

Supplementary Materials and Methods

Plasma Amino Acid Analysis

Plasma sample amino acid concentrations were measured by high-performance liquid chromatography-electrospray ionization-mass spectrometry followed by derivatization.¹ An MSQ Plus LC/MS system (Thermo Fischer Scientific, Waltham, MA) equipped with an electrospray ionization source was used in positive ionization mode for selected ion monitoring. Xcalibur version 1.4 SR1 software (Thermo Fischer Scientific, Yokohama, Japan) was used for data collection and processing. The high-performance liquid chromatography separation system consisted of an L-2100 pump, L-2200 autosampler, and L-2300 column oven (Hitachi High-Technologies Corporation, Tokyo, Japan). A Wakosil-II 3C8-100HG column (100, 2.1, 3 mm; Wako Pure Chemical Industries, Osaka, Japan) was used for the separation, and the mobile phase consisted of eluent A (25-mmol/L ammonium formate in water, pH 6.0) and eluent B (water:acetonitrile = 40:60).

Western Blotting

The expression of HCV core protein, Socs3, Foxo3a, phospho-Foxo3a (Ser253) (pFoxo3a), STAT1, pSTAT1 (Tyr701), S6K, pS6K, p-mTOR (Ser2448), Raptor, and β -actin were evaluated with mouse anti-core (Affinity BioReagents, Golden, CO), mouse anti-Socs3 (Santa Cruz Biotechnology, Santa Cruz, CA), rabbit anti-Foxo3a, rabbit anti- β -actin (Sigma-Aldrich, St Louis, MO), rabbit anti-phospho-Foxo3a (Ser253), rabbit anti-STAT1, rabbit anti-p-STAT1 (Tyr701), rabbit anti-p70 S6K, rabbit anti-pS6K, rabbit anti-p-mTOR (Ser2448), and rabbit anti-Raptor (Cell Signaling Technology, Beverly, MA), respectively. Densitometric analysis was conducted directly on the blotted membrane using a charge coupled device camera system (LAS-3000 Mini; Fujifilm, Tokyo, Japan) and Scion Image software (Frederick, MD).

Primer Sequences for PCR and siRNA

Primer sequences for PCR and siRNA were as follows: 2'S'OAS: forward 5'- CTC AGA AAT ACC CCA GCC AAA TC-3', reverse 5'-GTG GTG AGA GGA CTG AGG AA-3'; Socs3: forward 5'-TAC CAC CTG AGT CTC CAG CTT CTC-3', reverse 5'-CCT GGC AGT TCT CAT TAG TTC AGC ATT C-3'; Foxo3a: forward 5'-TGC TGT ATG CAA GAA CTT TCC AGT AGC AG-3', reverse 5'-ACT CTA GCC CCC ATG CTA CTA GTG-3'; glyceraldehyde-3-phosphate dehydrogenase: forward 5'-GAA GGT GAA GGT CGG AGT-3', reverse 5'-GAA GAT GGT GAT GGG ATT TC-3', siFoxo3a (SASI_Hs01_00119127; Sigma) sense: 5'-GAA UGA UGG GCU GAC UGA AdTdT-3', antisense: 5'-UUC AGU CAG CCC AUC AUU CdTdT-3'. Small interfering Raptor was purchased as

part of KIAA1303 siGENOME SMART pool siRNA reagents from Dharmacon, Inc (Lafayette, CO).

Construction of ISRE-Luc Reporter and FBEmut-luc Reporter Plasmids

Oligonucleotides containing the ISRE tandem repeat sequence (sense 5'-TCG AGA ACT GAA ACT GAA ACT GAA ACT GAA ACT GAA ACT GAA ACT GAA ACT GAA ACT GAA A-3', antisense 5'-AGC TTT TCA GTT TCA GTT TCA GTT TCA GTT TCA GTT TCA GTT TCA GTT TCA GTT TCA GTT C-3', consensus 5'-GAA Ann GAA ACT-3') were annealed, and integrated into Xho I and Hind III sites of the pGL4.23 luciferase vector (Promega). The human Socs3 promoter region (-109/+217) was amplified by genomic PCR using specific primers (forward, 5'-TGC TGC GAG TAG TGA CTA AAC ATT ACA AG-3' and reverse, 5'-CCG TGA AGT CCA CAA AGG AGC CTT C-3') and cloned into the EcoR V site of the pGL4.10-luc2 reporter vector (Promega). The Socs3 FBE mutant reporter vector was created by substituting 2 adenines in the putative FBE with guanines (wild-type sequence 5'-CTAAACA-3', mutated sequence 5'-CTGAGCA-3').

ChIP Assay

For the ChIP assay using the anti-ISGF3 γ antibody, 1×10^6 Huh-7 cells were treated with IFN- α (0 or 100 U/mL) and BCAA (2 mmol/L) in low-amino-acid medium for 6 hours. For ChIP using the anti-Foxo3a antibody, 1×10^6 Huh-7 cells were cultured in low-amino-acid medium for 24 hours.

Cells were cross-linked with 1% formaldehyde in PBS for 10 minutes at 37°C, and the reaction was stopped with 250 mmol/L glycine for 10 minutes. Cells were suspended in sodium dodecyl sulfate-lysis buffer (1% sodium dodecyl sulfate, 10 mmol/L ethylenediaminetetraacetic acid [EDTA], 50 mmol/L Tris-HCl [pH 8.1]), complete protease inhibitor cocktail (Roche Applied Science), and incubated for 30 minutes at 10°C. Cell lysate was sonicated with Bioruptor (Cosmo Bio, Tokyo, Japan) to obtain chromatin fragments and diluted 10-fold in ChIP dilution buffer (0.01% sodium dodecyl sulfate, 1.1% Triton-X 100, 1.2 mmol/L EDTA, 16.7 mmol/L Tris-HCl [pH 8.1], 150 mmol/L NaCl, complete protease inhibitor cocktail). Chromatin fragments were incubated with 2 μ g ISGF3 γ antibody (Santa Cruz Biotechnology), 2 μ g Foxo3a antibody (H-144; Santa Cruz Biotechnology), or normal rabbit immunoglobulin G for 18 hours at 4°C. Dynabeads (30 μ L) protein G (Invitrogen) was added and incubated for 1 hour at 4°C. The beads were washed with low-salt-wash buffer (0.1% sodium dodecyl sulfate, 1% Triton-X 100, 2.0 mmol/L EDTA, 20 mmol/L Tris-HCl [pH 8.1], 150 mmol/L NaCl), high-salt-wash buffer (0.1% sodium dodecyl sulfate, 1% Triton-X 100, 2.0 mmol/L EDTA, 20 mmol/L Tris-HCl [pH 8.1], 500 mmol/L NaCl), LiCl wash buffer (250 mmol/L LiCl, 1% NP-40, 1% de-

oxycholate, 1.0 mmol/L EDTA, 1.0 mmol/L Tris-HCl [pH 8.1]) and Tris-EDTA buffer. Immunoprecipitated chromatin fragments were eluted with elution buffer (1% sodium dodecyl sulfate, 100 mmol/L NaHCO₃, 10 mmol/L dithiothreitol), and reverse cross-linked by incubating for 6 hours at 65°C in elution buffer containing 200 mmol/L NaCl. DNA fragments were purified and quantified by real-time detection PCR with primers for putative ISRE in the 2'5'OAS promoter region (forward, 5'-AAA TGC ATT TCC AGA GCA GAG TTC AGA G-3', reverse, 5'-GGG TAT TTC TGA GAT CCA TCA TTG ACA GG-3') or putative FBE in the Socs3 promoter region (forward, 5'-TGC TGC GAG TAG TGA CTA AAC ATT ACA AG-3', reverse, 5'-AGC GGA GCA GGG AGT CCA AGT C-3'). Values were normalized by the measurement of input DNA.

pH77S.3/GLuc2A

pH77S.2 is a modification of pH77S² containing an additional mutation within the E2 protein (N476D in the polyprotein) that promotes infectious virus yields from RNA-transfected cells (Yi et al, unpublished data). To monitor replication, the GLuc sequence, fused at its C terminus to the foot-and-mouth disease virus 2A autoprotease, was inserted between p7 and NS2 of pH77S.2 (Supplementary Figure 4). To insert the GLuc-coding sequence between p7 and NS2 in pH77S.2, followed by the foot-and-mouth disease virus 2A protein-coding sequence, Mlu I, EcoR V, and Spe I restriction sites were created between the p7 and NS2 coding sequences by site-directed mutagenesis. DNA coding for GLuc was subcloned into the Mlu I and EcoR V sites of the modified plasmid after PCR amplification using the primers: 5'-ATA ATA TTA CGC GTA TGG GAG TCA AAG TTC TGT TTG CC-3' (sequence corresponding to the N-terminal GLuc is italicized and that corresponding to Mlu I is underlined) and 5'-ATA AAT AGAT ATC GTC ACC ACC GGC CCC CTT GAT CTT-3' (C terminal GLuc is italicized and EcoR V is underlined). A DNA fragment encoding the 17 amino acids of the foot-and-mouth disease virus 2A protein was generated by annealing the following complementary oligonucleotides: 5'-ATA TGA TAT CAA CTT TGA CCT TCT CAA GTT GGC CGG CGA CGT

CGA GTC CAA CCC AGG GCC CAC TAG CAT AT-3' and 5'-ATA TGC TAG TGG GCC CTG GGT TGG ACT CGA CGT CGC CGG CCA ACT TGA GAA GGT CAA AGT TGA TAT CAT AT-3' (underlined sequences indicate EcoR V and Spe I sites). The annealed oligonucleotides were digested by both restriction enzymes and the product inserted into the corresponding sites of pH77S.2 containing GLuc to generate pH77S.2/GLuc2A. Q41R is a cell-culture adaptive mutation within the NS3 protease domain of pH77S. Because it is not essential for production of infectious virus (Yi et al, unpublished data), pH77S.2 and pH77S.2/GLuc2A constructs underwent this mutation by site-directed mutagenesis of a PCR fragment spanning the Afe I and BsrG I sites to replace Gln₄₁ with wild-type Arg. The resulting plasmids (pH77S.2/R41Q and pH77S.2/GLuc2A/R41Q) were redesignated pH77S.3 and pH77S.3/GLuc2A, respectively.^{3,4} GLuc has several advantages over other luciferase reporter enzymes in that it is smaller and allows more sensitive detection than either firefly or Renilla luciferase.^{3,4} In addition, a signal sequence directs its secretion into cell-culture media, allowing real-time dynamic measurements of GLuc expression without the need for cell lysis. H77S.3/GLuc2A RNA produces infectious virus, although with lower efficiency than H77S.3 RNA (10-fold less).

References

1. Shimbo K, Kubo S, Harada Y, et al. Automated precolumn derivatization system for analyzing physiological amino acids by liquid chromatography/mass spectrometry. *Biomed Chromatogr* 2009; 24:683–691.
2. Yi M, Villanueva RA, Thomas DL, et al. Production of infectious genotype 1a hepatitis C virus (Hutchinson strain) in cultured human hepatoma cells. *Proc Natl Acad Sci U S A* 2006;103:2310–2315.
3. Shetty S, Kim S, Shimakami T, et al. Hepatitis C virus genomic RNA dimerization is mediated via a kissing complex intermediate. *RNA* 2010;16:913–925.
4. Shimakami T, Welsch C, Yamane D, et al. Protease inhibitor-resistant hepatitis C virus mutants with reduced fitness from impaired production of infectious virus. *Gastroenterology* 2011; 140:667–675.

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
© Asian Pacific Association for the Study of the Liver 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

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

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

hepatocellular carcinoma [1]. In HCV-infected patients with chronic hepatitis, treatment with interferon (IFN) can result in viral clearance and biochemical and histological improvements [2]. The response to the therapy varies according to HCV genotype and pretreatment HCV RNA level [3, 4].

The currently recommended treatment for patients infected genotype 2a HCV with high viral load is pegylated-interferon (PEG-IFN) plus ribavirin (RBV) for 24 weeks [1]. Approximately 80% of patients infected with genotype 2a HCV can achieve a sustained virological response (SVR) with this regimen [5, 6], although much lower percentages of patients infected with other genotypes can achieve SVR, especially with genotype 1 [1]. Because of its high response rate, shorter treatment duration was suggested by some studies, although an agreement has not been reached yet [7, 8]. On the other hand, about 20% of patients infected with this genotype cannot achieve SVR and it remains elusive which patients show poor responses.

Previous studies have reported that amino acid variations in the NS5A-interferon sensitivity determinant region (ISDR) [9], NS5A-interferon ribavirin resistance determinant region (IRRDR) [10], NS5B [11], and PKR-eIF2 phosphorylation homology domain (PePHD) of E2 [12], and core [13, 14] correlate with clinical outcome of IFN-based therapy, including PEG-IFN/RBV therapy in patients infected with genotype 1b HCV. Recent full HCV open reading frame (ORF) analysis for genotype 1 also has reported that core, NS3, and NS5A were associated with early viral response and the outcome in PEG-IFN/RBV therapy [15, 16]. However, in genotype 2a infection, only a few studies have investigated the association between HCV sequence variation and treatment response [17–19], and the role of viral factors has not been established yet, especially in the era of PEG-IFN/RBV therapy. Moreover, these previous studies investigated only several isolated HCV genomic regions, and comprehensive analysis of the full HCV ORF has not been undertaken so far.

In the present study, to assess comprehensively the influence of viral variations on response to the PEG-IFN/RBV therapy in genotype 2a HCV infection, we determined the complete pretreatment HCV ORFs from Japanese patients and investigated viral amino acid variation and their correlation with the response to the combination therapy of PEG-IFN plus RBV.

Patients and methods

Study population

A total of 103 adult Japanese patients infected with genotype 2a HCV, who received the combination therapy with

PEG-IFN (PEGINTRON[®], Schering-Plough, Tokyo, Japan) plus RBV (REBETOL[®], Schering-Plough) between 2005 and 2008 at the University of Yamanashi, Tokyo Medical and Dental University, and related institutions were first included in the study. They all fulfilled the following criteria: (1) negative for hepatitis B surface antigen; (2) high viral load (≥ 100 KIU/ml); (3) absence of hepatocellular carcinoma; (4) no other form of hepatitis, such as primary biliary cirrhosis, autoimmune liver disease, or alcoholic liver disease; and (5) free of co-infection with human immunodeficiency virus. Informed consent was obtained from each patient. The study was approved by the ethics committees of all the participating universities and hospitals. The therapy was performed according to the standard treatment protocol of PEG-IFN/RBV therapy for Japanese patients established by a hepatitis study group of the Ministry of Health, Labour, and Welfare, Japan (PEG-IFN α -2b 1.5 μ g/kg body weight, once weekly subcutaneously, and RBV 600–800 mg daily per os for 24 weeks). To clearly disclose the non-SVR viral characteristics, we have considered those patients who achieved total drug administration of 60% or more for both PEG-IFN and RBV, with the completion of the standard treatment duration. Moreover, although we excluded the patients with extended therapy to make the studied population uniform, we have included non-SVR patients with extended therapy to clarify the specific characteristics of non-SVR patients, a minor population group. As a result, 25 patients were excluded for the following reasons: 4 patients received insufficient dose, 8 patients were discontinued from the therapy within 12 weeks, and 13 SVR patients received extended therapy. Finally, 78 patients were considered as eligible for the study. During the combination therapy, blood samples were obtained at least once every month before, during, and after treatment and were analyzed for blood count, ALT, and HCV RNA levels. Liver biopsy specimens were obtained from most of the patients.

The 78 patients belonging to the different institutions were separately analyzed: 43 patients registered in Y-PERS (Yamanashi Pegintron Ribavirin Study Group) were included in group 1 (test group), and the 35 patients from Tokyo Medical and Dental University and related institutions (Ochanomizu Liver Conference Group) were included in group 2 (validation group). We divided the patients into these two groups to exclude false positives (type I errors) which might arise in successive HCV-ORF study. Since genotype-2a HCV contains as many as 3,033 amino acids, it was possible that incorrect amino acids can be judged as significant in full HCV-ORF comparison study as a result of type I errors. Therefore, to guard against false positives, HCV-ORF comparison study was undertaken in group 1, group 2, and combined group.

Complete HCV-ORF sequence determination by direct sequencing from pretreatment sera

HCV RNA was extracted from pretreatment serum samples by the AGPC method using Isogen (Wako, Osaka, Japan) according to the manufacturer's protocol. Complementary deoxyribonucleic acid (DNA) was synthesized with Superscript II (Invitrogen, Tokyo, Japan) using random primers (Invitrogen) and then amplified by two-step nested PCR using the primers newly designed for this study. All samples were initially denatured at 95°C for 7 min, followed by 40 cycles with denaturation at 95°C for 15 s, annealing at 55°C for 15 s, and extension at 72°C for 45 s with BD AdvantageTM 2 PCR Enzyme System (BD Biosciences Clontech, CA, USA).

PCR amplicons were sequenced directly by Big Dye Terminator Version 3.1 (ABI, Tokyo, Japan) with universal M13 forward/M13 reverse primers using an ABI prism 3130 sequencer (ABI). Generated sequence files were assembled using Vector NTI software (Invitrogen) and base-calling errors were corrected following inspection of the chromatogram.

Sliding window analysis

A sliding window analysis was introduced to search through HCV amino acid “regions”, rather than single amino acid positions, related to the final outcome of PEG-IFN/RBV therapy. Briefly, the total number of amino acid substitutions compared to the consensus sequence within a given amino acid length was counted in each amino acid position in each HCV sequence. Then the relation of substitution numbers and the final outcome was compared statistically between the SVR and non-SVR groups by Mann–Whitney's *U* test for each amino acid position. In this study, we changed the window length from 1 to 50 to search for those HCV regions. To visualize the result, significantly lower *p* values were colored in red and non-significant *p* values were colored in green to generate a “heat map” appearance using Microsoft Excel software. In the present study, *p* value of 1/1,000 or lower was colored in the maximum red.

Statistical analysis

Statistical differences in the parameters, including all available patients' demographic, biochemical, hematological, and virological data, such as sequence variation factors, were determined between the various groups by Student *t* test or Mann–Whitney's *U* test for numerical variables and Fisher's exact probability test for categorical variables. To evaluate the optimal threshold of variations for SVR prediction, the receiver operating characteristic

curve was constructed. Variables that achieved statistical significance ($p < 0.05$) in univariate analysis were entered into multiple logistic regression analysis to identify significant independent factors. We also calculated the odds ratios and 95% confidence intervals. All *p* values < 0.05 by the two-tailed test were considered significant.

Results

Characteristics of the patients studied

Of the patients analyzed, the SVR rate was 78.3% (58/74) with the standard therapy (four non-SVR patients received an extended therapy). The baseline characteristics of the patients (group 1, group2, and combined) classified according to SVR achievement are shown in Table 1. Fibrosis score ($p = 0.047$) and HCV RNA levels ($p = 0.002$) were significantly higher in non-SVR patients, but the cumulative RBV dose $\geq 80\%$ ($p = 0.003$) and rapid virological response (RVR) rate ($p = 0.011$) were significantly higher in SVR patients. In addition, patients with non-SVR had a tendency to be older ($p = 0.058$). Achievement of RVR reached 61.5% when all patients were included, and this rate was extremely high compared to achievement of RVR in patients with genotype 1b infection ($\sim 10\%$) observed in Yamanashi University Hospital (data not shown). The early virological response (EVR) rate was equally high in the SVR (100%) and non-SVR (89%) groups, showing that relapse to be the characteristic feature of the non-SVR patients with genotype 2a HCV. Actually, 18 patients in non-SVR were relapsers, while two patients were null responders.

Comparison of amino acid variations between the SVR and non-SVR in the complete HCV polyprotein and each HCV protein

To determine whether the sequence variations differed between the SVR and non-SVR groups, we first compared amino acid variations that were unique, relative to a population consensus, to either the SVR or non-SVR patients for the complete HCV polyprotein and each HCV protein. The number of amino acid variations in the sequences from the SVR patients was significantly higher than in those from the non-SVR patients, when the entire HCV polyprotein was analyzed (Fig. 1, left). These differences were especially significant in E1 and NS3 (Fig. 1, right). This result demonstrated that HCV sequences from patients with SVR comprised a heterogeneous population, while HCV sequences from patients with non-SVR comprised a rather homogeneous population, indicating the existence of unique non-responsive HCV sequences.

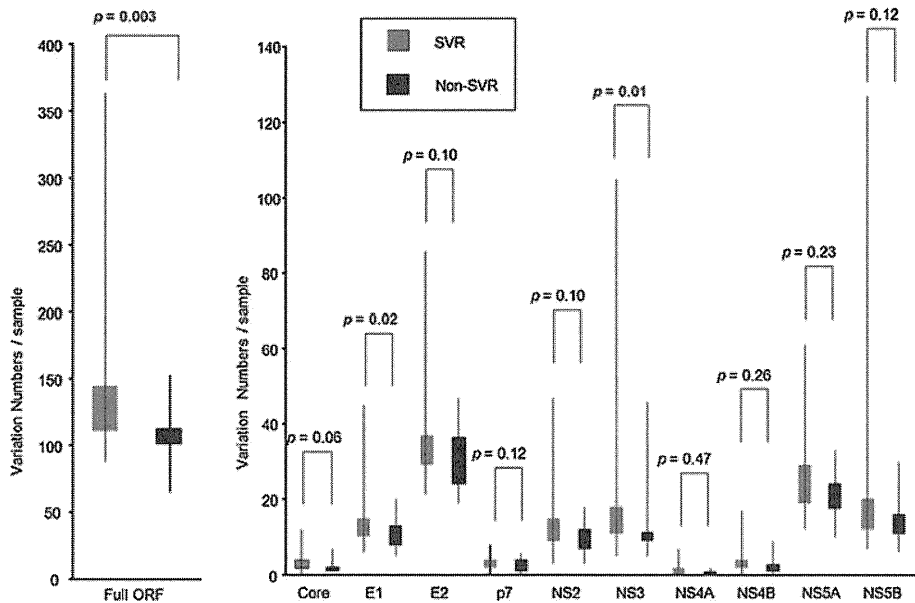
Table 1 Baseline characteristics of all patients (groups 1 and 2)

Characteristic	SVR (<i>n</i> = 58)			Non-SVR (<i>n</i> = 20)			<i>p</i> value ^Δ
	Group 1 (<i>n</i> = 36)	Group 2 (<i>n</i> = 22)	Combined (<i>n</i> = 58)	Group 1 (<i>n</i> = 7)	Group 2 (<i>n</i> = 13)	Combined (<i>n</i> = 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 (×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
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; ^{||} Mann–Whitney's *U* test; ^Δ *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 ($p = 5\text{E}-05$). 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

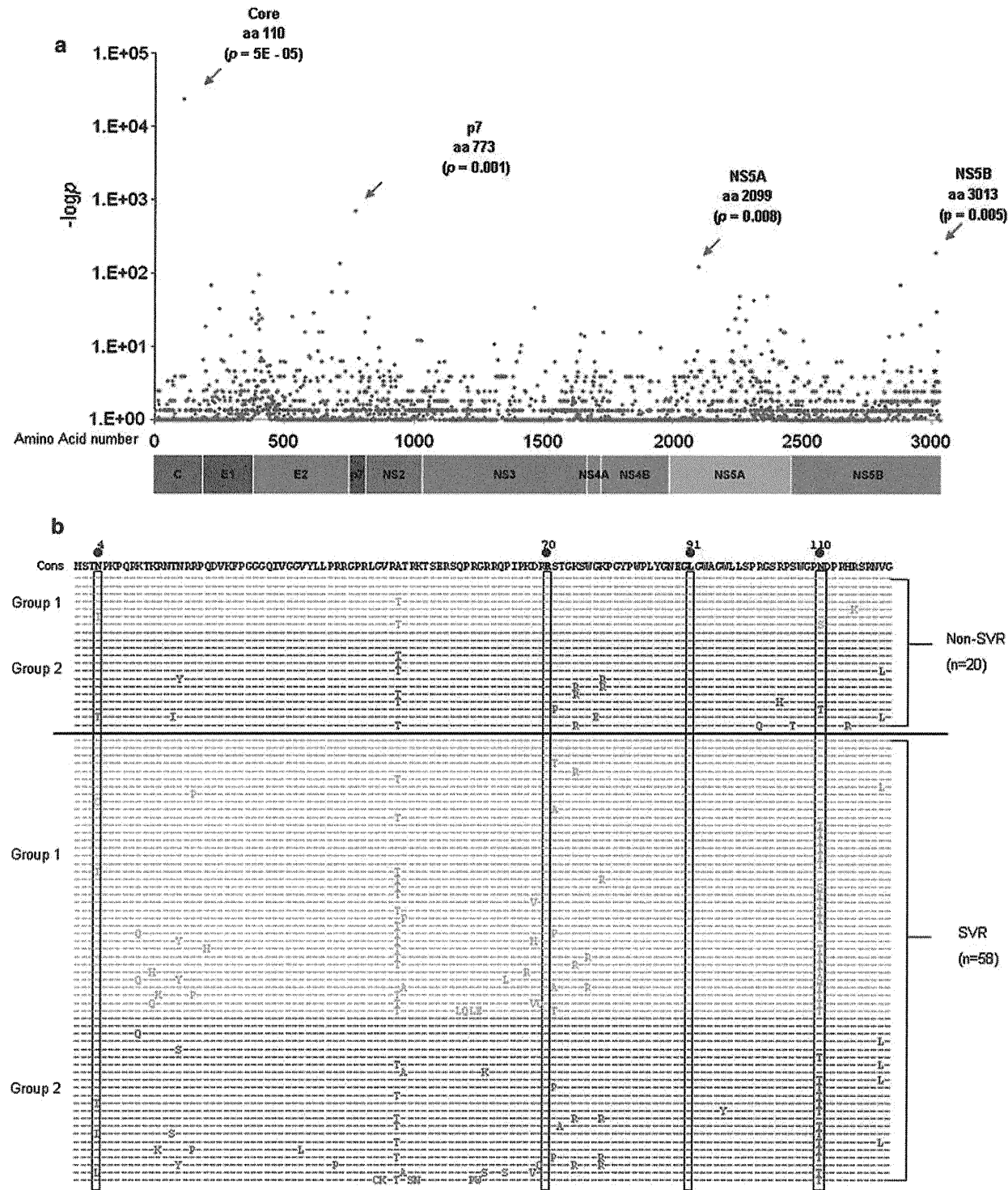


Fig. 2 **a** Different amino acid usages at each viral amino acid position between the sustained viral responders (SVR) and the non-sustained viral responders (non-SVR) patients. Amino acid variation was determined between SVR and non-SVR patients by Fisher's exact probability test. The longitudinal 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*

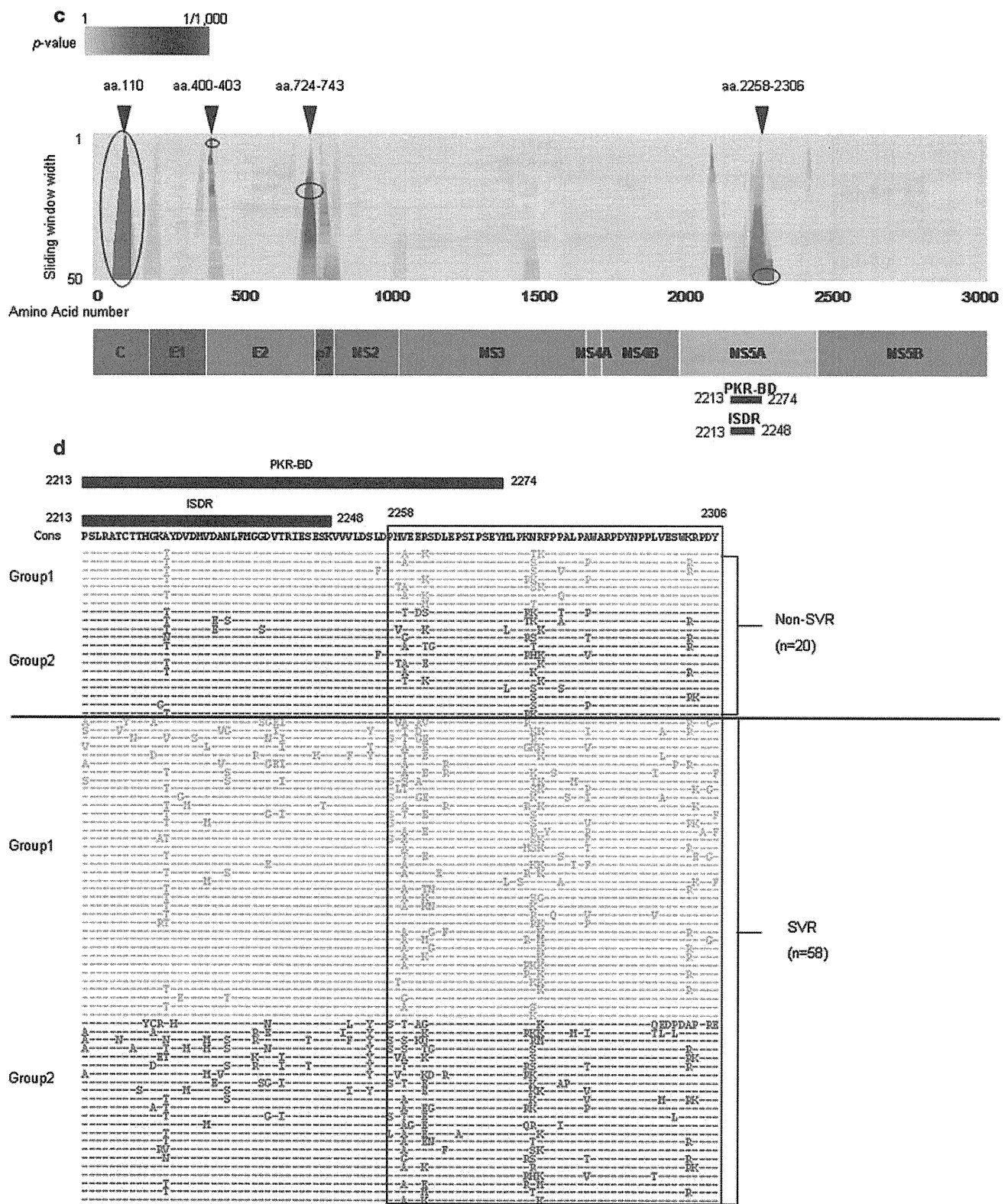


Fig. 2 continued

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 (<i>n</i> = 43)	<i>p</i> value	Group 2 (<i>n</i> = 35)	<i>p</i> value	Combined (<i>n</i> = 78)	<i>p</i> 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 (<i>n</i> = 43)	<i>p</i> value	Group 2 (<i>n</i> = 35)	<i>p</i> value	Combined (<i>n</i> = 78)	<i>p</i> 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

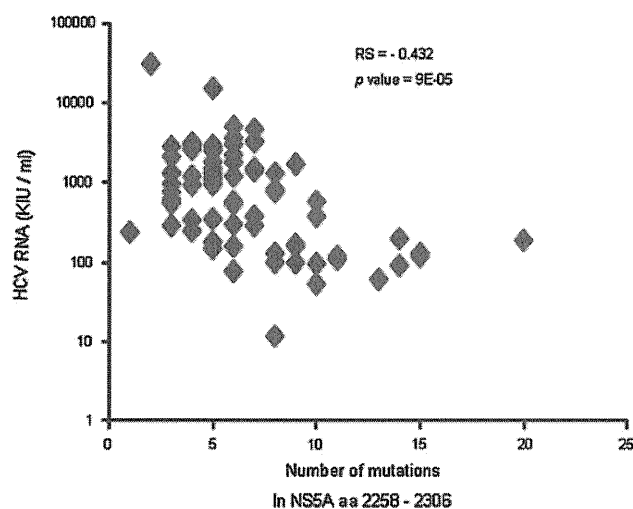
correlation with the response as above, the region was excluded from further analysis. Among the other three regions, only NS5A aa 2258–2306 showed significant differences in the two independent study groups (Table 3). Interestingly, the NS5A region overlapped the PKR-binding domain, which includes the 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)	<i>p</i> value
Age	1.01 (0.91–1.13)	0.85
HCV RNA	1.00 (1.00–1.00)	0.09
Fibrosis score $\geq 3/0-2$	2.37 (0.21–26.7)	0.48
RVR achievement	3.46 (0.54–22.1)	0.19
Ribavirin dose $\geq 80\%$	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 ($p = 9\text{E}-05$, 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 ($p = 5\text{E}-05$). 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 (<i>n</i> = 40)	Mutation ≥ 5 (<i>n</i> = 38)	<i>p</i> 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 (×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 (<i>n</i> = 33)	Core 110 N/S (<i>n</i> = 45)	<i>p</i> 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 (×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

- Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. *Hepatology* 2009;49(4):1335–1374
- 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
- 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
- Simmonds P. Clinical relevance of hepatitis C virus genotypes. *Gut* 1997;40(3):291–293
- 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
- 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
- Dalgard O, Bjoro K, Hellum KB, Myrvang B, Ritland S, Skaug K, Raknerud N, Bell H. Treatment with pegylated interferon and ribavirin in HCV infection with genotype 2 or 3 for 14 weeks: a pilot study. *Hepatology* 2004;40(6):1260–1265
- Mangia A, Santoro R, Minerva N, Ricci GL, Carretta V, Persico M, Vinelli F, Scotto G, Bacca D, Annesse 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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
- 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

- load and response to interferon. *Am J Gastroenterol* 2002;97(4): 988–998
20. Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y, Hosaka T, Kobayashi M, Saitoh S, Arase Y, Ikeda K, Kumada H. Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 2a high viral load and virological response to interferon–ribavirin combination therapy. *Intervirology* 2009;52(6):301–309
 21. Saito T, Ito T, Ishiko H, Yonaha M, Morikawa K, Miyokawa A, Mitamura K. Sequence analysis of PePHD within HCV E2 region and correlation with resistance of interferon therapy in Japanese patients infected with HCV genotypes 2a and 2b. *Am J Gastroenterol* 2003;98(6):1377–1383
 22. Watanabe H, Nagayama K, Enomoto N, Itakura J, Tanabe Y, Sato C, Izumi N, Watanabe M. Amino acid substitutions in PKR-eIF2 phosphorylation homology domain (PePHD) of hepatitis C virus E2 protein in genotype 2a/2b and 1b in Japan and interferon efficacy. *Hepatol Res* 2003;26(4):268–274
 23. Gale M Jr, Blakely CM, Kwieciszewski B, Tan SL, Dossett M, Tang NM, Korth MJ, Polyak SJ, Gretch DR, Katze MG. Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. *Mol Cell Biol* 1998;18(9):5208–5218
 24. Hiramatsu N, Oze T, Yakushijin T, Inoue Y, Igura T, Mochizuki K, Imanaka K, Kaneko A, Oshita M, Hagiwara H, Mita E, Nagase T, Ito T, Inui Y, Hijioka T, Katayama K, Tamura S, Yoshihara H, Imai Y, Kato M, Yoshida Y, Tatsumi T, Ohkawa K, Kiso S, Kanto T, Kasahara A, Takehara T, Hayashi N. Ribavirin dose reduction raises relapse rate dose-dependently in genotype 1 patients with hepatitis C responding to pegylated interferon alpha-2b plus ribavirin. *J Viral Hepat* 2009;16(8):586–594
 25. Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009;461(7262):399–401
 26. Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, Hino K, Hige S, Ito Y, Mita E, Tanaka E, Mochida S, Murawaki Y, Honda M, Sakai A, Hiasa Y, Nishiguchi S, Koike A, Sakaida I, Imamura M, Ito K, Yano K, Masaki N, Sugauchi F, Izumi N, Tokunaga K, Mizokami M. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* 2009;41(10):1105–1109
 27. Rauch A, Kutalik Z, Descombes P, Cai T, di Iulio J, Mueller T, Bochud M, Battegay M, Bernasconi E, Borovicka J, Colombo S, Cerny A, Dufour JF, Furrer H, Gunthard HF, Heim M, Hirschel B, Malinverni R, Moradpour D, Mullhaupt B, Witteck A, Beckmann JS, Berg T, Bergmann S, Negro F, Telenti A, Bochud PY. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure—a genome-wide association study. *Gastroenterology* 2010 Jan 7

Analysis of the Complete Open Reading Frame of Genotype 2b Hepatitis C Virus in Association with the Response to Peginterferon and Ribavirin Therapy

Makoto Kadokura^{1*}, Shinya Maekawa^{1*}, Ryota Sueki¹, Mika Miura¹, Kazuki Komase¹, Hiroko Shindo¹, Fumitake Amemiya¹, Tomoyoshi Uetake¹, Taisuke Inoue¹, Minoru Sakamoto¹, Mina Nakagawa², Naoya Sakamoto², Mamoru Watanabe², Nobuyuki Enomoto¹

¹ First Department of Internal Medicine, Faculty of Medicine, University of Yamanashi, Chuo, Yamanashi, Japan, ² Department of Gastroenterology and Hepatology, Tokyo Medical and Dental University, Bunkyo, Tokyo, Japan

Abstract

Background and Aims: Patients infected with genotype 2b hepatitis C virus (HCV) generally can achieve favorable responses to pegylated-interferon plus ribavirin therapy (PEG-IFN/RBV). However, a proportion of patients show poorer responses and the correlation between viral sequence variation and treatment outcome remains unclear.

Methods: The pretreatment complete open reading frame (ORF) sequences of genotype 2b HCV determined by direct sequencing were investigated for correlation with the final outcome in a total of 60 patients.

Results: In this study group, 87.5% (14/16) of non-sustained virological response (non-SVR) patients (n = 16) were relapsers. Compared to sustained virological response (SVR) patients (n = 44), non-SVR patients were older and could not achieve prompt viral clearance after the therapy induction. Comparing each viral protein between the two groups, viral sequences were more diverse in SVR patients and that diversity was found primarily in the E1, p7, and NS5A proteins. In searching for specific viral regions associated with the final outcome, several regions in E2, p7, NS2, NS5A, and NS5B were extracted. Among these regions, part of the interferon sensitivity determining region (ISDR) was included. In these regions, amino acid substitutions were associated with the final outcome in an incremental manner, depending upon the number of substitutions.

Conclusions: Viral sequences are more diverse in SVR patients than non-SVR patients receiving PEG-IFN/RBV therapy for genotype-2b HCV infection. Through systematic comparison of viral sequences, several specific regions, including part of the ISDR, were extracted as having significant correlation with the final outcome.

Citation: Kadokura M, Maekawa S, Sueki R, Miura M, Komase K, et al. (2011) Analysis of the Complete Open Reading Frame of Genotype 2b Hepatitis C Virus in Association with the Response to Peginterferon and Ribavirin Therapy. PLoS ONE 6(9): e24514. doi:10.1371/journal.pone.0024514

Editor: John E. Tavis, Saint Louis University, United States of America

Received: March 10, 2011; **Accepted:** August 12, 2011; **Published:** September 15, 2011

Copyright: © 2011 Kadokura et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: 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). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding was received for this study.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: maekawa@yamanashi.ac.jp

These authors contributed equally to this work.

Introduction

Worldwide, 180 million people are estimated to be infected with hepatitis C virus (HCV), a major cause of chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma (HCC) [1]. In HCV-infected patients with chronic hepatitis, treatment with interferon (IFN)-based therapy can result in viral clearance as well as biochemical and histological improvements [2]. In this IFN-based therapy, HCV genotype is the most significant factor affecting treatment responses [3,4].

In genotype 2b HCV infection, 80% of patients with high viral titers can achieve a sustained virological response (SVR) to the regimen of pegylated-interferon (PEG-IFN) plus ribavirin (RBV) for 24 weeks [5,6]. This response is high considering that much

lower percentages of patients infected with other genotypes can achieve SVR, especially with genotype 1 [1]. However, in other words, 20% of patients infected with genotype 2b HCV still cannot clear the virus and remain at risk of developing HCC. On the other hand, although various studies have been undertaken to clarify the factors contributing to the response to IFN-based therapy in genotype 1 infection, it remains poorly understood which patients with genotype 2b HCV infection will show unfavorable responses. Recently, the significance of IL28B single nucleotide polymorphisms (SNPs) in determining the response to PEG-IFN/RBV therapy was demonstrated in genotype 1 HCV infection [7,8]. However, the significance of IL28B SNPs was rather weak in genotype 2 HCV infection [9].

In terms of the association between HCV sequence variation and treatment responses, previous studies have reported that

Table 1. Baseline Characteristics of Studied Patients.

Characteristic	SVR (n = 44)	non-SVR (n = 16)	P value
Gender (Male/Female)	26/18	9/7	NS [‡]
Age (yrs)	56 (22–72) [*]	59 (30–80)	0.04 [‡]
BMI	23.5 (16.6–30.3)	24.7 (18.5–31.7)	NS [‡]
ALT (IU/l)	51 (19–380)	41 (17–390)	NS [‡]
GGTP (IU/l)	36 (11–133)	40 (17–292)	NS [‡]
T.Chol (mg/dl)	169 (119–225)	178 (145–217)	NS [‡]
WBC (/μl)	4600 (2620–7200)	5080 (3270–8600)	NS [‡]
Hb (g/dl)	14.2 (11.5–17.3)	14.6 (11.8–16.4)	NS [‡]
Platelet (×10 ⁴ /mm ³)	19 (7.1–31.8)	17.8 (8–36.7)	NS [‡]
Fibrosis score (0–2/≥3) [§]	38/5	7/3	NS [†]
HCV RNA (KIU/ml)	2050 (100–16000)	1800 (140–6300)	NS [‡]
IFN dose (≥80%/60–80%)	36/8	13/3	NS [†]
Ribavirin dose (≥80%/60–80%)	32/12	10/6	NS [†]
RVR rate (%)	55.8	6.3	0.0008 [†]
EVR rate (%)	97.7	68.8	0.004 [†]
ETR rate (%)	100	87.5	NS [†]

§: SVR : n = 43, non-SVR : n = 10.
*: median (range).
†: Fisher's exact probability test.
‡: Mann-Whitney's U test.
doi:10.1371/journal.pone.0024514.t001

amino acid variation in the NS5A-ISDR [10], NS5A-IRRDR [11], NS5B [12], PKR-eIF2 phosphorylation homology domain (PePHD) of E2 [13], and Core [14] correlate with the clinical outcome of IFN-based therapy, including PEG-IFN/RBV therapy for genotype 1b HCV infection. In the meantime, these viral sequence studies have been controversial regarding their true clinical importance, because the results of different studies were not always coincident [15,16,17]. On this background, recent studies trying to analyze the correlation of complete HCV open reading frame diversity, clinical characteristics, and the response to PEG-IFN/RBV therapy for genotype 1 HCV infection, in the most comprehensive approach yet attempted, have clarified that viral amino acid variation is associated with treatment responses, with consideration of racial background [18,19]. In genotype 2 infection, however, only a few studies have investigated the association of HCV sequence variation and treatment response [20,21] and the clinical significance has been yet established. We reported recently that variation of amino acid (aa) 110 in Core and amino acids (aa) 2258–2308 in NS5A were significantly associated with treatment outcome of the PEG-IFN/RBV therapy for genotype 2a HCV infection, through the analysis of the complete HCV ORFs in Japanese patients [22].

In this study, to assess comprehensively the influence of viral sequence variation on the response to the PEG-IFN/RBV therapy in genotype 2b HCV infection, we determined the complete pretreatment HCV ORFs from Japanese patients and investigated amino acid variation and its correlation with the response to combination therapy with PEG-IFN plus RBV.

Methods

Patients

A total of 77 adult Japanese patients infected with genotype 2b HCV, who received the combination therapy with PEG-IFN

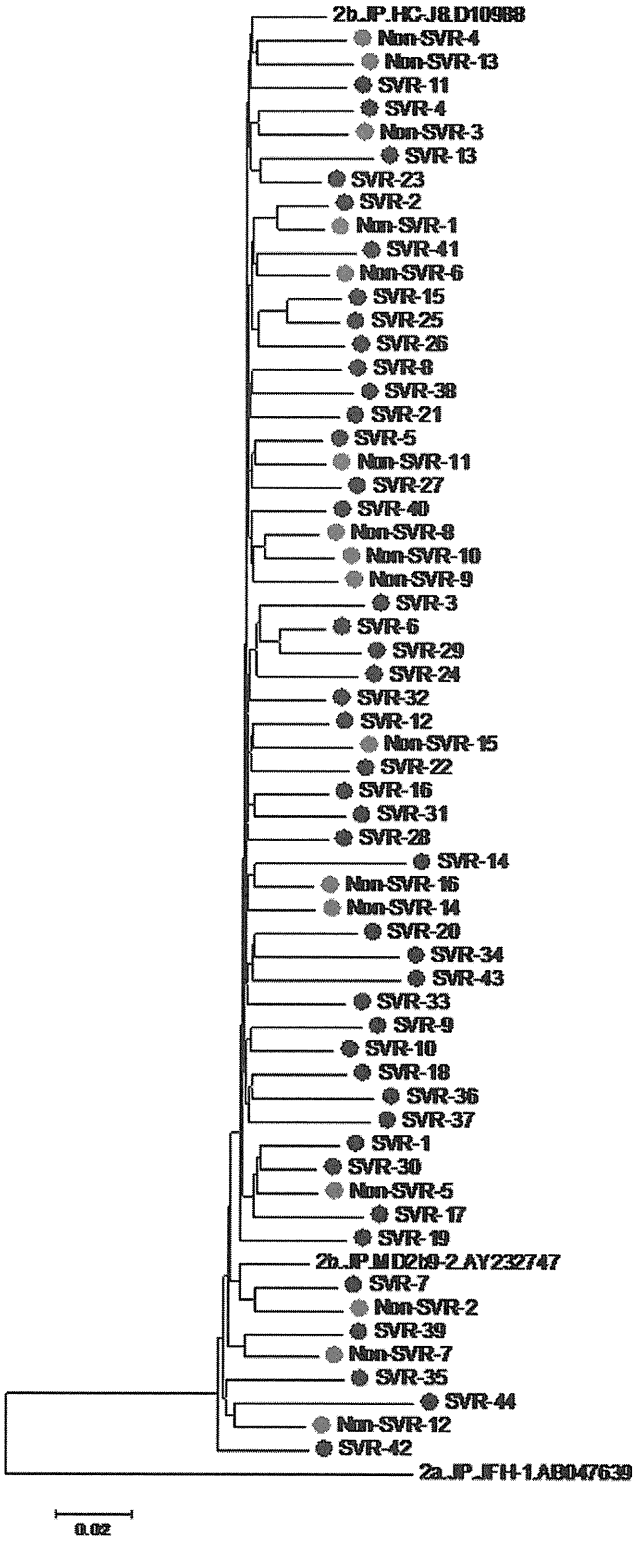


Figure 1. Phylogenetic analysis of the genotype-2b polyprotein sequences. In order to perform the phylogenetic analysis, we first aligned all 60 HCV complete ORF amino acid sequences obtained from the patients along with reference sequences (2b.HC-J8.D10988, 2.JP.MD2b9-2, and 2a.JP.JFH-1.AB047639), using the ClustalW program, and constructed the phylogenetic tree using the Neighbor-Joining method with MEGA version 4 software. Blue circles indicate SVR patients and red circles indicate non-SVR patients.
doi:10.1371/journal.pone.0024514.g001

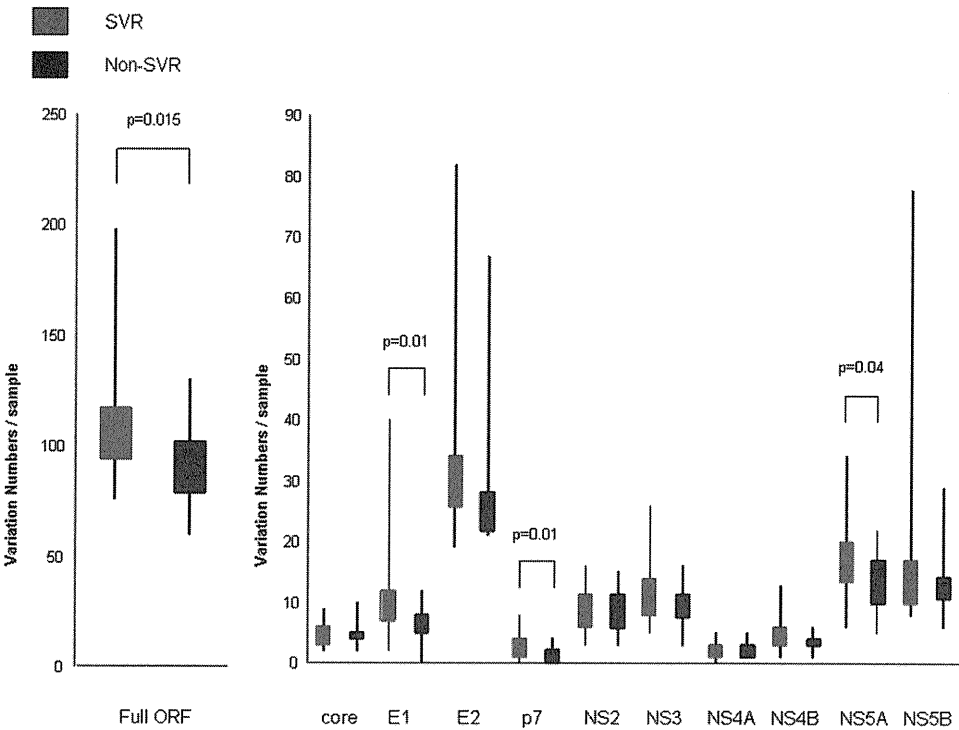


Figure 2. Number of amino acid substitutions per sample in the sustained virological responders (SVR) and the non-sustained virological responders (non-SVR) group. The numbers of variations, relative to a population consensus, that were unique to either SVR or non-SVR patients are shown for the complete open reading frame (ORF) (Fig. 1, left) and for each HCV protein (Fig. 1, right).
doi:10.1371/journal.pone.0024514.g002

(PEGINTRON®, Schering-Plough, Tokyo, Japan) plus RBV (REBETOL®, Schering-Plough) between 2005 and 2009 at University of Yamanashi, Tokyo Medical and Dental University,

and related institutions were first included in the study. They all fulfilled following criteria: (1) negative for hepatitis B surface antigen, (2) high viral load (≥ 100 KIU/ml), (3) absence of

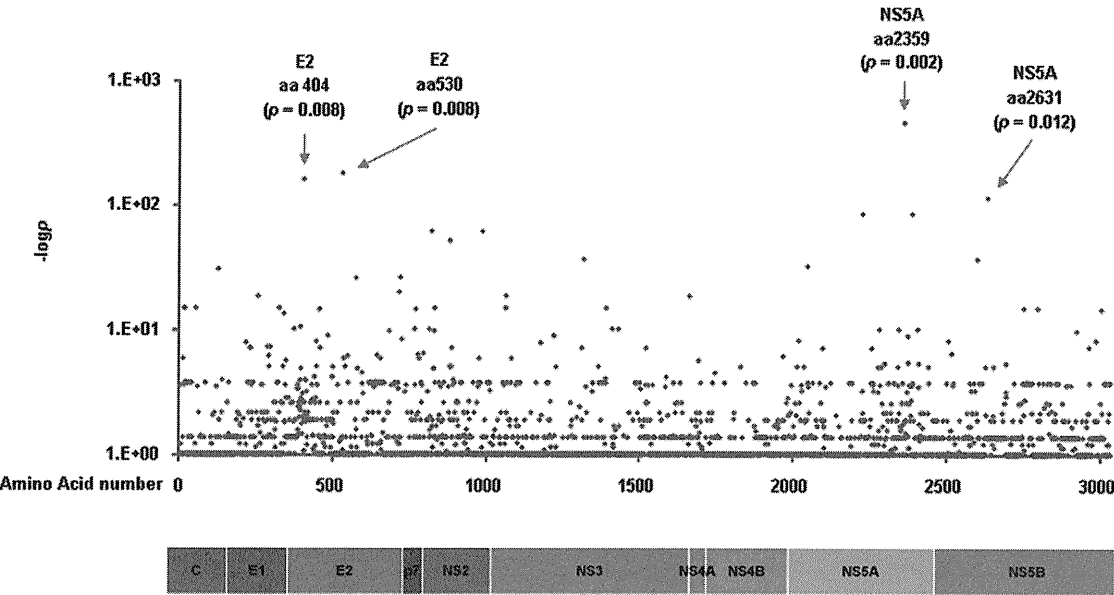


Figure 3. Different amino acid usage at each viral amino acid position between the sustained virological responders (SVR) and the non-sustained virological responders (non-SVR) patients. (a) Different amino acid usage at each viral amino acid position between the SVR and the non-SVR patients was analyzed by Fisher's exact probability test. The longitudinal axis shows the $-\log P$ value. (b) Sequence alignment in the Core region is demonstrated. Dashes indicate amino acids identical to the consensus sequence and substituted amino acids are shown by standard single letter codes.
doi:10.1371/journal.pone.0024514.g003

Table 2. Variation at each Amino Acid Position and SVR rate.

	E2 aa 404 non T	E2 aa 530 non T	NS5A aa 2359 N	NS5B aa 2631 non P
SVR rate	86.1% (31*/36**, p=0.008)	87.9% (29/33, p=0.008)	82% (41/50, p=0.002)	94.7% (18/19, p=0.012)

*SVR number in patients fulfilling the criteria.
**Number of patients fulfilling the criteria.
doi:10.1371/journal.pone.0024514.t002

hepatocellular carcinoma, (4) no other form of hepatitis, such as primary biliary cirrhosis, autoimmune liver disease, or alcoholic liver disease, (5) free of co-infection with human immunodeficiency virus. To clearly disclose the non-SVR viral characteristics, we have considered only those patients who achieved total drug administration of 60% or more for both PEG-IFN and RBV, with the completion of the standard treatment duration. Moreover, although we excluded patients with extended therapy to make the studied population uniform, we have included non-SVR patients with extended therapy to clarify the specific characteristics of non-SVR patients, a minor population group. As a result, 17 patients were excluded for the following reasons: 1 patient received insufficient dose, 4 patients were discontinued from the therapy within 12 weeks, and 12 SVR patients received extended therapy. Finally, 60 patients were considered as eligible for the study. During the combination therapy, blood samples were obtained at least once every month before, during and after treatment and were analyzed for blood count, ALT and HCV RNA levels. Liver biopsy specimens were obtained from most of the patients. All patients gave written informed consent to the study. The study was approved by the ethics committees of University of Yamanashi,

Tokyo Medical and Dental University, and related institutions. The therapy was performed according to the standard treatment protocol of PEG-IFN/RBV therapy for Japanese patients established by a hepatitis study group of the Ministry of Health, Labour, and Welfare, Japan (PEG-IFNα-2b 1.5 µg/kg body weight, once weekly subcutaneously, and RBV 600–800 mg daily per os for 24 weeks).

Complete HCV-ORF Sequence Determination by Direct Sequencing from Pretreatment Sera

HCV RNA was extracted from pretreatment serum samples by the AGPC method using Isogen (Wako, Osaka, Japan) according to the following protocol. Briefly, 150 µl of serum were mixed with 700 µl of Isogen, and an aqueous phase was extracted with 150 µl of chloroform. RNA was precipitated with 600 µl of isopropanol and with 2 µl of Glyco Blue (Ambion, Tokyo, Japan) as a carrier. The purified RNA was washed once with ethanol and finally dissolved in 15 µl of distilled water and stored at –70°C until use. Complementary DNA was synthesized according to the following protocol. 30 µl of the reverse transcription mixture were adjusted to contain 3 µl of the RNA solution, 300 U of Superscript

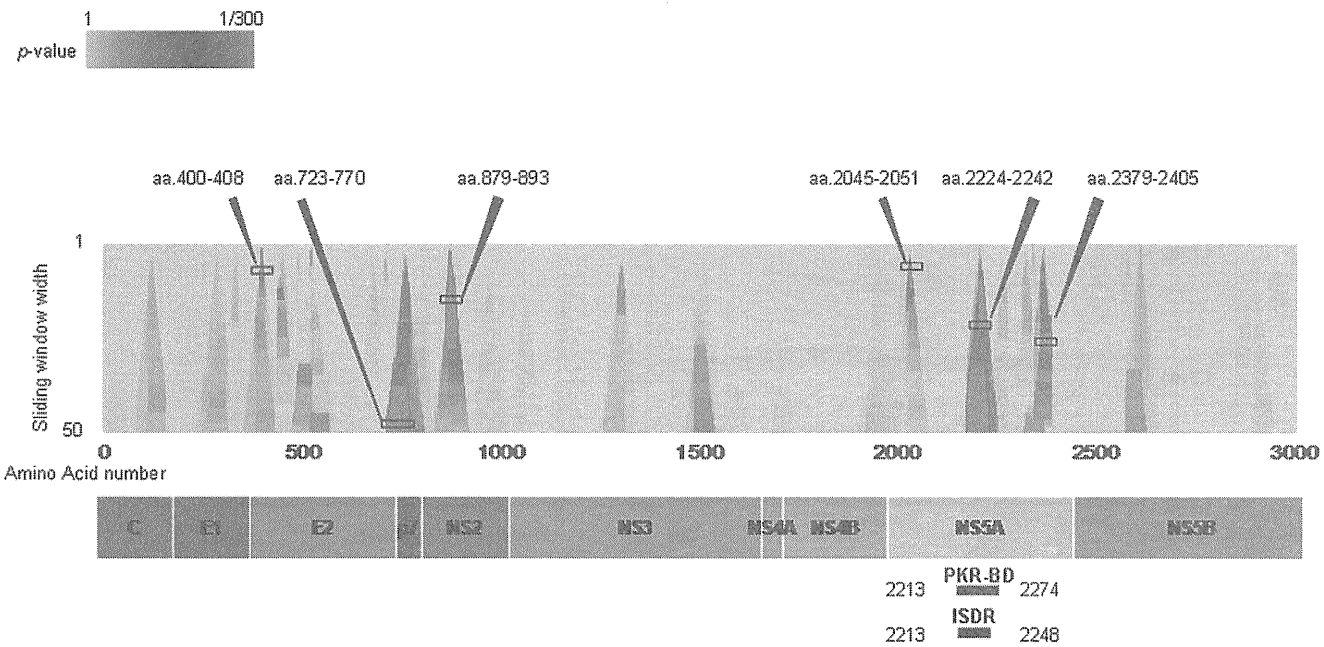


Figure 4. Sliding window analysis. (a) Comparison of amino acid variation between the SVR and non-SVR patients across HCV “regions” using sliding window analysis was performed. Viral regions affecting treatment outcome are shown as red areas. There are six hot areas: amino acid 400–408 and 723–770 in the E2 region, amino acid 879–893 in the NS2 region and, amino acid 2045–2051, 2224–2242 and 2379–2405 in the NS5A region. (b) Sequence alignment in the nonstructural (NS)5A around amino acids 2213 to 2274 is demonstrated. Dashes indicate amino acids identical to the consensus sequence and substituted amino acids are shown by standard single letter codes.
doi:10.1371/journal.pone.0024514.g004

Table 3. Number of Amino Acid Substitutions in each Region and SVR rate.

	E2 aa 400–408 mutation ≥4	E2 aa 723–770 mutation ≥2	NS2 aa 879–893 mutation ≥2	NS5A aa 2045–2051 absence of mutation	NS5A ISDR (aa 2213–2248) mutation ≥1	NS5A aa 2224–2242 mutation ≥1	NS5A aa 2379–2405 mutation ≥2
SVR rate	86.5% (32*/37**) p = 0.006	100% (18/18) p = 0.001	94.7% (18/19) p = 0.01	89.7% (35/39) p = 0.0002	86.1% (31/36) p = 0.008	90.9% (30/33) p = 0.001	90.9% (20/22) p = 0.03

*SVR number in patients fulfilling the criteria.
**Number of patients fulfilling the criteria.
doi:10.1371/journal.pone.0024514.t003

II (Invitrogen, Tokyo, Japan) with an accompanied buffer according to the manufacturer's instructions, 60 units of RNase inhibitor (Promega Corp., Madison, WI), and 300 pg of random primers (Invitrogen). The mixture was incubated at 37°C for 30 min. The HCV genome was amplified with 24 partially overlapping primer (Table S6) sets, designed specifically for this study, to perform two-step nested PCR. As previously reported, a M13 forward primer (5'-TGTAACGACGGCCAGT-3') and a M13 reverse primer (5'-CAGGAAACAGCTATGACC-3') were attached to the 5' termini of the sense and antisense second-round PCR primers, respectively, to facilitate direct sequencing. All samples were initially denatured at 95°C for 7 min., followed by 40 cycles with denaturation at 95°C for 15 seconds, annealing at 55°C for 15 seconds, and extension at 72°C for 45 seconds with BD Advantage™ 2 PCR Enzyme System (BD Biosciences Clontech, CA, USA). PCR amplicons were sequenced directly by Big Dye Terminator Version 3.1 (ABI, Tokyo, Japan) with universal M13 forward/M13 reverse primers using an ABI prism 3130 sequencer (ABI). The sequence files generated were assembled using Vector NTI software (Invitrogen) and base-calling errors were corrected following visual inspection of the chromatogram. When several peaks were observed at the same nucleotide position in the chromatogram, the highest chromatogram peak was read as the dominant nucleotide. In sequence analysis, multiple sequence alignment was performed with ClustalW, and the mean genetic distance was calculated using the p-distance algorithm in the MEGA version 4 DNA software. As a result, 60 genotype-2b HCV full open reading frame sequences were determined. In Table S1, obtained GenBank accession numbers for these sequences determined in this study are listed.

Table 4. Multivariate Logistic Regression Analysis.

Factor	odds (95% CI)	p value
Age	0.94 (0.85–1.04)	0.20
E2 aa 530 non T	4.33 (0.48–39.3)	0.19
NS5A aa 2359 N	3.22 (0.18–57.7)	0.43
NS5B 2631 non P	5.14 (0.29–91.2)	0.26
NS2 aa 879–893 mutations ≥2	9.77 (0.52–182)	0.13
NS5A aa 2045–2051 no mutations	4.46 (0.39–50.6)	0.23
NS5A aa 2224–2242 mutations ≥1	11.0 (1.13–107)	0.04
NS5A aa 2379–2405 mutations ≥1	7.03 (0.62–79.8)	0.12

To evaluate the optimal threshold of amino acid variations for SVR prediction in each viral region extracted, a receiver operating characteristic curve was constructed and the most optimal cut off value was determined for each region.
doi:10.1371/journal.pone.0024514.t004

Sliding Window Analysis

A sliding window analysis was introduced to search through HCV amino acid “regions”, rather than single amino acid positions, related to the final outcome of PEG-IFN/RBV therapy. Briefly, the total number of amino acid substitutions compared to the consensus sequence within a given amino acid length were counted at each amino acid position in each HCV sequence. The consensus sequence was generated from these 60 patients. Then the relation of substitution numbers and the final outcome was compared statistically between the SVR and non-SVR groups by Mann-Whitney’s U test for each amino acid position. In this study, we changed the window length from 1 to 50 to search for those HCV regions. To visualize the result, significantly lower p-values were colored in red and non-significant p-values were colored in green using Microsoft Excel software to generate a “heat map” appearance. In the present study, p-value of 1/300 or lower was colored in the maximum red.

Statistical Analysis

Statistical differences in the parameters, including all available patients’ demographic, biochemical, hematological, and virological data such as sequence variation factors, were determined between the various groups by Mann-Whitney’s U test for numerical variables and Fisher’s exact probability test for categorical variables. To evaluate the optimal threshold of variations for SVR prediction, a receiver operating characteristic curve was constructed and the area under the curve as well as the sensitivity and specificity were calculated. Variables that achieved statistical significance ($p < 0.05$) in univariate analysis were entered into multiple logistic regression analysis to identify significant independent factors. We also calculated the odds ratios and 95% confidence intervals. All p values of < 0.05 by the two-tailed test were considered significant.

Results

Characteristics of the patients studied

The SVR rate of the patients analyzed was 75.9% (44/58) with the standard therapy (two non-SVR patients received extended therapy). The baseline characteristics of the patients classified according to achievement of SVR are shown in Table 1. Rapid virological response (RVR; undetectable serum HCV RNA within 4 weeks) and early virological response (EVR; undetectable serum HCV RNA within 12 weeks) rates were significantly higher in SVR patients ($p = 0.0008$ and 0.004). In addition, patients with non-SVR were older ($p = 0.04$). Pretreatment HCV RNA titer, which is known to affect the treatment outcome in genotype 1 and 2a HCV infection, did not differ significantly between two groups. Achievement of RVR reached 42.4% when all patients were included, and this rate was high compared to achievement of RVR in patients with genotype 1b infection (~10%) observed in

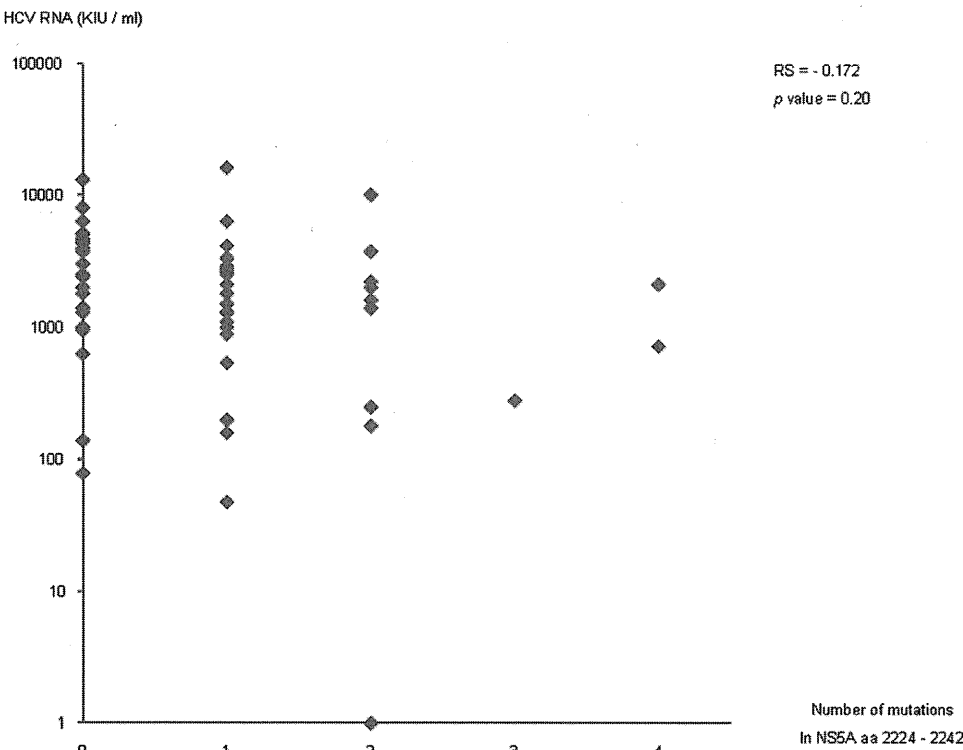


Figure 5. Correlation between pretreatment HCV RNA levels and the number of substitutions in the NS5A region aa 2224 to 2242. Spearman's correlation coefficient by rank test is demonstrated. doi:10.1371/journal.pone.0024514.g005

University of Yamanashi (data not shown). The early virological response (EVR) rate was equally high in the SVR (97.7%) and non-SVR (68.8%) groups. Interestingly, most of the non-SVR patients (14/16, 87.5%) in genotype-2b HCV infection showed end-of-treatment response (ETR; undetectable serum HCV RNA at the end of therapy), demonstrating that the main cause of non-SVR was relapse (reappearance of hepatitis C viremia during the follow-up period after stopping therapy in patients with an ETR,

n = 14), and not null response (detectable serum HCV RNA at the end of therapy, n = 2).

Phylogenetic analysis of SVR and non-SVR patients using the complete HCV amino acid sequence

To determine the viral sequence characteristics in the SVR and non-SVR groups, we first aligned all 60 HCV complete ORF amino acid sequences obtained from the patients' pretreatment sera along

Table 5. Baseline Characteristics of patients with NS5A aa 2224–2242 variations none or 1≤.

Characteristic	Variation 1≤ (n = 33)	No variation (n = 27)	P value
Gender (Male/Female)	17/16	18/9	NS [†]
Age (yrs)	57 (29–72) [‡]	57 (22–80)	NS [‡]
ALT (IU/l)	72 (19–380)	47 (17–390)	NS [‡]
Platelet (×10 ⁶ /mm ³)	19.3 (7.1–31.8)	17.5 (10.4–36.7)	NS [‡]
Fibrosis score (0–2/≥3) [§]	26/5	19/3	NS [†]
HCV RNA (KIU/ml)	1600 (100–16000)	2450 (140–13000)	NS [‡]
IFN dose (≥80%/60–80%)	26/7	23/4	NS [†]
Ribavirin dose (≥80%/60–80%)	24/9	19/8	NS [†]
RVR rate (%)	53.1	29.6	NS [†]
EVR rate (%)	96.9	81.5	NS [†]
SVR rate (%)	90.9	51.9	0.001 [†]
Replapse rate (%)	40.7	9.1	0.006 [†]

§: 1≤ : n = 31, 0 : n = 22.

‡: median (range).

†: Fisher's exact probability test.

‡: Mann-Whitney's U test.

doi:10.1371/journal.pone.0024514.t005

with reference sequences (2b.HC-J8.D10988, 2JP.MD2b9-2, and 2a.JP.JFH-1.AB047639 obtained from the Los Alamos HCV Database as representative sequences for genotype 2b and genotype 2a HCV) and constructed a phylogenetic tree (Fig. 1). As demonstrated in the tree, no evident clustering was apparent according to the difference of responses.

Comparison of amino acid variation between the SVR and non-SVR in the complete HCV polyprotein and each HCV protein

Next, we compared amino acid variations that were unique, relative to a population consensus, to either the SVR or non-SVR patients for the complete HCV polyprotein and each HCV protein. The number of amino acid variations in the sequences from the SVR patients was significantly higher than in those from the non-SVR patients, when the entire HCV polyprotein was analyzed (Fig. 2, left). These differences were especially significant in E1, p7 and NS5A (Fig. 2, right). This result demonstrated that HCV sequences from patients with SVR comprised a heterogeneous population, while HCV sequences from patients with non-SVR comprised a rather homogeneous population, indicating the existence of unique non-responsive HCV sequences in those regions in E1, p7, and NS5A.

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

Each amino acid position in the HCV ORF was compared to detect any differences between the SVR and non-SVR patients. In Fig. 3a, differences in amino acid residues at each position are shown as dots demonstrating $-\log P$ values. As shown in Table 2, four points were extracted: amino acid (aa) 404 in the E2 region ($p = 0.008$), aa 530 in the E2 region ($p = 0.008$), aa 2359 in the NS5A region ($p = 0.002$) and aa 2631 in the NS5B region ($p = 0.012$). Among them, the residue at aa 2359 in the NS5A region differed most frequently between the SVR and non-SVR patients. Amino acids 4 and 110 in the Core region, residues that have been reported to vary according to the virological responses in genotype 2a infection [22,23], did not differ significantly in this genotype 2b HCV study. Meanwhile, amino acids 70 and 91, which have been reported to vary according to virological response to PEG-IFN/RBV therapy in genotype 1b infection, were conserved irrespective of the outcome (Fig. 3b).

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

Fig. 4a and Table 3 shows the result of sliding window analysis. 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, six regions were associated with the final outcome (p -values less than 1/20): aa 400–408 in the E2 region ($p = 0.006$), aa 723–770 in the E2 and the N-terminus of p7 region ($p = 0.001$), aa 879–893 in the NS2 region ($p = 0.01$), aa 2045–2051 in the NS5A region ($p = 0.0002$), aa 2224–2242 in the NS5A region ($p = 0.001$) and aa 2379–2405 in the NS5A region ($p = 0.03$). Interestingly, aa 2224–2242 in the NS5A was located in the interferon sensitivity determining region (ISDR). Fig. 4b shows the aligned sequences of amino acids around 2213–2274 of HCV NS5A. Among these 6 regions, aa 723–770, aa 879–893, aa 2224–2242, and aa 2379–2405 were correlated with the final outcome in an incremental manner according to the number of substitutions in those regions (Table S2, S3, S4, S5). The number of substitutions in the ISDR

was also correlated to the final outcome in an incremental step-up manner (data not shown).

Multivariate analysis to detect independent predictive factors contributing to the SVR

Next, multivariate analysis was undertaken to identify pretreatment variables correlated with the final outcome. To evaluate the optimal threshold of amino acid variations for SVR prediction in each viral region extracted, a receiver operating characteristic curve was constructed and the most optimal cut off value was determined for each region. E2 aa404–408 was excluded from the analysis because we considered that the region was unlikely to be truly associated to the outcome as it is located in the hypervariable region, the region of the highest mutation rate in the HCV genome as a result of host's immune attack. E2 aa 723–770 was excluded from the analysis because all the patients above the cut-off value in the region achieved SVR and an odds calculation was not possible. The ISDR was also excluded because NS5A aa2224–2242 was completely contained in the ISDR. In addition, variables of EVR and RVR were excluded because they were post treatment variables. The multivariate analysis revealed that only NS5A aa 2224–2242 (odds ratio 11.0, $p = 0.039$) was finally identified as the independent variable predicting the final outcome (Table 4).

Biological relevance of variation in NS5A in this study group

Because NS5A aa 2224–2242 is located within the ISDR, for which the amino acid substitution numbers have been reported to be correlated with the HCV RNA titer in genotype 1 and 2a HCV infection [13], we analyzed the relationship between amino acid variations in that region and pretreatment HCV RNA titers. Contrary to our expectation, no evident relationship was found between variations in the NS5A region aa 2224–2242 and HCV RNA titer (Fig. 5). On the other hand, as shown in Table 5, although the initial viral responses (RVR or EVR) did not show evident association with the amino acid variations in the region, treatment relapse was significantly correlated with the amino acid variations in the region. In addition to NS5A aa 2224–2242, there was no evident relationship between HCV RNA level and variations in the other regions found in this study (data not shown).

Discussion

In this study, we showed that genotype 2b HCV sequences from Japanese patients who achieved SVR were more diverse than the sequences from patients with non-SVR. The result that SVR patients were more diverse in their HCV sequences than non-SVR patients is in accordance with previous studies of genotype 1 HCV infection, although the diverse viral genes varied according to genotype [18,19]. We found that these diversities were primarily found in E1, p7 and NS5A.

In systemic searching for single amino acid positions or consecutive amino acid regions in the HCV ORF associated with the treatment outcome, several regions were extracted in E2, p7, NS2, NS5A and NS5B. Among those identified regions, E2 aa 723–770, NS2 aa 879–893, NS5A aa2224–2242, and NS5A aa2379–2405 were correlated with the final outcome in an incremental manner according to the number of amino acid substitutions. Specifically, the sequences of those regions in non-SVR patients were almost homogeneous, while the sequences of the region in SVR patients were significantly diverse and multiple amino acid substitutions were found compared to the consensus sequence. Interestingly, among those regions, aa 2224–2242 was completely included in the ISDR, in which the number of amino acid substitutions is known to show significant correlation with