

1  
2 extended therapy. Finally, 78 patients were considered as eligible for the study. During  
3  
4 the combination therapy, blood samples were obtained at least once every month before,  
5  
6 during and after treatment and were analyzed for blood count, ALT and HCV RNA  
7  
8 levels. Liver biopsy specimens were obtained from most of the patients.  
9  
10

11  
12 The 78 patients belonging to the different institutions were separately analyzed:  
13  
14 43 patients registered in Y-PERS (Yamanashi Pegintron Ribavirin Study Group) were  
15  
16 included in group 1 (test group), and the 35 patients from Tokyo Medical and Dental  
17  
18 University and related institutions (Ochanomizu Liver Conference Group) were  
19  
20 included in group 2 (validation group). We divided the patients into these two groups in  
21  
22 order to exclude the false positives (type I errors) which might arise in successive  
23  
24 HCV-ORF study. Since genotype-2a HCV contains as many as 3033 amino acids, it was  
25  
26 possible that incorrect amino acids to be judged as significant in full HCV-ORF  
27  
28 comparison study as a result of type I errors. Therefore, to guard against false positives,  
29  
30 HCV-ORF comparison study was undertaken in group 1, group 2, and combined group.  
31  
32  
33  
34  
35  
36  
37  
38  
39

#### 40 Complete HCV-ORF Sequence Determination by Direct Sequencing from 41 42 Pretreatment Sera 43

44  
45 HCV RNA was extracted from pretreatment serum samples by the AGPC  
46  
47 method using Isogen (Wako, Osaka, Japan) according to the manufacturer's protocol.  
48  
49 Complementary DNA was synthesized with Superscript II (Invitrogen, Tokyo, Japan)  
50  
51 using random primers (Invitrogen) and then amplified by two-step nested PCR using the  
52  
53 primers newly designed for this study. All samples were initially denatured at 95°C for 7  
54  
55 min., followed by 40 cycles with denaturation at 95°C for 15 seconds, annealing at 55°C  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2 for 15 seconds, and extension at 72°C for 45 seconds with BD Advantage™ 2 PCR  
3  
4 Enzyme System (BD Biosciences Clontech, CA, USA).  
5  
6

7 PCR amplicons were sequenced directly by Big Dye Terminator Version 3.1  
8 (ABI, Tokyo, Japan) with universal M13 forward / M13 reverse primers using an ABI  
9  
10 prism 3130 sequencer (ABI). Generated sequence files were assembled using Vector  
11  
12 NTI software (Invitrogen) and base-calling errors were corrected following inspection  
13  
14 of the chromatogram.  
15  
16  
17  
18  
19  
20

### 21 **Sliding Window Analysis**

22 A sliding window analysis was introduced to search through HCV amino acid  
23  
24 “regions”, rather than single amino acid positions, related to the final outcome of  
25  
26 PEG-IFN/RBV therapy. Briefly, the total number of amino acid substitutions compared  
27  
28 to the consensus sequence within a given amino acid length were counted in each amino  
29  
30 acid position in each HCV sequence. Then the relation of substitution numbers and the  
31  
32 final outcome was compared statistically between the SVR and non-SVR groups by  
33  
34 Mann-Whitney's U test for each amino acid position. In this study, we changed the  
35  
36 window length from 1 to 50 to search for those HCV regions. To visualize the result,  
37  
38 significantly lower p-values were colored in red and non-significant p-values were  
39  
40 colored in green to generate a “heat map” appearance using Microsoft Excel software.  
41  
42 In the present study, p-value of 1/1000 or lower was colored in the maximum red.  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53

### 54 **Statistical Analysis**

55  
56  
57  
58 Statistical differences in the parameters, including all available patients'  
59  
60  
61  
62  
63  
64  
65

1 demographic, biochemical, hematological, and virological data such as sequence  
2 variation factors, was determined between the various groups by Student t test or  
3  
4  
5  
6  
7 Mann-Whitney's U test for numerical variables and Fisher's exact probability test for  
8  
9  
10 categorical variables. To evaluate the optimal threshold of variations for SVR  
11  
12 prediction, the receiver operating characteristic curve was constructed. Variables that  
13  
14 achieved statistical significance ( $p < 0.05$ ) in univariate analysis were entered into  
15  
16 multiple logistic regression analysis to identify significant independent factors. We also  
17  
18 calculated the odds ratios and 95% confidence intervals. All p values of  $< 0.05$  by the  
19  
20 two-tailed test were considered significant.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## 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 ribavirin 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 relapser, while two patients were null responder.

### 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

1  
2 sequences from the SVR patients was significantly higher than in those from the  
3  
4 non-SVR patients, when the entire HCV polyprotein was analyzed (Fig.1, left). These  
5  
6 differences were especially significant in E1 and NS3 (Fig.1, right). This result  
7  
8 demonstrated that HCV sequences from patients with SVR comprised a heterogeneous  
9  
10 population, while HCV sequences from patients with non-SVR comprised a rather  
11  
12 homogeneous population, indicating the existence of unique non-responsive HCV  
13  
14 sequences.  
15  
16  
17  
18  
19  
20

#### 21 **Comparison of HCV sequence variation between the SVR and non-SVR patients at** 22 **each amino acid position** 23 24

25  
26 Next, each amino acid position in the HCV ORF was compared to detect any  
27  
28 differences between the SVR and non-SVR patients after determination of the  
29  
30 consensus sequence from all 78 patients. In Fig.2a, the final differences of the two  
31  
32 independent studies combined are shown as dots demonstrating  $-\log P$  values. As shown  
33  
34 in the figure, amino acid usage at amino acid 110 in the core region differed strikingly  
35  
36 between the two groups ( $p=5E-05$ ). The site was detected in group 1 ( $p=0.01$ ) and was  
37  
38 validated in group 2 ( $p=0.004$ ) (Table 2a), and the final p-value became remarkably  
39  
40 high, making the p-value at this site most significantly low. Variations of aa 773 in p7,  
41  
42 aa 2099 in the NS5A, and aa 3013 in NS5B were also shown to differ significantly  
43  
44 between the SVR and the non-SVR patients when the two studies were combined;  
45  
46 however, they were not confirmed by one of the studies (Table 2a). Fig.2b shows the  
47  
48 aligned sequences of amino acids 1-120 of the core region. Substitutions at aa 110 from  
49  
50 non-T (N/S) to T were significantly more frequent in SVR (32/58, 55.2%) than in  
51  
52 non-SVR (1/20, 3.6%,  $p=5E-05$ ). Amino acid 4, the site reported recently to vary  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2 according to the viral response in genotype 2a infection, did not differ significantly in  
3  
4 our study. Amino acid 70 and 91, which have been reported to vary according to viral  
5  
6 response to PEG-IFN/RBV therapy in genotype 1b infection, were conserved  
7  
8 irrespective of the outcome.  
9  
10

### 11 12 13 14 **Comparison of amino acid variation between the SVR and non-SVR patients** 15 16 **across HCV “regions” using sliding window analysis** 17

18  
19 Fig.2c shows the combined result of sliding window analysis for study groups  
20  
21 1 and 2, this approach was used to detect differing HCV amino acid “regions”, rather  
22  
23 than single amino acid positions, between the SVR and the non-SVR patients.  
24  
25 According to the result, four regions were notably associated with the final outcome  
26  
27 (p-values less than 1/1000). Core aa 110, detected as a single amino acid position  
28  
29 discriminating between the SVR and the non-SVR patients, was also identified as one  
30  
31 of these regions. Because core aa 110 was already known for its strong correlation with  
32  
33 the response as above, the region was excluded from further analysis. Among the other  
34  
35 three regions, only NS5A aa 2258-2306 showed significant differences in the two  
36  
37 independent study groups (Table 2b). Interestingly, the NS5A region overlapped the  
38  
39 PKR-binding domain, which includes the interferon sensitivity determining region  
40  
41 (ISDR). Fig.2d shows the aligned sequences of amino acids around 2258-2306 of HCV  
42  
43 NS5A. As with previous studies, variations in the ISDR were also significantly more  
44  
45 frequent in SVR patients.  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55

### 56 **Multivariate analysis to detect independent factors contributing to the SVR** 57

58  
59 Multivariate analysis revealed that variation of core aa 110, the total number of  
60  
61  
62  
63  
64  
65

1  
2 substitutions within NS5A aa 2258-2306, and total ribavirin dose  $\geq 80\%$  were finally  
3  
4 identified as the independent variables influencing the final outcome (odds ratio 24.7,  
5  
6 11.5 and 16.0;  $p = 0.02, 0.03$  and  $0.02$ , Table 3).  
7  
8  
9

### 10 11 12 13 14 15 **Biological relevance of variation in core and NS5A in this study group**

16  
17 To determine biological relevance of core aa110 and NS5A aa2258-2306, we  
18 investigated their relationship with clinical background factors. Multiple variations in  
19 the NS5A region aa 2258-2306 were significantly related to pretreatment HCV RNA  
20 titer ( $p=9E-05$ , Fig. 3 and Table 4a). Interestingly, variation of the core aa110 was  
21 significantly associated with the patients' age ( $p=0.03$ , Table 4b).  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## 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 affect 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 relapser, but were not null-responders as observed frequently among genotype 1b patients treated with PEG-IFN/RBV therapy. Therefore, we compared the different viral responses according to the final outcome of SVR or non-SVR.

Variation of core aa 110 was identified as the single amino acid residue most significantly related to the final outcome ( $p=5E-05$ ). In recent studies of treatment of genotype 1b infection with PEG-IFN/RBV, amino acid variation in the core region was reported to be associated with response. It is interesting that the core region was also identified as an 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

1  
2 correlation was reported between amino acid variations in the core region and viral  
3  
4 responses of genotype 2a HCV infection (20). *Though the result seems discrepant from*  
5  
6 *our study, we suspect the inconsistent results were at least partially attributable to the*  
7  
8 *different groups used in comparison: we compared the difference between non-SVR*  
9  
10 *patients and SVR patients while they compared the difference between non-SVR and*  
11  
12 *RVR patients.*  
13  
14  
15

16  
17 In systemic searching for the viral “regions” associated with the treatment  
18  
19 outcome, NS5A aa2258-2306 was identified by two independent studies. Interestingly,  
20  
21 the region overlaps the PKR-binding domain (PKR-BD), including the ISDR, in which  
22  
23 the number of amino acid substitutions is known to be related to the response to  
24  
25 IFN-based therapy in genotype 1b, and also in genotype 2a (17-18). Therefore, we also  
26  
27 confirmed that total number of substitutions in the ISDR and PKR-BD is significantly  
28  
29 associated with the final outcome in this group of patients when the two studies were  
30  
31 combined.  
32  
33  
34  
35

36  
37 Some viral regions other than core and NS5A also showed the potential  
38  
39 association with the final outcome. Viral single amino acid substitutions of aa 773 in p7,  
40  
41 aa 2099 in the NS5A, and aa 3013 in NS5B, or viral regions in E1 aa 400-403 and in E2  
42  
43 aa 724-744 were more frequent in SVR. However, because these were not extracted as  
44  
45 significant in one of the two studies when analyzed separately, additional studies are  
46  
47 needed to confirm the association with the final outcome. On the other hand, we could  
48  
49 not find an association with the final outcome and the PePHD or IRRDR, including the  
50  
51 V3 regions (data not shown) reported 1b HCV infection (21-22).  
52  
53  
54

55  
56 It is interesting that the variation of the core region showed the clear  
57  
58 association with age. The younger patients with core aa 110T showing favorable  
59  
60  
61  
62  
63  
64  
65

1  
2 responses while the older patients with core aa 110 non-T showed unfavorable  
3  
4 responses. It is possible that different response rates according to the patients' ages in  
5  
6 genotype 2a infection might have been related to the core substitutions, although further  
7  
8 study is needed. In NS5A, it was reported that the variations within the PKR-binding  
9  
10 region, including those within the ISDR, can disrupt the NS5A-PKR interaction,  
11  
12 possibly rendering HCV sensitive to the antiviral effects of interferon (23). Clinically,  
13  
14 the number of substitutions within the region has been reported to correlate with the  
15  
16 serum HCV RNA level (12). We also confirmed that the number of substitutions within  
17  
18 the NS5A aa 2258-2306 was significantly associated with the pretreatment HCV RNA  
19  
20 titers.  
21  
22  
23  
24  
25

26  
27 Multivariate analysis of the combined group of patients showed that variation  
28  
29 of core aa 110, NS5A aa 2258-2306, and total ribavirin dose  $\geq 80\%$  were independent  
30  
31 variables associated with the final outcome (Table. 3). The association of ribavirin dose  
32  
33 and HCV relapse rate was reported previously (24) and that result was confirmed in this  
34  
35 study. On the other hand, the total PEG-IFN dosage was not identified when it was  
36  
37 administered at greater than 60% of the initially scheduled amount. Indeed, when the  
38  
39 drug dosage is excluded, the strongest association was seen in the viral elements of core  
40  
41 and NS5A, revealing the importance of these two regions in the treatment of genotype  
42  
43 2a HCV infection with PEG-IFN/RBV therapy.  
44  
45  
46  
47  
48

49 On the other hand, our study still has some limitations. In recent studies, IL28B  
50  
51 single nucleotide polymorphisms were reported to be correlated significantly with the  
52  
53 treatment response in genotype 1b HCV infections (25-26). In genotype 2a HCV  
54  
55 infection, a correlation was also reported to exist between the IL28B SNP and the  
56  
57 treatment response (27). However, we could not investigate the association of the IL28B  
58  
59  
60  
61  
62  
63  
64  
65

1  
2 single nucleotide polymorphisms in the treatment response in genotype 2a HCV  
3  
4 infections. In addition, the number of analyzed patients was rather small, especially in  
5  
6 non-SVR patients.  
7

8  
9  
10 In conclusion, by comprehensive investigation of the complete HCV ORF in  
11 patients showing different responses to PEG-IFN/RBV therapy, we have demonstrated  
12 that amino acid variation in the core and NS5A are significantly associated with the  
13 final outcome of treatment of genotype 2a chronic hepatitis C. Considering this result,  
14 determination of those HCV regions before treatment might provide further benefits for  
15  
16 the patients infected with genotype 2a HCV.  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

**ACKNOWLEDGMENT**

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

## REFERENCES

1. Ghany MG, Strader DB, Thomas DL, Seeff LB. Diagnosis, management, and treatment of hepatitis C: an update. *Hepatology* 2009 Apr;49(4):1335-74.
2. Di Bisceglie AM, Martin P, Kassianides C, Lisker-Melman M, Murray L, Waggoner J, Goodman Z, Banks SM, Hoofnagle JH. Recombinant interferon alfa therapy for chronic hepatitis C. A randomized, double-blind, placebo-controlled trial. *N Engl J Med* 1989 Nov 30;321(22):1506-10.
3. Haydon GH, Jarvis LM, Blair CS, Simmonds P, Harrison DJ, Simpson KJ, Hayes PC. Clinical significance of intrahepatic hepatitis C virus levels in patients with chronic HCV infection. *Gut* 1998 Apr;42(4):570-5.
4. Simmonds P. Clinical relevance of hepatitis C virus genotypes. *Gut* 1997 Mar;40(3):291-3.
5. Hadziyannis SJ, Sette H, Jr., Morgan TR, Balan V, Diago M, Marcellin P, Ramadori G, Bodenheimer H, Jr., Bernstein D, Rizzetto M, Zeuzem S, Pockros PJ, Lin A, Ackrill AM. Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann Intern Med* 2004 Mar 2;140(5):346-55.
6. Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, Reindollar R, Goodman ZD, Koury K, Ling M, Albrecht JK. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* 2001 Sep 22;358(9286):958-65.
7. 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 Dec;40(6):1260-5.
8. Mangia A, Santoro R, Minerva N, Ricci GL, Carretta V, Persico M, Vinelli F, Scotto G, Bacca D, Annese M, Romano M, Zechini F, Sogari F, Spirito F, Andriulli A. Peginterferon alfa-2b and ribavirin for 12 vs. 24 weeks in HCV genotype 2 or 3. *N Engl J Med* 2005 Jun

1  
2 23:352(25):2609-17.  
3

4 9. Enomoto N, Sakuma I, Asahina Y, Kurosaki M, Murakami T, Yamamoto C, Ogura  
5 Y, Izumi N, Marumo F, Sato C. Mutations in the nonstructural protein 5A gene and response  
6 to interferon in patients with chronic hepatitis C virus 1b infection. *N Engl J Med*1996 Jan  
7 11:334(2):77-81.  
8

9  
10  
11  
12 10. El-Shamy A, Nagano-Fujii M, Sasase N, Imoto S, Kim SR, Hotta H. Sequence  
13 variation in hepatitis C virus nonstructural protein 5A predicts clinical outcome of pegylated  
14 interferon/ribavirin combination therapy. *Hepatology*2008 Jul;48(1):38-47.  
15  
16

17  
18  
19 11. Hamano K, Sakamoto N, Enomoto N, Izumi N, Asahina Y, Kurosaki M, Ueda E,  
20 Tanabe Y, Maekawa S, Itakura J, Watanabe H, Kakinuma S, Watanabe M. Mutations in the  
21 NS5B region of the hepatitis C virus genome correlate with clinical outcomes of  
22 interferon-alpha plus ribavirin combination therapy. *J Gastroenterol Hepatol*2005  
23 Sep;20(9):1401-9.  
24  
25

26  
27  
28 12. Chayama K, Suzuki F, Tsubota A, Kobayashi M, Arase Y, Saitoh S, Suzuki Y,  
29 Murashima N, Ikeda K, Takahashi N, Kinoshita M, Kumada H. Association of amino acid  
30 sequence in the PKR-eIF2 phosphorylation homology domain and response to interferon  
31 therapy. *Hepatology*2000 Nov;32(5):1138-44.  
32  
33

34  
35  
36 13. Akuta N, Suzuki F, Sezaki H, Suzuki Y, Hosaka T, Someya T, Kobayashi M, Saitoh  
37 S, Watahiki S, Sato J, Matsuda M, Arase Y, Ikeda K, Kumada H. Association of amino acid  
38 substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and  
39 non-virological response to interferon-ribavirin combination therapy. *Intervirology*2005  
40 Nov-Dec;48(6):372-80.  
41  
42

43  
44  
45 14. Toyoda H, Kumada T, Tada T, Arakawa T, Hayashi K, Honda T, Katano Y, Goto H.  
46 Association between HCV amino acid substitutions and outcome of peginterferon and  
47 ribavirin combination therapy in HCV genotype 1b and high viral load. *J Gastroenterol*  
48 *Hepatol*2010 Jun;25(6):1072-8.  
49  
50

51  
52  
53 15. Donlin MJ, Cannon NA, Aurora R, Li J, Wahed AS, Di Bisceglie AM, Tavis JE.  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2 Contribution of genome-wide HCV genetic differences to outcome of interferon-based  
3  
4 therapy in Caucasian American and African American patients. *PLoS One*2010;5(2):e9032.

5  
6 16. Donlin MJ, Cannon NA, Yao E, Li J, Wahed A, Taylor MW, Belle SH, Di Bisceglie  
7  
8 AM, Aurora R, Tavis JE. Pretreatment sequence diversity differences in the full-length  
9  
10 hepatitis C virus open reading frame correlate with early response to therapy. *J Virol*2007  
11  
12 Aug;81(15):8211-24.

13  
14  
15 17. Murakami T, Enomoto N, Kurosaki M, Izumi N, Marumo F, Sato C. Mutations in  
16  
17 nonstructural protein 5A gene and response to interferon in hepatitis C virus genotype 2  
18  
19 infection. *Hepatology*1999 Oct;30(4):1045-53.

20  
21 18. Hayashi K, Katano Y, Honda T, Ishigami M, Itoh A, Hirooka Y, Nakano I, Urano F,  
22  
23 Yoshioka K, Toyoda H, Kumada T, Goto H. Mutations in the interferon  
24  
25 sensitivity-determining region of hepatitis C virus genotype 2a correlate with response to  
26  
27 pegylated-interferon-alpha 2a monotherapy. *J Med Virol*2009 Mar;81(3):459-66.

28  
29 19. Kobayashi M, Watanabe K, Ishigami M, Murase K, Ito H, Ukai K, Yano M, Takagi  
30  
31 K, Hattori M, Kakumu S, Yoshioka K. Amino acid substitutions in the nonstructural region  
32  
33 5A of hepatitis C virus genotypes 2a and 2b and its relation to viral load and response to  
34  
35 interferon. *Am J Gastroenterol*2002 Apr;97(4):988-98.

36  
37 20. Akuta N, Suzuki F, Hirakawa M, Kawamura Y, Yatsuji H, Sezaki H, Suzuki Y,  
38  
39 Hosaka T, Kobayashi M, Saitoh S, Arase Y, Ikeda K, Kumada H. Association of Amino Acid  
40  
41 Substitution Pattern in Core Protein of Hepatitis C Virus Genotype 2a High Viral Load and  
42  
43 Virological Response to Interferon-Ribavirin Combination Therapy. *Intervirology*2009 May  
44  
45 5:52(6):301-9.

46  
47 21. Saito T, Ito T, Ishiko H, Yonaha M, Morikawa K, Miyokawa A, Mitamura K.  
48  
49 Sequence analysis of PePHD within HCV E2 region and correlation with resistance of  
50  
51 interferon therapy in Japanese patients infected with HCV genotypes 2a and 2b. *Am J*  
52  
53 *Gastroenterol*2003 Jun;98(6):1377-83.

54  
55 22. Watanabe H, Nagayama K, Enomoto N, Itakura J, Tanabe Y, Sato C, Izumi N,  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

1  
2 Watanabe M. Amino acid substitutions in PKR-eIF2 phosphorylation homology domain  
3 (PePHD) of hepatitis C virus E2 protein in genotype 2a/2b and 1b in Japan and interferon  
4 efficacy. *Hepatol Res*2003 Aug;26(4):268-74.  
5  
6

7  
8 23. Gale M, Jr., Blakely CM, Kwieciszewski B, Tan SL, Dossett M, Tang NM, Korth  
9 MJ, Polyak SJ, Gretch DR, Katze MG. Control of PKR protein kinase by hepatitis C virus  
10 nonstructural 5A protein: molecular mechanisms of kinase regulation. *Mol Cell Biol*1998  
11 Sep;18(9):5208-18.  
12  
13

14  
15 24. Hiramatsu N, Oze T, Yakushijin T, Inoue Y, Igura T, Mochizuki K, Imanaka K,  
16 Kaneko A, Oshita M, Hagiwara H, Mita E, Nagase T, Ito T, Inui Y, Hijioka T, Katayama K,  
17 Tamura S, Yoshihara H, Imai Y, Kato M, Yoshida Y, Tatsumi T, Ohkawa K, Kiso S, Kanto T,  
18 Kasahara A, Takehara T, Hayashi N. Ribavirin dose reduction raises relapse rate  
19 dose-dependently in genotype 1 patients with hepatitis C responding to pegylated interferon  
20 alpha-2b plus ribavirin. *J Viral Hepat*2009 Aug;16(8):586-94.  
21  
22

23  
24 25. Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu  
25 P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB. Genetic variation  
26 in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature*2009 Sep  
27 17;461(7262):399-401.  
28  
29

30  
31 26. Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N,  
32 Nakagawa M, Korenaga M, Hino K, Hige S, Ito Y, Mita E, Tanaka E, Mochida S, Murawaki  
33 Y, Honda M, Sakai A, Hiasa Y, Nishiguchi S, Koike A, Sakaida I, Imamura M, Ito K, Yano K,  
34 Masaki N, Sugauchi F, Izumi N, Tokunaga K, Mizokami M. Genome-wide association of  
35 IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic  
36 hepatitis C. *Nat Genet*2009 Oct;41(10):1105-9.  
37  
38

39  
40 27. Rauch A, Kutalik Z, Descombes P, Cai T, di Iulio J, Mueller T, Bochud M, Battegay  
41 M, Bernasconi E, Borovicka J, Colombo S, Cerny A, Dufour JF, Furrer H, Gunthard HF,  
42 Heim M, Hirschel B, Malinverni R, Moradpour D, Mullhaupt B, Witteck A, Beckmann JS,  
43 Berg T, Bergmann S, Negro F, Telenti A, Bochud PY. Genetic variation in IL28B Is  
44  
45  
46  
47  
48  
49  
50

1  
2 Associated with Chronic Hepatitis C and Treatment Failure - A Genome-Wide Association  
3  
4 Study. *Gastroenterology* 2010 Jan 7.  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## FIGURE LEGENDS

**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 open reading frame (ORF) ( Fig.1, left) and for each HCV protein (Fig.1, right).

**Fig.2a 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.

**Fig.2b 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.

**Fig.2c Sliding window analysis.**

Viral regions affecting treatment outcome are shown as red spots. There are four hot spots: at core amino acid 110, amino acids 400-403 (i.e. the hyper variable region) in Envelope2 (E2) region, amino acids 724-743 in E2 and amino acids 2258-2306 in the nonstructural (NS)5A.

1  
2 **Fig.2d Sequence alignment amino acids in the nonstructural (NS)5A around amino**  
3 **acids 2258 to 2306.**  
4

5  
6  
7 Dashes indicate amino acids identical to the consensus sequence and substituted amino  
8  
9 acids are shown by standard single letter codes.  
10

11  
12  
13  
14 **Fig.3 Correlation between pretreatment HCV RNA levels and the number of**  
15 **substitutions in the NS5A region aa 2258 to 2306.**  
16

17  
18  
19 Spearman's correlation coefficient by rank test is demonstrated.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

Table 1. Baseline Characteristics of All Patients (Group 1 and 2)

Characteristic	SVR (n = 58)			non-SVR (n = 20)		P value <sup>‡</sup>	
	Group 1 (n = 36)	Group 2 (n = 22)	Combined (n = 58)	Group 1 (n = 7)	Group 2 (n = 13)		Combined (n = 20)
Gender(Male/Female)	20 / 16	9 / 13	29 / 29	4 / 3	5 / 8	9 / 11	0.80 <sup>†</sup>
Age(yrs)	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 <sup>‡</sup>
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 <sup>‡</sup>
Platelet( $\times 10^4/mm^3$ )	20.8 ± 6.2	19.0 ± 5.2	20.1 ± 5.8	14.7 ± 7.1	19.1 ± 4.9	17.6 ± 6.0	0.11 <sup>‡</sup>
Fibrosis score(0-2 / $\geq 3$ ) <sup>§</sup>	34 / 1	19 / 2	53 / 3	4 / 3	11 / 2	15 / 5	0.049 <sup>†</sup>
HCV RNA(KIU/ml)	760(2-3100)**	340(54-3600)	550(12-3600)	1300 (350-30000)	1400 (180-5000)	1300 (180-30000)	0.002 <sup>  </sup>
IFN dose( $\geq 80\%$ / 60-80%) <sup>¶</sup>	28 / 4	21 / 1	49 / 5	4 / 3	11 / 2	15 / 5	0.12 <sup>†</sup>
Ribavirin dose( $\geq 80\%$ / 60-80%) <sup>¶</sup>	27 / 5	17 / 5	44 / 10	4 / 3	5 / 8	9 / 11	0.003 <sup>†</sup>
RVR rate (%)	87.5	54.5	74.1	33.3	46.1	42.1	0.022 <sup>†</sup>
EVR rate (%)	100	100	100	66.7	100	89.4	0.07 <sup>†</sup>

\* : mean  $\pm$  SD \*\* : median (range) † : Fisher's exact probability test ‡ : Student t test || : Mann-Whitney's U test ¶ : P values between all SVR (n = 58) vs. all non-SVR (n = 20)

Several clinical characteristics listed below were unavailable in some patients

§ : SVR : n = 56 (35 in group1, 21 in group2), non-SVR : n = 17 (7 in group1, 10 in group2) ¶ : SVR : n = 54 (32 in group1, 22 in group2)