SVR patients when the two studies were combined; however, they were not confirmed by one of the studies (Table 2). Figure 2b shows the aligned sequences of amino acids 1–120 of the core region. Substitutions at aa 110 from

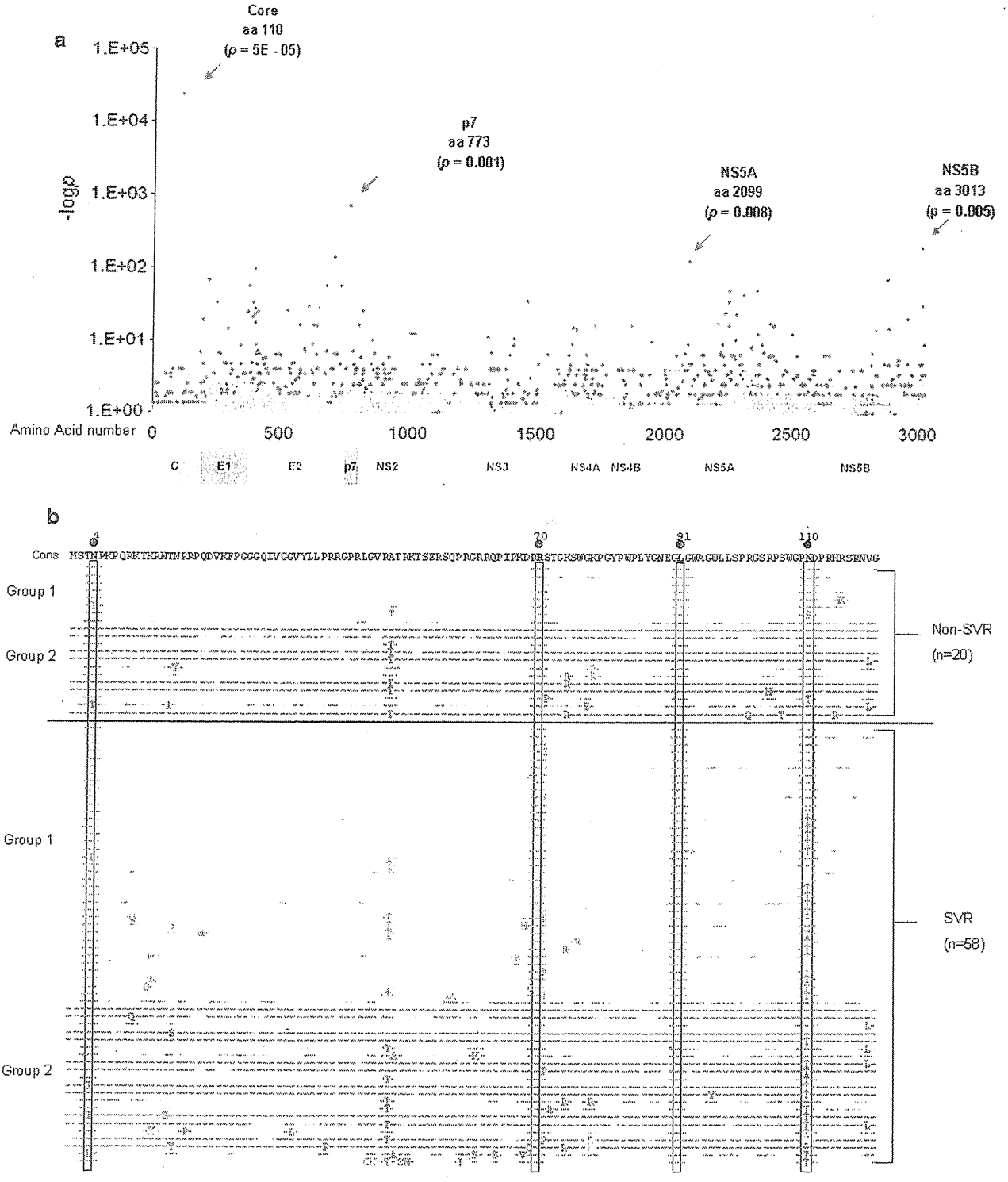


Fig. 2 **a** Different amino acid usages at each viral amino acid position between the sustained viral responders (SVR) and the non-sustained viral responders (non-SVR) patients. Amino acid variation was determined between SVR and non-SVR patients by Fisher's exact probability test. The longitudinal axis shows the –log *P* value. **b** Sequence alignment in the core region. *Dashes* indicate amino acids identical to the consensus sequence and substituted amino acids are shown by standard single letter codes. **c** Sliding window analysis. Viral regions affecting treatment outcomes are shown in *dark spots*. There are four hot spots: at core amino acid 110, amino acids 400–403 (i.e., the hypervariable region) in Envelope 2 (*E2*) region, amino acids 724–743 in E2, and amino acids 2258–2306 in the nonstructural (*NS*) 5A. **d** Sequence alignment amino acids in the nonstructural (*NS*) 5A around amino acids 2258–2306. *Dashes* indicate amino acids identical to the consensus sequence and substituted amino acids are shown by standard *single letter* codes

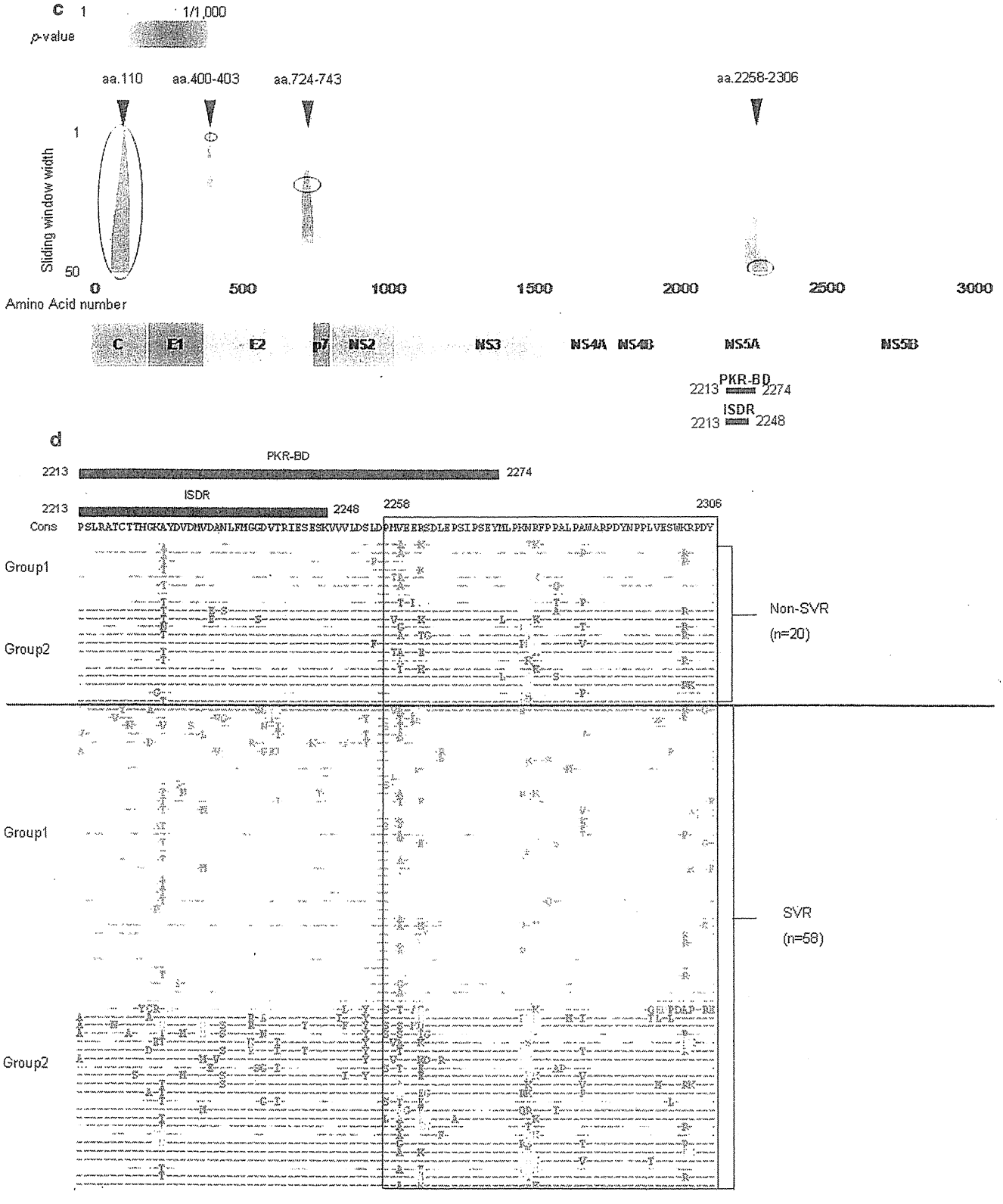


Fig. 2 continued

non-T (N/S) to T were significantly more frequent in SVR (32/58, 55.2%) than in non-SVR (1/20, 3.6%, $p = 5E{-}05$). Amino acid 4, the site reported recently to vary according to the viral response in genotype 2a infection, did not differ significantly in our study. Amino acid 70 and 91, which have been reported to vary according to viral response to

Table 2 Variations in each amino acid position and SVR rate

Position	Group 1 ($n = 43$)	p value	Group 2 ($n = 35$)	p value	Combined ($n = 78$)	p value
Core aa 110						
T	100% (19/19)	0.01	92.9% (13/14)	0.004	97% (32/33)	5E−05
Non T	70.8% (17/24)		42.9% (9/21)		57.8% (26/45)	
p7 aa 773						
V	77.4% (24/31)	0.16	53.6% (15/28)	0.03	66.1% (39/59)	0.002
Non V	100% (12/12)		100% (7/7)		100% (19/19)	
NS5A aa 2099						
R	92.9% (13/14)	0.40	91.7% (11/12)	0.01	92.3% (24/26)	0.01
Non R	79.3% (23/29)		47.8% (11/23)		65.4% (34/52)	
NS5B aa 3013						
L	78.9% (26/33)	0.17	47.8% (11/23)	0.01	66.1% (37/56)	0.008
Non L	100% (10/10)		91.7% (11/12)		95.5% (21/22)	

Table 3 Number of amino acid substitutions in each region and SVR rate

Region	Group 1 ($n = 43$)	p value	Group 2 ($n = 35$)	p value	Combined ($n = 78$)	p value
E2 aa 400–403						
Mutation ≥ 2	89.3% (25/28)	0.22	100% (11/11)	0.002	92.3% (36/39)	0.0005
Mutation 0–1	73.3% (11/15)		45.8% (11/24)		56.4% (22/39)	
E2 aa 724–743						
Mutation ≥ 1	100% (28/28)	0.0002	72% (18/25)	0.12	86.8% (46/53)	0.0006
No mutation	53.3% (8/15)		40% (4/10)		48% (12/25)	
ISDR(aa 2213–2248)						
Mutation ≥ 2	100% (15/15)	0.08	86.7% (13/15)	0.02	93.3% (28/30)	0.003
Mutation 0–1	75% (21/28)		45% (9/20)		62.5% (30/48)	
NS5A aa 2258–2306						
Mutation ≥ 5	100% (19/19)	0.01	84.2% (16/19)	0.006	92.1% (35/38)	0.0006
Mutation 0–4	70.8% (17/24)		37.5% (6/16)		57.5% (23/40)	

PEG-IFN/RBV therapy in genotype 1b infection, were conserved irrespective of the outcome.

Comparison of amino acid variation between the SVR and non-SVR patients across HCV "regions" using sliding window analysis

Figure 2c shows the combined result of sliding window analysis for study groups 1 and 2. This approach was used to detect differing HCV amino acid "regions", rather than single amino acid positions, between the SVR and the non-SVR patients. According to the result, four regions were notably associated with the final outcome (p values less than 1/1,000). Core aa 110, detected as a single amino acid position discriminating between the SVR and the non-SVR patients, was also identified as one of these regions. Because core aa 110 was already known for its strong correlation with the response as above, the region was excluded from further analysis. Among the other three regions, only NS5A aa 2258–2306 showed significant differences in the two independent study groups (Table 3). Interestingly, the NS5A region overlapped the PKR-binding domain, which includes the IFN sensitivity determining region (ISDR). Figure 2d shows the aligned sequences of amino acids around 2258–2306 of HCV NS5A. As with previous studies, variations in the ISDR were also significantly more frequent in SVR patients.

Multivariate analysis to detect independent factors contributing to the SVR

Multivariate analysis revealed that variation of core aa 110, the total number of substitutions within NS5A aa 2258–2306, and total RBV dose ≥80% were finally

Table 4 Multivariate logistic regression analysis

Factor	Odds (95% CI)	*p* value
Age	1.01 (0.91–1.13)	0.85
HCV RNA	1.00 (1.00–1.00)	0.09
Fibrosis score ≥3/0–2	2.37 (0.21–26.7)	0.48
RVR achievement	3.46 (0.54–22.1)	0.19
Ribavirin dose ≥80%	16.0 (1.66–153)	0.02
Core aa 110 T	24.7 (1.72–353)	0.02
NS5A aa 2258–2306 mutations 0–4/≥5	11.5 (1.23–108)	0.03

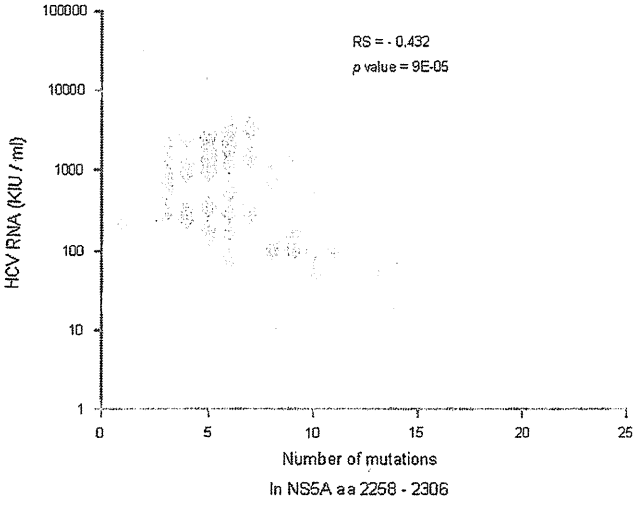


Fig. 3 Correlation between pretreatment HCV RNA levels and the number of substitutions in the NS5A region aa 2258–2306. Spearman's correlation coefficient by rank test is demonstrated

identified as the independent variables influencing the final outcome (odds ratio 24.7, 11.5, and 16.0; $p = 0.02$, 0.03 and 0.02; Table 4).

Biological relevance of variation in core and NS5A in this study group

To determine biological relevance of core aa110 and NS5A aa2258–2306, we investigated their relationship with clinical background factors. Multiple variations in the NS5A region aa 2258–2306 were significantly related to pretreatment HCV RNA titer ($p = 9E-05$, Fig. 3; Table 5). Interestingly, variation of the core aa110 was significantly associated with the patients' age ($p = 0.03$, Table 6).

Discussion

In this study, based on analysis of complete HCV-ORF sequences and comparison of SVR and non-SVR patients in two independent study groups, we have shown that amino acid variations in the core and NS5A correlate most significantly with the final outcome in the treatment for genotype 2a chronic hepatitis C. The study is unique in that the patients studied were all Japanese, excluding any effect of racial differences and providing a clearer analysis of the viral differences.

From the analysis of the characteristics of patients infected with genotype 2a HCV, it was clear that most non-SVR patients responded to the PEG-IFN/RBV therapy at least transiently, given that most of these non-SVR patients (89%) achieved EVR. This result demonstrated that most non-SVR patients were relapsers, but were not null-responders as observed frequently among genotype 1b patients treated with PEG-IFN/RBV therapy. Therefore, we compared the different viral responses according to the final outcome of SVR or non-SVR.

Variation of core aa 110 was identified as the single amino acid residue most significantly related to the final outcome ($p = 5E-05$). In recent studies of treatment of genotype 1b infection with PEG-IFN/RBV, amino acid variation in the core region was reported to be associated with response. It is interesting that the core region was also identified as a HCV gene associated with the response to PEG-IFN/RBV therapy of genotype 2a infection, although the amino acid residues of core in genotype 1b were different, being aa 70 and aa 91. It is also interesting that amino acids aa 70 and aa 91 are conserved as arginine and leucine, respectively, in genotype 2a, as reported to be associated with favorable PEG-IFN/RBV responses in genotype 1b infection, consistent with the association with a high SVR rate in genotype 2a infection. Very recently, a correlation was reported between amino acid variations in the core region and viral responses of genotype 2a HCV infection [20]. Though the result seems discrepant from our study, we suspect the inconsistent results were at least partially attributable to the different groups used in comparison: we compared the difference between non-SVR patients and SVR patients while they compared the difference between non-SVR and RVR patients.

In systemic searching for the viral "regions" associated with the treatment outcome, NS5A aa2258–2306 was identified by two independent studies. Interestingly, the region overlaps the PKR-binding domain (PKR-BD), including the ISDR, in which the number of amino acid substitutions is known to be related to the response to IFN-based therapy in genotype 1b, and also in genotype 2a [17, 18]. Therefore, we also confirmed that total number of substitutions in the ISDR and PKR-BD is significantly associated with the final outcome in this group of patients when the two studies were combined.

Some viral regions other than core and NS5A also showed the potential association with the final outcome. Viral single amino acid substitutions of aa 773 in p7, aa

Table 5 Baseline characteristics of patients with NS5A aa 2258–2306 mutations 0–4 or ≥5 (groups 1 and 2)

Characteristic	Mutation 0–4 ($n = 40$)	Mutation ≥ 5 ($n = 38$)	p value
Gender (male/female)	22/18	16/22	NS†
Age (years)	54.3 ± 11.4*	53.5 ± 11.5	NS‡
ALT (IU/l)	73.8 ± 70.3	85.3 ± 78.7	NS‡
Platelet ($\times 10^4/mm^3$)	18.0 ± 5.9	21.0 ± 5.7	0.03‡
Fibrosis score (0–2/≥3)§	33/5	33/2	NS†
HCV RNA (KIU/ml)	1,100 (99–30,000)**	380 (12–5,000)	0.02‖
IFN dose (≥80%/60–80%)¶	31/8	33/2	NS†
Ribavirin dose (≥80%/60–80%)¶	25/14	28/7	NS†
RVR rate (%)	65.8	62.9	NS†
EVR rate (%)	94.7	100	NS†
Relapse rate (%)	35.9	7.9	0.002†
SVR rate (%)	57.5	92.1	0.0006†

* Mean ± SD; † Fisher's exact probability test; ‡ Student t test; § mutation 0–4 $n = 38$, mutation ≥5: $n = 35$; ** median (range); ‖ Mann–Whitney's U test; ¶ mutation 0–4: $n = 39$, mutation ≥5: $n = 35$

Table 6 Baseline characteristics of patients with core 110 T or N/S (groups 1 and 2)

Characteristic	Core 110 T ($n = 33$)	Core 110 N/S ($n = 45$)	p value
Gender (male/female)	18/15	20/25	NS†
Age (years)	50.4 ± 13.0*	56.4 ± 9.5	0.032‡
ALT (IU/l)	64.5 ± 48.2	88.8 ± 86.2	NS‡
Platelet ($\times 10^4/mm^3$)	19.3 ± 4.9	19.5 ± 6.6	NS‡
Fibrosis score (0–2/≥3)§	30/1	36/6	NS†
HCV RNA (KIU/ml)	580 (54–3,600)**	980 (12–30,000)	NS‖
IFN dose (≥80%/60–80%)¶	26/3	38/7	NS†
Ribavirin dose (≥80%/60–80%)¶	23/6	30/15	NS†
RVR rate (%)	72.4	59.1	NS†
EVR rate (%)	100	95.5	NS†
Relapse rate (%)	3.0	38.6	9E−05†
SVR rate (%)	97.0	57.8	5E−05†

* Mean ± SD; † Fisher's exact probability test; ‡ Student t test; § core 110 T: $n = 31$, core 110 N/S: $n = 42$; ** median (range), ‖ Mann–Whitney's U test; ¶ core 110 T: $n = 29$

2099 in the NS5A, and aa 3013 in NS5B, or viral regions in E1 aa 400–403 and in E2 aa 724–744 were more frequent in SVR. However, because these were not extracted as significant in one of the two studies when analyzed separately, additional studies are needed to confirm the association with the final outcome. On the other hand, we could not find an association with the final outcome and the PePHD or IRRDR, including the V3 regions (data not shown) reported 1b HCV infection [21, 22].

It is interesting that the variation of the core region showed clear association with age. Younger patients with core aa 110T showed favorable responses, while older patients with core aa 110 non-T showed unfavorable responses. It is possible that different response rates according to the patients' ages in genotype 2a infection might have been related to the core substitutions, although further study is needed. In NS5A, it was reported that the variations within the PKR-binding region, including those within the ISDR, can disrupt the NS5A-PKR interaction, possibly rendering HCV sensitive to the antiviral effects of IFN [23]. Clinically, the number of substitutions within the region has been reported to correlate with the serum HCV RNA level [12]. We also confirmed that the number of substitutions within the NS5A aa 2258–2306 was significantly associated with the pretreatment HCV RNA titers.

Multivariate analysis of the combined group of patients showed that variation of core aa 110, NS5A aa 2258–2306, and total RBV dose ≥80% were independent variables associated with the final outcome (Table 4). The association of RBV dose and HCV relapse rate was reported previously [24] and that result was confirmed in this study. On the other hand, the total PEG-IFN dosage was not identified when it was administered at greater than 60% of the initially scheduled amount. Indeed, when the drug dosage was excluded, the strongest association was seen in the viral elements of core and NS5A, revealing the

importance of these two regions in the treatment of genotype 2a HCV infection with PEG-IFN/RBV therapy.

On the other hand, our study still has some limitations. In recent studies, IL28B single nucleotide polymorphisms were reported to be correlated significantly with the treatment response in genotype 1b HCV infections [25, 26]. In genotype 2a HCV infection, a correlation was also reported to exist between the IL28B SNP and the treatment response [27]. However, we could not investigate the association of the IL28B single nucleotide polymorphisms in the treatment response in genotype 2a HCV infections. In addition, the number of analyzed patients was rather small, especially in non-SVR patients.

In conclusion, by comprehensive investigation of the complete HCV ORF in patients showing different responses to PEG-IFN/RBV therapy, we have demonstrated that amino acid variation in the core and NS5A are significantly associated with the final outcome of treatment of genotype 2a chronic hepatitis C. Considering this result, determination of those HCV regions before treatment might provide further benefits for the patients infected with genotype 2a HCV.

Acknowledgements This study was supported in part by a grant-in-aid scientific research fund of the Ministry of Education, Science, Sports and Culture number 20390206 and in part by a grant-in-aid from the Ministry of Health, Labour, and Welfare of Japan (H19-kanen-002). We are grateful to Yamanashi Pegintron Ribavirin Study Group and Ochanomizu Liver Conference Group for their cooperation and advices. Especially, we thank Asuka Kanayama and Takako Ohmori for their technical assistance, and Takatoshi Kitamura and Shunichi Okada for their cooperation and advices.

References

1. Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. Hepatology 2009;49(4):1335–1374
2. Di Bisceglie AM, Martin P, Kassianides C, Lisker-Melman M, Murray L, Waggoner J, Goodman Z, Banks SM, Hoofnagle JH. Recombinant interferon alfa therapy for chronic hepatitis C A randomized, double-blind, placebo-controlled trial. N Engl J Med 1989;321(22):1506–1510
3. Haydon GH, Jarvis LM, Blair CS, Simmonds P, Harrison DJ, Simpson KJ, Hayes PC. Clinical significance of intrahepatic hepatitis C virus levels in patients with chronic HCV infection. Gut 1998;42(4):570–575
4. Simmonds P. Clinical relevance of hepatitis C virus genotypes. Gut 1997;40(3):291–293
5. Hadziyannis SJ, Sette H Jr, Morgan TR, Balan V, Diago M, Marcellin P, Ramadori G, Bodenheimer H Jr, Bernstein D, Rizzetto M, Zeuzem S, Pockros PJ, Lin A, Ackrill AM. Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. Ann Intern Med 2004;140(5):346–355
6. Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, Goodman ZD, Koury K, Ling M, Albrecht JK. 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
7. Dalgard O, Bjoro K, Hellum KB, Myrvang B, Ritland S, Skaug K, Raknerud N, Bell H. Treatment with pegylated interferon and ribavarin in HCV infection with genotype 2 or 3 for 14 weeks: a pilot study. Hepatology 2004;40(6):1260–1265
8. Mangia A, Santoro R, Minerva N, Ricci GL, Carretta V, Persico M, Vinelli F, Scotto G, Bacca D, Annese M, Romano M, Zechini F, Sogari F, Spirito F, Andriulli A. Peginterferon alfa-2b and ribavirin for 12 vs 24 weeks in HCV genotype 2 or 3. N Engl J Med 2005;352(25):2609–2617
9. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, Ogura Y, Izumi N, Marumo F, Sato C. 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
10. El-Shamy A, Nagano-Fujii M, Sasase N, Imoto S, Kim SR, Hotta H. Sequence variation in hepatitis C virus nonstructural protein 5A predicts clinical outcome of pegylated interferon/ribavirin combination therapy. Hepatology 2008;48(1):38–47
11. Hamano K, Sakamoto N, Enomoto N, Izumi N, Asahina Y, Kurosaki M, Ueda E, Tanabe Y, Maekawa S, Itakura J, Watanabe H, Kakinuma S, Watanabe M. Mutations in the NS5B region of the hepatitis C virus genome correlate with clinical outcomes of interferon-alpha plus ribavirin combination therapy. J Gastroenterol Hepatol 2005;20(9):1401–1409
12. Chayama K, Suzuki F, Tsubota A, Kobayashi M, Arase Y, Saitoh S, Suzuki Y, Murashima N, Ikeda K, Takahashi N, Kinoshita M, Kumada H. Association of amino acid sequence in the PKR-eIF2 phosphorylation homology domain and response to interferon therapy. Hepatology 2000;32(5):1138–1144
13. Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, Kobayashi M, Saitoh S, Watahiki S, Sato J, Matsuda M, Arase Y, Ikeda K, Kumada H. 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
14. Toyoda H, Kumada T, Tada T, Arakawa T, Hayashi K, Honda T, Katano Y, Goto H. Association between HCV amino acid substitutions and outcome of peginterferon and ribavirin combination therapy in HCV genotype 1b and high viral load. J Gastroenterol Hepatol 2010;25(6):1072–1078
15. Donlin MJ, Cannon NA, Aurora R, Li J, Wahed AS, Di Bisceglie AM, Tavis JE. Contribution of genome-wide HCV genetic differences to outcome of interferon-based therapy in Caucasian American and African American patients. PLoS One 2010;5(2):e9032
16. Donlin MJ, Cannon NA, Yao E, Li J, Wahed A, Taylor MW, Belle SH, Di Bisceglie AM, Aurora R, Tavis JE. Pretreatment sequence diversity differences in the full-length hepatitis C virus open reading frame correlate with early response to therapy. J Virol 2007;81(15):8211–8224
17. Murakami T, Enomoto N, Kurosaki M, Izumi N, Marumo F, Sato C. Mutations in nonstructural protein 5A gene and response to interferon in hepatitis C virus genotype 2 infection. Hepatology 1999;30(4):1045–1053
18. Hayashi K, Katano Y, Honda T, Ishigami M, Itoh A, Hirooka Y, Nakano I, Urano F, Yoshioka K, Toyoda H, Kumada T, Goto H. Mutations in the interferon sensitivity-determining region of hepatitis C virus genotype 2a correlate with response to pegylated-interferon-alpha 2a monotherapy. J Med Virol 2009;81(3):459–466
19. Kobayashi M, Watanabe K, Ishigami M, Murase K, Ito H, Ukai K, Yano M, Takagi K, Hattori M, Kakumu S, Yoshioka K. Amino acid substitutions in the nonstructural region 5A of hepatitis C virus genotypes 2a and 2b and its relation to viral