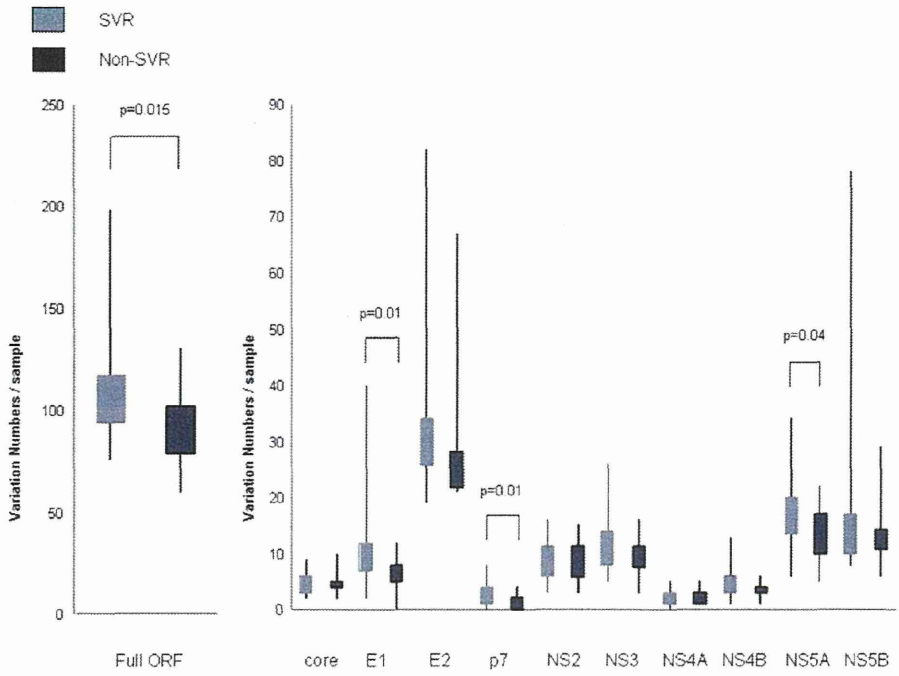

Figure 2. Number of amino acid substitutions per sample in the sustained virological responders (SVR) and the non-sustained virological responders (non-SVR) group. The numbers of variations, relative to a population consensus, that were unique to either SVR or non-SVR patients are shown for the complete 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

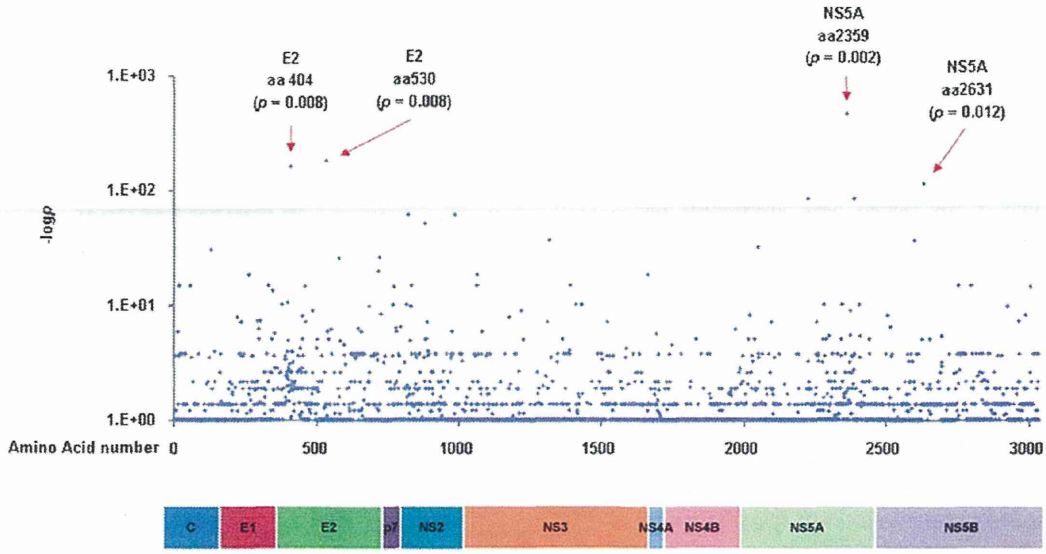

Figure 3. Different amino acid usage at each viral amino acid position between the sustained virological responders (SVR) and the non-sustained virological responders (non-SVR) patients. (a) Different amino acid usage at each viral amino acid position between the SVR and the non-SVR patients was analyzed by Fisher's exact probability test. The longitudinal axis shows the −logP 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

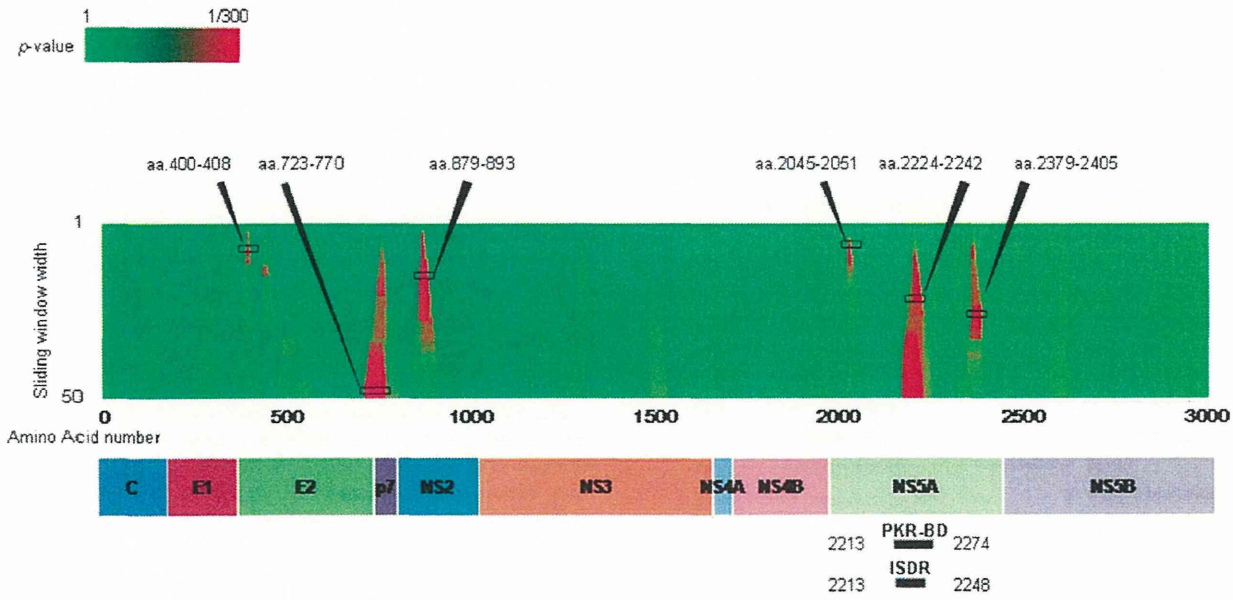

Figure 4. Sliding window analysis. (a) Comparison of amino acid variation between the SVR and non-SVR patients across HCV "regions" using sliding window analysis was performed. Viral regions affecting treatment outcome are shown as red areas. There are six hot areas: amino acid 400–408 and 723–770 in 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 absense 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′-TGTAAAACGACGGCCAGT-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

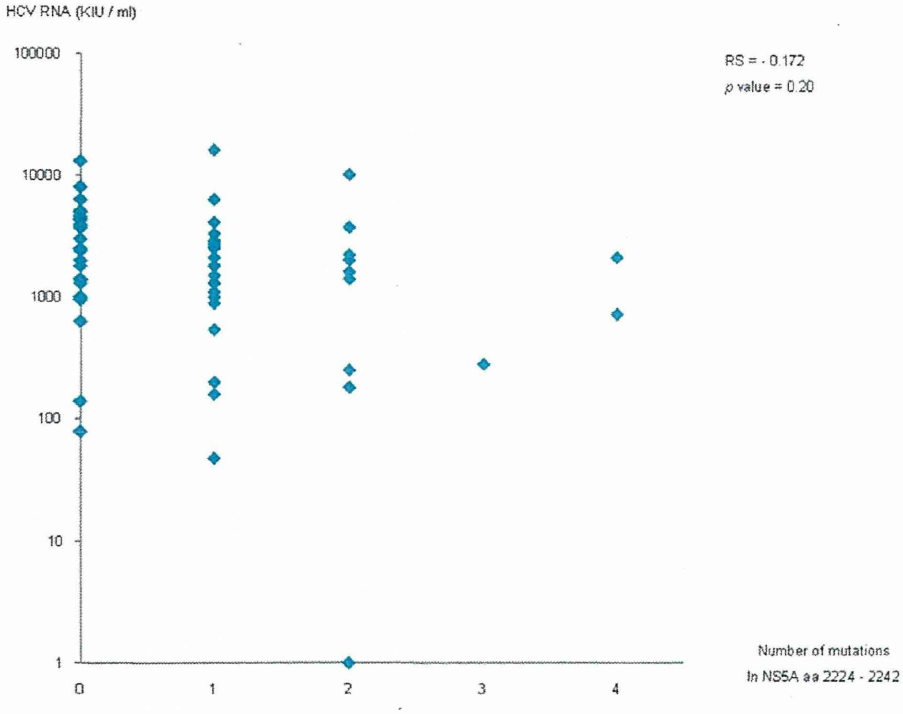

Figure 5. Correlation between pretreatment HCV RNA levels and the number of substitutions in the NS5A region aa 2224 to 2242. Spearman's correlation coefficient by rank test is demonstrated.
doi:10.1371/journal.pone.0024514.g005

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 ($\times 10^4$/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, 2.JP.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 resides at each position are shown as dots demonstrating −logP 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

the treatment response to IFN-based therapy in genotype 1b, and also in genotype 2 [21,24].

In recent studies of genotype 1b infection, amino acid variation of residues 70 and 91 in the Core were reported to be associated with the treatment response to IFN-based therapy. The correlation of amino acid variation in the Core (residues 4 and 110) with the response to PEG-IFN/RBV therapy was also identified in genotype 2a infection [22,23]. In genotype 2b infection, however, we could not find such associations between amino acid variation in the core region and the response to PEG-IFN/RBV therapy (Fig. 3b). Amino acid residues of aa 70 and 91 were conserved irrespective of differences in the PEG-IFN/RBV responses. On the other hand, although amino acid variations were also sometimes found at residues 4 and 110 in genotype 2b HCV, their frequency was low, and no evident association between the variation and the treatment response was found. Although the reason of the lack of association between the Core and the PEG-IFN/RBV treatment response in genotype-2b HCV infection is unknown, it suggests that a different mechanism affecting the treatment response might exist, depending on genotype-specific viral features.

In genotype 1 HCV, variations within the PKR-binding region of NS5A, including those within the ISDR, were reported to disrupt the NS5A-PKR interaction, possibly rendering HCV sensitive to the antiviral effects of interferon [25]. Clinically, the number of substitutions within the ISDR has been reported to correlate with the serum HCV RNA level in genotype 1 and 2a infections [13]. In addition, a recent study reported that mutations in the ISDR also show the correlation with the relapse in the PEG-IFN/RBV therapy in genotype 1b infection [26]. Because NS5A aa2224–2242, part of ISDR, was extracted as one of those regions related to the treatment response in genotype 2b infection, we undertook further analysis to investigate the correlation between amino acid variation numbers and serum HCV RNA level. Though the reason is unknown, we could not find evidence of a relationship between variation in the NS5A aa 2224–2242 and HCV RNA titer in genotype 2b infection, unlike genotypes 1 and 2a. Of note, a high SVR rate in genotype1 and genotype 2a infection is known to be closely correlated with a low HCV RNA level and multiple substitutions in ISDR. However, in genotype 2b infection in our study, there was no significant difference in the HCV RNA level between SVR and non-SVR patients, as shown in Table 1. Previously, the role of the ISDR in the contribution to SVR in genotype 1 and 2a has been discussed in detail in the context of serum HCV RNA level, and multiple substitutions in the ISDR are related to a low HCV RNA level and high SVR rate. However, it is not known which of these two factors is directly associated with viral clearance. Consideration of this three-sided relationship of ISDR, HCV RNA level and SVR rate in genotype-2b infection leads to the suggestion that amino acid variation in ISDR to be more direct contributor for SVR.

In spite of these findings, there were still limitations in our study. First, because genotype 2b infection only accounts for 10% of all HCV infection in Japan, the number of studied patients was rather small, especially non-SVR patients. In addition, because genotype 2b HCV contains as many as 3033 amino acids, it is possible that incorrect amino acids or regions were judged as significant in the complete HCV ORF comparison study as a result of type I errors. Therefore, if more patients were available for the analysis, the statistical power detecting the meaningful differences would be greater. Secondly, we could not include the IL28B SNP analysis in this study. If we could have combined the information of IL28B SNPs with the full HCV ORF information, a more comprehensive analysis would have been achieved.

In conclusion, we have shown that viral sequences were more diverse in SVR patients infected with genotype 2b HCV. Through systematic comparison between SVR and non-SVR patients, we have also shown that several localized regions were extracted as hot spots whose amino acid substitutions were closely related to the final outcome by affecting the relapse rate in the PEG-IFN/RBV therapy.

Supporting Information

Table S1 GenBank Accession Numbers. Obtained GenBank accession numbers for 60 genotype-2b HCV full open reading frame sequences are listed.
(DOC)

Table S2 Substitutions in NS5A aa 2224–2242 Amino Acid Regions and SVR rate. SVR rate increased with the number of substitutions in this region.
(DOC)

Table S3 Substitutions in NS5A aa 2379–2405 Amino Acid Regions and SVR rate. SVR rate increased with the number of substitutions in this region.
(DOC)

Table S4 Substitutions in NS2 aa 879–893 Amino Acid Regions and SVR rate. SVR rate increased with the number of substitutions in this region.
(DOC)

Table S5 Substitutions in E2 aa 723–770 Amino Acid Regions and SVR rate. SVR rate increased with the number of substitutions in this region.
(DOC)

Table S6 PCR Primer List. Primers designed to perform two-step nested PCR for this study are listed. Dominant genotype-2b HCV full open reading frame sequences was determined by the 24 partially overlapping amplicons amplified by these primers.
(XLS)

Acknowledgments

We are grateful to Yamanashi Pegintron Ribavirin Study Group and Ochanomizu Liver Conference Group for their cooperation and advice. Especially, we thank Asuka Kanayama and Takako Ohmori for their technical assistance, and Takatoshi Kitamura and Shunichi Okada for their cooperation and advice.

Author Contributions

Conceived and designed the experiments: MK SM NE. Performed the experiments: MK. Analyzed the data: MK SM NE. Contributed reagents/materials/analysis tools: RS MM HS KK. Wrote the paper: MK SM NE. Critical revision of the manuscript for important intellectual content: FA TU TI MS MN NS MW.

References

1. Ghany MG, Strader DB, Thomas DL, Seeff LB (2009) Diagnosis, management, and treatment of hepatitis C: an update. Hepatology 49: 1335–1374.
2. Di Bisceglie AM, Martin P, Kassianides C, Lisker-Melman M, Murray L, et al. (1989) Recombinant interferon alfa therapy for chronic hepatitis C. A randomized, double-blind, placebo-controlled trial. N Engl J Med 321: 1506–1510.
3. Haydon GH, Jarvis LM, Blair CS, Simmonds P, Harrison DJ, et al. (1998) Clinical significance of intrahepatic hepatitis C virus levels in patients with chronic HCV infection. Gut 42: 570–575.
4. Simmonds P (1997) Clinical relevance of hepatitis C virus genotypes. Gut 40: 291–293.

5. Hadziyannis SJ, Sette H, Jr., Morgan TR, Balan V, Diago M, et al. (2004) Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. Ann Intern Med 140: 346–355.
6. Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, et al. (2001) Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet 358: 958–965.
7. Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, et al. (2009) Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 461: 399–401.
8. Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, et al. (2009) Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. Nat Genet 41: 1105–1109.
9. Rauch A, Kutalik Z, Descombes P, Cai T, di Iulio J, et al. (2010) Genetic variation in IL28B Is Associated with Chronic Hepatitis C and Treatment Failure - A Genome-Wide Association Study. Gastroenterology.
10. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, et al. (1996) Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. N Engl J Med 334: 77–81.
11. El-Shamy A, Nagano-Fujii M, Sasase N, Imoto S, Kim SR, et al. (2008) Sequence variation in hepatitis C virus nonstructural protein 5A predicts clinical outcome of pegylated interferon/ribavirin combination therapy. Hepatology 48: 38–47.
12. Hamano K, Sakamoto N, Enomoto N, Izumi N, Asahina Y, et al. (2005) 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 20: 1401–1409.
13. Chayama K, Suzuki F, Tsubota A, Kobayashi M, Arase Y, et al. (2000) Association of amino acid sequence in the PKR-eIF2 phosphorylation homology domain and response to interferon therapy. Hepatology 32: 1138–1144.
14. Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, et al. (2005) Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. Intervirology 48: 372–380.
15. Duverlie G, Khorsi H, Castelain S, Jaillon O, Izopet J, et al. (1998) Sequence analysis of the NS5A protein of European hepatitis C virus 1b isolates and relation to interferon sensitivity. J Gen Virol 79(Pt 6): 1373–1381.
16. Hofgartner WT, Polyak SJ, Sullivan DG, Carithers RL, Jr., Gretch DR (1997) Mutations in the NS5A gene of hepatitis C virus in North American patients infected with HCV genotype 1a or 1b. J Med Virol 53: 118–126.
17. Zeuzem S, Lee JH, Roth WK (1997) Mutations in the nonstructural 5A gene of European hepatitis C virus isolates and response to interferon alfa. Hepatology 25: 740–744.
18. Donlin MJ, Cannon NA, Aurora R, Li J, Wahed AS, et al. (2010) Contribution of genome-wide HCV genetic differences to outcome of interferon-based therapy in Caucasian American and African American patients. PLoS One 5: e9032.
19. Donlin MJ, Cannon NA, Yao E, Li J, Wahed A, et al. (2007) Pretreatment sequence diversity differences in the full-length hepatitis C virus open reading frame correlate with early response to therapy. J Virol 81: 8211–8224.
20. Kobayashi M, Watanabe K, Ishigami M, Murase K, Ito H, et al. (2002) 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 97: 988–998.
21. Murakami T, Enomoto N, Kurosaki M, Izumi N, Marumo F, et al. (1999) Mutations in nonstructural protein 5A gene and response to interferon in hepatitis C virus genotype 2 infection. Hepatology 30: 1045–1053.
22. Kadokura M, Maekawa S, Sueki R, Miura M, Komase K, et al. (2011) 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. Hepatol Int.In Press.
23. Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Yatsuji H, et al. (2009) 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 52: 301–309.
24. Hayashi K, Katano Y, Honda T, Ishigami M, Itoh A, et al. (2009) 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 81: 459–466.
25. Gale M, Jr., Blakely CM, Kwieciszewski B, Tan SL, Dossett M, et al. (1998) Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. Mol Cell Biol 18: 5208–5218.
26. Kurosaki M, Tanaka Y, Nishida N, Sakamoto N, Enomoto N, et al. (2010) Pre-treatment prediction of response to pegylated-interferon plus ribavirin for chronic hepatitis C using genetic polymorphism in IL28B and viral factors. J Hepatol.

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

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-IFNα-2b 1.5 μg/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 Advantage^{TM} 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 ≥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 (~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.